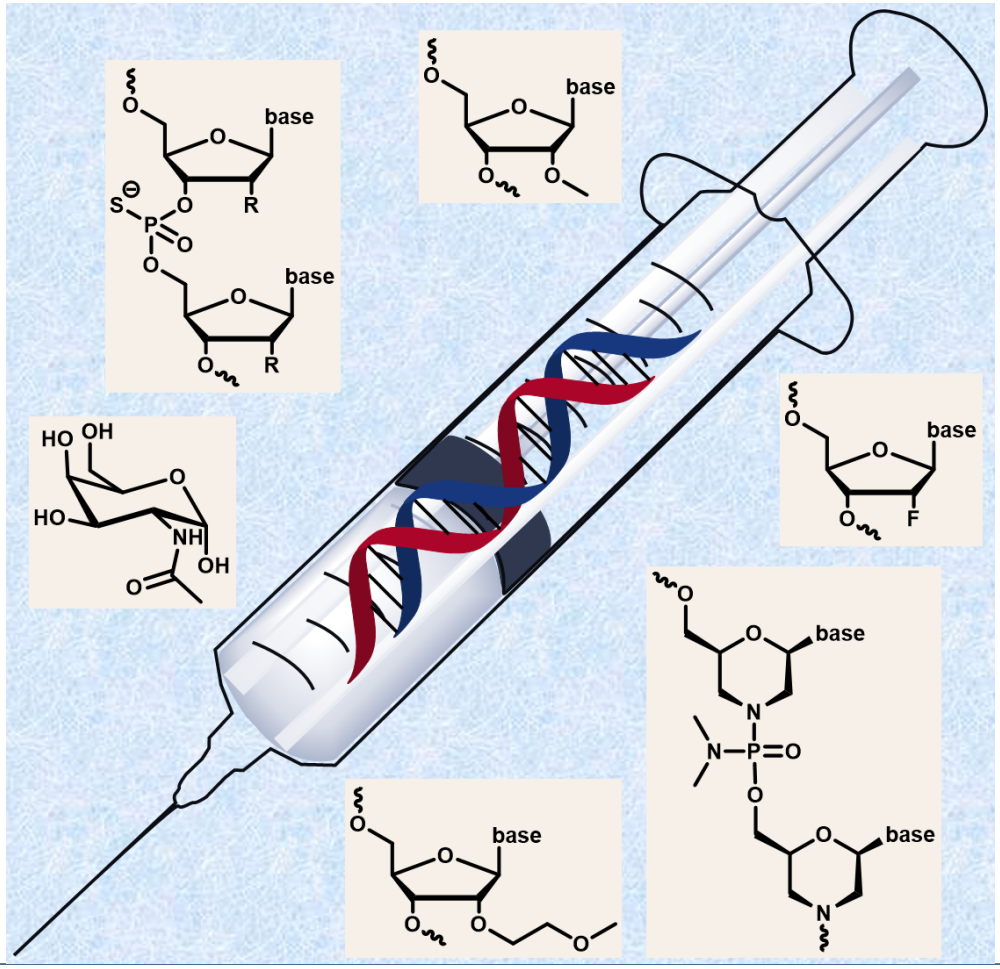
BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Celebrating the role of chemistry in the success of oligonucleotides as therapeutics

Edited by Pawan Kumar and Tom Brown



Imprint

Beilstein Journal of Organic Chemistry
www.bjoc.org
ISSN 1860-5397
Email: journals-support@beilstein-institut.de

The *Beilstein Journal of Organic Chemistry* is published by the Beilstein-Institut zur Förderung der Chemischen Wissenschaften.

Beilstein-Institut zur Förderung der Chemischen Wissenschaften
Trakehner Straße 7–9
60487 Frankfurt am Main
Germany
www.beilstein-institut.de

The cover image illustrates that oligonucleotides incorporating chemically modified nucleotides are now approved medicines.

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

The role of chemistry in the success of oligonucleotides as therapeutics

Pawan Kumar*[1,§] and Tom Brown*[2]

Editorial

Open Access

Address:
[1]Takeda Development Center Americas, Inc. (TDCA), 9625 Town Centre Drive, San Diego, CA 92121, USA and [2]Chemistry Research Laboratory, University of Oxford, 12 Mansfield Road, Oxford, OX1 3TA, UK

Email:
Pawan Kumar* - pawan.kumar1@takeda.com; Tom Brown* - tom.brown@chem.ox.ac.uk

* Corresponding author
§ alternative email address: pkumar80@hotmail.co.uk

Keywords:
antisense oligonucleotides; chemically modified nucleotides; siRNAs

Beilstein J. Org. Chem. **2022,** *18,* 197–199.
https://doi.org/10.3762/bjoc.18.22

Received: 27 January 2022
Accepted: 02 February 2022
Published: 14 February 2022

This article is part of the thematic issue "Celebrating the role of chemistry in the success of oligonucleotides as therapeutics".

Guest Editors: P. Kumar and T. Brown

RNA-targeting oligonucleotides (e.g., antisense, siRNA, and anti-miR) are widely explored as fundamental research tools and are gaining increasing promise as therapeutic agents, particularly against diseases of genetic origin. The idea of treating a disease by targeting the molecular messenger (mRNA) to stop the synthesis of proteins using short strands of DNA, now known as antisense oligonucleotides, was first coined about 40 years ago [1]. Almost 20 years later, another endogenous mechanism, known as RNA interference (RNAi) was discovered when it was shown that short stretches of double-stranded nucleotides, which are called short interfering RNA or "siRNA," can target mRNA and prevent it from being translated to make proteins [2]. While the mechanism by which antisense oligonucleotides (single stranded oligonucleotide) and siRNA (short RNA duplexes) work are completely different, both of them target mRNA to disrupt protein synthesis. In the following text we will refer both antisense oligonucleotides and siRNAs collectively as therapeutic oligonucleotides.

More than 10 oligonucleotide drugs have received regulatory approval by the FDA and are now helping patients suffering from conditions that were previously seen as untreatable. However, the road from bench to bedside for therapeutic oligonucleotides has not been straightforward. It took about four decades (from discovery of the antisense mechanism) to reach the point where it is now possible to treat previously undruggable conditions using therapeutic oligonucleotides. There were many challenges including, but not limited to, poor stability of unmodified DNA strands and short RNA duplexes in cells, large anionic charge, poor drug-like properties, and a tendency to trigger the immune response in the body. Chemists have been at forefront of solving these issues, and have introduced many chemically modified nucleotides into oligonucleotides to increase their binding affinity toward RNA targets, and to improve their stability against nucleases to slow down degradation. This strategy has been successful, and most oligonucleotide-based drugs that have been approved by the FDA contain

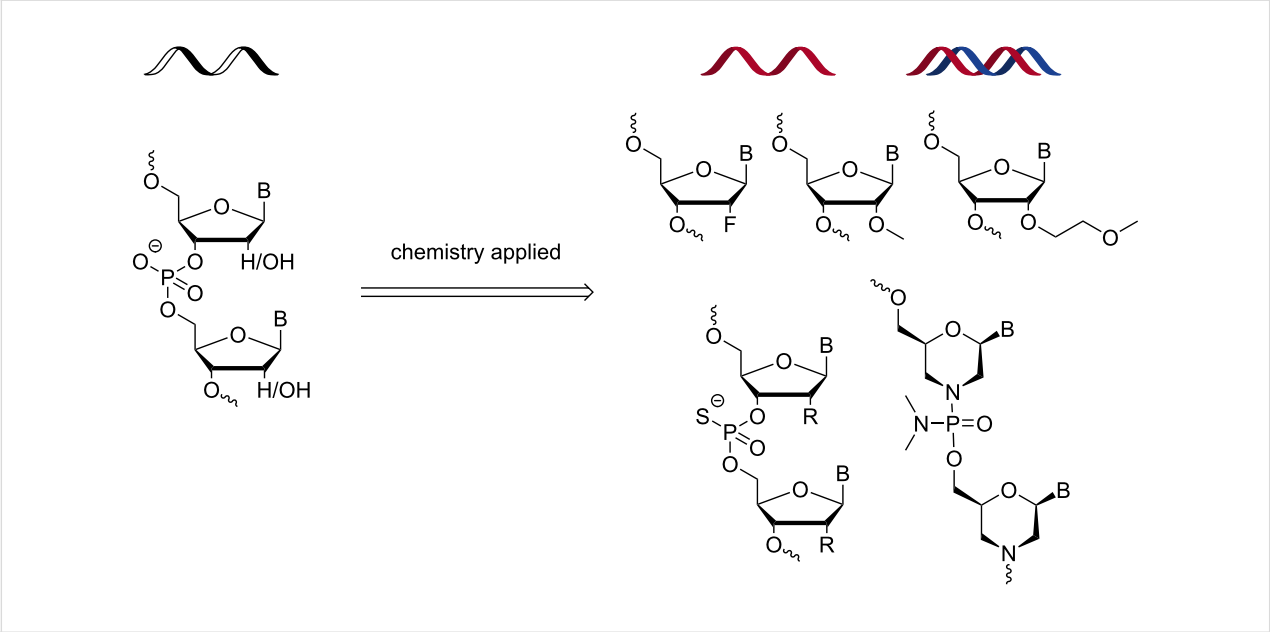


Figure 1: Introduction of chemically modified nucleotides into oligonucleotides is an essential part of their development as therapeutics.

chemically modified nucleotides (Figure 1) indicating the critical role chemists have played in bring oligonucleotides from bench to bedside. Importantly, a plethora of different chemically modified nucleotides have been described in the literature [3], and the field owes a big thanks to all the researchers who have contributed to expanding chemical space of modified nucleotides (building blocks of oligonucleotides). However, discussing these modifications in detail is beyond the scope of this editorial.

The use of chemically modified nucleotides massively improves the drug-like properties of oligonucleotides. However, their efficient delivery to the desired tissue/organ also needs to be addressed. Conjugation of oligonucleotides to GalNAc (*N*-acetylgalactosamine) has successfully been used for targeted delivery of oligonucleotides (both antisense and siRNAs) to the liver. Currently, there are three siRNA-GalNAc conjugates approved by the FDA and many others are in late-stage clinical trials. The success of GalNAc-oligonucleotide conjugates highlights the power of therapeutic oligonucleotides in treating previously untreatable conditions, once a means to deliver them into a desired tissue has been achieved. However, challenges remain in finding suitable GalNAc equivalents for delivering oligonucleotides to tissues other than the liver. Delivery of oligonucleotides to extra hepatic tissues is certainly an area where many academic and biotech laboratories are focused. In this context antibody-oligonucleotide conjugates have shown promise in targeted delivery and are entering into clinical trials. Chemistry is at the forefront of discovering new conjugates/modifications and it is hoped that solutions will be found to address the huge unmet medical need in the CNS space, and to treat diseases that have been elusive until now.

Developing lipid nanoparticles that can selectively deliver oligonucleotides to a desired tissue is also an attractive strategy. Importantly, siRNA encapsulated into lipid nanoparticles has been shown to be effective in patisiran, the first RNAi drug to reach patients. Lipid nanoparticles are also being used in the new generation of RNA vaccines for tackling the COVID pandemic.

Another noteworthy advancement is the ease and scale with which oligonucleotides are being produced today. Without access to larger quantities of oligonucleotides it would not have been possible to develop them as therapeutic agents. Forty years ago, the synthesis of an oligonucleotide in the lab was a huge task. Since then, seminal work from the Caruthers lab has solved this problem with the introduction of phosphoramidite chemistry [4]. This proved an outstanding breakthrough and is currently catalyzing the development of many new technologies including next generation siRNAs, antisense oligonucleotides, and CRISPR-based gene editing systems. Thanks to this phosphoramidite approach, it has also been possible to mass-produce oligonucleotide primers and probes for use in diagnostic testing kits for the detection of COVID-19 for tackling the pandemic.

With this thematic issue, we express our sincere gratitude to all the scientists for their ground-breaking work to bring oligonucleotide therapeutics to the bedside. We also express our

sincere thanks to the authors and reviewers who have contributed despite the challenges posed by the COVID pandemic. Support from the editorial team of the *Beilstein Journal of Organic Chemistry* is also greatly appreciated. The issue comprises many excellent contributions in the form of original research and review articles from world-leading experts in the field. We hope the thematic issue will inspire readers and provide a state of the art background on emerging areas in the rapidly evolving therapeutic oligonucleotide field.

Pawan Kumar and Tom Brown

San Diego and Oxford, January 2022

References

1. Zamecnik, P. C.; Stephenson, M. L. *Proc. Natl. Acad. Sci. U. S. A.* **1978,** *75,* 280–284. doi:10.1073/pnas.75.1.280
2. Elbashir, S. M.; Harborth, J.; Lendeckel, W.; Yalcin, A.; Weber, K.; Tuschl, T. *Nature* **2001,** *411,* 494–498. doi:10.1038/35078107
3. Wan, W. B.; Seth, P. P. *J. Med. Chem.* **2016,** *59,* 9645–9667. doi:10.1021/acs.jmedchem.6b00551
4. Beaucage, S. L.; Caruthers, M. H. *Tetrahedron Lett.* **1981,** *22,* 1859–1862. doi:10.1016/s0040-4039(01)90461-7

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Synthesis and properties of oligonucleotides modified with an *N*-methylguanidine-bridged nucleic acid (GuNA[Me]) bearing adenine, guanine, or 5-methylcytosine nucleobases

Naohiro Horie[1], Takao Yamaguchi[1], Shinji Kumagai[2] and Satoshi Obika[*1,3]

Full Research Paper

Open Access

Address:
[1]Graduate School of Pharmaceutical Sciences, Osaka University, 1-6 Yamadaoka, Suita, Osaka 565-0871, Japan, [2]Sohyaku. Innovative Research Division, Mitsubishi Tanabe Pharma Corporation, Shonan Health Innovation Park, 2-26-1 Muraoka-Higashi, Fujisawa, Kanagawa 251-8555, Japan and [3]National Institutes of Biomedical Innovation, Health and Nutrition (NIBIOHN), 7-6-8 Saito-Asagi, Ibaraki, Osaka 567-0085, Japan

Email:
Satoshi Obika[*] - obika@phs.osaka-u.ac.jp

* Corresponding author

Keywords:
artificial nucleic acid; duplex-forming ability; oligonucleotide synthesis

Beilstein J. Org. Chem. **2021,** *17,* 622–629.
https://doi.org/10.3762/bjoc.17.54

Received: 06 November 2020
Accepted: 18 February 2021
Published: 04 March 2021

This article is part of the thematic issue "Celebrating the role of chemistry in the success of oligonucleotides as therapeutics".

Guest Editors: P. Kumar and T. Brown

Abstract

Chemical modifications have been extensively used for therapeutic oligonucleotides because they strongly enhance the stability against nucleases, binding affinity to the targets, and efficacy. We previously reported that oligonucleotides modified with an *N*-methylguanidine-bridged nucleic acid (GuNA[Me]) bearing the thymine (T) nucleobase show excellent biophysical properties for applications in antisense technology. In this paper, we describe the synthesis of GuNA[Me] phosphoramidites bearing other typical nucleobases including adenine (A), guanine (G), and 5-methylcytosine (^{m}C). The phosphoramidites were successfully incorporated into oligonucleotides following the method previously developed for the GuNA[Me]-T-modified oligonucleotides. The binding affinity of the oligonucleotides modified with GuNA[Me]-A, -G, or -^{m}C toward the complementary single-stranded DNAs or RNAs was systematically evaluated. All of the GuNA[Me]-modified oligonucleotides were found to have a strong affinity for RNAs. These data indicate that GuNA[Me] could be a useful modification for therapeutic antisense oligonucleotides.

Introduction

The efficacy and safety of therapeutic oligonucleotides can be controlled by chemical modifications. For applications in antisense technology, chemical modifications aimed at enhancing the duplex-forming ability toward a target RNA (i.e., a complementary single-stranded RNA) and improving the stability against enzymatic degradations are commonly utilized. For instance, antisense oligonucleotides (ASOs) modified with 2',4'-bridged nucleic acid/locked nucleic acid (2',4'-BNA/LNA; Figure 1) are now widely used for gene regulation in vitro and in vivo because 2',4'-BNA/LNA greatly increases the affinity

toward the target RNAs, thus enhancing the efficacy of the modified ASOs [1-6]. Notably, the biophysical and pharmacological properties of 2',4'-BNA/LNA-modified ASOs can be further altered with subtle structural changes. Seth and co-workers developed the *S*-2',4'-constrained-2'-*O*-ethyl (*S*-cEt; Figure 1) derivative, which has an exocyclic methyl group in its bridged structure [7,8]. The *S*-cEt-modified ASOs displayed a higher nuclease resistance and lower hepatotoxicity in in vivo experiments than the corresponding 2',4'-BNA/LNA-modified ASOs [9]; the reduction in hepatotoxicity might be a sequence-dependent phenomenon. Currently, a number of *S*-cEt-modified ASOs with low hepatotoxicity have been confirmed to be effective for gene regulations in vivo [10,11]. We previously developed amido-bridged nucleic acids (AmNA[R]s) (Figure 1), in which the *N*-alkyl substituent groups were found to modulate nuclease resistance and hepatic distributions [12]. Wengel's group reported the synthesis of 2'-amino-LNA (Figure 1) functionalized with a peptide or sugar at the N2'-position, with the aim of modulating the physicochemical properties and specific organ distributions of the therapeutic oligonucleotides [13,14]. A more favorable example is the covalent attachment of a guanidine moiety, which is a common approach to partially neutralize the polyanionic property of oligonucleotides [15-18]. In our previous study, a guanidine-bridged nucleic acid (GuNA[H]; Figure 1) bearing a thymine (T) nucleobase was synthesized as a novel artificial nucleic acid for antisense applications [19]. The modification of oligonucleotides with GuNA[H]-T improved the nuclease resistance, cell membrane permeability, and binding affinity toward complementary single-stranded DNAs (ssDNAs) and RNAs (ssRNAs). We also synthesized and evaluated a GuNA[H]-T analog bearing a methyl group in the guanidine moiety (GuNA[Me]-T; Figure 1) [20]. The GuNA[Me]-T exhibited a similar duplex-forming ability and nuclease resistance as GuNA[H]-T. Since a subtle change in the structure of the 2',4'-BNA/LNA modulated its biophysical and pharmacological properties, in vivo experiments with GuNA[H] and GuNA[Me] are expected to provide further mechanistic insights into how small substituents affect the efficacy and safety of therapeutic oligonucleotides. Thus, the synthesis of GuNA[Me] phosphoramidites bearing other typical nucleobases, i.e., adenine (A), guanine (G), or 5-methylcytosine (^{m}C), instead of the immunologically unfavorable cytosine (C), is needed.

The preparation of all four phosphoramidites (A, G, ^{m}C, and T) is generally not easy because each nucleobase differs in the sensitivity to reactions, and appropriate protecting groups need to be selected [8,21-23]. We recently achieved the synthesis of all four GuNA[H] phosphoramidites, where transglycosylations of the 2'-amino-LNA analog with the corresponding nucleobases were performed as the key reactions [24,25]. The transglycosylation is a powerful strategy that simplifies the preparation of phosphoramidites at the late stages of the syntheses [26,27]. Here, we describe the synthesis of GuNA[Me]-A, -G, and -^{m}C phosphoramidites and their incorporations into oligonucleotides. The duplex-forming abilities of all the GuNA[Me]-modified oligonucleotides toward their ssDNA and ssRNA complements were systematically evaluated.

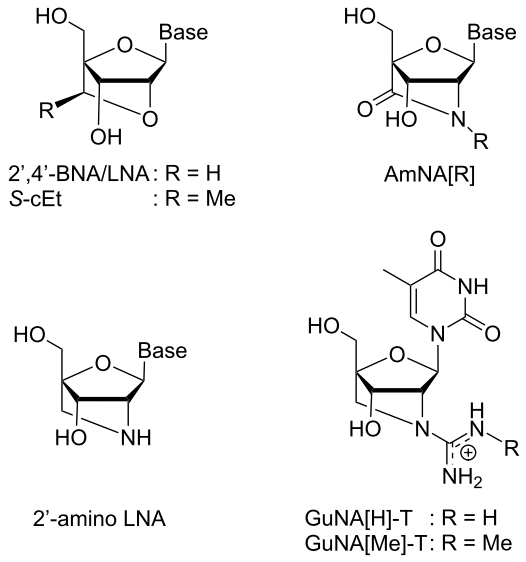
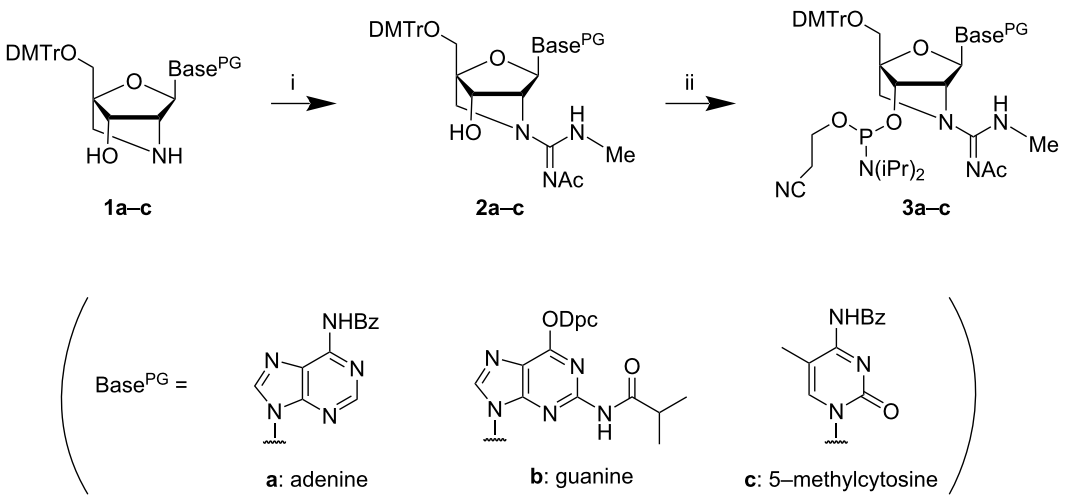


Figure 1: Structures of 2',4'-BNA/LNA analogs.

Results and Discussion

Synthesis of the GuNA[Me] phosphoramidites bearing either an A, G, or ^{m}C nucleobase

The preparation of the GuNA[Me]-A, -G, and -^{m}C phosphoramidites **3a–c** needed for the synthesis of the GuNA[Me]-modified oligonucleotides is detailed in Scheme 1. The acetyl group was selected as a protecting group for the guanidine moiety because it can be easily removed under the basic conditions (ammonia/methylamine solution) used for the DNA synthesis [20]. The phosphoramidite synthesis was started from 2'-amino-LNAs **1a–c**, which were rapidly prepared via the transglycosylations of 2'-amino-LNA-T [25]. First, the 2'-amino groups of **1a–c** were converted into guanidine moieties with a methyl group using *N*-acetyl-*S*,*N'*-dimethylisothiourea [28], which yielded 65–83% of the products **2a–c**. Subsequently, the designed GuNA[Me] phosphoramidites **3a–c** were successfully obtained following the phosphitylation of the 3'-hydroxy groups of **2a–c**. Notably, since the nucleobases were introduced at the late stage of the synthesis, we had no difficulty preparing these phosphoramidites.



Scheme 1: The preparation of the GuNA[Me]-A, -G, and -^{m}C phosphoramidites **3a–c**. Reagents and conditions: (i) *N*-acetyl-*S*,*N*′-dimethylisothiourea, AgOTf, DIPEA, THF, rt, 72% (**2a**), 65% (**2b**), 83% (**2c**); (ii) $(iPr)_2NP(Cl)O(CH_2)_2CN$, DIPEA, CH_2Cl_2, rt, 87% (**3a**); 65% (**3b**); 72% (**3c**).

Synthesis of oligonucleotides modified with GuNA[Me]-A, -G, or -^{m}C

The prepared GuNA[Me]-A, -G, and -^{m}C phosphoramidites were incorporated into the middle position of 12-mer oligonucleotides (Table 1). The oligonucleotide synthesis was performed using an automated DNA synthesizer following the established synthetic method for GuNA[Me]-T-modified oligonucleotides [20]. 5-(Ethylthio)-1*H*-tetrazole (ETT) was used as an activator for the coupling, and the coupling time was extended from 40 s to 20 min for the GuNA[Me] phosphoramidites. Other conditions were the same as those used for general DNA synthesis. After the elongation, the oligonucleotides were treated with ammonia/methylamine solution (7 M NH_3 in methanol/40% aqueous methylamine 1:1) at 60 °C for 5 h. Under these conditions, we obtained the GuNA[Me]-^{m}C-modified oligonucleotide **ON3** with high purity. In the case of the GuNA[Me] having a purine nucleobase (**ON1** and **ON2**), the acetyl group in the guanidine moiety remained in a considerable amount. This means that we should give attention to the reactivity of each nucleobase. Finally, the acetyl group was successfully removed by extending the deprotection time to 10 h. The yield range of the designed oligonucleotides **ON1**–**ON3** was 12–25%, as shown in Table 1.

Duplex-forming ability of oligonucleotides modified with GuNA[Me]-A, -G, or -^{m}C

The binding affinity of the GuNA[Me]-modified oligonucleotides **ON1**–**ON3** toward ssDNAs or ssRNAs was evaluated by measuring UV melting temperatures (T_m values), and the obtained values were compared with those of the corresponding unmodified oligonucleotides (**ON6**–**ON8**). The results are shown in Table 2. As expected, all of the GuNA[Me]-modified oligonucleotides **ON1**–**ON3** exhibited markedly higher T_m values toward ssRNAs than their unmodified counterparts **ON6**–**ON8** (ΔT_m = 5–6 °C). These results are similar to those obtained for the GuNA[Me]-T-modified oligonucleotide **ON4** (ΔT_m = 5 °C). Additionally, the modified **ON1**–**ON3** showed an enhanced duplex-forming ability toward the complementary ssDNAs (ΔT_m = 3–6 °C). Among them, GuNA[Me]-A-modified **ON1** exhibited a slightly lower ΔT_m value than others. This type of nucleobase-dependent difference in ΔT_m values is also seen in other GuNA[H]-modified oligonucleotides [25]. Since

Table 1: Synthetic yields and mass spectral data of the GuNA[Me]-modified oligonucleotides **ON1**–**ON3**.

oligonucleotides[a] (5′–3′)		yield [%]	MALDI–TOF mass	
			found $[M - H]^-$	calcd. $[M - H]^-$
d(GCG TTA TTT GCT)	(**ON1**)	12	3723.9	3724.5
d(GCG TTG TTT GCT)	(**ON2**)	14	3738.9	3740.5
d(GCG TT^{m}C TTT GCT)	(**ON3**)	25	3714.4	3714.5

[a]A, G, and ^{m}C indicate GuNA[Me] modifications.

Table 2: T_m values of duplexes formed between GuNA[Me]-modified oligonucleotides and complementary ssRNAs or ssDNAs.[a]

oligonucleotides[a] (5'–3')		T_m (ΔT_m) [°C]	
		vs ssRNA	vs ssDNA
d(GCG TTT TTT GCT)[b]	(**ON5**)	47	51
d(GCG TT<u>T</u> TTT GCT)[b]	(**ON4**)	52 (+5)	56 (+5)
d(GCG TTA TTT GCT)	(**ON6**)	45	49
d(GCG TT<u>A</u> TTT GCT)	(**ON1**)	50 (+5)	52 (+3)
d(GCG TTG TTT GCT)	(**ON7**)	51	54
d(GCG TT<u>G</u> TTT GCT)	(**ON2**)	57 (+6)	59 (+5)
d(GCG TTC TTT GCT)	(**ON8**)	52	53
d(GCG TT<u>^{m}C</u> TTT GCT)	(**ON3**)	58 (+6)	59 (+6)

[a]Conditions: 10 mM sodium phosphate buffer (pH 7.2), 100 mM NaCl, 4 μM each oligonucleotide, 0.5 °C/min at 260 nm. Sequences of the complementary ssRNA and ssDNA are 5'-r(AGC AAA NAA CGC)-3' and 5'-d(AGC AAA NAA CGC)-3', respectively. <u>T</u>, <u>A</u>, <u>G</u>, and <u>^{m}C</u> indicate GuNA[Me] modifications. [b]See reference [20].

oligonucleotides modified with 2',4'-BNA/LNA or its analog scpBNA show different nucleobase dependency [1,23], these results could be considered characteristic of the GuNA-modified oligonucleotides. Interactions between the guanidine moiety and nearby base pairing(s) might have affected the ΔT_m values, though further investigations are needed for the details.

CD spectral analyses of duplexes modified with GuNA[Me]-G

To analyze the structures of the duplexes containing GuNA[Me], circular dichroism (CD) spectra were measured for **ON2**/ssRNA and **ON2**/ssDNA duplexes (Figure 2). The CD spectra of **ON2**/ssRNA and **ON2**/ssDNA were found to be similar to those of the **ON7**/ssRNA and **ON7**/ssDNA duplexes, demonstrating that one modification with GuNA[Me] does not affect the whole duplex structures. Similar results were observed for **ON4**/ssRNA and **ON4**/ssDNA (Figure S15 in Supporting Information File 1). In our previous studies, DNA/RNA (A-form) duplexes containing a multiple GuNA[H] modification displayed similar spectral patterns to the natural and the 2',4'-BNA/LNA-modified counterparts [19]. Since GuNA[Me] showed similar results to GuNA[H] in terms of the duplex-forming ability [25], a multiple GuNA[Me] modification to A-form duplexes is also believed not to affect the structures.

Conclusion

We successfully synthesized GuNA[Me] phosphoramidites bearing either an A, G, or ^{m}C nucleobase. Each monomer was derived from the corresponding 2'-amino-LNA in two steps and introduced into oligonucleotides. By protecting the guanidine moieties with an acetyl group, we could obtain the oligonucleotides within a 12–25% yield range under the basic conditions (ammonia/methylamine solution) commonly used in oligonucleotide synthesis. The synthesized GuNA[Me]-modified oligonucleotides showed a high binding affinity toward the complementary ssRNAs and ssDNAs. Considering the facile synthesis of the GuNA[Me] monomers and the ability of the GuNA[Me]-modified oligonucleotides to form stable duplexes with ssRNAs, we expect that a modification using GuNA[Me] could be useful for antisense applications. In our ongoing studies, we are evaluating the efficacy of ASOs modified with GuNAs in vitro and in vivo, and the results will be reported in due course.

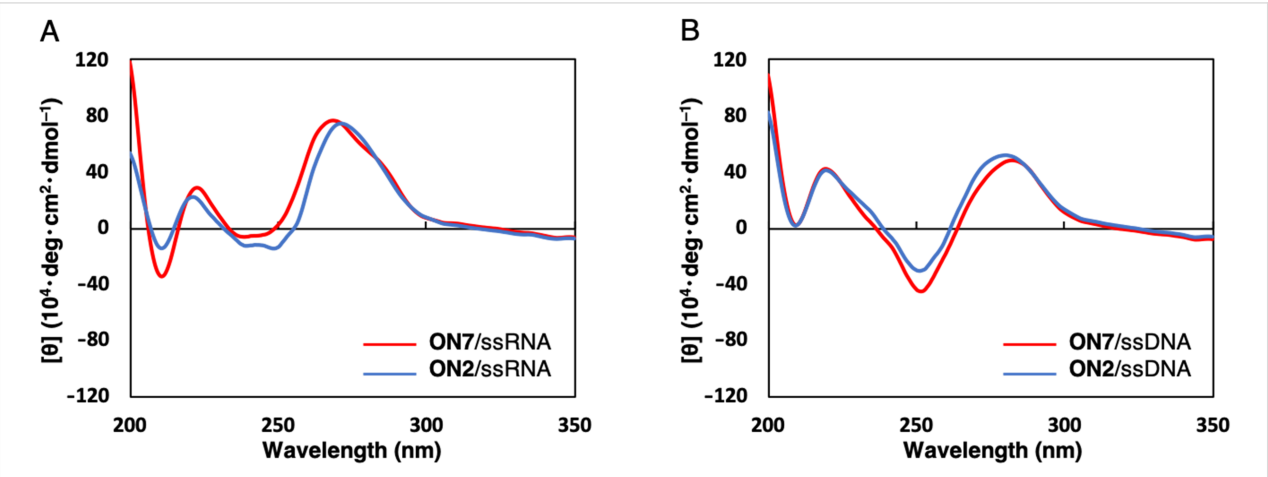


Figure 2: The CD spectra of the **ON7**/ssRNA, **ON2**/ssRNA, **ON7**/ssDNA, and **ON2**/ssDNA duplexes. Conditions: 10 mM sodium phosphate buffer (pH 7.2), 100 mM NaCl, 4 μM each oligonucleotide. Sequences of the complementary ssRNA and ssDNA are 5'-r(AGC AAA CAA CGC)-3' and 5'-d(AGC AAA CAA CGC)-3', respectively.

Experimental

Chemicals and instrumentation

All moisture-sensitive reactions were carried out in well-dried glassware under N_2 atmosphere. Dehydrated acetonitrile, dichloromethane, and tetrahydrofuran (THF) were used as purchased. ^{1}H, ^{13}C, and ^{31}P NMR spectra were recorded using a JEOL JNM-ECS300 spectrometer. The chemical shift values are expressed in δ values (ppm) relative to tetramethylsilane as an internal standard, $CHCl_3$ (δ = 7.26 ppm) for ^{1}H NMR, $CDCl_3$ (δ = 77.0 ppm) for ^{13}C NMR, and 5% H_3PO_4 (δ = 0 ppm) for ^{31}P NMR. Infrared (IR) spectra were recorded using a JASCO FT/IR-4200 spectrometer. The optical rotation was recorded using a JASCO P-2200 instrument. A MALDI–TOF mass spectrometer (SpiralTOF JMS-S3000) was used to measure the mass spectra of all compounds. For column chromatography, silica gel PSQ 60B or 100B was used. The progress of the reactions was monitored by analytical thin-layer chromatography on glass plates (TLC Silica gel 60 F_{254}), and the products were visualized using UV light.

Synthesis of phosphoramidites

(1*R*,3*R*,4*R*,7*S*)-5-(*N'*-Acetyl-*N*-methylcarbamimidoyl)-3-(*N*^{6}-benzoyladenine-9-yl)-1-(4,4'-dimethoxytrityl)oxymethyl-2-oxa-5-azabicyclo[2.2.1]heptan-7-ol (2a): This compound was synthesized in a similar manner as described in reference [20]. To the mixture of compound **1a** (841 mg, 1.23 mmol) and *N*-acetyl-*S*,*N'*-dimethylisothiourea (271 mg, 1.84 mmol), anhydrous THF (12 mL) was added and the mixture placed in an ice bath under stirring. Subsequently, *N*,*N*-diisopropylethylamine (0.35 mL, 2.0 mmol) and silver triflate (507 mg, 1.97 mmol) were added, and the mixture was stirred at room temperature overnight. Upon completion of the reaction, the mixture was diluted with ethyl acetate, after which sat. aq. NaCl was added. Following filtration, the product was extracted with ethyl acetate, washed with water and brine, dried (using Na_2SO_4), and concentrated. The product was then purified using column chromatography to yield **2a** (698 mg, 72%) as a yellow solid substance. **2a**: $[\alpha]_D^{23}$ −26.4 (*c* 1.0, $CHCl_3$); IR (KBr): 2999, 2952, 2837, 1696, 1606, 1509, 1451, 1410, 1297, 1251, 1177, 1155, 1074, 1035 cm^{-1}; ^{1}H NMR ($CDCl_3$) δ 2.00 (s, 3H), 2.80 (s, 3H), 3.48, 3.57 (AB, *J* = 10.7 Hz, 2H), 3.67 (s, 2H), 3.74 (s, 3H), 3.74 (s, 3H), 4.37 (s, 1H), 5.02 (s, 1H), 6.14 (s1H), 6.79 (d, *J* = 8.9 Hz, 4H), 7.15–7.33 (m, 7H), 7.42–7.61 (m, 5H), 7.98 (dd, *J* = 1.4 Hz, 7.2 Hz, 2H), 8.25 (s, 1H), 8.70 (s, 1H), 9.15 (s, 1H); ^{13}C NMR ($CDCl_3$) δ 26.44, 29.70, 54.23, 55.16, 60.27, 63.88, 71.06, 85.51, 86.46, 88.50, 113.18, 123.47, 126.95, 127.86, 127.90, 128.03, 128.84, 129.94, 130.01, 132.87, 133.31, 135.21, 135.47, 140.48, 144.26, 149.38, 150.80, 152.55, 158.52, 164.76; HRMS–MALDI (*m*/*z*): $[M + Na]^+$ calcd for $C_{43}H_{42}N_8O_7Na$, 805.3069; found, 805.3063.

(1*R*,3*R*,4*R*,7*S*)-5-(*N'*-Acetyl-*N*-methylcarbamimidoyl)-1-(4,4'-dimethoxytrityl)oxymethyl-3-(*O*^{6}-diphenylcarbamoyl-*N*^{2}-isobutyrylguanine-9-yl)-2-oxa-5-azabicyclo[2.2.1]heptan-7-ol (2b): This compound was synthesized in a similar manner as described in reference [20]. To the mixture of compound **1b** (2.15 g, 2.49 mmol) and *N*-acetyl-*S*,*N'*-dimethylisothiourea (402 mg, 2.75 mmol), anhydrous THF (25 mL) was added and the mixture placed in an ice bath under stirring. Subsequently, *N*,*N*-diisopropylethylamine (0.57 mL, 3.3 mmol) and silver triflate (835 mg, 3.25 mmol) were added, and the mixture was stirred at room temperature for 2 h. Upon completion of the reaction, the mixture was diluted with ethyl acetate and washed with sat. aq. $NaHCO_3$, after which sat. aq. NH_4Cl was added. Following filtration, the product was extracted with ethyl acetate, washed with water and brine, dried (using Na_2SO_4), and concentrated. The product was purified using column chromatography to yield **2b** (1.56 g, 65%) as a yellow solid substance. **2b**: $[\alpha]_D^{23}$ −15.2 (*c* 1.0, $CHCl_3$); IR (KBr): 3350, 2971, 2837, 1750, 1712, 1587, 1509, 1444, 1411, 1335, 1284, 1249, 1226, 1176, 1116, 1068, 1035 cm^{-1}; ^{1}H NMR ($CDCl_3$) δ 1.23 (d, *J* = 6.5 Hz, 3H), 1.25 (d, *J* = 6.2 Hz, 3H), 2.05 (s, 3H), 2.55–2.66 (m, 1H), 3.05 (d, *J* = 4.2 Hz, 3H), 3.48, 3.53 (AB, *J* = 10.8 Hz, 2H), 3.59, 3.74 (AB, *J* = 10.3 Hz, 2H), 3.78 (s, 3H), 3.78 (s, 3H), 4.29 (s, 1H), 5.07 (s, 1H), 6.00 (s, 1H), 6.83 (d, *J* = 7.9 Hz, 4H), 7.16–7.45 (m, 18H), 8.13 (s, 1H), 8.16 (s, 1H); ^{13}C NMR ($CDCl_3$) δ 19.19, 19.34, 26.03, 29.96, 36.61, 54.14, 55.14, 60.15, 63.64, 72.12, 85.49, 86.35, 88.17, 113.19, 121.57, 126.91, 127.90, 128.02, 129.18, 129.94, 129.99, 135.31, 135.52, 141.55, 144.34, 150.26, 151.60, 153.12, 156.07, 158.50, 162.61, 174.91; HRMS (MALDI) (*m*/*z*): $[M + Na]^+$ calcd. for $C_{53}H_{53}N_9O_9Na$, 982.3858; found, 982.3856.

(1*R*,3*R*,4*R*,7*S*)-5-(*N'*-Acetyl-*N*-methylcarbamimidoyl)-1-(4,4'-dimethoxytrityl)oxymethyl-3-(*O*^{6}-diphenylcarbamoyl-*N*^{2}-isobutyrylguanine-9-yl)-2-oxa-5-azabicyclo[2.2.1]heptan-7-ol (2c): This compound was synthesized in a similar manner as described in reference [20]. To the mixture of compound **1c** (679 mg, 1.01 mmol) and *N*-acetyl-*S*,*N'*-dimethylisothiourea (194 mg, 1.33 mmol), anhydrous THF (10 mL) was added and the mixture placed in an ice bath under stirring. Subsequently, *N*,*N*-diisopropylethylamine (0.28 mL, 1.6 mmol) and silver triflate (411 mg, 1.60 mmol) were added, and the mixture was stirred at room temperature for 1 h. Upon completion of the reaction, the mixture was diluted with ethyl acetate, after which sat. aq. NH_4Cl was added. Following filtration, the product was extracted with ethyl acetate, washed with water and brine, dried (using Na_2SO_4), and concentrated. The product was purified using column chromatography to yield **2c** (641 mg, 83%) as a white solid substance. **2c**: ^{1}H NMR ($CDCl_3$) δ 1.82 (s, 3H), 2.03 (d, *J* = 4.2 Hz, 3H), 2.81 (s, 3H), 3.31, 3.51 (AB, *J* = 9.7 Hz, 2H), 3.48, 3.57 (AB, *J* = 10.9 Hz,

2H), 3.76 (s, 3H), 3.77 (s, 3H), 4.31 (s, 1H), 4.60 (s, 1H), 5.56 (s, 1H), 6.83 (dd, J = 8.6 Hz, 1.7 Hz, 4H), 7.21–7.55 (m, 12H), 7.72 (s, 1H), 8.30 (d, J = 7.2 Hz, 2H); ^{13}C NMR ($CDCl_3$) δ 13.62, 26.30, 29.47, 53.81, 54.04, 55.17, 58.97, 63.18, 69.93, 86.08, 86.66, 88.70, 111.89, 113.24, 127.03, 128.00, 128.07, 129.86, 129.94, 130.11, 132.52, 135.14, 135.57, 136.87, 144.23, 147.92, 158.59, 159.66, 161.16, 179.48; HRMS–MALDI (*m*/*z*): $[M + Na]^+$ calcd for $C_{43}H_{44}N_6O_8Na$, 795.3113; found, 795.3106.

(1*R*,3*R*,4*R*,7*S*)-5-(*N'*-Acetyl-*N*-methylcarbamimidoyl)-3-(*N*^{6}-benzoyladenine-9-yl)-7-[2-cyanoethoxy(diisopropylamino)phosphanyl]oxyl-1-(4,4'-dimethoxytrityl)oxymethyl-2-oxa-5-azabicyclo[2.2.1]heptane (3a): This compound was synthesized in a similar manner as described in reference [20]. To a solution of **2a** (1.47 g, 1.9 mmol) in dichloromethane (19 mL), *N*,*N*-diisopropylethylamine (0.7 mL, 4.1 mmol) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (0.8 mL, 3.8 mmol) were added, and the mixture was stirred at room temperature for 6 h. Upon completion of the reaction, sat. aq. $NaHCO_3$ was added, and the product was extracted with dichloromethane. The organic phase was washed with water and brine, dried (using Na_2SO_4), and concentrated. The product was purified using column chromatography to yield **3a** (1.62 g, 87%) as a yellow solid substance. **3a**: ^{31}P NMR ($CDCl_3$) δ 149.15, 149.31; HRMS–MALDI (*m*/*z*): $[M + Na]^+$ calcd for $C_{52}H_{59}N_{10}O_8NaP$, 1005.4147; found, 1005.4143.

(1*R*,3*R*,4*R*,7*S*)-5-(*N'*-Acetyl-*N*-methylcarbamimidoyl)-7-[2-cyanoethoxy(diisopropylamino)phosphanyl]oxyl-1-(4,4'-dimethoxytrityl)oxymethyl-3-(*O*^{6}-diphenylcarbamoyl-*N*^{2}-isobutyrylguanine-9-yl)-2-oxa-5-azabicyclo[2.2.1]heptane (3b): This compound was synthesized in a similar manner as described in reference [20]. To a solution of **2b** (149 mg, 0.155 mmol) in dichloromethane (1.5 mL), *N*,*N*-diisopropylethylamine (56 µL, 0.32 mmol) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (69 µL, 0.31 mmol) were added, and the mixture was stirred at room temperature for 6 h. Upon completion of the reaction, sat. aq. $NaHCO_3$ was added, and the product was extracted with dichloromethane. The organic phase was washed with water and brine, dried (using Na_2SO_4), and concentrated. The product was purified using column chromatography to yield **3b** (117 mg, 65%) as a yellow solid substance. **3b**: ^{31}P NMR ($CDCl_3$) δ 148.80, 149.55; HRMS–MALDI (*m*/*z*): $[M + Na]^+$ calcd for $C_{62}H_{70}N_{11}O_{10}NaP$, 1182.4937; found, 1182.4955.

(1*R*,3*R*,4*R*,7*S*)-5-(*N'*-Acetyl-*N*-methylcarbamimidoyl)-7-[2-cyanoethoxy(diisopropylamino)phosphanyl]oxyl-1-(4,4'-dimethoxytrityl)oxymethyl-3-(*O*^{6}-diphenylcarbamoyl-*N*^{2}-isobutyrylguanine-9-yl)-2-oxa-5-azabicyclo[2.2.1]heptane (3c): This compound was synthesized in a similar manner as described in reference [20]. To a solution of **2c** (1.08 g, 1.40 mmol) in dichloromethane (14 mL), *N*,*N*-diisopropylethylamine (0.8 mL, 4.3 mmol) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (0.6 mL, 2.8 mmol) were added, and the mixture was stirred at room temperature for 6 h. Upon completion of the reaction, sat. aq. $NaHCO_3$ was added, and the product was extracted with dichloromethane. The organic phase was washed with water and brine, dried (using Na_2SO_4), and concentrated. The product was purified using column chromatography to yield **3c** (0.98 g, 72%) as a yellow solid substance. **3c**: ^{31}P NMR ($CDCl_3$) δ 148.61, 148.85; HRMS–MALDI (*m*/*z*): $[M + Na]^+$ calcd for $C_{52}H_{61}N_8O_9NaP$, 995.4191; found, 995.4181.

Oligonucleotide synthesis and purification

The synthesis of the oligonucleotides modified with GuNA[Me]-A, -G, or -^{m}C (0.2 µmol scale) was performed using the nS-8 oligonucleotide synthesizer (GeneDesign, Inc.) according to the standard phosphoramidite protocol with 0.5 M 5-ethylthiotetrazole as an activator. The protocol is similar to that described in reference [20]. A Custom Primer Support™ T 40s (GE Healthcare) was used as a solid support. The amidite solution was dehydrated. The standard synthesis cycle was used for the assembly of the reagents except that the coupling time was extended to 16 min. The synthesis was carried out in the trityl-on mode. The oligonucleotides were treated with a 1:1 mixture of 7 N ammonia solution in methanol and 40% aq. methylamine at room temperature for 10 h to remove the solid support, and then the mixture was heated at 60 °C for 10 h (^{m}C) or 15 h (A and G). After deprotection, the oligonucleotides were rapidly purified using a Sep-Pac® Plus C18 Cartridge. Subsequently, the desired oligonucleotides were further purified using reversed-phase HPLC with Waters XBridge™ C18 (4.6 × 50 mm analytical and 10 mm × 50 mm preparative) columns, with a linear gradient of MeCN (2.5–5% over 5 min, then 5–7.5% over 20 min) in 0.1 M triethylammonium acetate buffer (pH 7.0). The purity and structure of the oligonucleotides were confirmed by HPLC and MALDI–TOF mass spectrometry, respectively.

UV melting experiments and melting profiles

The UV melting experiments were carried out using SHIMADZU UV-1650PC and SHIMADZU UV-1800 spectrometers equipped with a T_m analysis accessory. Equimolecular amounts of the target ssRNAs or ssDNAs and the oligonucleotides were dissolved in 10 mM sodium phosphate buffer (pH 7.2) containing 100 mM NaCl to achieve a final strand concentration of 4 µM. The samples were annealed by heating at 95 °C followed by slow cooling to room temperature. The melting profile was recorded at 260 nm from 0 to 90 °C at a

scan rate of 0.5 °C/min. The T_m values were taken as the temperatures at which the formed duplexes were half dissociated, determined by the midline of the melting curves.

CD spectrum measurement

The CD spectra were recorded at 10 °C in a quartz cuvette of 1 cm optical path length. The samples were prepared in the same manner as described in the UV melting experiments. The molar ellipticity was calculated from the equation $[\theta] = \theta/cl$, where θ, c, and l indicate the relative intensity, sample concentration, and path length in centimeters, respectively.

Supporting Information

Supporting Information File 1

^{1}H, ^{13}C, and ^{32}P NMR spectra for all new compounds, HPLC charts and MALDI–TOF mass data for all new oligonucleotides, UV melting curves of the duplexes formed between GuNA[Me]-modified oligonucleotides and ssDNAs (or ssRNAs), and CD spectra of **ON4**/ssRNA and **ON4**/ssDNA.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-17-54-S1.pdf]

Funding

This work was supported in part by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant Number 20K05748 and by the Japan Agency for Medical Research and Development (AMED) under Grant Numbers JP19am0101084, JP18am0301004, and 19am0401003.

ORCID® iDs

Takao Yamaguchi - https://orcid.org/0000-0003-3180-0257
Satoshi Obika - https://orcid.org/0000-0002-6842-6812

References

1. Kaur, H.; Babu, B. R.; Maiti, S. *Chem. Rev.* **2007,** *107,* 4672–4697. doi:10.1021/cr050266u
2. Amodio, N.; Stamato, M. A.; Juli, G.; Morelli, E.; Fulciniti, M.; Manzoni, M.; Taiana, E.; Agnelli, L.; Cantafio, M. E. G.; Romeo, E.; Raimondi, L.; Caracciolo, D.; Zuccalà, V.; Rossi, M.; Neri, A.; Munshi, N. C.; Tagliaferri, P.; Tassone, P. *Leukemia* **2018,** *32,* 1948–1957. doi:10.1038/s41375-018-0067-3
3. Papachristodoulou, A.; Silginer, M.; Weller, M.; Schneider, H.; Hasenbach, K.; Janicot, M.; Roth, P. *Clin. Cancer Res.* **2019,** *25,* 7189–7201. doi:10.1158/1078-0432.ccr-17-3024
4. Chery, J.; Petri, A.; Wagschal, A.; Lim, S.-Y.; Cunningham, J.; Vasudevan, S.; Kauppinen, S.; Näär, A. M. *Nucleic Acid Ther.* **2018,** *28,* 273–284. doi:10.1089/nat.2018.0722
5. Javanbakht, H.; Mueller, H.; Walther, J.; Zhou, X.; Lopez, A.; Pattupara, T.; Blaising, J.; Pedersen, L.; Albæk, N.; Jackerott, M.; Shi, T.; Ploix, C.; Driessen, W.; Persson, R.; Ravn, J.; Young, J. A. T.; Ottosen, S. *Mol. Ther.–Nucleic Acids* **2018,** *11,* 441–454. doi:10.1016/j.omtn.2018.02.005
6. Shimo, T.; Tachibana, K.; Saito, K.; Yoshida, T.; Tomita, E.; Waki, R.; Yamamoto, T.; Doi, T.; Inoue, T.; Kawakami, J.; Obika, S. *Nucleic Acids Res.* **2014,** *42,* 8174–8187. doi:10.1093/nar/gku512
7. Seth, P. P.; Siwkowski, A.; Allerson, C. R.; Vasquez, G.; Lee, S.; Prakash, T. P.; Kinberger, G.; Migawa, M. T.; Gaus, H.; Bhat, B.; Swayze, E. E. *Nucleic Acids Symp. Ser.* **2008,** *52,* 553–554. doi:10.1093/nass/nrn280
8. Seth, P. P.; Vasquez, G.; Allerson, C. A.; Berdeja, A.; Gaus, H.; Kinberger, G. A.; Prakash, T. P.; Migawa, M. T.; Bhat, B.; Swayze, E. E. *J. Org. Chem.* **2010,** *75,* 1569–1581. doi:10.1021/jo902560f
9. Seth, P. P.; Siwkowski, A.; Allerson, C. R.; Vasquez, G.; Lee, S.; Prakash, T. P.; Wancewicz, E. V.; Witchell, D.; Swayze, E. E. *J. Med. Chem.* **2009,** *52,* 10–13. doi:10.1021/jm801294h
10. Carroll, J. B.; Warby, S. C.; Southwell, A. L.; Doty, C. N.; Greenlee, S.; Skotte, N.; Hung, G.; Bennett, C. F.; Freier, S. M.; Hayden, M. R. *Mol. Ther.* **2011,** *19,* 2178–2185. doi:10.1038/mt.2011.201
11. Pandey, S. K.; Wheeler, T. M.; Justice, S. L.; Kim, A.; Younis, H. S.; Gattis, D.; Jauvin, D.; Puymirat, J.; Swayze, E. E.; Freier, S. M.; Bennett, C. F.; Thornton, C. A.; MacLeod, A. R. *J. Pharmacol. Exp. Ther.* **2015,** *355,* 329–340. doi:10.1124/jpet.115.226969
12. Yamamoto, T.; Yahara, A.; Waki, R.; Yasuhara, H.; Wada, F.; Harada-Shiba, M.; Obika, S. *Org. Biomol. Chem.* **2015,** *13,* 3757–3765. doi:10.1039/c5ob00242g
13. Johannsen, M. W.; Crispino, L.; Wamberg, M. C.; Kalra, N.; Wengel, J. *Org. Biomol. Chem.* **2011,** *9,* 243–252. doi:10.1039/c0ob00532k
14. Kumar, R.; Ries, A.; Wengel, J. *Molecules* **2017,** *22,* 852. doi:10.3390/molecules22050852
15. Prakash, T. P.; Püschl, A.; Lesnik, E.; Mohan, V.; Tereshko, V.; Egli, M.; Manoharan, M. *Org. Lett.* **2004,** *6,* 1971–1974. doi:10.1021/ol049470e
16. Brzezinska, J.; D'Onofrio, J.; Buff, M. C. R.; Hean, J.; Ely, A.; Marimani, M.; Arbuthnot, P.; Engels, J. W. *Bioorg. Med. Chem.* **2012,** *20,* 1594–1606. doi:10.1016/j.bmc.2011.12.024
17. Deglane, G.; Abes, S.; Michel, T.; Prévot, P.; Vives, E.; Debart, F.; Barvik, I.; Lebleu, B.; Vasseur, J.-J. *ChemBioChem* **2006,** *7,* 684–692. doi:10.1002/cbic.200500433
18. Barman, J.; Gurav, D.; Oommen, O. P.; Varghese, O. P. *RSC Adv.* **2015,** *5,* 12257–12260. doi:10.1039/c4ra14721a
19. Shrestha, A. R.; Kotobuki, Y.; Hari, Y.; Obika, S. *Chem. Commun.* **2014,** *50,* 575–577. doi:10.1039/c3cc46017g
20. Horie, N.; Kumagai, S.; Kotobuki, Y.; Yamaguchi, T.; Obika, S. *Org. Biomol. Chem.* **2018,** *16,* 6531–6536. doi:10.1039/c8ob01307a
21. Koshkin, A. A.; Singh, S. K.; Nielsen, P.; Rajwanshi, V. K.; Kumar, R.; Meldgaard, M.; Olsen, C. E.; Wengel, J. *Tetrahedron* **1998,** *54,* 3607–3630. doi:10.1016/s0040-4020(98)00094-5
22. Mitsuoka, Y.; Kodama, T.; Ohnishi, R.; Hari, Y.; Imanishi, T.; Obika, S. *Nucleic Acids Res.* **2009,** *37,* 1225–1238. doi:10.1093/nar/gkn1062
23. Horiba, M.; Yamaguchi, T.; Obika, S. *J. Org. Chem.* **2016,** *81,* 11000–11008. doi:10.1021/acs.joc.6b02036
24. Sawamoto, H.; Arai, Y.; Yamakoshi, S.; Obika, S.; Kawanishi, E. *Org. Lett.* **2018,** *20,* 1928–1931. doi:10.1021/acs.orglett.8b00476

25. Kumagai, S.; Sawamoto, H.; Takegawa-Araki, T.; Arai, Y.; Yamakoshi, S.; Yamada, K.; Ohta, T.; Kawanishi, E.; Horie, N.; Yamaguchi, T.; Obika, S. *Org. Biomol. Chem.* **2020,** *18,* 9461–9472. doi:10.1039/d0ob01970d
26. Umemoto, T.; Masada, S.; Miyata, K.; Ogasawara-Shimizu, M.; Murata, S.; Nishi, K.; Ogi, K.; Hayase, Y.; Cho, N. *Tetrahedron* **2017,** *73,* 1211–1218. doi:10.1016/j.tet.2017.01.010
27. Fujisaka, A.; Hari, Y.; Takuma, H.; Rahman, S. M. A.; Yoshikawa, H.; Pang, J.; Imanishi, T.; Obika, S. *Bioorg. Med. Chem.* **2019,** *27,* 1728–1741. doi:10.1016/j.bmc.2019.02.034
28. Michael, J. D.; Ross, B. C.; Rees, P. M. *Tetrahedron Lett.* **1985,** *26,* 4149–4152. doi:10.1016/s0040-4039(00)89316-8

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

DNA with zwitterionic and negatively charged phosphate modifications: Formation of DNA triplexes, duplexes and cell uptake studies

Yongdong Su[1], Maitsetseg Bayarjargal[1], Tracy K. Hale[1,2] and Vyacheslav V. Filichev[*1,2,§]

Full Research Paper

Open Access

Address:
[1]School of Fundamental Sciences, Massey University, Private Bag 11-222, 4442 Palmerston North, New Zealand and [2]Maurice Wilkins Centre for Molecular Biodiscovery, Auckland 1142, New Zealand

Email:
Vyacheslav V. Filichev[*] - v.filichev@massey.ac.nz

* Corresponding author
§ Fax: (+) 64 6 3557953

Keywords:
cell uptake; charge neutral modification; DNA; modified phosphates; Staudinger reaction

Beilstein J. Org. Chem. **2021,** *17,* 749–761.
https://doi.org/10.3762/bjoc.17.65

Received: 08 December 2020
Accepted: 11 March 2021
Published: 29 March 2021

This article is part of the thematic issue "Celebrating the role of chemistry in the success of oligonucleotides as therapeutics".

Guest Editors: P. Kumar and T. Brown

Abstract

Two phosphate modifications were introduced into the DNA backbone using the Staudinger reaction between the 3',5'-dinucleoside β-cyanoethyl phosphite triester formed during DNA synthesis and sulfonyl azides, 4-(azidosulfonyl)-*N,N,N*-trimethylbutan-1-aminium iodide (N+ azide) or *p*-toluenesulfonyl (tosyl or Ts) azide, to provide either a zwitterionic phosphoramidate with N+ modification or a negatively charged phosphoramidate for Ts modification in the DNA sequence. The incorporation of these N+ and Ts modifications led to the formation of thermally stable parallel DNA triplexes, regardless of the number of modifications incorporated into the oligodeoxynucleotides (ONs). For both N+ and Ts-modified ONs, the antiparallel duplexes formed with complementary RNA were more stable than those formed with complementary DNA (except for ONs with modification in the middle of the sequence). Additionally, the incorporation of N+ modifications led to the formation of duplexes with a thermal stability that was less dependent on the ionic strength than native DNA duplexes. The thermodynamic analysis of the melting curves revealed that it is the reduction in unfavourable entropy, despite the decrease in favourable enthalpy, which is responsible for the stabilisation of duplexes with N+ modification. N+ONs also demonstrated greater resistance to nuclease digestion by snake venom phosphodiesterase I than the corresponding Ts-ONs. Cell uptake studies showed that Ts-ONs can enter the nucleus of mouse fibroblast NIH3T3 cells without any transfection reagent, whereas, N+ONs remain concentrated in vesicles within the cytoplasm. These results indicate that both N+ and Ts-modified ONs are promising for various in vivo applications.

Introduction

The ability to detect and modify the genome of living organisms is important for the diagnosis, prevention, and treatment of many diseases [1]. The site-specific targeting and manipulation of genomic DNA or RNA using chemically modified short oligodeoxynucleotides (ONs) is considered to be a viable therapeutic strategy [2-5]. Antigene strategies use ONs to specifi-

cally bind native DNA, induce genomic changes, and/or interfere with gene expression. Apart from strategies that use modular enzymes such as zinc-finger nucleases [6] or transcription activator-like effector nucleases (TALENs) [7] to recognise and cut DNA sequences, or CRISPR-CAS9 [8-10] and CAS9-constructs [11-14] which rely on large proteins to open the target duplex, triplex-forming oligonucleotides (TFOs) [15] can be designed to bind in a sequence-specific manner to double-stranded DNA (dsDNA) [16]. In forming the parallel triple-helix structure, a polypyrimidine TFO binds to dsDNA through Hoogsteen base-pairing [17], in which the cytosine bases in the TFO are protonated at the N3 atom (Figure 1B).

In antisense strategies, antisense ONs (AOs) interact with RNA molecules to interfere with protein expression [18,19]. The major challenge in designing chemically modified ONs as antigene/antisense agents is to ensure an efficient cellular uptake and nuclease resistance while still maintaining, or ideally increasing, binding affinity and specificity of the ONs towards their DNA or RNA target.

Many synthetic analogues of natural ONs, such as peptide nucleic acids (PNA) [20], locked nucleic acids [21] (LNA, also known as bridged nucleic acids (BNA) [22]) and phosphorothioate (PS) ONs [23,24] have been evaluated for antigene/antisense applications, however, each of the analogues did not meet all the requirements. For example, both PNA and modified PNAs have excellent chemical stability, are resistant to enzymatic degradation, and have high binding affinity towards complementary DNA and RNA, but have a tendency to aggregate, require high salt conditions, and have low solubility in water [1,25,26]. LNA (BNA) have an enhanced thermal stability in DNA triplexes and duplexes, a high binding affinity to RNA, and are nuclease resistant [22,26-28]. These properties have led to LNA (BNA) being used in various therapeutic ONs that have reached clinical trials [29]. However, the multistep synthesis of LNA and increased hepatotoxicity of some modified AOs ensure that further optimisation is required [30]. Chemical modification of ONs with a PS linkage resulted in ONs resistant to nuclease degradation but with several side effects due to non-specific interactions with cellular components [31].

Modifications of the phosphate backbone of DNA and RNA, especially charge neutral modifications, have gained attention in recent years because such modifications not only improve the nuclease resistance of ONs but also enhance their affinity towards complementary DNA/RNA/dsDNA and improve cell uptake. The lack of a negatively charged backbone also improved the binding of PNA to DNA or RNA strands. It has been shown that positively charged PNA bind more strongly to DNA

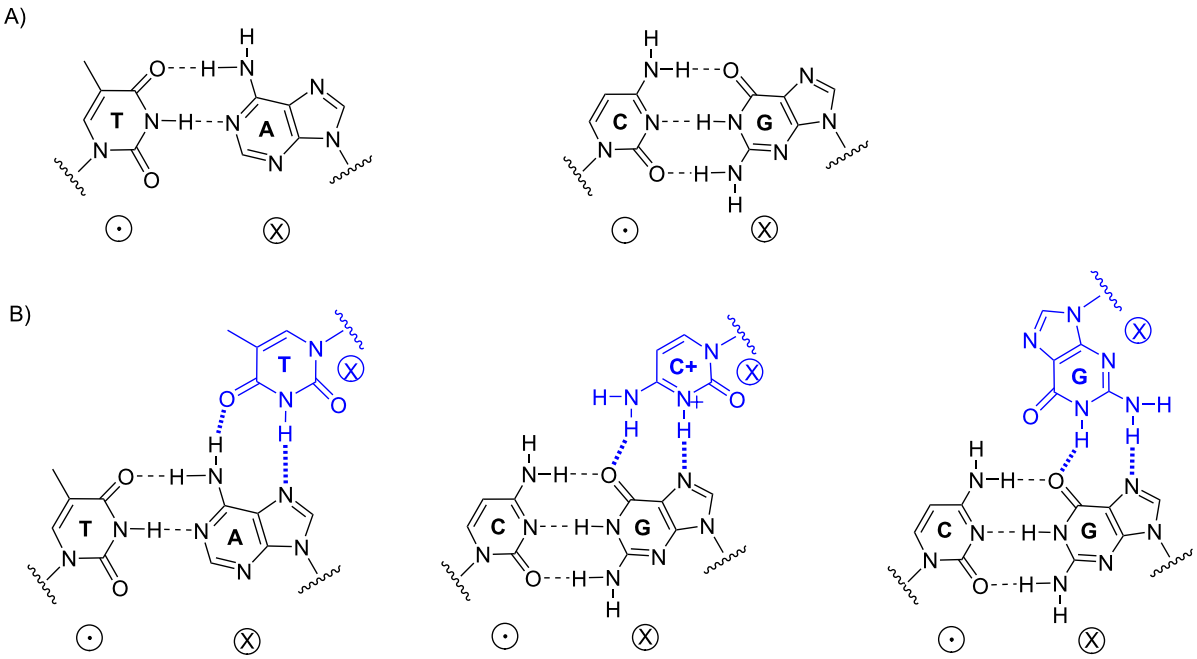


Figure 1: Illustration of H-bonding in a DNA duplex and a parallel triplex. A) Depiction of Watson–Crick base-paring (left: T-A and right: C-G); B) parallel triple helices: pyrimidine-rich third strand interactions are stabilised by Hoogsteen hydrogen bonds (the duplex is in black, TFO is in blue, Watson–Crick base-paring is shown with dashed bonds, and Hoogsteen base-paring is shown with hashed bonds). The relative orientation of phosphodiester backbones is indicated by the symbols "⊙" and "Ⓧ".

and RNA than negatively charged PNA at low salt concentrations (0–100 mM Na^+) whereas at medium to high salt concentrations (250–1000 mM Na^+) the trend is reversed [32].

As a charge-neutral phosphate mimic, the methylphosphonate linkage (PMe) has been introduced into the DNA backbone to improve stability of ONs towards enzymic digestion as well as DNA duplex and triplex binding affinity [33]. However, the poor aqueous solubility [34], reduced binding affinity with complementary RNA [35], and a destabilising effect on the thermal stability of G-quadruplexes [36] hinders its application. In contrast, a phosphate methylated linkage (POMe,) marginally destabilised complementary DNA but improved sequence specificity [37].

Recently, we synthesised a G-rich ON (TG_4T) with all phosphates replaced by a 4-(trimethylammonio)butylsulfonyl phosphoramidate group (N+, Scheme 1). The sequence was designed to obtain the formally charge-neutral zwitterionic N+TG_4T [38]. Each negatively charged phosphoramidate is neutralised by the positively charged quaternary ammonium group, providing a zwitterionic phosphate mimic. The resistance to enzymatic degradation, a higher thermal stability, and a faster association that was independent of ionic strength was observed for this N+-modified G-quadruplex $(TG_4T)_4$. These properties encouraged us to evaluate the N+ modification in the context of DNA duplexes and triplexes and to perform cell-uptake studies. For comparison, we also evaluated the properties of ONs modified with a tosyl sulfonyl phosphoramidate (Ts) that results in a negatively charged phosphate mimic [39].

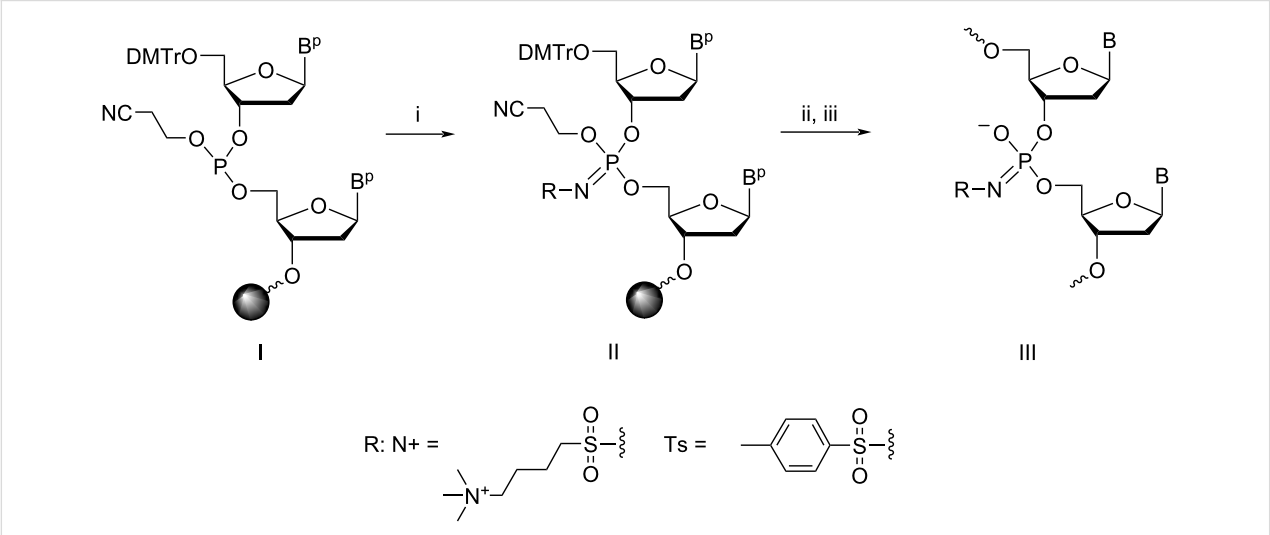
The introduction of each, N+ or Ts modification creates a chiral center at the phosphorus atom resulting in a mixture of 2^n diastereomers, where n is the number of modified phosphate groups. The reverse-phase (RP) HPLC purification occasionally results in the separation of individual diastereomers (usually for ONs with a single modification), which supports the insignificant difference in the lipophilicity of the diastereomeric ONs. We hypothesised that, in comparison with native ONs, the N+ONs should hybridise with higher affinity to complementary single-stranded DNA (ssDNA) or RNA due to both a reduced repulsion between negatively charged phosphates and a thermal stability being less dependent on the ionic strength of the solution. Moreover, N+ONs carrying zwitterionic phosphates could lead to an increased cell uptake.

For both N+ and Ts modifications, we synthesised 14-mer ONs with either one, two, three, or four modifications introduced at various positions in the sequence. The thermal stability of a parallel DNA triplex and duplexes of DNA and RNA formed with these ONs where then evaluated. Thermal denaturation experiments, nuclease resistance and cell-uptake assays were also conducted to evaluate these chemically modified ONs.

Results

Synthesis and purification of modified ONs

4-(Azidosulfonyl)-*N,N,N*-trimethylbutan-1-aminium iodide [38] and tosyl azide (*p*-toluenesulfonyl azide, TsN_3) [39] were synthesised and used for the synthesis of the modified ONs using an automated DNA synthesiser as described. The solution of sulfonyl azide (0.5 M TsN_3 in MeCN or 0.7 M 4-(azidosulfonyl)-*N,N,N*-trimethylbutan-1-aminium iodide in DMF) was



Scheme 1: The synthesis of ONs with Ts and N+ modification using the Staudinger reaction during the solid-phase DNA synthesis. Conditions: (i) 0.5 M TsN_3, MeCN, 37 °C, 30 min for Ts modification; 0.7 M 4-(azidosulfonyl)-*N,N,N*-trimethylbutan-1-aminium iodide, DMF, 37 °C, 30 min for the N+ modification; (ii) DNA synthesis; (iii) conc. aq NH_3, 55 °C, 12 h; B^p/B: protected/deprotected heterocyclic base; DMTr: 4,4'-dimethoxytrityl.

introduced as replacement of a standard iodine/pyridine oxidation step to react with 3',5'-dinucleoside β-cyanoethyl phosphites (Scheme 1, I), forming the *N*-modified iminophosphorane (Scheme 1, II). ONs bearing one or more *N*-modified phosphoramidate groups (Scheme 1, III) were obtained after the removal of the protective groups and β-cyanoethyl groups using ≈28% ammonia. We found out that a higher conversion was observed when performing the Staudinger reaction at 37 °C rather than at room temperature. The yield was also improved by minimising the handling of the solid support and performing the reaction using a microtube pump to deliver the sulfonyl azide solution onto the column with CPG support. The cleaved and deprotected N+- and Ts-ONs were initially purified using reversed-phase (RP) HPLC. However, the separation of ONs with varying numbers of modifications was not ideal as there were only marginal changes in the retention time in RP-HPLC. Therefore, ion-exchange (IE) HPLC was used for purifying these ONs. The substitution of each phosphate with N+ modification, resulted in a shorter retention time (τ) in IE-HPLC (Δτ = −3 min/modification, Table 1). For Ts-modified ONs, the incorporation of the Ts modifications, as a result of increased hydrophobicity, led to an increased retention time (Δτ = +1.5 to +2 min/modification, Table 1) compared to the native sequence. The composition of the ONs was confirmed by electrospray ionization mass spectrometry (ESIMS) in the negative mode (Table 1). For clarity, we introduced the following nomenclature of the ONs synthesised. The prefix 5'- or 3'- with either N+ or Ts- means that the first phosphate at the 5'- or 3'-end was modified; m-N+ or m-Ts- indicates that the named modification was incorporated in the middle of the sequence; 2N+, 3N+, 4N+ or 2Ts-, 3Ts-, 4Ts- indicates that two, three, or four modifications were distributed evenly in the sequence.

The solubility of the ONs was not influenced by the introduction of Ts and N+ modifications, as the purified, desalted, and lyophilised ONs were fully dissolved in 50 µL H_2O. The

Table 1: Names of the ONs synthesised, their sequences, retention times on the ion-exchange column[a], compositions, and isolated yields.

	sequence	retention time (min)	calculated MW	observed MW[b]	isolated yield (%)
ON1	5'-CCCCTTTCTTTTTT[c]	31.53	4121.7	–	
5'-N+ON2	5'-C_{N+}CCCTTTCTTTTTT	27.76	4296.7958	4297.7588[d]	9[e]
m-N+ON3	5'-CCCCTTT$_{N+}$CTTTTTT	27.43	4296.7958	4297.7455[d]	11[e]
3'-N+ON4	5'-CCCCTTTCTTTTT$_{N+}$T	27.74	4296.7958	4297.7466[d]	20[f]
2N+ON5	5'-C_{N+}CCCTTTCTTTTT$_{N+}$T	23.42	4473.9012	4473.8248[d]	8[e]
3N+ON6	5'-C_{N+}CCCTTT$_{N+}$CTTTTT$_{N+}$T	20.77	4649.0067	4650.1412[d]	10[e]
4N+ON7	5'-C_{N+}CCT$_{N+}$TTCT$_{N+}$TTTT$_{N+}$T	17.09	4826.0965	4826.0516[d]	23[f]
4N+{FAM}	5'-C_{N+}CCT$_{N+}$TTCT$_{N+}$TTTT$_{N+}$T{FAM}	–[g]	5396.4012	5396.1460[h]	20
5'-Ts-ON8	5'-C_{Ts}CCCTTTCTTTTTT	32.80	4272.7151	4274.6664[d]	4[e]
m-Ts-ON9	5'-CCCCTTT$_{Ts}$CTTTTTT	32.75	4272.7151	4273.4728[d]	5[e]
3'-Ts-ON10	5'-CCCCTTTCTTTTT$_{Ts}$T	32.81	4272.7151	4273.4694[h]	8[e]
2Ts-ON11	5'-C_{Ts}CCCTTTCTTTTT$_{Ts}$T	35.67	4425.7398	4427.7401[h]	6[e]
3Ts-ON12	5'-C_{Ts}CCCTTT$_{Ts}$CTTTTT$_{Ts}$T	36.76	4578.7646	4578.6076[i]	30[f]
4Ts-ON13	5'-C_{Ts}CCT$_{Ts}$TTCT$_{Ts}$TTTT$_{Ts}$T	40.07	4731.7893	4731.7620[j]	39[f]
4Ts-{FAM}	5'-C_{Ts}CCT$_{Ts}$TTCT$_{Ts}$TTTT$_{Ts}$T{FAM}	–[g]	5301.2398	5302.8380[h]	26
m-N+ON14	5'-CCCCTTTCTTT$_{N+}$TTT	27.50	4296.7958	4296.7520[d]	19[f]
m-N+ON15	5'-CCCCTTT C_{N+}TTTTTT	27.43	4296.7958	4296.7300[d]	16[f]
m-N+ON16	5'-CCCC$_{N+}$TTTCTTTTTT	27.45	4296.7958	4296.7440[d]	17[f]
m-N+ON17	5'-CC$_{N+}$CCTTTCTTTTTT	27.62	4296.7958	4296.7350[d]	21[f]
3N+ON18	5'-C_{N+}CCCTT$_{N+}$TCTTTTT$_{N+}$T	20.44	4649.0067	4649.937[d]	17[f]

[a]IE-HPLC was performed on an IE-column (TSKgel Super Q-5PW) using a gradient of NaCl concentration (0 → 0.5 M) in 20 mM Tris-HCl, 1 mM EDTA, pH 9.0 over 30 min; [b]based on ESIMS in the negative mode; [c]obtained from Integrated DNA Technologies; [d]calculated for $[M - 6H]^{6-}$; [e]synthesised in a 1 µmol scale using a previously reported procedure with transferring the solid support from a column into a vial for reaction with sulfonyl azide for 30 min at room temperature. Afterwards, the solid support was transferred back to the column to continue DNA synthesis [39]. Some amount of the solid support was lost during the transfer and washing steps, especially for multiple modifications, which was the main reason for the low yields of these ONs. [f]synthesised in a 1 µmol scale following a modified procedure using a microtube pump to deliver the sulfonyl azide solution onto the column with CPG-support at 37 °C [38]; [g]synthesised in a 3–4 µmol scale, purified by 20% denaturing PAGE (7 M urea), followed by extraction from the gel and desalting. [h]for $[M - 7H]^{5-}$; [i]for $[M + K^+ - 9H]^{8-}$; [j]for $[M - 4H]^{4-}$; the ESIMS spectra are provided in Supporting Information File 1.

Ts-modified ONs have previously been shown to marginally destabilise duplexes with complementary DNA and RNA [39]. The chemical stability of **5'-N+ON2** at various pH (5.5, 7.0, and 8.5) was evaluated by incubation in 10 mM Na phosphate buffer (140 mM NaCl, 0.1 mM Na_2-EDTA) at 50 °C for 24 h. No degradation was observed according to IE-HPLC analysis (see Figure S16 in Supporting Information File 1), which ensures that the N+-modified ONs will be chemically stable during the evaluation of the thermal stability of complexes with complementary DNA and RNA.

Thermal denaturation experiments

The thermal stability of antiparallel ON/RNA and ON/DNA duplexes as well as parallel DNA triplexes was assessed in thermal denaturation experiments and the results are summarised in Table 2.

The sequences possessing a different number of N+ and Ts modifications were studied initially in an antiparallel duplex formed with complementary RNA and compared with the corresponding antiparallel DNA duplexes at pH 7.0. Apart from the ONs possessing modifications in the middle of the sequence (entries 6, 9, and 12 in Table 2), stabilised ON/RNA duplexes were obtained for both N+ONs (ΔT_m = +1 – +12 °C, entries 2–7, Table 2) and Ts-ONs (ΔT_m = +7 to +11 °C, entries 8–13, Table 2). The highest thermal stabilisation against RNA induced by a single modification was observed for ONs with one modification at 3'-end (ΔT_m = +12 °C and +11 °C for N+ and Ts modifications, respectively). The corresponding antiparallel DNA duplexes were less thermally stable with ΔT_m = −1 to +2 °C. The same trend was seen for ONs with a modification at both the 5'- and 3'- ends: ON/RNA duplexes were more stable (ΔT_m = +7 °C for **2N+ON5** and ΔT_m = +10 °C for **2Ts-ON11**) than the corresponding antiparallel DNA duplexes (ΔT_m = +2 °C for **2N+ON5** and ΔT_m = +1 °C for **2Ts-ON11**). The thermal stability of the ON/RNA duplexes was not improved by increasing the number of N+ or Ts modifications in the ONs.

Table 2: T_m [°C, ± 0.5 °C] data for triplex and duplex melting, taken from UV melting curves (λ = 260 nm).

entry		antiparallel duplex			triplex[c]	
		RNA[a]	DNA[b]			
		pH 7.0	pH 5.0	pH 7.0	pH 5.0[d]	pH 6.0
1	**ON1**	46	48	50	45	28
2	**5'-N+ON2**	53 (+ 7.0)	44 (–4.0)	51 (+1.0)	40 (–5.0)	25 (–3.0)
3	**m-N+ON3**	47 (+ 1.0)	43 (–5.0)	48 (–2.0)	55 (+10.0)	28
4	**3'-N+ON4**	58 (+ 12.0)	46 (–2.0)	52 (+2.0)	56 (+11.0)	29 (+1.0)
5	**2N+ON5**	53 (+ 7.0)	44 (–4.0)	52 (+2.0)	56 (+11.0)	28
6	**3N+ON6**	41 (– 5.0)	45 (–3.0)	51 (+1.0)	48 (+3.0)	<15
7	**4N+ON7**	55 (+ 9.0)	48 (0.0)	51 (+1.0)	48 (+3.0)	28
8	**5'-Ts-ON8**	54 (+ 8.0)	39 (–9.0)	46 (–4.0)	51 (+6.0)	24 (–4.0)
9	**m-Ts-ON9**	44 (–2.0)	31 (–17.0)	37 (–13.0)	51 (+6.0)	<15
10	**3'-Ts-ON10**	57 (+11.0)	44 (–4.0)	49 (–1.0)	54 (+9.0)	27 (–1.0)
11	**2Ts-ON11**	56 (+10.0)	43 (–5.0)	51 (+1.0)	53 (+8.0)	25 (–2.0)
12	**3Ts-ON12**	38 (–8.0)	40 (–8.0)	45 (–5.0)	49 (+4.0)	<15
13	**4Ts-ON13**	53 (+7.0)	39 (–9.0)	44 (–6.0)	47 (+2.0)	20 (–8.0)
14	**m-N+ON14**	54 (+8.0)	–[e]	51 (+1.0)	–	–
15	**m-N+ON15**	56 (+10.0)	–	51 (+1.0)	–	–
16	**m-N+ON16**	54 (+8.0)	–	55 (+5.0)	–	–
17	**m-N+ON17**	56 (+10.0)	–	52 (+2.0)	–	–
18	**3N+ON18**	54 (+8.0)	–	51 (+1.0)	–	–

[a]The RNA sequence for the antiparallel duplex formation is **ON19**: 3'-rGGGGAAAGAAAAAA; *c* = 1.0 μM of each strand in 20 mM sodium cacodylate, 100 mM NaCl, 10 mM $MgCl_2$, pH 7.0; the T_m values for the ON/RNA duplexes were confirmed by CD melting experiments (Figures S7 and S8, and Table S2 in Supporting Information File 1); [b]the DNA sequence for the antiparallel duplex formation is **ON20**: 3'-GGGGAAAGAAAAAA; *c* = 1.0 μM of each strand in 20 mM sodium cacodylate, 100 mM NaCl, 10 mM $MgCl_2$, pH 5.0 and pH 7.0; [c]*c* = 1.5 μM of **ON1–13** and 1.0 μM of each strand of dsDNA (**D1**: 3'-CTGCCCCTTTCTTTTTT/5'-GACGGGGAAAGAAAAAA) in 20 mM sodium cacodylate, 100 mM NaCl, 10 mM $MgCl_2$, pH 5.0, 6.0 and 7.0; duplex T_m = 56.5 °C (pH 5.0), 58.5 °C (pH 6.0), and 57.0 °C (pH 7.0); triplex formation was confirmed by size-exclusion HPLC (SE-HPLC) in sodium cacodylate buffer (pH 5.0 and pH 6.0, Figure S15 in Supporting Information File 1), no triplex was formed at pH 7.0; [d]the T_m for triplex melting was determined by subtraction of the duplex melting curve from the overlaid melting curve (Figure S6 in Supporting Information File 1); [e]not performed.

For Watson–Crick-type duplexes, both the N+ and Ts modifications destabilised the DNA duplex at pH 5.0. The destabilising effect was more pronounced for Ts than for the N+ modification (ΔT_m = −17 to −4 °C for Ts and −5 to 0 °C for N+ modifications, respectively). For the antiparallel ON/DNA duplexes formed at pH 7.0, when comparing ONs with the same number of modifications, the N+ modifications led to higher T_m values than Ts-ONs. The incorporation of three and four Ts moieties led to a further decrease in the T_m whereas the corresponding N+ONs did not disrupt the duplex thermal stability. These results indicate that the N+ and Ts modifications can be viewed as RNA-like modifications, because their use in ONs led to higher ΔT_m values for ON/RNA than for ON/DNA duplexes.

The same modified sequences (**ON1**–**13**) were also studied in a pH-dependent Hoogsteen-type base-pairing towards the duplex **D1** forming a parallel triplex [40]. As can be seen in Table 2, all parallel triplexes formed at pH 5.0 were more thermally stable than at pH 6.0 and no triplex was formed at pH 7.0, which is consistent with the trend for parallel triplexes based on CT-TFOs [41]. Some fluctuations were observed for Hoogsteen-type triplexes formed by N+ONs. A modification at the 5' end destabilised triplexes at both pH 5.0 and 6.0 (ΔT_m = −5 °C and −3 °C, respectively, Table 2). All other N+ONs formed more stable triplexes with **D1** at pH 5.0, while marginal changes were observed for triplexes at pH 6.0 except for **3N+ON6** with three modifications that did not form a triplex. The incorporation of Ts modifications led to stabilised Hoogsteen-type triplexes at pH 5.0 (ΔT_m = +2 to +9 °C, Table 2), whereas triplexes at pH 6.0 were less stable (ΔT_m = −1 to −8 °C). For Ts-ONs with the modification in the center of the sequence (**m-Ts-ON9** and **3Ts-ON12**), no triplex formation was observed at pH 6.0/room temperature. These results show that Hoogsteen-type triplexes with single N+ or Ts modifications at the 3'-end are more thermally stable at the 5'-end, and that increasing the number of modifications showed no advantage for T_m of parallel triplexes.

A position-dependent influence of the N+ and Ts moieties on the T_m is suggested by the less thermally stable duplexes formed by the ONs with a single modification in the middle of the sequence (in TC motif) compared to native DNA. We synthesised another set of N+ONs, with single (**ON14–17**, entries 14–17, Table 2) and triple modifications (**ON18**, entry 18, Table 2) that had no modifications in the center of the sequence and evaluated the thermal stability of their antiparallel duplexes formed with complementary RNA and DNA at pH 7.0. The results in Table 2 show that **ON14–17** form more stable duplexes with RNA (ΔT_m = +8 to +10 °C) and DNA (ΔT_m = +1 to +5 °C). It is interesting that sequences with the N+ modification in the CT motif (**m-N+ON15** and **m-N+ON16**) did not destabilise the antiparallel duplexes unlike the N+ modification in the TC motif (**m-N+ON3** and **3N+ON6**). One possible reason for this position-dependent influence of the N+ and Ts moieties on the duplex stability might be due to a propeller twist [42] in the TC dinucleotide interfering with the N+ and Ts moieties and destabilising the DNA and RNA duplexes.

Next, we evaluated the binding affinity of the N+ and Ts-modified ONs for complementary DNA and RNA at different salt concentrations (25, 50, and 100 mM NaCl, Table 3). It has been reported that the thermal stability of DNA duplexes decreases as salt concentrations are reduced due to the increased electrostatic repulsion between the negatively charged phosphates [43]. The native DNA duplex **ON1**/**ON20** showed a decline in the T_m values from 50 to 37 °C and to 19 °C, when the NaCl concentration was decreased from 100 to 50 mM and to 25 mM, respectively (Table 3). A similar trend was observed for the Ts-modified ONs as the backbone is still negatively charged. In contrast, for the N+ONs the decrease in the T_m with decreasing NaCl concentration was not as significant as for the negatively charged ONs, and the T_m value at a 25 mM salt concentration was 20 °C higher than the T_m for the unmodified duplex, and 12 °C higher than the Ts-modified duplex. However, such behaviour was not as noticeable for the duplexes formed with complementary RNA. The T_m value of the control ON/RNA duplex (**ON1/ON19**) decreased by 2 °C when the NaCl concentration was reduced from 100 to 25 mM, whereas the T_m for the N+ON/RNA duplex (**4N+ON7/ON19**) decreased by 10 °C, although it was still more thermally stable than the control ON/RNA duplex. In contrast, the duplex formed by Ts-ON and RNA (**4Ts-ON13/ON19)** was destabilised (ΔT_m = −6 °C) at the lowest salt concentration tested.

We analysed the melting profiles and obtained the thermodynamic parameters of the duplexes at different salt concentrations ignoring the changes in DNA and salt concentrations induced by solution evaporation, the change of pH during heating, and assuming that there is no change in the heat capacity ($\Delta c_p = 0$) [45]. We also assumed a two-state transition between duplex and single-stranded DNA and a linear relationship between the CD/UV signal and fraction of molecules unfolded (see Supporting Information File 1 for the analysis of the melting curves). As the thermal stability of DNA duplexes increased with increasing concentrations of salt, we expected a favourable ΔH of duplex formation at a higher salt concentration. However, in all cases studied, ΔH was less favourable at 100 mM than at 25 mM NaCl. Recently, similar observations have been reported for DNA modified with methyl phosphotriester linkage (POMe) using isothermal titration calorimetry (ITC) measurements which provides ΔH values directly [37].

Table 3: The T_m [°C] and thermodynamic data at 298 K for the antiparallel duplexes at different NaCl concentrations, taken from UV melting curves (λ = 260 nm).[a]

	antiparallel duplex	NaCl (mM)	T_m (°C)[b]	ΔH[c] (kJ/mol)	$T\Delta S$ (kJ/mol)	ΔG_{298} (kJ/mol)
ON/DNA	**ON1/ON20**	25	19	–430 (±20)	–400 (±20)	–30 (±28)
		50	37	–400 (±9)	–350 (±8)	–50 (±12)
		100	50	–368 (±8)	–305 (±7)	–63 (±10)
	4N+ON7/ON20	25	41 (+22.0)	–388 (±9)	–334 (±8)	–54 (±12)
		50	44 (+7.0)	–322 (±15)	–266 (±14)	–56 (±20)
		100	51 (+1.0)	–320 (±8)	–260 (±7)	–60 (±10)
	4Ts-ON13/ON20	25	29 (+10.0)	–382 (±14)	–340 (±14)	–42 (±19)
		50	34 (–3.0)	–372 (±12)	–326 (±11)	–46 (±16)
		100	45 (–5.0)	–354 (±7)	–297 (±7)	–57 (±10)
ON/RNA	**ON1/ON19**	25	44	–390 (±10)	–332 (±9)	–58 (±13)
		100	46	–306 (±17)	–248 (±17)	–58 (±24)
	4N+ON7/ON19	25	45 (+1.0)	–419 (±7)	–359 (±6)	–60 (±9)
		100	55 (+9.0)	–329 (±13)	–264 (±12)	–66 (±17)
	4Ts-ON13/ON19	25	38 (–6.0)	–420 (±20)	–370 (±20)	–50 (±28)
		100	53 (+7.0)	–407 (±13)	–337 (±12)	–70 (±17)

[a]One μM of each strand in 20 mM sodium cacodylate buffer (pH 7.0, supplemented with 25, 50, or 100 mM NaCl, respectively); [b]T_m values are reported with ±0.5 °C uncertainties as determined from several experiments; values in parentheses are ΔT_m values calculated as T_m (sample) – T_m (unmodified duplex) at the same salt concentration; [c]thermodynamic parameters are calculated as described in Supporting Information File 1 (see also Figures S9–S14) at 298 K, errors were calculated as described in reference [44].

For the native DNA duplex (**ON1/ON20**), the more favourable ΔH at the low salt concentration (25 mM NaCl) was deprived by an even higher entropy penalty leading to a loss in ΔG ($\Delta\Delta G_{298}$ = 33 kJ/mol), thus lowering the T_m value at 25 mM NaCl. For the unmodified RNA duplex (**ON1/ON19**), $\Delta\Delta H$ between 25 mM and 100 mM NaCl is 84 kJ/mol, and the corresponding entopic factor $\Delta(T\Delta S)$ is 84 kJ/mol. As a result, changes in ΔG_{298} were negligible as reflected by the small decrease in the T_m value (ΔT_m = 2 °C) when the salt concentration was reduced from 100 mM to 25 mM NaCl.

For the N+ or Ts-modified ONs, ΔH for ON/DNA duplexes was less favourable at the same salt concentration than for the unmodified duplex, whereas $T\Delta S$ was more favourable. According to Kuo et al. [37], an increase in the salt concentration stabilises the DNA duplex by reducing the entropy costs of duplex formation rather than by reducing the strand charge repulsion. This reduction of entropy costs for duplex formation is due to the endothermic release of DNA-hydrating ordered water molecules into the bulk solvent. Since the introduction of the N+ modification compensated for the negative charge on the DNA backbone, the change in the entropy costs for N+ON/DNA between 100 mM and 25 mM NaCl solutions was less than that for the unmodified duplex ($\Delta(T\Delta S)$ = −74 kJ/mol for N+ON/DNA versus −95 kJ/mol for the native DNA duplex, respectively). A similar trend was seen for the Ts-modified ONs ($\Delta(T\Delta S)$ = −43 kJ/mol), but the change in ΔH between 25 mM and 100 mM NaCl was the lowest for the duplexes with DNA ($\Delta\Delta H$ = 28 kJ/mol). This indicates that the hydrophobicity of Ts results in less water molecules involved in the formation of hydrogen bonds with dsDNA. However, it does not improve the interaction between two DNA strands, possibly due to the large size of the Ts moiety.

For duplexes of the N+ and Ts-ONs formed with complementary RNA, both ΔH and $T\Delta S$ terms were more negative at the same salt concentration than those for the unmodified ON/RNA complex, which is the opposite to duplexes formed with DNA. This led to even larger enthalpy–entropy compensation at the medium salt concentration (100 mM NaCl). When the NaCl concentration decreased from 100 mM to 25 mM, even though the reduction of entropy costs for the Ts-ON/RNA duplex was minimal for the RNA ($\Delta(T\Delta S)$ = −33 kJ/mol), the loss in ΔG for Ts-ON/RNA duplexes was larger than that of unmodified and N+ON/RNA duplexes ($\Delta\Delta G_{298}$ = 0, 6, and 20 kJ/mol for unmodified DNA/RNA, N+ON/RNA, and Ts-ON/RNA, respectively). This resulted in an unstable Ts-ON/RNA duplex at 25 mM NaCl. However, it should be noted that a significantly improved ΔH for Ts-ON/RNA duplex in comparison with the native ON/RNA at 100 mM NaCl (ΔH = −407 kJ/mol versus −306 kJ/mol, respectively) which is accountable for the more favourable ΔG_{298} and higher T_m value.

Evaluation of N+ and Ts-modified ONs towards enzymatic digestion

The nuclease resistance of the modified ONs was evaluated using snake venom phosphodiesterase (phosphodiesterase I, Sigma) and compared to the unmodified sequence **ON1**. Under the conditions used in this experiment, **ON1** was completely degraded within 30 min (Figure 2). Both, N+ and Ts-modified ONs showed an enhanced nuclease resistance when modifications were present at the 3'-end and /or in the middle of the sequence. A single N+ or Ts modification at the 5'-end of the ON did not provide protection against phosphodiesterase I. However, the resistance of the modified ONs towards phosphodiesterase increased with the number of modifications present. N+ONs, with the same number of modifications, showed a higher resistance to nuclease degradation than Ts-ONs. For example, 92.0 ± 1.8% of **4N+ON7** remained intact, whereas only 54 ± 3% of **4Ts-ON13** was intact after 120 min of enzymatic digestion (Figure 2, see also Figure S17 in Supporting Information File 1). N+ONs possessing more than four modifications showed a full enzymatic resistance after 120 min [38].

Cell-uptake study

The cellular uptake of three modified ONs synthesised possessing four or five N+ or four Ts modifications and a fluorescent label (6-FAM) at the 3'-end (Table 1) was tested. Asynchronously growing NIH3T3 mouse fibroblasts were incubated with the ONs for 12 hours, fixed in 4% paraformaldehyde before the cells were processed for fluorescent confocal microscopy.

Figure 3 shows that FAM-labelled Ts and N+-modified DNAs are concentrated in vesicles (punctate foci in the oligo/FAM panel) that accumulate around the edge of the nucleus. Interestingly, the Ts-modified oligo is also present in the nucleus as indicated by the colocalisation of the ON (Figure 3E) with the nuclear DNA (Figure 3D). The diffuse nuclear pattern of the Ts-modified ON suggests they can escape the endocytic vesicles and enter the nucleus via the nuclear pores. This is in contrast to the lack of colocalisation of the FAM signal with the nuclear DNA in the negative control (no oligo) and for the N+-modified oligos (Figure 3B, H, and K, respectively). Confocal microscope sections that dissect the nucleus were collected showing that the FAM-ON imaged were in the cytoplasm when localised adjacent to the nucleus and not on the cell surface. Staining of the cell membrane, along with the nuclear DNA, confirmed that the Ts-ON foci are present within the cytoplasm as shown in Figure 4.

Discussion

Chemical modification provides an effective and efficient way of obtaining therapeutic antigene/antisense agents based on the nucleic acid scaffold. To regulate transcription or translation, chemically modified ONs need to be able to enter the cell, resist nuclease degradation, not be toxic to the cell, and importantly, bind to the target DNA or RNA in a sequence-specific manner with high affinity [46]. The electrostatic repulsion between negatively charged phosphates is considered to be one of the factors that determines the thermodynamic stability of nucleic acid secondary structures. Neutral or positively charged oligonucleotide analogues should bind more tightly with complementary DNA or RNA. Several studies have focused on the introduction of positively charged groups to a nucleobase [47,48], a sugar [49,50], or the DNA backbone [51-53] leading to the formation of more stable duplexes and triplexes [54]. The

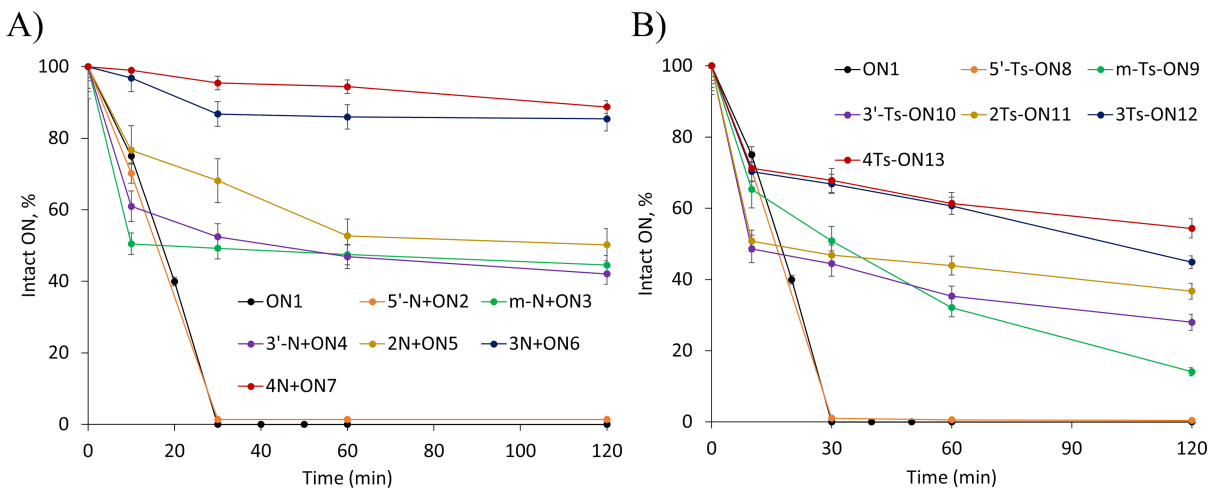


Figure 2: Percentage of intact ONs after 120 min. A) N+ONs; B) Ts-ONs. Percentage of intact ONs was determined by the ratio of full-length ONs at individual time in comparison with the sample at 0 min.

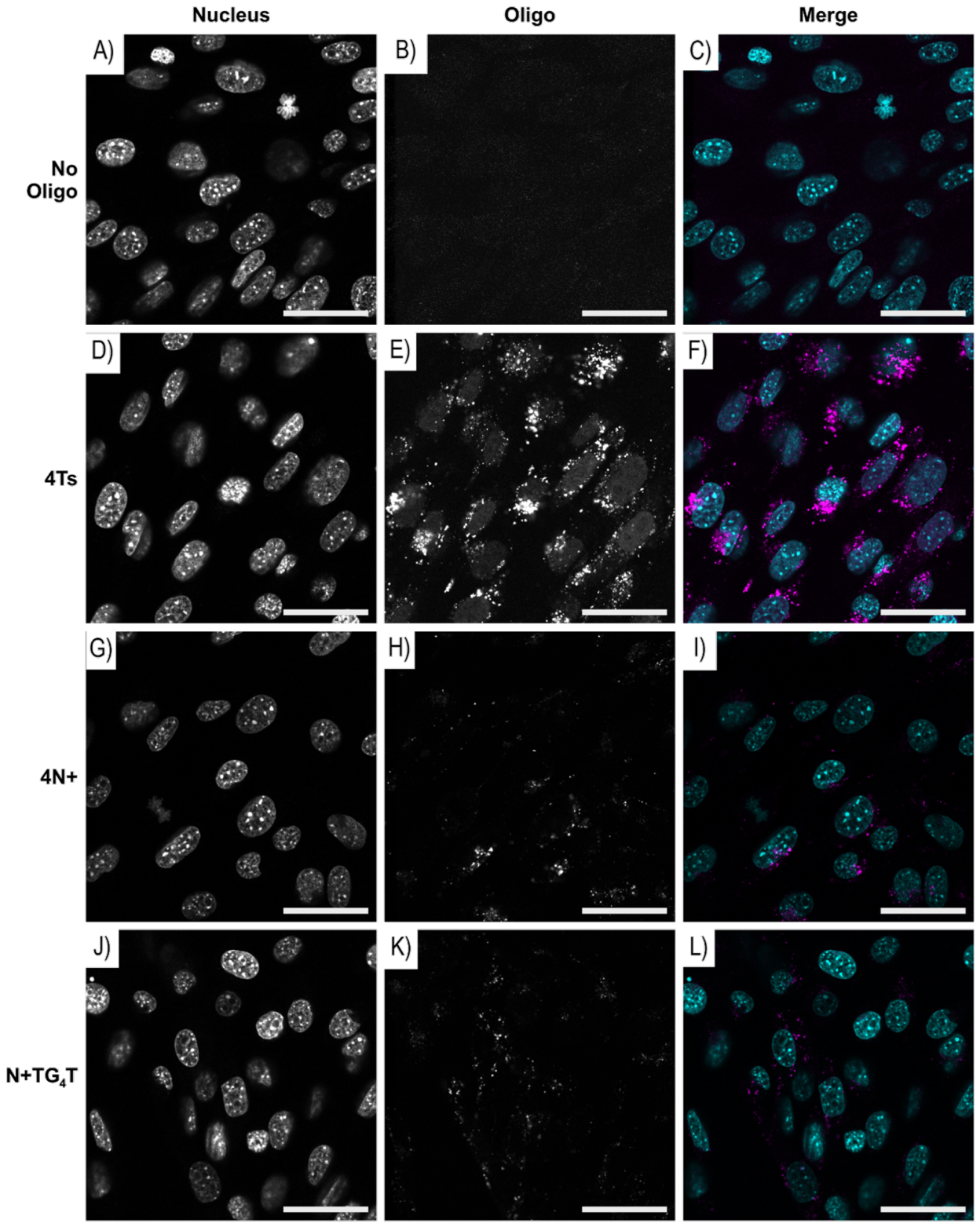


Figure 3: Representative images of mouse NIH 3T3 fibroblasts incubated with either (A–C) no oligo or 20 μM of (D–F) 4Ts, (G–I) 4N+, and (J–L) N+TG_4T FAM labelled ONs. Asynchronously growing NIH3T3 cells were incubated for 12 hours with 20 μM of the stated FAM-labelled ONs or without ON, then fixed with 4% paraformaldehyde before staining with Hoechst 3342 to identify nuclear DNA. The images were collected with a Leica SP5 DM6000B scanning confocal microscope. Individual panels, nucleus/Hoechst 3342 and oligo/FAM are shown for each section, along with merge where pseudo-coloured panels are overlaid, nucleus (blue) and oligo (magenta). Scale bar: 40 μm.

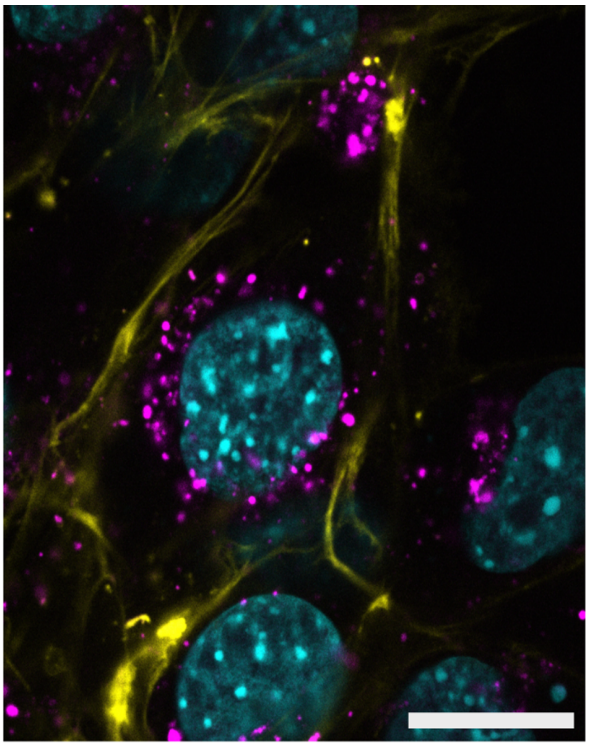


Figure 4: Representative confocal microscopy section showing the FAM vesicles inside the cell. Mouse NIH 3T3 fibroblasts were incubated with 20 μM of the FAM-labelled 4Ts-ON for 12 hours, then stained with CellBrite Fix 640 (Biotium) to identify the cell membrane before fixation with 4% paraformaldehyde. Nuclear DNA was then stained with Hoechst 3342 and images collected with a Leica SP5 DM6000B scanning confocal microscope. Overlaid pseudo-coloured panels of a section are shown: nuclear DNA (blue), 4Ts-FAM/ON (magenta), cell membrane (yellow). Scale bar: 20 μm.

introduction of sulfonamide RNA (SaRNA monomers) to replace the phosphodiester backbone led to charge-neutral sulfonamide antisense oligonucleotides (SaASOs), which resulted in a lower destabilisation (a more stable) DNA–RNA duplex compared to a DNA–DNA duplex [55]. In contrast, the incorporation of branched, charge-neutralising sleeve (BCNS) groups onto the DNA backbone led to self-neutralising ONs that did not induce a change of T_m when binding complementary DNA sequences [56], regardless of the number of BCNSs incorporated. This is in line with our results where increasing the number of N+ or Ts modifications showed no benefit in T_m values of the antiparallel duplexes formed with complementary RNA or DNA.

In order to be compatible with standard automated solid-phase DNA or RNA synthesis, the introduction of SaRNA monomers into an RNA backbone involved a 14-step preparation of the phosphoramidite for the SaRNA-TT dinucleotide. Similarly, the incorporation of BCNS groups on a DNA backbone requires the synthesis of thymidine and 2'-OMe-uridine phosphoramidites bearing BCNS groups comprising nine and ten synthetic steps, respectively. In comparison, the synthesis of an N+ monomer requires only four synthetic steps starting from commercially available 1,4-butane sultone that does not have any silica gel purification [38]. The incorporation of the N+ modification onto DNA is performed during DNA synthesis instead of a standard oxidation step. Moreover, N+ and Ts modifications can be introduced into any position in the sequence, which is not the case for SaRNA and BCNS nucleic acid analogues.

Unlike PNAs and some of BCNS groups, N+ and Ts-modified ONs demonstrated excellent chemical stability and solubility in buffer solutions. The presence of the N+ modification enhanced the stability of parallel triplexes at pH 5.0. The Ts modification also stabilises parallel triplexes at pH 5.0, but the stability decreased with increasing number of Ts moieties incorporated. Apart from ONs with N+ or Ts modifications in the middle of the sequence, both types of modified ONs hybridised with complementary RNA with a higher thermal stability than with DNA, suggesting that the N+ and Ts modifications can be used in antisense strategies. This is in contrast with reports that ONs with Ts groups destablised the duplex formation with complementary RNA (ΔT_m = −1.6 to −1.2 °C/modification) [39], which suggests that the effect of the Ts modification on the T_m is dependent on the sequence.

The thermal stability of duplexes formed by N+ONs and their complementary DNA sequence was less dependent on the ionic strength, which was predicted for zwitterionic nucleotides that can bind to natural DNA at low ionic strength as well or better than natural DNA [57]. Similar results have been reported for ONs with BCNS groups: the T_m of 2'-OMe duplexes was increased with increasing numbers of BCNSs at low ionic strength (25 mM HEPES buffer, pH 7.3). When binding to complementary RNA sequences, such behaviour was not as noticeable. It has been reported in the past that T_m values for ON/RNA duplexes are less sensitive to changes in the ionic strength in comparison with ON/DNA duplexes [58]. The Ts modification stabilised the duplex formation with RNA at a salt concentration of 100 mM , but destabilised the RNA duplex at low salt concentration.

The thermodynamic analysis of the melting curves revealed that the N+ modification stabilises the duplex with DNA because of a significantly reduced loss in entropy but stabilises the duplex formation with RNA because of the improved enthalpy at the same salt concentrations. A similar trend for ΔH and $T\Delta S$ is observed for the Ts modification compared to native DNA.

In line with the recent report [37], the loss of the thermodynamic stability at low salt concertations for the native DNA duplex was caused by the large entropic penalty that was not

compensated by the improved enthalpy. For the native RNA duplex, entropic penalty and improved enthalpy cancelled each other out resulting in a similar thermodynamic stability in the presence of 25 and 100 mM NaCl.

The polyelectrolyte ion condensation theory can be used to explain how an N+ modification stabilises duplex formation: For natural DNA, the double-helical form has a higher charge density in comparison with the single-stranded form. During denaturation, a portion of the counterions bound to DNA are lost to the bulk solvent due to the reduction in charge density. For a DNA duplex with one zwitterionic strand, the charge density of duplex and single stranded states is balanced, and only a fraction of the counterions should be lost during denaturation. As a result, the thermal stability of zwitterionic N+ DNA duplexes was less dependent on the ionic strength [57,59]. This is in line with our thermodynamic analysis that the dsDNA having N+ modifications showed less entropy costs when the ionic strength changed.

Native DNA and RNA sequences are highly susceptible to nuclease degradation within the cell. A modification on the phosphate group reduces the possibility of enzymatic digestion, which will be useful for cellular applications of N+ and Ts-modified ONs. The introduction of even a single N+ and Ts modification at the 3'-end, but not at the 5'-end, leads to the resistance of the modified ONs to enzymatic digestion by snake venom phosphodiesterase I.

The FAM-labelled ONs were shown to enter cells without the use of a transfection reagent. After a 12 hour incubation, the Ts-ON was present in both the cytoplasm and nucleus. In comparison, there was less cellular uptake of 4N+{FAM} ON and it was only present in the cytoplasm. These results indicate that ONs with phosphate modifications such as N+ or Ts might be suitable tools for the application of DNA and RNA vaccines [60], for the treatment of cancer [61], infectious diseases [62], and neurological disorders [63].

Conclusion

ONs possessing N+ and Ts modifications have good aqueous solubility and chemical stability, which allowed the assessment of these modifications in the context of DNA triplexes and duplexes. The presence of N+ or Ts modifications on the internucleotidic phosphates enhanced the binding affinity of the ONs for complementary RNA and increased their resistance to digestion by phosphodiesterase I. Fluorescently labelled Ts-ONs penetrate the cell and enter the nucleus, while N+ONs remain trapped in vesicles in the cytoplasm. These properties make the N+ and Ts-modified ONs promising candidates for cell-based applications.

Supporting Information

Supporting Information File 1

Experimental part.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-17-65-S1.pdf]

Acknowledgements

Manawatu Microscopy and Imaging Centre, NMR and mass spectrometry facilities at Massey University, and the assistance of Mr. Raoul Solomon, Dr. Patrick J. B. Edwards, and Mr. David Lun are gratefully acknowledged. We thank Dr. Maksim V. Kvach for valuable discussions and suggestions. We also thank Mr. Raoul Solomon for assistance in the preparation of Figure 3 and Figure 4.

Funding

V.V.F. and Y.S. are grateful for the financial support provided by the School of Fundamental Sciences (Massey University) and Cancer Society of New Zealand (grant No 15/07). TKH and MB thank the Health Research Council of New Zealand (grant 19/771).

ORCID® iDs

Yongdong Su - https://orcid.org/0000-0002-2092-8872
Maitsetseg Bayarjargal - https://orcid.org/0000-0001-9235-9149
Tracy K. Hale - https://orcid.org/0000-0002-9483-4588
Vyacheslav V. Filichev - https://orcid.org/0000-0002-7383-3025

Preprint

A non-peer-reviewed version of this article has been previously published as a preprint: https://doi.org/10.3762/bxiv.2020.138.v1

References

1. Muangkaew, P.; Vilaivan, T. *Bioorg. Med. Chem. Lett.* **2020,** *30,* 127064. doi:10.1016/j.bmcl.2020.127064
2. Xiao, Y.; Sheng, Y.; Zhou, J.; Chen, M.; Wen, W.; Zhang, X.; Wang, S. *Analyst* **2017,** *142,* 2617–2623. doi:10.1039/c7an00553a
3. Qadir, M. I.; Bukhat, S.; Rasul, S.; Manzoor, H.; Manzoor, M. *J. Cell. Biochem.* **2020,** *121,* 898–929. doi:10.1002/jcb.29364
4. Bennett, C. F.; Swayze, E. E. *Annu. Rev. Pharmacol. Toxicol.* **2010,** *50,* 259–293. doi:10.1146/annurev.pharmtox.010909.105654
5. Uil, T. G.; Haisma, H. J.; Rots, M. G. *Nucleic Acids Res.* **2003,** *31,* 6064–6078. doi:10.1093/nar/gkg815
6. Wu, J.; Kandavelou, K.; Chandrasegaran, S. *Cell. Mol. Life Sci.* **2007,** *64,* 2933–2944. doi:10.1007/s00018-007-7206-8
7. Boch, J. *Nat. Biotechnol.* **2011,** *29,* 135–136. doi:10.1038/nbt.1767
8. Deltcheva, E.; Chylinski, K.; Sharma, C. M.; Gonzales, K.; Chao, Y.; Pirzada, Z. A.; Eckert, M. R.; Vogel, J.; Charpentier, E. *Nature* **2011,** *471,* 602–607. doi:10.1038/nature09886

9. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J. A.; Charpentier, E. *Science* **2012,** *337,* 816–821. doi:10.1126/science.1225829
10. Hsu, P. D.; Lander, E. S.; Zhang, F. *Cell* **2014,** *157,* 1262–1278. doi:10.1016/j.cell.2014.05.010
11. Hsu, P. D.; Scott, D. A.; Weinstein, J. A.; Ran, F. A.; Konermann, S.; Agarwala, V.; Li, Y.; Fine, E. J.; Wu, X.; Shalem, O.; Cradick, T. J.; Marraffini, L. A.; Bao, G.; Zhang, F. *Nat. Biotechnol.* **2013,** *31,* 827–832. doi:10.1038/nbt.2647
12. Jensen, E. D.; Ferreira, R.; Jakočiūnas, T.; Arsovska, D.; Zhang, J.; Ding, L.; Smith, J. D.; David, F.; Nielsen, J.; Jensen, M. K.; Keasling, J. D. *Microb. Cell Fact.* **2017,** *16,* 46. doi:10.1186/s12934-017-0664-2
13. Sundaresan, R.; Parameshwaran, H. P.; Yogesha, S. D.; Keilbarth, M. W.; Rajan, R. *Cell Rep.* **2017,** *21,* 3728–3739. doi:10.1016/j.celrep.2017.11.100
14. Lee, J. K.; Jeong, E.; Lee, J.; Jung, M.; Shin, E.; Kim, Y.-h.; Lee, K.; Jung, I.; Kim, D.; Kim, S.; Kim, J.-S. *Nat. Commun.* **2018,** *9,* 3048. doi:10.1038/s41467-018-05477-x
15. Moser, H. E.; Dervan, P. B. *Science* **1987,** *238,* 645–650. doi:10.1126/science.3118463
16. Wu, Q.; Gaddis, S. S.; MacLeod, M. C.; Walborg, E. F.; Thames, H. D.; DiGiovanni, J.; Vasquez, K. M. *Mol. Carcinog.* **2007,** *46,* 15–23. doi:10.1002/mc.20261
17. Hoogsteen, K. *Acta Crystallogr.* **1963,** *16,* 907–916. doi:10.1107/s0365110x63002437
18. Jain, M. L.; Bruice, P. Y.; Szabó, I. E.; Bruice, T. C. *Chem. Rev.* **2012,** *112,* 1284–1309. doi:10.1021/cr1004265
19. Crooke, S. T.; Liang, X.-h.; Baker, B. F.; Crooke, R. M. *J. Biol. Chem.* **2021,** 100416. doi:10.1016/j.jbc.2021.100416
20. Hansen, M. E.; Bentin, T.; Nielsen, P. E. *Nucleic Acids Res.* **2009,** *37,* 4498–4507. doi:10.1093/nar/gkp437
21. Obika, S.; Nanbu, D.; Hari, Y.; Morio, K.-i.; In, Y.; Ishida, T.; Imanishi, T. *Tetrahedron Lett.* **1997,** *38,* 8735–8738. doi:10.1016/s0040-4039(97)10322-7
22. Koshkin, A. A.; Singh, S. K.; Nielsen, P.; Rajwanshi, V. K.; Kumar, R.; Meldgaard, M.; Olsen, C. E.; Wengel, J. *Tetrahedron* **1998,** *54,* 3607–3630. doi:10.1016/s0040-4020(98)00094-5
23. Padmapriya, A. A.; Tang, J. Y.; Agrawal, S. *Antisense Res. Dev.* **1994,** *4,* 185–199. doi:10.1089/ard.1994.4.185
24. Eckstein, F. *Antisense Nucleic Acid Drug Dev.* **2000,** *10,* 117–121. doi:10.1089/oli.1.2000.10.117
25. Lohse, J.; Dahl, O.; Nielsen, P. E. *Proc. Natl. Acad. Sci. U. S. A.* **1999,** *96,* 11804–11808. doi:10.1073/pnas.96.21.11804
26. Wan, W. B.; Seth, P. P. *J. Med. Chem.* **2016,** *59,* 9645–9667. doi:10.1021/acs.jmedchem.6b00551
27. Koshkin, A. A.; Nielsen, P.; Meldgaard, M.; Rajwanshi, V. K.; Singh, S. K.; Wengel, J. *J. Am. Chem. Soc.* **1998,** *120,* 13252–13253. doi:10.1021/ja9822862
28. Singh, S. K.; Koshkin, A. A.; Wengel, J.; Nielsen, P. *Chem. Commun.* **1998,** 455–456. doi:10.1039/a708608c
29. Seth, P. P.; Siwkowski, A.; Allerson, C. R.; Vasquez, G.; Lee, S.; Prakash, T. P.; Wancewicz, E. V.; Witchell, D.; Swayze, E. E. *J. Med. Chem.* **2009,** *52,* 10–13. doi:10.1021/jm801294h
30. Swayze, E. E.; Siwkowski, A. M.; Wancewicz, E. V.; Migawa, M. T.; Wyrzykiewicz, T. K.; Hung, G.; Monia, B. P.; Bennett, C. F. *Nucleic Acids Res.* **2007,** *35,* 687–700. doi:10.1093/nar/gkl1071
31. Nakamura, A.; Ali, S. A.; Kapoor, M. *Bone* **2020,** *138,* 115461. doi:10.1016/j.bone.2020.115461
32. De Costa, N. T. S.; Heemstra, J. M. *PLoS One* **2013,** *8,* e58670. doi:10.1371/journal.pone.0058670
33. Miller, P. S.; McParland, K. B.; Jayaraman, K.; Tso, P. O. P. *Biochemistry* **1981,** *20,* 1874–1880. doi:10.1021/bi00510a024
34. Shoji, Y.; Akhtar, S.; Periasamy, A.; Herman, B.; Juliano, R. L. *Nucleic Acids Res.* **1991,** *19,* 5543–5550. doi:10.1093/nar/19.20.5543
35. Nagahama, K.; Veedu, R. N.; Wengel, J. *Bioorg. Med. Chem. Lett.* **2009,** *19,* 2707–2709. doi:10.1016/j.bmcl.2009.03.116
36. Saccà, B.; Lacroix, L.; Mergny, J.-L. *Nucleic Acids Res.* **2005,** *33,* 1182–1192. doi:10.1093/nar/gki257
37. Kuo, T.-C.; Wu, M.-W.; Lin, W.-C.; Matulis, D.; Yang, Y.-S.; Li, S.-Y.; Chen, W.-Y. *J. Taiwan Inst. Chem. Eng.* **2020,** *110,* 1–7. doi:10.1016/j.jtice.2020.02.023
38. Su, Y.; Edwards, P. J. B.; Stetsenko, D. A.; Filichev, V. V. *ChemBioChem* **2020,** *21,* 2455–2466. doi:10.1002/cbic.202000110
39. Prokhorova, D. V.; Chelobanov, B. P.; Burakova, E. A.; Fokina, A. A.; Stetsenko, D. A. *Russ. J. Bioorg. Chem.* **2017,** *43,* 38–42. doi:10.1134/s1068162017010071
40. Felsenfeld, G.; Miles, H. T. *Annu. Rev. Biochem.* **1967,** *36,* 407–448. doi:10.1146/annurev.bi.36.070167.002203
41. Doluca, O.; Withers, J. M.; Filichev, V. V. *Chem. Rev.* **2013,** *113,* 3044–3083. doi:10.1021/cr300225q
42. El Hassan, M. A.; Calladine, C. R. *J. Mol. Biol.* **1996,** *259,* 95–103. doi:10.1006/jmbi.1996.0304
43. Kotin, L. *J. Mol. Biol.* **1963,** *7,* 309–311. doi:10.1016/s0022-2836(63)80009-1
44. Xia, T.; SantaLucia, J., Jr.; Burkard, M. E.; Kierzek, R.; Schroeder, S. J.; Jiao, X.; Cox, C.; Turner, D. H. *Biochemistry* **1998,** *37,* 14719–14735. doi:10.1021/bi9809425
45. Mergny, J.-L.; Lacroix, L. *Oligonucleotides* **2003,** *13,* 515–537. doi:10.1089/154545703322860825
46. Rayburn, E. R.; Zhang, R. *Drug Discovery Today* **2008,** *13,* 513–521. doi:10.1016/j.drudis.2008.03.014
47. Robles, J.; Grandas, A.; Pedroso, E. *Tetrahedron* **2001,** *57,* 179–194. doi:10.1016/s0040-4020(00)00991-1
48. Roig, V.; Asseline, U. *J. Am. Chem. Soc.* **2003,** *125,* 4416–4417. doi:10.1021/ja029467v
49. Kanazaki, M.; Ueno, Y.; Shuto, S.; Matsuda, A. *J. Am. Chem. Soc.* **2000,** *122,* 2422–2432. doi:10.1021/ja9934706
50. Puri, N.; Majumdar, A.; Cuenoud, B.; Natt, F.; Martin, P.; Boyd, A.; Miller, P. S.; Seidman, M. M. *Biochemistry* **2002,** *41,* 7716–7724. doi:10.1021/bi025734y
51. Bailey, P.; Sartorelli, V.; Hamamori, Y.; Muscat, G. E. O. *Nucleic Acids Res.* **1998,** *26,* 5501–5510. doi:10.1093/nar/26.23.5501
52. Michel, T.; Debart, F.; Vasseur, J.-J. *Tetrahedron Lett.* **2003,** *44,* 6579–6582. doi:10.1016/s0040-4039(03)01694-0
53. Vasquez, K. M.; Dagle, J. M.; Weeks, D. L.; Glazer, P. M. *J. Biol. Chem.* **2001,** *276,* 38536–38541. doi:10.1074/jbc.m101797200
54. Fox, K. R. *Curr. Med. Chem.* **2003,** *7,* 17–37. doi:10.2174/0929867003375506
55. Korotkovs, V.; Reichenbach, L. F.; Pescheteau, C.; Burley, G. A.; Liskamp, R. M. J. *J. Org. Chem.* **2019,** *84,* 10635–10648. doi:10.1021/acs.joc.9b00941
56. Yanachkov, I.; Zavizion, B.; Metelev, V.; Stevens, L. J.; Tabatadze, Y.; Yanachkova, M.; Wright, G.; Krichevsky, A. M.; Tabatadze, D. R. *Org. Biomol. Chem.* **2017,** *15,* 1363–1380. doi:10.1039/c6ob02576e
57. Hashimoto, H.; Nelson, M. G.; Switzer, C. *J. Am. Chem. Soc.* **1993,** *115,* 7128–7134. doi:10.1021/ja00069a009
58. Bentley, J.; Brazier, J. A.; Fisher, J.; Cosstick, R. *Org. Biomol. Chem.* **2007,** *5,* 3698–3702. doi:10.1039/b713292a

59. Tomac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P. E.; Nordén, B.; Gräslund, A. *J. Am. Chem. Soc.* **1996,** *118,* 5544–5552. doi:10.1021/ja960495l
60. Leitner, W. W.; Ying, H.; Restifo, N. P. *Vaccine* **1999,** *18,* 765–777. doi:10.1016/s0264-410x(99)00271-6
61. Lopes, A.; Vandermeulen, G.; Préat, V. *J. Exp. Clin. Cancer Res.* **2019,** *38,* 146. doi:10.1186/s13046-019-1154-7
62. Rezaei, T.; Khalili, S.; Baradaran, B.; Mosafer, J.; Rezaei, S.; Mokhtarzadeh, A.; de la Guardia, M. *J. Controlled Release* **2019,** *316,* 116–137. doi:10.1016/j.jconrel.2019.10.045
63. Petrushina, I.; Hovakimyan, A.; Harahap-Carrillo, I. S.; Davtyan, H.; Antonyan, T.; Chailyan, G.; Kazarian, K.; Antonenko, M.; Jullienne, A.; Hamer, M. M.; Obenaus, A.; King, O.; Zagorski, K.; Blurton-Jones, M.; Cribbs, D. H.; Lander, H.; Ghochikyan, A.; Agadjanyan, M. G. *Neurobiol. Dis.* **2020,** *139,* 104823. doi:10.1016/j.nbd.2020.104823

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Enhanced target cell specificity and uptake of lipid nanoparticles using RNA aptamers and peptides

Roslyn M. Ray[‡1], Anders Højgaard Hansen[‡2], Maria Taskova[2], Bernhard Jandl[3], Jonas Hansen[2], Citra Soemardy[1], Kevin V. Morris[1,4,5] and Kira Astakhova[*2]

Full Research Paper

Open Access

Address:
[1]Center for Gene Therapy, Beckman Research Institute, City of Hope, Duarte, CA, United States of America, [2]Department of Chemistry, Technical University of Denmark, Lyngby, Denmark, [3]Institute of Biological Chemistry, Faculty of Chemistry, University of Vienna, 1090 Vienna, Austria, [4]School of Medical Sciences, Griffith University, Gold Coast, Australia 4222 and [5]Menzies Health Institute Queensland, Griffith University, Gold Coast, QLD 4222, Australia

Email:
Kira Astakhova[*] - kiraas@kemi.dtu.dk

* Corresponding author ‡ Equal contributors

Keywords:
aptamer; blood–brain barrier; gene therapy; HIV-1; lipid nanoparticle

Beilstein J. Org. Chem. **2021,** *17,* 891–907.
https://doi.org/10.3762/bjoc.17.75

Received: 26 January 2021
Accepted: 09 April 2021
Published: 26 April 2021

This article is part of the thematic issue "Celebrating the role of chemistry in the success of oligonucleotides as therapeutics".

Guest Editors: P. Kumar and T. Brown

Abstract

Lipid nanoparticles (LNPs) constitute a facile and scalable approach for delivery of payloads to human cells. LNPs are relatively immunologically inert and can be produced in a cost effective and scalable manner. However, targeting and delivery of LNPs across the blood–brain barrier (BBB) has proven challenging. In an effort to target LNPs composed of an ionizable cationic lipid (DLin-MC3-DMA), cholesterol, the phospholipid 1,2-distearoyl-*sn*-glycero-3-phosphocholine (DSPC), and 1,2-dimyristoyl-*rac*-glycero-3-methoxypolyethylene glycol-2000 (DMG-PEG 2000) to particular cell types, as well as to generate LNPs that can cross the BBB, we developed and assessed two approaches. The first was centered on the BBB-penetrating trans-activator of transcription (Tat) peptide or the peptide T7, and the other on RNA aptamers targeted to glycoprotein gp160 from human immunodeficiency virus (HIV) or C-C chemokine receptor type 5 (CCR5), a HIV-1 coreceptor. We report herein a CCR5-selective RNA aptamer that acts to facilitate entry through a simplified BBB model and that drives the uptake of LNPs into CCR5-expressing cells, while the gp160 aptamer did not. We further observed that the addition of cell-penetrating peptides, Tat and T7, did not increase BBB penetration above the aptamer-loaded LNPs alone. Moreover, we found that these targeted LNPs exhibit low immunogenic and low toxic profiles and that targeted LNPs can traverse the BBB to potentially deliver drugs into the target tissue. This approach highlights the usefulness of aptamer-loaded LNPs to increase target cell specificity and potentially deliverability of central-nervous-system-active RNAi therapeutics across the BBB.

Introduction

Lipid nanoparticles (LNPs) represent an effective platform for delivering small molecules, RNA, or DNA into target cells [1]. LNPs have been successfully deployed via different administration routes in vivo to distribute cargo into target tissues [2-8]. By changing lipid composition [6] and/or including short peptides [9] and ligands [10], one can modulate the biodistribution of the LNP in the body. However, despite these advances, targeting of LNPs to the brain tissue remains a challenge [11].

In order to reach safer therapeutic options for treatment of brain diseases and disorders, a productive drug transport across the blood–brain barrier (BBB) is critical. For example, despite successful implementation of antiretroviral drugs for the treatment of human immunodeficiency virus 1 (HIV-1), HIV-1-associated neurological disorders persist due to the poor uptake of antiretroviral drugs across the BBB [12-14]. There are two ways to traverse the BBB, one is through temporary disruption of the physical barrier, which impairs BBB function, and the other is to use nanocarriers or particles [11]. The latter presents a noninvasive route that is safer than physical disruption [11]. One approach to increase transport of LNPs through the notoriously protective BBB is to use short positively charged peptides or receptor-specific ligands, both of which have shown to be effective at increasing transport of LNPs, nucleotides, and small molecules through the BBB [9,15-17]. For example, the short positively charged peptide Tat has previously been demonstrated to be effective as an excipient species to increase the uptake through the negatively charged BBB [9,18]. Tat (sequence: H-YGRKKRRQRRR-NH_2) is an arginine-rich short cell-penetrating peptide derived from the natural nuclear Tat protein of HIV-1 [19,20]. The HIV-1 Tat protein itself has been shown to traverse the BBB by acting as a cell-penetrating peptide [9,20]. Other small positively charged molecules used for BBB penetration include transferrin and corresponding peptide derivatives or analogs that act as ligands for the transferrin receptor. The transferrin receptor is highly expressed in brain capillaries, nucleated cells, and in rapidly dividing cells [21], and its endogenous ligand transferrin has previously been used to increase transport of small molecules and oligonucleotides across the BBB [21-23]. The peptide T7 consisting of seven amino acids (H-HAIYPRH-NH_2) was identified via phage display [24] and has a high affinity (≈10 nM) for the transferrin receptor [24,25]. This peptide does not compete with endogenous transferrin binding and has been used to successfully enhance drug delivery to brain tissue [15,22,24-26]. Both peptides were included in this study and modified with an N-terminal lipid anchor for LNP postinsertion. The design of the lipid anchor includes two palmitoyl chains that are attached through a 1,2-diaminopropanoic acid moiety (Dap) on the N-terminus of each peptide, providing the lipidated peptides dipalmitoyl-Dap-T7 and dipalmitoyl-Dap-Tat (Figure 1A). Double lipidation ensures a more stable lipid-membrane-anchoring compared to a single fatty acid chain or cholesteryl variant [27-29]. The careful choice of Dap and palmitic acid allows for the entire synthesis to be performed on solid support with no need for additional reactions after cleavage [27-29].

One approach to generate LNP formulations with higher specificity for antigen-expressing cells is to use RNA aptamers. RNA aptamers are short oligonucleotides that are evolved using a process called systematic evolution of ligands by exponential enrichment (SELEX) [30]. SELEX is an iterative process that begins with a large oligonucleotide library that, through a process of negative and positive selections, ends with a few

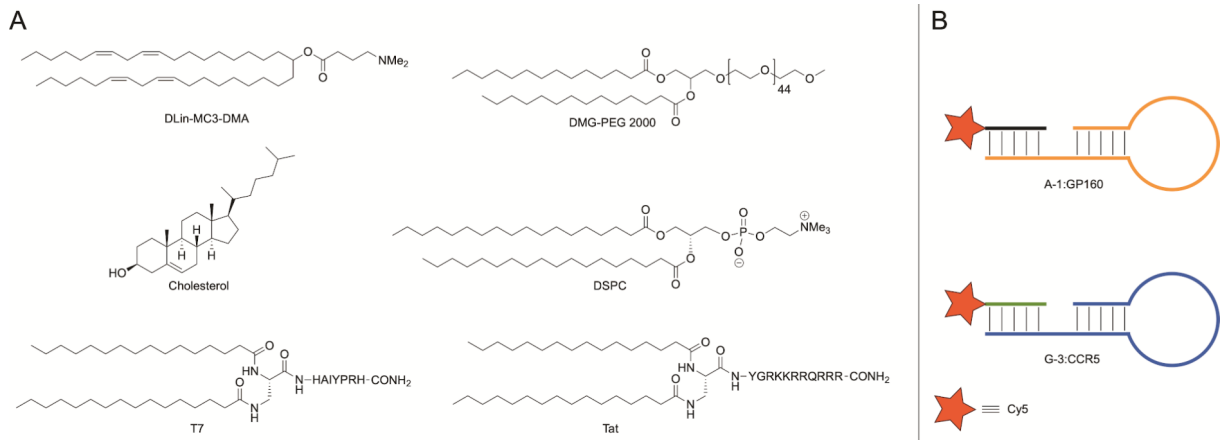


Figure 1: Components of the LNPs. A) Lipid species and lipidated cell-penetrating peptides applied by postinsertion. B) Graphical representation of aptamer–probe hybrids.

candidates that are specific for a particular protein [30,31]. Using HIV-1 as our model, we explored the use of two RNA aptamers as a mean to increase the specificity of LNPs for HIV-1-infected and/or target cells [31]. RNA aptamers are ideal candidates due to the lower immunogenicity profile than the DNA counterparts [30,32,33]. RNA aptamers are also highly amenable to form complex and dynamic secondary structures, which makes them ideal molecules for novel ligand development [31]. Zhou et al. previously reported on an RNA aptamer specific for the HIV-1 entry coreceptor C-C chemokine receptor type 5 (CCR5) [34] and an RNA aptamer specific for the HIV-1 envelope protein gp160 [35]. The CCR5 RNA aptamer G-3 has been shown to be specific for, and internalized by the CCR5 receptor [34]. Similarly, it has been found that the A-1 aptamer specifically recognizes gp160 and that it may be internalized through receptor-mediated endocytosis [35]. Both the G-3 and A-1 aptamers have been conjugated to small interfering RNA (siRNA), through a stick bridge motif, to deliver siRNAs into the respective target cells. The G-3 siRNA conjugate had the highest efficacy with 70% delivery into target T-cells, while the A-1 siRNA conjugate showed a 20% delivery into target gp160-expressing cells [34,35]. Thus, both aptamers present an additional potential route for LNP internalization and target cell specificity. In order to assess the ability of aptamers to drive LNP internalization, short complementary Cy5-DNA oligonucleotides specific for each aptamer were used as probes to detect LNP uptake in different cells.

In this study, we employed lipid compositions and formulation procedures previously reported in literature [4]. Specifically, the cationic and ionizable DLin-MC3-DMA lipid is a constituent of the FDA-approved LNP-formulated siRNA drug Patisiran® for treatment of familial transthyretin amyloidosis [36,37]. Clinical trial safety assessments of this formulation showed no liver toxicity and no immune stimulation, with ≈10% of trial participants experiencing mild to moderate adverse events upon administration [38]. It includes encapsulation of siRNA by a mixture of lipid components, such as an ionizable cationic lipid, 1,2-distearoyl-*sn*-glycero-3-phosphocholine (DSPC), cholesterol, and PEG-lipid, each with an essential role in the design (Figure 1). These lipids promote the effective distribution of the LNP in vivo as well as aid in effective cargo release from the endosome [1,37]. To this end, we herein report the efficacy, delivery capability, and functionality of the addition of peptides and RNA aptamers in facilitating entry through a simplified BBB model as well as to determine whether inclusion of these molecules could facilitate cell specific uptake. We further show that LNPs generally exhibit a low immunogenic and toxic profile and that RNA aptamers can act as potential enhancers to effectuate the delivery of LNPs into the central nervous system (CNS).

Results

Lipid nanoparticle development and characteristics

In accordance with a previously published procedure, we generated LNPs using a mixture of DLin-MC3-DMA, DSPC, cholesterol, and 1,2-dimyristoyl-*rac*-glycero-3-methoxypolyethylene glycol-2000 (DMG-PEG 2000). Lipids were first extruded and then complexed with negatively charged aptamers annealed with fluorescently tagged complementary DNA oligonucleotides (GP160:A-1 or CCR5:G-3) to simultaneously assemble the LNPs (the formulation list is provided in Table 1). At this stage, the LNPs were examined by dynamic light scattering (DLS, Table 2). While noncomplexed (empty) LNPs had an average size of 62.4 nm and a zeta potential (ZP) of −2.9 mV, LNPs mixed with GP160:A-1 and CCR5:G-3 displayed average sizes of 57.3 nm and 91.9 nm, respectively, as well as a more negative ZP (−11 mV and −9.4 mV, respectively, Table 2). These ZP values indicate that complexation leads from a neutral to anionic LNP product [39], a property that typically confers with low to no cytotoxicity in vivo [40]. Further, the additional decrease in the ZP indicates efficient aptamer loading into the LNPs. Additionally, low polydispersity index (PDI) values reported for both formulations (Table 2) indicate a high degree of monodispersity.

Table 1: Formulations used in the present study.[a]

LNP sample	Cy5 DNA probe/aptamer	lipopeptide
LNP B9	—	—
LNP B9 A-1	A-1:GP160	—
LNP B9 G-3	G-3:CCR5	—
LNP B9 T7	—	T7
LNP B9 Tat	—	Tat
LNP B9 A-1 T7	A-1:GP160	T7
LNP B9 A-1 Tat	A-1:GP160	Tat
LNP B9 G-3 T7	G-3:CCR5	T7
LNP B9 G-3 Tat	G-3:CCR5	Tat

[a]T7 (H-HAIYPRH-NH_2) is targeting transferrin receptor. Tat (H-YGRKKRRQRRR-NH_2) is derived from the natural nuclear Tat protein of HIV-1.

Next, LNPs were incubated with either Tat or T7 and the physical characteristics assessed by nanoparticle tracking analysis (NTA, Table 3) and transmission electron microscopy (TEM, Supporting Information File 1, Figure S2). After postinsertion, LNP sizes were found by NTA to range from 54–66 nm (Table 3), while TEM analysis revealed average sizes between 45–52 nm (Supporting Information File 1, Figure S2B). While

Table 2: DLS data listing particle size, PDI, and ZP of LNP formulations.

LNP formulation	physical characterization by DLS		
	mean diameter (nm)	PDI	zeta potential (mV)
LNP B9	62.4 ± 0.7	0.2 ± 0.01	–2.9 ± 1.1
LNP B9 A-1	57.3 ± 0.9	0.1 ± 0.03	–11.0 ± 1.4
LNP B9 G-3	91.9 ± 4.2	0.3 ± 0.03	–9.4 ± 1.0

there appears to be a ≈10 nm discrepancy when comparing DLS and NTA with TEM, this size difference was found to be consistent between these methods of analyses for all samples.

For example, LNP B9 A-1 Tat was characterized by the smallest average size using both NTA (≈54 nm) and TEM (≈45 nm). Thus, the average sizes obtained by NTA are in agreement with the average size observed using TEM (Supporting Information File 1, Figure S2A and Table S2). Similarly, while the mean diameter of LNP B9 G-3 was found to be larger by DLS (91.9 nm) than the values reported for NTA (67.2 nm) and TEM (52 nm), the sizes of the LNP B9 and LNP B9 A-1 samples via DLS are also in agreement with the reported NTA and TEM sizes. These discrepancies may be indicative of the inherent differences between these three analytical methods and highlight the need to confirm LNP sizes using more than one technique. Nevertheless, the small size of these nanoparticles (<100 nm) is ideal for in vivo applications as they may bypass the reticuloendothelial system and thereby increase LNP circulation time in vivo [41].

LNPs with postinsertion T7 peptide

Previous studies have demonstrated the ability of the T7 peptide to increase LNP transport across the BBB [22-24,42]. In order to test this, we used hCMEC/D3, HEK293Ts, HeLa, and TZM-bL cell lines. hCMEC/D3 is a human brain endothelial cell line that mimics a simplified BBB and, using a transwell assay, allows to study the BBB penetration potential of compounds [43]. To assess the specific uptake of the G-3 aptamer, we used TZM-bLs. TZM-bL is a HeLa-derived cell line that was engineered to express CD4 and CCR5 receptors on the cell surface [44]. HeLa cells were used as a negative control. To investigate the specific uptake of the A-1 aptamer, we used a HEK293T cell line engineered to express gp160 [45], HEK293T-gp160, and the parental HEK293T served as a control cell line.

hCMEC/D3 cells were cultured on a 0.4 μM transwell mesh until a trans-endothelial electrical resistance of above 30 $\Omega \cdot cm^2$ was reached. This measure is an indicator that a tight junction barrier has formed within these cells and can be used to determine the ability of the LNPs to pass through the BBB (Supporting Information File 1, Figure S3A). Additionally, we further confirmed our junctions using fluorescent microscopy on the barrier layers to confirm expression of claudin-5, a known tight junction protein (Supporting Information File 1, Figure S3B). We observed that LNPs were readily taken up by both HeLa and TZM-bls in the absence of a transwell insert (Figure 2A and Figure 2B). With the addition of the hCMEC/D3 cells in the apical chamber, we found that HeLa cells were less Cy5-positive (≈60%) than the target TZM-bl cells (≈100%, Figure 2A). Further, when examining the intensity of Cy5 in these cell populations, we found that the addition of the T7 peptide increases uptake by 1.2-fold through the hCMEC/D3 cellular

Table 3: NTA analysis listing size and concentration of LNPs.

LNP formulation	physical characterization by DLS		
	mean diameter (nm)	standard deviation (nm)	particle concentration ± SEM (particles/mL)
LNP B9	69.2 ± 0.3 nm	30.8 ± 1.5 nm	$3.13 \times 10^{11} \pm 1.75 \times 10^{10}$
LNP B9 A-1	66.6 ± 1.4 nm	25.2 ± 1.4 nm	$3.82 \times 10^{11} \pm 6.20 \times 10^{9}$
LNP B9 A-1 T7	65.7 ± 1.1 nm	26.3 ± 2.4 nm	$3.25 \times 10^{11} \pm 2.82 \times 10^{10}$
LNP B9 A-1 Tat	54.2 ± 0.6 nm	22.1 ± 1.4 nm	$8.90 \times 10^{11} \pm 7.23 \times 10^{10}$
LNP B9 G-3	67.2 ± 0.3 nm	30.2 ± 0.8 nm	$2.71 \times 10^{11} \pm 1.45 \times 10^{10}$
LNP B9 G-3 T7	66.5 ± 1.7 nm	32.2 ± 5.0 nm	$3.30 \times 10^{11} \pm 2.60 \times 10^{10}$
LNP B9 G-3 Tat	57.3 ± 0.5 nm	29.2 ± 1.7 nm	$8.05 \times 10^{11} \pm 7.83 \times 10^{10}$
LNP B9 T7	75.1 ± 1.5 nm	32.0 ± 1.4 nm	$2.19 \times 10^{11} \pm 1.65 \times 10^{10}$
LNP B9 Tat	61.2 ± 0.7 nm	15.2 ± 1.5 nm	$2.19 \times 10^{11} \pm 1.69 \times 10^{10}$

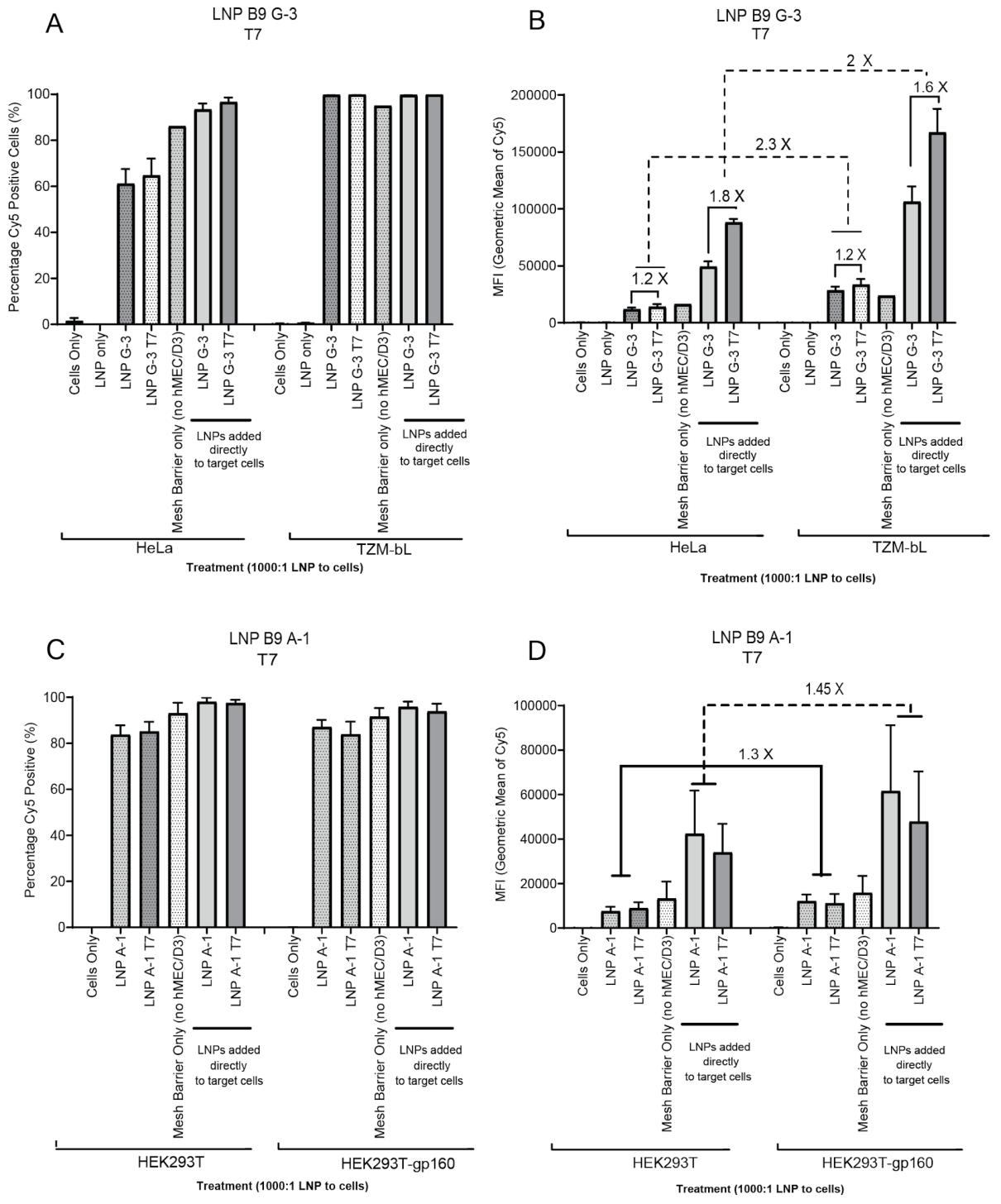


Figure 2: LNPs with T7 pass through the transwell cell barrier and are taken up by target cells. HeLa (CCR5-negative control cell) or TZM-bl (CCR5-positive cell type) cells (A and B) as well as HEK293T or gp160 positive HEK293T cells (C and D) were seeded at a density of 50,000 cells/well. The next day, transwell inserts containing confluent hMEC/D3 cells at trans-endothelial electrical resistance (TEER) above 30 Ω·cm^2 were placed into experimental wells, LNPs (1000:1) were added to the apical surface, and 24 h later, the target cells were processed for Cy5 detection using fluorescence-activated single-cell sorting (FACS). A) Percentage of cells positive for Cy5 detection in HeLa and TZM-bls. B) MFI of Cy5 in each cell population in HeLa and TZM-bls. C) Percentage of cells positive for Cy5 detection in HEK cell types. D) MFI of Cy5 in each cell population in HEK cell types. Histograms are representative of three independent biological experiments, each containing duplicate technical replicates.

barrier while also increasing uptake through direct addition by 1.6–1.8-fold (Figure 2B). Additionally, the mean fluorescent intensity (MFI) was found to be 2–2.3-fold higher in the target TZM-bl cells in both barrier and nonbarrier treatment groups compared to the control HeLa cells, indicating a higher accumulation of LNPs in the target cells (Figure 2B and Supporting Information File 1, Figure S4). Passive diffusion of the LNPs with the G-3 aptamer alone through a transwell insert without hCMEC/D3 cells appears to show higher uptake in the HeLa cell line but lower uptake in the TZM-bl cell line in comparison to the transwell insert with hCMEC/D3 cells (Figure 2A and Figure 2B).

In contrast, we found that formulating LNPs with the gp160-specific A-1 aptamer did not result in any significant increase in percentage uptake in the target gp160-positive HEK293T cells compared to HEK293T cells alone (Figure 2C). However, we did observe the MFI in gp160-positive HEK293T to be 1.3- and 1.45-fold higher (barrier and nonbarrier groups, respectively) than in the HEK293T cells alone (Figure 2D), suggesting higher levels of LNPs in gp160-expressing HEK293T cells. We also observed that direct addition of the LNPs resulted in a higher percentage of Cy5-positive cell detection and a higher MFI compared to the hCMEC/D3 barrier (Figure 2D and Supporting Information File 1, Figure S4).

Collectively, these data suggest that the candidate LNPs, particularly LNP B9 G-3 T7, may increase uptake through tight junctions and prove useful in transiting drugs and small cargo through the BBB in vivo.

LNPs with postinserted Tat peptide

Tat is a cationic peptide that is known to increase transport of molecules through the BBB and increase uptake into cells [18]. In a similar manner to the transferrin peptide T7, we investigated the ability of Tat to drive LNP uptake in cell lines. Interestingly, we found that the addition of the Tat peptide to either the A-1- or G-3-complexed LNPs did not have any effect on BBB penetration (Figure 3). Rather, we observed that LNPs containing the G-3 aptamer showed an increased uptake in target cells expressing CCR5 (Figure 3A and Figure 3B). We observed that TZM-bls had a ≈98% uptake of LNPs via the hCMEC/D3 barrier compared to ≈63% in HeLa cells (Figure 3A). We also observed a similar increase (1.75-fold, barrier and 1.65-fold, nonbarrier) in MFI in TZM-bl target cells compared to the nontarget HeLa cells (Figure 3B and Supporting Information File 1, Figure S4). Further, we observed similar trends for the A-1 aptamer, where Tat had no effect on BBB penetration (Figure 3C, Figure 3D, and Supporting Information File 1, Figure S4). Interestingly, in this group, percentage uptake was lower across all groups compared to the LNP A-1 T7 group (Figure 3C and Figure 3D). This may be due to differences in hCMEC/D3 barrier formation or LNP counting error using NTA.

Collectively, these data suggest that the addition of Tat to LNPs has no effects on BBB transit when compared to the T7 peptide. We further found that A-1 aptamer incorporation into the LNP formulation does not appear to enhance specific targeting of gp160-expressing cells either through the hCMEC/D3 barrier or through direct addition, suggesting that it may not be an ideal candidate moving forward.

LNPs do not stimulate an immune response

In order to further characterize LNPs, we decided to evaluate the immunogenic profile. We stimulated monocytes obtained from whole blood for 6 days with 10 ng/mL granulocyte-macrophage colony-stimulating factor (GM-CSF). This programs the monocytes to form macrophages that are primed to respond in a type-1 manner. After 24 h of stimulation with either the LNPs or positive controls for an RNA/DNA response (poly I:C) or a bacterial response (LPS), we found that the LNPs did not increase secretion of any of the cytokines tested (IL-1β, IL-10, IL-6, IFN-γ, TNFα, IL-2, IL-4, IL-8, and IL-5) above basal (phosphate-buffered saline, PBS) conditions (Figure 4A). Additionally, we confirmed LNP uptake by the monocyte-derived macrophages (MDMs) using fluorescent microscopy (Figure 4B). We found that all LNPs containing the Cy5 oligonucleotide were observable under the microscope (Figure 4B) and that all macrophages were 100% positive for Cy5. Additionally, using QuPath analysis software, we determined the Cy5 MFI for each image. Interestingly, we found that the LNP A-1 and the LNP G-3 had higher MFI values in all the donors assessed compared to the Tat and T7 counterparts (Figure 4C). Further, we found that the LNP G-3 exhibited the highest uptake in all the donors assessed (Figure 4C). These observations suggest that the candidate LNPs are relatively immunologically inert and may prove to be well-tolerated in vivo.

Aptamer and peptide LNPs have modest effects on cell viability in a cell-specific manner

We next assessed whether LNPs could affect cell viability in HeLa and HEK293Ts cells. Cells were treated with the LNPs for 24 h prior to performing the alamarBlue viability assay. In HeLa cells, we found that the LNP B9 alone had no effect on cell viability compared to the PBS control (Figure 5A). Interestingly, we observed that cell viability was reduced by ≈20% in HeLa cells treated with LNPs containing either A-1 or G-3 aptamer or LNPs with the Tat or T7 peptide alone (Figure 5A). However, LNPs containing both the aptamer and a peptide (Tat

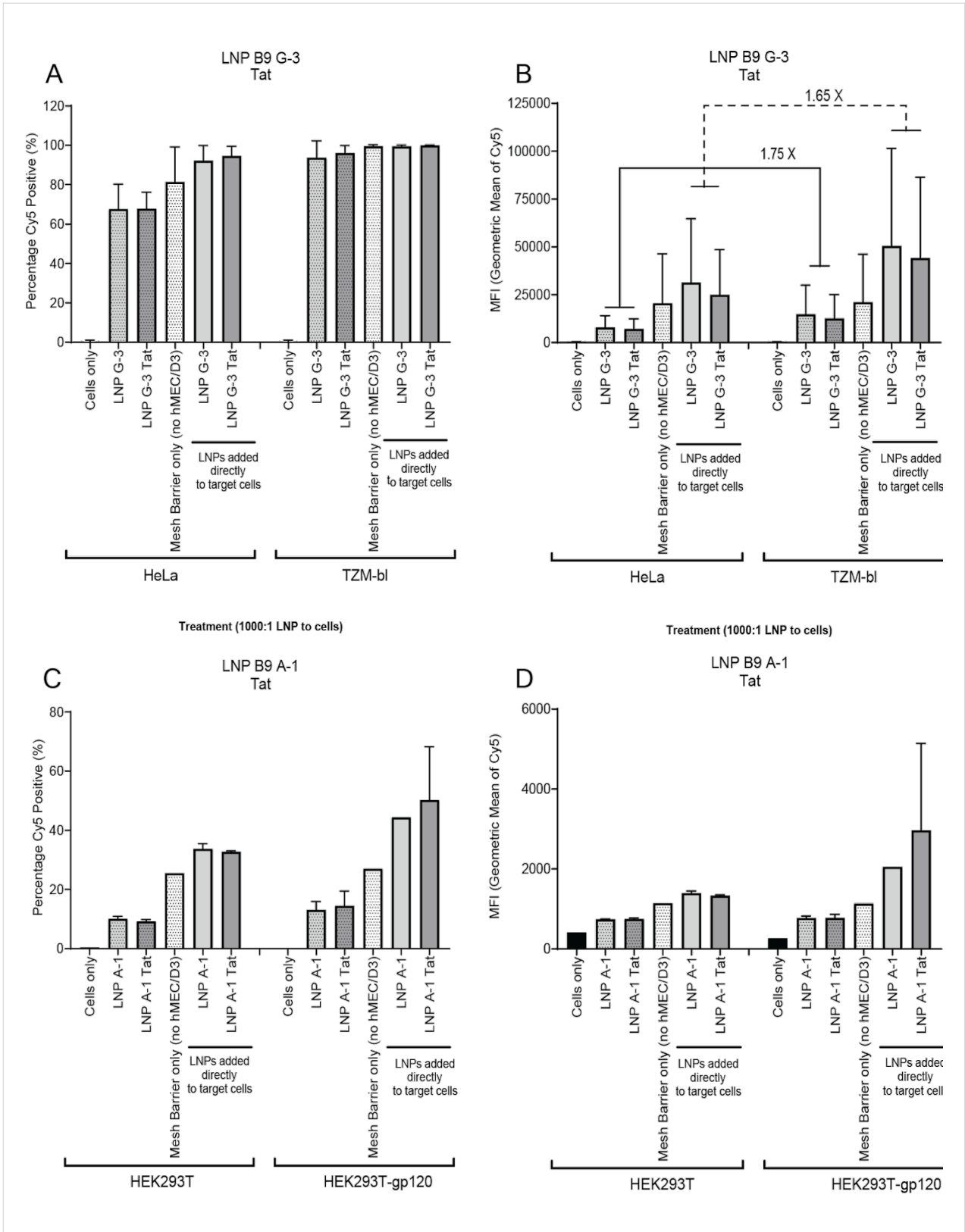


Figure 3: LNPs with Tat pass through the transwell cell barrier and are taken up by target cells. A) Percentage of cells positive for Cy5 detection in HeLa and TZM-bls. B) MFI of Cy5 in each cell population in HeLa and TZM-bls. C) Percentage cells positive for Cy5 detection in HEK cell types. D) MFI of Cy5 in each cell population in HEK cell types. Histograms are representative of two independent biological experiments, each containing duplicate technical replicates.

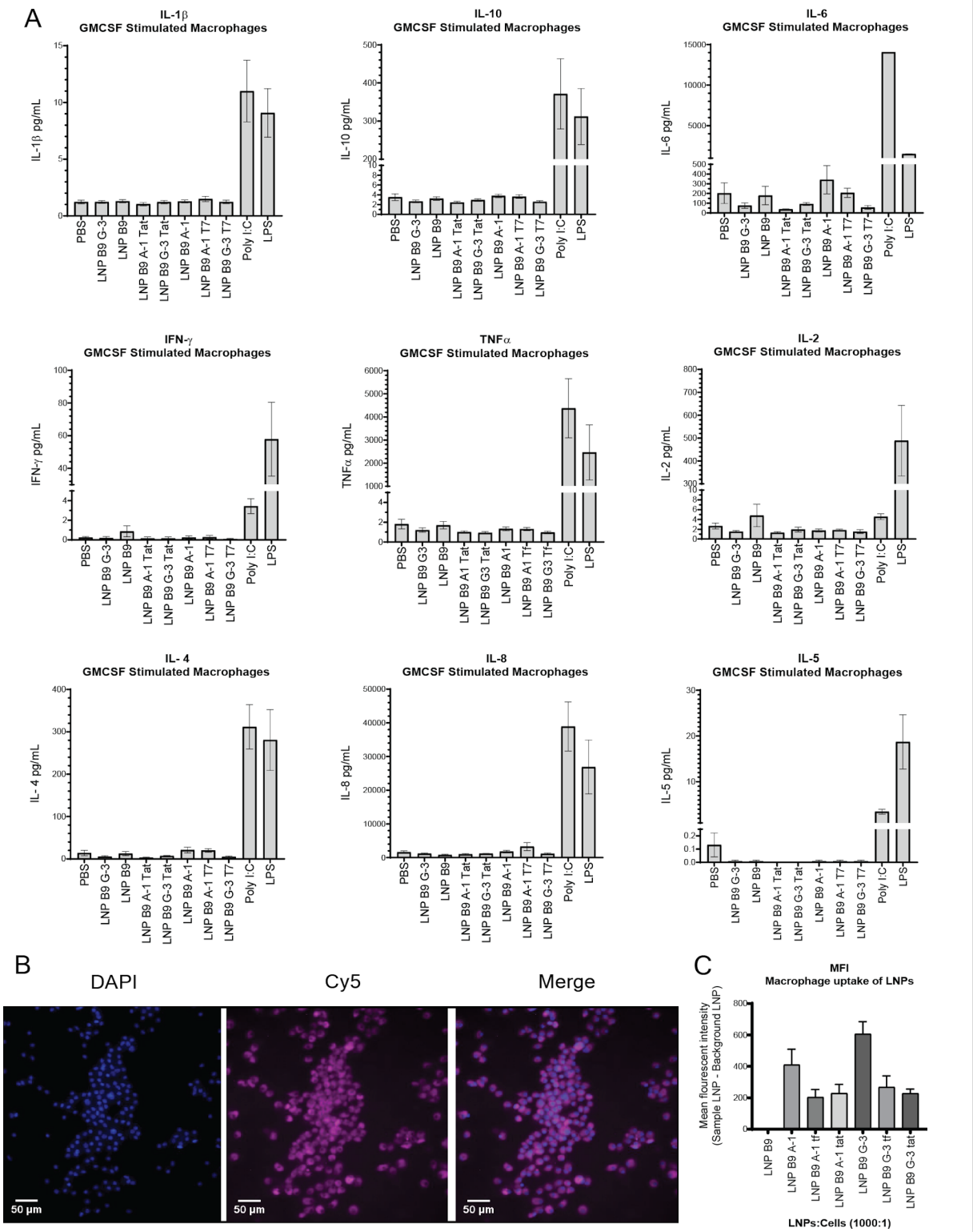


Figure 4: LNPs do not stimulate secretion of proinflammatory cytokines. A) GMCSF-primed MDMs were treated with LNPs at a ratio of LNPs/cells 1000:1 or with poly I:C or LPS for 24 h. Thereafter, supernatants were harvested, clarified, and processed for cytokine detection by Luminex. Analytes included IL-1β, IL-10, IL-6, IFN-γ, TNFα, IL-2, IL-4, IL-8, and IL-5. Histograms are representative of three biological experiments, each containing duplicate technical replicates. B) Representative fluorescent images (4′,6-diamidino-2-phenylindole (DAPI), Cy5, and merged) of macrophages and LNP G-3 after 24 h. All macrophages were 100% positive for LNP uptake independent of aptamer and peptide composition. C) However, MFI analysis using QuPath v0.2.2 suggests that the LNP G-3 had the highest uptake compared to the other LNP formulations in type-1 MDMs.

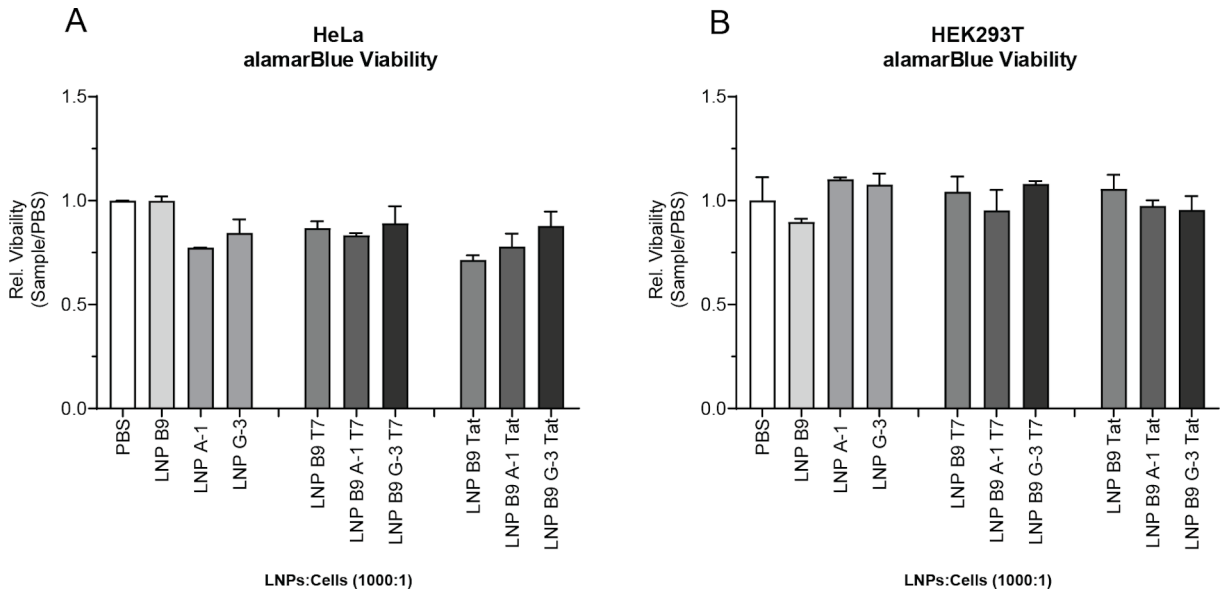


Figure 5: LNPs modestly affect cell viability in a cell-specific manner. HeLa (A) or HEK293T cells (B) were treated overnight with the LNPs at a ratio of 1000:1. Next, the alamarBlue viability assay was performed and viability as fraction of control (using PBS) was determined. Histograms are representative of the mean ± SEM. Data representative of two independent experiments performed in quadruplicates.

or T7) did not further affect cell viability (Figure 5A). This suggests that the aptamer and the peptides may contribute towards the loss of cell viability observed in this cell type. Conversely, we observed no loss of cell viability in HEK293T cells treated with LNPs containing either A-1 or G-3 aptamer, Tat or T7 alone, or the combination of aptamers and peptides (Figure 5B). As observed in the HeLa cell line, the LNP formulation alone had no effect on cell viability (Figure 5B). These data suggest that there may be some cell-specific sensitivity toward the LNPs formulations and that further studies are required to determine the optimal concentrations of aptamers and peptides within the LNPs or to optimize the ratio of LNPs to cells in order to reduce toxicity in any cell line tested.

Discussion

LNPs represent an increasingly popular modality for cargo delivery. The vast improvements in lipid design and architecture have resulted in several successful LNP-driven vaccines and therapeutics, including two RNA-based severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) vaccines [46], as well as an siRNA–LNP for the treatment of a transthyretin amyloidosis [36]. However, further improvements in toxicity profiles, cargo delivery, and cell or organ specificity are needed to expand the use of LNPs for gene and drug delivery.

LNPs and aptamers have previously been used with great success to increase cell specificity. In 2015, Liang et al. reported on a novel aptamer–LNP targeting osteoblasts. The authors conjugated aptamers to a 1,2-distearoyl-*sn*-glycero-3-phosphoethanolamine-*N*-[amino(polyethylene glycol)-2000] and observed that their LNP–DNA aptamer was able to deliver target siRNA into osteoblasts via macropinocytosis, increasing bone formation in vitro and in rodents [47]. In 2017, Kim et al. used a postinsertion method to incorporate an aptamer–maleimide–PEG into their LNPs to target epidermal growth factor receptor (EGFR)-positive cancer cells. The authors showed an increased delivery of siRNAs and fluorescent quantum dot nanocrystals both in vitro and in EGFR-positive tumor xenografts in mice [48]. In 2020, Chandra et al. used a maleimide–PEG in their LNP formulation to functionalize LNPs with an aptamer specific for the human epidermal growth factor 2 (HER2) receptor. Therein, functionalized LNPs increased siRNA delivery and subsequent sensitivity of the doxorubicin-resistant HER2-positive breast cancer cell lines by ≈2-fold over LNPs with no aptamers [49]. Taken together, the work reported by these authors, as well as by others, demonstrates the ability of aptamers to increase cellular specificity and uptake of LNPs into the target cells. In the present study, we observed that LNPs containing the G-3 aptamer targeting CCR5 resulted in a 40% increase in cellular uptake through the BBB and into target cells and that these cells had higher LNP uptake (measured by a higher MFI) than the non-antigen-expressing counterparts, while the gp160 aptamer (A-1) had no apparent effect on target cell uptake. One could speculate that this may be the result of

the nature of the target proteins. CCR5, a cell surface receptor, is internalized upon ligand binding before recycling back to the cell surface or processed for degradation in the lysosome [34]. On the other hand, gp120 is a viral surface protein that is involved in viral entry through complexation with cluster of differentiation 4 (CD4) and CCR5 or C-X-C motif chemokine receptor 4 (CXCR4) host cell surface receptors [35]. As such, gp160 expression on the host cell surface receptor may not be as adept at facilitating cell entry via receptor-mediated endocytosis. Although in 2009, Zhou et al. observed by confocal microscopy that the A-1 aptamer entered gp160-positive cells and suggested that receptor-mediated endocytosis could be the mechanism of entry, such a notion was not definitively demonstrated as the mechanism of uptake [35]. In addition, observed differences between these aptamers could also be due to differences in target receptor expression in the cell types and/or differences in the affinity and specificity of these aptamers for the target receptors and/or differences in the mechanisms of uptake. Finally, the formulation procedure also likely influences the ability of the aptamers to act as productive ligands for the respective receptors, although more studies will be needed to fully delineate these effects.

One important aspect we set out to address was to identify proxies for successful LNP-mediated cargo delivery through the BBB and into the brain. As previously stated, effective transport systems for brain drug delivery are highly warranted. Herein, we find that the LNP platform can be applied as a vehicle to circumvent the BBB and effectively deliver oligonucleotide probes to antigen-expressing cell lines. For HIV-1 there is currently a need for more effective delivery platforms compatible with antiretroviral drugs. Specifically, a productive CNS delivery of such compounds is expected to reduce HIV-1-associated neurological disorders as well as to reduce HIV-1 replication at this sanctuary site [13,50,51].

We investigated the use of T7 and Tat peptides and evaluated the ability of these to aid delivery of LNP–aptamer species across the BBB. We found that LNPs with either T7 or the Tat peptide did not significantly increase cellular uptake through the BBB above the LNPs-containing aptamers alone. T7 appeared to have an effect on cellular uptake when the LNPs were directly added to the cells and a small effect when applied through the apical chamber of the hCMEC/D3 cell line, while Tat had no effect. It may be prudent to dose the amount of postinserted Tat or T7 peptide used in these formulations. For example, in 2007, Duchardt et al. used 2–40 μM Tat peptide as a cell-penetrating peptide to facilitate siRNA entry. In particular, the authors observed that clathrin-dependent endocytosis increased with increasing concentrations of Tat peptide, suggesting that a high concentration may be needed to elicit an effective endocytosis mechanism [18]. In 2011, Qin et al. conjugated Tat to PEG 2000 and found that in their liposomal formulations, those containing 10% PEG 2000–Tat had the most efficient uptake in a BBB model [9]. Several studies have used transferrin-conjugated PEG analogs. In 2011, in a series of papers, Pang et al. observed that liposomes comprised of 5–10% PEG–transferrin increased brain delivery by 2.8-fold compared to liposomes without transferrin in a BBB model and in vivo [15] and further when loaded with the chemotherapeutic agent doxorubicin. They observed increased delivery of this compound and subsequently a significant tumor regression in mouse xenografts [22]. In 1997 and 2002, Kircheis et al. developed a polyethyleneimine (PEI)-conjugated transferrin molecule at a ratio of PEI/transferrin 21.4 nmol:270 nmol and observed that transferrin shielded the PEI, decreasing toxicity and increasing target cell uptake through binding to the transferrin receptor both in vitro and in vivo [52,53].

In the work presented here, we immobilized Tat or transferrin onto the LNP formulations using a postinsertional technique. It could be that it would be more prudent to make the LNP formulation with the addition of Tat and T7 peptide during the initial synthesis. Based on the literature [22], we assume that endocytosis plays a role in the uptake of the decorated LNPs developed herein. Further, it may be important to increase the amount of postinserted Tat and T7 used in future experiments, considering the concentration we used was relatively low (≈0.1% postaddition). Another approach is to use a next generation of short peptides that also bind to the transferrin receptor at noncompeting regions to endogenous transferrin in vivo [54]. These molecules are known as cysteine-dense peptides (CDPs) and have been shown to bind to the transferrin receptor in the picomolar range to facilitate BBB crossing in mouse models [54]. These short peptides may be advantageous to use when approaching an in vivo strategy, especially considering that the concentration of the peptide needed in the formulation may be lower compared to the T7 peptide used in this study; however, its safety profile must still be fully evaluated.

Nevertheless, our LNPs, particularly the ones containing the G-3 aptamer alone, resulted in BBB transport ranging from 50–65% in nontarget cell lines to 80–100% uptake in target cell lines, suggesting that this is a viable approach to improve uptake. Importantly, the hCMEC/D3 model represents a simplified representation of the BBB, which does not account for the full complexities of the BBB in vivo [43,55]. One could perform more complex in vitro assays that include a multicellular reconstruction of the BBB to also include astrocytes and microglial cells [56,57]. Cellular studies could also reveal in

detail the uptake mechanism of our LNPs. However, it may be more effective to perform further studies in nonprimate animal models to determine the efficacy of these LNPs in passing through the BBB. Assessing and quantitating the percentage of LNP B9 to traverse the BBB is a critical step to determine the use as an effective LNP able to deliver small molecules or oligonucleotides into the brain. One important caveat to note is that the aptamers are species-specific, and thus the use of a xenograft model with human cells in a nonprimate animal model is needed to determine the specificity of the LNP–aptamer tested.

Furthermore, while the LNP B9 alone had no effect on cellular viability, it appeared that the LNPs containing either the A-1 or G-3 aptamers, or the peptides, reduced cellular viability in HeLa cells by 20%, suggesting that there may be some toxicity when delivered to cells. However, these effects were not observed in HEK293T cells. It could be that the HeLa cell line is more sensitive than the HEK293T cell line. Nevertheless, the data suggest that further testing is required to determine the safety profile of these LNP aptamer and/or peptide formulations. One way we could reduce the toxicity profile is to chemically modify the RNA aptamers [33,58], or by reducing the aptamer concentration per LNP to thereby alleviate some of the observed cellular toxicity. It could be that the RNA aptamer itself could contribute towards cell death, possibly through stimulating the retinoic acid-inducible gene 1 (RIG-1) pathway, and it may thus be prudent to assess type-I interferons (IFN-α and IFN-β) in the future [32,59]. Importantly, the LNP B9 formulation alone had no effect on cell viability, suggesting that the ratio of cationic and ionizable lipids is optimal and does not present acute toxicity issues. However, more work is needed to assess the toxicity in vivo and, in particular, to evaluate the effect on the liver [60]. Importantly, the LNPs reported herein did not appear to stimulate an immune response in primary human monocyte-derived macrophages. Further, the addition of the aptamers and/or the peptides in the LNP formulations had no effect on immune stimulation, suggesting that these LNPs and the modifications may be well-tolerated in vivo. Importantly, both IL-6 and IFN-γ cytokines were not stimulated after exposure to the LNPs, suggesting that this LNP formulation may not induce cytokine release syndrome in vivo [60,61].

Conclusion

Taken together, we have shown that the LNP B9 formulation is safe, can traverse the BBB, and is readily taken up in multiple cell types. In the future, it will be interesting to explore whether increased uptake may also lead to increased delivery of target molecules, such as siRNA, mRNA, or small molecules. Further, having LNPs that are specific for HIV-1-infected cells or HIV-1 target cells, may help to facilitate HIV-1 drug treatment to regions of poor drug accessibility, such as the brain. More effective delivery of antiretroviral drugs may help to reduce HIV-1-associated neurological disorders that are present in HIV-1-positive individuals as well as to reduce populations of HIV-1-positive cells that are poorly accessible through current systemic drug treatment strategies.

Experimental

Materials

(6*Z*,9*Z*,28*Z*,31*Z*)-Heptatriaconta-6,9,28,31-tetraen-19-yl 4-(dimethylamino)butanoate (DLin-MC3-DMA, >98%) was purchased from D&C Chemicals (China), DSPC and cholesterol were purchased from Echelon Biosciences, Inc. (USA), and DMG-PEG 2000 was purchased from Avanti Polar Lipids, Inc. (USA). Ethanol (BioUltra, ≥99.8%), citric acid monohydrate, sodium chloride, Na_2HPO_4, and KH_2PO_4 were purchased from Sigma-Aldrich (Germany).

RNA and DNA oligonucleotides

The RNA aptamers and Cy5 DNA oligonucleotides were synthesized and purified using ion-paired and ion-exchange HPLC at the RNA/DNA synthesis core at City of Hope (Duarte, CA). The RNA aptamers, A-1 [35] and G-3 [34], were developed by Dr. Jiehua Zhou at City of Hope (Duarte, CA), and include the addition of a 3 carbon linker (*XXXXXX*), a sticky-bridge motif, and a 3' amino linker C6, 3aminoC6, at the 3' end. Annealing of the Cy5 DNA oligonucleotide to the sticky-bridge motif of the RNA aptamers was confirmed using an electromobility shift assay (EMSA) using an 8% tris/borate/ethylenediaminetetraacetic acid (TBE)-buffered gel (Novex™ Thermo Fisher Scientific, MA), under native conditions (Supporting Information File 1, Figure S1A).

RNA aptamers

- A-1 or GP160: 5'- GGG AGG **A**CG **A**UG **C**GG AA**U** **U**GA GGG **A**CC **A**CG **C**GC **U**GC **U**UG **U**UG **U**GA **U**AA G**C**A G**U**U **U**GU CG**U** GA**U** GG**C** AGA **C**GA **CUC** G**C**C **C**GA *XXXXXX G*U*A* C*A*U UCU *AGA* U*AG* **CC** /3aminoC6 -3'
- G-3 or CCR5: 5'- GGG AGG A**C**G A**U**G **C**GG G**C**C **U**UC G**U**U **U**GU **U**UC G**U**C **C**AC AGA **C**GA **CUC** G**C**C **C**GA *XXXXXX* **U***GA* **U***AG A***U**U *GA***U** *AGA* / 3aminoC6 -3'

Bold = 2'-flouronated base; italics and underlined = 2'-*O*-methyl base

Complementary DNA

- A-1: (Cy5/AGG CTA TCT AGA ATG TAC)
- G-3: (Cy5/TCT ATC AAT CTA TCA)

Peptide synthesis, purification, and characterization

Peptide assembly was carried out by solid-phase peptide synthesis in standard solid-phase extraction filtration columns. Initially, fluorenylmethyloxycarbonyl (Fmoc) group removal from the Rink linker was achieved by applying 20% piperidine in DMF (2 × 30 min). Preactivation of Fmoc amino acid (4 equiv) prior each coupling was performed with 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxide hexafluorophosphate (HATU, 4 equiv) and *N*,*N*-diisopropylethylamine (DIPEA, 6 equiv) in DMF. Then, the activated mixture was added to the resin swollen in DMF, and manual stirring was applied approximately every 15 min over a total reaction time of 2 h. The first amino acid was installed via double coupling. Fmoc deprotection was achieved via 20% piperidine in DMF (1 × 2 min and 1 × 18 min) to prepare the resin for the next coupling step. The resin was washed three times with each solvent in the given order DMF, DCM, and DMF after every reaction step.

Peptide sequences

- T7: H-HAIYPRH-NH_2
- Modified T7: dipalmitoyl-Dap-HAIYPRH-NH_2
- Tat: H-YGRKKRRQRRR-NH_2
- Modified Tat: dipalmitoyl-Dap-YGRKKRRQRRR-NH_2

Peptide conjugation with a lipid reagent

The peptides were N-terminally modified on solid support by coupling of Fmoc-Dap(Fmoc)-OH, followed by the coupling of palmitic acid to afford the complete peptide–lipid conjugates. Coupling of Fmoc-Dap(Fmoc)-OH and Fmoc deprotection were carried out as described above. To ensure the complete lipidation of the two free amines of Dap, 8 equiv of palmitic acid, 8 equiv of HATU, and 12 equiv of DIPEA in DMF were used. Cleavage of the peptide–lipid conjugates from the solid support and removal of the side-chain protecting groups was achieved by using trifluoroacetic acid (TFA)/phenol/water/triisopropylsilane (TIPS) 88:5:5:2 (3 × 60 min). After cleavage, the remaining resin was extracted with DCM (2 × 10 min). All DCM extracts and TFA cleavages were combined, and the resulting mixture was reduced under nitrogen flow. The received solid product was dissolved in DCM and subsequently reduced under nitrogen flow. This procedure was repeated two more times, followed by a lyophilization step to receive the crude peptide. The crude T7–lipid conjugate was purified by normal-phase chromatography utilizing gradient elution (2–50% MeOH in DCM). The desired modified T7 peptide was characterized via MALDI-TOF spectrometry (Bruker, MA, Supporting Information File 1, Figure S1B) and isolated as a colorless powder (9 mg, 6 μmol, 6% yield). MS (*m*/*z*): $[M + H]^+$ calcd, 1455.00; found, 1455.20. The crude Tat–lipid conjugate was precipitated from DMF as a white power and used without further purification. The modified Tat peptide was characterized via MALDI–TOF spectrometry (Supporting Information File 1, Figure S1C, 31 mg, 14 μmol, 15% yield). MS (*m*/*z*): $[M + H]^+$ calcd, 2121.48; found, 2121.17.

Lipid nanoparticle synthesis

The formulation protocol was largely adapted from Jayaraman et al. (2012) [4]. Freshly prepared lipid stocks (in chloroform) were mixed to obtain the desired mole fractions (DLin-MC3-DMA/DSPC/Cholesterol/DMG-PEG 2000 0.4:0.1:0.4:0.1), and the lipid mixture was concentrated under vacuum. The lipid film was dissolved in ethanol (20.3 mg/mL) and added dropwise to stirring 50 mM citrate buffer at pH 4.0 and preheated to 35 °C to get a final lipid concentration of 6.1 mg/mL. The lipid solution was stirred for an additional 20 min at 35 °C, after which the lipid solution was allowed to slowly reach rt, transferred to a 1 mL Hamilton syringe, and extruded 10 times at rt through two 100 nm Nucleopore membrane filters (Whatman) using Avanti Mini Extruder (Avanti Polar Lipids, Inc., USA).

Complementary oligonucleotides GP160:A-1 (1.4 nmol, 30 μL 1× PBS pH 7.4) and CCR5:G-3 (1.4 nmol, 30 μL 1× PBS pH 7.4) underwent annealing (85 °C for 10 min, 25 °C for 20 min, 4 °C for 20 min). GP160:A-1 (30 μL) and CCR5:G-3 (30 μL) were each added to a stirring LNP suspension (6.1 mg/mL, 165 μL) preheated to 35 °C, and LNP–DNA lipoplex suspensions were further diluted with 50 mM citrate buffer at pH 4.0 and 30% EtOH (120 μL), to get a final lipid concentration of 3.2 mg/mL and a DNA/lipid ratio of roughly 0.05 w/w. The LNP–DNA lipoplexes were allowed to form over 30 min at 35 °C (no stirring). Buffer exchange was performed using 3K Amicon Ultra-0.5 Centrifugal Filter Unit (Merck Millipore, USA), providing the final LNP–DNA lipoplexes in 1× PBS pH 7.4 (3 mg/mL final lipid concentration). Postinsertion of peptides was carried out by diluting the peptides to a final concentration of 1.7 μg/mL Tat lipid and 3.0 μg/mL T7 lipid in 1× PBS pH 7.4. Thereafter, diluted lipopeptides (18 μL T7, 31.8 μL Tat) were added to the LNPs (90 μL). Samples were incubated on a thermomixer for 30 minutes (25 °C at 250 rpm) for postinsertion addition. Thereafter, samples were stored at 4 °C until further use.

DLS and ZP

Particle size, polydispersity, and ZP were analyzed by DLS instrument model Zetasizer Nano ZS (Malvern Instruments, UK), having He–Ne 633 nm laser at an angle of detection of 90°, with an incubation time of 60 s. Samples were diluted 50-fold in Milli-Q water and placed into the disposable plastic cuvettes for measurement performed in triplicates ($n \geq 3$) to obtain a mean value.

NTA

Concentration and size of LNPs with and without peptides were additionally confirmed using the NanoSight NS300 device (Malvern Panalytical, UK) with the NTA software (Version 3.44, Malvern Panalytical, UK). Samples were run at a 1:1000 dilution, with three technical replicates per sample. A blue 488 nm laser was used to detect the LNPs, with a slide shutter level set to 1232 and the slider gain set to 219, and the syringe pump speed set to 30 using a flow-cell top plate module.

Cell lines and maintenance

HeLa and HEK293T cells were purchased from American Type Culture Collection (ATCC, VA). TZM-bls were acquired through the NIH AIDS reagent program and were engineered to express high levels of the HIV-1 coreceptor CCR5 [44]. HEK293T-gp160 cells were a kind gift from Dr. Bing Chen (Harvard, MA), and stably expressed the 92UG037.8 strain of the viral envelope protein, Env [45]. The human brain endothelial cell line hCMEC/D3 was purchased from Millipore Sigma (MA). HeLa, TZM-bls, HEK293T, and HEK293T-gp160 cell lines were all cultured in Dulbecco's Modified Eagle's Medium (DMEM, Corning™, NY) in 10% fetal bovine serum (FBS, GeminiBio, CA). The hCMEC/D3 cell line was maintained in EndoGRO-MV complete culture medium (Millipore Sigma, MA) on collagen (collagen type 1, rat tail, Millipore Sigma, MA)-coated flasks. hCMEC/D3 cells were cultured to a maximum of 10 passages to ensure proper tight junction formation. All cells were maintained in a water jacket incubator at 37 °C. All cell lines were routinely tested and found negative for mycoplasma.

Inflammation assay

Blood from consented and deidentified donors was used in this study under an approved IRB 19582 (City of Hope, Duarte, CA). To obtain monocytes, we followed the methodology by Menck et al. of 2014 [62]. Briefly, blood was initially processed using a Histopaque®-1077 (Millipore Sigma, MA) density separation to collect the buffy coat. Thereafter, the buffy coat was subject to a Percoll® (Cytiva, MA) density separation to enrich for the monocyte population in the buffy coat. Monocytes were counted and stored in Cyrostor-C5 (BioLife Solutions, WA) at −80 °C until further use. Monocytes were plated at a density of 1×10^5 cells per 96-well plate and stimulated for 6 days with 10 ng/mL GM-CSF (Gibco™, Thermo Fisher Scientific, MA) in Roswell Park Memorial Institute (RPMI) medium (Corning™, NY) supplemented with 5% FBS, 1% AB normal human serum (Millipore Sigma, MA), and 1% penicillin/streptomycin (Millipore Sigma, MA). The medium was replaced every 3 days. After 6 days, the medium was replaced without GM-CSF, and LNPs (ratio 1000:1), poly I:C (25 μg/mL, Millipore Sigma, MA), or LPS (1 μg/mL, Millipore Sigma, MA) was added to the macrophages. 24 h later, the supernatant was collected and centrifuged at 300*g* for 5 minutes to remove cellular debris. Harvested supernatant was stored at −80 °C until processed for cytokine expression using a 10-Plex Human Cytokine Panel (LHC6004M, Thermo Fisher Scientific, MA). The Luminex assay was processed on a Luminex® 200 machine (Luminexcorp, TX) by the Analytical Pharmacology Core (City of Hope, Duarte, CA).

Transwell assay

The transwell assay was adapted from Weksler et al. (2005) [55]. Briefly, hCMEC/D3 were cultured on presoaked 0.4 μM transwell filters (Greiner Bio-One Thincert™ CellCoat™, Austria) at a density of 5×10^4 cells/cm^2 in a 24-well culture dish. After 6 h, the medium was removed from the apical chamber and replaced with 200 μL fresh EnoGRO-MV complete culture medium (Millipore Sigma, MA). The basolateral chamber was filled with 600 μL medium. The next day, the medium was changed to a low-supplement endothelial cell growth medium-2 (EGM-2) basal medium (Lonza Walkerville, MD) supplemented with 2.5% FBS, 0.55 μM hydrocortisone (Stemcell Technologies, Canada), 1% penicillin/streptomycin (Millipore Sigma, MA), and 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, Gibco™, Thermo Fisher Scientific, MA). The culture was maintained and the medium replaced every 2nd day until a TEER of ≈30 Ω·cm^2 was reached. TEER was measured using an EVOM2 with a chopstick electrode (World Precision Instruments, FL). Resistivity was calculated using the formula given in Equation 1.

Once the integrity of the barrier was assessed, the apical transwell chambers were transferred to new 24-well culture dishes with 50,000 cells per well of HeLa, TZM-bl, HEK293T, or HEK293T-gp160 cells that had been plated 24 h previously. An LNPs/cells ratio of 1000:1 was added to each well. 24 h later, the apical layer was removed, and the basolateral cells were washed, trypsinized, and resuspended in 1× PBS. Detection of Cy5 was measured by flow cytometry on a BD Accuri™ C6 device (Becton, Dickinson and Company, NJ), and the data was analyzed using FlowJo™ Version 10.7.1 (Becton, Dickinson and Company, NJ). Cells were first gated on forward scatter-area (FSC-A) vs side scatter-area (SSC-A), followed by side scatter-height (SSC-H) vs SSC-A to gate on single cells, before designating negative and positive population gates using a histogram.

$$\text{resistivity } (\Omega \cdot \text{cm}^2) = (\text{sample } \Omega - \text{control } \Omega) \times \text{surface area of insert} \quad (1)$$

Viability assay

The alamarBlue assay was performed according to the manufacturer's instructions (Thermo Fisher Scientific, MA). Briefly, 10,000 HeLa cells and 40,000 HEK293T cells were seeded in a 96-well plate. The next day, LNPs at a ratio of 1000:1 were added to the cells. 24 h later, 0.1 volume of 10× alamarBlue was added and the cells incubated for 1 h at 37 °C. Fluorescence was measured on a GloMax® Explorer multimode microplate reader (Promega, WI). Background measurements from a medium-only control were subtracted from all the measurements before calibrating to the PBS control.

Light microscopy

To assess tight junction formation, we adapted the protocol from Vu et al. of 2009 [43]. Briefly, the apical chamber was washed with 1× PBS and fixed with ice-cold 4% paraformaldehyde for 15 minutes at 4 °C before washing two times with ice-cold PBS. The chambers were blocked with 1% bovine serum albumin (BSA)–PBS for 60 min at 4 °C and subsequently incubated overnight at 4 °C with Claudin 5–Alexa Fluor 488 (catalog number 35-258-8, Thermo Fisher Scientific, MA) at 5 µg/mL in 1% BSA–PBS. Thereafter, cells were washed three times with ice-cold PBS. The membrane was subsequently cut out of the insert with a scalpel blade and, using tweezers, placed on a slide and air-dried. Once dried, a small drop of Diamond Anti-Fade Mountant with DAPI (Invitrogen™, Thermo Fisher Scientific, MA) was added and a coverslip placed over the membrane. Slides were cured overnight at 4 °C before being visualized using a Zeiss Axio Vert A.1 light microscope with a Zeiss AxioCam 503 color camera (Carl Zeiss Microscopy GmbH, Germany). Images were processed using ZEN blue software (Version 2.3, Carl Zeiss Microscopy GmbH, Germany) and merged using ImageJ Version 1.53a (Wayne Rasband, NIH, USA).

To assess uptake of the LNPs in primary macrophages, samples were washed once with PBS and fixed with 4% ice-cold paraformaldehyde (in PBS) for 15 minutes at 4 °C. The formaldehyde was removed and the cells washed twice with ice-cold PBS. Thereafter, PBS containing DAPI (10 ng/mL) was added and the cells visualized using a Zeiss Observer II light microscope with a Zeiss AxioCam 506 Mono camera (Carl Zeiss Microscopy GmbH, Germany). Images were acquired using the ZEN blue software (Version 2.3, Carl Zeiss Microscopy GmbH, Germany). Images were processed using ImageJ Version 1.53a (Wayne Rasband, NIH, USA). To analyze the mean fluorescent intensity, we used QuPath v0.2.2 [63] (The University of Edinburgh, UK). We analyzed two different fields of view per treatment group for each donor (n = 3). For the analysis, we used the positive cell detection software with the following parameters: detection channel set to DAPI with a requested pixel size of 0.45 µm. Nucleus parameters were set to a background radius of 8 µm, a media filter radius of 1 µm, a sigma value of 3 µm, a minimum area of 10 µm^2, and a maximum area of 400 µm^2. Intensity parameters were set to a threshold of 150. Cell expansion was set to 5 µm. "Split by shape", "Include cell nucleus", "Smooth boundaries", and "Make measurements" boxes were all checked. Intensity threshold parameters were set to a single threshold with the score compartment set to cytoplasm: Alexa Fluor 647 mean. Mean cytoplasm Alexa Flour 647 values were used and represented as mean ± SEM.

Negative staining electron microscopy of LNPs

Specimens diluted at 1:1000 were absorbed onto glow-discharged, carbon-coated 200 mesh electron microscopy (EM) grids. Samples were prepared by conventional negative staining with 1% (w/v) uranyl acetate. EM images were collected with an FEI Tecnai 12 transmission electron microscope (Thermo Fisher Scientific, MA) equipped with a LaB6 filament and operated at an acceleration voltage of 120 kV. Images were recorded with a Gatan 2 × 2 k CCD camera (Gatan, Inc., CA) at a magnification of 21,000–26,000× and a defocus value of ≈1.5 µm. TEM images were analyzed using ImageJ version 1.53a (Wayne Rasband, NIH, USA). Briefly, the scale was set to the scale bar on the image, and the diameter for entire nanoparticles was measured in each image. At least 3 images per LNP formulation were used to determine the size distribution of the LNPs. Data are represented as a box and whisker plot, with min and max values representing the error bars.

Statistical analysis

Experiments are representative of two or three biological repeats performed in technical duplicates, unless otherwise stated. Data are represented as histograms with mean ± SEM. Data was prepared and analyzed using GraphPad Prism for Windows Version 8.3 (GraphPad Software, CA).

Supporting Information

Supporting Information File 1

EMSA and MALDI–TOF of oligonucleotides, TEM data for LNPs, hCMEC/D3 cell images, and FACS images.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-17-75-S1.pdf]

Acknowledgements

Research reported in this publication includes work performed by the Electron Microscopy, the RNA/DNA Synthesis, and the Analytical Pharmacology core facilities at City of Hope, Duarte,

CA. We would like to thank Dr. Zhuo Li and Ricardo Zerda at City of Hope Electron Microscopy core for their help. We would like to thank the Light Microscopy core for the use of their microscopes as well as the Analytical Cytometry core for the use of their flow cytometers (City of Hope, Duarte, CA). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Galina Shevchenko (Center for Gene Therapy and Irell & Manella Graduate School of Biological Sciences, Beckman Research Institute, City of Hope, Duarte, CA) is acknowledged for assisting with LNP preparation and tests.

Funding

This project was supported by NIMH R01 113407-01 to K. V. M. and by Villum YIP grant 40851 to K. A. The Analytical Pharmacology is supported by the National Cancer Institute of the National Institutes of Health under grant number P30CA033572.

ORCID® iDs

Roslyn M. Ray - https://orcid.org/0000-0001-6060-6969
Maria Taskova - https://orcid.org/0000-0002-9727-2496
Kevin V. Morris - https://orcid.org/0000-0002-0157-0553
Kira Astakhova - https://orcid.org/0000-0003-4878-0301

Preprint

A non-peer-reviewed version of this article has been previously published as a preprint: https://doi.org/10.3762/bxiv.2021.7.v1

References

1. Kulkarni, J. A.; Cullis, P. R.; van der Meel, R. *Nucleic Acid Ther.* **2018,** *28,* 146–157. doi:10.1089/nat.2018.0721
2. Rungta, R. L.; Choi, H. B.; Lin, P. J.; Ko, R. W.; Ashby, D.; Nair, J.; Manoharan, M.; Cullis, P. R.; Macvicar, B. A. *Mol. Ther.–Nucleic Acids* **2013,** *2,* e136.
3. Zhao, Y.; Huang, L. *Adv. Genet.* **2014,** *88,* 13–36. doi:10.1016/b978-0-12-800148-6.00002-x
4. Jayaraman, M.; Ansell, S. M.; Mui, B. L.; Tam, Y. K.; Chen, J.; Du, X.; Butler, D.; Eltepu, L.; Matsuda, S.; Narayanannair, J. K.; Rajeev, K. G.; Hafez, I. M.; Akinc, A.; Maier, M. A.; Tracy, M. A.; Cullis, P. R.; Madden, T. D.; Manoharan, M.; Hope, M. J. *Angew. Chem., Int. Ed.* **2012,** *51,* 8529–8533. doi:10.1002/anie.201203263
5. Whitehead, K. A.; Dorkin, J. R.; Vegas, A. J.; Chang, P. H.; Veiseh, O.; Matthews, J.; Fenton, O. S.; Zhang, Y.; Olejnik, K. T.; Yesilyurt, V.; Chen, D.; Barros, S.; Klebanov, B.; Novobrantseva, T.; Langer, R.; Anderson, D. G. *Nat. Commun.* **2014,** *5,* No. 4277. doi:10.1038/ncomms5277
6. Cheng, Q.; Wei, T.; Farbiak, L.; Johnson, L. T.; Dilliard, S. A.; Siegwart, D. J. *Nat. Nanotechnol.* **2020,** *15,* 313–320. doi:10.1038/s41565-020-0669-6
7. Pardi, N.; Tuyishime, S.; Muramatsu, H.; Kariko, K.; Mui, B. L.; Tam, Y. K.; Madden, T. D.; Hope, M. J.; Weissman, D. *J. Controlled Release* **2015,** *217,* 345–351. doi:10.1016/j.jconrel.2015.08.007
8. Sago, C. D.; Lokugamage, M. P.; Paunovska, K.; Vanover, D. A.; Monaco, C. M.; Shah, N. N.; Gamboa Castro, M.; Anderson, S. E.; Rudoltz, T. G.; Lando, G. N.; Munnilal Tiwari, P.; Kirschman, J. L.; Willett, N.; Jang, Y. C.; Santangelo, P. J.; Bryksin, A. V.; Dahlman, J. E. *Proc. Natl. Acad. Sci. U. S. A.* **2018,** *115,* E9944–E9952. doi:10.1073/pnas.1811276115
9. Qin, Y.; Chen, H.; Yuan, W.; Kuai, R.; Zhang, Q.; Xie, F.; Zhang, L.; Zhang, Z.; Liu, J.; He, Q. *Int. J. Pharm.* **2011,** *419,* 85–95. doi:10.1016/j.ijpharm.2011.07.021
10. Gonzalez-Carter, D.; Liu, X.; Tockary, T. A.; Dirisala, A.; Toh, K.; Anraku, Y.; Kataoka, K. *Proc. Natl. Acad. Sci. U. S. A.* **2020,** *117,* 19141–19150. doi:10.1073/pnas.2002016117
11. Guarnieri, D.; Falanga, A.; Muscetti, O.; Tarallo, R.; Fusco, S.; Galdiero, M.; Galdiero, S.; Netti, P. A. *Small* **2013,** *9,* 853–862. doi:10.1002/smll.201201870
12. Bertrand, L.; Velichkovska, M.; Toborek, M. *J. Neuroimmune Pharmacol.* **2021,** *16,* 74–89. doi:10.1007/s11481-019-09858-x
13. Osborne, O.; Peyravian, N.; Nair, M.; Daunert, S.; Toborek, M. *Trends Neurosci.* **2020,** *43,* 695–708. doi:10.1016/j.tins.2020.06.007
14. Rao, K. S.; Ghorpade, A.; Labhasetwar, V. *Expert Opin. Drug Delivery* **2009,** *6,* 771–784. doi:10.1517/17425240903081705
15. Pang, Z.; Gao, H.; Yu, Y.; Chen, J.; Guo, L.; Ren, J.; Wen, Z.; Su, J.; Jiang, X. *Int. J. Pharm.* **2011,** *415,* 284–292. doi:10.1016/j.ijpharm.2011.05.063
16. Pulgar, V. M. *Front. Neurosci.* **2019,** *12,* No. 1019. doi:10.3389/fnins.2018.01019
17. Rui, Y.; Wilson, D. R.; Green, J. J. *Trends Biotechnol.* **2019,** *37,* 281–293. doi:10.1016/j.tibtech.2018.08.010
18. Duchardt, F.; Fotin-Mleczek, M.; Schwarz, H.; Fischer, R.; Brock, R. *Traffic* **2007,** *8,* 848–866. doi:10.1111/j.1600-0854.2007.00572.x
19. Fang, S.-l.; Fan, T.-c.; Fu, H.-W.; Chen, C.-J.; Hwang, C.-S.; Hung, T.-J.; Lin, L.-Y.; Chang, M. D.-T. *PLoS One* **2013,** *8,* e57318. doi:10.1371/journal.pone.0057318
20. Vivès, E.; Brodin, P.; Lebleu, B. *J. Biol. Chem.* **1997,** *272,* 16010–16017. doi:10.1074/jbc.272.25.16010
21. Qian, Z. M.; Li, H.; Sun, H.; Ho, K. *Pharmacol. Rev.* **2002,** *54,* 561–587. doi:10.1124/pr.54.4.561
22. Pang, Z.; Gao, H.; Yu, Y.; Guo, L.; Chen, J.; Pan, S.; Ren, J.; Wen, Z.; Jiang, X. *Bioconjugate Chem.* **2011,** *22,* 1171–1180. doi:10.1021/bc200062q
23. Sharma, G.; Lakkadwala, S.; Modgil, A.; Singh, J. *Int. J. Mol. Sci.* **2016,** *17,* No. 806. doi:10.3390/ijms17060806
24. Lee, J. H.; Engler, J. A.; Collawn, J. F.; Moore, B. A. *Eur. J. Biochem.* **2001,** *268,* 2004–2012. doi:10.1046/j.1432-1327.2001.02073.x
25. Liang, M.; Gao, C.; Wang, Y.; Gong, W.; Fu, S.; Cui, L.; Zhou, Z.; Chu, X.; Zhang, Y.; Liu, Q.; Zhao, X.; Zhao, B.; Yang, M.; Li, Z.; Yang, C.; Xie, X.; Yang, Y.; Gao, C. *Drug Delivery* **2018,** *25,* 1652–1663. doi:10.1080/10717544.2018.1494223
26. Li, Y.; An, C.; Han, D.; Dang, Y.; Liu, X.; Zhang, F.; Xu, Y.; Zhong, H.; Sun, X. *New J. Chem.* **2018,** *42,* 19043–19061. doi:10.1039/c8nj04819c
27. Etzerodt, T. P.; Trier, S.; Henriksen, J. R.; Andresen, T. L. *Soft Matter* **2012,** *8,* 5933–5939. doi:10.1039/c2sm25075f
28. Gleue, L.; Schupp, J.; Zimmer, N.; Becker, E.; Frey, H.; Tuettenberg, A.; Helm, M. *Cells* **2020,** *9,* No. 2213. doi:10.3390/cells9102213
29. Fritz, T.; Voigt, M.; Worm, M.; Negwer, I.; Müller, S. S.; Kettenbach, K.; Ross, T. L.; Roesch, F.; Koynov, K.; Frey, H.; Helm, M. *Chem. – Eur. J.* **2016,** *22,* 11578–11582. doi:10.1002/chem.201602758

30. Catuogno, S.; Esposito, C. L. *Biomedicines* **2017,** *5,* No. 49. doi:10.3390/biomedicines5030049
31. Ye, M.; Hu, J.; Peng, M.; Liu, J.; Liu, J.; Liu, H.; Zhao, X.; Tan, W. *Int. J. Mol. Sci.* **2012,** *13,* 3341–3353. doi:10.3390/ijms13033341
32. Hwang, S.-Y.; Sun, H.-Y.; Lee, K.-H.; Oh, B.-H.; Cha, Y. J.; Kim, B. H.; Yoo, J.-Y. *Nucleic Acids Res.* **2012,** *40,* 2724–2733. doi:10.1093/nar/gkr1098
33. Odeh, F.; Nsairat, H.; Alshaer, W.; Ismail, M. A.; Esawi, E.; Qaqish, B.; Bawab, A. A.; Ismail, S. I. *Molecules* **2019,** *25,* No. 3. doi:10.3390/molecules25010003
34. Zhou, J.; Satheesan, S.; Li, H.; Weinberg, M. S.; Morris, K. V.; Burnett, J. C.; Rossi, J. J. *Chem. Biol.* **2015,** *22,* 379–390. doi:10.1016/j.chembiol.2015.01.005
35. Zhou, J.; Swiderski, P.; Li, H.; Zhang, J.; Neff, C. P.; Akkina, R.; Rossi, J. J. *Nucleic Acids Res.* **2009,** *37,* 3094–3109. doi:10.1093/nar/gkp185
36. Zhang, X.; Goel, V.; Robbie, G. J. *J. Clin. Pharmacol.* **2020,** *60,* 573–585. doi:10.1002/jcph.1553
37. Kulkarni, J. A.; Witzigmann, D.; Leung, J.; Tam, Y. Y. C.; Cullis, P. R. *Nanoscale* **2019,** *11,* 21733–21739. doi:10.1039/c9nr09347h
38. Suhr, O. B.; Coelho, T.; Buades, J.; Pouget, J.; Conceicao, I.; Berk, J.; Schmidt, H.; Waddington-Cruz, M.; Campistol, J. M.; Bettencourt, B. R.; Vaishnaw, A.; Gollob, J.; Adams, D. *Orphanet J. Rare Dis.* **2015,** *10,* No. 109. doi:10.1186/s13023-015-0326-6
39. Clogston, J. D.; Patri, A. K. Zeta Potential Measurement. In *Characterization of Nanoparticles Intended for Drug Delivery;* McNeil, S., Ed.; Methods in Molecular Biology (Methods and Protocols), Vol. 697; Humana Press, 2011; pp 63–70. doi:10.1007/978-1-60327-198-1_6
40. Skoglund, S.; Hedberg, J.; Yunda, E.; Godymchuk, A.; Blomberg, E.; Odnevall Wallinder, I. *PLoS One* **2017,** *12,* e0181735. doi:10.1371/journal.pone.0181735
41. Masserini, M. *ISRN Biochem.* **2013,** 238428. doi:10.1155/2013/238428
42. Jefferies, W. A.; Brandon, M. R.; Hunt, S. V.; Williams, A. F.; Gatter, K. C.; Mason, D. Y. *Nature* **1984,** *312,* 162–163. doi:10.1038/312162a0
43. Vu, K.; Weksler, B.; Romero, I.; Couraud, P.-O.; Gelli, A. *Eukaryotic Cell* **2009,** *8,* 1803–1807. doi:10.1128/ec.00240-09
44. Derdeyn, C. A.; Decker, J. M.; Sfakianos, J. N.; Wu, X.; O'Brien, W. A.; Ratner, L.; Kappes, J. C.; Shaw, G. M.; Hunter, E. *J. Virol.* **2000,** *74,* 8358–8367. doi:10.1128/jvi.74.18.8358-8367.2000
45. Chen, J.; Kovacs, J. M.; Peng, H.; Rits-Volloch, S.; Lu, J.; Park, D.; Zablowsky, E.; Seaman, M. S.; Chen, B. *Science* **2015,** *349,* 191–195. doi:10.1126/science.aaa9804
46. Dong, Y.; Dai, T.; Wei, Y.; Zhang, L.; Zheng, M.; Zhou, F. *Signal Transduction Targeted Ther.* **2020,** *5,* No. 237. doi:10.1038/s41392-020-00352-y
47. Liang, C.; Guo, B.; Wu, H.; Shao, N.; Li, D.; Liu, J.; Dang, L.; Wang, C.; Li, H.; Li, S.; Lau, W. K.; Cao, Y.; Yang, Z.; Lu, C.; He, X.; Au, D. W. T.; Pan, X.; Zhang, B.-T.; Lu, C.; Zhang, H.; Yue, K.; Qian, A.; Shang, P.; Xu, J.; Xiao, L.; Bian, Z.; Tan, W.; Liang, Z.; He, F.; Zhang, L.; Lu, A.; Zhang, G. *Nat. Med.* **2015,** *21,* 288–294. doi:10.1038/nm.3791
48. Kim, M. W.; Jeong, H. Y.; Kang, S. J.; Choi, M. J.; You, Y. M.; Im, C. S.; Lee, T. S.; Song, I. H.; Lee, C. G.; Rhee, K.-J.; Lee, Y. K.; Park, Y. S. *Sci. Rep.* **2017,** *7,* No. 9474. doi:10.1038/s41598-017-09555-w
49. Chandra, S.; Nguyen, H. M.; Wiltz, K.; Hall, N.; Chaudhry, S.; Olverson, G.; Mandal, T.; Dash, S.; Kundu, A. *J. Cancer Treat. Diagn.* **2020,** *4,* 1–13.
50. Gong, Y.; Chowdhury, P.; Nagesh, P. K. B.; Rahman, M. A.; Zhi, K.; Yallapu, M. M.; Kumar, S. *Sci. Rep.* **2020,** *10,* No. 3835. doi:10.1038/s41598-020-60684-1
51. Nowacek, A.; Gendelman, H. E. *Nanomedicine (London, U. K.)* **2009,** *4,* 557–574. doi:10.2217/nnm.09.38
52. Kircheis, R.; Kichler, A.; Wallner, G.; Kursa, M.; Ogris, M.; Felzmann, T.; Buchberger, M.; Wagner, E. *Gene Ther.* **1997,** *4,* 409–418. doi:10.1038/sj.gt.3300418
53. Kircheis, R.; Wightman, L.; Kursa, M.; Ostermann, E.; Wagner, E. *Gene Ther.* **2002,** *9,* 731–735. doi:10.1038/sj.gt.3301748
54. Crook, Z. R.; Girard, E.; Sevilla, G. P.; Merrill, M.; Friend, D.; Rupert, P. B.; Pakiam, F.; Nguyen, E.; Yin, C.; Ruff, R. O.; Hopping, G.; Strand, A. D.; Finton, K. A. K.; Coxon, M.; Mhyre, A. J.; Strong, R. K.; Olson, J. M. *J. Mol. Biol.* **2020,** *432,* 3989–4009. doi:10.1016/j.jmb.2020.04.002
55. Weksler, B. B.; Subileau, E. A.; Perrière, N.; Charneau, P.; Holloway, K.; Leveque, M.; Tricoire-Leignel, H.; Nicotra, A.; Bourdoulous, S.; Turowski, P.; Male, D. K.; Roux, F.; Greenwood, J.; Romero, I. A.; Couraud, P. O. *FASEB J.* **2005,** *19,* 1872–1874. doi:10.1096/fj.04-3458fje
56. Stone, N. L.; England, T. J.; O'Sullivan, S. E. *Front. Cell. Neurosci.* **2019,** *13,* No. 230. doi:10.3389/fncel.2019.00230
57. Bagchi, S.; Chhibber, T.; Lahooti, B.; Verma, A.; Borse, V.; Jayant, R. D. *Drug Des., Dev. Ther.* **2019,** *13,* 3591–3605. doi:10.2147/dddt.s218708
58. Lee, Y.; Urban, J. H.; Xu, L.; Sullenger, B. A.; Lee, J. *Nucleic Acid Ther.* **2016,** *26,* 173–182. doi:10.1089/nat.2015.0575
59. Kohlway, A.; Luo, D.; Rawling, D. C.; Ding, S. C.; Pyle, A. M. *EMBO Rep.* **2013,** *14,* 772–779. doi:10.1038/embor.2013.108
60. Moss, K. H.; Popova, P.; Hadrup, S. R.; Astakhova, K.; Taskova, M. *Mol. Pharmaceutics* **2019,** *16,* 2265–2277. doi:10.1021/acs.molpharmaceut.8b01290
61. Lee, D. W.; Gardner, R.; Porter, D. L.; Louis, C. U.; Ahmed, N.; Jensen, M.; Grupp, S. A.; Mackall, C. L. *Blood* **2014,** *124,* 188–195. doi:10.1182/blood-2014-05-552729
62. Menck, K.; Behme, D.; Pantke, M.; Reiling, N.; Binder, C.; Pukrop, T.; Klemm, F. *J. Visualized Exp.* **2014,** No. 91, e51554. doi:10.3791/51554
63. Bankhead, P.; Loughrey, M. B.; Fernández, J. A.; Dombrowski, Y.; McArt, D. G.; Dunne, P. D.; McQuaid, S.; Gray, R. T.; Murray, L. J.; Coleman, H. G.; James, J. A.; Salto-Tellez, M.; Hamilton, P. W. *Sci. Rep.* **2017,** *7,* No. 16878. doi:10.1038/s41598-017-17204-5

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Beyond ribose and phosphate: Selected nucleic acid modifications for structure–function investigations and therapeutic applications

Christopher Liczner[1], Kieran Duke[1], Gabrielle Juneau[1], Martin Egli[*2] and Christopher J. Wilds[*1]

Review

Open Access

Address:
[1]Department of Chemistry and Biochemistry, Concordia University, Montréal, Québec H4B 1R6, Canada and [2]Department of Biochemistry, Vanderbilt Institute of Chemical Biology, and Center for Structural Biology, School of Medicine, Vanderbilt University, Nashville, Tennessee 37232, United States

Email:
Martin Egli[*] - martin.egli@vanderbilt.edu; Christopher J. Wilds[*] - chris.wilds@concordia.ca

* Corresponding author

Keywords:
antisense; chemically modified oligonucleotides; crystallography; siRNA; structure

Beilstein J. Org. Chem. **2021,** *17,* 908–931.
https://doi.org/10.3762/bjoc.17.76

Received: 12 February 2021
Accepted: 14 April 2021
Published: 28 April 2021

This article is part of the thematic issue "Celebrating the role of chemistry in the success of oligonucleotides as therapeutics".

Guest Editors: P. Kumar and T. Brown

Abstract

Over the past 25 years, the acceleration of achievements in the development of oligonucleotide-based therapeutics has resulted in numerous new drugs making it to the market for the treatment of various diseases. Oligonucleotides with alterations to their scaffold, prepared with modified nucleosides and solid-phase synthesis, have yielded molecules with interesting biophysical properties that bind to their targets and are tolerated by the cellular machinery to elicit a therapeutic outcome. Structural techniques, such as crystallography, have provided insights to rationalize numerous properties including binding affinity, nuclease stability, and trends observed in the gene silencing. In this review, we discuss the chemistry, biophysical, and structural properties of a number of chemically modified oligonucleotides that have been explored for gene silencing.

Introduction

The natural nucleic acids sugar-phosphate backbone comes in two flavors, 2'-deoxyribose in DNA and ribose in RNA. However, this relative simplicity combined with the five natural bases, adenine (A), cytosine (C), guanine (G), thymine (T) and uracil (U, in RNA) belies the fact that both DNA and RNA are decorated with chemical modifications. For a catalogue of natural modifications in DNA, see https://dnamod.hoffmanlab.org/ [1], and in RNA, see https://iimcb.genesilico.pl/modomics/ [2]. In DNA, base modifications are much more common than those in the backbone and play a

central role in epigenetics, such as, for example, the 'fifth base' 5-methylcytosine (5mC) [3]. In the backbone, chemical modification appears to be limited to the phosphorothioate *R*p-stereoisomer (*R*p-PS, i.e., phosphate with one of the non-bridging oxygens replaced by sulfur) in bacterial genomes, where it may serve a protective role against nucleases [4] and its loss results in genomic instability [5]. There are over a hundred known base modifications in RNA and the *R*p-PS backbone modification occurs in ribosomal RNA (rRNA) of both pro- and eukaryotes [6]. A very common natural modification that concerns the ribose moiety is 2'-*O*-methylation (2'-*O*Me). 2'-*O*Me nucleotides are scattered throughout all types of RNA, including mRNA, tRNA, rRNA, snRNA, snoRNA, miRNA and viral RNA [7-9]. Moreover, the modification occurs irrespective of the nature of the base and is therefore also referred to as Nm (N = A, C, G, 5mU, U, ψU, I, etc.) [10].

The specific role(s) an individual modification plays is often not known, but we can surmise involvements in transcription, translation, replication, splicing and other fundamental processes in biological information transfer. More specifically, they can affect chemical and thermodynamic stability, folding, secondary and tertiary structure, activity and interactions between nucleic acids, proteins and receptors. Particularly, as far as improving metabolic stability, pairing properties (RNA affinity), protein binding and transport/cellular uptake are concerned, chemical modifications are a prerequisite for the discovery and development of oligonucleotide therapeutics [11-15]. Thus, the natural PS and 2'-*O*Me backbone modifications provide improved resistance to degradation by exo- and endonucleases and they both affect protein binding [16,17]. Eight of the now approved 13 oligonucleotide drugs feature the PS modification in the backbone and all four approved siRNA therapeutics: ONPATTRO® (patisiran, 2018), GIVLAARI® (givosiran, 2019), OXLUMO® (lumasiran, 2020) and LEQVIO® (inclisiran, 2020) have 2'-*O*Me modifications [18-21] (https://www.oligotherapeutics.org/20th-anniversary-of-rna-interference-in-mammalian-cells/). Interestingly, both 2'-*O*Me [22] and PS [23] date back to the 1960s and constitute the earliest modifications reported by chemists along with the synthesis of 2'-deoxy-2'-fluoro-nucleosides (FRNA) [24].

The negatively charged phosphodiester linkages in the backbones of DNA and RNA are of fundamental importance for reactivity, stability, conformation and hydration [25,26]. The sugar moieties in DNA and RNA determine the shape of the double helix, i.e., the facile flip between the C2'-*endo* (B-form DNA) and C3'-*endo* (A-form DNA) puckers by deoxyribose and the shift toward the C3'-*endo* pucker due to the presence of the 2'-OH in RNA [27,28]. As well, the seemingly small difference of a single hydroxy group between the sugars in DNA and RNA is at the origin of the vastly expanded fold [29-32] and functional spaces of RNA [33-39]. Perhaps less known is the fact that the sugar moiety in the backbone of a nucleic acid determines the base pairing priorities. For example, in DNA G:C > A:T whereas in homo-DNA (2',3'-β-D-dideoxyglucopyranose nucleic acid) G:C > A:A ≈ G:G > A:T (reverse Hoogsteen A:A and G:G pairs) ([40] and cited references). Messenger RNA is the target of both the antisense and RNAi strategies to interfere with biological information transfer prior to production of proteins, enzymes and receptors that may be inhibited by small-molecule and antibody therapeutics. However, native RNA oligonucleotides do not possess sufficient metabolic stability for in vivo applications. Therefore, chemical modification is absolutely essential to re-engineer RNA into a therapeutic tool [15].

The chemical make-up of RNA, i.e., the ribose-phosphate backbone, has inspired countless strategies to chemically modify either the sugar [12,41-44], or the phosphate (e.g., amide-RNA [45]), or both [46,47]. In addition, the ribose has been replaced with alternative sugar moieties, such as a tetrose (L-α-threofuranose, TNA [48]), and hexoses (e.g., hexitol, HNA [49]; altritol AtNA [50]; xylol XyNA [51]), or cyclohexene (CeNA [52]), a morpholino moiety (PMO [53]), and an acyclic, chiral glycol linker (GNA [54]), to generate so-called xeno nucleic acids (XNAs [55,56]). In arguably the most radical alternative nucleic acid pairing system, peptide nucleic acid (PNA), the sugar-phosphate backbone is replaced by an amide-based, neutral and achiral scaffold that allows cross-pairing with both DNA and RNA as well as formation of double- and triple-stranded species [57]. Despite this growing universe of modifications, 2'-modifications, such as the original 2'-*O*Me, 2'-*O*-(2-methoxyethyl) (MOE [58,59]), and locked nucleic acid (LNA [44,60]) as well as the FRNA analogue [61-63] along with the phosphorothioates, will likely remain critical for the development of new oligonucleotide-based therapeutics. In the present review, we will summarize the properties of selected backbone modifications (Figure 1) and discuss investigations regarding their structure and function and, if applicable, their importance for therapeutic applications.

Review

Internucleotide linkage modifications

N3' → P5' phosphoramidate

The N3' → P5' phosphoramidate DNA (3'-NP DNA) contains a negatively charged internucleotide linkage, but one of the bridging oxygens is replaced by a nitrogen (Figure 1A). The 3'-NP linkage is generated during solid-phase synthesis where the incoming protected 5'-DMT-3'-aminonucleoside couples to the 5'-H-phosphonate in the presence of a base (Scheme 1) [64].

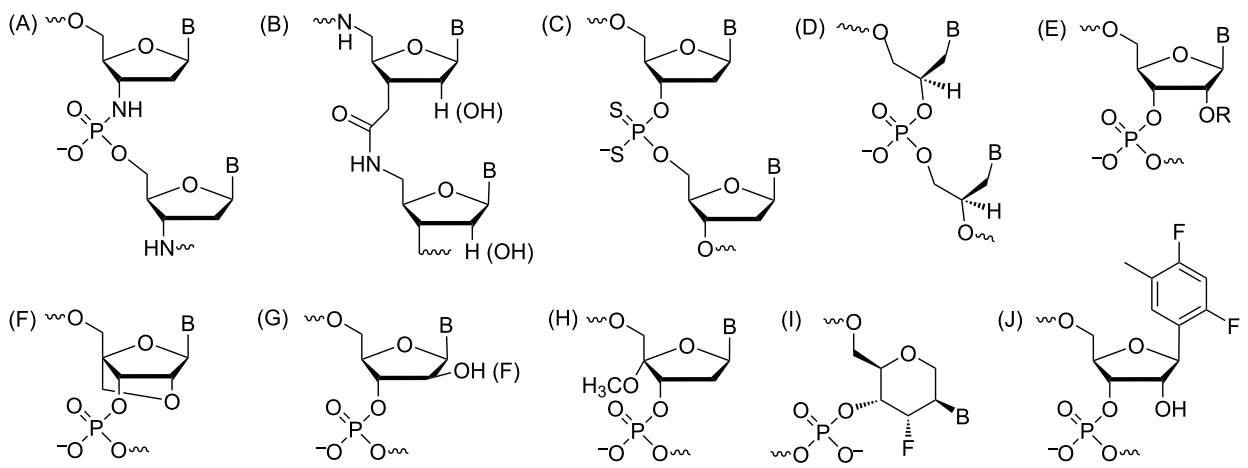
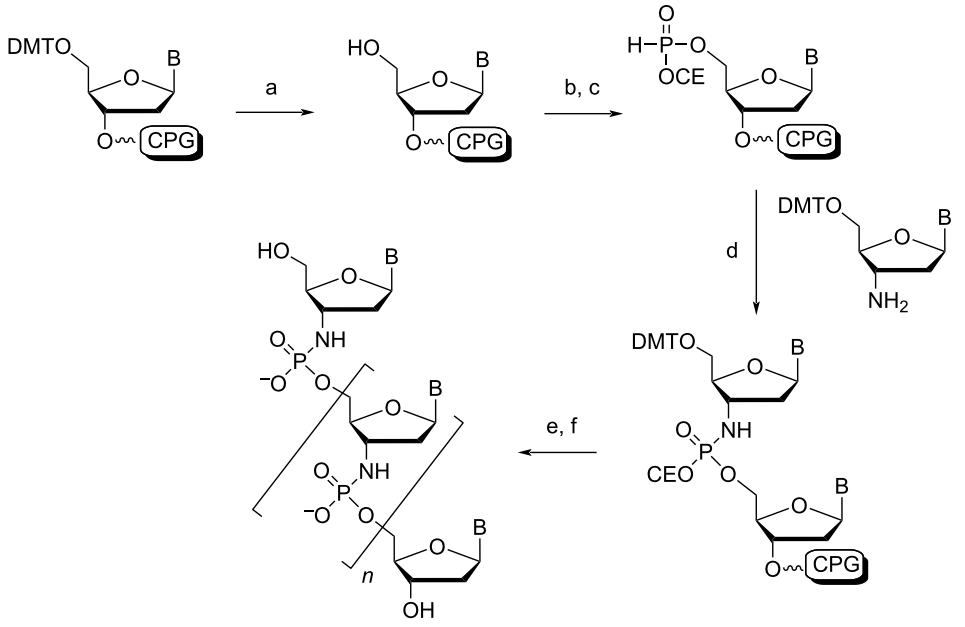


Figure 1: Structures of the chemically modified oligonucleotides (A) N3' → P5' phosphoramidate linkage, (B) amide (AM1) linkage, (C) phosphorodithioate (PS2), (D) glycol nucleic acid (*R*-isomer), (E) 2'-*O*-alkyl modifications (R = -CH_3, -$CH_2CH_2OCH_3$), (F) locked nucleic acids (LNA)/bridged nucleic acids (BNA), (G) arabinose (ANA) and arabinofluoro (FANA) nucleic acids, (H) C4'-modified nucleic acids, (I) 3'-fluorohexitol nucleic acid, (J) ribo-difluorotoluyl-modified nucleic acid.



Scheme 1: Synthesis of a N3' → P5' phosphoramidate linkage by solid-phase synthesis. (a) dichloroacetic acid; (b) ClP($NiPr_2$)(OCE); (c) tetrazole/water; (d) triethylamine/carbon tetrachloride; (e) repeat steps a–d; (f) detritylate then deprotect with NH_3. DMT = dimethoxytrityl, CPG = succinyl-linked long chain alkylamine controlled pore glass solid support, CE = 2-cyanoethyl. Adapted from [64].

In comparison with natural phosphodiester oligonucleotides, these modified oligonucleotides display improved nuclease resistance and an enhanced duplex thermal stability of 2.3–2.6 °C per linkage independent of nucleotide sequence and base composition [65]. The presence of alternating phosphodiester and phosphoramidate linkages within an oligonucleotide resulted in improved binding to RNA relative to DNA. Homopyrimidine 3'-NP DNA forms a stable triplex at neutral pH with double-stranded DNA and RNA [64-66].

These attributes, nuclease stability, and hybridization to single and double stranded nucleic acid targets have led to studies to investigate 3'-NP DNA for antisense and antigene purposes. For example, as an antisense agent in the treatment of human leukemia [67], as an inhibitor of transcription elongation targeted to proviral HIV DNA [68], and as a triplex-forming oligonucleotide that selectively binds a sequence within the chromatin structure of cell nuclei [69]. Remarkably, 3'-NP DNA can also act as an RNA mimic in interactions with binding pro-

teins despite lacking a ribose moiety, making them useful nuclease-resistant probes for studying RNA–protein interactions [70].

To better elucidate the structural features of 3'-NP DNA responsible for this enhanced selective binding and stability, the Egli group determined the crystal structure of the fully modified 3'-NP DNA duplex with the sequence 5'-d(CnpGnpCnpGnpAnpApTnpTnpCnpGnpCnpG)-3' at 2 Å resolution [71]. It was found that the overall duplex structure adopted by 3'-NP DNA resembles that of an RNA-like A-form double helix. The deoxyribose ring of phosphoramidate DNA is locked in a northern (C3'-*endo*) conformation due to the decreased gauche effect between 4'-O and the 3'-N compared to the 4'-O and 3'-O interactions in DNA. The 3'-amino moieties in the structure's backbone were found to coordinate a larger amount of water molecules, on both the backbone and at groove sites. This increased hydration, as well as the configuration of the 3'-amino group enables the hydrogen atom to orient towards anions (chloride) in the vicinity and the 3'-nitrogen lone pair engages in a lp → σ* anomeric effect with the antibonding orbital from the adjacent P–O5' bond (Figure 2A). This conjugation is surmised to cause considerably increased rigidity of the phosphoramidate sugar-phosphate backbone relative to native phosphodiester oligomers. This N-type sugar puckering and increased hydration of the sugar phosphate backbone could also account for the triplex-favoring properties of this modification [72].

Amide

While many amide backbone oligonucleotide variants exist, the focus of this review will be on the AM1-type shown in Figure 1B, as this is the most studied and therapeutically promising modification of its class (a summary of other amide variations can be found elsewhere [74,75]). The strategy used to incorporate this modification into DNA or RNA has been to first synthesize the nucleoside dimer phosphoramidite with the appropriate amide linkage, which can then be introduced into the strand by solid-phase synthesis. These dimers are synthesized by using an amide coupling reagent to condense a 3'-carboxylic acid nucleoside with a 5'-amine nucleoside, where the necessary protecting groups are present on the nucleobase and sugar moieties [76,77].

Unlike the phosphodiester linkage of natural DNA, the AM1 modification is an example of a non-ionic backbone. The crystal structure of a 13-mer RNA duplex with a single central AM1 modification revealed that this modification is accommodated in an A'-form duplex [75]. Interestingly, an unconventional C–H···O hydrogen bond was observed between the amide's carbonyl oxygen and the nearby uracil C6–H6. The thermal stability of this modified duplex was, however, quite similar to native RNA. Typically, there is a decrease of 0.2–0.8 °C in the thermal stability of RNA/DNA hybrid duplexes for each AM1 modification [11,78]. NMR structural studies have shown that the AM1 modification is well tolerated in an RNA duplex, with

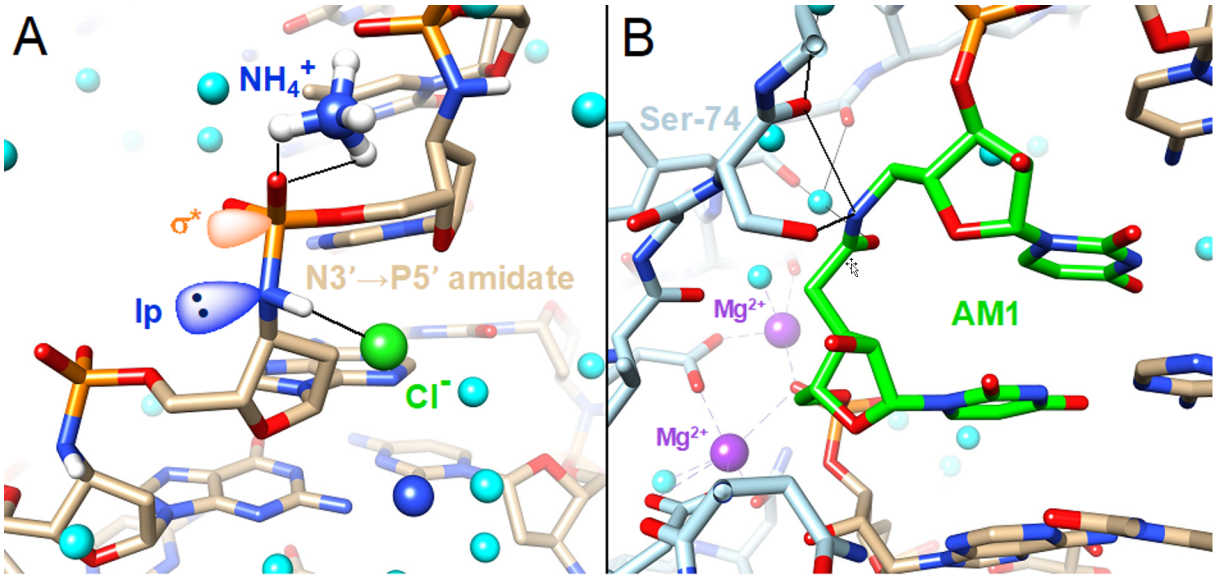


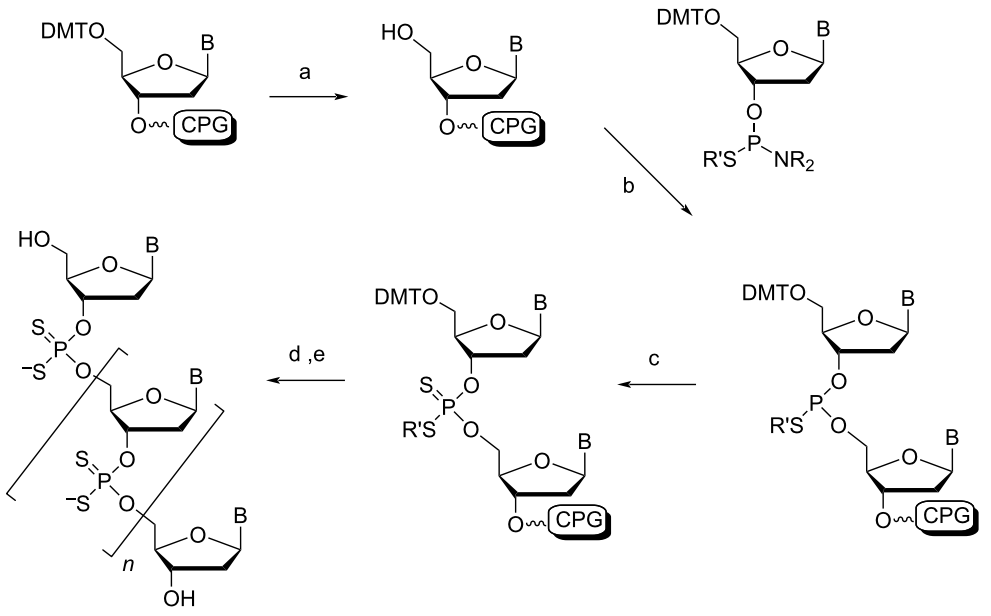
Figure 2: Crystal structures of (A) N3' → P5' phosphoramidate DNA (PDB ID 363D) [71] and (B) amide (AM1) RNA in complex with *Bacillus halodurans* RNase H (PDB ID 5VAJ) [73]. The relative orientation of the N3' n and P–O5' σ* orbitals in the backbone of 3'-NP DNA are consistent with an anomeric effect. The 3'-nitrogen is H-bonded to a chloride anion (green sphere) and the phosphate group forms a salt bridge to ammonium. Water molecules are cyan spheres and H-bonds are drawn with thin lines.

little effect on the global structure [79]. Furthermore, siRNA duplexes with amide modifications at the 3'-overhang region show enhanced endonuclease and 3'-exonuclease resistance [80]. Thus far, the AM1 modification has not found great success in antisense therapeutics, owing to RNAse H not recognizing a uniformly modified AM1-DNA:RNA heteroduplex. Recently, however, an 18-mer AM1-DNA gapmer was synthesized, with 4 AM1 linkages on each flank of the oligonucleotide [81]. Once bound to its RNA target, RNAse H was able to completely degrade the RNA in just 30 minutes, demonstrating the effectiveness of AM1 modifications in chimeric oligonucleotides for antisense therapeutics.

While this lack of charge was also believed to render AM1-RNA incompatible with siRNA therapeutic strategies, as there was crystallographic data [82] that showed the main interaction between the phosphates of the RNA duplex and the Ago2 protein is electrostatic in nature, this was, however, not the case, owing to the observed increase in silencing activity for AM1-modified siRNAs with amide linkages at specific sites [75]. Structural insight into this observation was obtained using the crystal structure of the complex between *Bacillus halodurans* RNase H and the r(GAC ACC UGA UAM1UC) - d(GAA TCA GGT GTC) hybrid duplex [73]. Compared to the native complex, conformational changes in the RNA and protein were only observed around the site of the AM1 modification. Not only was the amide an ideal structural mimic of phosphate, it also possessed stabilizing hydrogen bonds between the amide N–H and the main chain oxygen and side chain Oγ of S74 (Figure 2B), explaining their tolerance towards efficient recognition by Ago2. Interestingly, however, disfavoring stabilizing interactions with Ago2 through an amide backbone modification can be therapeutically beneficial when placed in the proper site. This was exemplified by a recent study that placed a single AM1 backbone modification between nucleotides 1 and 2 at the 5'-end of the siRNA passenger strand, whereby the off-target effects of that strand were abolished and the activity of the guide strand was restored [83].

Phosphorodithioate

The synthesis of phosphorodithioate (PS2)-modified oligonucleotides was first described in 1991 by the Caruthers group [84]. Typically, each 2'-deoxynucleoside 3'-phosphorothioamidite is prepared by phosphitylating the protected nucleosides with tris(pyrrolidino)phosphine under tetrazole catalysis, followed by immediate treatment with monobenzoylethanedithiol. The 3'-phosphorothioamidites are incorporated into an oligonucleotide by standard solid-phase synthesis conditions, however, the oxidation step is replaced with sulfurization by elemental sulfur (Scheme 2) [85]. It should be noted that more efficient sulfurization agents exist with faster kinetics and higher solubility in organic solvents, useful for automated synthesis, such as the Beaucage reagent [86]. Conveniently, during deprotection of the support-bound oligonucleotide, aminolysis removes the β-thiobenzoylethyl group from the backbone to generate the free PS2-modified oligonucleotide. This modification is achiral at the phosphorus atom (Figure 1C), and thus, unlike the phosphoromonothioate (PS) analogues (extensively



Scheme 2: Synthesis of a phosphorodithioate linkage by solid-phase synthesis. (a) detritylation; (b) tetrazole; (c) sulfurization, capping, then washing; (d) repeat steps a–c; (e) detritylate then deprotect with NH_4OH. R = pyrrolidino, R' = β-thiobenzoylethyl. Adapted from [85].

covered in other reviews [18,42,87,88]), the synthesized oligonucleotide is stereochemically pure. This simplifies their purification, as there is no longer the need to separate biochemically distinct diastereomers in order to make meaningful conclusions about the modification in a therapeutic or crystallographic context (although individual PS diastereoisomeric linkages can be resolved in electron density maps at sufficiently high resolution [18,89]). This modification has been attractive in antisense therapeutics as these altered oligonucleotides can form a hybrid duplex with unmodified RNA, which is recognized by RNase H [89,90].

While the thermal stability of PS2-modified RNA duplexes slightly decreases compared to the unmodified duplex, there is an increase in nuclease stability, even relative to PS-modified duplexes [91]. Crystal structures of PS2-modified RNA duplexes were determined to be isomorphous to their native RNA counterpart, causing no perturbation in the ribose sugar conformation, nor the torsion angles of the backbone [92]. More interestingly, siRNA duplexes with PS2-modified sense strands showed an increase in binding affinity towards the Ago2 protein of the RISC complex [92,93]. The model based on the crystal structure of human Ago2 bound to an siRNA duplex demonstrated that PS2 moieties near the 3'-terminus of the sense strand lie in the vicinity of a hydrophobic patch that is surrounded by lysine and arginine residues [15]. The latter generate an electric field that could polarize sulfur atoms (the PS2 group still carries a negative charge), thereby enhancing the interaction of the PS2 moiety with the edge of phenylalanine as seen in the complex between PS2-modified anti-thrombin aptamer and thrombin [94] (Figure 3).

Commonly, internucleotide-modified oligonucleotides are coupled with 2'-substitutions in order to enhance or regain desirable therapeutic properties. For example, not only did introducing a 2'-*O*Me modification at the PS2 nucleotide sites of an siRNA duplex sense strand increase the thermal stability of the duplex to levels comparable to the unmodified variant, it also further improved the binding affinity to the Ago2 protein, hypothesized to be in part caused by a superior hydrophobic effect [92].

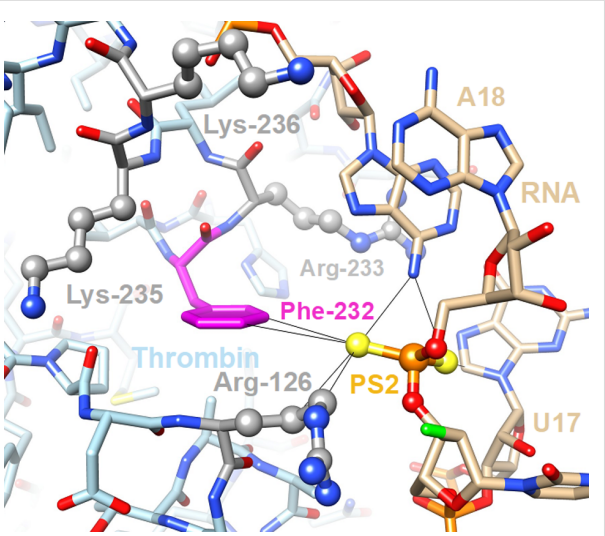


Figure 3: Close-up view of a key interaction between the PS2-modified antithrombin RNA aptamer and thrombin in the crystal structure of the complex (PDB ID 5DO4) [95]. An RNA-induced fit brings the PS2 moiety in close contact with the edge of Phe-232 (magenta carbon atoms) that forms a hydrophobic patch surrounded by four basic residues (side chains highlighted in ball-and-stick mode with carbon atoms colored in gray). These arginine and lysine residues generate an electric field that polarizes the thiophosphate moiety, thereby contributing to the 1000-fold tighter binding of the PS2-modified RNA to thrombin relative to the parent aptamer.

Glycol nucleic acid

Glycol nucleic acid (GNA) with its chiral, acyclic three-carbon backbone linked by phosphate is the simplest phosphodiester-based nucleic acid analogue (Figure 1D). It contains one stereocenter allowing for the synthesis of either (*S*)-GNA or (*R*)-GNA where chirality is fixed by use of either (*R*) or (*S*) starting material, respectively. These simple nucleic acid building blocks were first synthesized in 1971 by Ueda et al. [96]. The group was able to synthesize adenine, cytosine, and uracil GNA analogues by reacting these bases with glycerol α-chlorohydrin or glycidol. The following year, the Seita group showed that thymine and guanine analogues could be prepared in a similar fashion [97]. Interestingly, both groups found that condensation of purine bases to yield GNA derivatives gave two dihydroxypropylated isomers: the N3 (I) and the N9 (II) dihydroxypropylated isomers. Using glycerol α-chlorohydrin, the ratio of I/II was 1:4 with II being the preferred isomer but when using glycidol, this ratio shifted to 3:1 in the favor of the desired isomer [96,97]. From there on, the use of glycidol for the preparation of GNA analogues became the gold standard. The first GNA polymers were obtained through condensation with *N*,*N*-dicyclohexylcarbodiimide (DCC) giving rise to homopolymeric tetramers of either G-GNA or T-GNA [97]. In 1996, Acevedo and Andrews were the first to demonstrate the synthesis of GNA nucleoside phosphoramidite derivatives as well as the ability of the phosphoramidite derivatives to withstand solid-phase conditions, inevitably laying the groundwork for GNA solid-phase synthesis [98]. Using the glycidol approach, Zhang et al. synthesized 18-mer oligonucleotides containing GNA-T monomers [99]. Starting from (*R*)-glycidol, the free hydroxy group is tritylated. The resulting product is then reacted with unprotected thymine which, in the presence of stoichiometric amounts of sodium hydride, results in the epoxide ring opening and the formation of the glycol backbone. The preamidite is then phosphitylated yielding the desired GNA-T

amidite (Scheme 3). Recently, this simple acyclic nucleic acid backbone is of interest as a prospective evolutionary precursor of RNA [100]. Furthermore, GNA analogues with N2' → P3' phosphoramidate linkages have been studied as a potential alternative genetic system and they have been incorporated into siRNA duplexes to increase in vivo potency [54,100].

DNA oligomers containing GNA residues have been shown to form duplexes with DNA and RNA and to display self-pairing, whereby duplex formation was accompanied by hypochromicity [97,99]. In terms of stability, a single substitution from DNA to either (*S*)-GNA or (*R*)-GNA results in a decrease in T_m of 13 °C and 7 °C, respectively. As the number of substitutions is increased, the T_m decreases in a non-linear fashion. Replacement of all residues of a DNA strand by either (*S*)-GNA or (*R*)-GNA results in the complete loss of duplex formation, thereby confirming the detrimental effect of single and/or multiple GNA incorporations on duplex stability [101,102]. However, Zhang et al. demonstrated that an all-(*S*)-GNA can form a duplex with RNA [99]. It has been shown that a GNA/GNA duplex exceeds the thermal stability of DNA/DNA and RNA/RNA duplexes of the same sequence (increase in T_m of 18–25 °C) [99,101]. Moreover, (*S*)-GNA and (*R*)-GNA do not cross-pair either in a parallel or antiparallel fashion; thus GNA:GNA duplex formation is limited to homochiral pairing between either (*S*)-GNA or (*R*)-GNA strands [103]. With respect to nuclease stability, Nielson et al. showed that a 17mer oligonucleotide containing one T-GNA substitution has a nuclease half-life of 18–22 minutes in snake venom phosphodiesterase (SVPDE), thus exhibiting significantly higher stability compared to the parent strand [104]. Furthermore, Schlegel et al. showed that the position of the GNA substitution in a DNA/DNA duplex greatly influences its ability to resist 3'-exonucleases. Their work showed that a single or double (*S*)- or (*R*)-GNA substitution at the 3' end of a dT_{20} oligomer with a natural phosphodiester backbone greatly increases the oligonucleotide's ability to resist SVPDE. Furthermore, when moving the single or dinucleotide substitution to the penultimate position, a marked decrease in nuclease stability was observed. However, when these modifications where moved to the terminal positions, an 8- or 5-fold increase in nuclease resistance was observed for the (*S*)- or (*R*)-isomer, respectively [54].

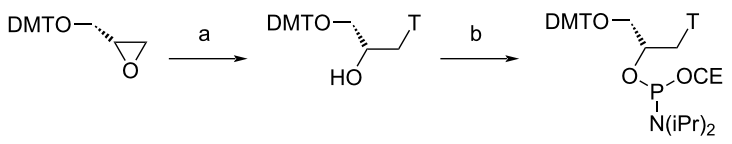
It is generally assumed that nucleic acid analogues require cyclic units in the backbone to generate the necessary conformational preorganization for duplex formation. This assumption does not hold true for GNA backbones where the destabilization caused by the shorter glycol moiety in DNA duplexes most likely stems from the structural incompatibility with the B-form deoxyribonucleotide-phosphate backbone. On the other hand, GNA–GNA duplexes form highly stable antiparallel duplexes that follow Watson–Crick base pairing rules [99]. GNA strands self-assemble into homochiral antiparallel right-handed ((*S*)-GNA) and left-handed ((*R*)-GNA) duplexes held together by Watson–Crick base pairs. Furthermore, these duplexes exhibit cross-strand base stacking consistent with A-form DNA and RNA duplexes [55].

Crystallographic studies have shown that (*S*)-GNA can form M-type helices (with metallo-base pairs) similar to A-form helices (with brominated base pairs). The M-type structure with 16 base pairs per turn and a helical pitch of 60 Å (ca. 3.8 Å helical rise) deviates significantly from the canonical A-form (11 base pairs/turn and ca. 2.6 Å rise) and B-form (10 base pairs/turn and ca. 3.4 Å rise) duplexes [54,55,105-107]. GNA duplexes possess only one large groove which corresponds to the canonical minor groove, the canonical major groove is a convex surface. Furthermore, the glycol backbone adopts two conformations alternating between *gauche* and *anti* conformations such that each base pair contains one nucleotide in the *gauche* conformation and one in the *anti* conformation. There is also a large backbone-base inclination (46° to −53°) which results in zipperlike interstrand and reduced intrastrand base stacking interactions [103]. The crystal structure of an RNA duplex containing (*R*)-GNA revealed that this modification disrupts both the phosphate backbone and hydrogen bonding of an adjacent base pair whereas (*S*)-GNA has a minimal influence on the structure of the duplex [54] (Figure 4). Moreover, incorporation of (*S*)-GNA residues in the seed region of the antisense strand of siRNA was observed to mitigate off target effects [54].

Sugar and nucleobase modifications

2'-O-Alkyl modifications

Historically, the 2'-*O*Me modification (Figure 5A) was the first of its class. The synthesis of each 2'-*O*Me ribonucleoside re-



Scheme 3: Synthesis of the (*S*)-GNA thymine phosphoramidite from (*S*)-glycidyl 4,4'-dimethoxytrityl ether. (a) Thymine, NaH, DMF; (b) $ClP(NiPr_2)(OCE)$, $(iPr_2)_2NEt$. T = thymine. Adapted from [99].

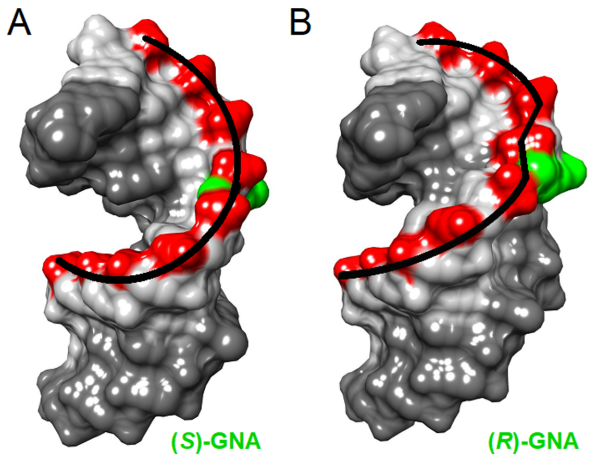


Figure 4: Surface models of the crystal structures of RNA dodecamers with single (A) (*S*)-GNA-T (PDB ID 5V1L) [54] and (B) (*R*)-GNA-T (PDB ID 5V1K) [54] nucleotides per strand. The presence of the (*R*)-GNA isomer introduces a kink in the backbone and causes local disruption of base stacking, in-line with a significantly reduced T_m relative to the (*S*)-GNA isomer.

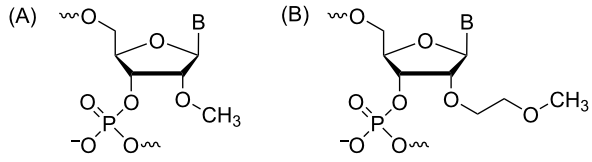
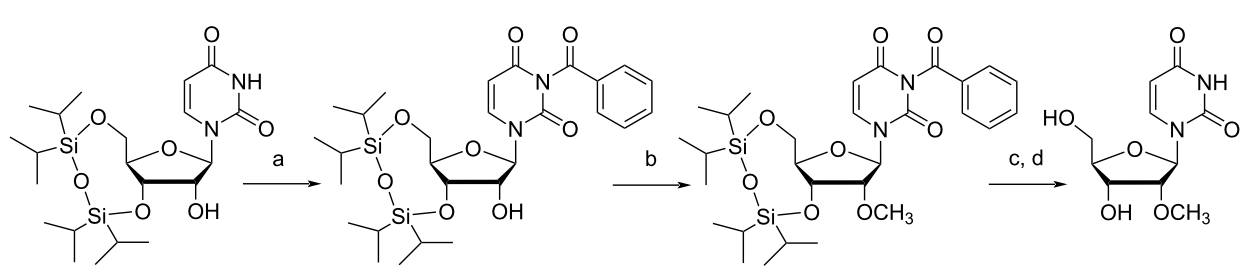


Figure 5: Structures of 2'-*O*-alkyl modifications. (A) 2'-*O*-methoxy RNA (2'-*O*Me RNA), (B) 2'-*O*-(2-methoxyethyl) RNA (2'-*O*-MOE RNA).

quired specific considerations [108]. Starting from 3',5'-*O*-(tetraisopropyldisiloxane-1,3-diyl) (TIPDS) protected uridine, protection of N3 was needed in order to prevent methylation at this position (Scheme 4). The N3-benzoylated derivative could then be treated with methyl iodide in the presence of silver oxide in order to methylate the 2'-OH. A similar strategy was employed to synthesize 3',5'-*O*-TIPDS-N^4-benzoyl-2'-*O*-methylcytidine. Next, 3',5'-*O*-TIPDS-N^6-benzoyladenosine suffered from methylation at the nucleobase and thus, 6-chloro-9-β-D-ribofuranosylpurine was instead used as the starting material. Once TIPDS protected, the 2'-OH could, once again, be selectively methylated with methyl iodide and silver oxide. The protected adenine base was regenerated by treatment with ammonia followed by benzoylation. Once the methyl group was incorporated into these ribonucleosides, the TIPDS group was selectively removed by tetrabutylammonium fluoride (TBAF) or hydrochloric acid treatment, followed by 5'-tritylation. In the case of guanosine, this strategy for 2'-OH methylation was unsuccessful, owing again to undesired methylation at the nucleobase. Instead, the 5'-*O*-monomethoxytrityl derivative of N^2-isobutyrylguanosine was treated with diazomethane in dimethylformamide in the presence of tin chloride, affording both 2'-*O*Me and 3'-*O*Me regioisomers. Fortunately, these isomers could be separated by silica gel column chromatography. Other synthetic approaches have since been developed [109-111], however, this pioneering work should be appreciated as nowadays, the 2'-*O*Me phosphoramidites of each protected ribonucleoside are all commercially available.

The study of 2'-*O*Me modified oligonucleotides was stimulated by the fact that they bind to RNA with higher affinity than unmodified RNA or DNA, as well as their improved nuclease resistance [112], promoting their usefulness in antisense therapies. Unfortunately, it was determined that uniformly 2'-*O*Me modified RNA:RNA duplexes were not substrates for RNAse H [113]. Structural insights of this modification were determined from the crystal structure of a duplex of self-complementary 10-mer DNA strands with a single internal 2'-*O*Me modified adenosine [114]. This duplex adopted an overall A-form, with the sugars in the C3'-*endo* orientation and the two, well solvated methoxy groups, pointing into the relatively wide minor groove of the duplex.

It was shown that as the number of carbons in the 2'-*O*-alkyl chain increased, so too did the destabilizing effect towards RNA binding affinity [115]. Thus, it was initially believed that even though nuclease resistance increased with chain length, this destabilizing effect would render 2'-*O*-alkyl-modified RNA a less potent therapeutic agent. In 1994, there was crystallo-



Scheme 4: Synthesis of the 2'-*O*Me uridine from 3',5'-*O*-(tetraisopropyldisiloxane-1,3-diyl)uridine. (a) Benzoyl chloride, triethylamine; (b) CH_3I, Ag_2O; (c) dilute NH_4OH; (d) 0.5 N HCl. Adapted from [108].

graphic evidence, however, that suggested the addition of a polarizable group in the longer 2'-*O*-alkyl chains that could hydrogen bond with nucleobases in the minor groove of the duplex would be well tolerated [114]. This supported the hypothesis that the 2'-*O*-[2-(methoxy)ethyl] (MOE) modification (Figure 5B) wouldn't lead to significant destabilization of the duplex, prompting its development.

The synthesis of 2'-*O*-MOE-modified ribonucleosides was first described in 1995 [116]. Since then, two practical strategies have been developed for synthesizing 2'-*O*-MOE ribonucleosides. For pyrimidines, this involves treating 2,2'-anhydrouridine with aluminum 2-methoxyethoxide, which attacks and inserts at the 2'-position, opening the ring and producing the nucleoside with the correct stereochemistry (Scheme 5) [117]. Conveniently, this 2'-*O*-MOE uridine can be converted to the cytidine derivative by 4-nitrophenylation, 3',5'-trimethylsilylation and finally, treatment with aqueous ammonia. In contrast, the purine synthetic route first uses the bis-silylating agent [methylene bis(diisopropylsilyl)chloride] (MDPS) to protect both the 5' and 3'-hydroxy groups [118]. The protected nucleoside can then be treated with 2-methoxyethyl bromide in the presence of NaHMDS in order to selectively alkylate the 2'-OH, followed by TBAF treatment to remove the MDPS protecting group.

The 2'-*O*-MOE soon became the gold standard alkyl modification, owing to its improvement in therapeutically relevant properties. Compared to 2'-*O*Me RNA, the 2'-*O*-MOE RNA analogue has similar or even increased RNA binding affinity, as well as a tenfold increase in nuclease resistance [119]. Moreover, compared the PS-DNA, 2'-*O*-MOE RNA has an increased thermal stability of 2 °C per modification, with similar nuclease resistance [11,41]. Rationale for the improved properties of the 2'-*O*-MOE modification was gained through the analysis of the crystal structure of a uniformly modified self-complementary 12-mer RNA duplex [58]. The duplex was observed to be in the A-form, with the sugar residues being in a C3'-*endo* conformation. The MOE substituents were in the gauche orientation, being well accommodated in the minor groove and making a stabilizing interaction with a trapped water molecule and the adjacent phosphate (Figure 6). It's this pre-organization of the MOE groups, making the duplex more rigid, which is hypothesized to cause the increase in RNA binding affinity. Furthermore, the increase in nuclease resistance is believed to be due to steric constraints from the MOE substituent and water molecule protecting the adjacent phosphate.

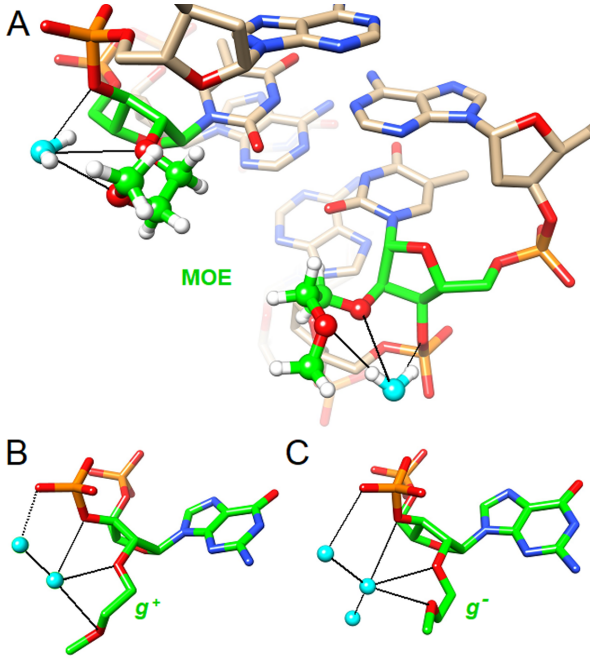
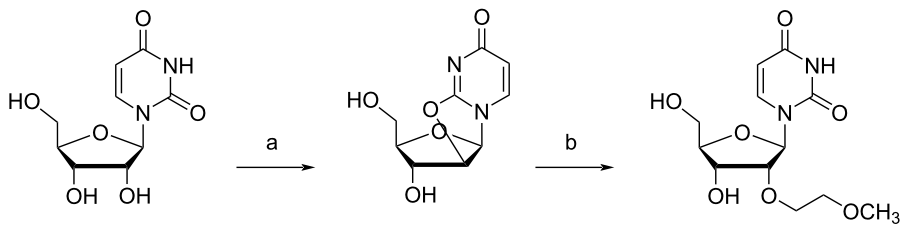


Figure 6: Structure of 2'-*O*-(2-methoxyethyl)-RNA (MOE-RNA). (A) View into the minor groove of an A-form DNA decamer with single MOE-T nucleotides per strand (PDB ID 411D, highlighted with green carbon atoms) [120]. Water molecules are trapped in a chelate-like manner between the O3', O2' and OC' (outer oxygen of the MOE substituent). (B) and (C) individual nucleotides from a crystal structure of an MOE-RNA dodecamer duplex (PDB ID 469D) [58]. Of the 24 MOE substituents, 22 adopt a gauche conformation, either *g*+ or *g*−, whereby both trap a water molecule that can be linked to the 3'-phosphate via a water bridge.

Many other 2'-*O*-alkyl modifications have been synthesized and studied extensively, and are summarized elsewhere [41,121]. Importantly, while 2'-*O*-alkyl-modified RNA cannot activate



Scheme 5: Synthesis of the 2'-*O*-MOE uridine from uridine. (a) $(PhO)_2CO$, $NaHCO_3$, DMA, 100 °C; (b) $Al(OCH_2CH_2OCH_3)_3$, reflux. Adapted from [117].

the RNAse H dependent degradation pathway, they can, however, act through a different therapeutic mechanism as steric blockers, inhibiting mRNA translation, RNA reverse transcription or RNA splicing [122-125].

Locked nucleic acids (LNA)/bridged nucleic acids (BNA)

Locked nucleic acids are a class of modified nucleosides which traditionally involve the incorporation of a methylene bridge between C4' and O2' of the ribose sugar (Figure 7A). This incorporation, as first reported by both Wengel and Obika, locks the nucleoside in the C3'-*endo* (north) conformation which allows for enhanced binding affinities towards both DNA and RNA targets [126,127]. Both ^{1}H NMR [127-129] and crystallographic studies [126] have been used to demonstrate the Northern puckering of the sugar and the *anti* orientation of the nucleobase. The key synthetic step in the synthesis of LNA involves the tosylation of a 4'-*C*-hydroxymethyl derivative, followed by a base-induced ring closure to afford the 2'-*O*,4'-*C*-linked bicyclic nucleoside derivative (Scheme 6) [127,128]. Incorporation of LNA into a variety of oligonucleotides with varying lengths and sequences has shown increased thermal stability when binding to either DNA or RNA complements with T_m increases of +1 to +8 and +2 to +10 °C, respectively [127,128,130-134]. The higher stabilization of RNA can be attributed to the preorganization of LNA nucleosides towards formation of A-form duplexes [128], whereas in DNA duplexes LNA residues steer the conformation of the neighboring DNA monomers into the C3'-*endo* conformation [135,136]. These modifications have also been shown to confer a higher level of nuclease resistance than isosequential DNA or phosphorothioate modifications [137-141]. In combination with the high selectivity for RNA sequences, this makes LNA-modified oligonucleotides well suited for use as antisense therapeutics. Recent publications have used LNA's high affinity for RNA sequences in gapmer-designed antisense oligonucleotides for successful targeting of a key gene involved in TGFβ inhibition [142]. The inclusion of LNA nucleosides within a larger single-stranded DNA oligonucleotide has also allowed for subtle gene modifications to be implemented while evading mismatch repair (MMR) [143]. Furthermore, Ju et al. recently reported the use of LNA-based suppressors for the inhibition of viral miRNA through carbon dot-mediated delivery [144]. A diastereomer of LNA, α-L-LNA (Figure 7B), also induces a higher affinity for both DNA and RNA complements in addition to providing a high stability against nucleases [145,146]. Unlike LNA, this diastereomer is a mimic of DNA instead of RNA and promotes a C2'-*endo* puckering of the sugar [147]. As a result, it has been shown to be better (fivefold) than other modified LNA analogues at knocking down target genes in vitro [145]. Also, these isomers have recently been shown to be useful in stabilizing streptavidin-binding aptamers [148], and in the use of antisense oligonucleotides for splice modulation through the induction of *Dmd* exon-23 skipping in mice in vitro [149].

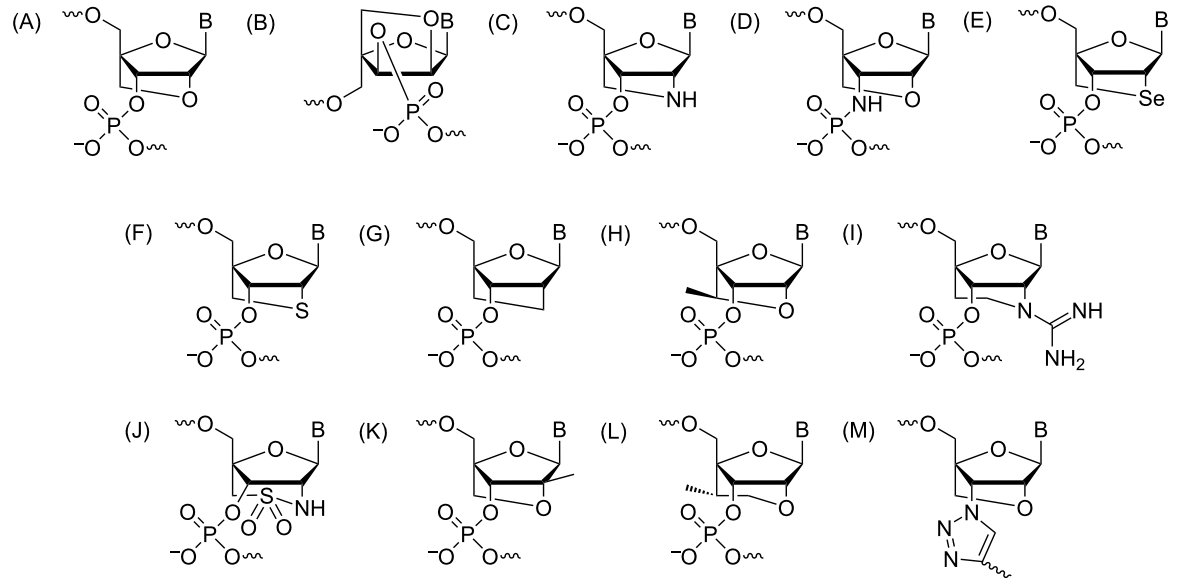
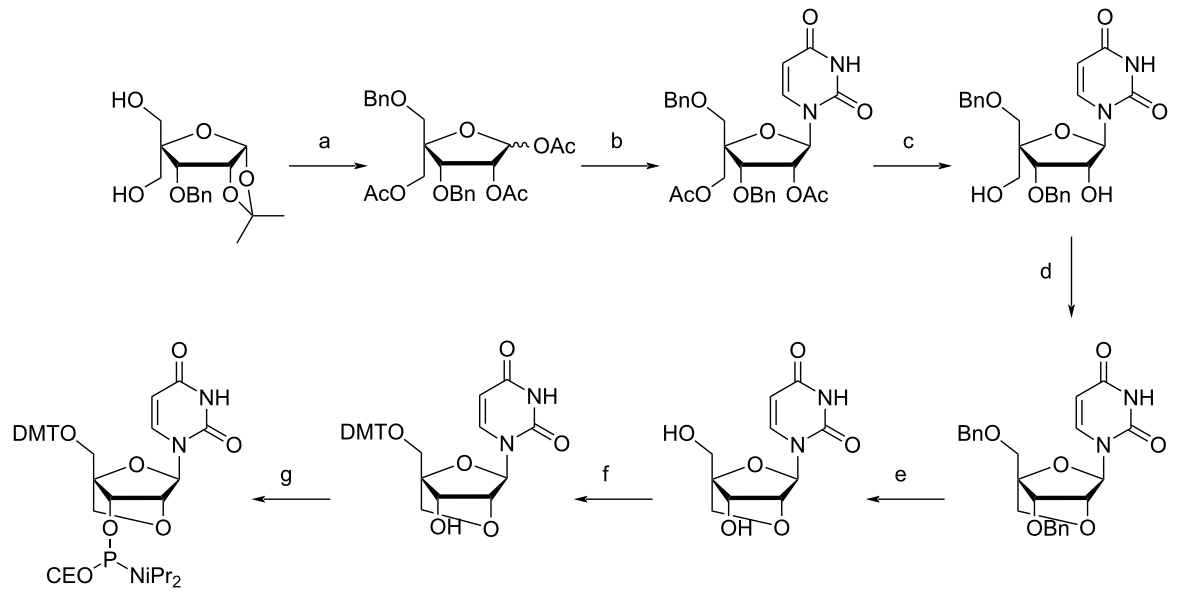


Figure 7: Structures of locked nucleic acids (LNA)/bridged nucleic acids (BNA) modifications. (A) LNA/BNA, (B) α-L-LNA, (C) C2'-amino-LNA, (D) 3'-amino-2',4'-LNA, (E) seleno-LNA, (F) thio-LNA, (G) carba-LNA, (H) *S*-constrained ethyl (cEt) nucleic acid, (I) 2'-*N*-guanidino,4'-*C*-ethylene nucleic acid (GENA), (J) sulfonamide-bridged nucleic acid (suNA), (K) 2'-Me-LNA, (L) 6'-*S*-Me-2'-*O*,4'-*C*-ethylene-bridged nucleic acid (6'-*S*-Me-ENA), (M) triazole linked LNA.



Scheme 6: Synthesis of the uridine LNA phosphoramidite. (a) i) NaH, BnBr, DMF, ii) acetic anhydride, pyridine, iii) 80% AcOH, iv) acetic anhydride, pyridine; (b) uracil, *N*,*O*-bis(trimethylsilyl)acetamide, TMS-triflate, acetonitrile; (c) $NaOCH_3$, methanol; (d) i) *p*-toluenesulphonyl chloride, pyridine, ii) NaH, DMF; (e) H_2, $Pd(OH)_2/C$, ethanol; (f) DMTCl, pyridine; (g) $ClP(NiPr_2)(OCE)$, $(iPr_2)_2NEt$, dichloromethane. Bn = benzyl, Ac = acetyl. Adapted from [128].

Recently, a lot of attention has been paid to modifying the LNA scaffold to incorporate various heteroatoms, modify the bicyclic framework, and to change the location of the methylene bridge to tailor the properties of these nucleosides. The incorporation of nitrogen at C2' has been explored for further functionalization while retaining the LNA scaffold. Singh et al. were the first to report the synthesis of C2'-amino-LNAs (Figure 7C) in 1998 [150], with the synthetic route being optimized over time [151,152]. The stability of these derivatives is similar to those of LNA [150-152], with the added advantage of additional coupling reactions to fluorescent groups [151], or small molecules being possible either during solid-phase synthesis (SPS) [153,154] or post synthetically [155,156]. Gapmer oligonucleotides that incorporate 2'-amino-LNA show increased uptake in organs such as the heart, liver, and lungs in comparison to other LNA modifications [145]. Nitrogen can also be incorporated at the C3' position in the form of a 3'-amino-2',4'-LNA (Figure 7D) monomer which has been shown to stabilize oligonucleotides similarly to unmodified LNA with a nuclease resistance greater than PS-modified oligonucleotides [157]. Incorporation of selenium at C2' in a thymine-bearing LNA nucleoside (Figure 7E) has been demonstrated to have a hybridization ability and a nuclease resistance that are highly reversible in response to redox changes of the selenium atom [158]. Recent work has also looked at this modification in LNA nucleosides bearing an adenine base [159], but this nucleoside was found to be highly sensitive to heat, making its incorporation into oligonucleotides challenging. Thio-LNA (Figure 7F), which has sulfur incorporated at the C2' position, has similar binding properties as amino-LNA and β-D-LNA, but with varying biodistribution patterns and a higher cellular uptake in mice [145]. Work looking at carba-LNA, which lacks the O2' functionality, has shown the importance of the oxygen atoms in hybridizing to complementary RNA [160]. Unsubstituted carba-LNA (Figure 7G), which lacks a hydrophilic substituent at C2', leads to a decrease in heteroduplex stability [160]. This agrees with the observation in the crystal structure of an LNA-modified DNA duplex where the 2'-oxygen acts as an H-bond acceptor for water, potentially making a favorable contribution to the increased pairing affinity of LNA [161].

Constrained ethyl (cEt) nucleic acids (Figure 7H), which contain a [2.2.1] tricyclic core, have been developed and show improved potency when compared to second generation 2'-*O*-MOE antisense oligonucleotides [162,163]. The cEt also demonstrate an improved toxicity profile in comparison to standard LNA ASOs [162]. The arduous synthesis of the nucleoside analogues has been refined to minimize the number of needed stereochemical adjustments and overall steps [164]. ASOs containing these modified nucleosides have demonstrated promising antitumor activity for lymphoma and lung cancer [165].

Numerous other LNA analogues have been constructed including, but not limited to, 2'-*N*-guanidino,4'-*C*-ethylene (GENA) (Figure 7I) [166], sulfonamide-bridged (suNA) (Figure 7J)

[167], 2'-Me LNAs (Figure 7K) [168,169], 6'-Me-2'-*O*,4'-*C*-ethylene-bridged (6'-Me-ENA) (Figure 7L) [170], and various triazole-linked LNA (Figure 7M) [171,172] that have all shown the ability to modulate LNA properties.

Arabinose and fluoroarabinose nucleic acids

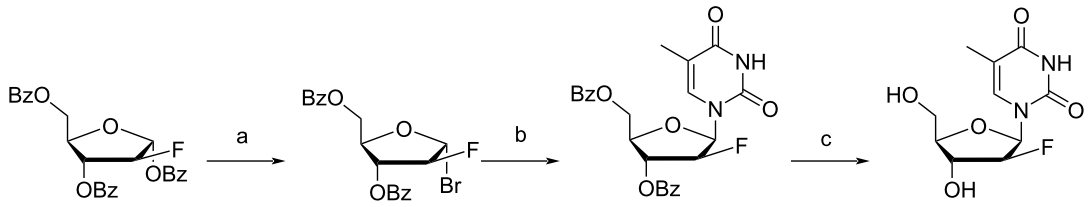
Arabino nucleic acids (ANA) are analogs of RNA where the hydroxy group at C2' is inverted (Figure 1G). In fluoroarabino nucleic acids (FANA) this C2' hydroxy group is replaced by fluorine. Arabino- and fluoroarabino nucleosides have demonstrated anticancer and antiviral activities (as reviewed in [173]). β-D-Arabinonucleosides of pyrimidines can be prepared from 2,2'-anydronucleosides [174] and purines from approaches which include condensation of the nucleobase with 2,3,5-tri-*O*-benzyl-D-arabinofuranosyl chloride [175]. The 2'-fluoro-β-D-arabinofuranose nucleosides can be prepared by coupling of the nucleobase with 3,5-di-*O*-benzoyl-2-deoxy-2-fluoro-α-D-arabinofuranosyl bromide (Scheme 7) [176-180]. Both β-D-arabino and 2'-fluoro-β-D-arabinofuranose nucleosides can be converted to phosphoramidite derivatives for incorporation into oligonucleotides for solid-phase synthesis [178,181-184].

Hybridization studies of uniformly modified ANA of mixed nucleobase composition to complementary RNA revealed reduced thermal stability relative to the corresponding DNA/RNA duplex by approximately 1.5 °C per modification [182,183]. A significant reduction in stability of the duplex was observed in the binding of ANA to complementary DNA relative to the DNA duplex [182,183]. In contrast, FANA of mixed nucleobase composition displayed improved binding with both complementary DNA and RNA, relative to DNA/DNA and DNA/RNA duplexes by approximately 1 °C and 0.5 °C per modification, respectively [178]. The 2'-stereoisomer of FANA, FRNA also demonstrates improved binding to RNA, relative to DNA [185]. Circular dichroism spectra of FANA/RNA and ANA/RNA duplexes show similarity to that of DNA/RNA [178,183]. Both ANA and FANA demonstrate good stability to nucleases [183,186]. Hybrid duplexes of ANA and FANA with complementary RNA were substrates of RNase H, with greater cleavage of the RNA strand observed for the latter, demonstrating the gene silencing potential of these analogs [183,186]. Uniformly modified phosphorothioate (PS) FANA forms a duplex with RNA with a higher T_m relative to the PS-DNA/RNA duplex, however, RNase H-mediated cleavage of RNA was diminished for the duplex formed with PS-FANA relative to PS-DNA [187]. Improved cleavage by RNase H was observed with chimeric PS-FANA/DNA [187]. PS-FANA/DNA chimera with either flanked or alternating segments of FANA residues, as demonstrated by knockdown of *c-MYB* mRNA with a persistent silencing effect [188].

A 1.55 Å crystal structure of a Dickerson–Drew dodecamer containing fluoroarabinothymine revealed that these modified nucleotides adopt an O4'-*endo* (east) conformation that is readily accommodated in a B-form duplex [189] (Figure 8A). Fluoroarabinothymine in an A-form DNA duplex had a northern conformation (Figure 8B,C) whereas arabinouridine in either an A- or B-form environment had a south-eastern conformation (Figure 8D,E), suggesting greater flexibility for FANA versus ANA [190]. NMR structures of hairpin duplexes consisting of RNA and either FANA or ANA stems suggested that both modifications adopt an O4'-*endo* sugar pucker [191,192]. The O4'-*endo* sugar conformation has been reported for the DNA strand in DNA/RNA hybrid duplexes, the natural substrate of RNase H [193,194]. Structures of duplexes containing FANA and FRNA (Figure 8F) have revealed that thermal stabilization may be attributed to nonconventional hydrogen bonds in the backbone [195-197]. Gene silencing by RNAi has also been explored with siRNA containing FANA residues [198]. These studies have shown that FANA is accommodated in the sense strand and 5'-end and 3'-termini of the antisense strand of the siRNA [198].

C4'-Modified nucleic acids

Modifications at the C4' sugar position (Figure 1H) have long been desirable as a means of modulating the properties of nucleic acids without interfering with Watson–Crick pairing. Incorporations at C4' are close in proximity to both the 3' and 5'-neighboring phosphate groups, allowing for a tailoring of the nuclease resistance [200]. In 2011, Rosenberg demonstrated the



Scheme 7: Synthesis of the 2'-fluoroarabinothymidine. (a) 30% HBr in acetic acid; (b) 2,4-bis-*O*-(trimethylsilyl)thymine, carbon tetrachloride; (c) NH_4OH, methanol. Bz = benzoyl. Adapted from [177].

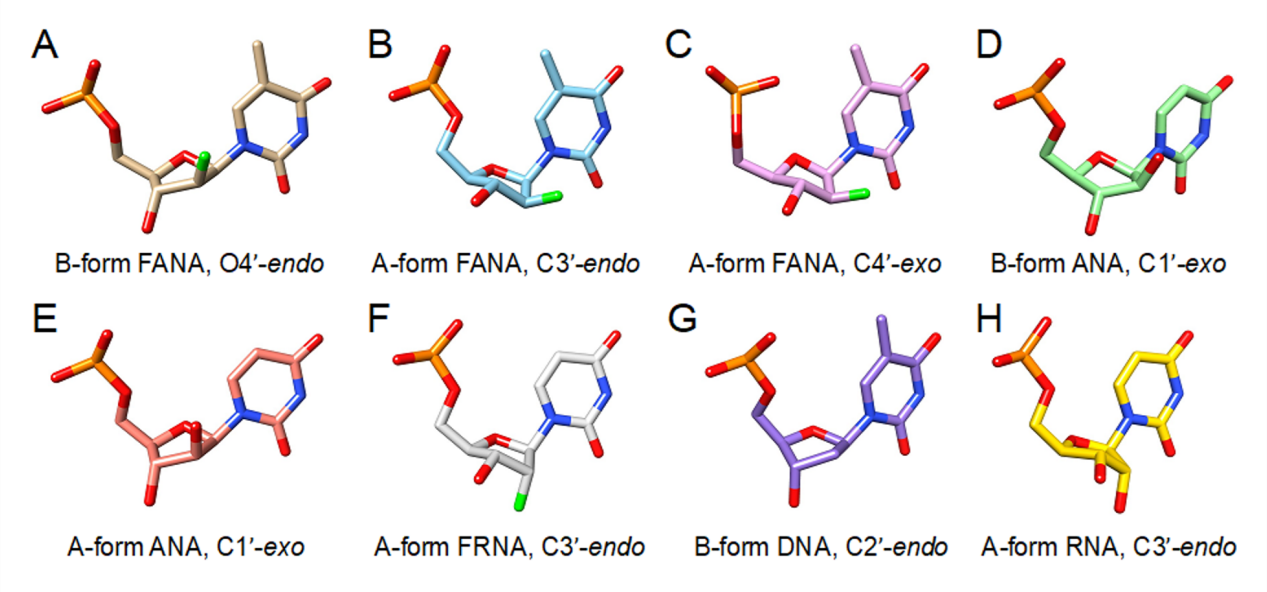


Figure 8: Sugar puckers of arabinose (ANA) and arabinofluoro (FANA) nucleic acids compared with the puckers of the fluoro-ribonucleic acid analog (FRNA) as well as DNA and RNA. (A) FANA-T in B-form DNA (PDB ID 388D) [189]. (B) FANA-T in A-form DNA (PDB ID 2FIL, duplex 1) [190]. (C) FANA-T in A-form DNA (PDB 2FIL, duplex 2) [190]. (D) ANA-U in B-form DNA (PDB ID 2FII) [190]. (E) ANA-U in A-form DNA (PDB ID 2FIJ) [190]. (F) FRNA-U in A-form RNA (PDB ID 3P4A) [62]. (G) B-form DNA (PDB ID 388D) [189]. (H) A-form RNA (PDB ID 5DEK) [199].

favorable binding properties of an oligothymidylate modified with 4'-methoxy or 4'-(2-methoxyethoxy) functionalities (Figure 9A,B) [201]. These modified nucleic acids were found to have superior hybridization behaviors towards both complementary DNA (see Figure 8G for pucker) and RNA (see Figure 8H for pucker) with sugar puckers in the northern (C3'-*endo*) and southern (C2'-*endo*) configurations for the respective alpha and beta isomers [201]. In 2015, this work was extended to incorporate these modifications into oligonucleotides containing all four bases [202]. *N*-Iodosuccinimide promoted the alkoxylation of the 4'–5'-enol acetates yielded the corresponding 5'-acetoxy-5'-iodo-4'-methoxy intermediates [202]. These intermediates were hydrolyzed with a mixture of triethylammonium bicarbonate (TEAB) and *N*,*N*-dimethylformamide (DMF) followed by a sodium borohydride reduction to give the 4'-alkoxy products [202]. The 4'-methoxy-2'-deoxynu-

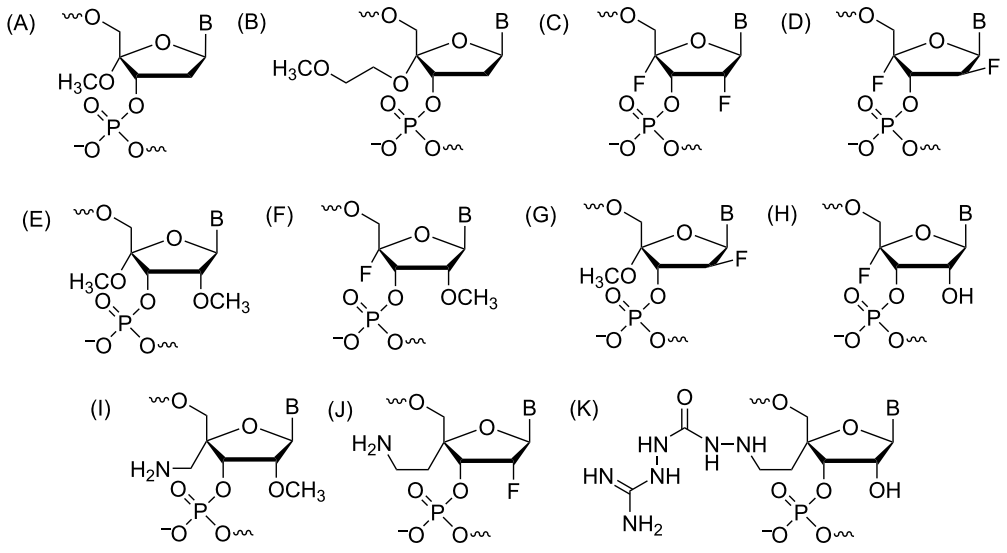


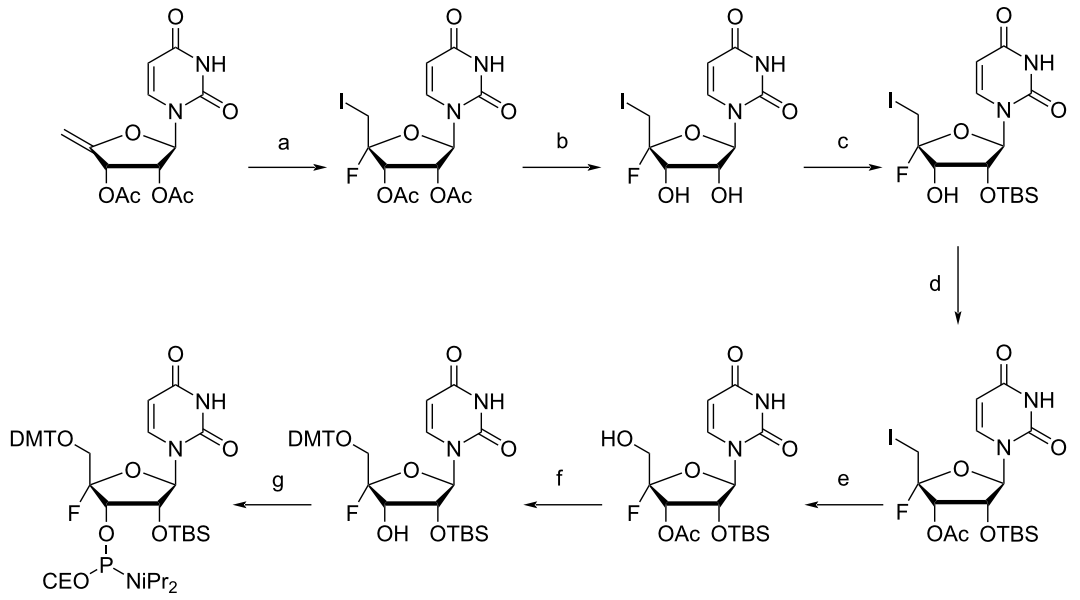
Figure 9: Structures of C4'-modified nucleic acids. (A) 4'-methoxy, (B) 4'-(2-methoxyethoxy), (C) 2',4'-difluoro (2',4'-diF) RNA, (D) 2',4'-difluoro (2',4'-diF) ANA, (E) 2',4'-dimethoxy RNA, (F) 2'-methoxy,4'-fluoro RNA, (G) 2'-fluoro,4'-methoxy ANA, (H) 4'-fluoro RNA, (I) 4'-*C*-aminoalkyl-2'-*O*-methyl, (J) 4'-*C*-aminoalkyl-2'-fluoro, (K) 4'-*C*-guanidinocarbohydrazidomethyl.

cleosides exhibited high resistance towards depurination under acidic conditions [202]. In contrast, nucleosides that are modified with 4'-fluoro modifications have more labile glycosidic linkages under similar conditions [203,204]. Rosenberg attributed this contrast to the electronegativity differences between the groups and the effect this would have on the stabilization of the resulting oxocarbenium ion [202]. Oligomers modified with the 4'-methoxy modification hybridized better to complementary RNA, rather than DNA, due to the N-type conformation of the sugar pucker, as confirmed by NMR [202]. These same oligomers exhibited half-lives of approximately 40 minutes in the presence of phosphodiesterase I [202]. In contrast, the natural DNA sequence had a half-life of 1 min [202].

The incorporation of fluorine at the C4' position has long constituted a challenge owing to the instability of the glycosidic bond in the resulting nucleosides. This modification is desirable due to its involvement in the mode of action of the natural antibiotic nucleocidin [203,205]. Damha reasoned that the incorporation of fluorine at both C2' and C4' could lead to a stable nucleoside due to the glycosidic bond stabilization brought about by 2'-fluorination [206] which turned out to be correct after successful isolation of both 2',4'-diF-rU and 2',4'-diF-rC nucleosides (Figure 9C) [206]. Through NMR, these nucleosides were found to be essentially locked in the northern (C3'-*endo*) sugar pucker, albeit without the need for the bicyclic structures typical for locked nucleic acids [206]. The 2',4'-diF-rU nucleoside was introduced into an RNA by way of an HCV polymerase and extended to give a full-length oligonucleotide product, whereas 2',4'-diF-rUTP inhibited RNA synthesis at the early stages of dinucleotide-primed reactions [206]. Standard solid-phase synthesis allowed for the incorporation of this modified nucleoside into both RNA and DNA oligonucleotides. The impact on stability was found to be minimal in the case of RNA/RNA duplexes; mildly destabilizing with RNA/DNA hybrid duplexes; and highly destabilizing when incorporated into the DNA strand of DNA/RNA or DNA/DNA duplexes [207]. Damha attributed this destabilization to structural distortions caused by A/B junctions within the helical structures [207].

2',4'-diF-modified siRNA sequences were capable of triggering RNAi with high efficiency, and the incorporation of multiple residues in the guide (antisense) strand yielded more potent siRNAs than those containing LNA or FANA modifications [207]. 2',4'-diF-ANA (Figure 9D) also adopted the northern (C3'-*endo*) sugar pucker despite the 2'-βF, which generally leads to the adoption of a southern or eastern pucker [208]. This monomer was found to have minimal effects on the thermal stability of nucleic acid duplexes. However, when incorporated into a DNA/RNA hybrid duplex it was shown to decrease the rate of both human and HIV reverse transcriptase-associated RNase H-mediated cleavage [208]. In 2018, the work was expanded to include 2',4'-diOMe-rU, 2'-OMe,4'-F-rU, and 2'-F,4'-OMe-araU nucleosides (Figure 9E,F,G) [209]. This work reinforced the notion that both 4'-OMe and 4'-F modifications steer the sugar pucker towards a C3'-*endo* (north) conformation [209], even in the presence of C2' groups that would favor a different puckering of the ribose sugar. The 4'-modifications provided either a small stabilizing or destabilizing effect depending on the type of underlying duplex, and these 4'-substituents were able to modulate the binding affinities for the parent 2'-modifed oligonucleotides [209]. siRNA containing inserts of the C4' α-epimer of 2'-F,4'-OMe-rU, in either the sense or antisense strands, triggered gene silencing with efficiencies comparable to that of 2'-F-rU [210].

Recently, Zhou provided the first synthesis of a 4'-F-rU (Figure 9H) phosphoramidite which was stable enough to then be incorporated into longer oligonucleotides through standard solid-phase synthesis (Scheme 8) [211]. They found that the modified 4'-F-rU ribonucleotide had a high resemblance to the unmodified uridine, allowing it to be used as a probe for RNA structure determination through ^{19}F NMR [211]. This modification led to RNA which was stable and predominantly in the C3'-*endo* (north) conformation [211], similar to the 2',4'-diF-RNA previously reported by Damha [208]. Zhou reasoned that because 3'-*O*-β-glucosylated nucleocidin, an intermediate in the biosynthetic pathway of nucleocidin, was stable, they may be able to successfully achieve the synthesis of the 4'-F-rU phosphoramidite through a selective protection of the hydroxy groups in stages [211]. Starting with a prepared 5'-iodo-4'-fluorouridine analogue that had been used in previous attempts of this synthesis, they removed the acetyl protecting groups at C3' and C2' with NH_3/MeOH to give 5'-iodo-4'-fluorouridine [211]. Selective protection of the 2'-OH with TBDMS-Cl followed by protection of the 3'-OH with an acetyl group gave the fully protected intermediate [211]. Treatment of this intermediate with m-CPBA in the presence of a phase-transfer catalyst in acidic medium gave the resulting 5'-OH compound [211]. The authors reported no transfer of the 2'-TBDMS group onto the 5'-OH, however, following removal of the 3'-*O*-acetyl group with NH_3/MeOH, some TBDMS transfer to the C3' position is seen [211]. 5'-DMT protection then led to the pre-amidite [211]. ^{19}F NMR results show that not only does this modification allow for discernment between ssRNA and dsRNA, but it also allows for the identification of mismatches and the binding of RNA-processing proteins with chemical shift dispersions as large as 4 ppm, suggesting that this modification has a wide use for the determination of a variety of RNA structures through NMR spectroscopy [211].



Scheme 8: Synthesis of the 4'-F-rU phosphoramidite. (a) AgF, I_2, dichloromethane, tetrahydrofuran; (b) NH_3, methanol; (c) TBS-Cl, $AgNO_3$, pyridine, tetrahydrofuran; (d) acetic anhydride, dimethylaminopyridine, pyridine; (e) tetrabutylammonium hydroxide, trifluoroacetic acid, *m*-chloroperoxybenzoic acid; (f) (i) NH_3, methanol (ii) DMTCl, pyridine; (g) ClP(NiPr$_2$)(OCE), 1-methylimidazole, (iPr$_2$)$_2$NEt, dichloromethane. TBS = *tert*-butyldimethylsilyl. Adapted from [211].

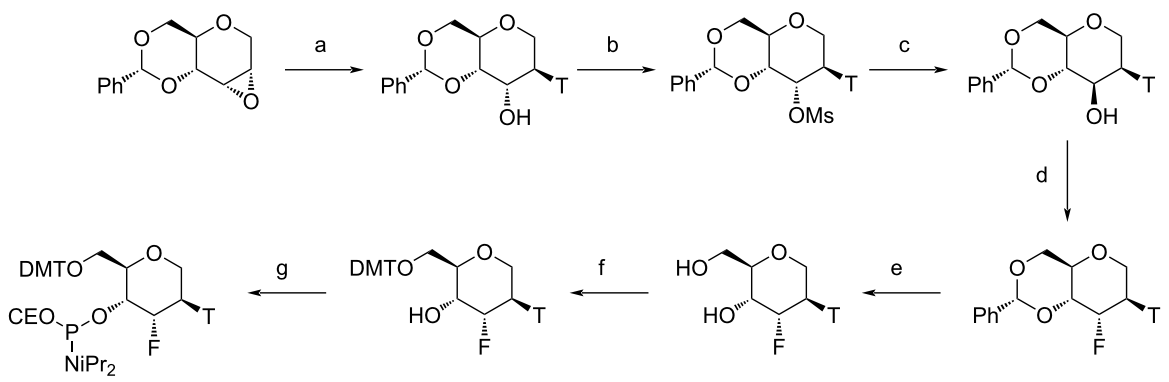
In contrast, the incorporation of 4'-*C*-aminoalkyl-2'-*O*-methyl (Figure 9I) nucleosides leads to a slight destabilization of helical structures due to the adoption of a C2'-*endo* (south) conformation [212,213]. When fluorine is incorporated at C2' instead of 2'-*O*Me (Figure 9J), these 4'-*C*-aminoalkyl nucleosides are found to stabilize both dsRNA and siRNA to a larger extent [214]. The incorporation of 8 nucleosides into an siRNA passenger strand showed RNAi activity identical to the unmodified siRNA, with 50% of the siRNA strands remaining intact after 48 h in 20% BSA [214]. Recent work on the synthesis of novel 4'-*C*-guanidinocarbohydrazidomethyl-5-methyluridine (GMU) (Figure 9K) has shown that functionalizing the C4' position with guanidinium leads to siRNAs with increased thermal stability (1–3 °C/mod) and improved stability in human serum [215]. These guanidinium-modified siRNAs also lead to sustained gene silencing with only picomolar concentrations after 96 h of transfection [215]. Their qPCR experiments show that the cause of this sustained gene silencing activity is due to enhanced guide strand recruitment within the RISC complex [215].

3'-Fluorohexitol nucleic acids (FHNA)

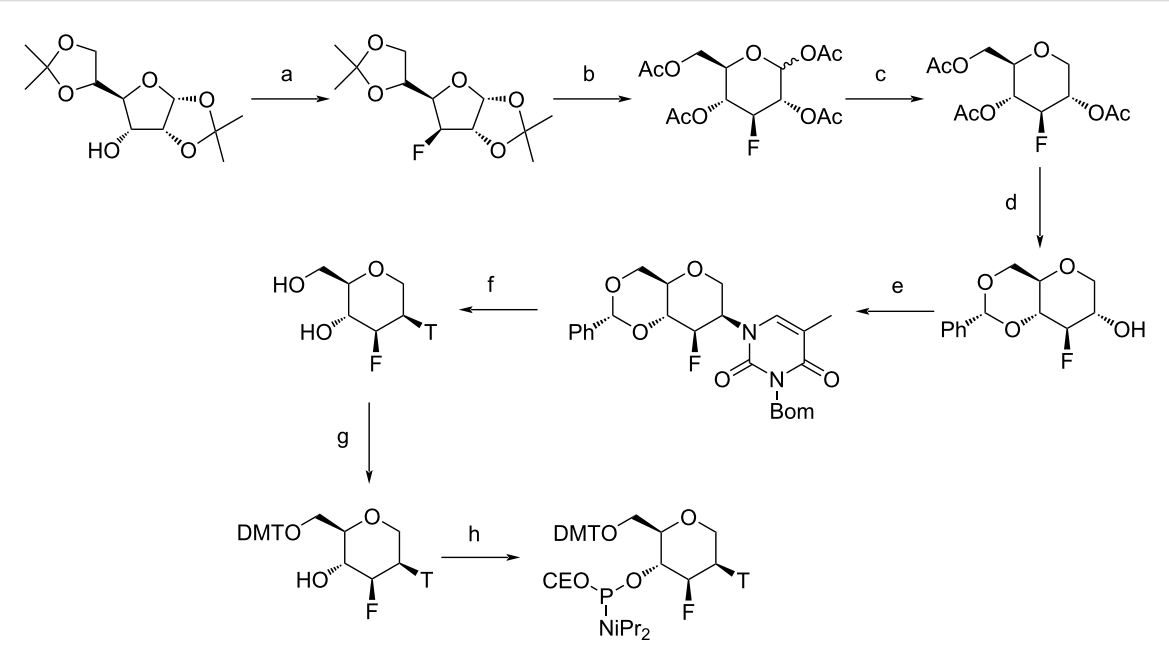
Herdewijn was the first to describe the synthesis as well as the biophysical, structural, and biological characterization of hexitol nucleic acids (HNA), mannitol nucleic acids (MNA), and altritol nucleic acids (AtNA) [216-220]. These carbohydrate-modified nucleosides incorporate a six-membered pyranose ring in place of the furanose ring found in unmodified DNA and RNA, with the nucleobase positioned at the C2' position in an axial orientation mimicking the C3'-*endo* (north) sugar puckering of furanose nucleosides [221]. MNA and AtNA possess an additional hydroxy group at the C3' position in the *R* and *S* configurations, respectively [219,220]. HNA was found to bind to complementary RNA in an antiparallel, sequence-dependent fashion, leading to the stabilization of HNA/RNA duplexes [218]. HNA also stabilizes HNA/DNA duplexes but to a smaller degree due to differences in minor groove solvation [222]. mRNA translation experiments have shown that HNA can function as a steric blocking agent of Ha-ras in cell-free experiments [223]. AtNA/RNA displays higher thermal stability when compared to HNA/RNA and natural nucleic acid controls [220]. In contrast, the introduction of MNA leads to duplex destabilization due to unfavorable steric clashes and limited nucleoside preorganization [219].

In 2011, a work detailing the first synthesis of both isomers of 3'-fluoro-modified hexitol nucleic acid (FHNA and Ara-FHNA) (Figure 1I) was published (Scheme 9 and Scheme 10) [221]. The incorporation of fluorine has long been used in siRNA [224], miRNA [225], and for ^{19}F NMR structural studies of nucleic acids [211]. It was proposed that the incorporation of fluorine at the C3' position of HNA could further expand its use as a potential antisense therapeutic [221].

The published data show that incorporation of a 3'-fluorine atom in the trans-diaxial orientation relative to the base in



Scheme 9: Synthesis of the thymine FHNA phosphoramidite. (a) thymine, 1,8-diazabicyclo[5.4.0]undec-7-ene, acetonitrile; (b) methanesulfonyl chloride, pyridine; (c) aq NaOH, 1,4-dioxane; (d) nonafluorobutanesulfonyl fluoride, 1,8-diazabicyclo[5.4.0]undec-7-ene, tetrahydrofuran; (e) H_2, $Pd(OH)_2/C$, methanol; (f) DMTCl, pyridine; (g) $P(NiPr_2)_2(OCE)$, 1*H*-tetrazole, NMI, DMF. Ms = methanesulfonyl, Ph = phenyl, T = thymine. Adapted from [221].



Scheme 10: Synthesis of the thymine Ara-FHNA phosphoramidite. (a) i) trifluoromethanesulfonic anhydride, pyridine, ii) CsF, *tert*-butanol; (b) i) Amberlite IR-120-H, 1,4-dioxane, water, ii) acetic anhydride, pyridine; (c) i) 33% HBr in acetic acid, dichloromethane, ii) tributyltin hydride, 2,2′-azobis(2-methylpropionitrile), toluene; (d) i) K_2CO_3, methanol, ii) benzaldehyde dimethyl acetal, *p*-toluenesulfonic acid, DMF; (e) i) trifluoromethanesulfonic anhydride, dichloromethane, pyridine, ii) N3-benzyloxymethylthymine, 1,8-diazabicyclo[5.4.0]undec-7-ene, dimethyl sulfoxide; (f) H_2, $Pd(OH)_2/C$, methanol; (g) DMTCl, pyridine; (h) $P(NiPr_2)_2(OCE)$, 1*H*-tetrazole, NMI, DMF. Bom = benzyloxymethyl. Adapted from [221].

FHNA (Figure 10A) leads to stabilization of the resulting nucleic acid duplex, whereas the incorporation of ara-FHNA leads to sequence-dependent destabilization of the duplex [221]. The FHNA modification is better at discerning G–T mismatches than DNA or LNA, and both FHNA and Ara-FHNA were more stable against exonuclease digestion in comparison to LNA and MOE-modified oligonucleotides [221]. X-ray crystallographic studies showed that the equatorial 3'-fluorine of Ara-FHNA-T in the A-form DNA decamer pushes away O4' from the 3'-adjacent 2'-deoxy-A within the minor groove of the duplex [221] (Figure 10B). To avoid a clash between the Ara-FHNA hexose and the 3'-adjacent deoxyribose, the duplex undergoes a slight conformational change that results in partial unstacking of the thymine and adenine bases [221], explaining the lower RNA affinity of Ara-FHNA compared to FHNA. Further experiments in vivo also demonstrated the effectiveness of FHNA-modified siRNA in the downregulation of mouse phosphatase and tensin homologue (PTEN) without inducing hepatotoxicity

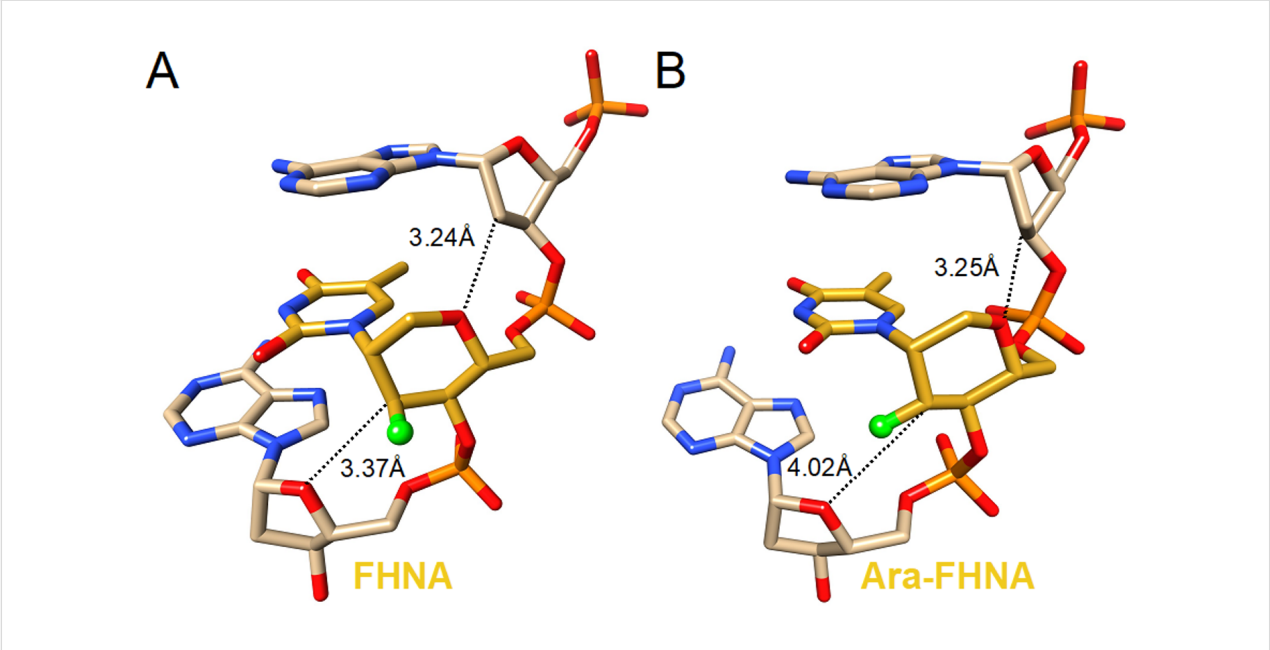


Figure 10: Crystal structures of (A) FHNA and (B) Ara-FHNA in modified A-form DNA decamers (PDB IDs 3Q61 and 3SD8, respectively) [221]. Unlike the trans-diaxial orientation of the fluorine in FHNA, the equatorial orientation of fluorine in Ara-FHNA pushes away the 3'-adjacent nucleotide (dashed lines) and causes local unstacking of bases.

[221]. Recent work has also shown that FHNA modifications improve the potency of GalNAc-conjugated gapmer ASOs [226].

Methylation at the C6' position further influences the RNA affinity of nucleic acids containing these modifications. *R*-6'-Me-FHNA is highly destabilizing, whereas *S*-6'-Me-FHNA leads to duplex stabilization [227]. This trend is identical to the C5' methylation of LNA [228]. The 1.24 Å crystal structures of A-form decamer duplexes containing these C6'-methylations show a small 1–5 intranucleoside contact between the C6' methyl group and the O4' in *R*-6'-Me-FHNA [227]. Additionally, *R*-6'-Me-FHNA perturbs the structure of water surrounding the O2P atoms which will further reduce the pairing affinity of the *R* isomer [227].

Herdewijn recently published the synthesis of 4'-aminotritylhexitol nucleosides for the eventual synthesis of N4' → P6' phosphoramidates of aminohexitol nucleic acids (AHNA) [229], as well as the synthesis of 3'-fluoro-4'-aminohexitol nucleosides which contain both the 3'-fluoro functionality and the N4' → P6' phosphoramidate linkage [230].

Ribo-difluorotoluyl

2'-Deoxydifluorotoluyl (dF) nucleoside derivatives (Figure 1J) were first synthesized by Schweitzer and Kool in 1994 in order to study the importance of H-bonding and base stacking in DNA. Specifically, they focused on the 2,4-difluorotoluene moiety as an isostere of the natural thymine base, albeit without the ability to form H-bonds [231]. A few years later, in 1997, Moran et al. showed that dF was a good template for enzymatic DNA synthesis, permitting production of the complementary DNA strand and hence suggesting that shape complementarity may be more important than H-bonding for fidelity and efficiency of DNA polymerases [232,233]. Recently, the rF nucleoside analogue has been investigated for its ability to efficiently silence gene expression when incorporated into short interfering RNA (siRNA) duplexes and to further investigate the fidelity of various RNA polymerases [234-236]. siRNA guide strands modified at the 5' end with rF showed similar silencing to the unmodified control. Furthermore, internal rF modifications showed lower affinity for their target but exhibited higher nuclease resistance [235,237]. Moreover, the rF/A pair lowers the T_m of the siRNA duplex but is less destabilizing than a mismatch (A/A, C/A and G/A) [235]. Several crystal structures of oligonucleotides containing the dF or rF nucleoside analogue alone and oligos with dF bound to DNA polymerases have been determined [235,237-240]. The 1.6 Å resolution structure of the Dickerson–Drew dodecamer (DDD) with dF replacing T8 (i.e., dCGCGAATFCGCG), solved with crystals of the duplex grown in the presence of *Bacillus halodurans* RNase H (which was bound to the duplex but did not exert an influence on its structure), revealed distances of 3.09 and 3.12 Å for the F^4(dF)···N^6(A) atoms of the two dF:A pairs similar to the O^4(T)···N^6(A) distance (2.96 and 3.11 Å) observed for the native DDD [240]. The 1.6 Å crystal structure of a duplex

containing the rF analog ([rCGCFAAUUAGCG]$_2$) revealed a F^4(rF)···N^6(A) distance of approximately 4 Å between the rF:A pairs [235].

Conclusion

Chemically modified oligonucleotides have come of age as a class of therapeutic agents for a number of diseases. Taking inspiration from the structure, properties and biological roles of nucleic acids, scientists have employed chemistry to prepare a diverse collection of modifications to the architecture of this molecule imbuing desirable characteristics for applications as a therapeutic agent. In addition, many nucleic acid analogs have been explored for additional studies including investigation of artificial genetic systems, catalysts, and sensors. Amongst the oligonucleotide-based therapeutics that have been approved as drugs, the dominating modifications are the phosphorothioate backbone and at the C2'-position (of ribose) including 2'-*O*Me, 2'-F, and 2'-*O*-MOE. Moreover, combinations of these modifications in an oligonucleotide leads to a synergistic effect enhancing their therapeutic properties. Such combinations of nucleotide and backbone modifications with numerous analogs that have been developed will continue as an exciting direction for the next generation of oligonucleotide-based therapeutics. Rational design of future modifications with improved properties may be gleaned from insights from structural techniques. For example, stability, gene silencing and structural studies of chemically modified oligonucleotides containing fluorine at the sugar and nucleobase have provided insights into the role of noncovalent interactions on the properties of these molecules. The partnership between organic synthesis, biophysical chemistry, biochemistry and structural biology continues to guide the design and drive the achievements for oligonucleotide-based therapeutics.

Acknowledgements

The authors are grateful to the numerous co-workers and collaborators, from academia and industry, with whom we have worked over the years.

Funding

We thank the Natural Sciences and Engineering Research Council of Canada (Grant No. RGPIN-2017-06683 to C.J.W.), Alnylam Pharmaceuticals Inc. Cambridge, MA (M.E.) and the US National Institutes of Health for financial support (Grants No. NIH P01 CA160032, R01 ES026966 and R01 ES029357 to M.E.). C.L. is the recipient of a NSERC Alexander Graham Bell Canada Graduate Scholarship - Doctoral (CGS-D) and Miriam Aaron Roland Graduate Fellowship. K.D. is the recipient of a NSERC Postgraduate Scholarship - Master's (PGS-M) and a Hydro-Quebec Master's Scholarship. G.J. is a trainee of the NSERC Collaborative Research and Training Experience Program in Programmed Molecules for Therapeutics, Sensing and Diagnostics (PROMOTE) and a recipient of a Concordia Merit Scholarship.

ORCID® iDs

Christopher Liczner - https://orcid.org/0000-0002-1741-0159
Martin Egli - https://orcid.org/0000-0003-4145-356X
Christopher J. Wilds - https://orcid.org/0000-0002-0336-4753

References

1. Sood, A. J.; Viner, C.; Hoffman, M. M. *J. Cheminf.* **2019,** *11,* 30. doi:10.1186/s13321-019-0349-4
2. Boccaletto, P.; Machnicka, M. A.; Purta, E.; Piątkowski, P.; Bagiński, B.; Wirecki, T. K.; de Crécy-Lagard, V.; Ross, R.; Limbach, P. A.; Kotter, A.; Helm, M.; Bujnicki, J. M. *Nucleic Acids Res.* **2018,** *46,* D303–D307. doi:10.1093/nar/gkx1030
3. Wagner, M.; Steinbacher, J.; Kraus, T. F. J.; Michalakis, S.; Hackner, B.; Pfaffeneder, T.; Perera, A.; Müller, M.; Giese, A.; Kretzschmar, H. A.; Carell, T. *Angew. Chem., Int. Ed.* **2015,** *54,* 12511–12514. doi:10.1002/anie.201502722
4. Wang, L.; Chen, S.; Xu, T.; Taghizadeh, K.; Wishnok, J. S.; Zhou, X.; You, D.; Deng, Z.; Dedon, P. C. *Nat. Chem. Biol.* **2007,** *3,* 709–710. doi:10.1038/nchembio.2007.39
5. Kellner, S.; DeMott, M. S.; Cheng, C. P.; Russell, B. S.; Cao, B.; You, D.; Dedon, P. C. *Nat. Chem. Biol.* **2017,** *13,* 888–894. doi:10.1038/nchembio.2407
6. Wu, Y.; Tang, Y.; Dong, X.; Zheng, Y. Y.; Haruehanroengra, P.; Mao, S.; Lin, Q.; Sheng, J. *ACS Chem. Biol.* **2020,** *15,* 1301–1305. doi:10.1021/acschembio.0c00163
7. Züst, R.; Cervantes-Barragan, L.; Habjan, M.; Maier, R.; Neuman, B. W.; Ziebuhr, J.; Szretter, K. J.; Baker, S. C.; Barchet, W.; Diamond, M. S.; Siddell, S. G.; Ludewig, B.; Thiel, V. *Nat. Immunol.* **2011,** *12,* 137–143. doi:10.1038/ni.1979
8. Motorin, Y.; Marchand, V. *Genes* **2018,** *9,* 642. doi:10.3390/genes9120642
9. Elliott, B. A.; Ho, H.-T.; Ranganathan, S. V.; Vangaveti, S.; Ilkayeva, O.; Abou Assi, H.; Choi, A. K.; Agris, P. F.; Holley, C. L. *Nat. Commun.* **2019,** *10,* 3401. doi:10.1038/s41467-019-11375-7
10. Dimitrova, D. G.; Teysset, L.; Carré, C. *Genes* **2019,** *10,* 117. doi:10.3390/genes10020117
11. Freier, S. M.; Altmann, K.-H. *Nucleic Acids Res.* **1997,** *25,* 4429–4443. doi:10.1093/nar/25.22.4429
12. Manoharan, M. *Biochim. Biophys. Acta, Gene Struct. Expression* **1999,** *1489,* 117–130. doi:10.1016/s0167-4781(99)00138-4
13. Manoharan, M.; Rajeev, K. G. Utilizing Chemistry to Harness RNA Interference Pathways for Therapeutics: Chemically Modified siRNAs and Antagomirs. In *Antisense Drug Technology,* 2nd ed.; Crooke, S. T., Ed.; CRC Press: Boca Raton, FL, USA, 2008; pp 437–464.
14. Ni, S.; Yao, H.; Wang, L.; Lu, J.; Jiang, F.; Lu, A.; Zhang, G. *Int. J. Mol. Sci.* **2017,** *18,* 1683. doi:10.3390/ijms18081683
15. Egli, M.; Manoharan, M. *Acc. Chem. Res.* **2019,** *52,* 1036–1047. doi:10.1021/acs.accounts.8b00650
16. Yamasaki, K.; Akutsu, Y.; Yamasaki, T.; Miyagishi, M.; Kubota, T. *Nucleic Acids Res.* **2020,** *48,* 4551–4561. doi:10.1093/nar/gkaa170

17. Hyjek-Składanowska, M.; Vickers, T. A.; Napiórkowska, A.; Anderson, B. A.; Tanowitz, M.; Crooke, S. T.; Liang, X.-h.; Seth, P. P.; Nowotny, M. *J. Am. Chem. Soc.* **2020,** *142,* 7456–7468. doi:10.1021/jacs.9b13524
18. Shen, X.; Corey, D. R. *Nucleic Acids Res.* **2018,** *46,* 1584–1600. doi:10.1093/nar/gkx1239
19. Crooke, S. T.; Witztum, J. L.; Bennett, C. F.; Baker, B. F. *Cell Metab.* **2018,** *27,* 714–739. doi:10.1016/j.cmet.2018.03.004
20. Roberts, T. C.; Langer, R.; Wood, M. J. A. *Nat. Rev. Drug Discovery* **2020,** *19,* 673–694. doi:10.1038/s41573-020-0075-7
21. Egli, M.; Manoharan, M. *Nucleic Acids Res.* submitted.
22. Furukawa, Y.; Kobayashi, K.; Kanai, Y.; Honjo, M. *Chem. Pharm. Bull.* **1965,** *13,* 1273–1278. doi:10.1248/cpb.13.1273
23. Eckstein, F. *J. Am. Chem. Soc.* **1966,** *88,* 4292–4294. doi:10.1021/ja00970a054
24. Codington, J. F.; Doerr, I.; Van Praag, D.; Bendich, A.; Fox, J. J. *J. Am. Chem. Soc.* **1961,** *83,* 5030–5031. doi:10.1021/ja01485a036
25. Westheimer, F. *Science* **1987,** *235,* 1173–1178. doi:10.1126/science.2434996
26. Kamerlin, S. C. L.; Sharma, P. K.; Prasad, R. B.; Warshel, A. *Q. Rev. Biophys.* **2013,** *46,* 1–132. doi:10.1017/s0033583512000157
27. Rich, A. *Nat. Struct. Mol. Biol.* **2003,** *10,* 247–249. doi:10.1038/nsb0403-247
28. Egli, M. Sugar Pucker and Nucleic Acid Structure. In *The Excitement of Discovery: Selected Papers of Alexander Rich: A Tribute to Alexander Rich;* Zhang, S., Ed.; Series in Structural Biology; World Scientific Publishers: Singapore, 2018; pp 309–315. doi:10.1142/9789813272682_others05
29. Tamura, M.; Holbrook, S. R. *J. Mol. Biol.* **2002,** *320,* 455–474. doi:10.1016/s0022-2836(02)00515-6
30. Butcher, S. E.; Pyle, A. M. *Acc. Chem. Res.* **2011,** *44,* 1302–1311. doi:10.1021/ar200098t
31. Denning, E. J.; MacKerell, A. D., Jr. *J. Am. Chem. Soc.* **2012,** *134,* 2800–2806. doi:10.1021/ja211328g
32. Darré, L.; Ivani, I.; Dans, P. D.; Gómez, H.; Hospital, A.; Orozco, M. *J. Am. Chem. Soc.* **2016,** *138,* 16355–16363. doi:10.1021/jacs.6b09471
33. Shi, Y. *Nat. Rev. Mol. Cell Biol.* **2017,** *18,* 655–670. doi:10.1038/nrm.2017.86
34. Fica, S. M.; Nagai, K. *Nat. Struct. Mol. Biol.* **2017,** *24,* 791–799. doi:10.1038/nsmb.3463
35. Peebles, C. L.; Perlman, P. S.; Mecklenburg, K. L.; Petrillo, M. L.; Tabor, J. H.; Jarrell, K. A.; Cheng, H.-L. *Cell* **1986,** *44,* 213–223. doi:10.1016/0092-8674(86)90755-5
36. Qin, P. Z.; Pyle, A. M. *Curr. Opin. Struct. Biol.* **1998,** *8,* 301–308. doi:10.1016/s0959-440x(98)80062-6
37. Cech, T. *Science* **1987,** *236,* 1532–1539. doi:10.1126/science.2438771
38. Emilsson, G. M.; Nakamura, S.; Roth, A.; Breaker, R. R. *RNA* **2003,** *9,* 907–918. doi:10.1261/rna.5680603
39. O'Rourke, S. M.; Scott, W. G. *Prog. Mol. Biol. Transl. Sci.* **2018,** *159,* 177–202. doi:10.1016/bs.pmbts.2018.07.006
40. Egli, M.; Lubini, P.; Pallan, P. S. *Chem. Soc. Rev.* **2007,** *36,* 31–45. doi:10.1039/b606807c
41. Egli, M.; Minasov, G.; Tereshko, V.; Pallan, P. S.; Teplova, M.; Inamati, G. B.; Lesnik, E. A.; Owens, S. R.; Ross, B. S.; Prakash, T. P.; Manoharan, M. *Biochemistry* **2005,** *44,* 9045–9057. doi:10.1021/bi050574m
42. Deleavey, G. F.; Damha, M. J. *Chem. Biol.* **2012,** *19,* 937–954. doi:10.1016/j.chembiol.2012.07.011
43. Seth, P. P.; Swayze, E. E. Unnatural Nucleoside Analogs for Antisense Therapy. In *Natural Products in Medicinal Chemistry;* Hanessian, S., Ed.; Wiley-VCH: Weinheim, Germany, 2014; Vol. 60, pp 403–440. doi:10.1002/9783527676545.ch12
44. Campbell, M. A.; Wengel, J. *Chem. Soc. Rev.* **2011,** *40,* 5680–5689. doi:10.1039/c1cs15048k
45. Kotikam, V.; Rozners, E. *Acc. Chem. Res.* **2020,** *53,* 1782–1790. doi:10.1021/acs.accounts.0c00249
46. Wan, W. B.; Migawa, M. T.; Vasquez, G.; Murray, H. M.; Nichols, J. G.; Gaus, H.; Berdeja, A.; Lee, S.; Hart, C. E.; Lima, W. F.; Swayze, E. E.; Seth, P. P. *Nucleic Acids Res.* **2014,** *42,* 13456–13468. doi:10.1093/nar/gku1115
47. Iwamoto, N.; Butler, D. C. D.; Svrzikapa, N.; Mohapatra, S.; Zlatev, I.; Sah, D. W. Y.; Meena; Standley, S. M.; Lu, G.; Apponi, L. H.; Frank-Kamenetsky, M.; Zhang, J. J.; Vargeese, C.; Verdine, G. L. *Nat. Biotechnol.* **2017,** *35,* 845–851. doi:10.1038/nbt.3948
48. Schöning, K.-U.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A. *Science* **2000,** *290,* 1347–1351. doi:10.1126/science.290.5495.1347
49. Maier, T.; Przylas, I.; Strater, N.; Herdewijn, P.; Saenger, W. *J. Am. Chem. Soc.* **2005,** *127,* 2937–2943. doi:10.1021/ja045843v
50. Kumar, P.; Degaonkar, R.; Guenther, D. C.; Abramov, M.; Schepers, G.; Capobianco, M.; Jiang, Y.; Harp, J.; Kaittanis, C.; Janas, M. M.; Castoreno, A.; Zlatev, I.; Schlegel, M. K.; Herdewijn, P.; Egli, M.; Manoharan, M. *Nucleic Acids Res.* **2020,** *48,* 4028–4040. doi:10.1093/nar/gkaa125
51. Maiti, M.; Maiti, M.; Knies, C.; Dumbre, S.; Lescrinier, E.; Rosemeyer, H.; Ceulemans, A.; Herdewijn, P. *Nucleic Acids Res.* **2015,** *43,* 7189–7200. doi:10.1093/nar/gkv719
52. Ovaere, M.; Herdewijn, P.; Van Meervelt, L. *Chem. – Eur. J.* **2011,** *17,* 7823–7830. doi:10.1002/chem.201003594
53. Moulton, H. M.; Moulton, J. D., Eds. *Morpholino Oligomers;* Methods in Molecular Biology, Vol. 1565; Humana Press, 2017. doi:10.1007/978-1-4939-6817-6
54. Schlegel, M. K.; Foster, D. J.; Kel'in, A. V.; Zlatev, I.; Bisbe, A.; Jayaraman, M.; Lackey, J. G.; Rajeev, K. G.; Charissé, K.; Harp, J.; Pallan, P. S.; Maier, M. A.; Egli, M.; Manoharan, M. *J. Am. Chem. Soc.* **2017,** *139,* 8537–8546. doi:10.1021/jacs.7b02694
55. Anosova, I.; Kowal, E. A.; Dunn, M. R.; Chaput, J. C.; Van Horn, W. D.; Egli, M. *Nucleic Acids Res.* **2016,** *44,* 1007–1021. doi:10.1093/nar/gkv1472
56. Devine, K. G.; Jheeta, S. *Life* **2020,** *10,* 346. doi:10.3390/life10120346
57. Eriksson, M.; Nielsen, P. E. *Q. Rev. Biophys.* **1996,** *29,* 369–394. doi:10.1017/s0033583500005886
58. Teplova, M.; Minasov, G.; Tereshko, V.; Inamati, G. B.; Cook, P. D.; Manoharan, M.; Egli, M. *Nat. Struct. Biol.* **1999,** *6,* 535–539. doi:10.1038/9304
59. Crooke, S. T.; Baker, B. F.; Kwoh, T. J.; Cheng, W.; Schulz, D. J.; Xia, S.; Salgado, N.; Bui, H.-H.; Hart, C. E.; Burel, S. A.; Younis, H. S.; Geary, R. S.; Henry, S. P.; Bhanot, S. *Mol. Ther.* **2016,** *24,* 1771–1782. doi:10.1038/mt.2016.136
60. Koch, T. *Curr. Phys. Chem.* **2013,** *3,* 55–68. doi:10.2174/1877946811303010009
61. Manoharan, M.; Akinc, A.; Pandey, R. K.; Qin, J.; Hadwiger, P.; John, M.; Mills, K.; Charisse, K.; Maier, M. A.; Nechev, L.; Greene, E. M.; Pallan, P. S.; Rozners, E.; Rajeev, K. G.; Egli, M. *Angew. Chem., Int. Ed.* **2011,** *50,* 2284–2288. doi:10.1002/anie.201006519

62. Pallan, P. S.; Greene, E. M.; Jicman, P. A.; Pandey, R. K.; Manoharan, M.; Rozners, E.; Egli, M. *Nucleic Acids Res.* **2011,** *39,* 3482–3495. doi:10.1093/nar/gkq1270
63. Patra, A.; Paolillo, M.; Charisse, K.; Manoharan, M.; Rozners, E.; Egli, M. *Angew. Chem., Int. Ed.* **2012,** *51,* 11863–11866. doi:10.1002/anie.201204946
64. Gryaznov, S.; Chen, J.-K. *J. Am. Chem. Soc.* **1994,** *116,* 3143–3144. doi:10.1021/ja00086a062
65. Gryaznov, S. M.; Lloyd, D. H.; Chen, J. K.; Schultz, R. G.; DeDionisio, L. A.; Ratmeyer, L.; Wilson, W. D. *Proc. Natl. Acad. Sci. U. S. A.* **1995,** *92,* 5798–5802. doi:10.1073/pnas.92.13.5798
66. Chen, J.-K.; Schultz, R. G.; Lioyd, D. H.; Gryaznov, S. M. *Nucleic Acids Res.* **1995,** *23,* 2661–2668. doi:10.1093/nar/23.14.2661
67. Skorski, T.; Perrotti, D.; Nieborowska-Skorska, M.; Gryaznov, S.; Calabretta, B. *Proc. Natl. Acad. Sci. U. S. A.* **1997,** *94,* 3966–3971. doi:10.1073/pnas.94.8.3966
68. Heidenreich, O.; Gryaznov, S.; Nerenberg, M. *Nucleic Acids Res.* **1997,** *25,* 776–780. doi:10.1093/nar/25.4.776
69. Giovannangeli, C.; Diviacco, S.; Labrousse, V.; Gryaznov, S.; Charneau, P.; Helene, C. *Proc. Natl. Acad. Sci. U. S. A.* **1997,** *94,* 79–84. doi:10.1073/pnas.94.1.79
70. Rigl, C. T.; Lloyd, D. H.; Tsou, D. S.; Gryaznov, S. M.; Wilson, W. D. *Biochemistry* **1997,** *36,* 650–659. doi:10.1021/bi961980w
71. Tereshko, V.; Gryaznov, S.; Egli, M. *J. Am. Chem. Soc.* **1998,** *120,* 269–283. doi:10.1021/ja971962h
72. Escude, C.; Giovannangeli, C.; Sun, J. S.; Lloyd, D. H.; Chen, J. K.; Gryaznov, S. M.; Garestier, T.; Helene, C. *Proc. Natl. Acad. Sci. U. S. A.* **1996,** *93,* 4365–4369. doi:10.1073/pnas.93.9.4365
73. Mutisya, D.; Hardcastle, T.; Cheruiyot, S. K.; Pallan, P. S.; Kennedy, S. D.; Egli, M.; Kelley, M. L.; van Brabant Smith, A.; Rozners, E. *Nucleic Acids Res.* **2017,** *45,* 8142–8155. doi:10.1093/nar/gkx558
74. Wilds, C. J.; Minasov, G.; Natt, F.; von Matt, P.; Altmann, K.-H.; Egli, M. *Nucleosides, Nucleotides Nucleic Acids* **2001,** *20,* 991–994. doi:10.1081/ncn-100002475
75. Mutisya, D.; Selvam, C.; Lunstad, B. D.; Pallan, P. S.; Haas, A.; Leake, D.; Egli, M.; Rozners, E. *Nucleic Acids Res.* **2014,** *42,* 6542–6551. doi:10.1093/nar/gku235
76. Idziak, I.; Just, G.; Damha, M. J.; Giannaris, P. A. *Tetrahedron Lett.* **1993,** *34,* 5417–5420. doi:10.1016/s0040-4039(00)73923-2
77. Rozners, E.; Katkevica, D.; Bizdena, E.; Strömberg, R. *J. Am. Chem. Soc.* **2003,** *125,* 12125–12136. doi:10.1021/ja0360900
78. von Matt, P.; De Mesmaeker, A.; Pieles, U.; Zürcher, W.; Altmann, K.-H. *Tetrahedron Lett.* **1999,** *40,* 2899–2902. doi:10.1016/s0040-4039(99)00389-5
79. Selvam, C.; Thomas, S.; Abbott, J.; Kennedy, S. D.; Rozners, E. *Angew. Chem., Int. Ed.* **2011,** *50,* 2068–2070. doi:10.1002/anie.201007012
80. Iwase, R.; Kurokawa, R.; Ueno, J. *Nucleic Acids Symp. Ser.* **2009,** *53,* 119–120. doi:10.1093/nass/nrp060
81. Epple, S.; Thorpe, C.; Baker, Y. R.; El-Sagheer, A. H.; Brown, T. *Chem. Commun.* **2020,** *56,* 5496–5499. doi:10.1039/d0cc00444h
82. Elkayam, E.; Kuhn, C.-D.; Tocilj, A.; Haase, A. D.; Greene, E. M.; Hannon, G. J.; Joshua-Tor, L. *Cell* **2012,** *150,* 100–110. doi:10.1016/j.cell.2012.05.017
83. Hardcastle, T.; Novosjolova, I.; Kotikam, V.; Cheruiyot, S. K.; Mutisya, D.; Kennedy, S. D.; Egli, M.; Kelley, M. L.; Smith, A. v. B.; Rozners, E. *ACS Chem. Biol.* **2018,** *13,* 533–536. doi:10.1021/acschembio.7b01012
84. Beaton, G.; Brill, W. K.-D.; Grandas, A.; Ma, Y.-X.; Nielsen, J.; Yau, E.; Caruthers, M. H. *Tetrahedron* **1991,** *47,* 2377–2388. doi:10.1016/s0040-4020(01)81775-0
85. Marshall, W.; Caruthers, M. *Science* **1993,** *259,* 1564–1570. doi:10.1126/science.7681216
86. Iyer, R. P.; Egan, W.; Regan, J. B.; Beaucage, S. L. *J. Am. Chem. Soc.* **1990,** *112,* 1253–1254. doi:10.1021/ja00159a059
87. Wickstrom, E. *Adv. Drug Delivery Rev.* **2015,** *87,* 25–34. doi:10.1016/j.addr.2015.04.012
88. Clavé, G.; Reverte, M.; Vasseur, J.-J.; Smietana, M. *RSC Chem. Biol.* **2021,** *2,* 94–150. doi:10.1039/d0cb00136h
89. Pallan, P. S.; Lybrand, T. P.; Schlegel, M. K.; Harp, J. M.; Jahns, H.; Manoharan, M.; Egli, M. *Biochemistry* **2020,** *59,* 4627–4637. doi:10.1021/acs.biochem.0c00685
90. Caruthers, M. H. *Acc. Chem. Res.* **1991,** *24,* 278–284. doi:10.1021/ar00009a005
91. Yang, X.; Sierant, M.; Janicka, M.; Peczek, L.; Martinez, C.; Hassell, T.; Li, N.; Li, X.; Wang, T.; Nawrot, B. *ACS Chem. Biol.* **2012,** *7,* 1214–1220. doi:10.1021/cb300078e
92. Pallan, P. S.; Yang, X.; Sierant, M.; Abeydeera, N. D.; Hassell, T.; Martinez, C.; Janicka, M.; Nawrot, B.; Egli, M. *RSC Adv.* **2014,** *4,* 64901–64904. doi:10.1039/c4ra10986d
93. Wu, S. Y.; Yang, X.; Gharpure, K. M.; Hatakeyama, H.; Egli, M.; McGuire, M. H.; Nagaraja, A. S.; Miyake, T. M.; Rupaimoole, R.; Pecot, C. V.; Taylor, M.; Pradeep, S.; Sierant, M.; Rodriguez-Aguayo, C.; Choi, H. J.; Previs, R. A.; Armaiz-Pena, G. N.; Huang, L.; Martinez, C.; Hassell, T.; Ivan, C.; Sehgal, V.; Singhania, R.; Han, H.-D.; Su, C.; Kim, J. H.; Dalton, H. J.; Kovvali, C.; Keyomarsi, K.; McMillan, N. A. J.; Overwijk, W. W.; Liu, J.; Lee, J.-S.; Baggerly, K. A.; Lopez-Berestein, G.; Ram, P. T.; Nawrot, B.; Sood, A. K. *Nat. Commun.* **2014,** *5,* 3459. doi:10.1038/ncomms4459
94. Egli, M.; Lybrand, T. P. *J. Am. Chem. Soc.* **2019,** *141,* 4445–4452. doi:10.1021/jacs.9b00104
95. Abeydeera, N. D.; Egli, M.; Cox, N.; Mercier, K.; Conde, J. N.; Pallan, P. S.; Mizurini, D. M.; Sierant, M.; Hibti, F.-E.; Hassell, T.; Wang, T.; Liu, F.-W.; Liu, H.-M.; Martinez, C.; Sood, A. K.; Lybrand, T. P.; Frydman, C.; Monteiro, R. Q.; Gomer, R. H.; Nawrot, B.; Yang, X. *Nucleic Acids Res.* **2016,** *44,* 8052–8064. doi:10.1093/nar/gkw725
96. Ueda, N.; Kawabata, T.; Takemoto, K. *J. Heterocycl. Chem.* **1971,** *8,* 827–829. doi:10.1002/jhet.5570080527
97. Seita, T.; Yamauchi, K.; Kinoshita, M.; Imoto, M. *Makromol. Chem.* **1972,** *154,* 255–261. doi:10.1002/macp.1972.021540123
98. Acevedo, O. L.; Andrews, R. S. *Tetrahedron Lett.* **1996,** *37,* 3931–3934. doi:10.1016/0040-4039(96)00745-9
99. Zhang, L.; Peritz, A.; Meggers, E. *J. Am. Chem. Soc.* **2005,** *127,* 4174–4175. doi:10.1021/ja042564z
100. Chen, J. J.; Cai, X.; Szostak, J. W. *J. Am. Chem. Soc.* **2009,** *131,* 2119–2121. doi:10.1021/ja809069b
101. Schlegel, M. K.; Peritz, A. E.; Kittigowittana, K.; Zhang, L.; Meggers, E. *ChemBioChem* **2007,** *8,* 927–932. doi:10.1002/cbic.200600435
102. Tsai, C.-H.; Chen, J.; Szostak, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104,* 14598–14603. doi:10.1073/pnas.0704211104
103. Meggers, E.; Zhang, L. *Acc. Chem. Res.* **2010,** *43,* 1092–1102. doi:10.1021/ar900292q

104. Nielsen, P.; Dreiøe, L. H.; Wengel, J. *Bioorg. Med. Chem.* **1995,** *3,* 19–28. doi:10.1016/0968-0896(94)00143-q
105. Schlegel, M. K.; Essen, L.-O.; Meggers, E. *J. Am. Chem. Soc.* **2008,** *130,* 8158–8159. doi:10.1021/ja802788g
106. Johnson, A. T.; Schlegel, M. K.; Meggers, E.; Essen, L.-O.; Wiest, O. *J. Org. Chem.* **2011,** *76,* 7964–7974. doi:10.1021/jo201469b
107. Schlegel, M. K.; Essen, L.-O.; Meggers, E. *Chem. Commun.* **2010,** *46,* 1094–1096. doi:10.1039/b916851f
108. Inoue, H.; Hayase, Y.; Imura, A.; Iwai, S.; Miura, K.; Ohtsuka, E. *Nucleic Acids Res.* **1987,** *15,* 6131–6148. doi:10.1093/nar/15.15.6131
109. Ross, B. S.; Springer, R. H.; Tortorici, Z.; Dimock, S. *Nucleosides Nucleotides* **1997,** *16,* 1641–1643. doi:10.1080/07328319708006245
110. Beigelman, L.; Haeberli, P.; Sweedler, D.; Karpeisky, A. *Tetrahedron* **2000,** *56,* 1047–1056. doi:10.1016/s0040-4020(00)00002-8
111. Roy, S. K.; Tang, J.-y. *Org. Process Res. Dev.* **2000,** *4,* 170–171. doi:10.1021/op990100t
112. Cummins, L. L.; Owens, S. R.; Risen, L. M.; Lesnik, E. A.; Freier, S. M.; McGee, D.; Guinosso, C. J.; Cook, P. D. *Nucleic Acids Res.* **1995,** *23,* 2019–2024. doi:10.1093/nar/23.11.2019
113. Inoue, H.; Hayase, Y.; Iwai, S.; Ohtsuka, E. *FEBS Lett.* **1987,** *215,* 327–330. doi:10.1016/0014-5793(87)80171-0
114. Lubini, P.; Zürcher, W.; Egli, M. *Chem. Biol.* **1994,** *1,* 39–45. doi:10.1016/1074-5521(94)90039-6
115. Lesnik, E. A.; Guinosso, C. J.; Kawasaki, A. M.; Sasmor, H.; Zounes, M.; Cummins, L. L.; Ecker, D. J.; Cook, P. D.; Freier, S. M. *Biochemistry* **1993,** *32,* 7832–7838. doi:10.1021/bi00081a031
116. Martin, P. *Helv. Chim. Acta* **1995,** *78,* 486–504. doi:10.1002/hlca.19950780219
117. Legorburu, U.; Reese, C. B.; Song, Q. *Tetrahedron* **1999,** *55,* 5635–5640. doi:10.1016/s0040-4020(99)00229-x
118. Chow, S.; Wen, K.; Sanghvi, Y. S.; Theodorakis, E. A. *Nucleosides, Nucleotides Nucleic Acids* **2003,** *22,* 583–587. doi:10.1081/ncn-120021959
119. Altmann, K.-H.; Fabbro, D.; Dean, N. M.; Geiger, T.; Monia, B. P.; Müllert, M.; Nicklin, P. *Biochem. Soc. Trans.* **1996,** *24,* 630–637. doi:10.1042/bst0240630
120. Tereshko, V.; Portmann, S.; Tay, E. C.; Martin, P.; Natt, F.; Altmann, K.-H.; Egli, M. *Biochemistry* **1998,** *37,* 10626–10634. doi:10.1021/bi980392a
121. Prakash, T. P. *Chem. Biodiversity* **2011,** *8,* 1616–1641. doi:10.1002/cbdv.201100081
122. Stephenson, M. L.; Zamecnik, P. C. *Proc. Natl. Acad. Sci. U. S. A.* **1978,** *75,* 285–288. doi:10.1073/pnas.75.1.285
123. Barabino, S. M. L.; Blencowe, B. J.; Ryder, U.; Sproat, B. S.; Lamond, A. I. *Cell* **1990,** *63,* 293–302. doi:10.1016/0092-8674(90)90162-8
124. Iribarren, A. M.; Sproat, B. S.; Neuner, P.; Sulston, I.; Ryder, U.; Lamond, A. I. *Proc. Natl. Acad. Sci. U. S. A.* **1990,** *87,* 7747–7751. doi:10.1073/pnas.87.19.7747
125. Boiziau, C.; Larrouy, B.; Sproat, B. S.; Toulmè, J.-J. *Nucleic Acids Res.* **1995,** *23,* 64–71. doi:10.1093/nar/23.1.64
126. Obika, S.; Nanbu, D.; Hari, Y.; Morio, K.-i.; In, Y.; Ishida, T.; Imanishi, T. *Tetrahedron Lett.* **1997,** *38,* 8735–8738. doi:10.1016/s0040-4039(97)10322-7
127. Singh, S. K.; Koshkin, A. A.; Wengel, J.; Nielsen, P. *Chem. Commun.* **1998,** 455–456. doi:10.1039/a708608c
128. Koshkin, A. A.; Singh, S. K.; Nielsen, P.; Rajwanshi, V. K.; Kumar, R.; Meldgaard, M.; Olsen, C. E.; Wengel, J. *Tetrahedron* **1998,** *54,* 3607–3630. doi:10.1016/s0040-4020(98)00094-5
129. Nielsen, C. B.; Singh, S. K.; Wengel, J.; Jacobsen, J. P. *J. Biomol. Struct. Dyn.* **1999,** *17,* 175–191. doi:10.1080/07391102.1999.10508352
130. Braasch, D. A.; Corey, D. R. *Chem. Biol.* **2001,** *8,* 1–7. doi:10.1016/s1074-5521(00)00058-2
131. Obika, S.; Nanbu, D.; Hari, Y.; Andoh, J.-i.; Morio, K.-i.; Doi, T.; Imanishi, T. *Tetrahedron Lett.* **1998,** *39,* 5401–5404. doi:10.1016/s0040-4039(98)01084-3
132. Petersen, M.; Wengel, J. *Trends Biotechnol.* **2003,** *21,* 74–81. doi:10.1016/s0167-7799(02)00038-0
133. Singh, S. K.; Wengel, J. *Chem. Commun.* **1998,** 1247–1248. doi:10.1039/a801571f
134. Wengel, J. *Acc. Chem. Res.* **1999,** *32,* 301–310. doi:10.1021/ar980051p
135. Bondensgaard, K.; Petersen, M.; Singh, S. K.; Rajwanshi, V. K.; Kumar, R.; Wengel, J.; Jacobsen, J. P. *Chem. – Eur. J.* **2000,** *6,* 2687–2695. doi:10.1002/1521-3765(20000804)6:15<2687::aid-chem2687>3.0.co;2-u
136. Petersen, M.; Bondensgaard, K.; Wengel, J.; Jacobsen, J. P. *J. Am. Chem. Soc.* **2002,** *124,* 5974–5982. doi:10.1021/ja012288d
137. Schmidt, K. S.; Borkowski, S.; Kurreck, J.; Stephens, A. W.; Bald, R.; Hecht, M.; Friebe, M.; Dinkelborg, L.; Erdmann, V. A. *Nucleic Acids Res.* **2004,** *32,* 5757–5765. doi:10.1093/nar/gkh862
138. Kaur, H.; Babu, B. R.; Maiti, S. *Chem. Rev.* **2007,** *107,* 4672–4697. doi:10.1021/cr050266u
139. Fluiter, K.; ten Asbroek, A. L. M. A.; de Wissel, M. B.; Jakobs, M. E.; Wissenbach, M.; Olsson, H.; Olsen, O.; Oerum, H.; Baas, F. *Nucleic Acids Res.* **2003,** *31,* 953–962. doi:10.1093/nar/gkg185
140. Grünweller, A.; Wyszko, E.; Bieber, B.; Jahnel, R.; Erdmann, V. A.; Kurreck, J. *Nucleic Acids Res.* **2003,** *31,* 3185–3193. doi:10.1093/nar/gkg409
141. Schubert, S.; Gül, D. C.; Grunert, H.-P.; Zeichhardt, H.; Erdmann, V. A.; Kurreck, J. *Nucleic Acids Res.* **2003,** *31,* 5982–5992. doi:10.1093/nar/gkg791
142. Kuespert, S.; Heydn, R.; Peters, S.; Wirkert, E.; Meyer, A.-L.; Siebörger, M.; Johannesen, S.; Aigner, L.; Bogdahn, U.; Bruun, T.-H. *Int. J. Mol. Sci.* **2020,** *21,* 1952. doi:10.3390/ijms21061952
143. van Ravesteyn, T. W.; Dekker, M.; Fish, A.; Sixma, T. K.; Wolters, A.; Dekker, R. J.; te Riele, H. P. J. *Proc. Natl. Acad. Sci. U. S. A.* **2016,** *113,* 4122–4127. doi:10.1073/pnas.1513315113
144. Ju, E.; Li, T.; Liu, Z.; da Silva, S. R.; Wei, S.; Zhang, X.; Wang, X.; Gao, S.-J. *ACS Nano* **2020,** *14,* 476–487. doi:10.1021/acsnano.9b06333
145. Fluiter, K.; Frieden, M.; Vreijling, J.; Rosenbohm, C.; De Wissel, M. B.; Christensen, S. M.; Koch, T.; Ørum, H.; Baas, F. *ChemBioChem* **2005,** *6,* 1104–1109. doi:10.1002/cbic.200400419
146. Rajwanshi, V. K.; Håkansson, A. E.; Dahl, B. M.; Wengel, J. *Chem. Commun.* **1999,** 1395–1396. doi:10.1039/a903189h
147. Nielsen, K. M. E.; Petersen, M.; Håkansson, A. E.; Wengel, J.; Jacobsen, J. P. *Chem. – Eur. J.* **2002,** *8,* 3001–3009. doi:10.1002/1521-3765(20020703)8:13<3001::aid-chem3001>3.0.co;2-1
148. Jørgensen, A. S.; Hansen, L. H.; Vester, B.; Wengel, J. *Bioorg. Med. Chem. Lett.* **2014,** *24,* 2273–2277. doi:10.1016/j.bmcl.2014.03.082
149. Raguraman, P.; Wang, T.; Ma, L.; Jørgensen, P. T.; Wengel, J.; Veedu, R. N. *Int. J. Mol. Sci.* **2020,** *21,* 2434. doi:10.3390/ijms21072434

150. Singh, S. K.; Kumar, R.; Wengel, J. *J. Org. Chem.* **1998,** *63,* 10035–10039. doi:10.1021/jo9814445
151. Astakhova, I. K.; Wengel, J. *Acc. Chem. Res.* **2014,** *47,* 1768–1777. doi:10.1021/ar500014g
152. Rosenbohm, C.; Christensen, S. M.; Sørensen, M. D.; Pedersen, D. S.; Larsen, L.-E.; Wengel, J.; Koch, T. *Org. Biomol. Chem.* **2003,** *1,* 655–663. doi:10.1039/b208864a
153. Sørensen, M. D.; Petersen, M.; Wengel, J. *Chem. Commun.* **2003,** 2130–2131. doi:10.1039/b307026c
154. Hrdlicka, P. J.; Babu, B. R.; Sørensen, M. D.; Wengel, J. *Chem. Commun.* **2004,** 1478–1479. doi:10.1039/b404446k
155. Valeur, E.; Bradley, M. *Chem. Soc. Rev.* **2009,** *38,* 606–631. doi:10.1039/b701677h
156. Baxter, E. W.; Reitz, A. B. Reductive Aminations of Carbonyl Compounds with Borohydride and Borane Reducing Agents. *Organic Reactions;* John Wiley & Sons, 2004; pp 1–714. doi:10.1002/0471264180.or059.01
157. Obika, S.; Onoda, M.; Morita, K.; Andoh, J.-i.; Koizumi, M.; Imanishi, T. *Chem. Commun.* **2001,** 1992–1993. doi:10.1039/b105640a
158. Morihiro, K.; Kodama, T.; Kentefu; Moai, Y.; Veedu, R. N.; Obika, S. *Angew. Chem., Int. Ed.* **2013,** *52,* 5074–5078. doi:10.1002/anie.201300555
159. Moai, Y.; Hiroaki, H.; Obika, S.; Kodama, T. *Nucleosides, Nucleotides Nucleic Acids* **2020,** *39,* 131–140. doi:10.1080/15257770.2019.1675169
160. Xu, J.; Liu, Y.; Dupouy, C.; Chattopadhyaya, J. *J. Org. Chem.* **2009,** *74,* 6534–6554. doi:10.1021/jo901009w
161. Egli, M.; Minasov, G.; Teplova, M.; Kumar, R.; Wengel, J. *Chem. Commun.* **2001,** 651–652. doi:10.1039/b009447l
162. Seth, P. P.; Siwkowski, A.; Allerson, C. R.; Vasquez, G.; Lee, S.; Prakash, T. P.; Kinberger, G.; Migawa, M. T.; Gaus, H.; Bhat, B.; Swayze, E. E. *Nucleic Acids Symp. Ser.* **2008,** *52,* 553–554. doi:10.1093/nass/nrn280
163. Seth, P. P.; Vasquez, G.; Allerson, C. A.; Berdeja, A.; Gaus, H.; Kinberger, G. A.; Prakash, T. P.; Migawa, M. T.; Bhat, B.; Swayze, E. E. *J. Org. Chem.* **2010,** *75,* 1569–1581. doi:10.1021/jo902560f
164. Blade, H.; Bradley, D.; Diorazio, L.; Evans, T.; Hayter, B. R.; Howell, G. P. *J. Org. Chem.* **2015,** *80,* 5337–5343. doi:10.1021/acs.joc.5b00607
165. Hong, D.; Kurzrock, R.; Kim, Y.; Woessner, R.; Younes, A.; Nemunaitis, J.; Fowler, N.; Zhou, T.; Schmidt, J.; Jo, M.; Lee, S. J.; Yamashita, M.; Hughes, S. G.; Fayad, L.; Piha-Paul, S.; Nadella, M. V. P.; Mohseni, M.; Lawson, D.; Reimer, C.; Blakey, D. C.; Xiao, X.; Hsu, J.; Revenko, A.; Monia, B. P.; MacLeod, A. R. *Sci. Transl. Med.* **2015,** *7,* 314ra185. doi:10.1126/scitranslmed.aac5272
166. Barman, J.; Gurav, D.; Oommen, O. P.; Varghese, O. P. *RSC Adv.* **2015,** *5,* 12257–12260. doi:10.1039/c4ra14721a
167. Mitsuoka, Y.; Aoyama, H.; Kugimiya, A.; Fujimura, Y.; Yamamoto, T.; Waki, R.; Wada, F.; Tahara, S.; Sawamura, M.; Noda, M.; Hari, Y.; Obika, S. *Org. Biomol. Chem.* **2016,** *14,* 6531–6538. doi:10.1039/c6ob01051b
168. Vejlegaard, K.; Wegeberg, C.; McKee, V.; Wengel, J. *Org. Biomol. Chem.* **2018,** *16,* 1312–1321. doi:10.1039/c7ob02663c
169. Seth, P. P.; Allerson, C. R.; Berdeja, A.; Siwkowski, A.; Pallan, P. S.; Gaus, H.; Prakash, T. P.; Watt, A. T.; Egli, M.; Swayze, E. E. *J. Am. Chem. Soc.* **2010,** *132,* 14942–14950. doi:10.1021/ja105875e
170. Ito, Y.; Tsutsui, N.; Osawa, T.; Hari, Y. *J. Org. Chem.* **2019,** *84,* 9093–9100. doi:10.1021/acs.joc.9b01035
171. Kumar, P.; El-Sagheer, A. H.; Truong, L.; Brown, T. *Chem. Commun.* **2017,** *53,* 8910–8913. doi:10.1039/c7cc05159j
172. Sharma, V. K.; Singh, S. K.; Krishnamurthy, P. M.; Alterman, J. F.; Haraszti, R. A.; Khvorova, A.; Prasad, A. K.; Watts, J. K. *Chem. Commun.* **2017,** *53,* 8906–8909. doi:10.1039/c7cc04092j
173. Seley-Radtke, K. L.; Yates, M. K. *Antiviral Res.* **2018,** *154,* 66–86. doi:10.1016/j.antiviral.2018.04.004
174. Fox, J. J. *Pure Appl. Chem.* **1969,** *18,* 223–256. doi:10.1351/pac196918010223
175. Glaudemans, C. P. J.; Fletcher, H. G., Jr. *J. Org. Chem.* **1963,** *28,* 3004–3006. doi:10.1021/jo01046a016
176. Tann, C. H.; Brodfuehrer, P. R.; Brundidge, S. P.; Sapino, C., Jr.; Howell, H. G. *J. Org. Chem.* **1985,** *50,* 3644–3647. doi:10.1021/jo00219a048
177. Howell, H. G.; Brodfuehrer, P. R.; Brundidge, S. P.; Benigni, D. A.; Sapino, C., Jr. *J. Org. Chem.* **1988,** *53,* 85–88. doi:10.1021/jo00236a017
178. Wilds, C. J.; Damha, M. J. *Nucleic Acids Res.* **2000,** *28,* 3625–3635. doi:10.1093/nar/28.18.3625
179. Elzagheid, M. I.; Viazovkina, E.; Damha, M. J. *Curr. Protoc. Nucleic Acid Chem.* **2002,** *10,* 1.7.1–1.7.19. doi:10.1002/0471142700.nc0107s10
180. Elzagheid, M. I.; Viazovkina, E.; Damha, M. J. *Nucleosides, Nucleotides Nucleic Acids* **2003,** *22,* 1339–1342. doi:10.1081/ncn-120022960
181. Damha, M. J.; Usman, N.; Ogilvie, K. K. *Can. J. Chem.* **1989,** *67,* 831–839. doi:10.1139/v89-129
182. Giannaris, P. A.; Damha, M. J. *Can. J. Chem.* **1994,** *72,* 909–918. doi:10.1139/v94-118
183. Noronha, A. M.; Wilds, C. J.; Lok, C.-N.; Viazovkina, K.; Arion, D.; Parniak, M. A.; Damha, M. J. *Biochemistry* **2000,** *39,* 7050–7062. doi:10.1021/bi000280v
184. Viazovkina, E.; Mangos, M. M.; Elzagheid, M. I.; Damha, M. J. *Curr. Protoc. Nucleic Acid Chem.* **2002,** *10,* 4.15.1–4.15.22. doi:10.1002/0471142700.nc0415s10
185. Kawasaki, A. M.; Casper, M. D.; Freier, S. M.; Lesnik, E. A.; Zounes, M. C.; Cummins, L. L.; Gonzalez, C.; Cook, P. D. *J. Med. Chem.* **1993,** *36,* 831–841. doi:10.1021/jm00059a007
186. Damha, M. J.; Wilds, C. J.; Noronha, A.; Brukner, I.; Borkow, G.; Arion, D.; Parniak, M. A. *J. Am. Chem. Soc.* **1998,** *120,* 12976–12977. doi:10.1021/ja982325+
187. Lok, C.-N.; Viazovkina, E.; Min, K.-L.; Nagy, E.; Wilds, C. J.; Damha, M. J.; Parniak, M. A. *Biochemistry* **2002,** *41,* 3457–3467. doi:10.1021/bi0115075
188. Kalota, A.; Karabon, L.; Swider, C. R.; Viazovkina, E.; Elzagheid, M.; Damha, M. J.; Gewirtz, A. M. *Nucleic Acids Res.* **2006,** *34,* 451–461. doi:10.1093/nar/gkj455
189. Berger, I.; Tereshko, V.; Ikeda, H.; Marquez, V. E.; Egli, M. *Nucleic Acids Res.* **1998,** *26,* 2473–2480. doi:10.1093/nar/26.10.2473
190. Li, F.; Sarkhel, S.; Wilds, C. J.; Wawrzak, Z.; Prakash, T. P.; Manoharan, M.; Egli, M. *Biochemistry* **2006,** *45,* 4141–4152. doi:10.1021/bi052322r
191. Trempe, J.-F.; Wilds, C. J.; Denisov, A. Y.; Pon, R. T.; Damha, M. J.; Gehring, K. *J. Am. Chem. Soc.* **2001,** *123,* 4896–4903. doi:10.1021/ja003859p
192. Denisov, A. Yu.; Noronha, A. M.; Wilds, C. J.; Trempe, J.-F.; Pon, R. T.; Gehring, K.; Damha, M. J. *Nucleic Acids Res.* **2001,** *29,* 4284–4293. doi:10.1093/nar/29.21.4284

193. Salazar, M.; Fedoroff, O. Y.; Miller, J. M.; Ribeiro, N. S.; Reid, B. R. *Biochemistry* **1993,** *32,* 4207–4215. doi:10.1021/bi00067a007
194. Fedoroff, O. Y.; Salazar, M.; Reid, B. R. *J. Mol. Biol.* **1993,** *233,* 509–523. doi:10.1006/jmbi.1993.1528
195. Watts, J. K.; Martín-Pintado, N.; Gómez-Pinto, I.; Schwartzentruber, J.; Portella, G.; Orozco, M.; González, C.; Damha, M. J. *Nucleic Acids Res.* **2010,** *38,* 2498–2511. doi:10.1093/nar/gkp1225
196. Martín-Pintado, N.; Yahyaee-Anzahaee, M.; Campos-Olivas, R.; Noronha, A. M.; Wilds, C. J.; Damha, M. J.; González, C. *Nucleic Acids Res.* **2012,** *40,* 9329–9339. doi:10.1093/nar/gks672
197. Martin-Pintado, N.; Deleavey, G. F.; Portella, G.; Campos-Olivas, R.; Orozco, M.; Damha, M. J.; González, C. *Angew. Chem., Int. Ed.* **2013,** *52,* 12065–12068. doi:10.1002/anie.201305710
198. Dowler, T.; Bergeron, D.; Tedeschi, A.-L.; Paquet, L.; Ferrari, N.; Damha, M. J. *Nucleic Acids Res.* **2006,** *34,* 1669–1675. doi:10.1093/nar/gkl033
199. Kel′in, A. V.; Zlatev, I.; Harp, J.; Jayaraman, M.; Bisbe, A.; O′Shea, J.; Taneja, N.; Manoharan, R. M.; Khan, S.; Charisse, K.; Maier, M. A.; Egli, M.; Rajeev, K. G.; Manoharan, M. *J. Org. Chem.* **2016,** *81,* 2261–2279. doi:10.1021/acs.joc.5b02375
200. Harp, J. M.; Guenther, D. C.; Bisbe, A.; Perkins, L.; Matsuda, S.; Bommineni, G. R.; Zlatev, I.; Foster, D. J.; Taneja, N.; Charisse, K.; Maier, M. A.; Rajeev, K. G.; Manoharan, M.; Egli, M. *Nucleic Acids Res.* **2018,** *46,* 8090–8104. doi:10.1093/nar/gky703
201. Liboska, R.; Snášel, J.; Barvík, I.; Buděšínský, M.; Pohl, R.; Točík, Z.; Páv, O.; Rejman, D.; Novák, P.; Rosenberg, I. *Org. Biomol. Chem.* **2011,** *9,* 8261–8267. doi:10.1039/c1ob06148h
202. Petrová, M.; Páv, O.; Buděšínský, M.; Zborníková, E.; Novák, P.; Rosenbergová, Š.; Pačes, O.; Liboska, R.; Dvořáková, I.; Šimák, O.; Rosenberg, I. *Org. Lett.* **2015,** *17,* 3426–3429. doi:10.1021/acs.orglett.5b01430
203. Jenkins, I. D.; Verheyden, J. P. H.; Moffatt, J. G. *J. Am. Chem. Soc.* **1976,** *98,* 3346–3357. doi:10.1021/ja00427a049
204. Owen, G. R.; Verheyden, J. P. H.; Moffatt, J. G. *J. Org. Chem.* **1976,** *41,* 3010–3017. doi:10.1021/jo00880a018
205. Thomas, S. O.; Singleton, V. L.; Lowery, J. A.; Sharpe, R. W.; Pruess, L. M.; Porter, J. N.; Mowat, J. H.; Bohonos, N. *Antibiot. Annu.* **1956,** 716–721.
206. Martínez-Montero, S.; Deleavey, G. F.; Kulkarni, A.; Martín-Pintado, N.; Lindovska, P.; Thomson, M.; González, C.; Götte, M.; Damha, M. J. *J. Org. Chem.* **2014,** *79,* 5627–5635. doi:10.1021/jo500794v
207. Martínez-Montero, S.; Deleavey, G. F.; Martín-Pintado, N.; Fakhoury, J. F.; González, C.; Damha, M. J. *ACS Chem. Biol.* **2015,** *10,* 2016–2023. doi:10.1021/acschembio.5b00218
208. Martínez-Montero, S.; Deleavey, G. F.; Dierker-Viik, A.; Lindovska, P.; Ilina, T.; Portella, G.; Orozco, M.; Parniak, M. A.; González, C.; Damha, M. J. *J. Org. Chem.* **2015,** *80,* 3083–3091. doi:10.1021/jo502948t
209. Malek-Adamian, E.; Patrascu, M. B.; Jana, S. K.; Martínez-Montero, S.; Moitessier, N.; Damha, M. J. *J. Org. Chem.* **2018,** *83,* 9839–9849. doi:10.1021/acs.joc.8b01329
210. Malek-Adamian, E.; Guenther, D. C.; Matsuda, S.; Martínez-Montero, S.; Zlatev, I.; Harp, J.; Burai Patrascu, M.; Foster, D. J.; Fakhoury, J.; Perkins, L.; Moitessier, N.; Manoharan, R. M.; Taneja, N.; Bisbe, A.; Charisse, K.; Maier, M.; Rajeev, K. G.; Egli, M.; Manoharan, M.; Damha, M. J. *J. Am. Chem. Soc.* **2017,** *139,* 14542–14555. doi:10.1021/jacs.7b07582
211. Li, Q.; Chen, J.; Trajkovski, M.; Zhou, Y.; Fan, C.; Lu, K.; Tang, P.; Su, X.; Plavec, J.; Xi, Z.; Zhou, C. *J. Am. Chem. Soc.* **2020,** *142,* 4739–4748. doi:10.1021/jacs.9b13207
212. Gore, K. R.; Nawale, G. N.; Harikrishna, S.; Chittoor, V. G.; Pandey, S. K.; Höbartner, C.; Patankar, S.; Pradeepkumar, P. I. *J. Org. Chem.* **2012,** *77,* 3233–3245. doi:10.1021/jo202666m
213. Koizumi, K.; Maeda, Y.; Kano, T.; Yoshida, H.; Sakamoto, T.; Yamagishi, K.; Ueno, Y. *Bioorg. Med. Chem.* **2018,** *26,* 3521–3534. doi:10.1016/j.bmc.2018.05.025
214. Kano, T.; Katsuragi, Y.; Maeda, Y.; Ueno, Y. *Bioorg. Med. Chem.* **2018,** *26,* 4574–4582. doi:10.1016/j.bmc.2018.08.001
215. Nawale, G. N.; Bahadorikhalili, S.; Sengupta, P.; Kadekar, S.; Chatterjee, S.; Varghese, O. P. *Chem. Commun.* **2019,** *55,* 9112–9115. doi:10.1039/c9cc04141a
216. Herdewijn, P. *Chem. Biodiversity* **2010,** *7,* 1–59. doi:10.1002/cbdv.200900185
217. Aerschot Van, A.; Verheggen, I.; Hendrix, C.; Herdewijn, P. *Angew. Chem., Int. Ed. Engl.* **1995,** *34,* 1338–1339. doi:10.1002/anie.199513381
218. Hendrix, C.; Rosemeyer, H.; Verheggen, I.; Van Aerschot, A.; Seela, F.; Herdewijn, P. *Chem. – Eur. J.* **1997,** *3,* 110–120. doi:10.1002/chem.19970030118
219. Hossain, N.; Wroblowski, B.; Van Aerschot, A.; Rozenski, J.; De Bruyn, A.; Herdewijn, P. *J. Org. Chem.* **1998,** *63,* 1574–1582. doi:10.1021/jo9718511
220. Allart, B.; Khan, K.; Rosemeyer, H.; Schepers, G.; Hendrix, C.; Rothenbacher, K.; Seela, F.; Van Aerschot, A.; Herdewijn, P. *Chem. – Eur. J.* **1999,** *5,* 2424–2431. doi:10.1002/(sici)1521-3765(19990802)5:8<2424::aid-chem2424>3.0.co;2-w
221. Egli, M.; Pallan, P. S.; Allerson, C. R.; Prakash, T. P.; Berdeja, A.; Yu, J.; Lee, S.; Watt, A.; Gaus, H.; Bhat, B.; Swayze, E. E.; Seth, P. P. *J. Am. Chem. Soc.* **2011,** *133,* 16642–16649. doi:10.1021/ja207086x
222. Winter, H. D.; Lescrinier, E.; Aerschot, A. V.; Herdewijn, P. *J. Am. Chem. Soc.* **1998,** *120,* 5381–5394. doi:10.1021/ja973721f
223. Vandermeeren, M.; Préveral, S.; Janssens, S.; Geysen, J.; Saison-Behmoaras, E.; Van Aerschot, A.; Herdewijn, P. *Biochem. Pharmacol.* **2000,** *59,* 655–663. doi:10.1016/s0006-2952(99)00367-6
224. Allerson, C. R.; Sioufi, N.; Jarres, R.; Prakash, T. P.; Naik, N.; Berdeja, A.; Wanders, L.; Griffey, R. H.; Swayze, E. E.; Bhat, B. *J. Med. Chem.* **2005,** *48,* 901–904. doi:10.1021/jm049167j
225. Davis, S.; Propp, S.; Freier, S. M.; Jones, L. E.; Serra, M. J.; Kinberger, G.; Bhat, B.; Swayze, E. E.; Bennett, C. F.; Esau, C. *Nucleic Acids Res.* **2009,** *37,* 70–77. doi:10.1093/nar/gkn904
226. Prakash, T. P.; Yu, J.; Kinberger, G. A.; Low, A.; Jackson, M.; Rigo, F.; Swayze, E. E.; Seth, P. P. *Bioorg. Med. Chem. Lett.* **2018,** *28,* 3774–3779. doi:10.1016/j.bmcl.2018.10.011
227. Pallan, P. S.; Yu, J.; Allerson, C. R.; Swayze, E. E.; Seth, P.; Egli, M. *Biochemistry* **2012,** *51,* 7–9. doi:10.1021/bi201810r
228. Seth, P. P.; Allerson, C. R.; Østergaard, M. E.; Swayze, E. E. *Bioorg. Med. Chem. Lett.* **2012,** *22,* 296–299. doi:10.1016/j.bmcl.2011.11.012
229. De, S.; Jabgunde, A. M.; Patil, R. S.; De Jonghe, S.; Beigelman, L.; Herdewijn, P. *J. Org. Chem.* **2018,** *83,* 15155–15169. doi:10.1021/acs.joc.8b02444
230. Jabgunde, A. M.; Patil, R. S.; De, S.; Lescrinier, E.; De Jonghe, S.; Beigelman, L.; Herdewijn, P. *Tetrahedron* **2019,** *75,* 1107–1114. doi:10.1016/j.tet.2019.01.026

231. Schweitzer, B. A.; Kool, E. T. *J. Org. Chem.* **1994,** *59,* 7238–7242. doi:10.1021/jo00103a013
232. Moran, S.; Ren, R. X.-F.; Rumney, S., IV; Kool, E. T. *J. Am. Chem. Soc.* **1997,** *119,* 2056–2057. doi:10.1021/ja963718g
233. Khakshoor, O.; Wheeler, S. E.; Houk, K. N.; Kool, E. T. *J. Am. Chem. Soc.* **2012,** *134,* 3154–3163. doi:10.1021/ja210475a
234. Kellinger, M. W.; Ulrich, S.; Chong, J.; Kool, E. T.; Wang, D. *J. Am. Chem. Soc.* **2012,** *134,* 8231–8240. doi:10.1021/ja302077d
235. Xia, J.; Noronha, A.; Toudjarska, I.; Li, F.; Akinc, A.; Braich, R.; Frank-Kamenetsky, M.; Rajeev, K. G.; Egli, M.; Manoharan, M. *ACS Chem. Biol.* **2006,** *1,* 176–183. doi:10.1021/cb600063p
236. Ulrich, S.; Kool, E. T. *Biochemistry* **2011,** *50,* 10343–10349. doi:10.1021/bi2011465
237. Li, F.; Pallan, P. S.; Maier, M. A.; Rajeev, K. G.; Mathieu, S. L.; Kreutz, C.; Fan, Y.; Sanghvi, J.; Micura, R.; Rozners, E.; Manoharan, M.; Egli, M. *Nucleic Acids Res.* **2007,** *35,* 6424–6438. doi:10.1093/nar/gkm664
238. Egli, M.; Pallan, P. S. *Annu. Rev. Biophys. Biomol. Struct.* **2007,** *36,* 281–305. doi:10.1146/annurev.biophys.36.040306.132556
239. Egli, M. *Acc. Chem. Res.* **2012,** *45,* 1237–1246. doi:10.1021/ar200303k
240. Pallan, P. S.; Egli, M. *J. Am. Chem. Soc.* **2009,** *131,* 12548–12549. doi:10.1021/ja905739j

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Antiviral therapy in shrimp through plant virus VLP containing VP28 dsRNA against WSSV

Santiago Ramos-Carreño[1], Ivone Giffard-Mena[*1], Jose N. Zamudio-Ocadiz[2,3], Alfredo Nuñez-Rivera[2,3], Ricardo Valencia-Yañez[1], Jaime Ruiz-Garcia[4], Maria Teresa Viana[5] and Ruben D. Cadena-Nava[*2]

Full Research Paper

Open Access

Address:
[1]Facultad de Ciencias Marinas, Universidad Autónoma de Baja California (UABC), Carretera Transpeninsular Ensenada-Tijuana No. 3917, Colonia Playitas, C.P. 22860 Ensenada, B.C., México, [2]Centro de Nanociencias y Nanotecnología, Universidad Nacional Autónoma de México (UNAM). Km 107 Carretera Tijuana-Ensenada, Col. Pedregal Playitas, C.P. 22860 Ensenada, B.C., México, [3]Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California, (CICESE), Carretera Ensenada - Tijuana No. 3918, Zona Playitas, C.P. 22860, Ensenada, B.C., México, [4]Instituto de Física, Universidad Autónoma de San Luis Potosí, Álvaro Obregón 64, San Luis Potosí 78000, México and [5]Instituto de Investigaciones Oceanológicas, Universidad Autónoma de Baja California (UABC), Carretera Transpeninsular Ensenada-Tijuana No. 3917, Colonia Playitas, C.P. 22860 Ensenada, B.C., México

Email:
Ivone Giffard-Mena[*] - igiffard@uabc.edu.mx;
Ruben D. Cadena-Nava[*] - rcadena74@gmail.com

* Corresponding author

Keywords:
antiviral therapy; CCMV; oral administration; *P. vannamei*; plant VLPs; RNAi; VP28; white spot syndrome virus

Beilstein J. Org. Chem. **2021,** *17,* 1360–1373.
https://doi.org/10.3762/bjoc.17.95

Received: 03 December 2020
Accepted: 20 May 2021
Published: 01 June 2021

This article is part of the thematic issue "Celebrating the role of chemistry in the success of oligonucleotides as therapeutics".

Guest Editors: P. Kumar and T. Brown

Abstract

The white spot syndrome virus (WSSV), currently affecting cultured shrimp, causes substantial economic losses to the worldwide shrimp industry. An antiviral therapy using double-stranded RNA interference (dsRNAi) by intramuscular injection (IM) has proven the most effective shrimp protection against WSSV. However, IM treatment is still not viable for shrimp farms. The challenge is to develop an efficient oral delivery system that manages to avoid the degradation of antiviral RNA molecules. The present work demonstrates that VLPs (virus-like particles) allow efficient delivery of dsRNAi as antiviral therapy in shrimp. In particular, VLPs derived from a virus that infects plants, such as cowpea chlorotic mottle virus (CCMV), in which the capsid protein (CP) encapsidates the dsRNA of 563 bp, are shown to silence the WSSV glycoprotein VP28 (dsRNAvp28). In experimental challenges in vivo, the VLPs- dsRNAvp28 protect shrimp against WSSV up to 40% by oral administration and 100% by IM. The novel research demonstrates that plant VLPs, which avoid zoonosis, can be applied to pathogen control in shrimp and also other organisms, widening the application window in nanomedicine.

Introduction

The white spot syndrome virus (WSSV) is recognized as one of the most severe epidemic pathogens of shrimp, causing severe economic losses to shrimp aquaculture. More than three decades ago Chou et al. [1] first described the emergence of this pathogen and since then, rapidly, it has spread globally [2,3]. The aquaculture industry still suffers productive and economic impacts from the outbreak, causing up to 100% mortality in shrimp farms within 3 to 10 days [1,4]. The rapid propagation and susceptibility of WSSV infection in several species, particularly the white shrimp *Penaeus vannamei* [5,6], have sparked intense research for its prevention and control [7].

So far several strategies have been reported to control the WSSV, including activation of the immune system, DNA vaccines, herbal extracts, and RNA interference (RNAi) [8,9]. Among them, the RNAi technology has shown great potential to protect shrimp against the WSSV in some lab-scale experiments [10,11]. The RNAi mechanism comprises a set of cellular processes of posttranscriptional gene silencing that begins with administering the double-stranded RNA (dsRNA). It concludes with a specific gene silencing based on sequence homology between the digested fragments of the dsRNA and the gene of interest [12-16]. The antiviral response of RNAi is triggered by double-stranded RNA (dsRNA) to block the synthesis of a specific viral protein, in the case of WSSV the targets being the structural proteins VP19, VP24, VP26, and VP28, as they are involved in cell recognition, virus entry, binding and assembly of the virion. Previous studies have shown that silencing these structural proteins in WSSV challenge assays, increases shrimp survival [10,11,17-21]. The VP28 glycoprotein plays an important role in systemic infection by interacting with cell membrane proteins, and it is one of the most abundant proteins along with VP26 (≈60%) in the external WSSV surface [21,22].

The RNAi trials using an intramuscular injection (IM) have shown that VP28 glycoprotein is the target of choice to block WSSV infection in shrimp [14,23,24]. However, RNAi intramuscular (IM) administration is limited to lab-scale experiments since its use is not yet viable for applications on a large scale, as found in salmon farms [25]. The naked RNA degrades quickly when supplied in feed [26,27], either due to feed processing or the digestion process [20]. The challenge is to develop a treatment through the oral route [11,28] instead of IM, yet one in which the RNA is nonetheless is protected.

One solution is a nanocarrier [11,27,29] to protect, stabilize and maintain the integrity of the RNAi in the environment [14]. Recently, dsRNA has been integrated into nanovehicles such as non-virulent capsids or virus-like particles (VLPs) [30-32] lacking the viral genome. Their small size (20–140 nm), allows them to permeate the cell membranes without causing toxicity or immune response in the treated organisms [30,32-36]. In particular, the VLPs derived from plant viruses are attractive, since any zoonotic possibility is eliminated, being biocompatible and biodegradable [34,36,37]. Its structure presents advantages over other synthetic nanomaterials, as it is simple and easy to purify for large scale production [34,37,38].

The plant virus cowpea chlorotic mottle virus (CCMV) has been extensively studied and characterized, due to its potential applications in nanomedicine [33,36,39-41]. Native CCMV has a positive single-stranded RNA. It is a Bromoviridae family member that infects cowpea *(Vigna unguiculata)* plants. The CCMV VLPs with heterologous RNA has already been in-vitro synthesized [32,42], being RNases resistant, and can release cargo in the cytoplasm of mammalian cells [32,33,43].

This work aims to evaluate the efficacy of CCMV VLP-VP28 dsRNA (VLP-dsRNAvp28) delivery against WSSV, by oral administration to shrimp through commercial feed pellets. Through in vivo bioassays, the antiviral efficacy of VLPs is assessed by intramuscular injection and per os, in *Penaeus vannamei* infected with WSSV.

To our knowledge, this is the first report where an oral VLPs are administered to treat infected shrimp against viruses. This is a novel technique in aquaculture.

Materials and Methods

dsRNAvp28. The VP28 dsRNA (dsRNAvp28) was generated based on the VP28 sequence of WSSV (GenBank: EU931451.1) [44]. The sequence is shown in Supporting Information File 1, Table ST1. The dsRNAvp28 was purchased from groRNA/Genolution company (South Korea).

CCMV capsid protein purification. The plant virus CCMV was produced in California cowpea plants (*Vigna ungiculata*). The plants were mechanically inoculated with a solution containing the virus. After two weeks, the infected leaves were collected and ground in a virus extraction buffer (0.5 M sodium acetate, 0.08 M magnesium acetate, pH 4.5) using a kitchen blender. The obtained homogeneous extract was filtered through a cheesecloth to remove solid material. Then the homogenate was mixed with a half-volume of chloroform and centrifuged at 15,000 rpm for 15 min using a JA-14 rotor (Beckman Coulter, USA). After that, the supernatant was recovered and stirred for at least 3 h. Then the sample was layered on a 10% sucrose cushion and ultracentrifuged for 2 hours at 30,000 rpm using an SW-32Ti rotor (Beckman Coulter, USA).

Later, the supernatant was discarded and the pellets were resuspended with a virus suspension buffer (50 mM sodium acetate, 8 mM magnesium acetate, pH 4.5). The solution was ultracentrifuged through a sucrose gradient at 30,000 rpm for 2 hours, at 4 °C. The virus was recovered from the blue band, and the sucrose was removed by ultracentrifugation. The pellets were resuspended in virus suspension buffer (50 mM sodium acetate, 8 mM magnesium acetate, pH 4.5). All the procedure was done at 4 °C. The virus's concentration and purity were determined by UV–vis spectrophotometry, and the virus aliquots were kept at −80 °C.

The protein purification was performed according to a previously described protocol [40]. Briefly, the CCMV was dialyzed in a disassembly buffer (0.5 M $CaCl_2$, 50 mM Tris, 1 mM EDTA, 1 mM DTT, 0.5 mM PMSF, pH 7.5) at 4 °C for 24 h. Then, the sample was ultracentrifuged at 50,000 rpm for 510 min at 4 °C, using a Beckman Type 90 Ti rotor. The pellet was discarded, and the supernatant containing the capsid protein (CP) was recovered. Later, the CP was dialyzed against a buffer (1 M NaCl, 20 mM Tris, 1 mM EDTA, pH 7.2) overnight. The protein concentration and purity were determined by UV–vis spectrophotometry; only CP samples with the wavelength ratio 280/260 ≥ 1.5 were used for the VLPs assembling. SDS-PAGE was used to verify the integrity of the capsid protein.

In vitro assembly of VLPs-dsRNAvp28. Dissociated CCMV CP and dsRNAvp28 were mixed in a mass ratio of 6:1 (CP/dsRNA) and dialyzed overnight against RNA assembly buffer (50 mM NaCl, 10 mM KCl, 5 mM $MgCl_2$, 50 mM Tris-HCl, pH 7.2) at 4 °C. The samples were acidified by dialysis in virus suspension buffer (50 mM sodium acetate, 8 mM magnesium acetate, pH 4.5) for at least 4 hours. Then, to disrupt the empty capsids, the sample was dialyzed against an RNA assembly buffer. The VLP-dsRNAvp28 was then purified and concentrated by ultrafiltration with reassembly buffer using a 100 kDa Amicon centrifuge filter (0.5 mL, Millipore) at 8,000*g* for 15 min, and the step was repeated at least three times.

VLPs assembly products were analyzed by gel electrophoresis mobility shift assay (EMSA) in native agarose gel at 1%. The electrophoresis was run in a horizontal agarose gel system (FBSB710 Fisher Scientific) for 4 h at 50 volts (virus buffer), 4 °C and then, the gel was stained with ethidium bromide. The image was captured using a documentation system (MS Major Science).

Transmission electron microscopy analysis of VLPs. 6 µL of VLP-dsRNAvp28 from the assembly stock solution was placed onto a carbon-coated grid (400 mesh Cu, Ted Pella) for 2.5 min. The excess solution was removed with a Whatman filter paper, and the sample was stained with 6 µL of 2% uranyl acetate for 1 min. The samples were analyzed with a JEOL JEM-2010 transmission electron microscope equipped with a digital camera operated at 200 keV. The size of the VLPs was measured using the ImageJ (U.S. NIH) software from digital recorded TEM images.

Shrimp and rearing conditions. *P. vannamei* postlarvae (PL) were grown in 2,500 L circular tanks containing seawater (34 ppt salinity) at 28 ± 1 °C, oxygen > 5.0 mg/L, pH 7.6 ± 0.16 and ammonium < 0.5 mg/L. The postlarvae were fed a commercial diet (Natural Force 35® VIMIFOS, Mexico) at 5% of the total biomass thrice a day. The seawater was filtered (10.5 and 5 µm sediment water filters, respectively), exposed to UV and aerated before use. Forty percent of water was replaced every three days to collect food waste and feces.

Once the PL reached a juvenile stage, they were transferred into 12 L aquariums. Each aquarium was equipped with a filter and a heating system (Titanium Heater HMO-200, JSK). The shrimp were immersed in a 0.002% formaldehyde solution in seawater for 30 min before transferring them to the aquariums to remove any fouling present. Six shrimp were placed per aquarium, containing seawater of 34 ppt at 28 ± 0.3 °C, oxygen between 5.0 to 8.0 mg / L, pH 7.6 ± 0.16 and total ammonium < 0.5 mg/L. A photoperiod of 12 h light and 12 h dark was used. The shrimp were fed with a commercial diet twice a day at 3% of their biomass. Shrimp were gradually acclimatized to 16 ppt and kept 15 days in observation before starting the experiment. Filters containing activated carbon were used to maintain an optimum seawater quality. Sixty percent of the water was replaced daily. At the end of the bioassay all materials were disinfected using granulated calcium hypochlorite at 1600 ppm and neutralized with sodium thiosulfate (Brenntag pacific Inc. Santa Fe Springs, CA 90670) at 872 ppm. The Infectious waste was sterilized before disposal.

WSSV inoculum preparation. The isolate of WSSV was used from a disease outbreak from Sonora, Mexico in 2008 (Son2008). The viral inoculum was prepared from frozen samples (−80 °C) of dying shrimp with WSSV positive diagnostic [45,46]. For this, 100 mg of gills from four individuals (25 mg each) were homogenized in 900 µL (1:10 ratio; mg/µL) of TN buffer (20 mM Tris-HCl, 400 mM NaCl, pH 7.4). The homogenized solution was centrifuged in two steps at 1800 and 3000*g* for 20 min, respectively, at 4 °C. The supernatant was recovered and filtered through a membrane filter (0.45 µm VWR®, Europe) [47]. This inoculum solution is referred to as the 1:10 dilution. The in vivo experiments were immediately initiated

after preparing the inoculums. Simultaneously, uninfected shrimp or free WSSV were parallel-used under the same procedure as a negative control (WSSV-negative).

Viral inoculum activation. Two groups of 15 shrimp were inoculated with the solution obtained from infected shrimp as previously described. Then, the shrimp were transferred into 60 L rectangular aquariums. A third group (n = 15) was used as a control. Shrimp inoculation was performed by intramuscular injection (IM), using a 0.5 mL insulin syringe (BD Micro-FineTM) (31G × 6 mm), injecting 20 µL of 10^{-1} viral inoculum (original stock 1:10 p/v) to each shrimp in the fifth abdominal segment, whereas for the control group a TN sterilized buffer (20 mM Tris/HCl, 400 mM NaCl, pH 7.4) was used. The shrimp were fed commercial pellets three times a day. Every four hours, moribund organisms were collected and euthanized using liquid nitrogen, and subsequently stored at −80 °C for further analysis. WSSV was confirmed by endpoint polymerase chain reaction (PCR), following Koch's postulates.

Minimum infectious dose determination. The IM minimum lethal dose of WSSV to generate mortality as per os infection was determined simultaneously. Three replicates per treatment were used with six organisms (3.6 g ± 0.66 g) per aquarium. Before the viral challenge, shrimp were acclimatized for seven days under similar conditions. Then, shrimp were injected with 20 µL of a 10-fold serial dilution (10^{-1},10^{-2}, 10^{-4}, 10^{-6}, 10^{-8}, 10^{-10}, 18 organisms per dilution) of WSSV inoculum (Son2008) stock 1:10 p/v. Shrimp were injected with virus-free gill homogenates, and TN buffer was used as control. The lethal dose 50% endpoint (LD_{50} mL^{-1}) was calculated using the formula: $\log_{10}$ 50% endpoint dilution = − [(total number of animals died/number of animals inoculated per dilution) + 0.5] × log dilution factor [48]. To establish the per os WSSV infection time, five replicate aquaria with five shrimp (3.6 g ± 0.66 g) per tank were orally challenged. Before the infection per os, fasted shrimp for 24 hours were fed twice a day with infected ground tissue (≈10 biomass) [46]. Six hours after the last dose, the unconsumed infected tissue was removed, and aquarium water was replaced, per Thomas et al. [49]. Mortality was recorded to register the dose effectiveness of the inoculum (infected tissue) [50]. All collected shrimp (alive, dying, or dead) were cryo-frozen in liquid nitrogen (LN_2) and stored at −80 °C for further analysis. All animal experimentation was supervised and authorized by the ethics committee of the institutional committee at UABC to comply with all the humanitarian protocols in handling to avoid animal suffering.

Optimal dose of dsRNAvp28. The optimal dose of the dsRNAvp28 (Genolution) was determined in a bioassay using different concentration doses. Five replicate aquaria with five juvenile shrimp (5.40 g ± 0.56 g) were used for the challenges. Organisms were acclimatized and fed as previously described. After seven days, 20 µL of WSSV inoculum (10^{-6} dilution) was applied (intramuscular injection) to each animal's left side, simultaneously on the right side dsRNAvp28 was injected in doses of 0.5, 1.0, 2.0, and 3.0 µg/shrimp per group. A positive WSSV infection control without dsRNAvp28 treatment and a WSSV-free group were then injected with healthy tissue homogenate (20 µL) and 3.0 µg of dsRNAvp28/shrimp were included; see Table ST2 in Supporting Information File 1, material section.

Administration of VLP-dsRNAvp28 by the oral cavity. The inhibition efficacy of dsRNAvp28 to WSSV by oral route was evaluated using free dsRNAvp28 and VLP-dsRNAvp28 administered directly into the shrimp's oral cavity. The procedure was standardized before the bioassay. In summary, 50 µL of TN solution containing 10% red food coloring (pigment red, McCormick4, USA) was administered through the oral cavity using an insulin syringe (BD Micro-FineTM) of 0.5 mL (31G × 6 mm). The distribution of the red-stained solution was observed with a stereoscopic microscope (Labomed, Model CZM6 Trinocular, Stereo Microscope) to determine the time and distribution of the product in the digestive tract of the shrimp.

After that, two sets of groups in four replicates with five shrimp each. In one of them, 6.0 µg (50 µL) of free dsRNAvp28 was administered, whereas in the second, 50 µL of VLP-dsRNAvp28 (6.0 µg of dsRNAvp28) was applied. After 18 hours both groups were challenged with WSSV by IM injection, with a dose of 10^{-6} Son2008 inoculum. (Herein "pellet feed" refers to when animals are fed with treatments, and "oral cavity" refers to when the VLPs treatment is given directly into the oral cavity through a needle to ensure intestinal functionality).

Feed pellets with VLP-dsRNAvp28. Two methods were used to prepare the pellet feed containing VLP-dsRNAvp28: first, coating the external surface of commercial pellets with the VLPs, and second, pulverizing the pellets, mixing the VLP's with them, and reconstituting them (The details are described in Supporting Information File 1). In all experiments, to follow the standard procedures in bioassays with shrimp, each treatment had at least three replicates [10]. The pellets with VLP-dsRNAvp28 were coated with industrial grade fish oil or salmon fish oil (see details in Supporting Information File 1). Pellets with VLP-dsRNAvp28 prepared with commercial binders (Dry Oil® and NutriKelp®) are described in Supporting Information File 1.

Detection of WSSV by real-time quantitative PCR. The real-time PCR (qPCR) for quantitation of WSSV was performed using DNA isolated from shrimp muscle tissue and TaqMan® Fast Advanced Master Mix kit (Applied Biosystems, USA). Amplification reactions of 20 µL were prepared by mixing 23.33 ng of DNA, 0.3 µM of each primer, and 0.15 µM of TaqMan probe, and the qPCR was performed following Durand and Lightner [51] methodology. In summary, 2.0 min at 50 °C for uracil-*N*-glycosylase (UNG) activation; 10 min at 95 °C to activate AmpliTaq Fast DNA Polymerase and then, 40 cycles of 15 seconds at 95 °C and 1 min at 60 °C.

For the WSSV quantification, a standard curve was obtained with the plasmid DNA with the *vp664* gene of 69 bp [45,51] at a 1:10 dilution factor. The concentration range of the standard curve was 3.9×10^9 to 3.9×10^4 copies/ng. The ABI StepOnePlus v2.0 sequence detection system software (Applied Biosystems, USA) was used. Amplification reactions included all shrimp were analyzed (alive, dying and dead) from each experimental group.

The viral load of WSSV obtained by qPCR from three independent experiments was analyzed by comparing the average number of copies (copies/ng) of two replicates from the same shrimp of each group (n = 4–9 samples), plus their confidence interval.

Statistical analysis. For each treatment, the protection against WSSV after feeding with the antiviral therapy was evaluated through the survival and mean lethal time (LT_{50}) [52]. A Log-Rank (Mantel–Cox) test was used to analyze the Kaplan–Meier survival curves generated with the GraphPad Prism version 5.01 software (San Diego California USA). In all cases, a value of $p < 0.05$ was considered significant. For the WSSV detection, an analysis of variance (ANOVA) was used to compare the average number of copies of WSSV and the average number of copies between treatments was analyzed with the Tukey's test (a = 0.05). The Student's t-test was used to obtain significant differences (t – 95%) between treatments (alive vs dying/dead).

Results

The dsRNA was efficiently encapsidated with CCMV CP using a mass ratio of 6:1 of CP/dsRNAvp28. The electrophoresis mobility shift assay (EMSA) of the assembly shows that most of the sample is close to the well, and a small sample portion migrated similarly to the wild type CCMV (lane 2 and 3, respectively, in Figure 1). In contrast, the free dsRNAvp28 ran faster (lane 4 in Figure 1) in comparison with the sample and wild type CCMV, as an indication of VLPs formation.

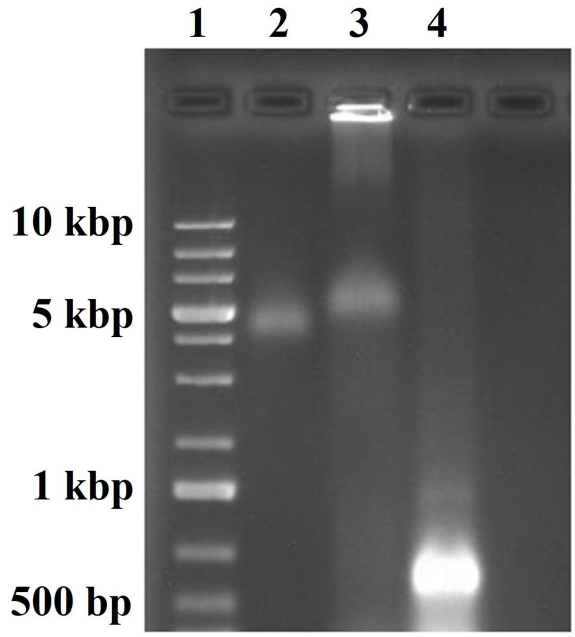


Figure 1: Analysis of the VLP-dsRNAvp28 assembly by electrophoresis mobility shift assay (EMSA) in a 1% native agarose gel. Lane 1 is the DNA ladder; 2: wild type CCMV; 3: self-assembly of dsRNA with CCMV CP; and lane 4: free dsRNA.

The VLPs assembly at each stage was analyzed by transmission electron microscopy (TEM). The procapsids and CP-dsRNA complexes obtained after dialysis in assembly buffer (pH 7.2) can be observed in images a and b (Figure 2A). After the second dialysis in virus buffer (pH 4.5), well-defined VLPs were formed (Figure 2A, c, and d). Finally, the dialyzed sample is shown in Figure 2A, e, and f. The morphology of the VLPs is maintained after this last step of the assembly process. The VLPs synthesized had two types of morphologies: icosahedral capsids and large rods (Figure 2A, c to f). Also, aggregations of spherical capsid can be observed at the last VLP assembly step. The distribution of the procapsids diameter, icosahedral VLPs, and nanotubes is shown in Figure 2B. According to the Gaussian fit for each of the distributions, the average diameter of the procapsids, icosahedral VLPs, and the rods were 21, 26, and 21 nm, respectively.

During the WSSV viral inoculum activation, the symptoms' onset times and mortality occurred between 18 and 22 hours post-infection (hpi) (Figure 3A). At 22 hpi, the first death was detected. The minimum survival rates at 29 hpi, for the first inoculum reactivation, and 44.5 for the second (referred to as 1-WSSV-2008 and 2-WSSV-2008, respectively) were recorded. After 53.5 hpi, both for 1-WSSV-2008 and 77 hpi for 2-WSSV-2008, all shrimp were dead. Similarly, all infected shrimp from the control groups (WSSV-Positive) for the different treatments

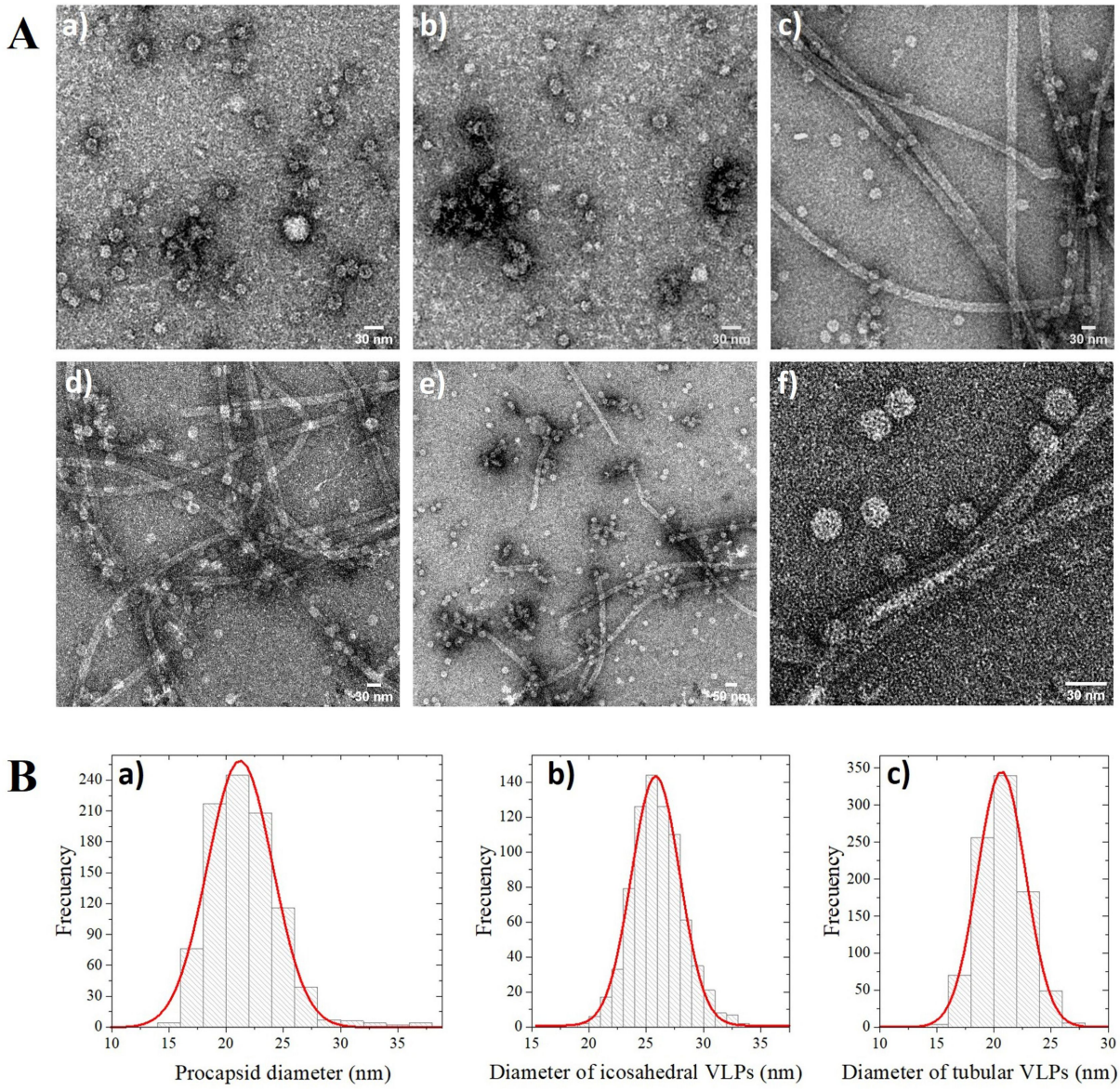


Figure 2: TEM micrographs of different stages of the assemblies of CCMV CP with dsRNAvp28. In section A, the images a) and b) correspond to the assembly in virus buffer; c) and d) are acidified assembly; e) and f) images correspond to the sample that was dialyzed again in assembly buffer. Section B shows the size distributions of the ensembles: a) diameter distribution of the procapsids with a mean diameter peak at 21 nm; b) diameters of the icosahedral VLP-dsRNAvp28 with a mean diameter peak at 26 nm and c) diameter of the tubular structures with a mean diameter peak at 21 nm.

were dead. In contrast, 100% survival was obtained for the WSSV-Negative control groups (WSSV free).

The minimum infectious dose of WSSV resulted in significant differences ($p < 0.001$). The dilutions 10^{-1} and 10^{-2} gave 0% survival at 56.2 and 57.3 hpi, respectively. Whereas shrimp inoculated with the dilutions 10^{-4} and 10^{-6} showed complete mortality at 73.4 and 88.0 hpi, respectively. Moreover, the last group using 10^{-8} and 10^{-10} inoculum showed complete mortality at 162.3 and 210 hpi, respectively (Figure 3B). The first deaths were recorded at 18 and 20.3 hpi for 10^2 and 10^4 dilutions, respectively. Regarding the per os infection, the first death was recorded at 47 hpi; all shrimp were dead at 139 hpi. The 10^{-6} dilution treatment resulted in an intermediate survival compared to the other dilution treatments, displaying a similar behavior as the per os infection. The calculated lethal dose at 50% endpoint dilution (LD_{50}/mL) was $10^{-6.5}$. Therefore, the 10^{-6} dilution was used for the successive tests. The WSSV-Negative group showed 100% survival.

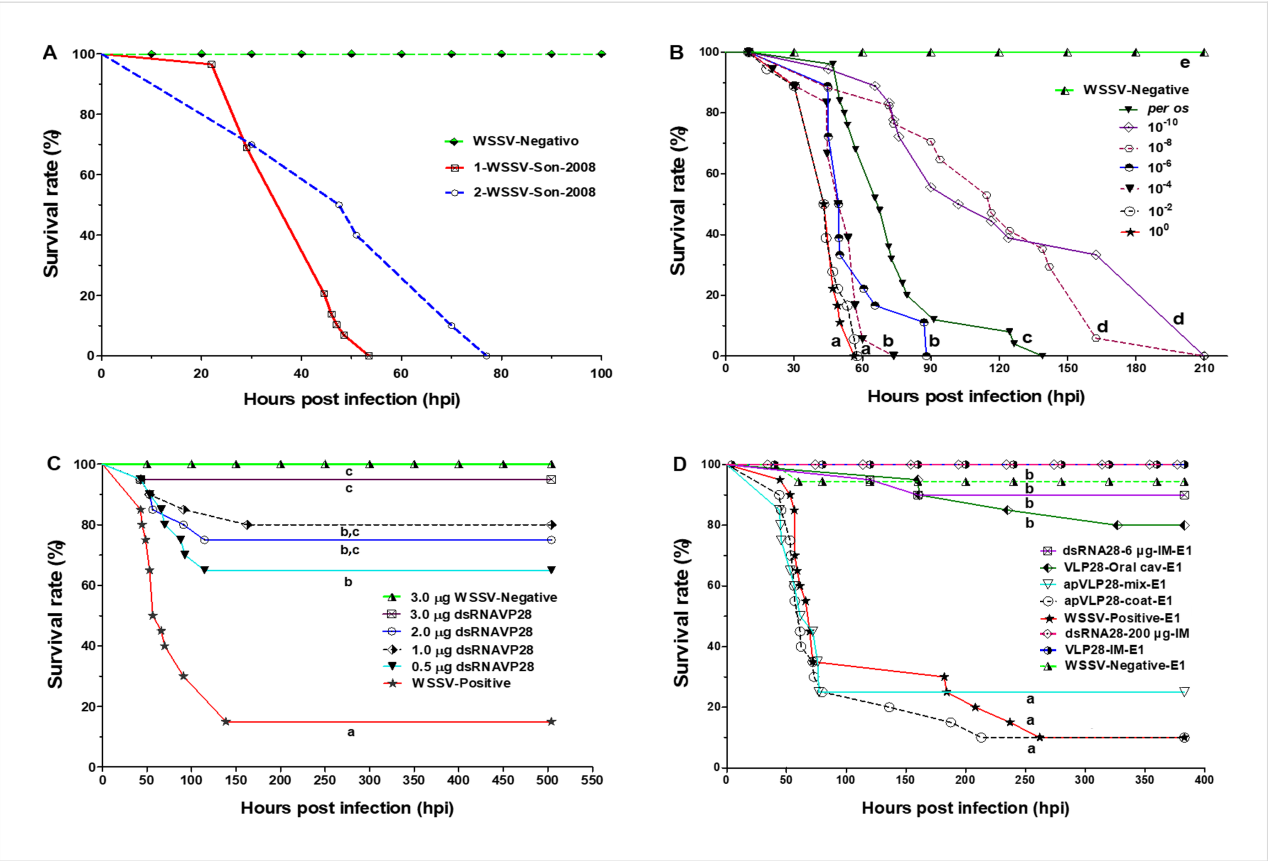


Figure 3: *P. vannamei* survival when exposed to WSSV and treatments. (A) IM inoculum activation in two consecutive experiments (1-WSSV-Son2008 and 2-WSSV-Son2008) (B) Per os infection with WSSV-Son2008 isolate to determine the LD_{50}/mL (C) IM dsRNAvp28 at 3.0, 2.0, 1.0, and 0.5 µg, WSSV-Negative control received a 3.0 µg dose. The survival was evaluated up to 504 hpi (21 days). (D) Oral antiviral treatment with VLP-dsRNAvp28 (6 µg per shrimp) in the pellet with fish oil (industrial grade) as a binding agent. Different letters (a–d) in each experiment (B, C, and D) on the curves indicate significant differences ($p < 0.0001$) among treatments using the Log-rank (Mantel–Cox) Test and not the final absolute survival percentage. IM, intramuscular injection, see Table ST4 (Supporting Information File 1) for treatment abbreviature details.

The WSSV-infected shrimp treated with different amounts (0.5, 1.0, 2.0, and 3.0 µg) of dsRNAvp28 through IM resulted in a significantly higher survival rate of >60% compared to the infected group without treatment (WSSV-Positive) with 15% survival in 21 days (Figure 3C). The high mortality of shrimp occurred between 70 and 100 hpi in all treatments. The survival curves resulted in a significant difference ($p < 0.0001$). When the WSSV-Negative (control non-infected) received 3.0 µg of dsRNAvp28/shrimp by IM, there was 100% shrimp survival. In comparison, the infected group treated with 3.0 µg of dsRNAvp28/shrimp showed only one death at 43 hpi (95% survival) during the 21 days of the experiment. As a result of the dose-response, 3.0 µg/shrimp was chosen as the subsequent dose for the IM treatments.

Different results were obtained with treated shrimp fed with pellets carrying the VLP-dsRNAvp28. When pellets were coated with VLP-dsRNAvp28 mixed with fish oil (ApVLP28-coat-E1) there was a 10% survival. Whereas those fed with ApVLP28-mix-E1 resulted in 25% survival (Table ST4, Exp. 1 in Supporting Information File 1) up to 384 hpi. However, the positive group resulted in 100% survival. The control group from the dsRNA28-6 µg-IM-E1 treatment achieved a 90% survival compared to the VLP28-IM-E1 group, with 100% survival (Figure 3D).

When the VLPs were administered via the oral cavity (VLP28-oral cav-E1), an 80% survival was obtained. Moreover, the group of shrimp that were given an IM dose of 200 µg of free dsRNAvp28 and infected with WSSV (dsRNA28-200 µg-IM) all survived up to the end of the experiment, 16 days post-infection (dpi), without showing any abnormal symptoms or behavior observable due to high dose of dsRNAvp28 (Figure 3D).

When VLP-dsRNAvp28 was used to coat pellets with salmon oil alone (ApsVLP28-coat-E2) or mixed (ApsVLP28-mix-E2), low survival was observed (50 and 31.25%, respectively) (Figure 4A). Simultaneously, the VLP28-IM-E2 group resulted in 100% survival until the end of the experiment (15 dpi). The groups treated by oral cavity (VLP28-oral-cav-E2) or naked

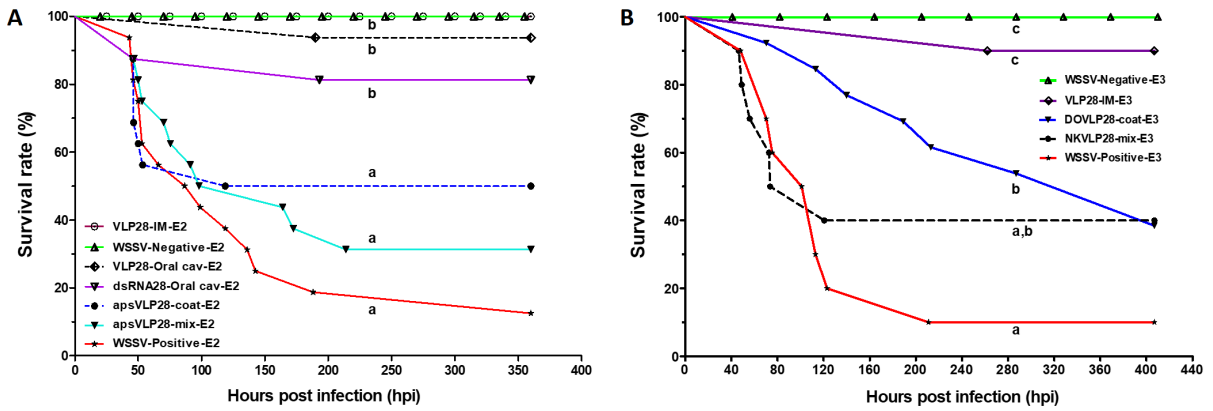


Figure 4: Cumulative survival curves of *P. vannamei* infected with WSSV and provided with VLPs antiviral treatment. (A) Pellets with VLP-dsRNAvp28 and covered with salmon oil (apsVLP28-coat or mixed). Administration by oral cavity (VLP28-oral cav-E2, and dsRNA28-oral cav-E2) means that the antiviral was given with a syringe right into the oral cavity. Note that these groups had high survival (93.75% and 81.25%) up to 350 hpi or 15 days. (B) Pellets with VLP-dsRNAvp28 prepared with commercial binders (Dry Oil® and NutriKelp®). The bioassay was ended 17 days post-treatment (or 400 hpi). Different letters (a–c) on the curves indicate significant differences ($p < 0.0001$) between treatments with Log-rank (Mantel–Cox), not the final absolute survival percentage. IM, intramuscular injection, see Table ST4 in Supporting Information File 1 for treatment abbreviature details.

dsRNA28 (dsRNA28-oral cav-E2) had 93.75 and 81.25% survival. In contrast, the WSSV-Positive-E2 showed a 12.5% survival rate until the end of the experiment (360 hpi, 15 days). Moreover, no significant differences between this group and those treated orally with ApsVLP28-coat-E2 and ApsVLP28-mix-E2 were found; whereas the WSSV-Negative-E2 treatment had a 100% survival rate.

While the Dry Oil® binder (DOVLP28-coat) and NutriKelp® binder (NKVLP28-mix) were used to incorporate the VLPs, a 38.5 and 40% survival rate was obtained, respectively. Whereas the control groups VLP28-IM-E3, WWSV-Negative-E3, and WSSV-Positive-E3 had a survival rate of 90, 100, and 10%, respectively. The cumulative survival curves of treatments with pellets VLP-dsRNAvp28 prepared using commercial binders are shown in Figure 4B.

The analysis of qPCR data showed that the viral load decreases significantly ($p < 0.05$) in WSSV-infected shrimp survival when orally treated with VLP-dsRNAvp28 (VLP28-mix and VLP28-coat), compared to positive controls (WSSV-Positive). However, similar results were obtained for shrimp-fed pellets prepared with different binders (fish oil and commercial binders). Organisms treated with VLPs by IM (VLP28-IM) or oral cavity (VLP28-Oral-cav) therapy were WSSV negative in more than 90% (15–17 dpi). After 60 dpi, the organisms treated IM with VLP-dsRNAvp28 had a slight degree of infection. Shrimp treated by oral antiviral therapy, with coated and mixed pellets (VLP28-coat and VLP28-mix) and collected dying or dead, resulted in higher viral load concentrations compared to those collected alive (Table 1). However, at the end of the experiment those shrimp collected alive were positive for WSSV, but with a slight degree of infection. The WSSV-Negative controls were free of virus.

Discussion

CCMV VLPs containing dsRNA were successfully synthesized to silence the WSSV VP28 protein expression. Here we used a 6:1 mass ratio of capsid protein to dsRNA, according to previous works for the encapsidation of ssRNA [42] and siRNA [32]. To our knowledge, this is the first report showing a long dsRNA encapsidation using a plant virus capsid protein.

The analysis by EMSA showed that the VLPs that self-assemble migrate differently than the free dsRNAi (Figure 1, lane 4). After dialysis in assembly buffer (pH 7.2), the sample analysis by TEM shows the spherical procapsids formation and CP-dsRNAi complexes (Figure 2A, images a and b). The Gaussian fit size distribution of the spherical procapsids gave an average diameter of 21.2 nm and corresponded to capsids with triangulation number $T = 2$. The sample's dialysis at pH 7.2 favors the electrostatic interactions between CCMV CP-dsRNA to form procapsids and CP-RNA complexes [42]. The dsRNA negative charges can be neutralized by the positive N-terminal protein in these procapsids [42]. However, procapsids are not suitable for any treatment because they do not efficiently protect their cargo. The dsRNA in the procapsids may be degraded by nucleases [42,53]. The appropriate synthesis of VLP-dsRNAvp28 was only obtained after the sample was acidified by dialysis in virus buffer (pH 4.5). After acidification no more aberrant and complex capsids were observed (Figure 2A, images c and d). The low pH promotes protein–protein interac-

Table 1: WSSV copies in shrimp abdominal tissue by real-time quantitative PCR (qPCR). Average copies of WSSV in ng^{-1} and SD values of shrimp treated with coated and mixed pellets using industrial-grade fish oil (ap), salmon fish oil (aps), and commercial binders (DO and NK).

treatment	live		dying/dead	
	average[a]	SD	average[a]	SD
apVLP28-mix[b]	2.39×10^{10}	3.10×10^{10}	3.027×10^{10}	2.93×10^{10}
apVLP28-coat[b]	9.36×10^{4}	3.89×10^{4}	1.23×10^{10}	5.82×10^{9}
apsVLP28-mix[c]	2.32×10^{4}	2.71×10^{4}	7.39×10^{9}	1.07×10^{10}
apsVLP28-coat[c]	2.01×10^{4}	1.48×10^{4}	4.79×10^{9}	4.04×10^{9}
DOVLP28-mix	1.11×10^{4}	6.85×10^{3}	6.33×10^{9}	6.76×10^{9}
NKVLP28-coat	7.87×10^{3}	7.75×10^{3}	7.91×10^{8}	6.98×10^{8}
WSSV-positive			1.30×10^{10}	2.60×10^{10}

[a]WSSV copies [ng^{-1}]; [b]ap = industrial grade fish oil; [c]aps = salmon oil.

tions and allows the stable forms of VLPs [54]. The TEM micrographs revealed two types of VLPs shapes: the icosahedral and the long tubular structures. The individual VLPs with a spherical (icosahedral) shape are likely to have few ssRNA molecules, due to the low contamination of the RNAi stock with ssRNA of 563 nts (according to the company that synthesized the RNA). The icosahedral VLPs are not empty, because in assembly buffer the CCMV CP form capsids only when anionic molecules are present.

On the other hand, the long tubular structures result from the experimental conditions during the VLPs formation. Due to the isoelectric point of the capsid protein (pH ≈ 4.8), the protein charge can easily modify the capsid protein dimers' spontaneous curvature, leading to the formation of tubular structures [55,56]. Also, the dsRNAvp28 is a long dsRNA with a persistence length of around 60 nm [57,58] that could be favorable for tube formation. The interaction of a rigid and quasi-long dsRNA molecule with the capsid protein dimers enables the elongated tubular structure formation under these experimental conditions [59].

The spherical VLPs have an average diameter of 25.8 nm corresponding to capsids with a triangulation number $T = 3$, similar to the wild type (wt) CCMV [53]. Whereas in the nanotubular VLPs, a diameter of 21.7 nm is revealed. In this work, the nanotubular length was not determined because the tubular synthesized VLPs are very long, and some are curved, making the measurement difficult. The correlation of the TEM and EMSA results suggests that the band in the agarose gel migrated slightly less than the wild type CCMV corresponds to, in multiple icosahedral capsids and short tubes with dsRNA. Similar results have been obtained with long ssRNA [42]. In contrast, the band that is close to the well corresponds to the long nanotubes. The individual icosahedral VLPs are not possible to visualize in the gel electrophoresis due to their low concentration in the sample. Most of dsRNAvp28 is self-assembled into long nanotubular VLPs, and similar results have been reported with dsDNA [60].

The CCMV has been reported to be biocompatible in mammals, testing the wild-type virus in mice [32,61]. However, there were no studies to demonstrated non-toxicity in other species such as crustaceans and fish. Therefore, before performing the bioassays with VLP-dsRNAvp28, this study evaluated the toxicity of wt CCMV in healthy shrimp. The bioassay was carried out for three weeks, and the shrimp showed no symptoms of any disease or apparent abnormality when treated by IM (up to 20 µg of CCMV per shrimp). Higher doses of dsRNAvp28 (200 µg) per WSSV infected shrimp by IM injection were also evaluated, showing no adverse effects or evident disease (Figure 4A).

The biocompatibility of CCMV in shrimp is of great commercial significance. The biocompatibility of CCMV suggests a broad potential to develop treatments for disease control in aquatic organisms and mammals.

Plant virus-based VLPs, in general, are particularly advantageous in aquaculture and medicine because they are biocompatible, biodegradable, and do not infect mammals [32,62] or marine organisms. To date, CCMV has shown the ability to be distributed widely in mouse organs and tissues using different administration routes [61]. Also, the CCMV VLPs are resistant to enzyme degradation through the digestive tract [32-34]. It is to be kept in mind that possibly the shrimp's virus, in contrast to CCMV VLP's, needs specific receptors to be internalized in the shrimp cells. For these reasons, CCMV VLPs

show quite an advantage over the VLPs derived from the shrimp virus.

The mortality rate of shrimp inoculated with WSSV is dose-dependent [52,63,64]. Dose dependency can be grouped in three virulence levels, according to the dilutions used: high 10^{-1}–10^{-2} (45–43.5 hpi), medium 10^{-4}–10^{-6} (51.4–49.5 hpi) and low 10^{-8}–10^{-10} (116.5–109.3 hpi). It is important to note that similar mortality behavior was observed between dilution 10^{-6} and infection per os. The median lethal dose obtained here ($10^{-6.5}$ LD_{50}/mL) is consistent with previous reports [46,52]. In our experiments, the cumulative mortality of 100% for the 10^{-6} (LD_{50}/mL) dose was at 88 hpi, and the median lethal time LT_{50} was 49.58 hpi.

The amount of inoculum orally ingested was estimated to be more than that of IM injection, because only a small proportion of the virus inoculated orally can infect shrimp [46]. However, we observed that challenged shrimp did not consume all the macerated infected tissue offered. Then, by inoculation per os, ≈10% of infected tissue biomass was used (for two days). It registered an accumulated mortality rate of 92% at 124.5 hpi and 100% at 139 hpi. In our study, the median survival time was 67.7 hpi. However, even if the results are consistent, the infection by IM injection is recommended in challenge bioassays, allowing greater viral dose control, compared to infection per os where it is difficult to calculate the consumption of infected shrimp tissue [65].

By IM injection, the dsRNAvp28 resulted in a great protective efficacy in *P. vannamei* against WSSV infection. Experimental results indicate that a minimum dose of 0.5 μg/shrimp is enough to protect up to 65% of the population against the virus. The maximum dose used in the present work was 3.0 μg/shrimp with 95% protection at 504 hpi (21 days). These evaluated doses are lower than those previously reported from 5.5 μg doses [66] up to 31 μg of dsRNA/shrimp [23]. This work has demonstrated the efficacy of the dose, and the sequence of the dsRNA used. According to our results, and considering possible losses by dispersing the VLP-dsRNAvp28 in the water or inside the shrimp, a maximum dose of 6 μg of dsRNA/shrimp as a single dose can be considered for oral administration.

The treatments using salmon fish oil to adhere the CCMV VLP28 to the feed pellet showed an increase in shrimp survival up to 50% (ApsVLP28-coat). On the other hand, therapy with VLP-dsRNAvp28 taken orally was more effective than when merely present in the feed as a coating. Taking the VLP-dsRNA28 orally assures capsid functionality by protecting the dsRNA structure. Administering VLP capsids inside the feed resulted in increased shrimp survival after challenged with the WSSV and treated per os. The survival rates obtained were 38.5% and 40% with DOVLP28-coat and NKVLP28-mix, respectively. Although the percentage with DOVLP28-mix is lower than NKVLP28-coat, the mortality was higher with the last treatment, reaching a 50% mortality rate at 73.6 hpi compared to 287 hpi that reached 53.8% mortality rate with DOVLP28-mix. This protection is significantly higher compared to the first results using fish oil. Both treatments using commercial binders indicate that it is possible to administer it in the pellets. However, it is crucial to state that pellets usually undergo pelleting or extrusion, damaging the VLPs. Therefore, further studies should be on how this can be administered in the pellets at industrial levels.

Other studies by IM injection of chemically modified chitosan nanoparticles loaded with anti-VP28 RNA [20] and antisense plasmid constructs for VP28 [24] have shown protection of 95% and 90%, respectively. However, in all these treatments the shrimp exposed to WSSV finally died at 14 dpi. To date, only two works have reported the use of VLPs with dsRNAvp28 against WSSV. In both cases, the VLPs were synthesized from viruses that infect shrimp. One was with the macrobrachium rosenbergii nodavirus (MrNv) [66], whereas the second was with the infectious hypodermal and hematopoietic necrosis virus (IHHNV) [31]; both studies showed a survival rate of 44.5 and 40% by IM injection (6 μg of dsRNAvp28 per shrimp), respectively. Here, we were able to obtain similar results when VLP-dsRNAvp28 was administered per os. However, we experiment with the same dose of VLP-dsRNAvp28 (6.0 μg/shrimp), equivalent to the same dilution of WSSV to infect them. But the bioassay was finished at 17 dpi.

However, one should not rule out possible differences in the shrimp origin line (genetics, immunology), feeding factors, manipulation (stress), the pathogenicity of the used WSSV isolate, and the infective dose, among others. By IM injection with the CCMV VLP-dsRNAvp28, we found survival rates of up to 100% with 17 dpi and up to 50% survival rates at 60 dpi using one single dose of 6 μg. In contrast with IM administration reports, we showed a good survival rate by oral antiviral therapy. It is important to note that our results show practically 100% protection through IM injection. Xie et al. [27] considered that the main difficulty in applying RNAi in shrimp in vivo is its intracellular release. Although naked dsRNA can penetrate cell membranes when injected locally, it is rapidly degraded by plasma nucleases.

The treated organisms with VLP-dsRNAvp28 by oral cavity obtained an 86% survival rate. However, during the oral cavity application treatment (VLP28-oral cav-3 and dsRNA28-oral cav), some shrimp regurgitated part of the treatment, so the efficacy

by this route was 86.8 and 81.2% survival, respectively (experiment E2). Although oral cavity and IM application showed a high survival compared to the administration of the VLP-dsRNAvp28 in pellet, it could indicate that VLPs: 1) were lost in the water by pellet detachment; 2) were not ingested by shrimp; 3) shrimp enzymes degraded it; or 4) a high concentration of VLPs was lost in feces. We hypothesized that an adequate amount is not being absorbed, since the observed survival rate does not exceed 50%. Thus, the problem is not the treatment itself but the dose that finally reaches the shrimp tissues. An investigation will be conducted testing higher doses.

Oral antiviral treatment in aquatic organisms is not straightforward because of the enormous challenges of breaking the water barrier. Therefore, for therapy or vaccine, it is essential to maintain, before ingestion, the compound's stability and the adherence to the pellets. When shrimp eat the pellet they have the peculiar tendency to fragment it (due to its size, and to food selectivity for palatability, hardness). This differs from fish, who swallow the whole pellet. Therefore, a considerable amount of VLP-dsRNAvp28 can be lost in the water while the shrimp is feeding.

The experiments presented were performed using different shrimp sizes from 3.6 ± 0.7 to 17.7 ± 2.7 g. However, no size effect could be detected on the amount of dsRNAvp28 administered IM and orally. The efficacy of dsRNAvp28 by IM from 3.0 to 6.0 μg per organism, was effective in small and large shrimp, indicating the possibility that doses used are higher than required.

The efficacy of CCMV VLP-dsRNAvp28 to protect WSSV infected shrimp was verified by qPCR. Viremia was reduced in orally treated organisms. Therefore, oral administration should be considered effective as antiviral therapy before viral infection, since extra doses will be necessary. (But keep in mind that infected shrimp will stop eating from three to four days after initial infection, so oral therapy at that point cannot cure them). The qPCR data analysis indicates that VLP-dsRNAvp28 by oral therapy reduces the mortality rate by reducing the WSSV infection.

Mejía-Ruiz et al. [28] reported that antiviral protection provided by a single IM administration of dsRNAvp28 is short-lived, 10 to 20 days post-treatment (dpt), with 63% and 87% mortality rate, respectively, being gradually lost after 30 dpt, Also Witteveldt et al. [67] observed that viral protection in *P. monodon* was reduced 21 days after administering orally VP28 expressed in bacteria as an antiviral treatment. Furthermore, Ufaz et al. [20] showed that the protective effect of treatment remains active at least two weeks after viral exposure. In shrimp farms usually, the WSSV is not detected until dead organisms are perceived, making it impossible to determine precisely the time of infection. However, it might be possible to protect neighbor ponds or farms once the onset of a local viremia is detected nearby.

We hypothesize that antiviral therapy based on CCMV VLP-dsRNAvp28 with a single dose by oral administration cannot exceed one month of protection. According to the survival results, the IM injection up to two months protection could be achieved. For this reason, the antiviral therapy would be based mainly on preventive therapy or at the first signs of infection, through continuous prophylaxis during the period of the shrimp culture. By this means, the risk of crop losses before a potential outbreak occurs could be avoided. Once shrimp are infected by the WSSV, they will stop eating within 18 to 24 hpi, so at that point, oral administration is no longer possible.

In this work, we have shown that VLPs derived from the CCMV have a high potential as a vehicle for RNAi delivery. Likewise, the brome mosaic virus (BMV) VLPs-dsRNAvp28 show similar results to those of the CCMV (data not shown). Furthermore, these VLPs can be chemically modified with a peptide or using protein engineering, to express on its external surface to better recognize (target) the WSSV infected cells increasing the antiviral therapy efficiency.

Because new viral outbreaks are the primary threat to aquaculture production, innovative biosecurity measures to limit production losses are essential [68,69]. Biosecurity programs do not always reduce the incidence of outbreaks in areas where the WSSV is prevalent in natural carriers [2]. Thus, current prevention strategies do not eradicate the virus. It is imperative to find prevention that works. Vaccines or antiviral therapies to effectively control or eliminate these outbreaks should be a priority in further investigations. The Government and private sector should work together to develop strategies to protect the profitability of the aquaculture sector [70].

Conclusion

This work represents the first study of long dsRNAs encapsidation using plant virus capsid proteins, such as CCMV, for WSSV treatment in shrimp. Our results indicate that intramuscular injection treatment revealed a survival rate of nearly 100%, while a 90% survival is shown by oral cavity administration using CCMV VLP-dsRNAvp28 in shrimp infected with WSSV. However, using the CCMV VLPs orally administered in feed pellets resulted in a survival rate of 40%.

Our preliminary results shown here with CCMV VLP-dsRNAvp28 offers adequate protection against WSSV. Al-

though the therapy proves effective protection, reinforcement to protect the organisms during a culture season or when an outbreak begins to occur in neighboring ponds or farms also can be applied.

We report the different strategies that provide a significant advance in methods for the delivery of therapeutic molecules. The antiviral therapy here presented could be applied, with further research, to other aquatic species or even terrestrial organisms, or within nanomedicine applications.

Supporting Information

Supporting Information File 1

Tables of detailed experimental assays and methods to prepare the pellet feed containing VLP-dsRNAvp28.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-17-95-S1.pdf]

Acknowledgements

We thank Francisco Ruiz-Medina, for TEM analysis, Dr. Noe Ruiz-Garcia for helping with the qPCR statistical analysis, Dr. Mario Galaviz from UABC, and Mochis Zazueta Urias from SRY Promotora Acuícola S.A. de C.V. Camahuiroa, Huatabampo, Sonora for providing shrimp, Dr. Katrin Quester, and M.Sc. Itandehui Betanzo Gutierrez for technical assistance. We also appreciate Nataly López Molina, Francisco Saucedo, Álvaro Hernández and Geronimo Avila, for technical assistance in laboratory trails. Finally, the authors thank William Johnson Dawson for editing.

Funding

This work was supported by National Council of Science and Technology of Mexico (CONACYT), Grants PN 247474, and CB 239878. RDCN acknowledges PASPA-DGAPA-UNAM for sabbatical support. SRC acknowledges CONACyT Ph.D. fellowship No. 215469.

ORCID® iDs

Santiago Ramos-Carreño - https://orcid.org/0000-0002-8137-0198
Ivone Giffard-Mena - https://orcid.org/0000-0003-1688-0703
Alfredo Nuñez-Rivera - https://orcid.org/0000-0001-6315-6801
Jaime Ruiz-Garcia - https://orcid.org/0000-0003-3730-3825
Maria Teresa Viana - https://orcid.org/0000-0002-3074-767X
Ruben D. Cadena-Nava - https://orcid.org/0000-0001-8428-6701

References

1. Chou, H.; Huang, C.; Wang, C.; Chiang, H.; Lo, C. *Dis. Aquat. Org.* **1995,** *23,* 165–173. doi:10.3354/dao023165
2. Flegel, T. W. *J. Invertebr. Pathol.* **2012,** *110,* 166–173. doi:10.1016/j.jip.2012.03.004
3. Lightner, D. V. *J. Invertebr. Pathol.* **2011,** *106,* 110–130. doi:10.1016/j.jip.2010.09.012
4. Flegel, T. W. *Aquaculture* **2006,** *258,* 1–33. doi:10.1016/j.aquaculture.2006.05.013
5. Joseph, T. C.; James, R.; Rajan, L. A.; Surendran, P. K.; Lalitha, K. V. *Biotechnol. Rep.* **2015,** *7,* 51–54. doi:10.1016/j.btre.2015.04.006
6. Vaseeharan, B.; Jayakumar, R.; Ramasamy, P. *Lett. Appl. Microbiol.* **2003,** *37,* 443–447. doi:10.1046/j.1472-765x.2003.01428.x
7. Thammasorn, T.; Sangsuriya, P.; Meemetta, W.; Senapin, S.; Jitrakorn, S.; Rattanarojpong, T.; Saksmerprome, V. *BMC Biotechnol.* **2015,** *15,* 110. doi:10.1186/s12896-015-0226-9
8. Escobedo-Bonilla, C. M. *J. Antivirals Antiretrovirals* **2013,** S9. doi:10.4172/jaa.s9-001
9. Verbruggen, B.; Bickley, L.; van Aerle, R.; Bateman, K.; Stentiford, G.; Santos, E.; Tyler, C. *Viruses* **2016,** *8,* 23. doi:10.3390/v8010023
10. Escobedo-Bonilla, C. M.; Vega-Peña, S.; Mejía-Ruiz, C. H. *J. King Saud Univ., Sci.* **2015,** *27,* 182–188. doi:10.1016/j.jksus.2014.11.004
11. Sarathi, M.; Simon, M. C.; Venkatesan, C.; Hameed, A. S. S. *Mar. Biotechnol.* **2008,** *10,* 242–249. doi:10.1007/s10126-007-9057-6
12. Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C. *Nature* **1998,** *391,* 806–811. doi:10.1038/35888
13. Montgomery, M. K.; Xu, S.; Fire, A. *Proc. Natl. Acad. Sci. U. S. A.* **1998,** *95,* 15502–15507. doi:10.1073/pnas.95.26.15502
14. Itsathitphaisarn, O.; Thitamadee, S.; Weerachatyanukul, W.; Sritunyalucksana, K. *J. Invertebr. Pathol.* **2017,** *147,* 76–85. doi:10.1016/j.jip.2016.11.006
15. Labreuche, Y.; Warr, G. W. *Fish Shellfish Immunol.* **2013,** *34,* 1002–1010. doi:10.1016/j.fsi.2012.06.008
16. Robalino, J.; Bartlett, T.; Shepard, E.; Prior, S.; Jaramillo, G.; Scura, E.; Chapman, R. W.; Gross, P. S.; Browdy, C. L.; Warr, G. W. *J. Virol.* **2005,** *79,* 13561–13571. doi:10.1128/jvi.79.21.13561-13571.2005
17. Chazal, N.; Gerlier, D. *Microbiol. Mol. Biol. Rev.* **2003,** *67,* 226–237. doi:10.1128/mmbr.67.2.226-237.2003
18. Leu, J.-H.; Tsai, J.-M.; Wang, H.-C.; Wang, A. H.-J.; Wang, C.-H.; Kou, G.-H.; Lo, C.-F. *J. Virol.* **2005,** *79,* 140–149. doi:10.1128/jvi.79.1.140-149.2005
19. Wu, W.; Wang, L.; Zhang, X. *Virology* **2005,** *332,* 578–583. doi:10.1016/j.virol.2004.12.011
20. Ufaz, S.; Balter, A.; Tzror, C.; Einbender, S.; Koshet, O.; Shainsky-Roitman, J.; Yaari, Z.; Schroeder, A. *Mol. Syst. Des. Eng.* **2018,** *3,* 38–48. doi:10.1039/c7me00092h
21. Chang, Y.-S.; Liu, W.-J.; Lee, C.-C.; Chou, T.-L.; Lee, Y.-T.; Wu, T.-S.; Huang, J.-Y.; Huang, W.-T.; Lee, T.-L.; Kou, G.-H.; Wang, A. H.-J.; Lo, C.-F. *PLoS One* **2010,** *5,* e10718. doi:10.1371/journal.pone.0010718
22. Tang, X.; Wu, J.; Sivaraman, J.; Hew, C. L. *J. Virol.* **2007,** *81,* 6709–6717. doi:10.1128/jvi.02505-06
23. Nilsen, P.; Karlsen, M.; Sritunyalucksana, K.; Thitamadee, S. *Sci. Rep.* **2017,** *7,* 1028. doi:10.1038/s41598-017-01181-w
24. Akhila, D. S.; Mani, M. K.; Rai, P.; Condon, K.; Owens, L.; Karunasagar, I. *Aquaculture* **2015,** *435,* 306–309. doi:10.1016/j.aquaculture.2014.10.005
25. Ma, J.; Bruce, T. J.; Jones, E. M.; Cain, K. D. *Microorganisms* **2019,** *7,* 569. doi:10.3390/microorganisms7110569
26. Schroeder, A.; Levins, C. G.; Cortez, C.; Langer, R.; Anderson, D. G. *J. Intern. Med.* **2010,** *267,* 9–21. doi:10.1111/j.1365-2796.2009.02189.x

27. Xie, F. Y.; Woodle, M. C.; Lu, P. Y. *Drug Discovery Today* **2006,** *11,* 67–73. doi:10.1016/s1359-6446(05)03668-8
28. Mejía-Ruíz, C. H.; Vega-Peña, S.; Alvarez-Ruiz, P.; Escobedo-Bonilla, C. M. *J. Invertebr. Pathol.* **2011,** *107,* 65–68. doi:10.1016/j.jip.2011.02.002
29. Whitehead, K. A.; Langer, R.; Anderson, D. G. *Nat. Rev. Drug Discovery* **2009,** *8,* 129–138. doi:10.1038/nrd2742
30. Kiatmetha, P.; Chotwiwatthanakun, C.; Jariyapong, P.; Santimanawong, W.; Ounjai, P.; Weerachatyanukul, W. *PeerJ* **2018,** *6,* e6079. doi:10.7717/peerj.6079
31. Jariyapong, P.; Chotwiwatthanakun, C.; Pooljun, C.; Weerachatyanukul, W. *Aquaculture* **2019,** *504,* 260–266. doi:10.1016/j.aquaculture.2019.02.001
32. Nuñez-Rivera, A.; Fournier, P. G. J.; Arellano, D. L.; Rodriguez-Hernandez, A. G.; Vazquez-Duhalt, R.; Cadena-Nava, R. D. *Beilstein J. Nanotechnol.* **2020,** *11,* 372–382. doi:10.3762/bjnano.11.28
33. Azizgolshani, O.; Garmann, R. F.; Cadena-Nava, R.; Knobler, C. M.; Gelbart, W. M. *Virology* **2013,** *441,* 12–17. doi:10.1016/j.virol.2013.03.001
34. Chen, Q.; Lai, H. *Hum. Vaccines Immunother.* **2013,** *9,* 26–49. doi:10.4161/hv.22218
35. Hasebe, R.; Suzuki, T.; Makino, Y.; Igarashi, M.; Yamanouchi, S.; Maeda, A.; Horiuchi, M.; Sawa, H.; Kimura, T. *BMC Microbiol.* **2010,** *10,* 165. doi:10.1186/1471-2180-10-165
36. Zhang, Y.; Dong, Y.; Zhou, J.; Li, X.; Wang, F. *Molecules* **2018,** *23,* 2311. doi:10.3390/molecules23092311
37. Steinmetz, N. F. *Nanomedicine (N. Y., NY, U. S.)* **2010,** *6,* 634–641. doi:10.1016/j.nano.2010.04.005
38. Pokorski, J. K.; Steinmetz, N. F. *Mol. Pharmaceutics* **2011,** *8,* 29–43. doi:10.1021/mp100225y
39. Yildiz, I.; Shukla, S.; Steinmetz, N. F. *Curr. Opin. Biotechnol.* **2011,** *22,* 901–908. doi:10.1016/j.copbio.2011.04.020
40. Konecny, R.; Trylska, J.; Tama, F.; Zhang, D.; Baker, N. A.; Brooks, C. L., III; McCammon, J. A. *Biopolymers* **2006,** *82,* 106–120. doi:10.1002/bip.20409
41. Steinmetz, N. F.; Evans, D. J. *Org. Biomol. Chem.* **2007,** *5,* 2891–2902. doi:10.1039/b708175h
42. Cadena-Nava, R. D.; Comas-Garcia, M.; Garmann, R. F.; Rao, A. L. N.; Knobler, C. M.; Gelbart, W. M. *J. Virol.* **2012,** *86,* 3318–3326. doi:10.1128/jvi.06566-11
43. Villagrana-Escareño, M. V.; Reynaga-Hernández, E.; Galicia-Cruz, O. G.; Durán-Meza, A. L.; De la Cruz-González, V.; Hernández-Carballo, C. Y.; Ruíz-García, J. *BioMed Res. Int.* **2019,** 1–11. doi:10.1155/2019/4630891
44. Ramos-Paredes, J.; Grijalva-Chon, J. M.; De la Rosa-Vélez, J.; Enríquez-Paredes, L. M. *Aquacult. Res.* **2012,** *43,* 339–348. doi:10.1111/j.1365-2109.2011.02836.x
45. Ramos-Carreño, S.; Valencia-Yáñez, R.; Correa-Sandoval, F.; Ruíz-García, N.; Díaz-Herrera, F.; Giffard-Mena, I. *Arch. Virol.* **2014,** *159,* 2213–2222. doi:10.1007/s00705-014-2052-0
46. Escobedo-Bonilla, C. M.; Wille, M.; Alday Sanz, V.; Sorgeloos, P.; Pensaert, M. B.; Nauwynck, H. J. *Dis. Aquat. Org.* **2005,** *66,* 163–170. doi:10.3354/dao066163
47. Prior, S.; Browdy, C. L.; Shepard, E. F.; Laramore, R.; Parnell, P. G. *Dis. Aquat. Org.* **2003,** *54,* 89–96. doi:10.3354/dao054089
48. Ramakrishnan, M. A. *World J. Virol.* **2016,** *5,* 85. doi:10.5501/wjv.v5.i2.85
49. Thomas, A.; Sudheer, N. S.; Viswanathan, K.; Kiron, V.; Bright Singh, I. S.; Narayanan, R. B. *J. Invertebr. Pathol.* **2014,** *123,* 17–24. doi:10.1016/j.jip.2014.08.004
50. Choi, M. R.; Kim, Y. J.; Jang, J.-S.; Kim, S.-K. *J. Microbiol. Biotechnol.* **2011,** *21,* 170–175. doi:10.4014/jmb.1005.05036
51. Durand, S. V.; Lightner, D. V. *J. Fish Dis.* **2002,** *25,* 381–389. doi:10.1046/j.1365-2761.2002.00367.x
52. Escobedo-Bonilla, C. M.; Audoorn, L.; Wille, M.; Alday-Sanz, V.; Sorgeloos, P.; Pensaert, M. B.; Nauwynck, H. J. *Dis. Aquat. Org.* **2006,** *68,* 181–188. doi:10.3354/dao068181
53. Comas-Garcia, M.; Cadena-Nava, R. D.; Rao, A. L. N.; Knobler, C. M.; Gelbart, W. M. *J. Virol.* **2012,** *86,* 12271–12282. doi:10.1128/jvi.01695-12
54. Bancroft, J. B. *Adv. Virus Res.* **1970,** *16,* 99–134. doi:10.1016/s0065-3527(08)60022-6
55. Vega-Acosta, J. R.; Cadena-Nava, R. D.; Gelbart, W. M.; Knobler, C. M.; Ruiz-García, J. *J. Phys. Chem. B* **2014,** *118,* 1984–1989. doi:10.1021/jp407379t
56. Lavelle, L.; Gingery, M.; Phillips, M.; Gelbart, W. M.; Knobler, C. M.; Cadena-Nava, R. D.; Vega-Acosta, J. R.; Pinedo-Torres, L. A.; Ruiz-Garcia, J. *J. Phys. Chem. B* **2009,** *113,* 3813–3819. doi:10.1021/jp8079765
57. Abels, J. A.; Moreno-Herrero, F.; van der Heijden, T.; Dekker, C.; Dekker, N. H. *Biophys. J.* **2005,** *88,* 2737–2744. doi:10.1529/biophysj.104.052811
58. Doose, S.; Barsch, H.; Sauer, M. *Biophys. J.* **2007,** *93,* 1224–1234. doi:10.1529/biophysj.107.107342
59. Mukherjee, S.; Pfeifer, C. M.; Johnson, J. M.; Liu, J.; Zlotnick, A. *J. Am. Chem. Soc.* **2006,** *128,* 2538–2539. doi:10.1021/ja056656f
60. de Ruiter, M. V.; van der Hee, R. M.; Driessen, A. J. M.; Keurhorst, E. D.; Hamid, M.; Cornelissen, J. J. L. M. *J. Controlled Release* **2019,** *307,* 342–354. doi:10.1016/j.jconrel.2019.06.019
61. Kaiser, C. R.; Flenniken, M. L.; Gillitzer, E.; Harmsen, A. L.; Harmsen, A. G.; Jutila, M. A.; Douglas, T.; Young, M. J. *Int. J. Nanomed.* **2007,** *2,* 715–733.
62. Masarapu, H.; Patel, B. K.; Chariou, P. L.; Hu, H.; Gulati, N. M.; Carpenter, B. L.; Ghiladi, R. A.; Shukla, S.; Steinmetz, N. F. *Biomacromolecules* **2017,** *18,* 4141–4153. doi:10.1021/acs.biomac.7b01196
63. Ngo, T. T. N.; Senior, A. M.; Culina, A.; Santos, E. S. A.; Vlak, J. M.; Zwart, M. P. *J. Fish Dis.* **2018,** *41,* 1733–1744. doi:10.1111/jfd.12877
64. Tang, K. F. J.; Lightner, D. V. *Aquaculture* **2000,** *189,* 11–21. doi:10.1016/s0044-8486(00)00367-7
65. Huang, Y.-C.; Yin, Z.-X.; Ai, H.-S.; Huang, X.-D.; Li, S.-D.; Weng, S.-P.; He, J.-G. *Aquaculture* **2011,** *311,* 54–60. doi:10.1016/j.aquaculture.2010.11.032
66. Jariyapong, P.; Chotwiwatthanakun, C.; Direkbusarakom, S.; Hirono, I.; Wuthisuthimethavee, S.; Weerachatyanukul, W. *Aquaculture* **2015,** *435,* 86–91. doi:10.1016/j.aquaculture.2014.09.034
67. Witteveldt, J.; Cifuentes, C. C.; Vlak, J. M.; van Hulten, M. C. W. *J. Virol.* **2004,** *78,* 2057–2061. doi:10.1128/jvi.78.4.2057-2061.2004
68. Bondad-Reantaso, M. G.; Subasinghe, R. P.; Arthur, J. R.; Ogawa, K.; Chinabut, S.; Adlard, R.; Tan, Z.; Shariff, M. *Vet. Parasitol.* **2005,** *132,* 249–272. doi:10.1016/j.vetpar.2005.07.005
69. *OIE Aquatic Animal Health Code 2018,* 21st ed.; OIE - Aquatic Animal Health Code; World Organisation for Animal Health: 12 rue de Prony, 75017 Paris, France, 2018. Available online: http://www.oie.int/en/standard-setting/aquatic-code/access-online/.
70. Moss, S. M.; Moss, D. R.; Arce, S. M.; Lightner, D. V.; Lotz, J. M. *J. Invertebr. Pathol.* **2012,** *110,* 247–250. doi:10.1016/j.jip.2012.01.013

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Double-headed nucleosides: Synthesis and applications

Vineet Verma[1], Jyotirmoy Maity[2], Vipin K. Maikhuri[1], Ritika Sharma[1], Himal K. Ganguly[3] and Ashok K. Prasad[*1]

Review

Open Access

Address:
[1]Bioorganic Laboratory, Department of Chemistry, University of Delhi, Delhi-110 007, India, [2]Department of Chemistry, St. Stephen's College, University of Delhi, Delhi-110 007, India and [3]Department of Biophysics, Bose Institute, P1/12 CIT Scheme VIIM, Kolkata-700 054, India

Email:
Ashok K. Prasad[*] - ashokenzyme@gmail.com

* Corresponding author

Keywords:
acyclic double-headed nucleosides; bicyclic double-headed nucleosides; furanosyl double-headed nucleosides; modified nucleosides; pyranosyl double-headed nucleosides

Beilstein J. Org. Chem. **2021,** *17,* 1392–1439.
https://doi.org/10.3762/bjoc.17.98

Received: 27 January 2021
Accepted: 27 May 2021
Published: 08 June 2021

This article is part of the thematic issue "Celebrating the role of chemistry in the success of oligonucleotides as therapeutics".

Guest Editors: P. Kumar and T. Brown

Abstract

Double-headed nucleoside monomers have immense applications for studying secondary nucleic acid structures. They are also well-known as antimicrobial agents. This review article accounts for the synthetic methodologies and the biological applications of double-headed nucleosides.

Introduction

Nucleosides are the constructional subunits of deoxyribonucleic acids (DNA) or ribonucleic acids (RNA), which contain either a purine or pyrimidine nucleobase and a furanosyl moiety of pentose sugars, 2′-deoxyribose or ribose [1,2]. Nucleotides are constituted by addition of a phosphate group at the 5′-position of the nucleosides and these monomeric units polymerize to construct nucleic acids (DNA or RNA). These macromolecules preserve and express genetic information in all living cells and viruses. Modified nucleosides are a class of organic compounds which are unnatural and have an altered/substituted nucleobase and/or a modified pentose sugar [3,4]. The synthetic accessibility of these organic molecules encouraged researchers to prepare sugar-modified nucleosides [5,6] and nucleobase-modified nucleosides [7,8]. Modified nucleoside monomers comprising more than one nucleobase are called double-headed nucleosides [9,10]. A thorough literature search regarding double-headed nucleosides disclosed that these modified nucleosides were constituted with any two naturally occurring nucleobases, i.e., adenine, guanine, thymine, uracil, and cytosine [9,10] or one naturally occurring nucleobase and one heterocyclic/carbocyclic moiety either attached directly to the sugar or via a linker. Further modifications were introduced by

the substitution of some of the naturally occurring nucleobases by halogens or alkyl groups. On the other hand, a variety of heterocyclic/carbocyclic moieties were considered as the head of these modified nucleosides. The heterocyclic structures which were found to be attached to these double-headed nucleosides include triazolophthalazine [11], 4,6-di-*tert*-butylbenzoxazole [12], mesitylisoxazole [13], 5-trimethylsilyl-1,2,3-triazole [14], 1-pivaloyloxymethyl-1*H*-1,2,3-triazole [15], 1,3,4-oxadiazino[6,5-*b*]indole [16], 6,7-dihydro-6-oxo-5*H*-1,2,4-triazolo[3,4-*b*][1,3,4]thiadiazine [17], 1,2,4-triazino[5,6-*b*]indole [18], 1,3,4-thiadiazoline [19], 1,3,4-oxadiazoline [19], 1,2,4-triazoline [19], 3-mercapto-1*H*-1,2,4-triazole [20], 1,3,4-oxadiazole-2(3*H*)-thione [20], 4-amino-5-mercapto-4*H*-1,2,4-triazole [20], and 1,2,4-triazolo[3,4-*b*](1,3,4)-thiadiazole moieties [21]. Additionally, selected examples of double-headed nucleosides comprising aromatic/polyaromatic/carbocyclic moieties such as phenyl [13-15,22], pyrene [23-25], adamant-1-yl [24], cholesteryl [24], perylen-3-yl [24], 4-(*tert*-butyldimethylsilyloxy/hydroxy)phenyl [26], 3/4-(*N*-((dimethylamino)methy-lidene)aminosulfonyl)phenyl [26,27], and sulfonamido-substituted benzothiazole [28] attached as an additional head are also reported in this review article. Literature data revealed that most of the double-headed nucleosides have the first nucleobase attached to the anomeric carbon of the pentofuranosyl/hexopyranosyl sugar moiety and an additional nucleobase/heterocyclic/carbocyclic moiety attached either directly or through a linker to any carbon of the sugar moiety either by C–N or C–C bonds. However, in case of base to base double-headed nucleosides, the additional nucleobase/substituent or unsubstituted phenyl moiety/polyaromatic moiety/carbocyclic moiety/heterocyclic moiety is attached to the first nucleobase with/without a linker. Whereas, all the acyclic double-headed nucleosides had natural nucleobases or heterocyclic moieties attached at the terminal carbons only.

Double-headed nucleosides are synthetically derived nucleoside scaffolds that are known to impact significantly secondary structures in nucleic acids [29]. Some oligonucleotides containing a particular double-headed nucleotide monomer have been found to form a three-way junction structure with a hairpin loop and two flanking sequences [30,31]. Moreover, these nucleotides have been found to orient the additional nucleobase towards the core of the duplex to participate in Watson–Crick base pairing [32-34]. The incorporation of the double-headed nucleosides into oligonucleotides followed by their duplex formation studies against complimentary oligonucleotide strands had described a very selective zipper-interaction [35], whereas a relative stabilization was observed due to stacking of these additional nucleobases across the minor groove [31]. The biological activity of the acyclic double-headed nucleosides was assessed through in vitro studies on Gram-positive bacteria *Staphylococcus aureus*, *Listeria inovanii* and Gram-negative bacteria *Klebsiella pneumoniae*, *Salmonella* sp., and *Escherichia coli* [20]. Triazolyl double-headed nucleosides showed efficacy against eosinophil-derived neurotoxin, which is an eosinophil secretion protein and a member of the Ribonuclease A (RNase A) superfamily [36]. Double-headed nucleosides were also found to be active against orthopox viruses, vaccinia virus, and cowpox virus under in vitro conditions [11], whereas few double-headed nucleoside analogues showed a moderate cytostatic activity against human cervix carcinoma HeLa cells [37].

It is pertinent to mention that Sharma et al. [29] have reviewed the double-headed nucleotides in the recent past with a focus on their effects in nucleic acid duplexes and other secondary structures. Herein, we focused on the synthetic protocols used for accessing a variety of double-headed nucleoside monomers. Thus, this review is the comprehensive compilation of the synthetic protocols available for the production of double-headed nucleoside monomers and their applications. For better overview the review has been structured based on the types of the sugar moiety of the nucleoside and the position of the attachment of the additional base, either directly or through a linker on the sugar.

Review

Furanosyl double-headed nucleosides

Based on literature reports most of the double-headed nucleosides comprised a pentofuranosyl sugar moiety. Various synthetic methodologies have been developed for the introduction of the additional nucleobase/heterocyclic system directly or via a linker at the C-2′/C-3′/C-4′/C-5′ position of the pentofuranosyl moieties. We have categorized the double-headed furanosyl-nucleosides depending on the position of the attachment of the additional nucleobase/heterocyclic system at the particular carbon of the pentofuranosyl moiety of the nucleoside (Figure 1).

1,2-Furanosyl double-headed nucleosides

Herein, all nucleosides comprising furanosyl ring structures are included, with the first nucleobase attached to the C-1′ position and the second nucleobase introduced at the C-2′ position either with or without a linker (Figure 1).

Nielsen and co-workers [38,39] synthesized 2′-(pyrimidin-1-yl)methyl- or 2′-(purin-9-yl)methyl-substituted double-headed nucleosides **4a**–**f** of arabinofuranosyluracil. The convergent synthesis of the double-headed nucleosides was achieved from uridine, which was first converted to the 3′,5′-(1,1,3,3-tetraisopropyldisiloxan-1,3-diyl)-protected (TIPDS) ketonucleoside **1** following a standard procedure [40]. The subsequent

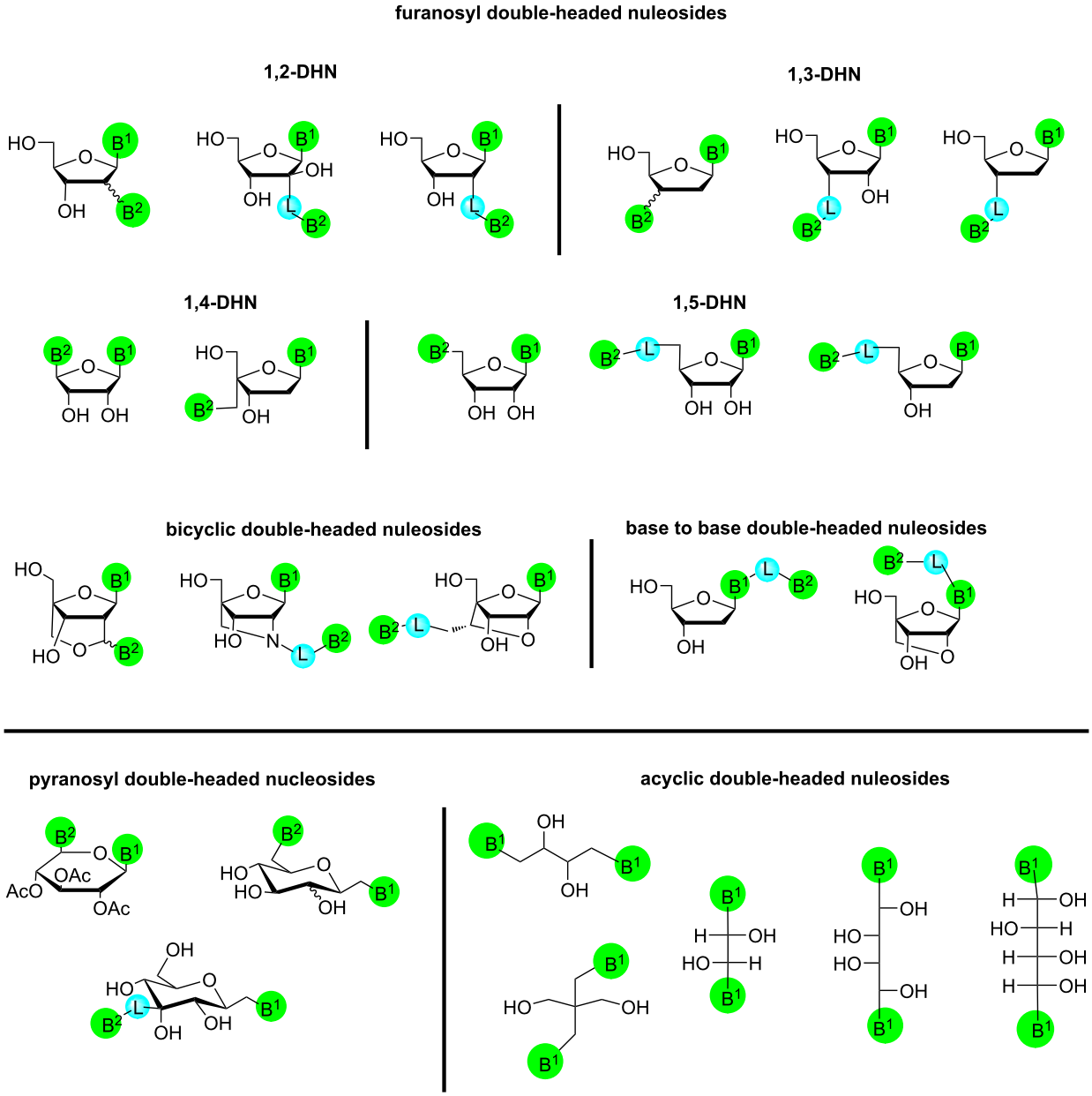
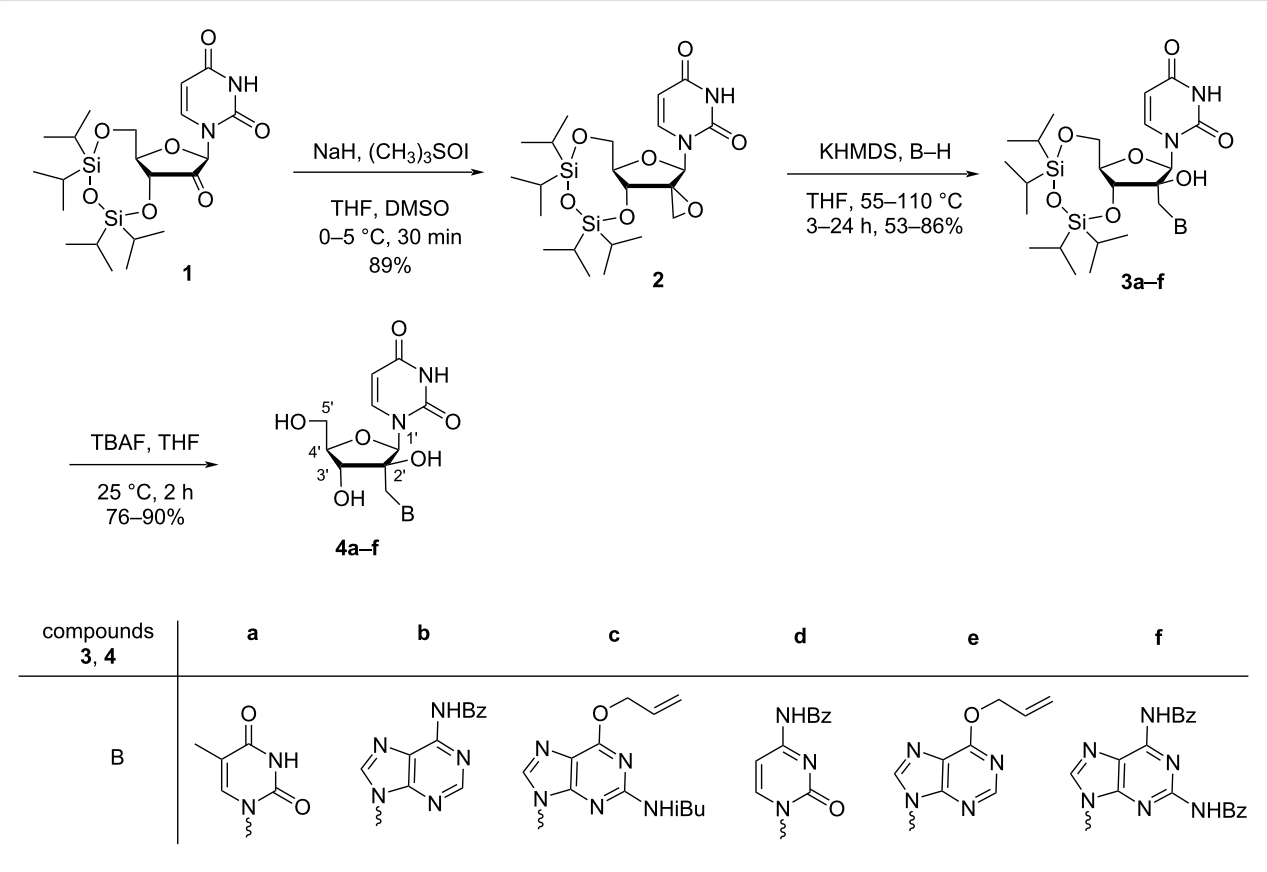


Figure 1: Double-headed nucleosides. B^1 and B^2 = nucleobases or heterocyclic/carbocyclic moieties; L = linker.

Corey–Chaykovsky epoxidation [41] of 2′-ketonucloside **1** with trimethylsulfoxonium iodide in DMSO afforded the spironucleoside **2**, which in turn was converted to the TIPDS-protected 2′-(pyrimidin-1-yl)methyl-/2′-(purin-9-yl)methylarabinofuranosyluracil derivatives **3a–f** by nucleophilic epoxide ring opening with thymine, *N*-benzoyladenine, 6-*O*-allyl-*N*-isobutyrylguanine, *N*-benzoylcytosine, 6-*O*-allylhypoxanthine or *N*,*N*-dibenzoyldiaminopurine in 53 to 83% yield. The desilylation of the nucleosides **3a–f** with tetrabutylammonium fluoride in tetrahydrofuran (THF) led to the formation of six different double-headed nucleosides **4a–f** (Scheme 1) [38,39].

The synthesized double-headed nucleosides **4a**,**b** were dimethoxytritylated (DMTr), phosphitylated, and incorporated into DNA oligonucleotides using the standard automated phosphoramidite method. The UV-based melting temperature (T_m) of hybrids of the modified oligonucleotides with complementary DNA strands were studied. The analysis of the melting temperature of the duplex and extensive molecular dynamics studies revealed that the synthesized double-headed nucleotides behave as functional dinucleotide mimics and hybridize with complementary targets neatly with their Watson–Crick faces compatible with natural DNA [39].

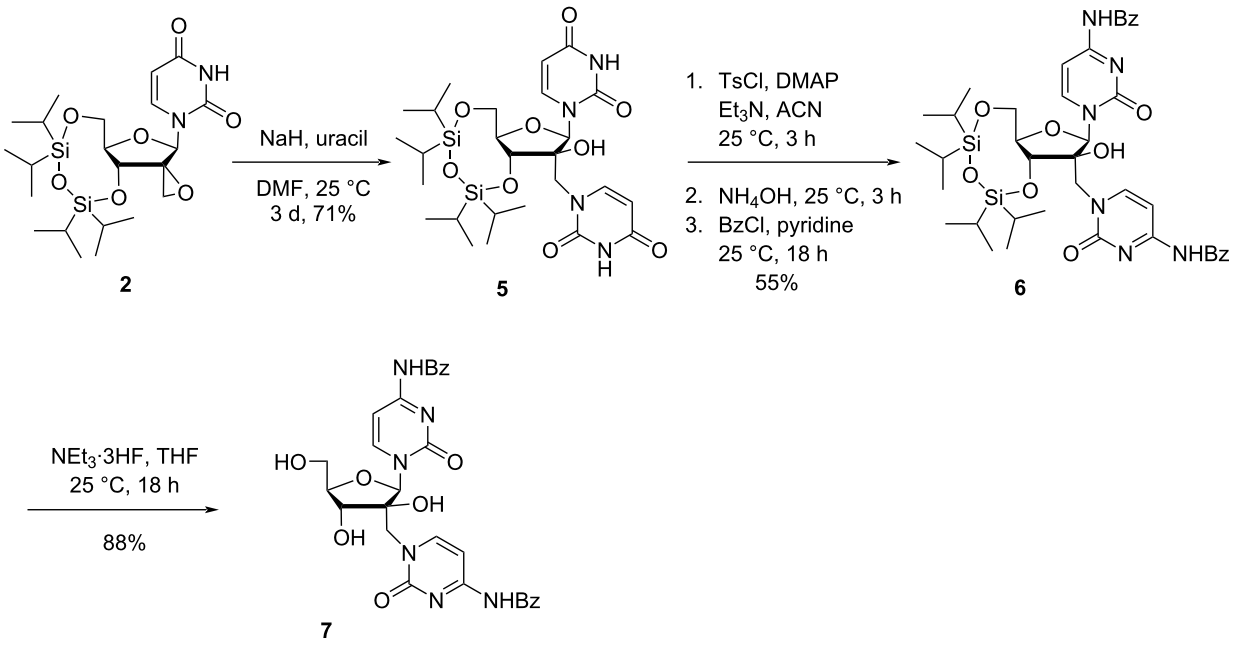


Scheme 1: Synthesis of 2′-(pyrimidin-1-yl)methyl- or 2′-(purin-9-yl)methyl-substituted double-headed nucleosides **4a–f** of arabinofuranosyluracil.

Nielsen and co-workers [42] additionally synthesized 2′-(*N*-benzoylcytosin-1-yl)methylarabinofuranosyl-*N*-benzoylcytosine (**7**) from uridine using a similar methodology. Thus, the nucleophilic epoxide ring opening in spironucleoside **2** with uracil in DMF in a N^1-regioselective manner afforded the TIPDS-protected double-headed nucleoside **5** having two uracil bases (the additional uracil being attached through a methylene linker to the 2′-position of arabinouracil). Subsequently, the two uracil bases of the TIPDS-protected double-headed nucleoside **5** were converted to *N*-benzoylated cytosines in a three-step one-pot procedure in 55% yield. For this conversion, the carbonyl group at the 4-position of uracil was first activated by tosylation, which was followed by conversion to the amine upon reaction with ammonia and protection of the newly introduced amino group with benzoyl chloride to afford the double-headed nucleoside **6**. The removal of the silyl protecting group with $NEt_3{\cdot}3HF$ in THF yielded 2′-(*N*-benzoylcytosin-1-yl)methylarabinofuranosyl-*N*-benzoylcytosine (**7**, Scheme 2) [42].

The double-headed nucleoside **7** was dimethoxytritylated and phosphitylated following the standard procedure and incorporated into oligonucleotides to study its effects on duplex stability. The single incorporation in oligonucleotides and study of the melting temperature (T_m) of its duplex hybridized with a complementary DNA strand revealed an increase in T_m by 4 °C with respect to the normal duplex. This indicated the participation of both nucleobases of the double-headed nucleotides in Watson–Crick base pairing. The same group also showed that a multiple incorporation of the double-headed nucleotide is also tolerated, but the double-headed nucleotides with the present design were not suitable as triplex-forming oligonucleotides [42].

Pedersen and Nielsen [35] synthesized a double-headed nucleoside with two different nucleobases, i.e., 2′-deoxy-2′-(thymine-1-yl)ethyluridine (**11**) (Scheme 3). The oxidative cleavage of the allyl group in TIPDS-protected 2-allyl-2-deoxyuridine **8** gave the TIPDS-protected hydroxynucleoside **9** as key intermediate. The treatment of **9** with benzoyl chloride under suitable conditions to selectively protect the 3-NH group of the uracil moiety afforded N^3-benzoyluridine (**10**). The reaction of TIPDS-protected hydroxynucleoside **9** with N^3-benzoylthymine under Mitsunobu reaction conditions, followed by deprotection with TBAF and aqueous methanolic ammonia resulted in the formation of 2′-deoxy-2′-(thymine-1-yl)ethyluridine (**11**) in 37% yield. When the same procedure was repeated with



Scheme 2: Synthesis of double-headed nucleoside **7** having two cytosine moieties.

N^3-benzoyluridine (**10**) the double-headed nucleoside **11** was obtained in 67% yield (Scheme 3) [35].

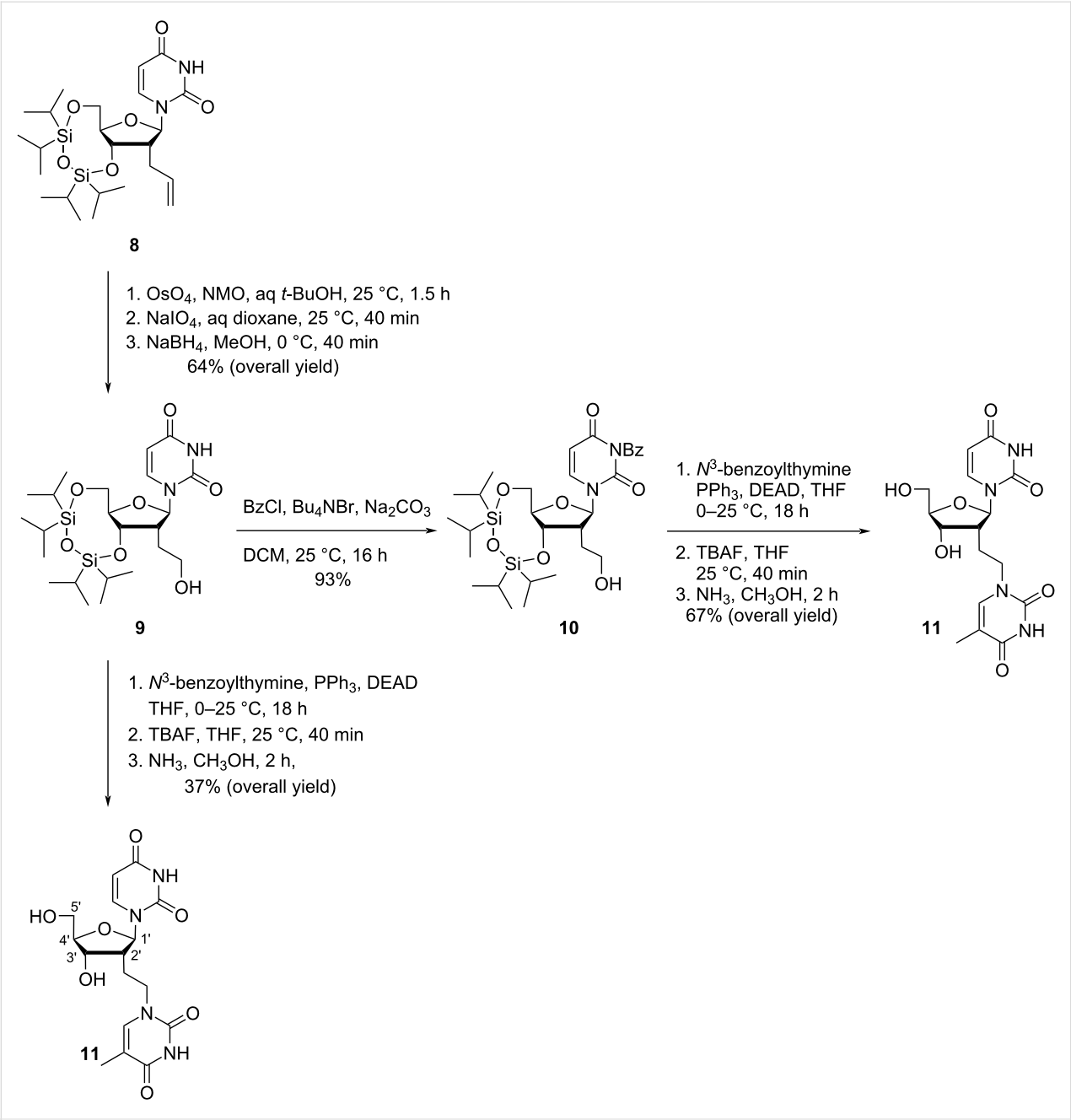
The double-headed nucleoside **11** was dimethoxytritylated and phosphitylated following the standard procedures and incorporated once into a 13-mer oligodeoxynucleotide and an LNA-modified oligodeoxynucleotide sequence, and four-times in the middle of a 12-mer oligodeoxynucleotide sequence in order to study the effect of the additional nucleobase in duplexes, bulged duplexes, and in three way junctions [35]. The designed double-headed nucleoside was found to be reasonably well tolerated in duplexes and stabilized three-way junctions. Significant conformational changes in these secondary structures have also been induced [35].

Nielsen and co-workers [43] synthesized 2′-(4-(thymin-1-ylmethyl)-1,2,3-triazole-1-yl)- and 2′-(4-(N^6-benzoyladenine-9-ylmethyl)-1,2,3-triazole-1-yl)-substituted double-headed nucleosides of 2′-deoxy-5′-*O*-(4,4′-dimethoxytrityl)uridine (**14** and **15**) from the nucleoside azide **12** which in turn was obtained by the nucleophilic opening of *O*-2,2′-anhydrouridine [44]. The azido nucleoside **12** was reacted with N^6-benzoyl-N^9-propargyladenine (**13a**) and N^1-propargylthymine (**13b**) via a CuAAC reaction where the triazole-containing linker connected the additional thymine or adenine to the 2′-position of 2′-deoxyuridine forming the double-headed nucleosides **14** and **15**, respectively (Scheme 4) [43].

Both double-headed nucleosides **14** and **15**, when incorporated into oligonucleotides were found to stabilize three-way junction in both DNA–DNA and DNA–RNA duplexes and when introduced into a (+1)-zipper motif, cross strand interactions were observed in a DNA–DNA duplex [43].

TIPDS protection of uridine (**16**), followed by the treatment of the product with acetic anhydride/acetic acid in DMSO produced the protected nucleoside **17** [45,46] (Scheme 5). Next, the fully protected nucleoside **17** was subjected to chlorination using thionyl chloride in dichloromethane, followed by the treatment of the product with N^3-benzoylthymine under basic conditions (K_2CO_3 in DMF) to produce the nucleoside **18**. The removal of the *tert*-butyldimethylsilyl group from nucleoside **18**, followed by dimethoxytritylation at the primary hydroxy and phosphitylation at the secondary hydroxy group afforded the double-headed nucleoside monomer **19** (Scheme 5) [45].

The synthesized double-headed nucleoside **19** was introduced in oligonucleotides and its impact on the secondary nucleic acid structure was studied. It was revealed that the double-headed nucleoside **19** was well accommodated in a hybrid DNA:RNA duplex and stabilized bulged duplex and three way junctions [45]. The potential of the double-headed nucleoside **19** in secondary nucleic acid structures was compared with the earlier reported monomer **11** and found to be inferior to double-headed nucleoside **11** due to the 3′-endo conformation which placed the

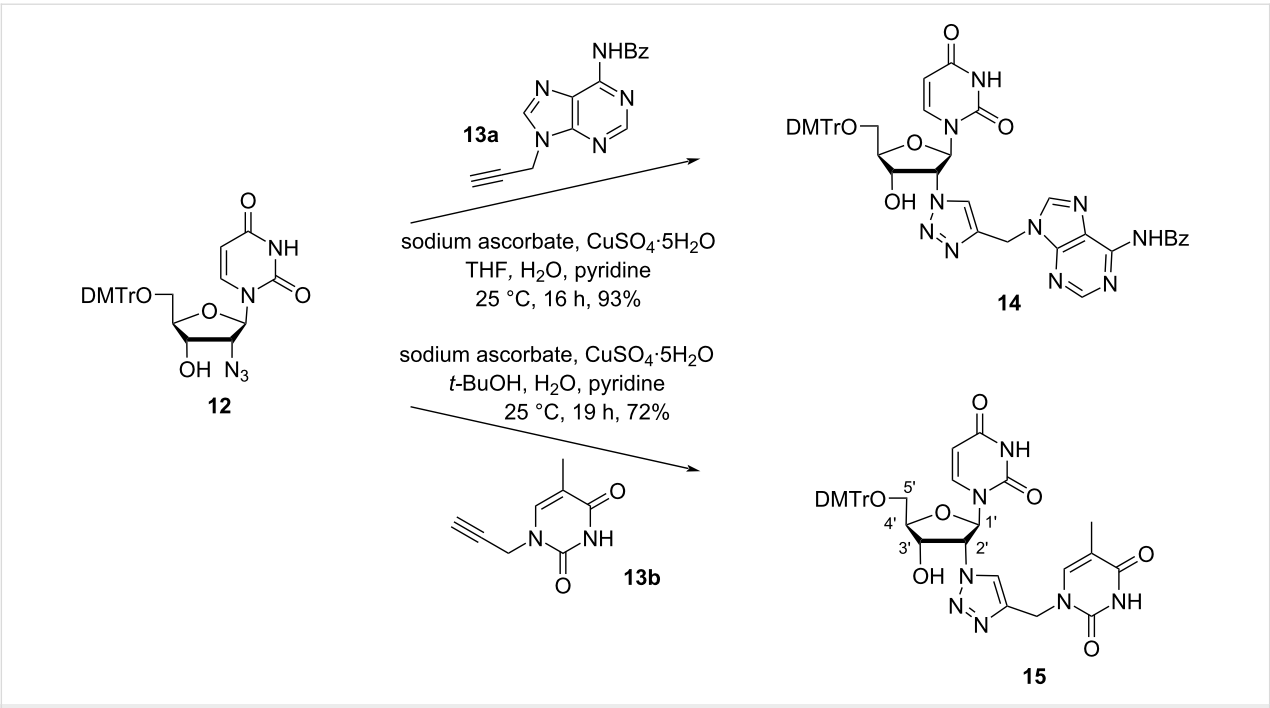


Scheme 3: Synthesis of double-headed nucleoside 2′-deoxy-2′-*C*-(2-(thymine-1-yl)ethyl)-uridine (**11**).

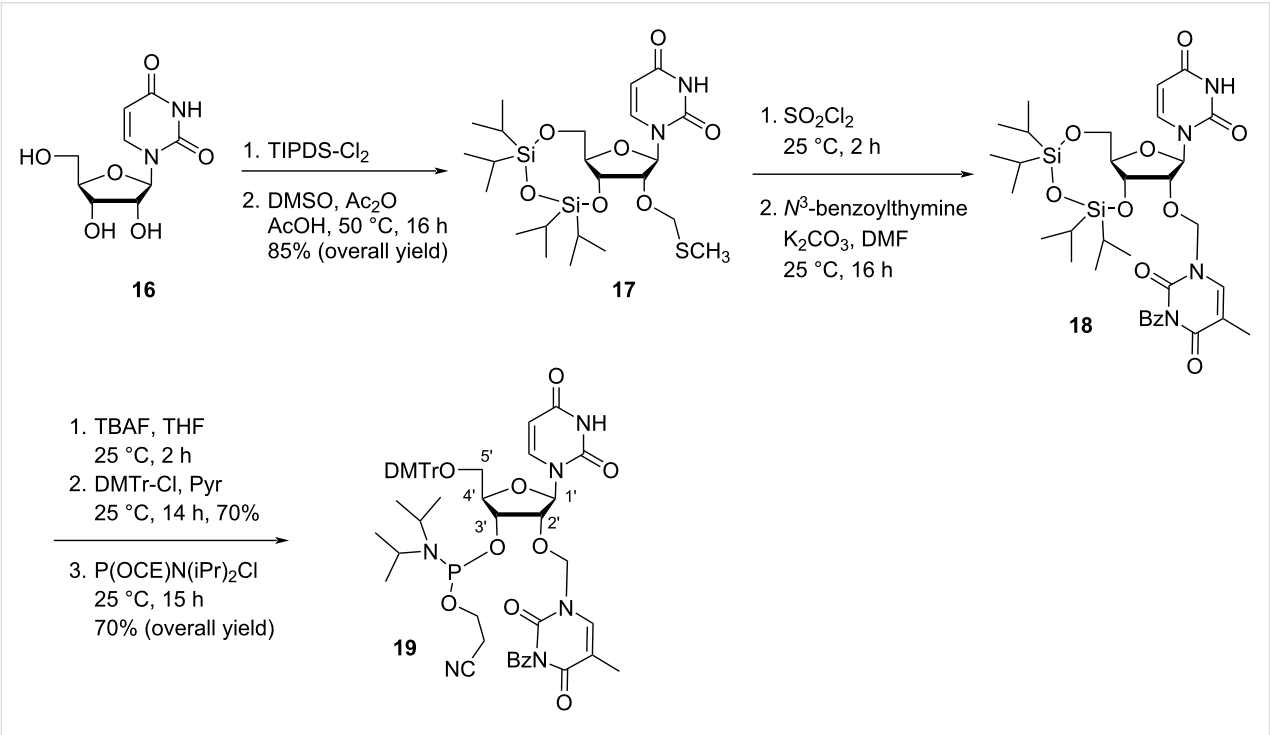
2′-substituent towards the minor groove rather than to the duplex core [35].

Vilarrasa and co-workers [47] synthesized 2′-uracil-1-yl and 2′-thymin-1-yl derivatives of 2′-deoxythymidine starting from uridine (**16**). The synthesis started with the TIPDS protection of **16** followed by introduction of an azide group in the C-2′ position of the molecule to afford nucleoside **22**. The treatment of azide **22** with pyrrolidine in acetonitrile followed by hydrogenation afforded aminonucleoside **23**, which was used as a key intermediate for the synthesis of the double-headed nucleosides **24** and **25** (Scheme 6) [47].

The same group [47] also synthesized the C-2′ isomeric nucleosides **28** and **29**, i.e., with inverted configuration at C-2′ as compared to nucleosides **24** and **25** (Scheme 7). The synthesis of these two nucleosides was carried out through the formation of the anhydro nucleoside **26** and its transformation into the aminonucleoside **27**. The key intermediate nucleoside **27** was then treated with 3-ethoxypropenoyl isocyanate or 3-methoxy-



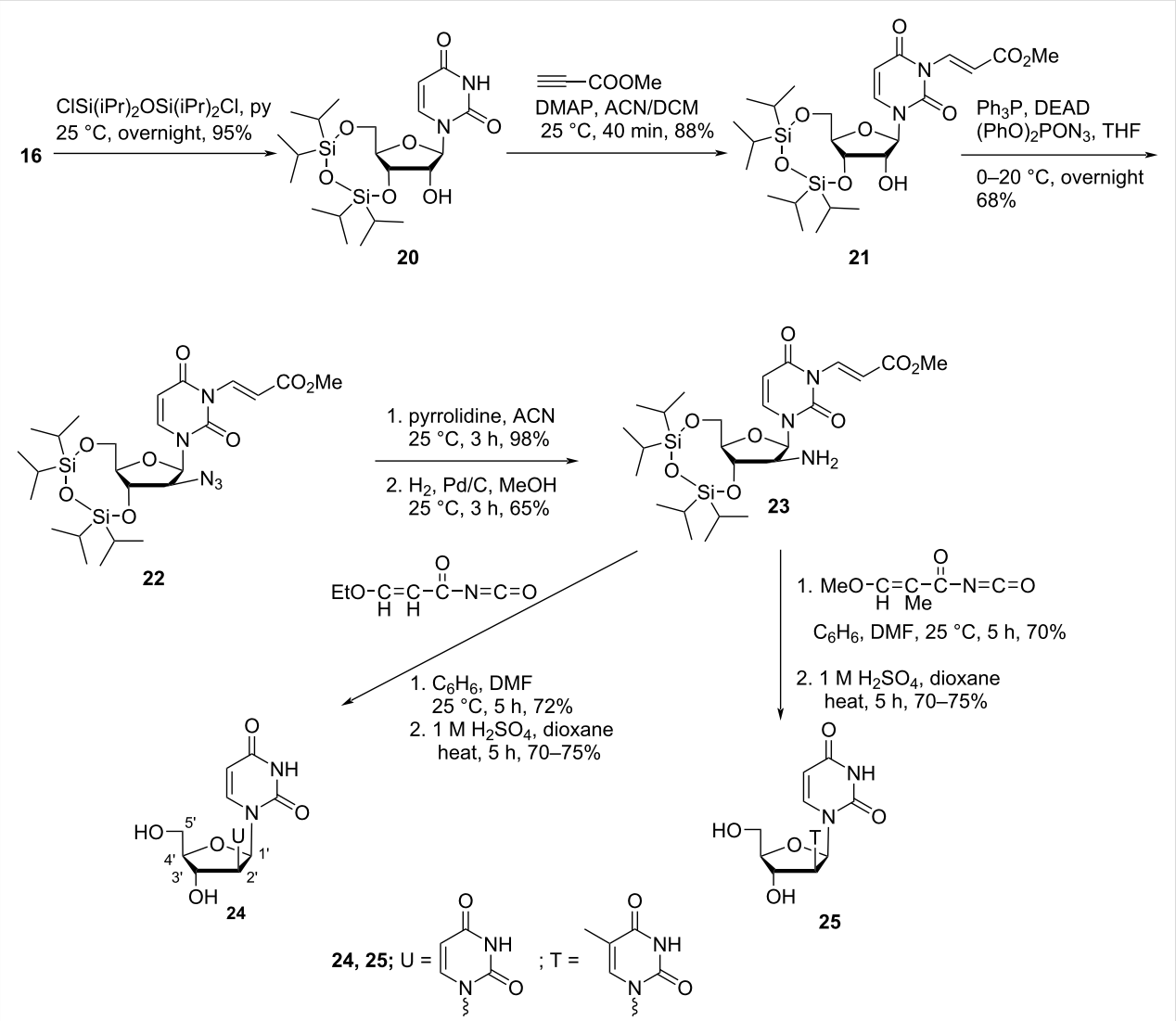
Scheme 4: Double-headed nucleosides **14** and **15** obtained by click reaction.



Scheme 5: Synthesis of the double-headed nucleoside **19**.

2-methylpropenoyl isocyanate in a mixture of benzene and DMF, followed by acidification with sulfuric acid affording the nucleosides **28** and **29**, respectively in high yields (Scheme 7) [47].

Nielsen and co-workers [33] synthesized the double-headed nucleoside 2′-*C*-(thymine-1-yl)methyl-2′-deoxyuridine (**33**) starting from the ribose derivative 3,5-bis-*O*-(*p*-chlorobenzyl)-2-deoxy-2-hydroxymethyl-α-D-ribofuranose (**30**) which in turn

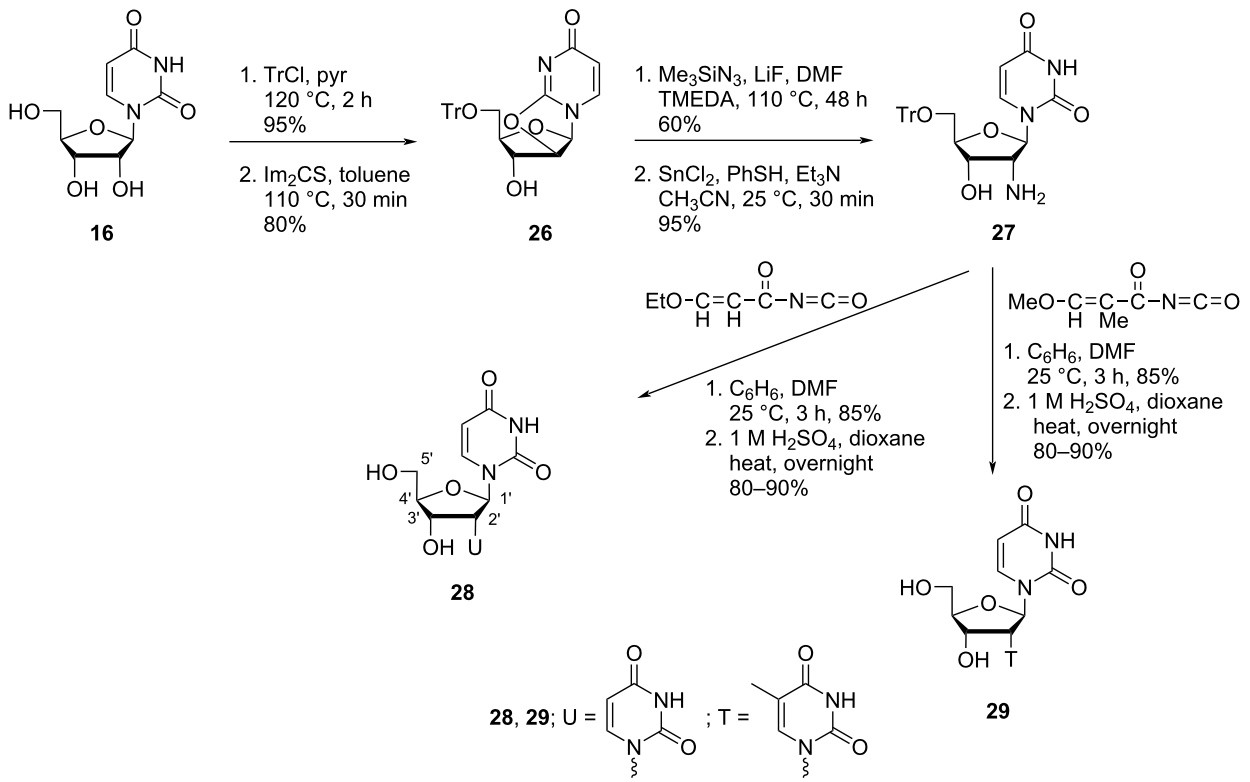


Scheme 6: Synthesis of the double-headed nucleosides **24** and **25**.

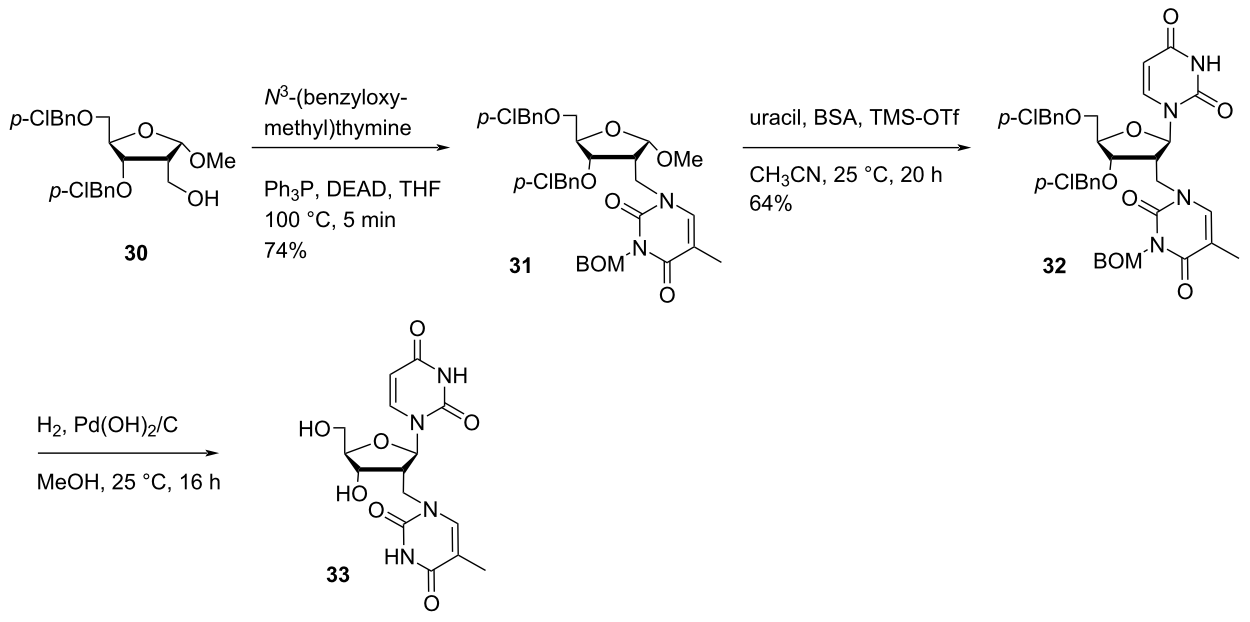
was synthesized from D-ribose in six steps following a procedure reported in the literature [48]. The Mitsunobu reaction of ribose derivative **30** and N^3-(benzyloxymethyl)thymine afforded nucleoside **31** which was subjected to Vorbrüggen coupling with silylated uracil to give the protected double-headed nucleoside **32**. Global deprotection of **32** using palladium-catalyzed hydrogenation conditions resulted in the formation of the targeted double-headed nucleoside **33** (Scheme 8) [33].

The double-headed nucleoside **33** was dimethoxytritylated, phosphitylated, and incorporated into duplex and its ability to recognize complementary base pairs was monitored by UV melting curve analysis [33]. Hybridization data revealed that the synthesized double-headed nucleotide recognized itself either through formation of Watson–Crick base pairs with two complementary adenosines or through the formation of T:T (thymine:thymine) base pairs that resulted in the formation of two novel nucleic acid motifs. The novel nucleic acid motifs could be incorporated either single or multiple times in dsDNA duplexes without altering its stability. It was revealed by molecular dynamics (MD) simulations that the DNA sugar–phosphate backbone accommodated modified nucleotide by stretching or curling up as required and all the four base pairs based upon the structure of the synthesized double-headed nucleotide could be accommodated in the similar way as the T:A (thymine:adenine) base pair in the motif. The nucleic acid motifs may also be used in designing nanoscale DNA structures where a specific duplex twist is required [33].

Nielsen and co-workers [34] also synthesized the double-headed nucleoside 2′-*C*-(N^6-benzoyladenine-9-yl)methyl-2′-



Scheme 7: Synthesis of double-headed nucleosides **28** and **29**.



Scheme 8: Synthesis of double-headed nucleoside **33**.

deoxyuridine (**37**) starting from ribose derivative 3,5-bis-*O*-(*p*-chlorobenzyl)-2-deoxy-2-hydroxymethyl-α-D-ribofuranose (**30**) [48]. The ribose derivative **30** was then reacted with triflic anhydride in the presence of pyridine followed by reaction with adenine in the presence of sodium hydride to afford nucleoside **34**. The nucleoside **34** was further reacted with benzoyl chloride to afford the fully protected nucleoside **35** which upon further reaction with silylated uracil in the presence of tin(IV)

chloride via Vorbrüggen coupling afforded the protected double-headed nucleoside **36**. The nucleoside **36** was finally deprotected in the presence of $Pd(OH)_2/C$ under hydrogen atmosphere to generate the double-headed nucleoside **37** (Scheme 9) [34].

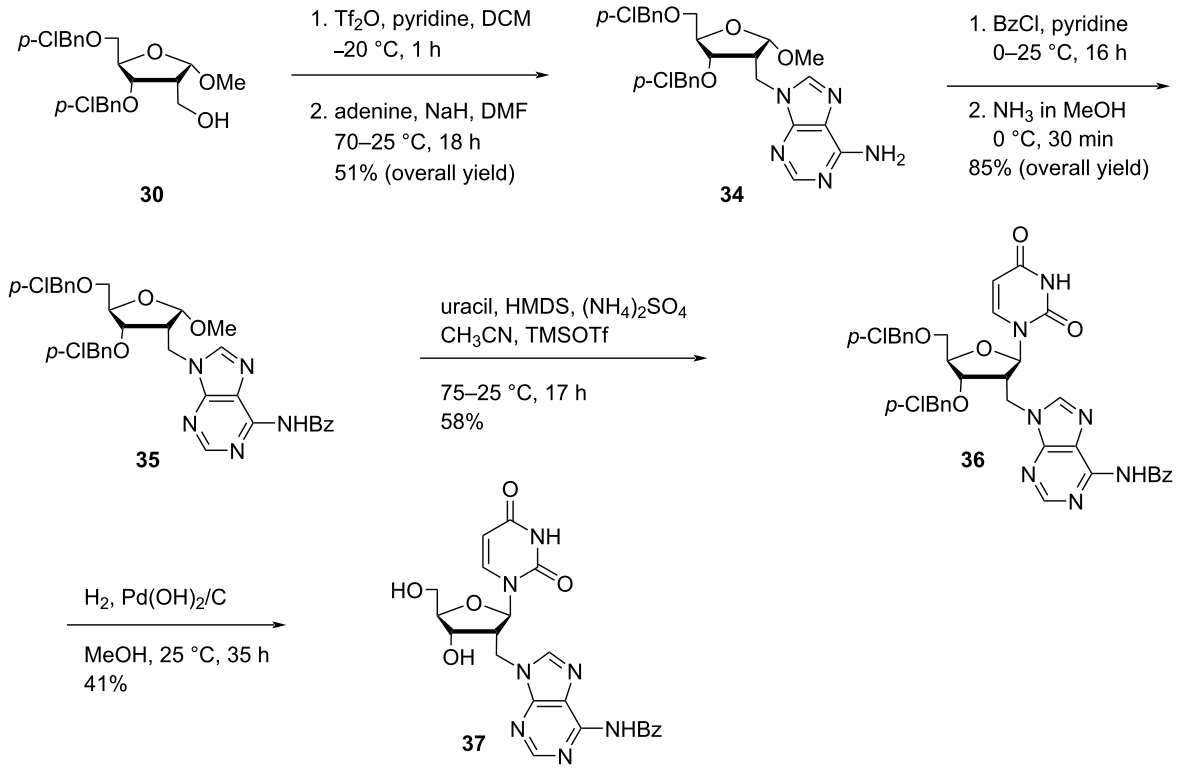
The double-headed nucleoside **37** was dimethoxytritylated, phosphitylated, and incorporated into 11- to 13-mer oligonucleotides using the standard automated phosphoramidite method. The UV-based melting temperature (T_m) of hybrids of the modified oligonucleotides with complementary DNA strands were studied. The analysis of the melting temperature of the resulting duplex revealed that the synthesized double-headed nucleotide behaved as a compressed dinucleotide and combination of all natural nucleobases on compressed scaffold can form Watson–Crick base pairs with complementary bases [34].

Nielsen and co-workers [23] synthesized the double-headed nucleoside 1-(5′-*O*-(4,4′-dimethoxytrityl)-2′-*C*-((4-(pyren-1-yl)-1,2,3-triazole-1-yl)methyl)arabinofuranosyl)uracil (**41**) starting from spironucleoside **2** which in turn was synthesized from uridine following a procedure reported in the literature [30,32,49]. The spironucleoside **2** was then reacted with sodium azide to afford the *arabino*-uridine **38** with an azidomethyl group in the C-2′ position. The *arabino*-uridine **38** was reacted with TBAF and 4,4′-dimethoxytrityl chloride to afford nucleoside **39** which was reacted with 1-ethynylpyrene (**40**) under copper-catalyzed alkyne–azide cycloaddition (CuAAC) reaction conditions to yield the double-headed nucleoside **41** (Scheme 10) [23].

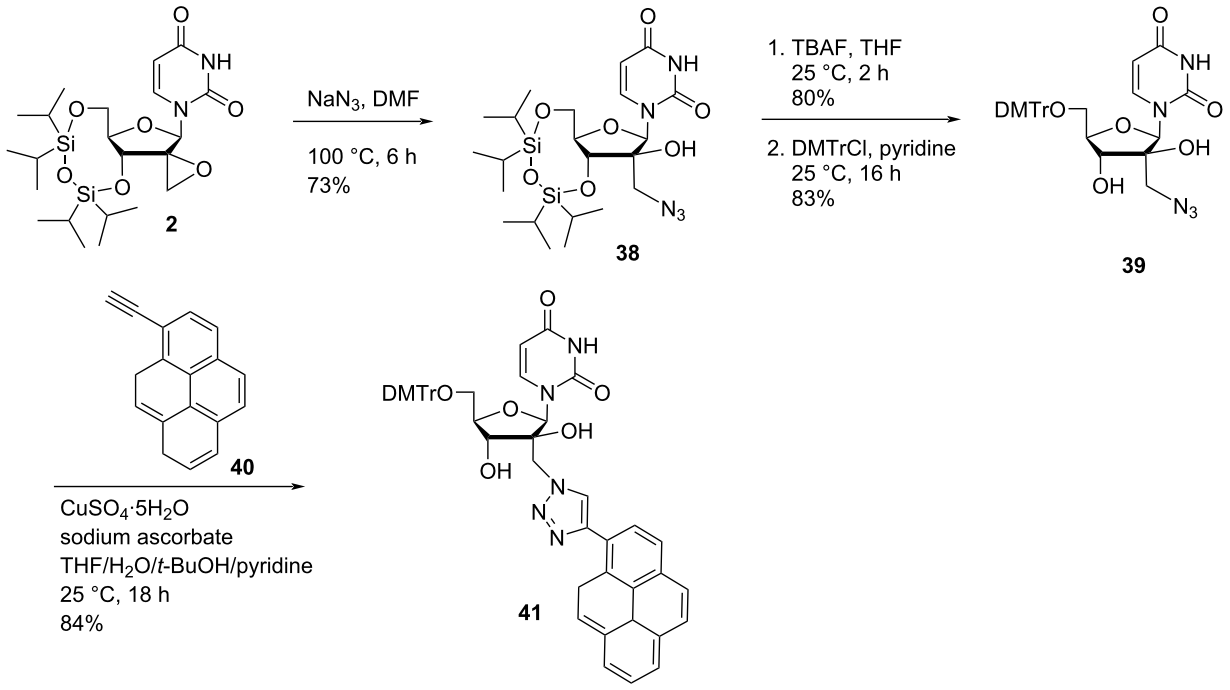
The double-headed nucleoside **41** was phosphitylated and then incorporated into oligonucleotides and was found to form highly stable DNA duplexes and three way junctions. There was a four-fold increase in the intensity of the pyrene excimer signal observed when an oligonucleotide containing two incorporations of the double-headed nucleoside **41** hybridized with an RNA target whereas the pyrene–pyrene excimer band almost vanished when the oligonucleotide was hybridized with a DNA target. The double-headed nucleoside **41** has potential in DNA invader probes as well as in RNA targeting and detection [23].

1,3-Furanosyl double-headed nucleosides

In this section, all double-headed nucleosides with furanosyl ring structures are collected. The first nucleobase is attached at the anomeric position of the furanosyl ring structure and the second nucleobase is connected to the C-3′ position with or without a linker (Figure 1).



Scheme 9: Synthesis of double-headed nucleoside **37**.



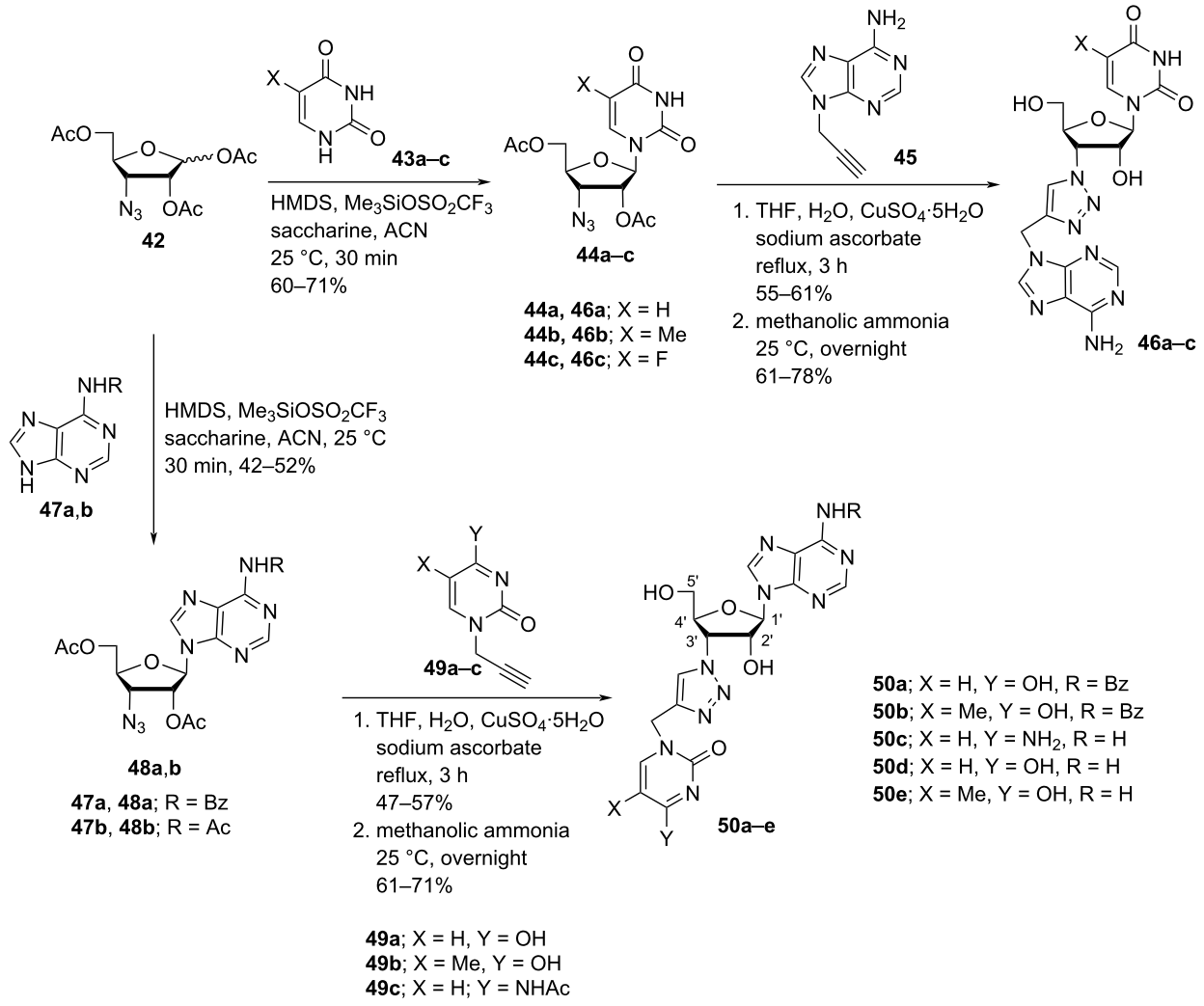
Scheme 10: Synthesis of the double-headed nucleoside 1-(5′-*O*-(4,4′-dimethoxytrityl)-2′-*C*-((4-(pyren-1-yl)-1,2,3-triazole-1-yl)methyl)arabinofuranosyl)uracil (**41**).

Leonaidas and co-workers [36] have synthesized 3′-(4-((adenine-9-yl)methyl)-1,2,3-triazol-1-yl)-substituted double-headed nucleosides of 1-(β-D-ribofuranosyl)uracil/thymine/5-fluorouracil **46a–c** and 3′-(4-((pyrimidin-1-yl)methyl)-1,2,3-triazol-1-yl)-substituted double-headed nucleosides of 9-(β-D-ribofuranosyl)adenine/N^6-benzoyladenine **50a–e**. The synthesis started from C-3-azidoribofuranose **42** which in turn was obtained from 1,2:5,6-di-*O*-isopropylidene-D-glucose [50]. Furanoside **42** was reacted with the silyl-protected nucleobases **43a–c** and **47a**,**b** in the presence of trimethylsilyl trifluoromethanesulfonate in acetonitrile to give the 3′-azido-3′-deoxy-β-D-ribonucleosides **44a–c** and **48a**,**b** via Vorbrüggen coupling reaction. The nucleosides were further reacted with propargylated nucleobases through a copper-catalyzed azide–alkyne cycloaddition (CuAAC) reaction followed by treatment with methanolic ammonia to give the C-3′-substituted double-headed ribofuranonucleosides **46a–c** and **50a–e** (Scheme 11) [36].

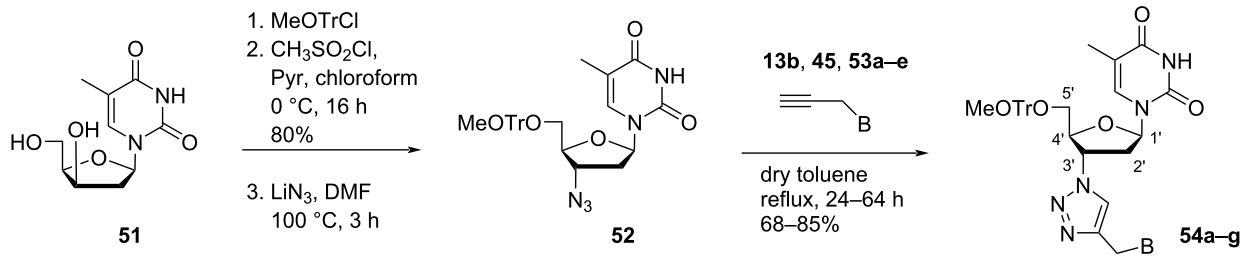
The double-headed nucleosides **46a–c** and **50a–e** were evaluated for their inhibitory potency towards RNase A and eosinophil-derived neurotoxin (EDN). Among all the nucleosides, the double-headed nucleoside **50c** showed a stronger preference for EDN than for ribonuclease A whereas all other derivatives were found to be more specific for ribonuclease A [36].

Lazrek et al. [51] synthesized C-3'-modified double-headed nucleosides **54a–g** where a 1,2,3-triazol ring acts as linker of the nucleobase and the sugar moiety. First, seven N^9/N^1-propargylpurine/pyrimidine nucleobases **13b**, **45**, and **53a–g** were synthesized by treating the nucleobases with propargyl bromide in the presence of K_2CO_3. The synthesis of compounds **54a–g** started with β-D-lyxofuranosylthymine (**51**), which was first methoxytritylated at the C-5′ primary hydroxy position followed by mesylation of the C-3′ secondary hydroxy position. The subsequent treatment with sodium azide in DMF afforded the corresponding nucleoside **52** [52]. Triazolylation of compound **52** with the nucleobases **13b**, **45**, and **53a–g** by refluxing the substrates in toluene afforded the targeted 5′-*O*-monomethoxytritylated nucleosides **54a–g** (Scheme 12) [51].

Vilarrasa and co-workers [47] synthesized 3′-uracil-1-yl and 3′-thymin-1-yl derivatives of 2′-deoxythymidine, i.e., compounds **59** and **60** starting from 5′-*O*-tritylthymidine (**55**). The tritylated thymidine **55** first was converted to the protected azide derivative **57** in two steps, followed by its reduction in the presence of tin(II) chloride, thiophenol and triethylamine and treatment with pyrrolidine in acetonitrile to afford the C-3'-aminonucleoside **58**. The reaction of this key intermediate with 3-ethoxypropenoyl isocyanate or 3-methoxy-2-methylpropenoyl isocyanate in a solvent mixture of benzene and DMF,



Scheme 11: Synthesis of triazole-containing double-headed ribonucleosides **46a–c** and **50a–e**.



compound	13b, 54a	53a, 54b	53b, 54c	53c, 54d	53d, 54e	53e, 54f	45, 54g
B	thymin-1-yl	uracil-1-yl	5-chlorouracil-1-yl	5-bromouracil-1-yl	5-fluorouracil-1-yl	5-iodouracil-1-yl	adenin-9-yl

Scheme 12: Synthesis of double-headed nucleosides **54a–g**.

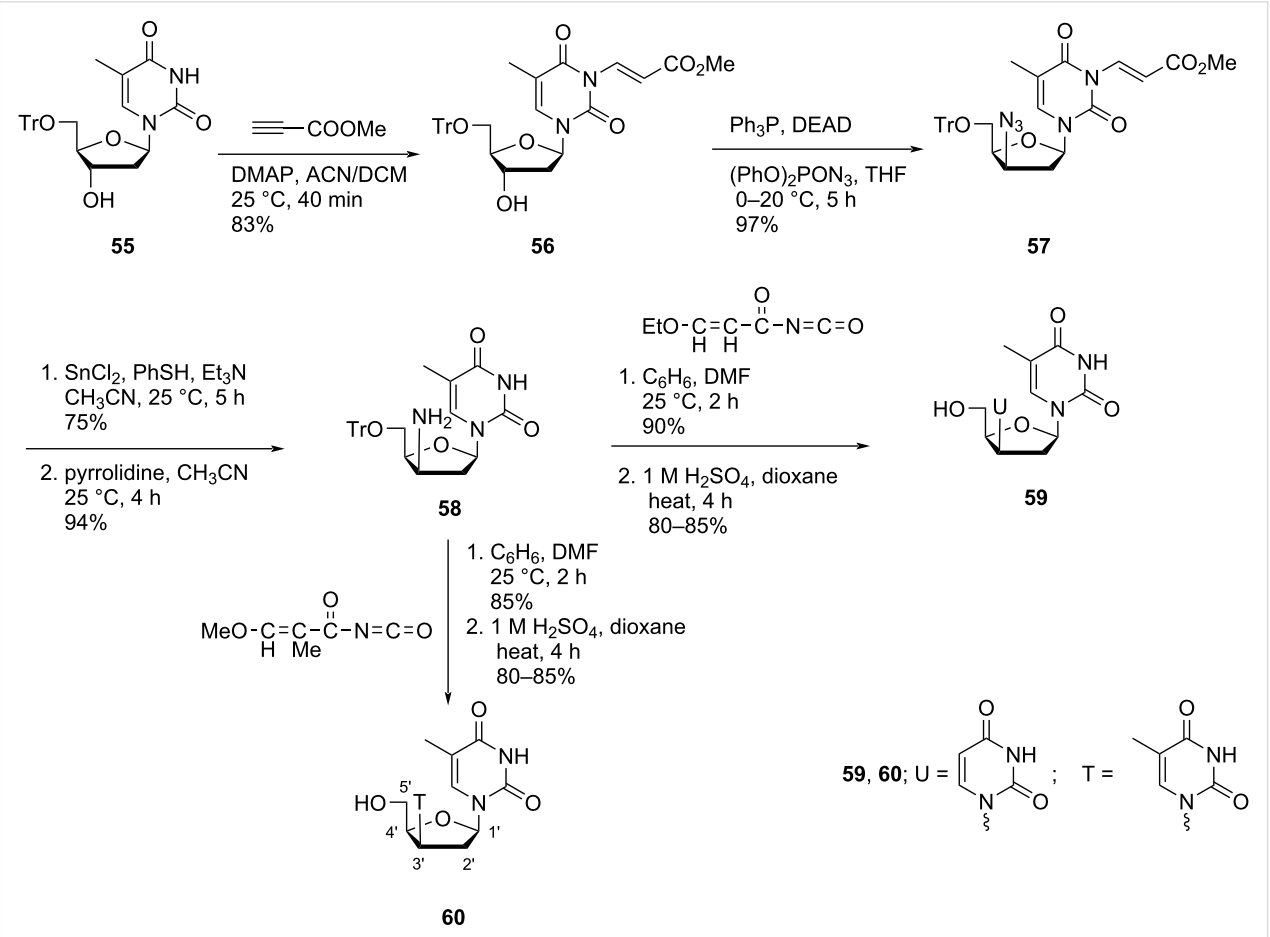
followed by acidification with sulfuric acid produced the desired nucleosides **59** and **60**, respectively (Scheme 13) [47].

Vilarrasa and co-workers [47] also synthesized the double-headed nucleosides **63** and **64** with downwards orientation of the additional nucleosides at the C-3′ position. The synthesis was carried out via formation of anhydride **61**. Azidation, followed by reduction of the corresponding nucleoside with tin chloride produced nucleoside **62** which was treated as a key intermediate for the production of the double-headed nucleosides **63** and **64**. Reaction of nucleoside **62** with 3-ethoxypropenoyl isocyanate or 3-methoxy-2-methylpropenoyl isocyanate in a solution mixture of benzene and DMF, followed by acidification with sulfuric acid produced nucleosides **63** and **64**, respectively (Scheme 14) [47].

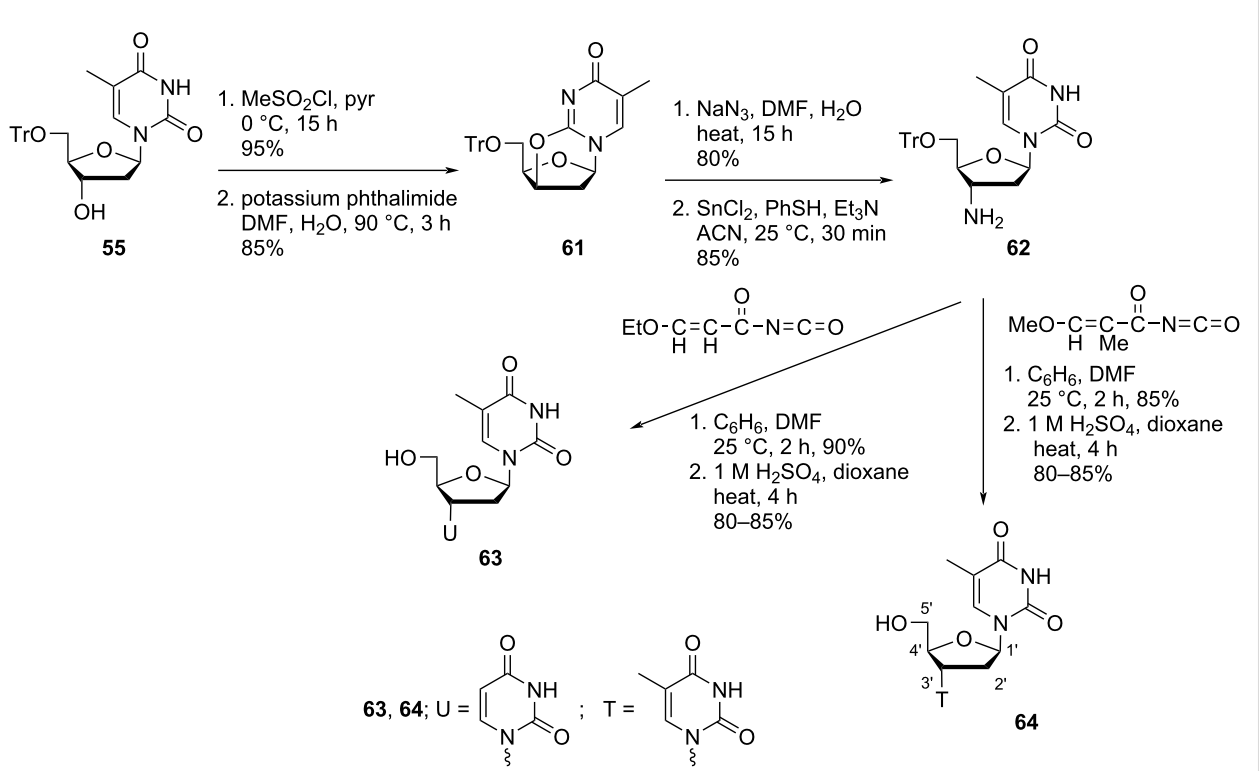
1,4-Furanosyl double-headed nucleosides

A literature search revealed two different categories of 1,4-furanosyl double-headed nucleosides. In the first category, the first nucleobase was a natural (attached at C-1′ position) and the second nucleobase was an aromatic moiety, which was attached at the C-4′ position without any linker (Figure 1). Whereas the second category of nucleosides contained first natural nucleobase at the C-1′ position and a second natural nucleobase attached at the C-4′ position with a methylene linker. The nucleosides of the second type may also contain a hydroxymethyl group at the C-4′ position.

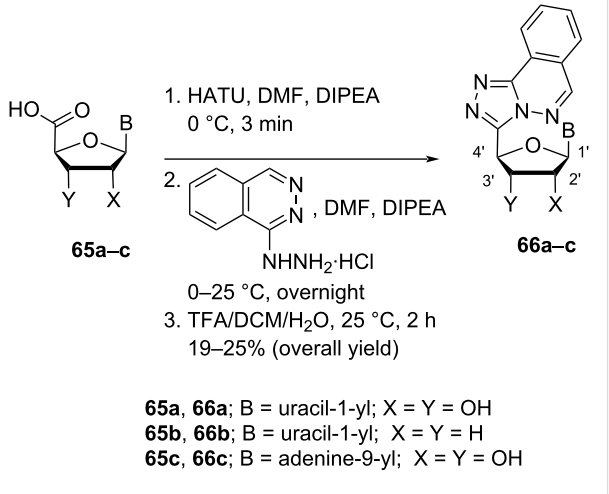
Torrence and co-workers [11] synthesized triazolophthalazine-substituted double-headed nucleosides **66a–c** from uridine/adenosine-5′-carboxylic acids **65a–c** which in turn were prepared through the (2,2,6,6-tetramethylpiperidin-1-oxyl) (TEMPO) and 1,1-bis(acetoxy)iodobenzene (BAIB)-assisted oxidation of the 5′-hydroxymethylene group in adenosine/uridine by following the methodology developed by Epp and Widlanski [53]. The nucleoside-5′-carboxylic acids **65a–c** were reacted with 2-(1*H*-7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) as coupling reagent followed by reaction with phthalazin-1-ylhydrazin hydrochloride in DMF in the presence of diisopropylethylamine (DIPEA) as base to afford the double-headed nucleosides **66a–c** (Scheme 15) [11].



Scheme 13: Synthesis of double-headed nucleosides **59** and **60**.



Scheme 14: Synthesis of the double-headed nucleosides **63** and **64**.

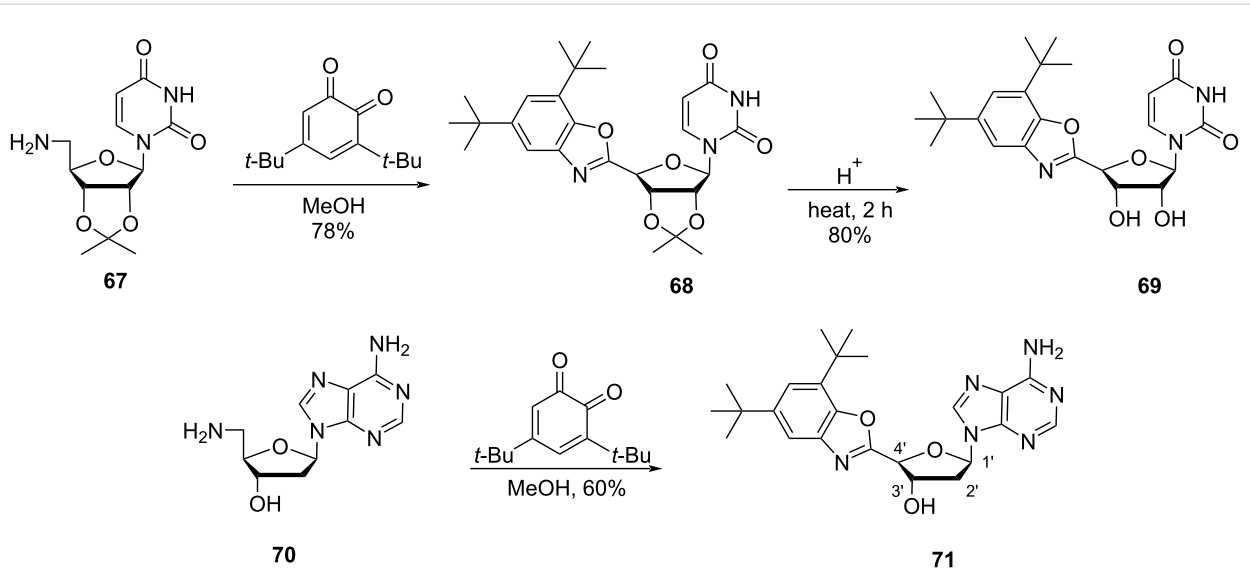


Scheme 15: Synthesis of double-headed nucleosides **66a–c**.

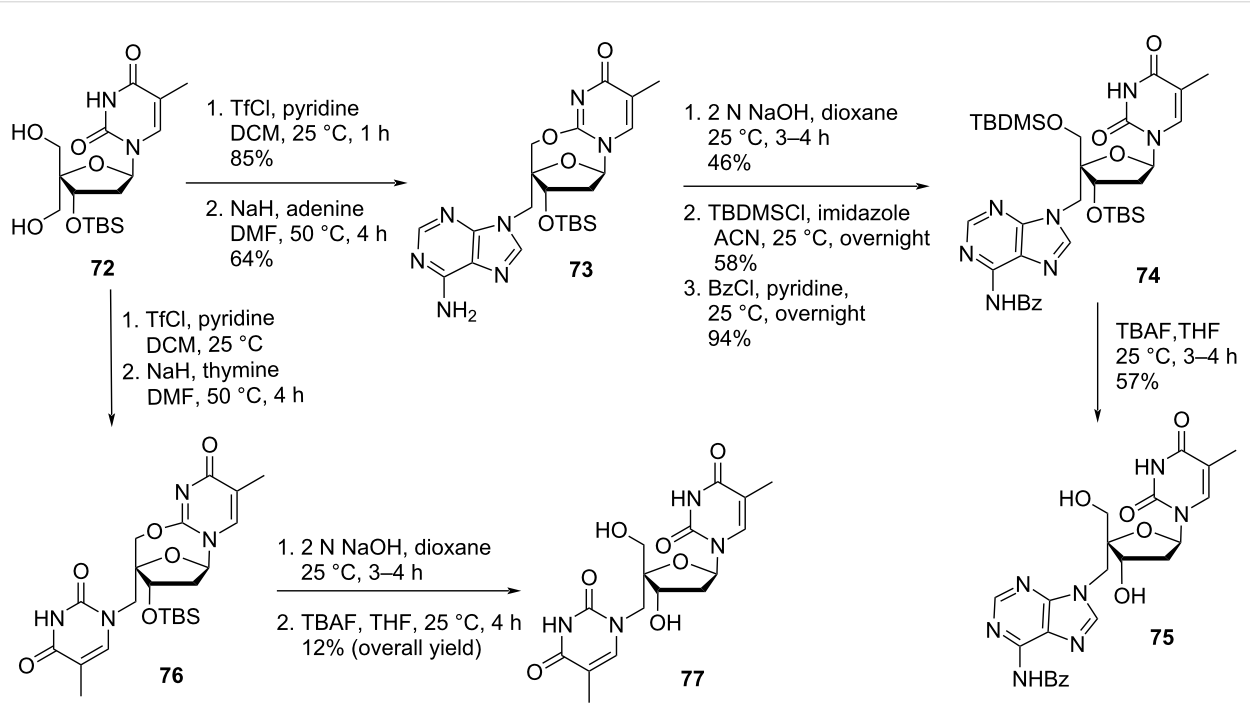
Timoshchuk and Hogrefe [12] have synthesized the double-headed nucleosides (*R*)-N^1-(4-(4,6-di-*tert*-butylbenzoxazol-2-yl)-2′,3′-*O*-isopropylidene-β-D-erythrofuranosyl)uracil (**69**) and (*R*)-N^9-(4-(4,6-di-*tert*-butylbenzoxazol-2-yl)-2′-deoxy-β-D-erythrofuranosyl)adenine (**71**) by the reaction of 3,5-di-*tert*-butyl-1,2-benzoquinone with 5′-amino-5′-deoxy-2′,3′-*O*-isopropylideneuridine (**67**) and 5′-amino-2′,5′-dideoxyadenosine (**70**). The unprotected double-headed nucleoside (*R*)-N^1-(4-(4,6-di-*tert*-butylbenzoxazol-2-yl)-β-D-erythrofuranosyl)uracil (**69**) was obtained by acidic hydrolysis of the intermediate benzoxazole derivative **68** (Scheme 16) [12].

Herdewijn and co-workers [54] synthesized the double-headed nucleoside monomers 4′-*C*-((N^6-benzoyladenin-9-yl)methyl)thymidine (**75**) and 4′-*C*-((thymin-1-yl)methyl)thymidine (**77**) starting from 3′-*O*-(*tert*-butyldimethylsilyl)-4′-(hydroxymethyl)thymidine (**72**) which was conveniently synthesized from thymidine in five steps as reported in the literature [54-56]. The nucleoside **72** was then converted into the corresponding triflate derivative which was further reacted with the nucleobases adenine or thymine to afford compounds **73** and **76**, respectively (Scheme 17) [54].

The *tert*-butyldimethylsilyl-protected (TBDMS) nucleoside **76** was first hydrolyzed using NaOH, which was followed by TBDMS deprotection using tetra-*n*-butylammonium fluoride (TBAF) in tetrahydrofuran (THF) to afford the double-headed nucleoside **77**. The TBDMS-protected nucleoside **73** was first hydrolyzed using NaOH followed by the reaction with TBDMSCl and benzoyl chloride to get the N^6-benzoyl-3',5'-*O*-diTBDMS-protected nucleoside **74**. Removal of the silyl-protecting groups in the double-headed nucleoside **74** with



Scheme 16: Synthesis of benzoxazole-containing double-headed nucleosides **69** and **71** from 5′-amino-5′-deoxynucleoside **67**.



Scheme 17: Synthesis of 4′-*C*-((*N*^{6}-benzoyladenin-9-yl)methyl)thymidine (**75**) and 4′-*C*-((thymin-1-yl)methyl)thymidine (**77**).

TBAF in THF resulted in the formation of the desired double-headed nucleoside **75** (Scheme 17) [54].

The double-headed nucleosides **75** and **77** were 4-methoxytritylated and phosphitylated following the standard procedures and incorporated into oligonucleotides. Extrahelical A-T base interactions were observed when these double-headed nucleoside monomers were placed in opposite strands of the duplex with separation of one regular base pair from each other [54].

1,5-Furanosyl double-headed nucleosides

In this category, the nucleosides contain the first natural nucleobase at the C-1′ position and the second natural nucleobase/aro-

matic moiety/heterocyclic ring attached at the C-5′ position with or without a linker (Figure 1).

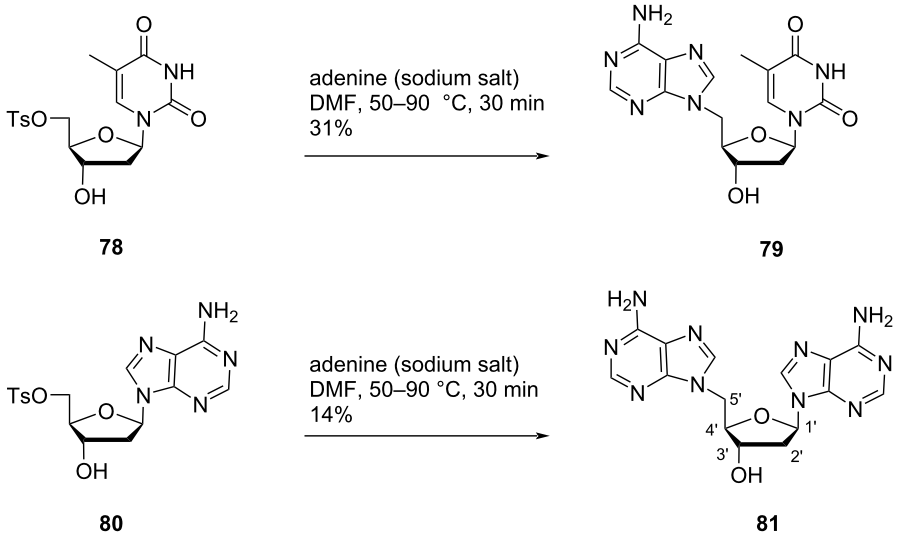
Shen and co-workers [57,58] proposed the synthesis of the double-headed nucleosides 5′-(adenine-9-yl)-5′-deoxythymidine (**79**) and 5′-(adenine-9-yl)-2′,5′-dideoxyadenosine (**81**) from 2′-deoxy-5′-*O*-tosylthymidine/adenosine **78** and **80**, respectively. The 2′-deoxy-5′-*O*-tosyl nucleosides were reacted with the sodium salt of adenine in DMF to afford the double-headed nucleosides **79** and **81**, respectively (Scheme 18) [57,58].

Žinić and co-workers [59] synthesized 5′-(5-iodouracil-1-yl)uridine (**85**), 5′-(5-iodouracil-1-yl)-5′-deoxyadenosine (**86**) and 5′-(uracil-1-yl)-5′-deoxyuridine (**87**) starting from the "reversed" 5-iodouracil-1-yl nucleoside **83** which in turn was synthesized by reacting the sodium salt of 5-iodouracil with isopropylidene-protected ribofuranoside **82** [55,59]. The reversed nucleoside **83** was next suitably protected to form nucleoside **84** which was then reacted with either uracil or N^6-benzoyladenine via Vorbrüggen's method of nucleobase coupling to produce the double-headed nucleosides **85** and **86**. Catalytic hydrogenolysis of the iodinated double-headed nucleoside **85** gave the nucleoside **87** (Scheme 19) [59].

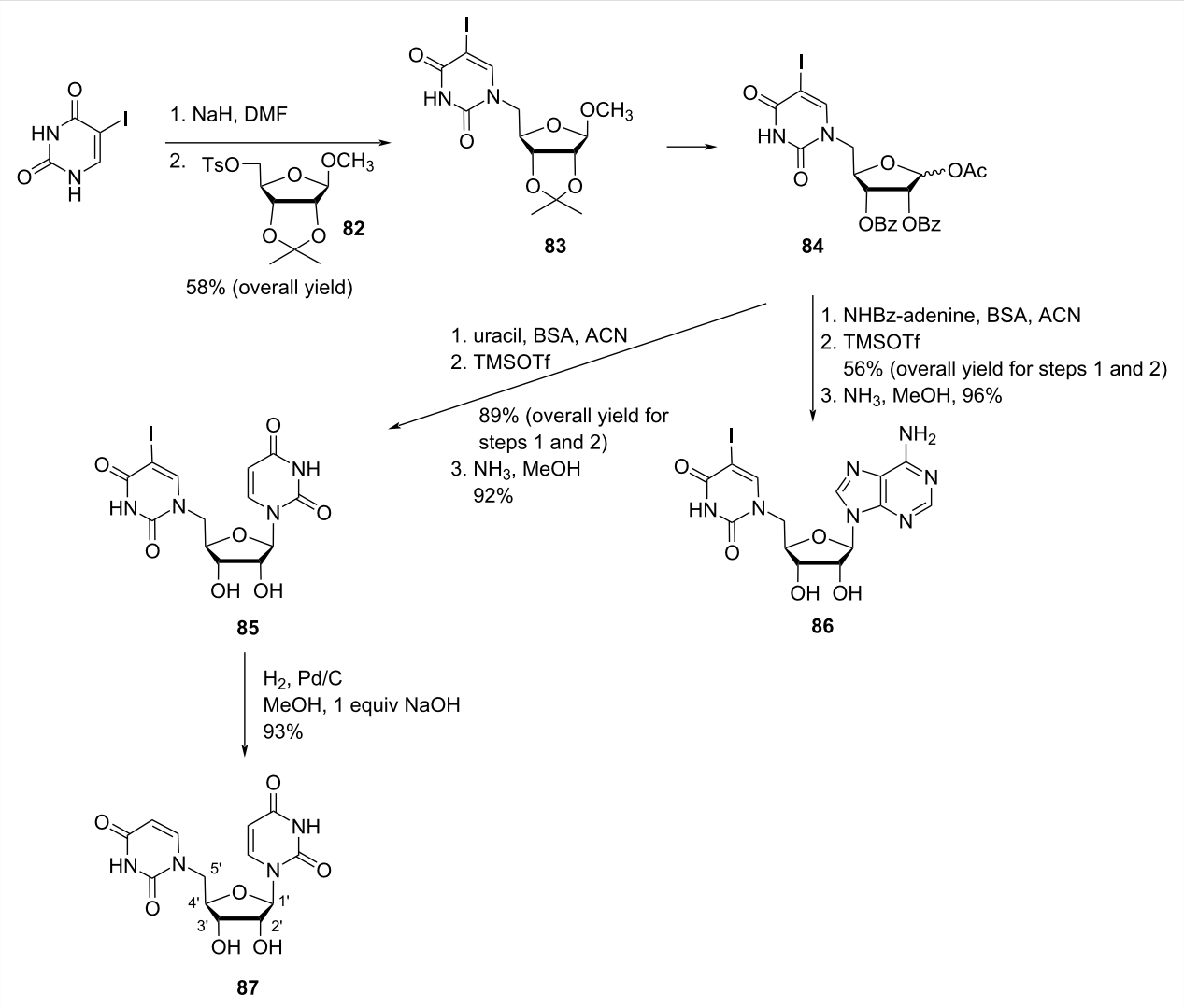
Horton and Tsai [13] synthesized double-headed nucleosides 2,6-dichloro-9-(2,3,5-tri-*O*-acetyl-5-*C*-(3-mesitylisoxazol-5-yl)-α-L-idopentofuranosyl)-9*H*-purine (**91**) and 2,6-dichloro-9-(2,3,5-tri-*O*-acetyl-5-*C*-(1-phenyl-1,2,3-triazol-4-yl)-β-D-glucopentofuranosyl)-9*H*-purine (**92**) starting from 3,5-di-*O*-acetyl-6,7-dideoxy-1,2-*O*-isopropylidene-L-ido/α-D-gluco-hept-6-ynofuranoses **88a**,**b** The L-ido- and D-gluco precursors **88a,b** were reacted with trifluoroacetic acid followed by acetic anhydride to afford the 1,2,3,5-tetra-*O*-acetyl nucleoside analogs **89a** and **89b**, respectively. Montgomery and Hewson base coupling reaction [60] of 1,2,3,5-tetra-*O*-acetyl nucleoside analogues **89a** and **89b** with 2,6-dichloropurine under acidic conditions resulted in the formation of mononucleoside analogs **90a,b**. The nucleoside **90a** was reacted with mesitylnitrile to give the double-headed nucleoside **91**, whereas nucleoside **90b** was reacted with phenylazide to give the double-headed nucleoside **92** (Scheme 20) [13].

Lazrek et al. [51] synthesized C-5′-modified double-headed nucleosides **96a–g**, where a 1,2,3-triazolo ring acted as the linker between the nucleobase and the sugar moiety. First, seven N^9/N^1-propargylpurine/pyrimidine nucleobases **13b, 45** and **53a–e** were synthesized by treating the nucleobases with propargyl bromide in the presence of K_2CO_3. Nucleoside **94** was synthesized from thymidine (**93**) which was first tritylated at the C-5′ primary hydroxy position followed by acetylation at the C-3′ secondary hydroxy group [61]. Next, detritylation and tosylation of the protected nucleoside **94** followed by treatment with lithium azide in DMF and saturated methanolic ammonia solution afforded nucleoside **95**. Refluxing of nucleoside **95** with **13b, 45** and **53a–e** in toluene produced the desired nucleosides **96a–g** (Scheme 21) [51].

Shaikh et al. [14] reported the synthesis of double-headed nucleosides where an aromatic moiety or a nucleobase is attached at the C-5′ position of the nucleoside. The synthetic methodology started with the 5′-epoxide **97**, which was synthesized from 3′-*O*-(*tert*-butyldimethylsilyl)thymidine in three steps, where the oxidation of the C-5′-hydroxy group followed



Scheme 18: Synthesis of double-headed nucleosides 5′-(adenine-9-yl)-5′-deoxythymidine (**79**) and 5′-(adenine-9-yl)-2′,5′-dideoxyadenosine (**81**).

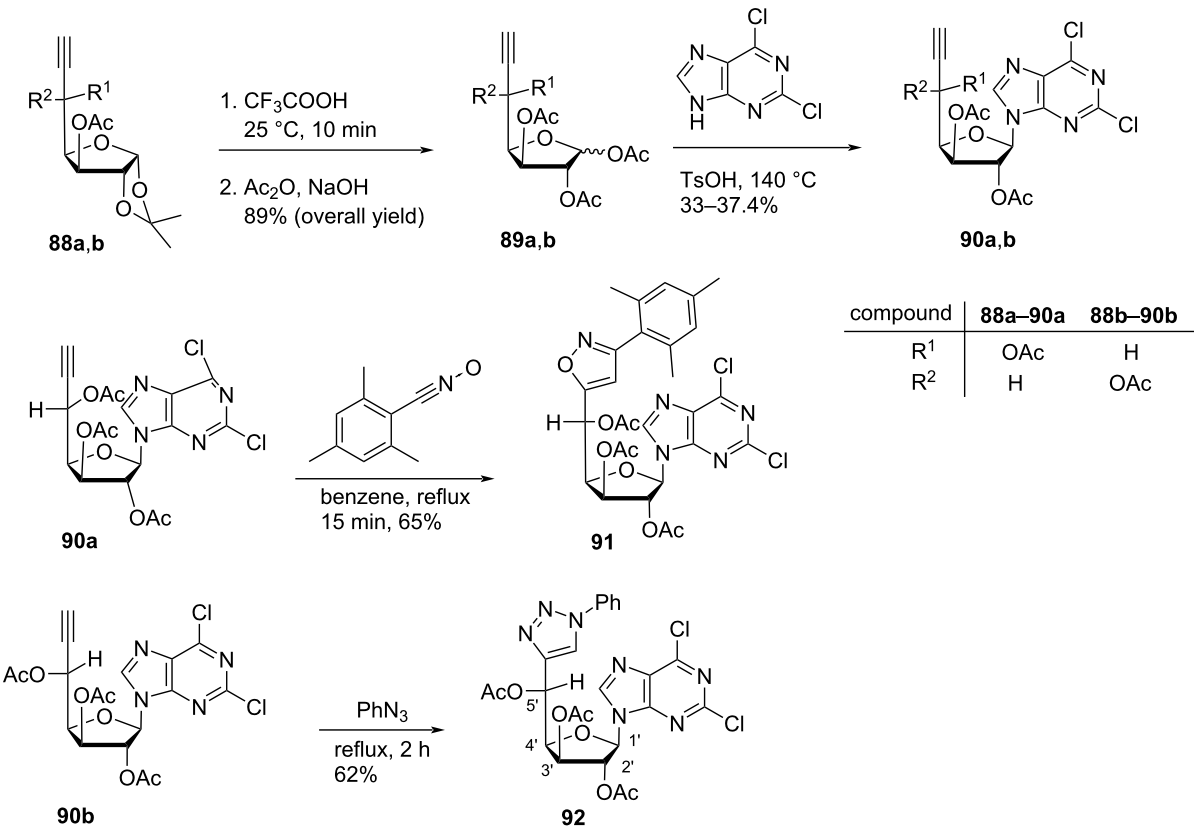


Scheme 19: Synthesis of double-headed nucleosides **85–87** via reversed nucleosides methodology.

by a Wittig reaction with methylenetriphenylphosphorane ($Ph_3P{=}CH_2$) produced the 5′-methylene derivative [62]. Finally, oxidation with *meta*-chloroperbenzoic acid (mCPBA) afforded the nucleoside **97**. Treatment of the nucleoside **97** with Grignard reagent PhMgBr in THF produced nucleoside **98**, whose secondary hydroxy group was protected by reaction with pixyl chloride to afford the nucleoside **99**. The removal of the *tert*-butyldimethylsilyl protecting group under standard conditions afforded the double-headed nucleoside **100** (Scheme 22) [14].

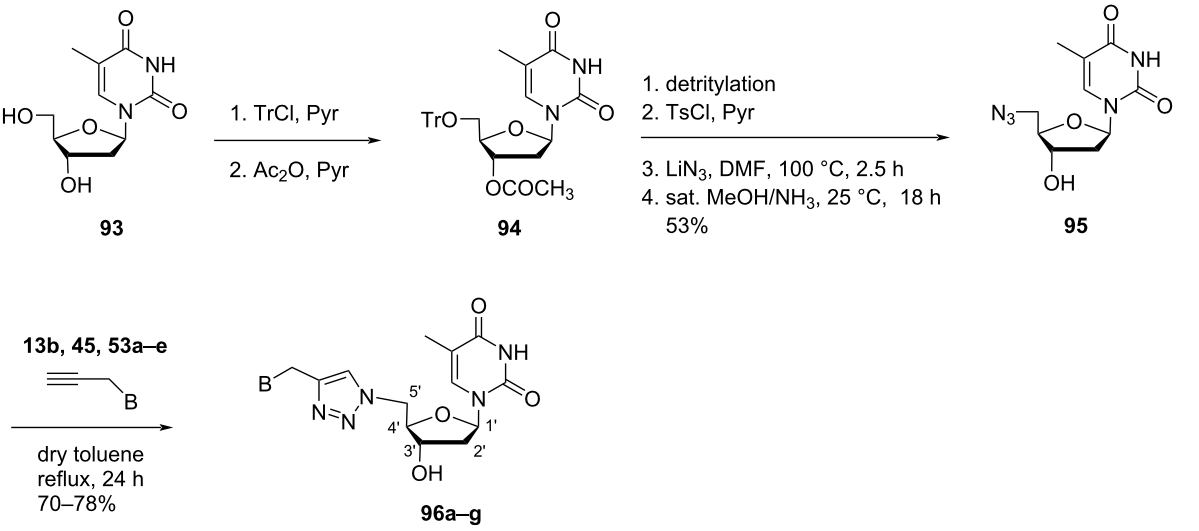
Opening of the epoxide ring in nucleoside **97** with sodium azide in DMF produced nucleoside **101**, whose secondary hydroxy group was protected by reaction with pixyl chloride to afford nucleoside **102**. The azido nucleoside **102** was a key intermediate, which was used for the synthesis of a variety of 1,2,3-triazolyl-linked double-headed nucleosides. Thus the treatment of azido nucleoside **102** with phenylacetylene in the presence of sodium ascorbate and copper sulfate in a solvent mixture of *t*-BuOH, water and pyridine, followed by the removal of the *tert*-butyldimethylsilyl protecting group gave nucleoside **103** (Scheme 22) [14]. Under similar reaction conditions, the treatment of nucleoside **102** with N^1-benzoyl-5-ethynyluracil followed by desilylation produced the double-headed nucleoside **104**, whereas the reaction of the azido nucleoside **102** with trimethylsilylacetylene (TMS-acetylene) followed by desilylation produced the nucleoside **105** (Scheme 23) [14].

Christensen et al. [31] synthesized the double-headed nucleoside 5′-(*S*)-*C*-(thymin-1-ylmethyl)-3′-*O*-(*tert*-butyldimethylsilyl)thymidine (**107**) by treating the 5′-olefinic nucleoside **106** with mCPBA in dichloromethane followed by reaction of the resulted product with thymine in the presence of K_2CO_3 in hot DMF. The targeted double-headed nucleoside 5′-(*R*)-*C*-(thymin-1-ylmethyl)-3′-*O*-(*tert*-butyldimethyl-silyl)thymidine (**108**) was



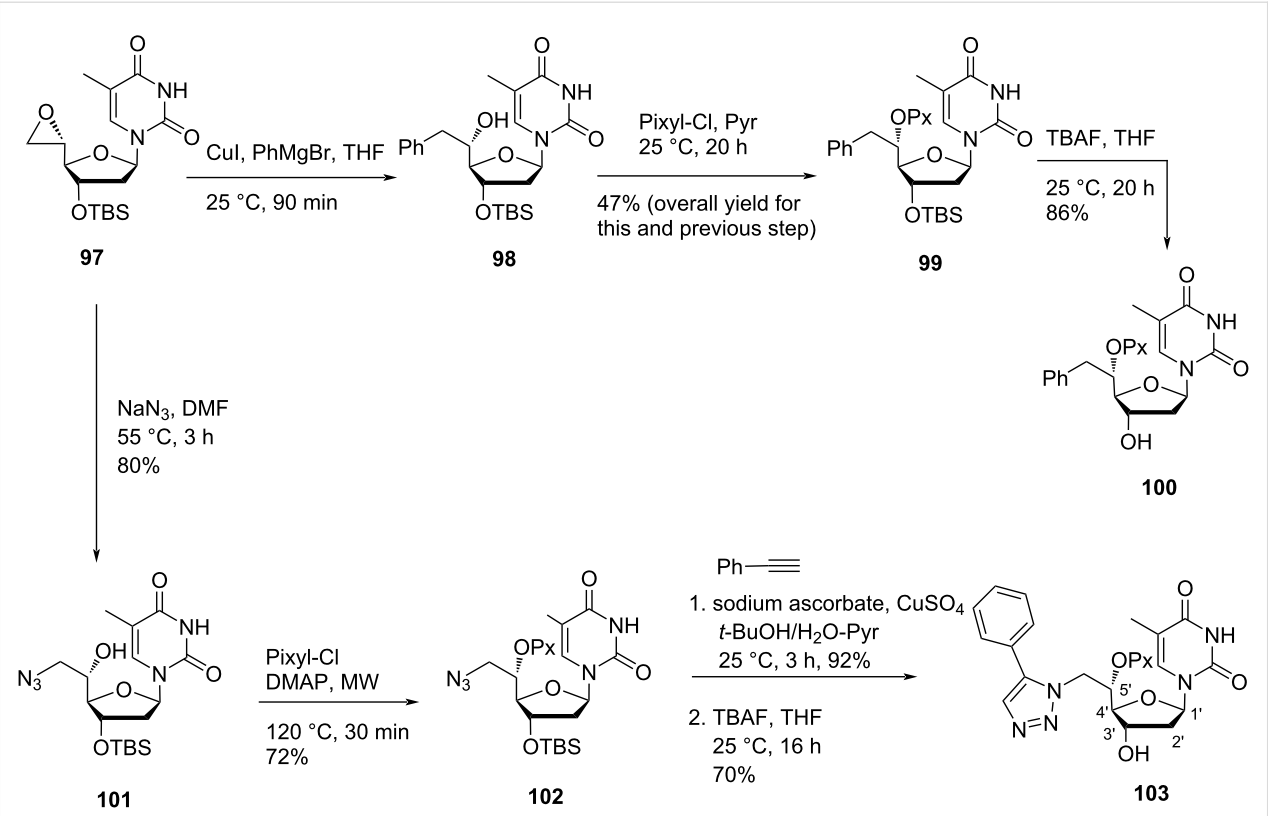
compound	88a–90a	88b–90b
R^1	OAc	H
R^2	H	OAc

Scheme 20: Double-headed nucleosides **91** and **92** derived from ω-terminal-acetylenic sugar derivatives **90a,b**.

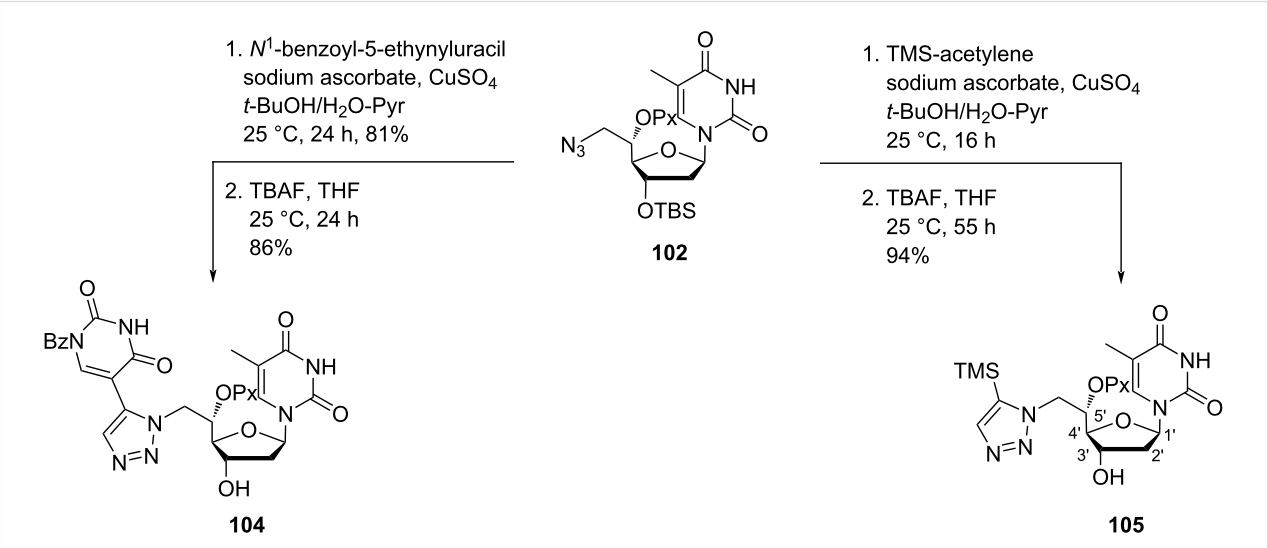


compound	13b, 96a	53a, 96b	53b, 96c	53c, 96d	53d, 96e	53e, 96f	45, 96g
B	thymin-1-yl	uracil-1-yl	5-chlorouracil-1-yl	5-bromouracil-1-yl	5-fluorouracil-1-yl	5-iodouracil-1-yl	adenin-9-yl

Scheme 21: Synthesis of double-headed nucleosides **96a–g**.



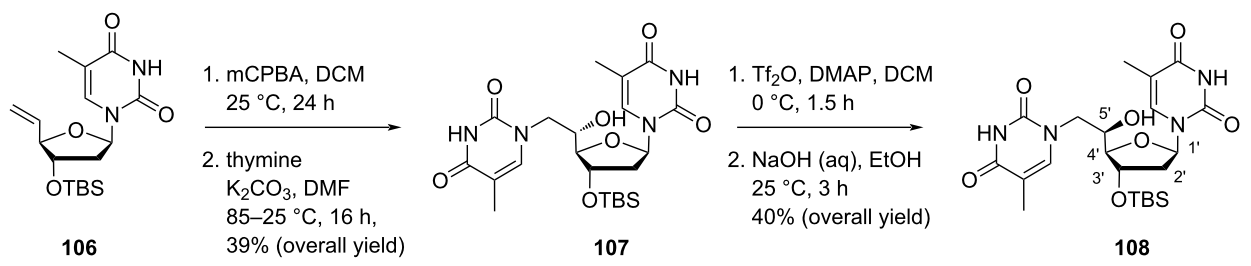
Scheme 22: Synthesis of double-headed nucleosides **100** and **103**.



Scheme 23: Double-headed nucleosides **104** and **105** with a triazole motif.

synthesized by treating the double-headed nucleoside **107** with triflic anhydride followed by basic hydrolysis (Scheme 24) [31].

Subsequently, the double-headed nucleoside **107** was incorporated into oligonucleotides [31,33,35,63,64] and when the duplex was generated with complementary DNA and RNA sequences, the additional nucleobase was positioned in the minor groove of the duplex. However, the presence of the additional nucleobase resulted in a thermal destabilization of the duplex as compared to unmodified duplexes. The introduction of two



Scheme 24: Synthesis of the double-headed nucleosides **107** and **108**.

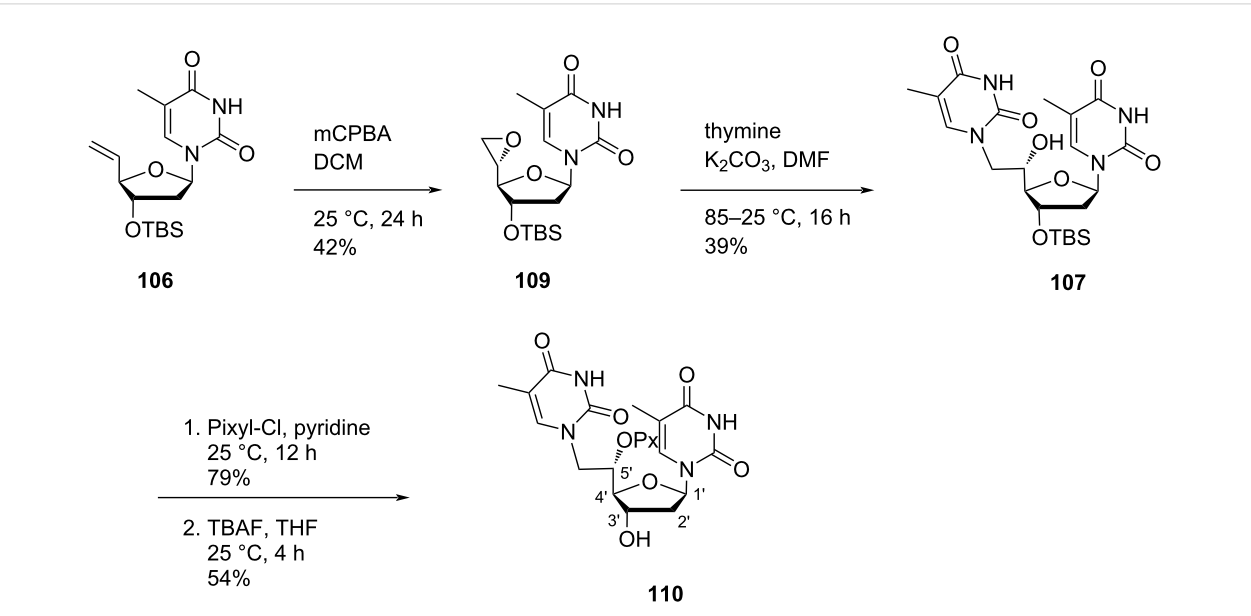
double-headed nucleosides in two complementary DNA sequences forming a DNA-zipper motif showed a stabilization of the duplex and increased base–base stacking interactions.

Nielsen and co-workers [30,31,65] synthesized double-headed nucleosides 5′-(*S*)-*C*-(thymine-1-yl/purin-9-yl)methyl-substituted double-headed nucleosides of thymidine **110–113** with additional nucleobase in the 5′(*S*)-C-position of thymidine. The double-headed nucleosides **110–113** were synthesized from the olefinic nucleoside **106**, which was converted into the epoxide **109** by treatment with mCPBA following the literature procedure [14,62]. The epoxide **109** so formed was reacted with thymine to afford nucleoside **107**, which on pixylation and removal of the *tert*-butyldimethylsilyl-protecting group in the presence of TBAF give double-headed nucleoside **110** (Scheme 25) [30,31,65].

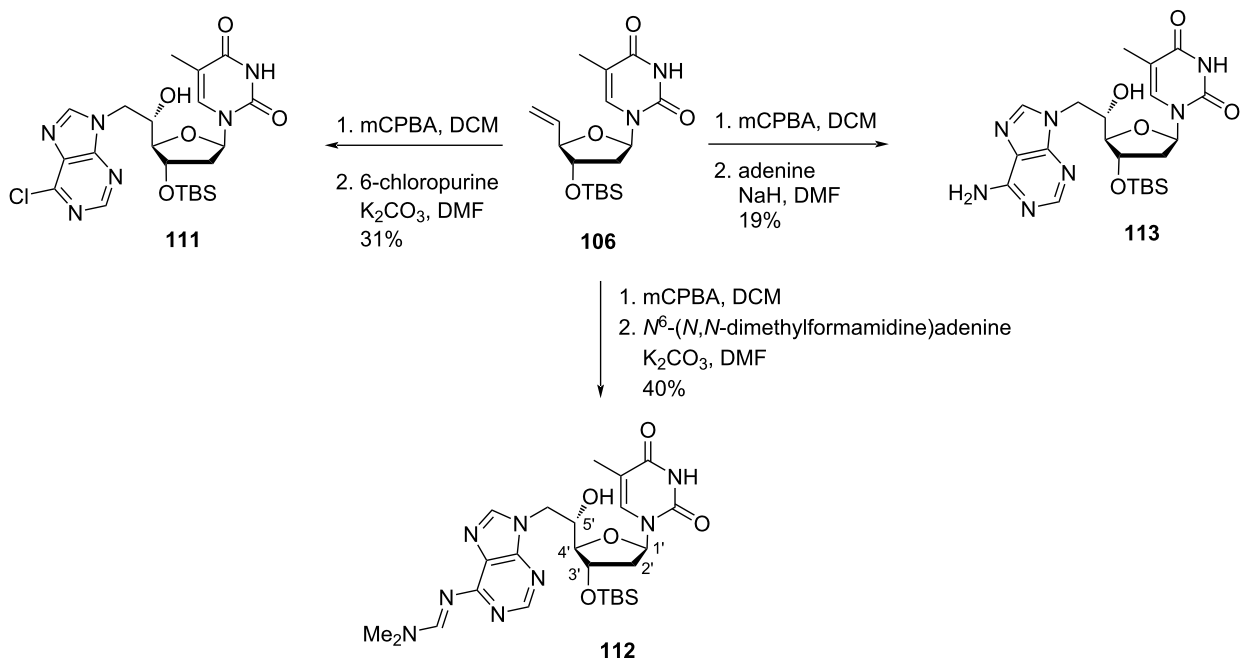
The epoxidation of the olefinic nucleoside **106** with mCPBA followed by reaction with 6-chloropurine, N^6-(*N*,*N*-dimethylformamidine)adenine, or adenine in the presence of K_2CO_3/NaH in hot DMF afforded the double-headed nucleosides **111–113** (Scheme 26) [30,31,65].

These nucleoside monomers were converted into phosphoramidites and then incorporated into oligonucleotide sequences, followed by thermal hybridization studies that indicated that the 5′-(*S*)-C-position is ideal for placing an additional nucleobase in the minor groove and interstrand stacking effects decreased with an increase in the length of the linker [31,65].

Nielsen and co-workers [43] synthesized the double-headed nucleoside 5′-*O*-pixyl-5′(*S*)-*C*-(4-(thymin-1-yl-methyl)-1,2,3-triazol-1-yl)methylthymidine (**114**) from 3′-TBS-protected 5′-(*S*)-*C*-azidomethylthymidine **102** which was synthesized from 3′-*O*-TBS-protected thymidine [31,65,66]. The nucleoside azide **102** was then reacted with propargylated thymine via CuAAC reaction, and subsequent removal of TBS group in the presence of TBAF and THF afforded the double-headed nucleoside **114**



Scheme 25: Synthesis of double-headed nucleoside **110** with additional nucleobase in 5′-(*S*)-*C*-position joined through methylene linker.



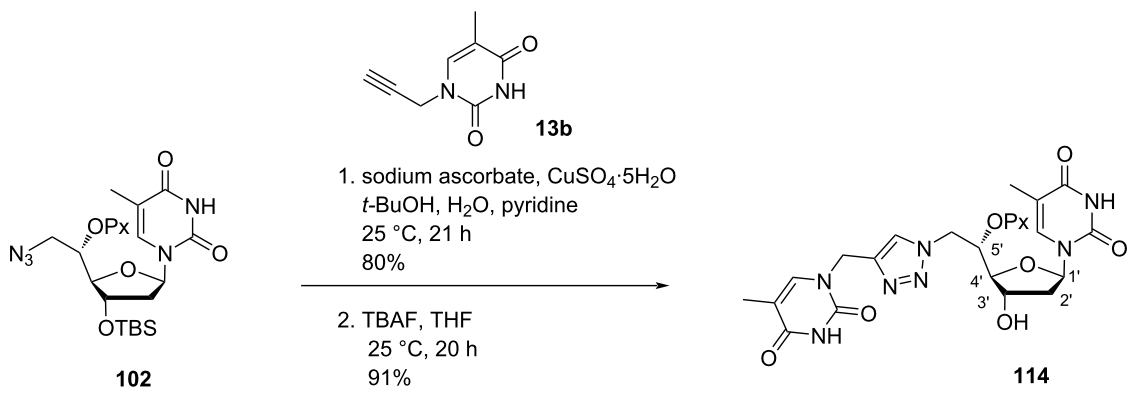
Scheme 26: Synthesis of double-headed nucleosides **111–113** with additional nucleobases in the 5′-(*S*)-*C*-position joined through methylene linkers.

where a triazole linker connected the additional thymine to the 5′-position of thymidine (Scheme 27) [43].

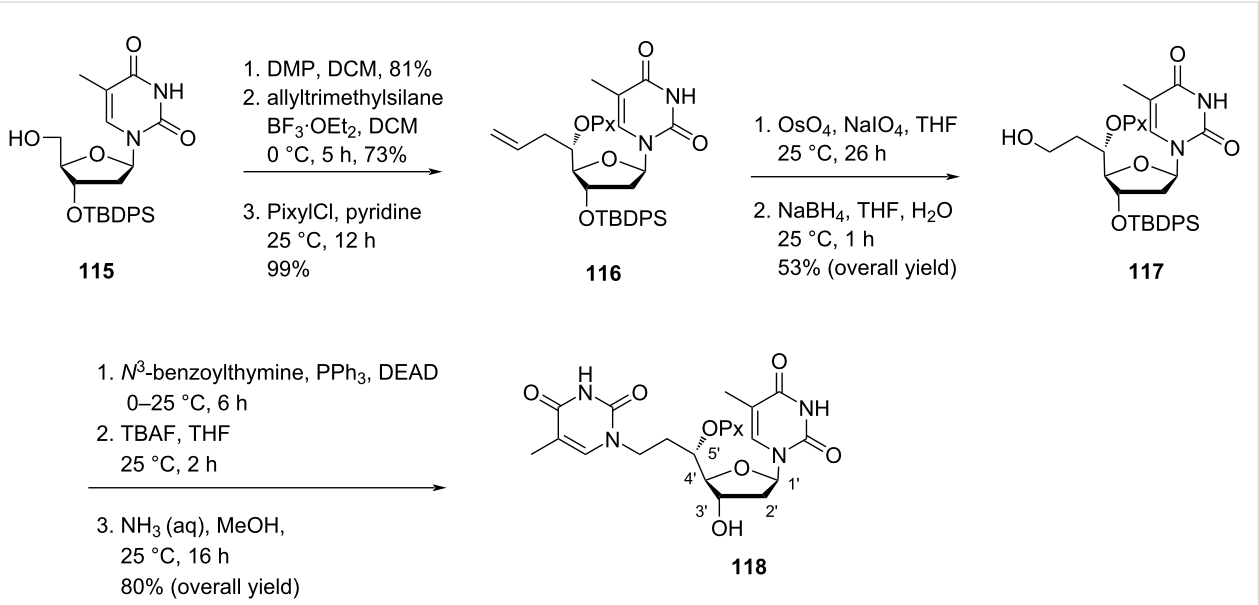
The incorporation of the double-headed nucleoside monomer **114** into oligonucleotides failed to stabilize three-way junctions [43] which is contrary to the double-headed nucleoside **11** which stabilized three-way junction very efficiently [35].

Nielsen and co-workers [30,65] synthesized the double-headed nucleoside 5′-(*S*)-*C*-(2-(thymine-1-yl)ethyl)thymidine (**118**) with an additional nucleobase at the 5′-(*S*)-*C*-position of thymidine. Double-headed nucleoside **118** was synthesized starting from 3′-*tert*-butyldiphenylsilyl (TBDPS)-protected thymidine **115** which was converted into the pixylated 5′(*S*)-*C*-allyl-substituted nucleoside **116** following previously reported procedures [65]. Then, the nucleoside **116** was converted into the primary alcohol **117** by treatment with OsO_4 and oxidative cleavage by $NaIO_4$ followed by reduction using $NaBH_4$. The primary alcohol **117** was further converted into nucleoside monomer **118** by introduction of the second nucleobase thymine through Mitsunobu reaction followed by deprotection steps in the presence of TBAF and methanolic ammonia (Scheme 28) [30,65].

The nucleoside monomer **118** was phosphitylated and then incorporated into oligodeoxynucleotides but stabilization in the



Scheme 27: Synthesis of double-headed nucleoside **114** by click reaction.



Scheme 28: Synthesis of double-headed nucleosides **118** with an additional nucleobase at the 5′-(*S*)-C-position.

secondary structures due to additional thymine in combination with the ethylene linker in double-headed nucleoside **118** was not observed because of an increase in the length of the linker which is contrary to the double-headed nucleoside **110** which stabilized secondary structures very well due to shorter length of linker [30,31,65].

Bicyclic double-headed nucleosides

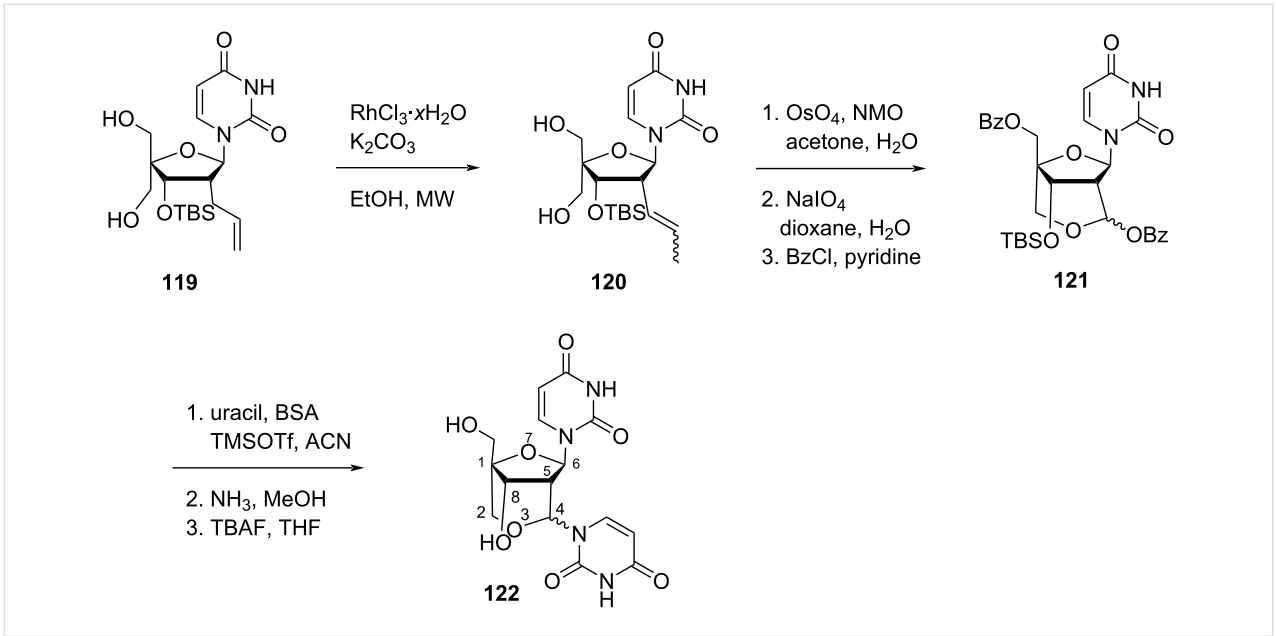
In this section we have included the double-headed nucleoside monomers, which have a locked nucleic acid type conformation and the additional nucleobase is attached at one of the carbon or nitrogen atoms constituting the bridge (Figure 1). All examples discussed herewith are constituted by furanosyl carbohydrate moiety.

Nielsen and co-workers [67] synthesized the bicyclic double-headed nucleoside (1*R*,4*R*/*S*,5*R*,6*R*,8*S*)-8-hydroxy-1-hydroxymethyl-4,6-di(uracil-1-yl)-3,7-dioxabicyclo[3.2.1]octane (**122**), where the additional nucleobase is attached at the bridge between C-2′ and C-4′. The synthesis of the aimed nucleoside **122** started from the nucleoside **119** which in turn was synthesized from uridine in six steps following a literature procedure [68]. The olefinic nucleoside **119** was subjected to a $RhCl_3$-mediated allyl rearrangement to give nucleoside **120** as a mixture of *E*/*Z* isomers. Further, double bond cleavage of nucleoside **120** followed by benzoylation produced the benzoic acid ester of the hemiacetal analogue **121** which was finally converted into double-headed nucleoside **122** via Vorbrüggen coupling reaction followed by deprotection using methanolic ammonia and TBAF (Scheme 29) [67].

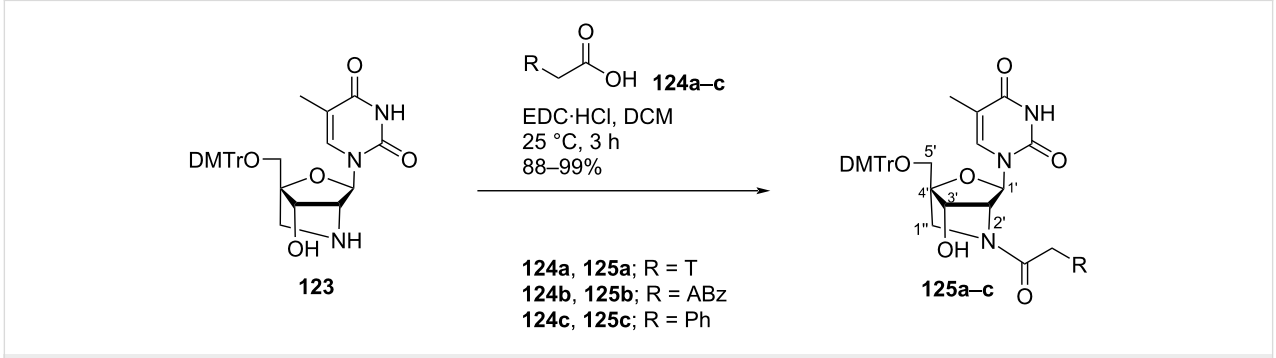
Madsen and co-workers [69] synthesized N^2-(thymin-1-ylacetyl)-, N^2-(N^6-benzoyladenin-9-ylacetyl)-, and N^2-phenylacetyl-substituted double-headed nucleosides of 1-(2′-amino-2′-deoxy-5-*O*-(4,4′-dimethoxytrityl)-2′-*N*,4′-*C*-methylene-β-D-ribofuranosyl)thymine **125a–c** by the reaction of 5′-*O*-dimethoxytritylated 2′-amino-LNA thymine nucleoside **123** with acetic acid derivatives, i.e., (thymin-1-yl)acetic acid (**124a**), (N^6-benzoyladenin-9-yl)acetic acid (**124b**), and phenylacetic acid (**124c**) in the presence of EDC·HCl as condensation reagent (Scheme 30) [69].

The double-headed nucleoside monomers **125a–c** were incorporated into oligodeoxyribonucleotides via phosphoramidite derivatization of the C-3′ hydroxy group present in the moiety. The oligonucleotides thus synthesized were found to stabilize the duplex formed with complementary DNA [69].

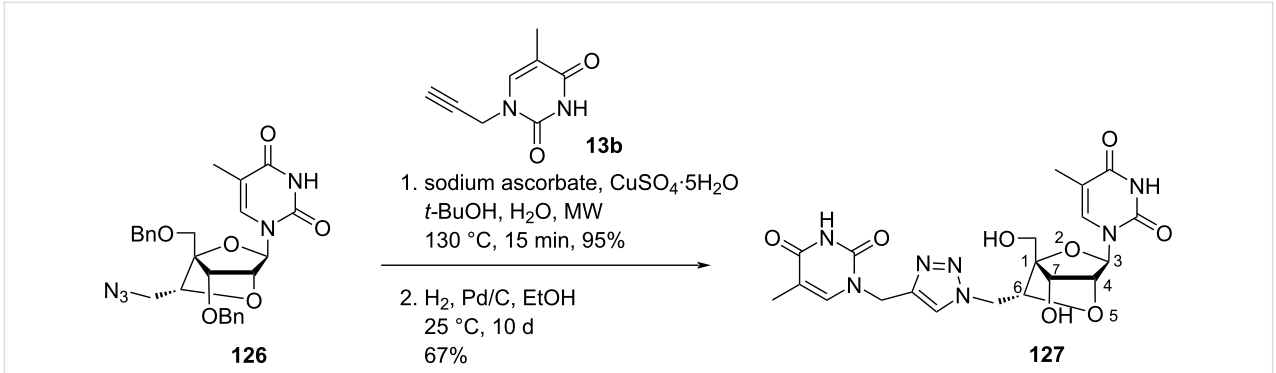
Nielsen and co-workers [43,70] synthesized the double-headed nucleoside (1*S*,3*R*,4*R*,6*R*,7*S*)-7-hydroxy-1-(hydroxymethyl)-3-(thymin-1-yl)-6-(4-(thymin-1-ylmethyl)-1,2,3-triazol-1-yl)methyl-2,5-dioxabicyclo[2.2.1]heptane (**127**). The CuAAc reaction of nucleoside azide **126** with propargylated thymine followed by Pd-catalyzed debenzylation under hydrogen atmosphere resulted in the formation of the double-headed nucleoside **127**. In nucleoside **127** the triazole ring is attached to the additional thymine moiety via a methylene linker and connected to the 6′-position of an LNA-thymidine monomer, via another methylene linker (Scheme 31) [43,70]. Interestingly, the incorporation of the double-headed nucleoside **127** into oligonucleotides failed to stabilize three-way junctions [43].



Scheme 29: Synthesis of bicyclic double-headed nucleoside **122**.



Scheme 30: Synthesis of double-headed nucleosides **125a–c** derived from 2′-amino-LNA.



Scheme 31: Double-headed nucleoside **127** obtained by click reaction.

Base to base double-headed nucleosides

Base to base double-headed nucleosides contain an additional natural nucleobase/substituted or unsubstituted phenyl moiety/polyaromatic moiety/carbocyclic moiety/heterocyclic moiety attached at the C-5 position of the pyrimidine nucleobase (first nucleoside) with or without a linker. The sugar moieties associ-

ated with these nucleosides were either a 2′-deoxyribofuranosyl moiety or a bicyclic moiety (Figure 1).

Nielsen and co-workers [71] synthesized the double-headed nucleoside 5′-*O*-(4,4′-dimethoxytrityl)-5-(thymin-1-yl)methyl-2′-deoxyuridine (**130**) with thymine attached to the C-5 position of 2′-deoxyuridine through a methylene linker. The double-headed nucleoside **130** was synthesized from 3′,5′-*O*-diacetyl-5-formyl-2′-deoxyuridine (**128**) which was reduced in the presence of $NaBH_4$ followed by the treatment with MsCl in pyridine to get the nucleoside salt **129**. Next, the pyridinium group was replaced by an N^3-protected thymine in basic medium followed by removal of the protecting groups and the selective DMTr protection of the C-5′-hydroxy group (Scheme 32) [71].

The double-headed nucleoside monomer 5′-*O*-(4,4′-dimethoxytrityl)-5-(thymin-1-yl)methyl-2′-deoxyuridine (**130**) was converted into the corresponding phosphoramidite at the C-3′-hydroxy group and then incorporated into oligonucleotides and was found to decrease the thermal stability of the duplexes [71].

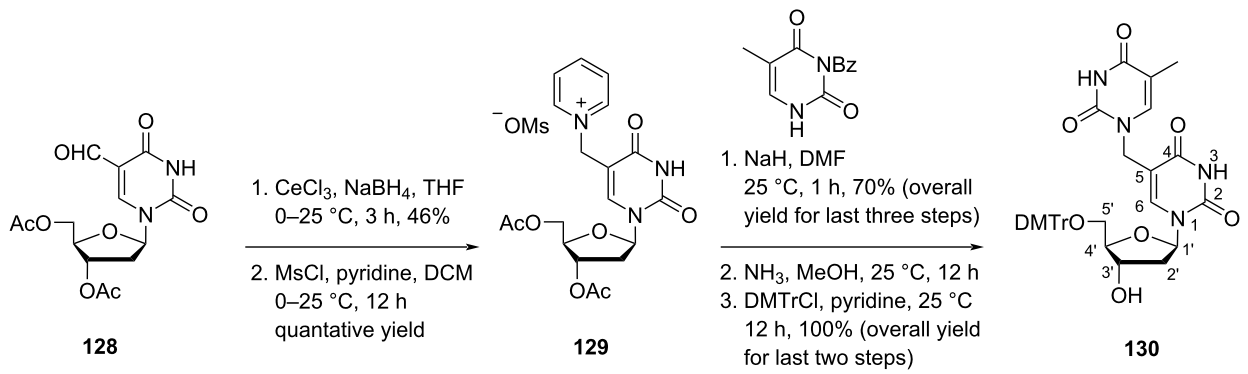
Nielsen and co-workers [72] synthesized 5-(3-(thymin-1-yl)propyn-1-yl)-, 5-(3-(N^4-acetylcytosin-1-yl)propyn-1-yl)-, 5-(3-(N^6-benzoyladenin-9-yl)propyn-1-yl), and 5-(3-(N^2-isobutyrylguanin-9-yl)-substituted double-headed nucleosides of 5′-*O*-DMTr-protected 2′-deoxyuridine (**132a–d**) and 5′-*O*-DMTr-protected N^4-(dimethylaminomethylene)-2′-deoxycytidine (**134a–d**) by using Sonogashira cross coupling reaction between the propargylated nucleobases, i.e., 1-propargylthymine, N^4-acetyl-1-propargylcytosine, N^6-benzoyl-9-propargyladenine, and N^2-isobutyryl-9-propargylguanine (**13a–d**) and 5′-*O*-DMTr-protected 5-iodo-2′-deoxyuridine (**131**) and N^4-(dimethylaminomethylene)-2′-deoxycytidine (**133**) in 65–81% yield (Scheme 33).

The synthesized double-headed nucleosides were phosphitylated and incorporated into oligonucleotides and the melting temperatures were evaluated against unmodified DNA strands. Oligonucleotides with fourteen consecutive incorporations of different double-headed nucleosides were synthesized and the DNA duplexes showed increased stability owing to increased stacking interactions among the nucleobases of the opposite strands [72]. Molecular dynamics simulations demonstrated the exposure of Watson–Crick/Hoogsteen faces of additional nucleobases for their recognition in the major groove.

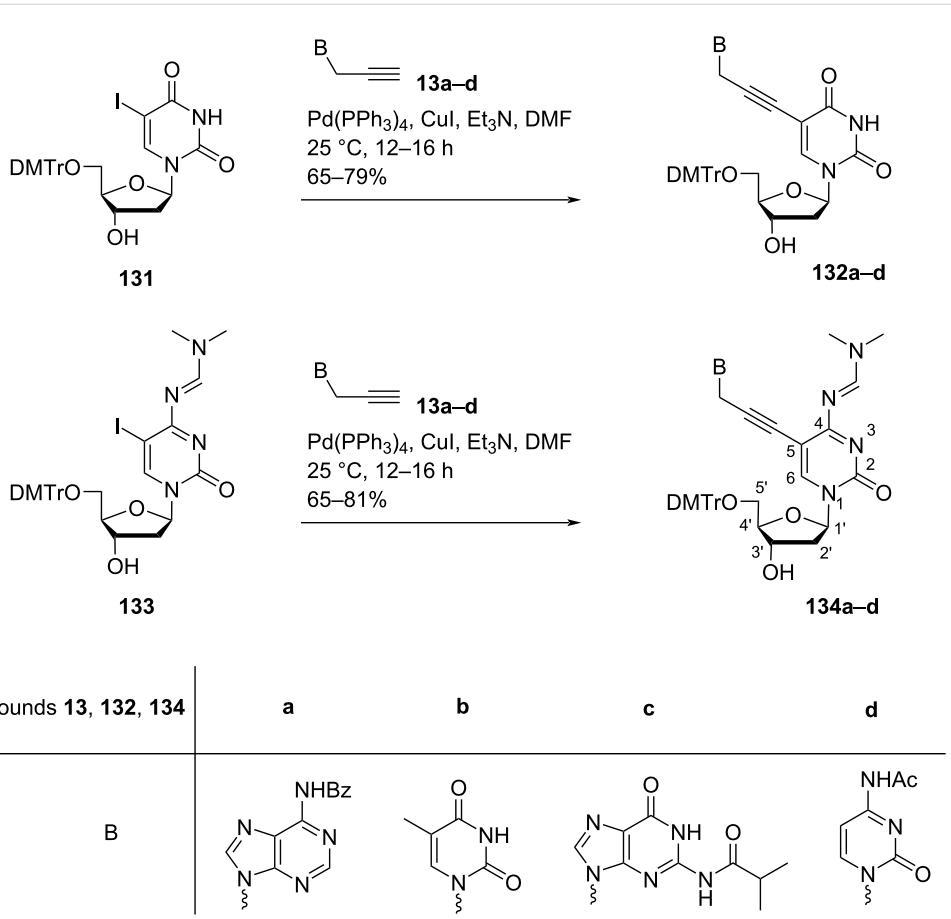
Sharma and co-workers [73] synthesized 5-(3-(thymin-1-yl)phenyl)- and 5-(4-(thymin-1-yl)phenyl)-substituted double-headed nucleosides of 5′-*O*-dimethoxytrityl-2′-deoxyuridine (**137**, **138**) from 5′-*O*-DMTr-2′-deoxy-5-iodouridine (**135**). Boronic esters N^1-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl) and N^1-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)-substituted N^3-benzoylthymine (**136a,b**) were synthesized by N^3-benzoylthymine from the procedure given by Gothelf and co-workers [74,75]. The boronic esters (**136a,b**) were coupled with 5′-*O*-DMTr-2′-deoxy-5-iodouridine (**135**) via Suzuki coupling to give double-headed nucleosides **137** and **138** (Scheme 34) [73].

The double-headed nucleosides 5′-*O*-dimethoxytrityl-5-(3-(thymin-1-yl)phenyl)ethynyl-2′-deoxyuridine (**140**) and 5′-*O*-dimethoxytrityl-5-(4-(thymin-1-yl)phenyl)ethynyl-2′-deoxyuridine (**141**) were synthesized via a Sonogashira cross coupling reaction between the N^1-(3/4-iodophenyl)thymine derivatives **136c** and **136d** and 2′-deoxy-5-ethynyluridine derivative **139** (Scheme 35) [75].

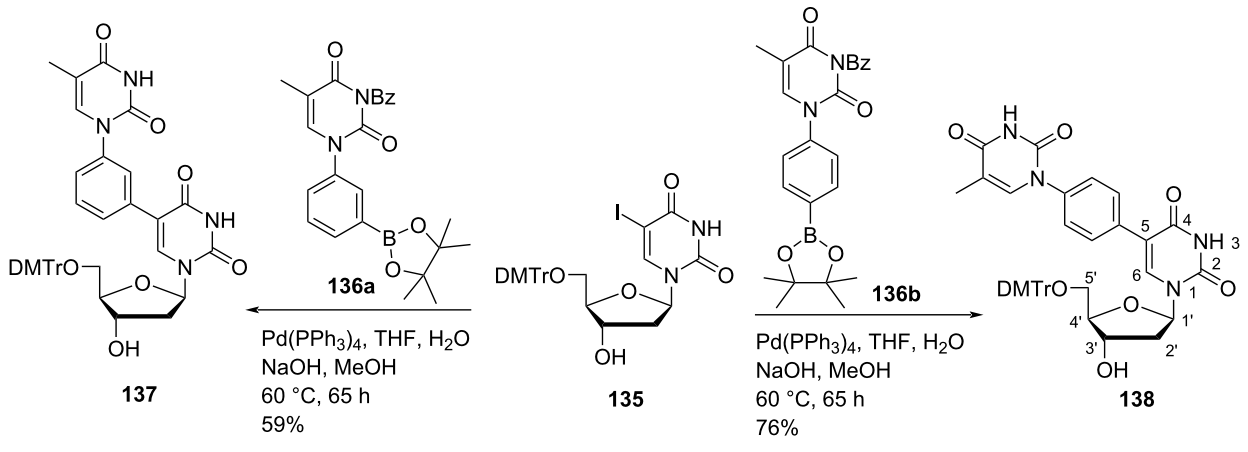
All four nucleoside monomers were converted into phosphoramidites and then introduced into oligonucleotides. The thermal stability of DNA:DNA and DNA:RNA duplexes was determined and it was found that duplexes with a *meta*-substitution and a phenylacetylene linker were more stable than the corresponding *para*-substituted and phenyl-linker containing derivatives.



Scheme 32: Synthesis of double-headed nucleoside **130**.

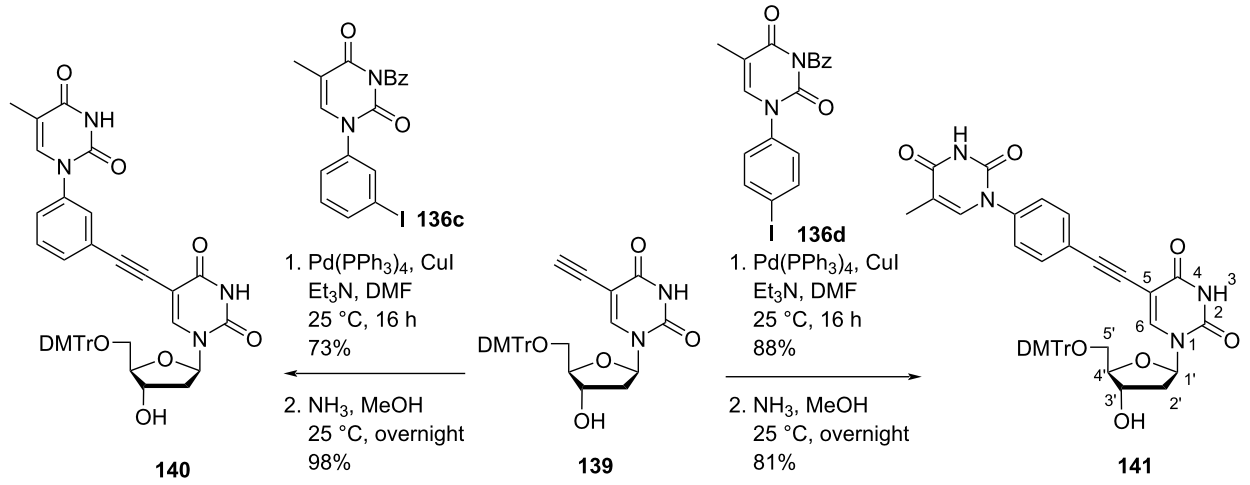


Scheme 33: Double-headed nucleosides **132a–d** and **134a–d** synthesized by Sonogashira cross coupling reaction.

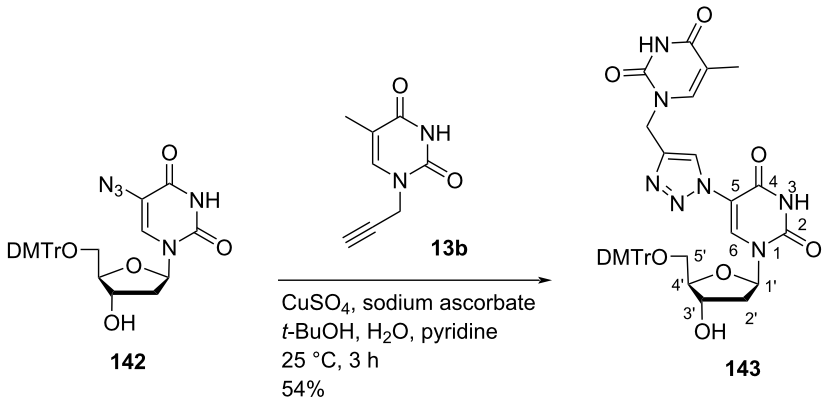


Scheme 34: Synthesis of double-headed nucleosides **137** and **138** via Suzuki coupling.

Nielsen and co-workers [71] synthesized the double-headed nucleoside 5′-*O*-(4,4′-dimethoxytrityl)-5-(4-(thymin-1-yl)methyl-1,2,3-triazol-1-yl)-2′-deoxyuridine (**143**) with an additional thymine attached to the 5-position of the 2′-deoxyuridine through a triazolomethylene linker. The double-headed nucleoside **143** was synthesized by the CuAAC reaction between 5-azido-5′-*O*-DMTr-2′-deoxyuridine (**142**) and 1-propargylthymine (**13b**) (Scheme 36) [71].



Scheme 35: Synthesis of double-headed nucleosides **140** and **141** via Sonogashira cross coupling reaction.



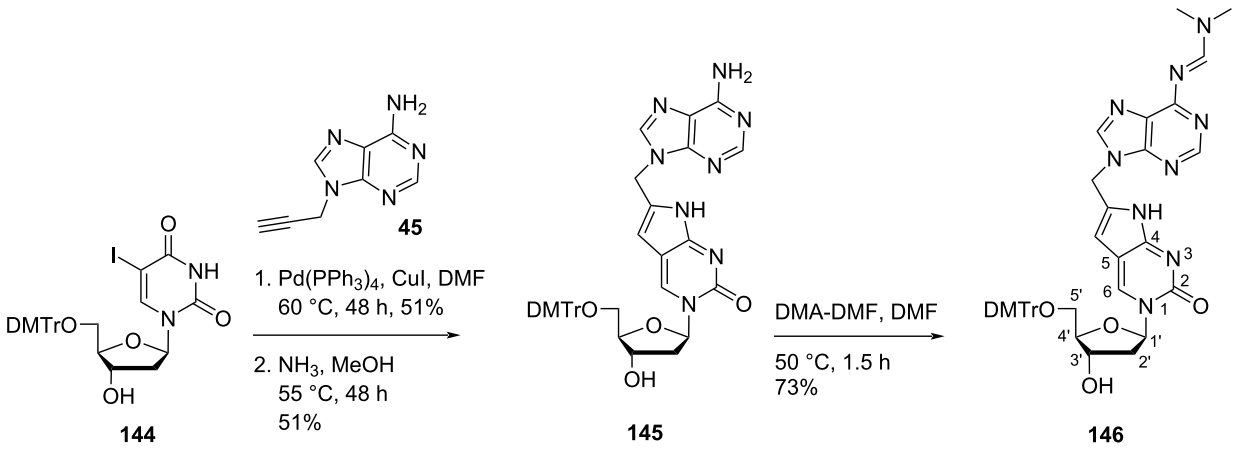
Scheme 36: Synthesis of double-headed nucleoside **143**.

Nielsen and co-workers [71] also synthesized the double-headed nucleoside 5′-*O*-(4,4′-dimethoxytrityl)-6-((N^6-(dimethylaminomethylidenyl)adenin-9-yl)methyl-2′-deoxypyrrolocytidine (**146**) which has adenine attached to the 6-position of the pyrrolo-2′-deoxycytidine through a methylene linker. The double-headed nucleoside **146** was synthesized through the Sonogashira coupling reaction between 5′-*O*-DMTr-5-iodo-2′-deoxyuridine (**144**) and N^9-propargyladenine (**45**) followed by treatment with methanolic ammonia and DMA–DMF (Scheme 37) [71].

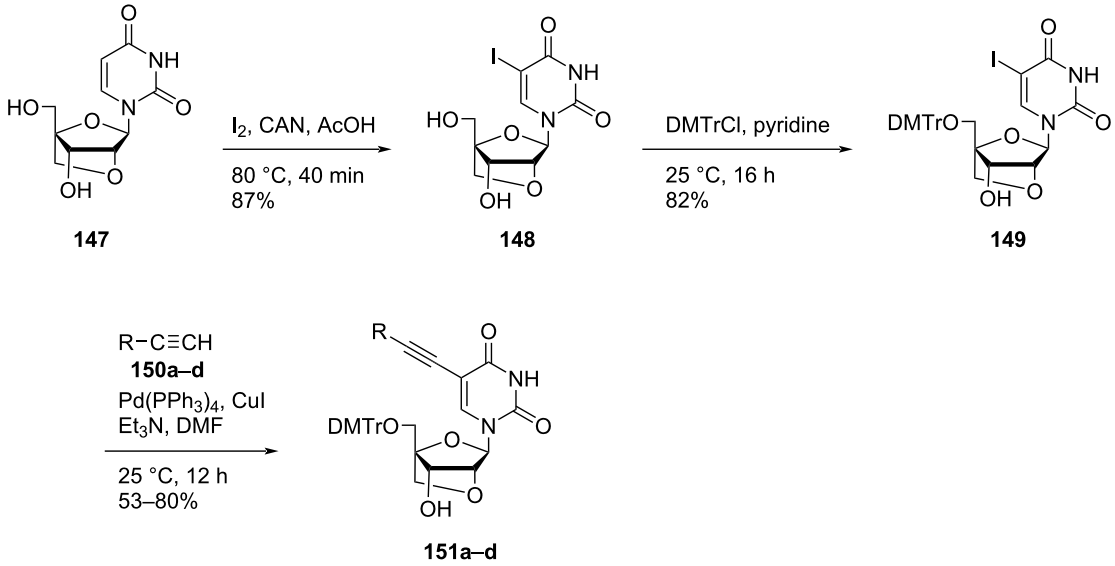
Both double-headed nucleoside monomers **143** and **146** were phosphoramidated at the C-3′ hydroxy group and incorporated into oligonucleotides. The synthesized oligonucleotides were found to decrease the thermal stability of the duplexes. However, their potential in triplex forming oligonucleotides was also studied which concluded the formation of most stable triplexes with single incorporations of additional pyrimidine nucleobases connected via a propylene linker [71].

Hrdlicka and co-workers [24] synthesized 5-*C*-alkynyl-functionalized double-headed nucleosides **151a–d** starting from LNA uridine diol **147** which in turn was synthesized from diacetone-α-D-allose following a procedure reported in the literature [76]. LNA uridine diol **147** was reacted with iodine and ceric ammonium nitrate (CAN) in acetic acid to afford the nucleoside **148**. Nucleoside **148** was then 5′-*O*-dimethoxytritylated in the presence of DMTrCl (4,4′-dimethoxytrityl chloride) and pyridine. The 5′-*O*-dimethoxytritylated nucleoside **149** was further coupled with terminal alkynes **150a–d** under Sonogashira conditions to afford the double-headed nucleosides **151a–d** (Scheme 38) [24].

Hrdlicka and co-workers [24] also synthesized 5-*C*-triazolyl-functionalized double-headed nucleosides **154a**,**b** starting from 5-*C*-ethynyl-functionalized LNA uridine **152**. The LNA uridine **152** was reacted with 1-azidopyrene (**153a**) and 1-azidomethylpyrene (**153b**) separately under copper-catalyzed alkyne azide cycloaddition (CuAAC) reaction conditions to yield the

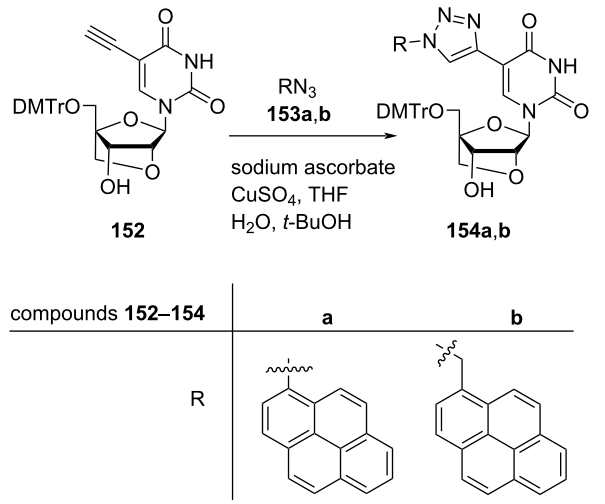


Scheme 37: Synthesis of the double-headed nucleoside **146**.



Scheme 38: Synthesis of 5-*C*-alkynyl-functionalized double-headed nucleosides **151a–d**.

double-headed nucleosides **154a** and **154b**, respectively (Scheme 39) [24].

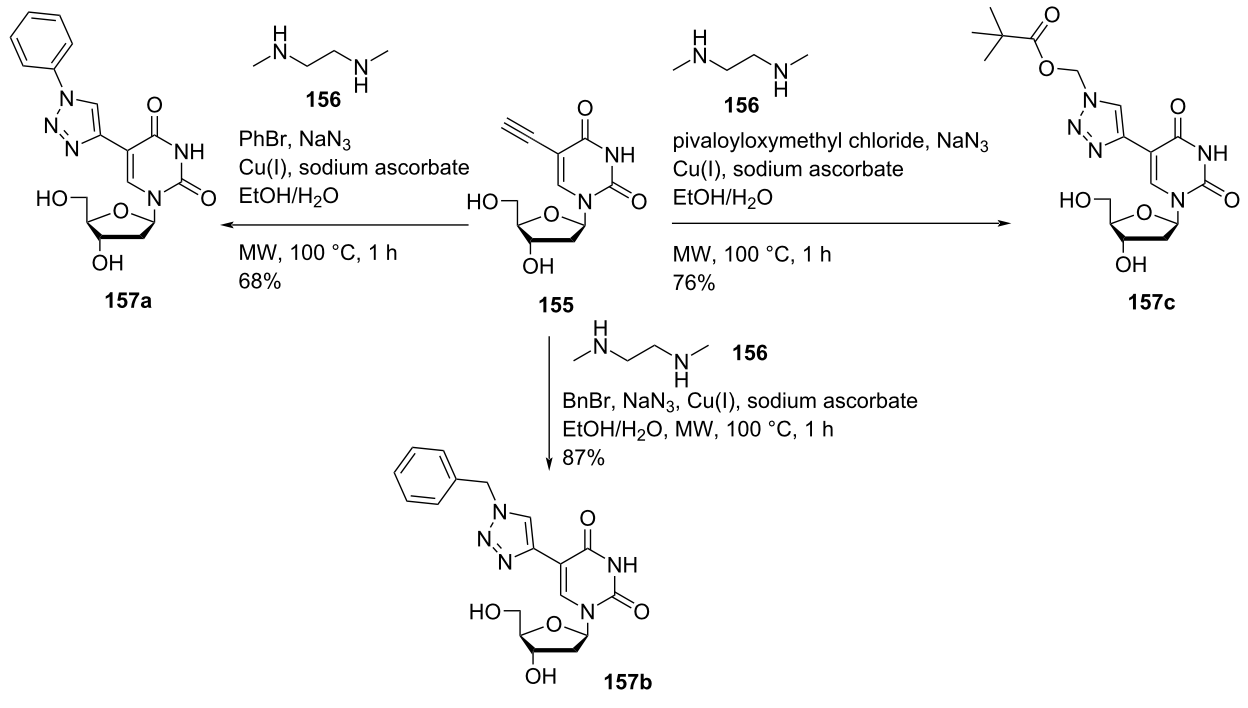


Scheme 39: Synthesis of 5-*C*-triazolyl-functionalized double-headed nucleosides **154a**, **b**.

The synthesized double-headed nucleosides **151a–d** and **154a**,**b** were phosphitylated, incorporated into oligonucleotides and characterized with respect to thermal denaturation, enzymatic stability, and fluorescence properties. The incorporation of the double-headed nucleosides **151a–d** and **154a**,**b** into oligonucleotides failed to form thermostable duplexes with complementary DNA and RNA strands but exhibited a potential resistance towards 3′-exonuclease. The synthesized double-headed nucleosides **151c**,**d** and **154a**,**b** when incorporated into oligonucleotides enabled fluorescent discrimination of targets with single nucleotide polymorphisms (SNPs) [24].

Nielsen and co-workers [15] synthesized a series of double-headed nucleosides 5-(1-phenyl-1*H*-1,2,3-triazol-4-yl)-2′-deoxyuridine (**157a**), 5-(1-benzyl-1*H*-1,2,3-triazol-4-yl)-2′-deoxyuridine (**157b**), and 5-(1-pivaloyloxymethyl-1*H*-1,2,3-triazol-4-yl)-2′-deoxyuridine (**157c**). The synthesis started from 5-ethynyl-2′-deoxyuridine (**155**) which in turn was synthesized from 5-iodo-2′-deoxyuridine following literature procedures [77-79]. The terminal alkyne **155** was reacted with bromobenzene and sodium azide under microwave heating in an EtOH/H_2O mixture in the presence of copper iodide, sodium ascorbate, and *N*,*N*-dimethylethylenediamine (**156**) to afford the double-headed nucleoside **157a**. The reaction of the terminal alkyne **155** with benzyl bromide and pivaloyloxymethyl chloride under similar conditions afforded the double-headed nucleosides **157b** and **157c**, respectively (Scheme 40) [15].

The double-headed nucleosides **157a–c** were introduced into nonamer oligonucleotides by phosphoramidite chemistry [15]. a single incorporation of double-headed nucleosides **157a–c** into oligonucleotides resulted in the formation of unstable duplexes



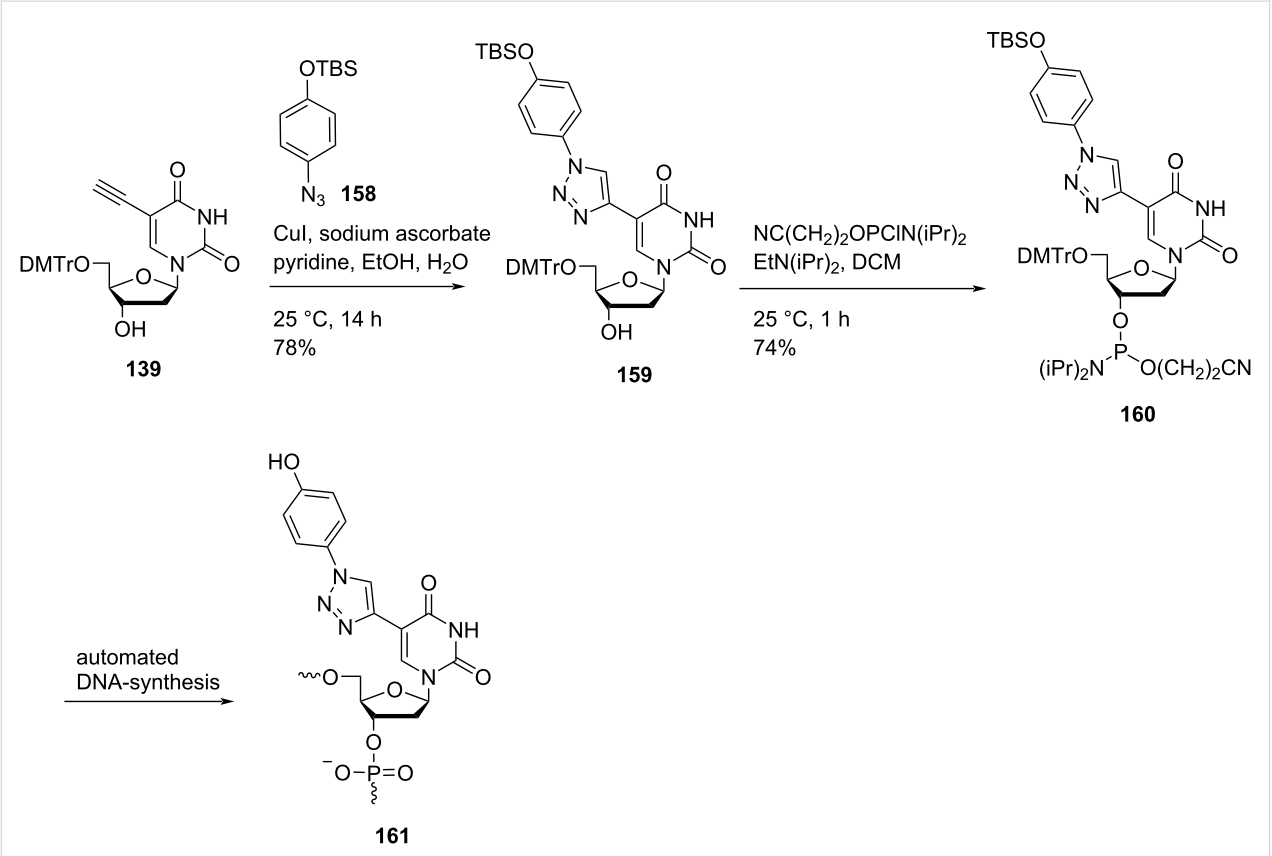
Scheme 40: Synthesis of double-headed nucleosides **157a–c**.

with complementary DNA and RNA strands whereas four consecutive incorporations led to increased duplex stability due to an efficient stacking of heteroaromatic triazoles as revealed by CD spectroscopy and molecular dynamics simulations [15,22]. The double-headed nucleoside **157a** was further used for the synthesis of 5-(phenyltriazol)-2′-deoxyuridine-modified 2′-*O*-methyl mixmer antisense oligonucleotides (AOs). The obtained AOs were investigated for their potential to induce exon skipping in DMD (Duchenne muscular dystrophy) transcript using *H2K mdx* mouse myotubes. It was found that exon-23 skipping potential of oligonucleotide containing 5-(phenyltriazole)-2′-deoxyuridine (**157a**) building blocks placed distantly was slightly better than oligonucleotides containing the 5-(phenyltriazole)-2′-deoxyuridine (**157a**) building blocks placed consecutively [80].

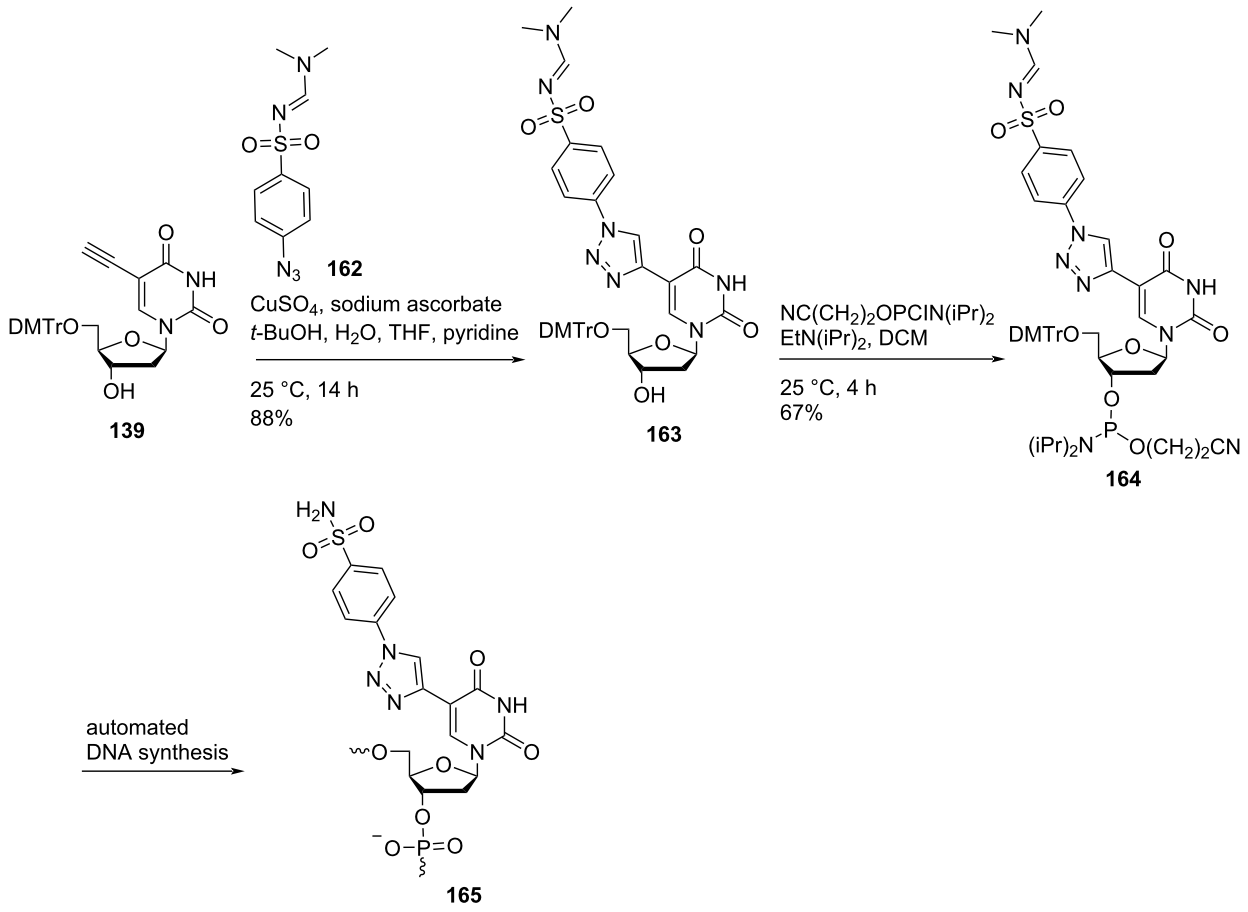
Nielsen and co-workers [26] synthesized triazole-containing double-headed nucleosides **159** and **163** by the reaction of 5′-*O*-dimethoxytritylated nucleoside **139** with *tert*-butyldimethylsilyl 4-azidophenylether (**158**) and *N*-(dimethylaminomethylidene)-4-azidobenzenesulfonamide (**162**), respectively under copper-catalyzed alkyne azide cycloaddition (CuAAC) reaction conditions (Scheme 41 and Scheme 42) [26].

The synthesized double-headed nucleosides **159** and **163** were reacted with 2-cyanoethyl-*N*,*N*-diisopropyl-phosphoramidochloridite in the presence of DIPEA (*N*,*N*-diisopropylethylamine) to afford phosphoramidites **160** and **164** which were then incorporated into oligodeoxynucleotides using automated solid phase synthesis. The synthesized oligonucleotides were removed from the solid support by treatment with concentrated aqueous ammonia which resulted in the formation of incorporated monomers **161** and **165** by simultaneous removal of *tert*-butyldimethylsilyl and amidine protecting groups, respectively (Scheme 41 and Scheme 42) [26].

The incorporation of the double-headed nucleosides **159** and **163** into oligonucleotides resulted in the formation of thermally stable DNA:RNA duplexes due to an efficient π–π stacking between two or more phenyltriazoles in the major groove. The more stable duplex was obtained when oligonucleotide containing monomer **165** was hybridized with the complementary RNA strand due to the best stacking shown by sulfonamide-substituted phenyltriazoles in the major groove [26,27]. Single incorporations of 5-*C*-triazolylbenzenesulfonamide-substituted monomer **165** at four positions within the gap region of RNase H gapmer antisense oligonucleotides (ASOs) reduced wild-type



Scheme 41: Synthesis of double-headed nucleoside **159**, phosphoramidite **160** and the corresponding nucleotide monomer **161**.



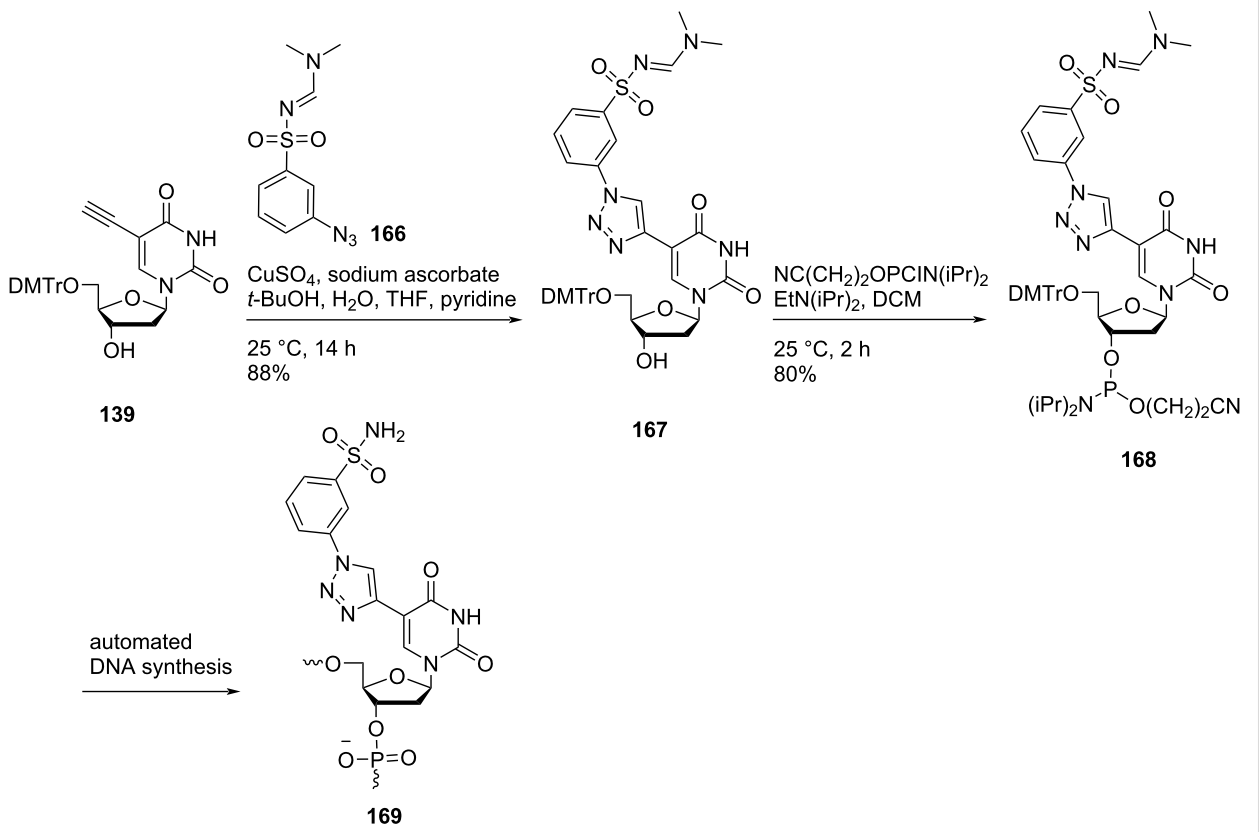
Scheme 42: Synthesis of double-headed nucleoside **163**, phosphoramidite **164** and the corresponding nucleotide monomer **165**.

and mutant huntingtin mRNA in human patient fibroblasts. A structural model of the catalytic domain of human RNase H bound to ASO:RNA heteroduplexes was created which was utilized for explaining the activity and selectivity observations in cells and in the biochemical assays [81].

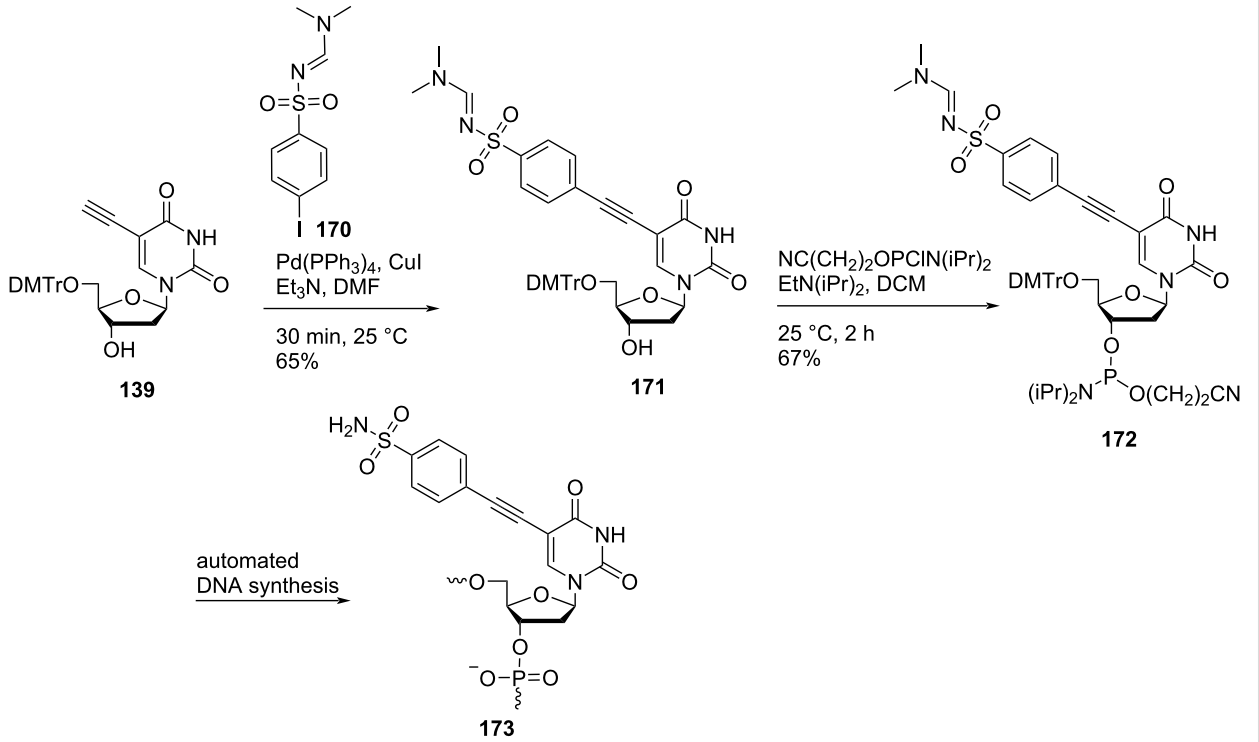
Sharma and co-workers [27] synthesized the double-headed nucleoside **167** by reacting 5′-*O*-dimethoxytritylated nucleoside **139** with *N*-(dimethylaminomethylidene)-3-azidobenzenesulfonamide (**166**) under copper-catalyzed alkyne–azide cycloaddition (CuAAC) reaction conditions (Scheme 43). The synthesized double-headed nucleoside **167** was further reacted with 2-cyanoethyl-*N*,*N*-diisopropyl-phosphoramidochloridite in the presence of DIPEA to afford phosphoramidite **168** which was then incorporated into oligodeoxynucleotides using automated solid phase synthesis. The synthesized oligonucleotides were removed from the solid support by treatment with concentrated aqueous ammonia which resulted in the formation of incorporated monomer **169** by removal of the amidine protection (Scheme 43) [27].

Sharma and co-workers [27] also synthesized double-headed nucleosides **171** and **175** by the Sonogashira coupling of 5′-*O*-dimethoxytritylated alkyne **139** with *N*-(dimethylaminomethylidene)-4-iodobenzenesulfonamide (**170**) and *N*-(dimethylaminomethylidene)-3-iodobenzenesulfonamide (**174**), respectively (Scheme 44 and Scheme 45).

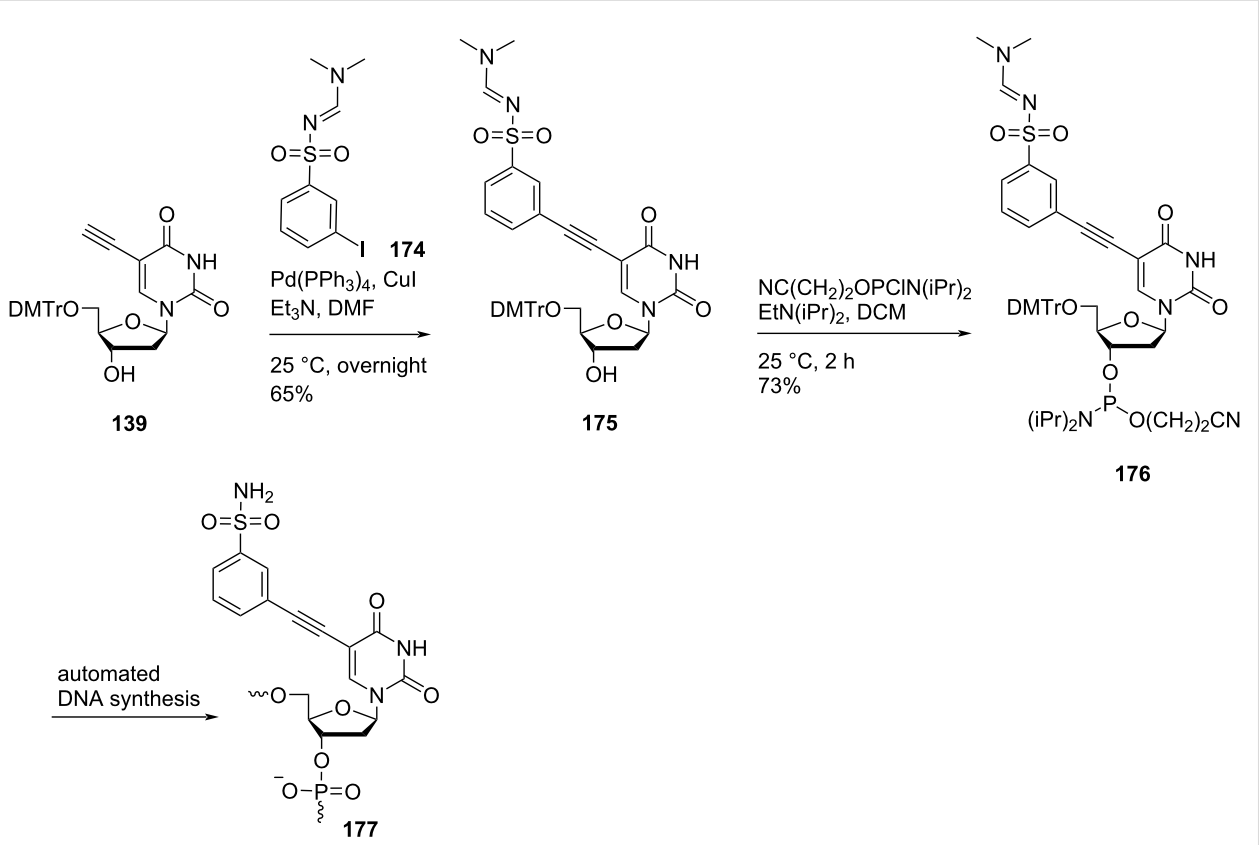
The synthesized double-headed nucleosides **171** and **175** were reacted with 2-cyanoethyl-*N*,*N*-diisopropyl-phosphoramidochloridite in the presence of DIPEA to afford phosphoramidites **172** and **176**, respectively. The phosphoramidites **172** and **176** were then incorporated into oligonucleotides using automated solid phase synthesis which after removal from the solid support by treatment with concentrated aqueous ammonia resulted in the formation of incorporated monomers **173** and **177**, respectively by removal of the amidine protection of sulfonamides (Scheme 44 and Scheme 45) [27]. The double-headed nucleoside **178** was also synthesized starting from 5′-*O*-dimethoxytritylated alkyne **139** under Sonogashira cross coupling reaction conditions (Scheme 46) [28].



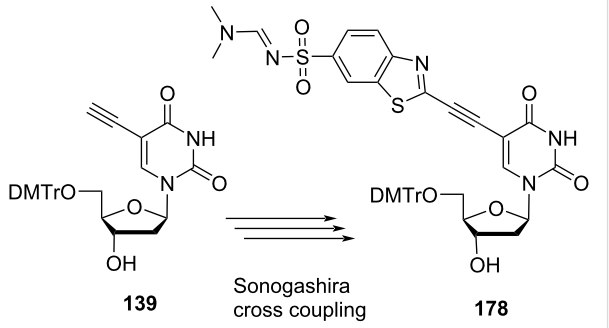
Scheme 43: Synthesis of double-headed nucleoside **167**, phosphoramidite **168**, and the corresponding nucleotide monomer **169**.



Scheme 44: Synthesis of double-headed nucleoside **171**, phosphoramidite **172**, and the corresponding nucleotide monomer **173**.



Scheme 45: Synthesis of double-headed nucleoside **175**, phosphoramidite **176**, and the corresponding nucleotide monomer **177**.

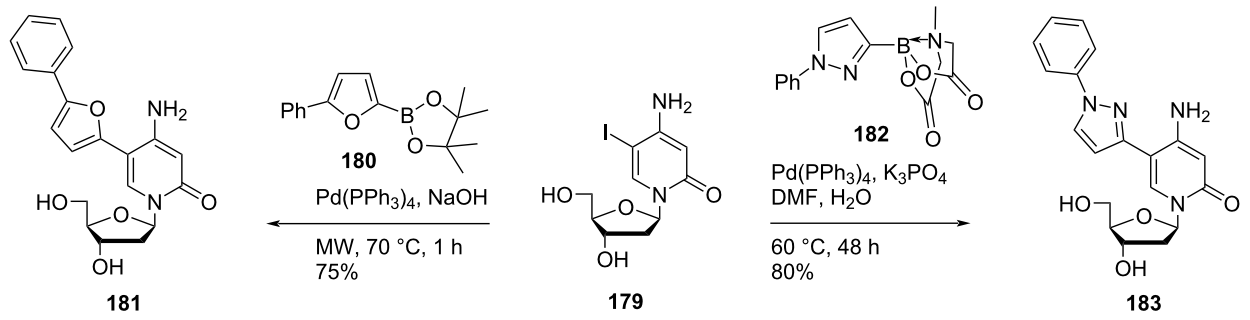


Scheme 46: Synthesis of double-headed nucleoside **178**.

The incorporation of the double-headed nucleoside **167** into oligonucleotides resulted in the formation of an equally stable DNA:RNA duplex as in the case of double-headed nucleoside **163** irrespective of the positional orientation of the sulfonamide group due to an efficient π–π stacking between two or more phenyltriazoles in the major groove [27]. On the other hand, the incorporation of the double-headed nucleosides **171** and **175** into oligonucleotides resulted in the formation of less stable DNA:RNA duplexes because of the poor stacking by the alkynyl group as compared to triazolyl groups in double-headed nucleosides **163** and **167** [27]. The double-headed nucleotide **173** was fully accepted by KOD (kodakaraensis), Phusion, and Klenow DNA polymerases as substrate which resulted in the formation of fully extended DNA. KOD DNA polymerase was found to be the best enzyme to produce DNA containing the double-headed nucleotide **173** in good yield and Phusion DNA polymerase amplified the template containing double-headed nucleotide **173** efficiently by PCR (polymerase chain reaction) [82].

Nielsen and co-workers [22] synthesized the double-headed nucleoside **181** by Suzuki–Miyaura cross coupling reaction of 5-iodo-2′-deoxycytidine (**179**) with 5-phenylfuran-2-boronic acid pinacol ester (**180**) in the presence of $Pd(PPh_3)_4$ and NaOH (Scheme 47).

They also synthesized the double-headed nucleoside **183** by K_3PO_4-mediated [83] Suzuki–Miyaura cross coupling reaction of 5-iodo-2′-deoxycytidine (**179**) with pyrazole MIDA (*N*-methyliminodiacetic acid) boronate **182** which in turn was synthesized by sydnone–alkyne cycloaddition reaction between ethynylboronic acid MIDA ester and *N*-phenylsydnone (Scheme 47) [22].

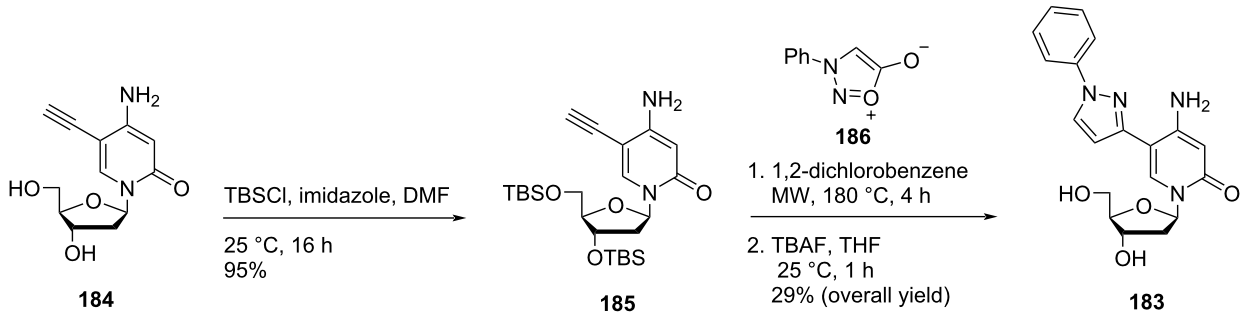


Scheme 47: Synthesis of the double-headed nucleosides **181** and **183**.

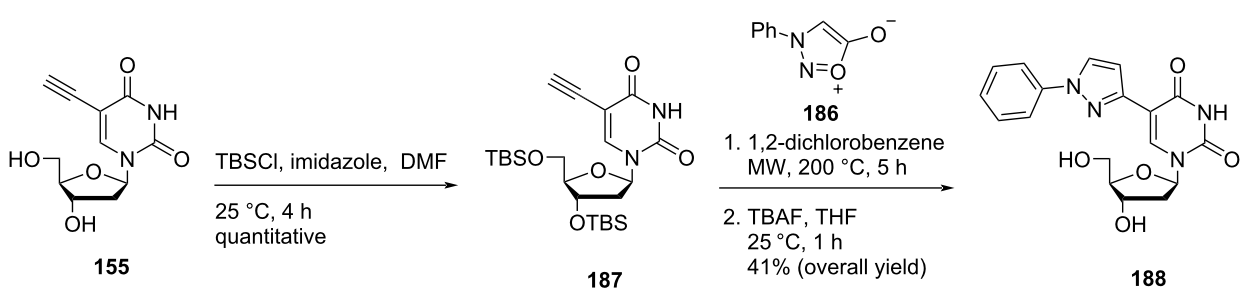
The double-headed nucleoside **183** was also synthesized from 5-ethynyl-2′-deoxycytidine (**184**) which was first converted into 3′,5′-bis-*O*-(*tert*-butyldimethylsilyl)-5-ethynyl-2′-deoxycytidine (**185**). The protected nucleoside **185** was reacted with *N*-phenylsydnone (**186**) via thermal [2 + 3] sydnone–alkyne cycloaddition [84] to afford the double-headed nucleoside **183** (Scheme 48) [22].

Nielsen and co-workers [22] also synthesized the double-headed nucleoside **188** via thermal [2 + 3] sydnone–alkyne cycloaddition reaction between 3′,5′-bis-*O*-(*tert*-butyldimethylsilyl)-5-ethynyl-2′-deoxyuridine (**187**) and *N*-phenylsydnone (**186**) (Scheme 49).

The incorporation of the double-headed nucleosides **181** and **183** multiple times into oligonucleotides resulted in the formation of stable DNA:RNA duplexes due to the perfect stacking of the aromatic moieties in the major groove of the duplex [22]. The double-headed nucleoside **183** containing a phenylpyrazole moiety exhibited better π–π stacking interactions in the major groove with itself and with an adjacent double-headed nucleoside (**157a**) incorporated as compared to the double-headed nucleoside **181** containing a flexible phenylfuran moiety. There was not any change in the geometry of the duplexes observed upon introduction of double-headed nucleosides **181** and **183** as revealed by CD spectroscopy and molecular modeling. The synthesized oligonucleotides containing



Scheme 48: Alternative synthesis of the double-headed nucleoside **183**.



Scheme 49: Synthesis of double-headed nucleoside **188** through thermal [2 + 3] sydnone–alkyne cycloaddition reaction.

consecutive triazole-functionalized double-headed nucleosides **183** and **157a** were found to form highly stable duplexes due to a large aromatic overlap of their substituents at the 5-position due to which they can be utilized as a simple tool in high affinity RNA targeting oligonucleotides [22].

Hrdlicka and co-workers [25] synthesized the double-headed nucleosides **190** and **191** starting from *C*5-ethynyl-5′-*O*-(4,4′-dimethoxytrityl)-2′-deoxyuridine (**139**) which in turn was synthesized from 5-iodo-2′-deoxyuridine following a procedure reported in the literature [77]. The nucleoside **139** was reacted with 1-azidomethylpyrene (**189a**) and 1-azidopyrene (**189b**) under copper-catalyzed alkyne–azide cycloaddition (CuAAC) reaction conditions to afford the double-headed nucleosides **190** and **191**, respectively (Scheme 50) [25].

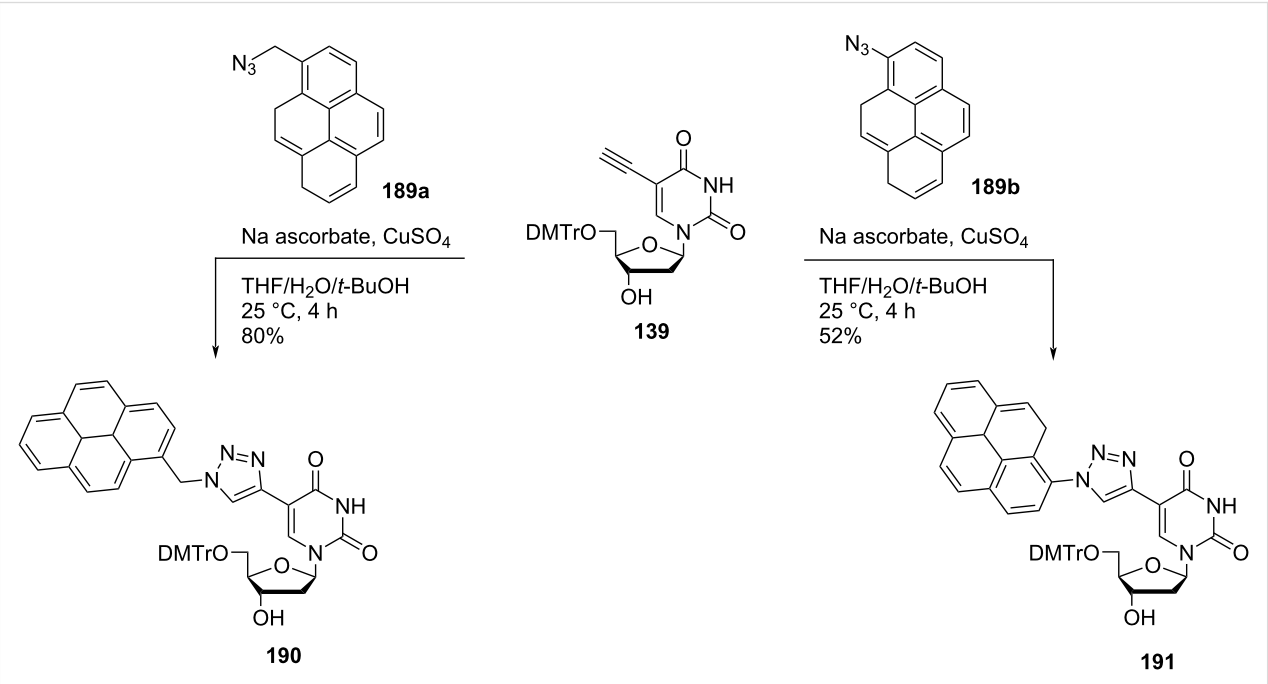
The incorporation of the double-headed nucleoside **190** into oligonucleotides displayed significant hybridization-induced increase in fluorescence emission whereas the double-headed nucleoside **191** allowed for efficient fluorescent discrimination of SNPs (single nucleotide polymorphisms) via a G-specific quenching mechanism when incorporated into oligonucleotides [25].

Pyranosyl double-headed nucleosides

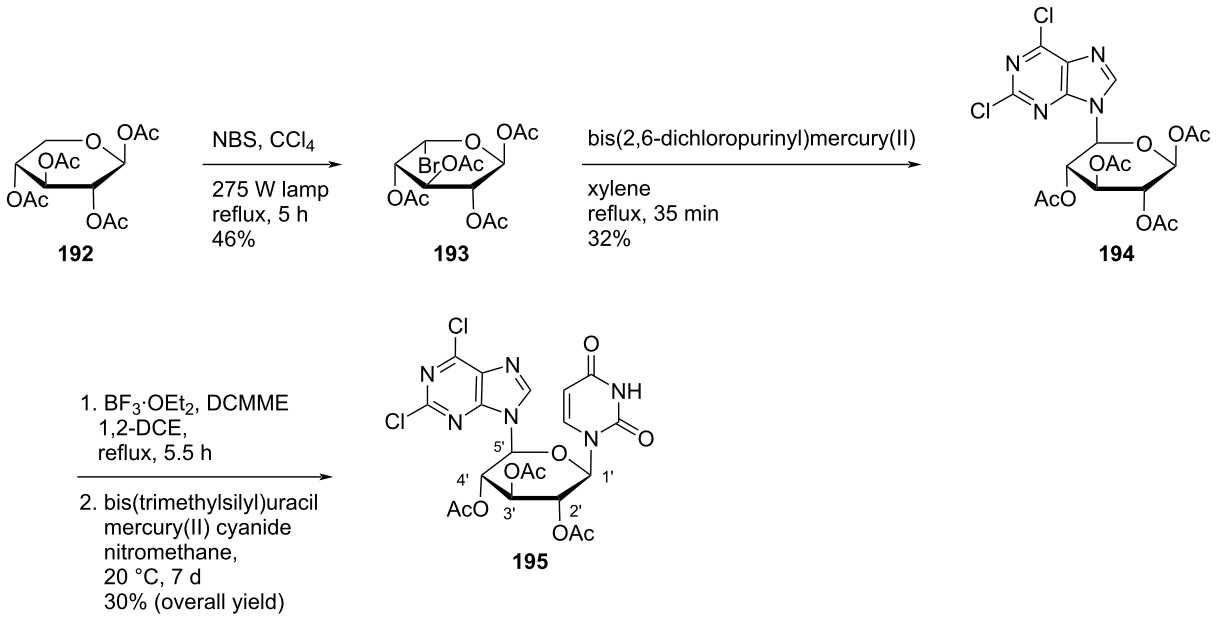
Synthetic methodologies have been developed for placing the additional nucleobase at various positions in the pyranonucleosides. Here, we have categorized the double-headed pyranosyl nucleoside monomers depending on the point of attachment at the pyranose sugar. The double-headed pyranosyl nucleosides have the nucleobase attached at the C-6′ position of the pyranose moiety, either directly or by a methylene linker (Figure 1). Pyranosyl nucleosides where the additional nucleobase is attached at the C-3′ position of the pyranosyl moiety through a triazolo-linker have also been synthesized.

Ferrier and Tyler [85] synthesized the double-headed nucleoside 1-[(5*S*)-2,3,4-tri-*O*-acetyl-5-(2,6-dichloropurin-9-yl)-β-D-xylopyranosyl]uracil (**195**) by photobromination of tetra-*O*-acetyl-β-D-xylopyranose (**192**) to afford the crystalline product **193** which upon reaction with bis(2,6-dichloropurinyl)mercury in xylene afforded crystalline compound **194**. Subsequently, the nucleoside analogue **194** was reacted with $BF_3{\cdot}OEt_2$ followed by reaction with silylated uracil to get the double-headed nucleoside **195** (Scheme 51) [85].

Prasad and co-workers [86] synthesized hexopyranosyl double-headed pyrimidine homonucleosides 1-[(6-deoxy-6-(uracil-1-yl)-β-D-glucopyranosyl)methyl]uracil (**200a**), 1-[(6-deoxy-6-(thymin-1-yl)-β-D-2,3,4-tri-*O*-benzylglucopyranosyl)methyl]thymine (**200b**) and 1-[(6-deoxy-6-(uracil-1-yl)-β-D-mannopyranosyl)methyl]uracil (**200c**) from dihydroxy 2,6-anhydro-3,4,5-tri-*O*-benzylheptitols (**196a**,**b**) which in turn were synthesized from D-glucose and D-mannose [87]. The benzylated 2,6-anhydroheptitols **196a**,**b** were reacted with tosyl chloride to form the ditosylated compounds **197a**,**b** which upon



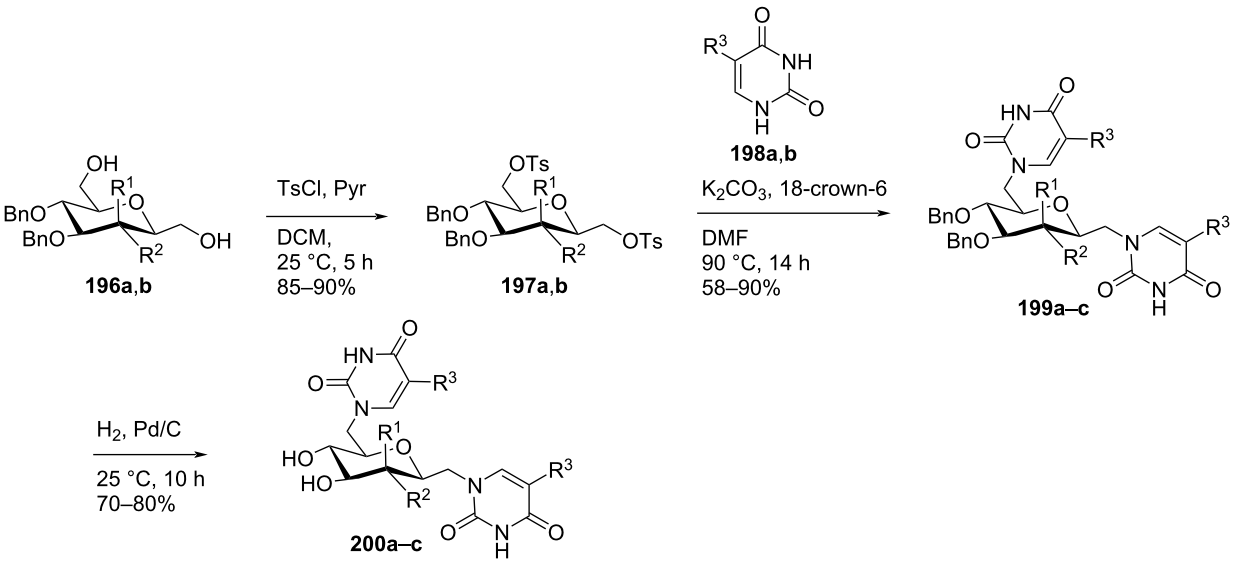
Scheme 50: Synthesis of the double-headed nucleosides **190** and **191**.



Scheme 51: Synthesis of 1-((5*S*)-2,3,4-tri-*O*-acetyl-5-(2,6-dichloropurin-9-yl)-β-D-xylopyranosyl)uracil (**195**).

reaction with substituted thymine and uracil **198a**,**b** afforded the benzylated double-headed nucleosides **199a**–**c**. Next, debenzylation of the nucleoside monomers **199a–c** afforded the final double-headed nucleoside monomers **200a–c**. (Scheme 52) [86].

Komiotis and co-workers [37] synthesized 3′-*C*-(1,4-disubstituted-1,2,3-triazolyl)-substituted double-headed pyranonucleosides **203–210** from 3′-*C*-ethynyl-β-D-allopyranonucleoside derivatives **201a–f** (Figure 2) which in turn were synthesized from 1,2:5,6-di-*O*-isopropylidene-α-D-ribohexofuranos-3-ulose [88].



compound	196a,b	197a,b	198a	198b	199a	199b	199c	200a	200b	200c
R^1	H	OBn	–	–	H	H	OBn	H	H	OH
R^2	OBn	H	–	–	OBn	OBn	H	OH	OH	H
R^3	–	–	H	CH_3	H	CH_3	H	H	CH_3	H

Scheme 52: Synthesis of hexopyranosyl double-headed pyrimidine homonucleosides **200a–c**.

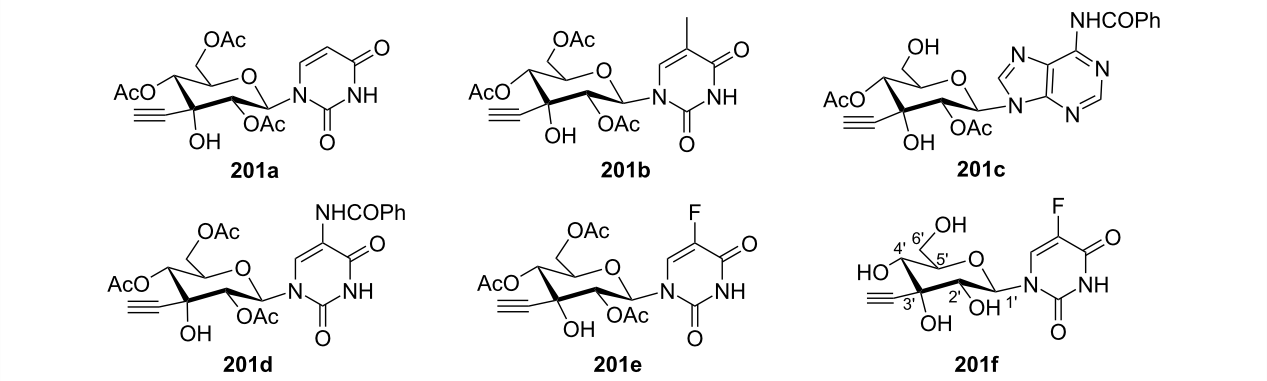
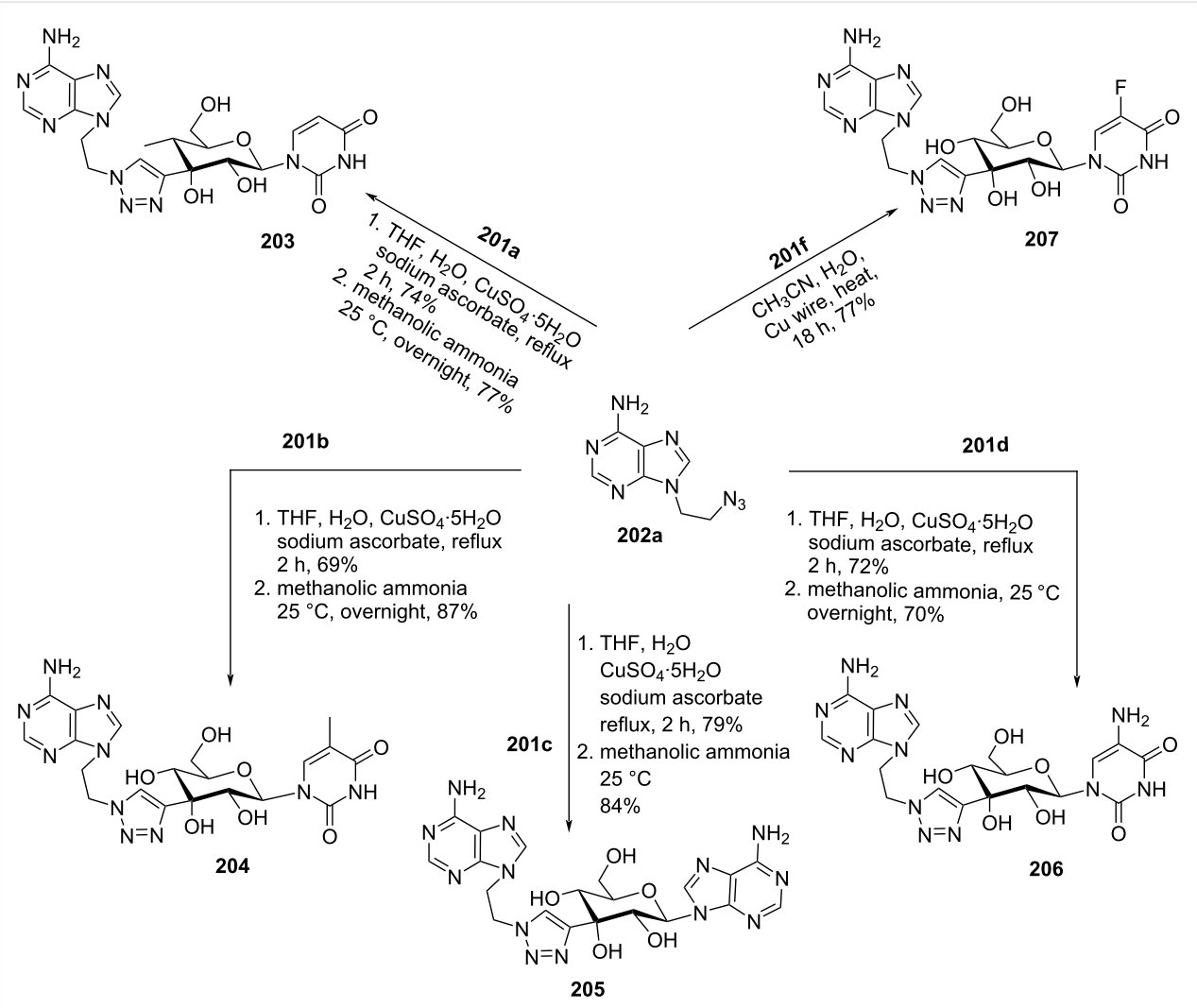
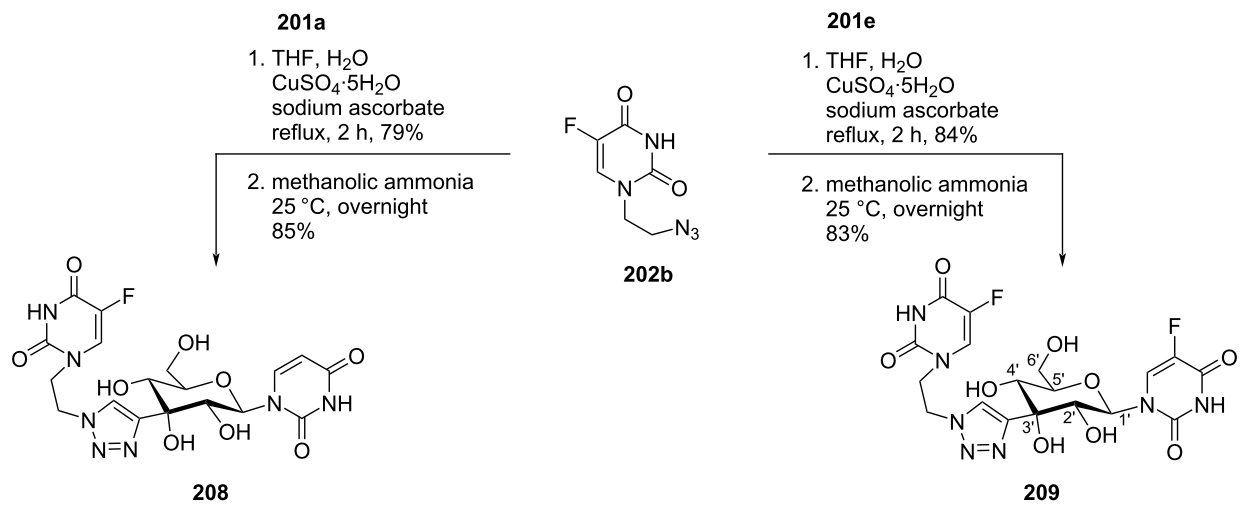


Figure 2: 3′-*C*-Ethynyl-β-D-allopyranonucleoside derivatives **201a–f**.

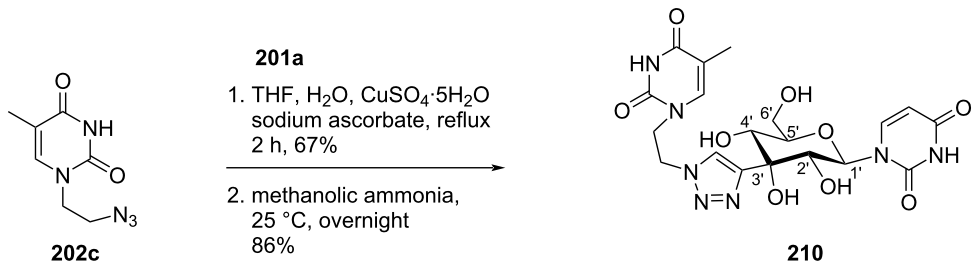
The 3′-*C*-ethynyl-substituted pyranonucleoside derivatives **201a–f** were reacted with azidoethyladenine, 5-fluorouracil and thymine **202a–c** via copper-catalyzed azide–alkyne cycloaddition (CuAAC) reaction followed by treatment with methanolic ammonia to afford the double-headed nucleosides **203–210** (Scheme 53, Scheme 54, and Scheme 55) [37].



Scheme 53: Synthesis of 3′-*C*-(1,4-disubstituted-1,2,3-triazolyl)-double-headed pyranonucleosides **203–207**.



Scheme 54: Synthesis of 3′-*C*-(1,4-disubstituted-1,2,3-triazolyl)-double-headed pyranonucleosides **208** and **209**.



Scheme 55: Synthesis of 3′-*C*-(1,4-disubstituted-1,2,3-triazolyl)-double-headed pyranonucleoside **210**.

The double-headed nucleosides **203–210** were evaluated for their antiviral and cytostatic activities, and the nucleosides **204**, **206**, and **207** showed moderate cytostatic activity against human cervix carcinoma HeLa cells [37].

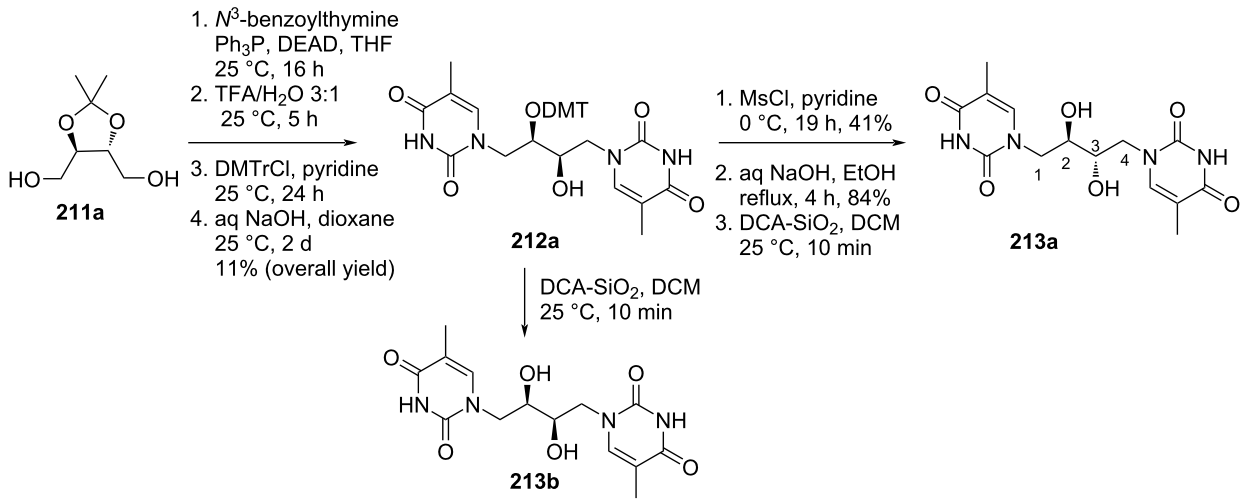
Acyclic double-headed nucleosides

In this section, double-headed nucleosides are included that have an acyclic carbohydrate moiety and the heterocyclic moieties/nucleobases are terminally attached at the sugar moiety (Figure 1).

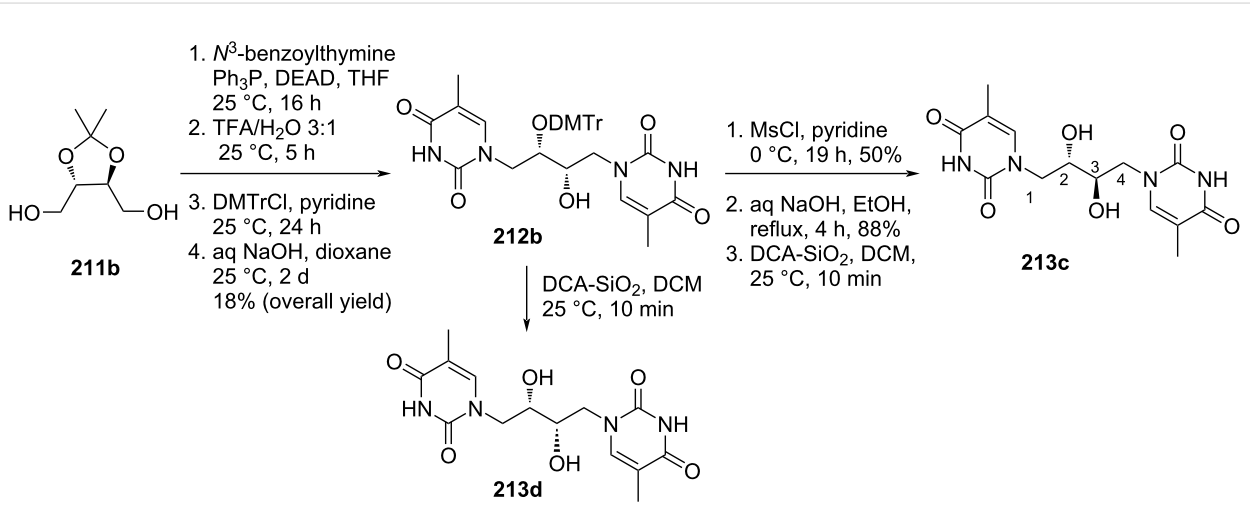
Nielsen and co-workers [89] synthesized four stereoisomers of double-headed acyclic nucleosides 1,4-bis(thymine-1-yl)butane-2,3-diols **213a–d** starting from either D- or L-2,3-*O*-isopropylidenethreitol **211a**,**b**. The dihydroxy compounds **211a**,**b** were reacted with N^3-benzoylthymine under Mitsunobu reaction conditions followed by DMTr protection to give two enantiopure compounds (2*R*,3*R*)-1,4-bis(thymin-1-yl)-3-*O*-DMTr-butan-2-ol (**212a)** and (2*S*,3*S*)-1,4-bis(thymin-1-yl)-3-*O*-DMTr-butan-2-ol (**212b**). The two compounds upon removal of the DMTr group gave the acyclic double-headed nucleosides (2*R*,3*R*)-1,4-bis(thymine-1-yl)butane-2,3-diol (**213b**) and (2*S*,3*S*)-1,4-bis(thymine-1-yl)butane-2,3-diol (**213d**). Further, the reaction of the 3-*O*-DMTr-protected nucleosides **212a**,**b** with mesyl chloride followed by treatment with aq. NaOH in ethanol afforded the nucleosides (2*S*,3*R*)-1,4-bis(thymine-1-yl)butane-2,3-diol (**213a**) and (2*R*,3*S*)-1,4-bis(thymine-1-yl)butane-2,3-diol (**213c) (**Scheme 56 and Scheme 57) [89].

These double-headed nucleosides when incorporated into oligonucleotides destabilized both DNA and RNA duplexes. However, nucleosides with 2′(*S*)-configuration were found to destabilize duplexes and bulged motifs to a lesser extent than the other stereoisomers [89].

Nasr [16] synthesized 1,4-bis(9-methyl-1,3,4-oxadiazino[6,5-*b*]indol-2-yl-1-ium) dichloride, 1,4-bis(9-ethyl-1,3,4-oxadiazino[6,5-*b*]indol-2-yl-1-ium) dichloride, and 1,4-bis(9-acetyl-1,3,4-oxadiazino[6,5-*b*]indol-2-yl-1-ium) dichloride-substituted double-headed nucleosides of 1,2,3,4-tetra-*O*-acetylgalactotetritol **218b–d** starting from 1,3-dihydro-2,3-dioxo-2*H*-indoles **214a–c**. The indoles were condensed with galactaric



Scheme 56: Synthesis of double-headed acyclic nucleosides (2*S*,3*R*)-1,4-bis(thymine-1-yl)butane-2,3-diol (**213a**) and (2*R*,3*R*)-1,4-bis(thymine-1-yl)butane-2,3-diol (**213b**).



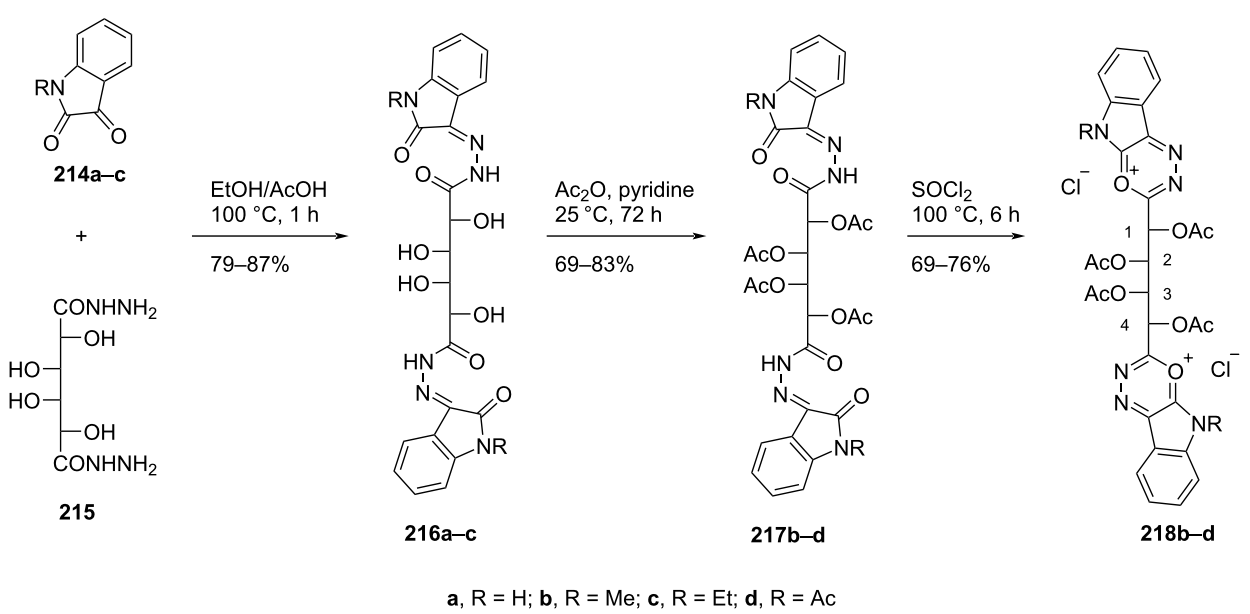
Scheme 57: Synthesis of double-headed acyclic nucleosides (2*R*,3*S*)-1,4-bis(thymine-1-yl)butane-2,3-diol (**213c**) and (2*S*,3*S*)-1,4-bis(thymine-1-yl)butane-2,3-diol (**213d**).

acid bishydrazide to give compounds **216a–c** which upon acetylation followed by heterocyclization in the presence of thionyl chloride afforded nucleosides **218b–d** (Scheme 58) [16].

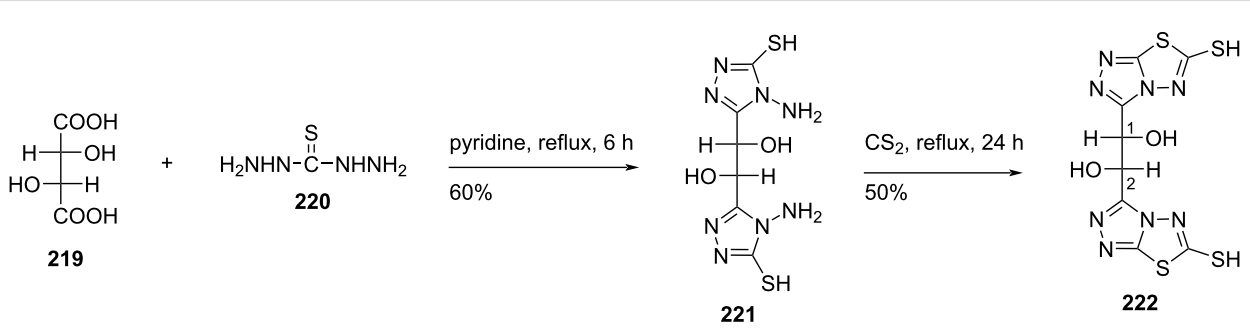
The synthesized double-headed nucleosides **218b–d** may exhibit potential biological activities due to the resistance of the C-glycosidic moiety towards hydrolytic or enzymatic cleavage [90] and the enhanced hydrophilicity which results in an increased transportation to biological systems [16].

El Ashry and co-workers [17] synthesized functionalized 1,2-bis(1,2,4-triazol-3-yl)ethane-1,2-diols **222** and **223a–f** starting from (1*R*,2*S*)-1,2-bis(4-amino-5-mercapto-1,2,4-triazol-3-yl)-ethane-1,2-diol (**221**) which in turn was synthesized by reacting L-tartaric acid (**219**) with thiocarbohydrazide. The reaction of 4-amino-5-mercapto-3-substituted-1,2,4-triazole **221** with carbon disulfide, ethyl bromoacetate, phenacyl bromide, benzoin, *p*-nitrobenzaldehyde, dimedone, and maleic anhydride afforded the double-headed nucleosides **222** and **223a–f** (Scheme 59 and Scheme 60) [17].

The double-headed nucleosides **222** and **223a–f** were synthesized with the aim to evaluate their biological activities due to the potent inhibitory effect of the precursor 4-amino-5-mercapto-1,2,4-triazole against glycosidase enzymes [17,91-93].



Scheme 58: Synthesis of double-headed acetylated 1,3,4-oxadiazino[6,5-*b*]indolium-substituted *C*-nucleosides **218b–d**.



Scheme 59: Synthesis of double-headed acyclic nucleoside **222**.

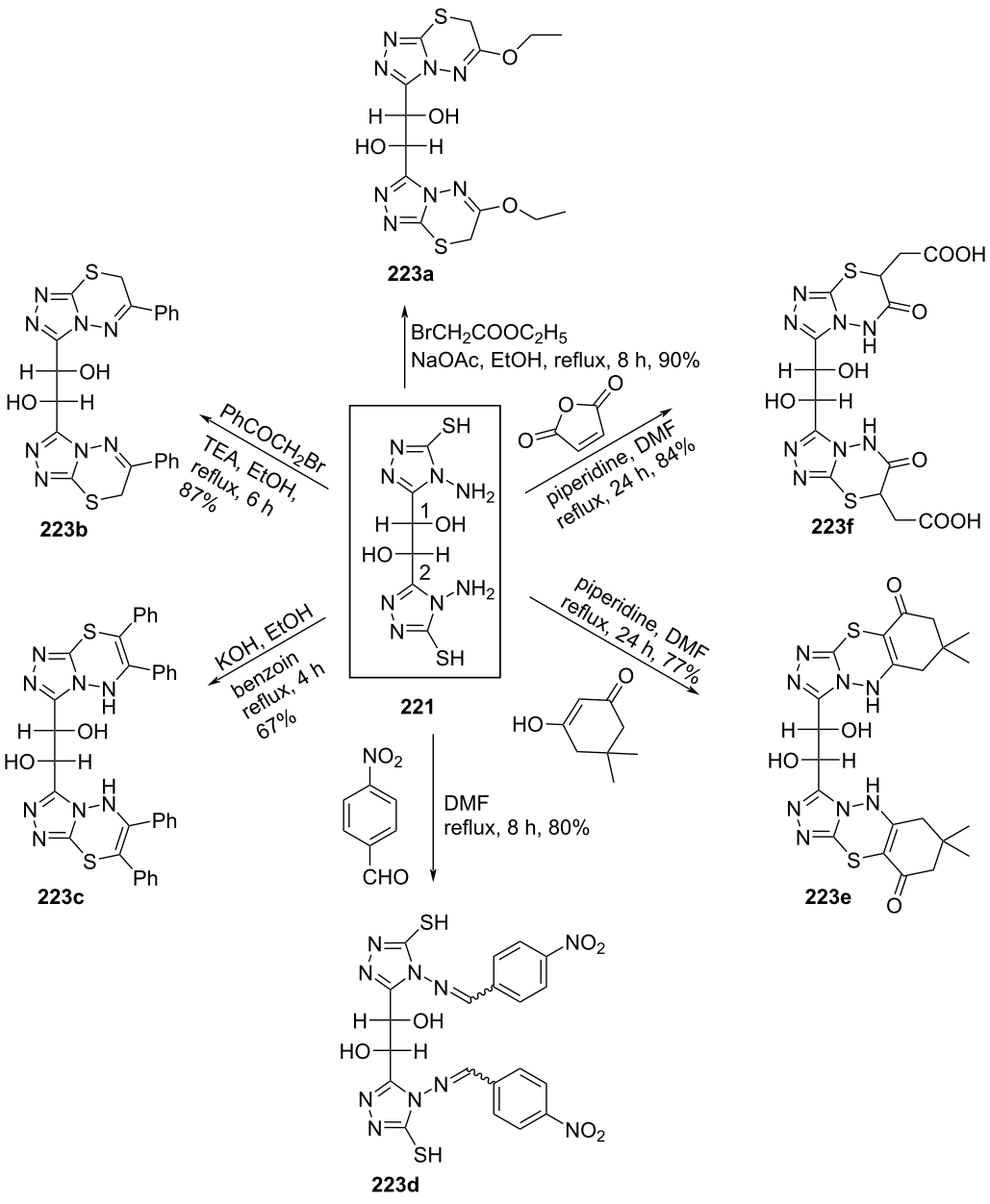
Nasr [18] synthesized the double-headed acyclic 1,2,4-triazino[5,6-*b*]indole *C*-nucleosides **226–231** through the heterocyclization of bis(2-oxoindolin-3-ylidene)galactaric acid hydrazide (**225**) with various one-nitrogen cyclizing agents (Scheme 61).

The synthesized double-headed nucleosides **226–231** were expected to possess potent biological activities due to the known antimicrobial [94-98], antiviral [99], antihypertensive [99,100], analgesic [101], and antitumor activities [102] exhibited by various derivatives of 1,2,4-triazino[5,6-*b*]indole [18].

Nasr and co-workers [19] also synthesized double-headed 1,2,4-triazoline (**232a,b**, **233**), 1,3,4-oxadiazoline (**234**), 1,3,4-thiadiazoline (**235**) acyclo *C*-nucleosides starting from galactaric acid hydrazide (**215**). The syntheses started with the condensation of compound **215** with carbon disulfide in the presence of ethanolic potassium hydroxide to give the dipotassium salt of galactaric acid bis(hydrazidocarbodithioic acid) which was then heterocyclized under different reaction conditions to give three types of double-headed nucleosides (Scheme 62) [19].

The acyclic double-headed nucleosides **232a** and **233–235** were screened for their in vitro antibacterial activity against the Gram-negative bacterium *Escherichia coli* and the Gram-positive bacterium *Staphylococcus aureus* and for their antifungal activity against *Candida albicans* using the agar diffusion method [103]. Among the tested compound, derivative **235** showed fair activity against *E. coli* and *C. albicans* but was inactive against *S. aureus* whereas compound **234** showed activity only against *S. aureus* [19].

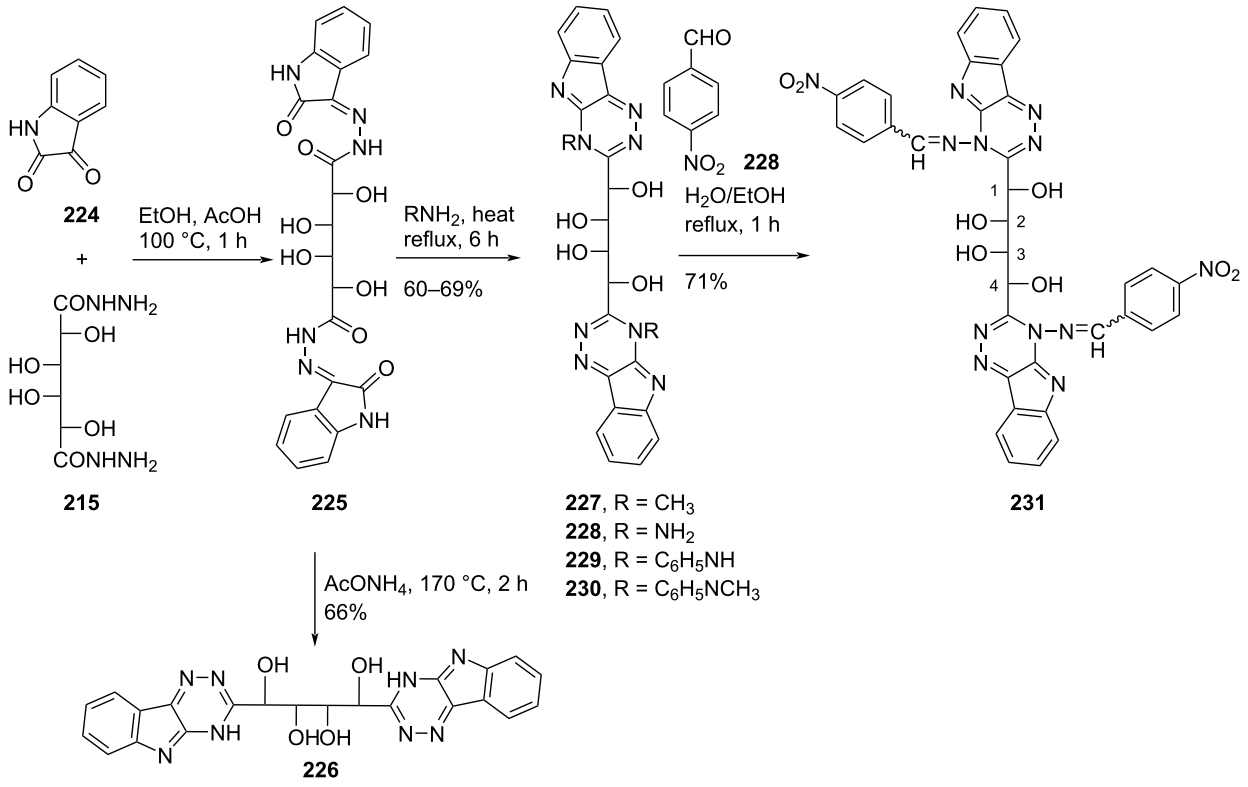
Amara and Othman [20] synthesized the double-headed acyclo-*C*-nucleosides 1,4-bis(3-mercapto-1*H*-1,2,4-triazol-5-yl)butane-



Scheme 60: Synthesis of functionalized 1,2-bis(1,2,4-triazol-3-yl)ethane-1,2-diols **223a–f**.

1,2,3,4-tetrol **240**, 5,5′-(1,2,3,4-tetrahydroxybutane-1,4-diyl)-bis(1,3,4-oxadiazole-2(3*H*)-thione) **241**, and 1,4-bis(4-amino-5-mercapto-4*H*-1,2,4-triazol-3-yl)butane-1,2,3,4-tetrol **242** starting from D-glucose (**236**). The inexpensive sugar **236** was converted into 2,3,4,5-tetrahydroxyhexanedihydrazide **238** in two steps which was further reacted with either ammonium thiocyanate or carbon disulfide to give the bishydrazinocarbothioamide **239** and the acyclic double-headed nucleoside **241**, respectively. The double-headed nucleosides **240** and **242** were obtained by treatment of compound **239** with NaOH and of compound **241** with hydrazine hydrate, respectively (Scheme 63) [20].

The double-headed *C*-nucleosides **240**–**242** were tested in vitro against Gram-positive bacteria *Staphylococcus aureus*, *Listeria inovanii* and Gram-negative bacteria *Klebsiella pneumoniae*, *Salmonella* sp., and *Escherichia coli*. All the double-headed nucleosides except derivative **242** showed moderate antibacterial activity in comparison with the known antibiotic combination amoxicillin/clavulanic acid (AMC) [20].



Scheme 61: Synthesis of acyclic double-headed 1,2,4-triazino[5,6-*b*]indole *C*-nucleosides **226–231**.

The structural and electronic properties of the double-headed nucleosides were explored theoretically by performing semi-empirical molecular orbital, ab initio Hartree–Fock (HF), and density functional theory (DFT) calculations and their geometries were optimized at the level of Austin Model 1 (AM1) [104].

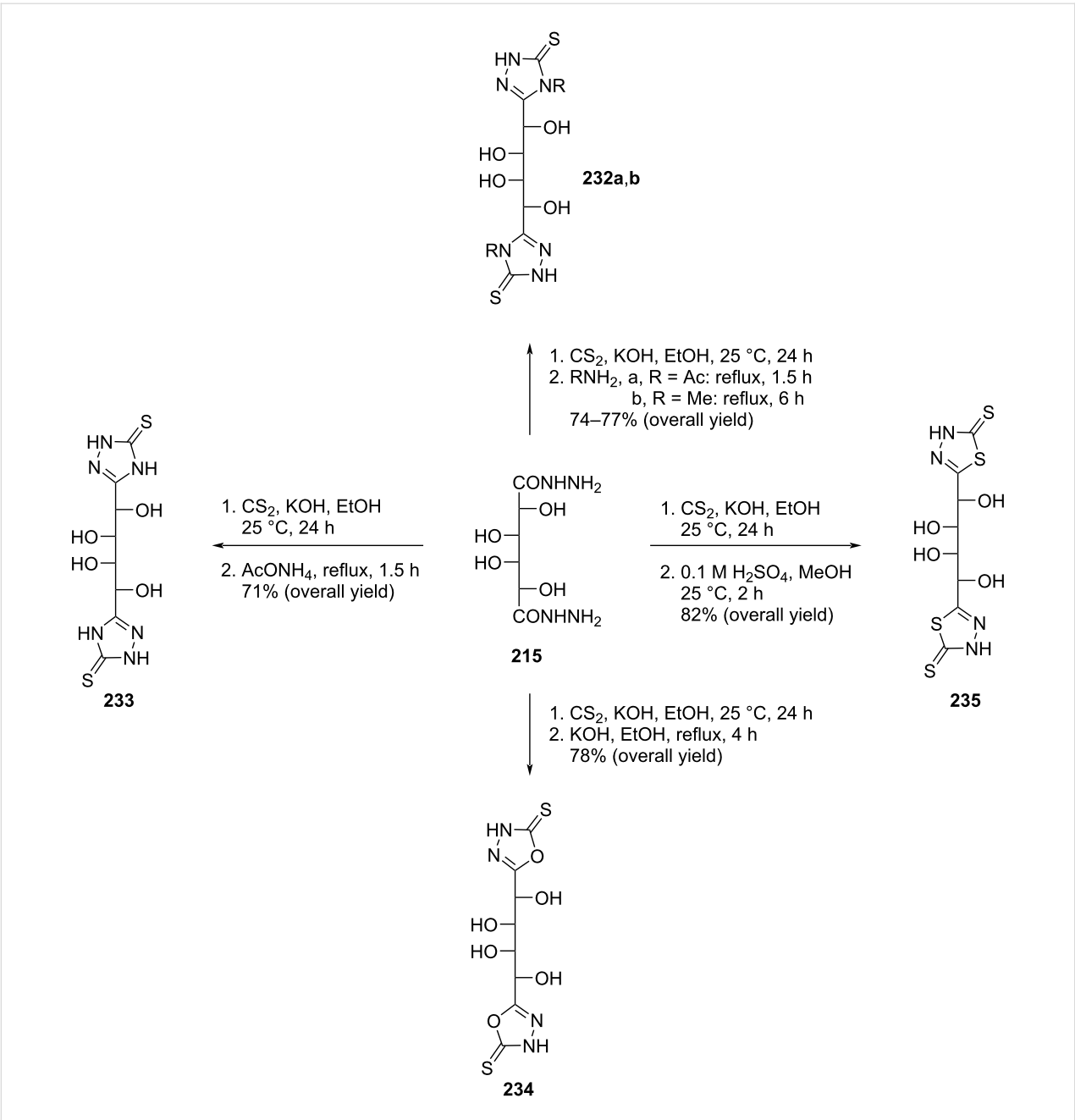
Galactaric acid (**243**) was diesterified with ethanol in the presence of conc. sulfuric acid. The corresponding diethyl ester **244** was treated with thiocarbohydrazide in a fusion reaction to produce compound **245**, which upon treatment with acetic anhydride under heating conditions, afforded the acyclic double-headed acyclic *C*-nucleoside **246** (Scheme 64) [21].

Some 3,6-disubstituted 1,2,4-triazolo[3,4-*b*]1,3,4-thiadiazole derivatives possess anti-HIV-I [105], anti-inflammatory [106,107], anticancer [108,109], and antibacterial properties [110]. The double-headed *C*-nucleoside **246** comprises an alditolyl moiety attached at position 3 of the 1,2,4-triazolo[3,4-*b*]1,3,4-thiadiazole core which can enhance the biological activity due to the hydrophilic nature of the alditolyl residue which may further increase the transportation into biological systems [21].

Compound **247** was treated with adenine in the presence of sodium hydride in DMF at 105 °C to incorporate two adenine moieties affording compound **248**. The benzoylation of compound **248**, followed by treatment with methanolic ammonia at low temperature produced the corresponding *N*-benzoylated adenine derivative **249** [111]. The cleavage of the diacetal in compound **249** was achieved with 75% TFA/water resulting in compound **250**, which was considered as the acyclic double-headed nucleoside without any protection of the primary hydroxy groups (Scheme 65) [111].

In a similar reaction sequence, compound **251** was treated with N^3-benzoylthymine to afford compound **252**, which was treated with 75% TFA–water for deprotection of the hydroxy groups to afford the final monomer **253** (Scheme 66) [111].

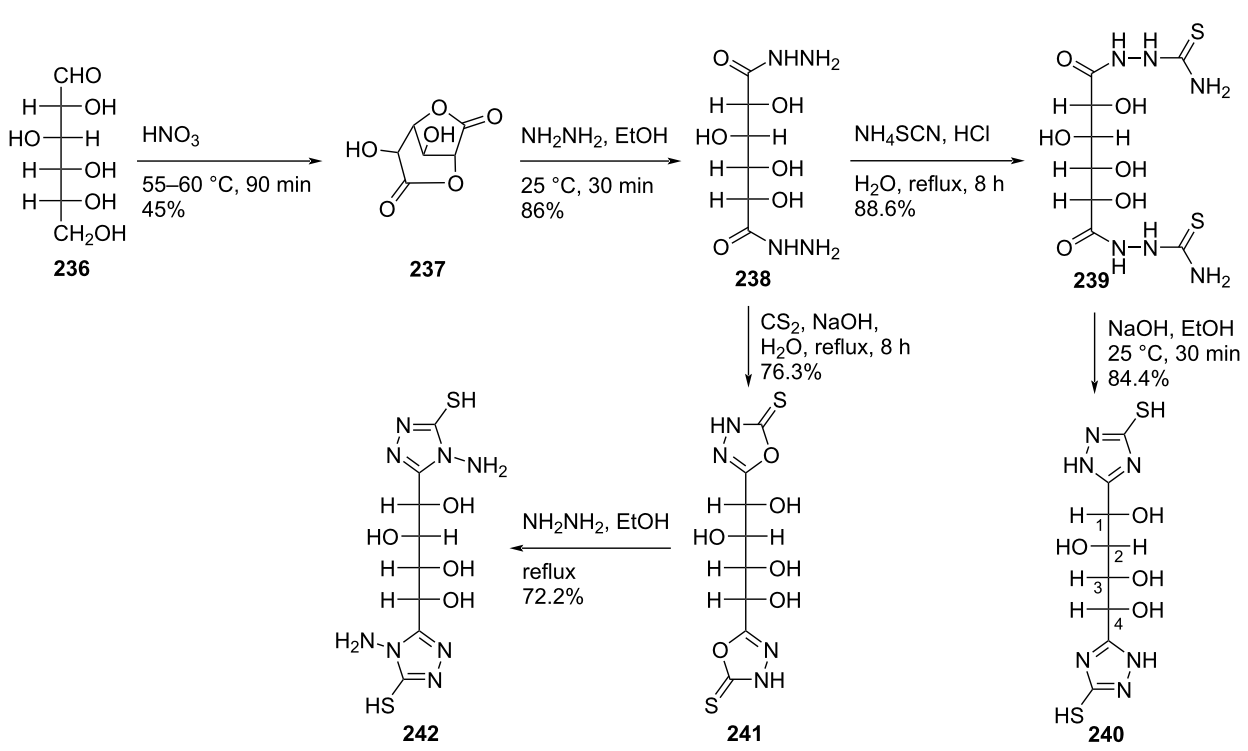
The synthesized acyclo nucleosides **250** and **253** were phosphitylated and incorporated into oligonucleotides to evaluate the effects on duplex stability. It was observed that the hybridization properties of the oligonucleotides with one acyclic achiral nucleoside, i.e., **250** or **253** when incorporated in the middle of a 12-mer or 13-mer decreased with complementary DNA or RNA [111].



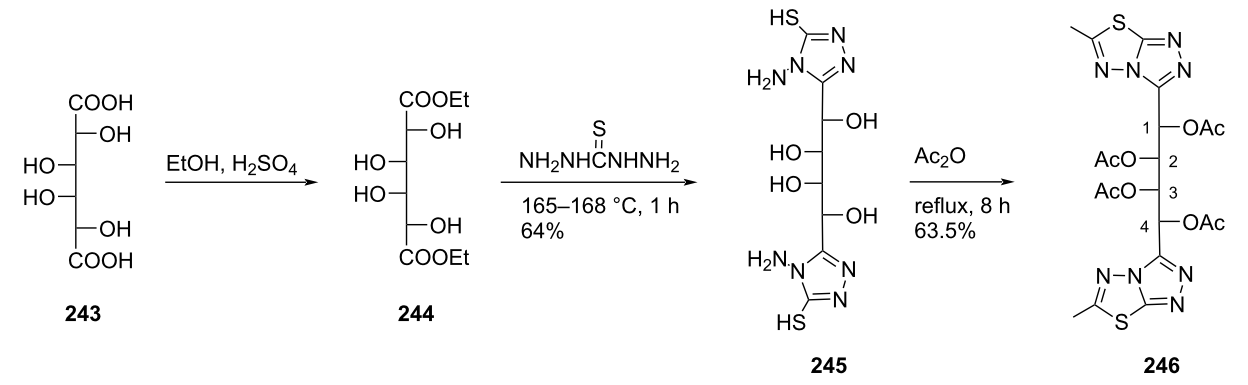
Scheme 62: Synthesis of double-headed 1,3,4-thiadiazoline, 1,3,4-oxadiazoline, and 1,2,4-triazoline acyclo *C*-nucleosides **232a**,**b** and **233**–**235**.

Four pyrimidine nucleobases **254a–d** were treated with methyl iodide in the presence of sodium hydroxide to get methylthio derivatives **255a–d**, which were treated with 2,2-bis(bromomethyl)-1,3-diacetoxypropane (**256**) in the presence of NaH in DMF to afford the mono-headed acyclic nucleosides **257a–d** [112]. The second nucleobase was introduced in compounds **257a**–**d** by repeating the reaction with the desired nucleobase under otherwise identical conditions (NaH/DMF) giving the acyclic double-headed pyrimidine nucleosides **258a–d**. Finally, the treatment of compounds **258a**–**d** with NaOMe in methanol produced the unprotected nucleosides **259a–d** (Scheme 67) [112].

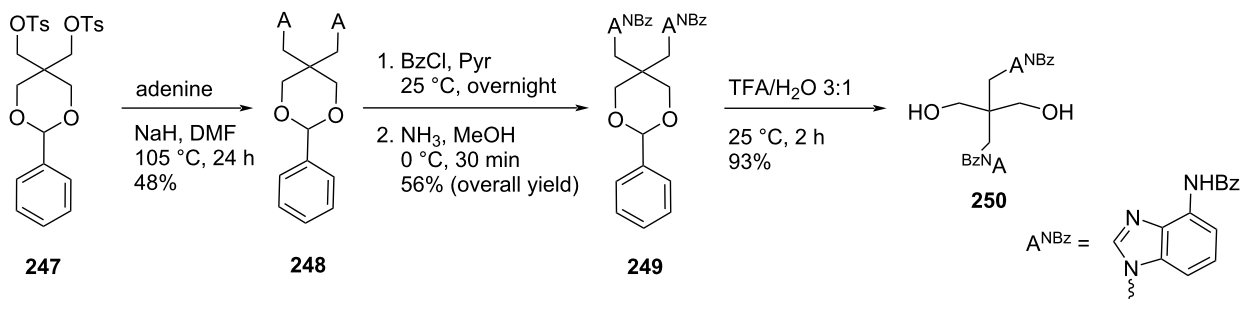
The double-headed nucleoside **261** was obtained by a two-step reaction sequence starting from compound **256**, which was first reacted with theophylline in DMF, to give the acyclic double-headed purine nucleoside **260** followed by treatment with NaOMe in methanol to get the unprotected product **261**. In the sequence, all reactions were carried out under microwave irradiation conditions (Scheme 68) [112].



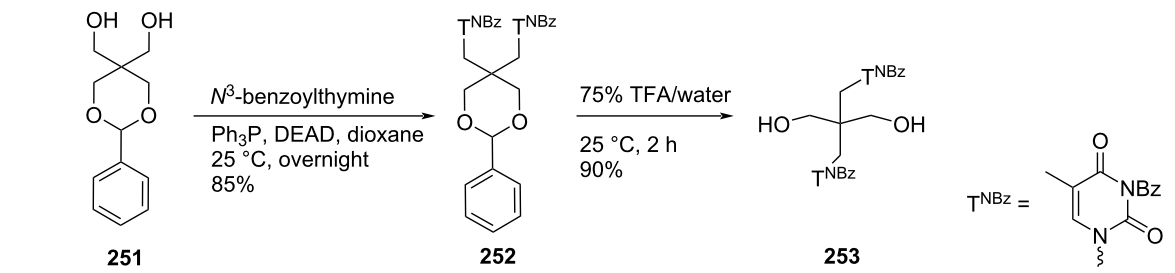
Scheme 63: Synthesis of double-headed acyclo *C*-nucleosides **240**–**242**.



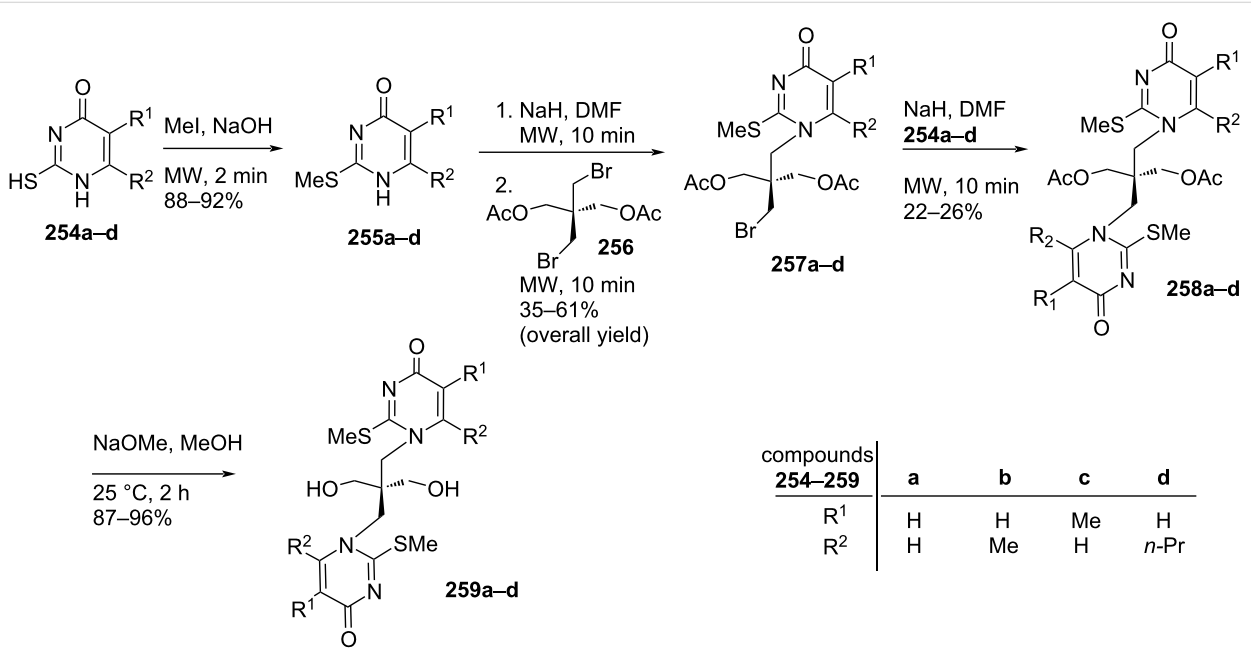
Scheme 64: Synthesis of double-headed acyclo *C*-nucleoside **246**.



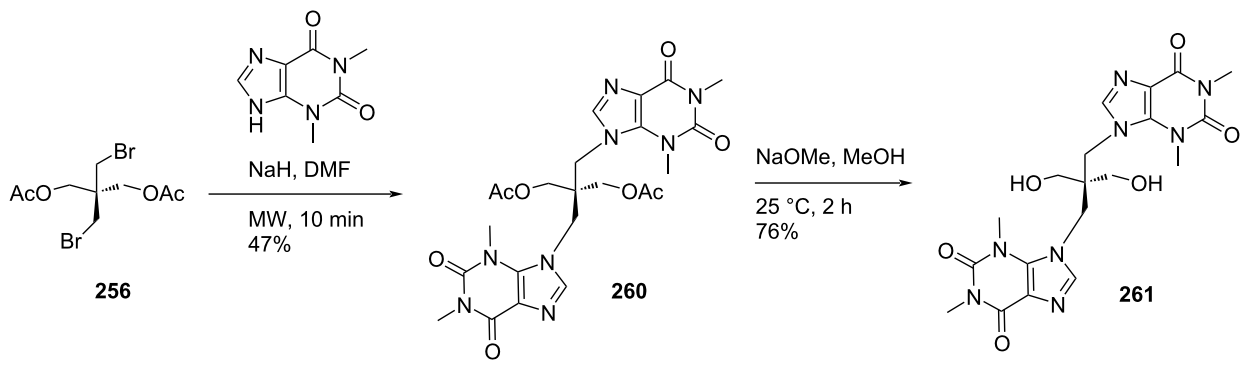
Scheme 65: Synthesis of acyclo double-headed nucleoside **250**.



Scheme 66: Synthesis of acyclo double-headed nucleoside **253**.



Scheme 67: Synthesis of acyclo double-headed nucleosides **259a–d**.



Scheme 68: Synthesis of acyclo double-headed nucleoside **261**.

The branched chain tetraseco-nucleosides **259a–d** and **261** were synthesized because acyclic nucleosides of tetraseco-type were found to possess interesting antiviral activities [21,113-115].

Conclusion

Among the variety of modified nucleosides, double-headed nucleoside monomers are an important class of compounds,

which have shown their importance in nucleoside chemistry. Here, we have focused on the available methodologies for the synthesis of several double-headed nucleosides. For a systematic discussion, we have classified them into three different categories, i.e., double-headed nucleosides with additional head/nucleobase on the sugar moieties, nucleobase moieties and on acyclic carbohydrate moieties. We have subdivided the category of monocyclic furanosyl double-headed nucleosides into 1,2-furanosyl-, 1,3-furanosyl-, 1,4-furanosyl-, and 1,5-furanosyl double-headed nucleosides depending on the position of the aglycon moiety in the furanosyl ring and systematically described their synthetic methodologies. Next, we elaborated the procedures for the synthesis of bicyclic furanosyl double-headed nucleosides, followed by procedures for the development of base to base double-headed nucleosides. The chemical strategies for the synthesis of pyranosyl double-headed nucleosides and acyclic double-headed nucleosides were also described. Along with the methodologies for the development of double-headed nucleoside monomers, the synthetic approach for their incorporation into the oligonucleotides was also elaborated in this review. Biological applications of the synthesized nucleosides were also described.

Future Direction

Double-headed nucleosides are important structural scaffolds that modulate nucleic acid structures. Rationally designed nucleosides can tune interstrand and intrastrand interactions that are exhibited in nucleic acids. As a consequence, these synthetic scaffolds can be exploited rationally in biomolecular designs and medicinal chemistry. These modified double-headed nucleosides could be incorporated into oligonucleotides to explore their potential as antisense nucleosides. Similarly, as some of these nucleosides have shown their potential as antimicrobial agents, they could be explored extensively for their biological activity. This review will help researchers to get an insight into the available procedures for the synthesis of double-headed nucleosides and briefly explores their role in modulating nucleic acid structures and in medicinal chemistry. The researchers working in the field of modified nucleosides will be encouraged further to take up challenges for the synthesis of currently unexplored double-headed nucleosides with extensive configurations, connectivity through different linkers, and exploration of different purine and pyrimidine moieties as nucleobases.

Funding

We are grateful to the Institute of Eminence, University of Delhi for providing financial support under R&D program. V.V. and V.K.M. thank CSIR, New Delhi and CFEES (DRDO), New Delhi, respectively for the award of Junior/Senior Research Fellowships.

ORCID® iDs

Himal K. Ganguly - https://orcid.org/0000-0001-6142-4155

References

1. Blackburn, G. M.; Gait, M. J.; Loakes, D.; Williams, D. M., Eds. *Nucleic Acids in Chemistry and Biology,* 3rd ed.; Royal Society of Chemistry: Cambridge, UK, 2006. doi:10.1039/9781847555380
2. Neidle, S. *Principles of Nucleic Acid Structure;* Elsevier: Amsterdam, Netherlands, 2008. doi:10.1016/b978-0-12-369507-9.x5001-8
3. Herdewijn, P., Ed. *Modified Nucleosides: in Biochemistry, Biotechnology and Medicine;* Wiley-VCH: Weinheim, Germany, 2008. doi:10.1002/9783527623112
4. Simons, C. *Nucleoside Mimetics: Their Chemistry and Biological Properties;* CRC Press: Boca Raton, FL, USA, 2000. doi:10.1201/9781482283143
5. Mangla, P.; Maity, J.; Rungta, P.; Verma, V.; Sanghvi, Y. S.; Prasad, A. K. *ChemistrySelect* **2019,** *4,* 3241–3246. doi:10.1002/slct.201900809
6. Ichikawa, E.; Kato, K. *Curr. Med. Chem.* **2001,** *8,* 385–423. doi:10.2174/0929867013373471
7. McKenzie, L. K.; El-Khoury, R.; Thorpe, J. D.; Damha, M. J.; Hollenstein, M. *Chem. Soc. Rev.* **2021,** *50,* 5126–5164. doi:10.1039/d0cs01430c
8. Jahnz-Wechmann, Z.; Framski, G. R.; Januszczyk, P. A.; Boryski, J. *Front. Chem. (Lausanne, Switz.)* **2016,** *4,* No. 19. doi:10.3389/fchem.2016.00019
9. Townsend, L. B. *Chemistry of Nucleosides and Nucleotides;* Springer Science + Business Media: New York, NY, USA, 1994; Vol. 3. doi:10.1007/978-1-4757-9667-4
10. Lamberth, C. *Org. Prep. Proced. Int.* **2002,** *34,* 149–167. doi:10.1080/00304940209355752
11. Zhang, X.; Amer, A.; Fan, X.; Balzarini, J.; Neyts, J.; De Clercq, E.; Prichard, M.; Kern, E.; Torrence, P. F. *Bioorg. Chem.* **2007,** *35,* 221–232. doi:10.1016/j.bioorg.2006.11.003
12. Timoshchuk, V. A.; Hogrefe, R. I. *Nucleosides, Nucleotides Nucleic Acids* **2009,** *28,* 464–472. doi:10.1080/15257770903044598
13. Horton, D.; Tsai, J.-H. *Carbohydr. Res.* **1979,** *75,* 141–150. doi:10.1016/s0008-6215(00)84634-0
14. Shaikh, K. I.; Madsen, C. S.; Nielsen, L. J.; Jørgensen, A. S.; Nielsen, H.; Petersen, M.; Nielsen, P. *Chem. – Eur. J.* **2010,** *16,* 12904–12919. doi:10.1002/chem.201001253
15. Kočalka, P.; Andersen, N. K.; Jensen, F.; Nielsen, P. *ChemBioChem* **2007,** *8,* 2106–2116. doi:10.1002/cbic.200700410
16. Nasr, A. Z. *Nucleosides, Nucleotides Nucleic Acids* **2004,** *23,* 1825–1832. doi:10.1081/ncn-200040617
17. Moustafa, A. H.; Haggam, R. A.; Younes, M. E.; El Ashry, E. S. H. *Nucleosides, Nucleotides Nucleic Acids* **2005,** *24,* 1885–1894. doi:10.1080/15257770500268962
18. Nasr, A. Z. *J. Chin. Chem. Soc.* **2005,** *52,* 519–524. doi:10.1002/jccs.200500075
19. Nasr, A. Z.; Mostafa, M. A. *Phosphorus, Sulfur Silicon Relat. Elem.* **2005,** *180,* 1769–1779. doi:10.1080/104265090888397
20. Amara, S.; Othman, A. A. *Arabian J. Chem.* **2016,** *9,* S1840–S1846. doi:10.1016/j.arabjc.2012.05.005
21. Awad, L. F.; El Ashry, E. S. H. *Carbohydr. Res.* **1998,** *312,* 9–22. doi:10.1016/s0008-6215(98)00205-5
22. Hornum, M.; Kumar, P.; Podsiadly, P.; Nielsen, P. *J. Org. Chem.* **2015,** *80,* 9592–9602. doi:10.1021/acs.joc.5b01577

23. Kumar, P.; Sharma, P. K.; Nielsen, P. *Bioorg. Med. Chem.* **2017,** *25,* 2084–2090. doi:10.1016/j.bmc.2017.02.011
24. Kumar, P.; Østergaard, M. E.; Baral, B.; Anderson, B. A.; Guenther, D. C.; Kaura, M.; Raible, D. J.; Sharma, P. K.; Hrdlicka, P. J. *J. Org. Chem.* **2014,** *79,* 5047–5061. doi:10.1021/jo500614a
25. Østergaard, M. E.; Guenther, D. C.; Kumar, P.; Baral, B.; Deobald, L.; Paszczynski, A. J.; Sharma, P. K.; Hrdlicka, P. J. *Chem. Commun.* **2010,** *46,* 4929–4931. doi:10.1039/c0cc01133a
26. Andersen, N. K.; Chandak, N.; Brulíková, L.; Kumar, P.; Jensen, M. D.; Jensen, F.; Sharma, P. K.; Nielsen, P. *Bioorg. Med. Chem.* **2010,** *18,* 4702–4710. doi:10.1016/j.bmc.2010.05.019
27. Kumar, P.; Chandak, N.; Nielsen, P.; Sharma, P. K. *Bioorg. Med. Chem.* **2012,** *20,* 3843–3849. doi:10.1016/j.bmc.2012.04.036
28. Kumar, P.; Chandak, N.; Nielsen, P.; Sharma, P. K. *Trends Carbohydr. Res.* **2016,** *8,* 9–14.
29. Sharma, P. K.; Kumar, P.; Nielsen, P. *Aust. J. Chem.* **2016,** *69,* 1094–1101. doi:10.1071/ch16021
30. Andersen, C.; Sharma, P. K.; Christensen, M. S.; Steffansen, S. I.; Madsen, C. M.; Nielsen, P. *Org. Biomol. Chem.* **2008,** *6,* 3983–3988. doi:10.1039/b810930c
31. Christensen, M. S.; Madsen, C. M.; Nielsen, P. *Org. Biomol. Chem.* **2007,** *5,* 1586–1594. doi:10.1039/b700852j
32. Hornum, M.; Sharma, P. K.; Reslow-Jacobsen, C.; Kumar, P.; Petersen, M.; Nielsen, P. *Chem. Commun.* **2017,** *53,* 9717–9720. doi:10.1039/c7cc05047j
33. Madsen, C. S.; Witzke, S.; Kumar, P.; Negi, K.; Sharma, P. K.; Petersen, M.; Nielsen, P. *Chem. – Eur. J.* **2012,** *18,* 7434–7442. doi:10.1002/chem.201103467
34. Kumar, P.; Sharma, P. K.; Madsen, C. S.; Petersen, M.; Nielsen, P. *ChemBioChem* **2013,** *14,* 1072–1074. doi:10.1002/cbic.201300086
35. Pedersen, S. L.; Nielsen, P. *Org. Biomol. Chem.* **2005,** *3,* 3570–3575. doi:10.1039/b510167k
36. Chatzileontiadou, D. S. M.; Parmenopoulou, V.; Manta, S.; Kantsadi, A. L.; Kylindri, P.; Griniezaki, M.; Kontopoulou, F.; Telopoulou, A.; Prokova, H.; Panagopoulos, D.; Boix, E.; Balatsos, N. A. A.; Komiotis, D.; Leonidas, D. D. *Bioorg. Chem.* **2015,** *63,* 152–165. doi:10.1016/j.bioorg.2015.10.007
37. Kiritsis, C.; Manta, S.; Papasotiriou, I.; Coutouli-Argyropoulou, E.; Trakossas, S.; Balzarini, J.; Komiotis, D. *Med. Chem.* **2012,** *8,* 320–329. doi:10.2174/157340612800786624
38. Kumar, P.; Sharma, P. K.; Nielsen, P. *J. Org. Chem.* **2014,** *79,* 11534–11540. doi:10.1021/jo502189h
39. Hornum, M.; Stendevad, J.; Sharma, P. K.; Kumar, P.; Nielsen, R. B.; Petersen, M.; Nielsen, P. *Chem. – Eur. J.* **2019,** *25,* 7387–7395. doi:10.1002/chem.201901077
40. Lemaire, S.; Houpis, I.; Wechselberger, R.; Langens, J.; Vermeulen, W. A. A.; Smets, N.; Nettekoven, U.; Wang, Y.; Xiao, T.; Qu, H.; Liu, R.; Jonckers, T. H. M.; Raboisson, P.; Vandyck, K.; Nilsson, K. M.; Farina, V. *J. Org. Chem.* **2011,** *76,* 297–300. doi:10.1021/jo101822j
41. Corey, E. J.; Chaykovsky, M. *J. Am. Chem. Soc.* **1965,** *87,* 1353–1364. doi:10.1021/ja01084a034
42. Beck, K.; Reslow-Jacobsen, C.; Hornum, M.; Henriksen, C.; Nielsen, P. *Bioorg. Med. Chem. Lett.* **2019,** *29,* 740–743. doi:10.1016/j.bmcl.2019.01.002
43. Jorgensen, A. S.; Shaikh, K. I.; Enderlin, G.; Ivarsen, E.; Kumar, S.; Nielsen, P. *Org. Biomol. Chem.* **2011,** *9,* 1381–1388. doi:10.1039/c0ob00438c
44. Kirschenheuter, G. P.; Zhai, Y.; Pieken, W. A. *Tetrahedron Lett.* **1994,** *35,* 8517–8520. doi:10.1016/s0040-4039(00)78425-5
45. Kumar, P.; Madsen, C. S.; Nielsen, P. *Bioorg. Med. Chem. Lett.* **2013,** *23,* 6847–6850. doi:10.1016/j.bmcl.2013.10.006
46. Cieślak, J.; Grajkowski, A.; Ausín, C.; Gapeev, A.; Beaucage, S. L. *Nucleic Acids Res.* **2012,** *40,* 2312–2329. doi:10.1093/nar/gkr896
47. Costa, A. M.; Faja, M.; Farràs, J.; Vilarrasa, J. *Tetrahedron Lett.* **1998,** *39,* 1835–1838. doi:10.1016/s0040-4039(98)00100-2
48. Li, N.-S.; Piccirilli, J. A. *J. Org. Chem.* **2004,** *69,* 4751–4759. doi:10.1021/jo0495337
49. Sano, T.; Shuto, S.; Inoue, H.; Ueda, T. *Chem. Pharm. Bull.* **1985,** *33,* 3617–3662. doi:10.1248/cpb.33.3617
50. Gao, Z.-G.; Duong, H. T.; Sonina, T.; Kim, S.-K.; Van Rompaey, P.; Van Calenbergh, S.; Mamedova, L.; Kim, H. O.; Kim, M. J.; Kim, A. Y.; Liang, B. T.; Jeong, L. S.; Jacobson, K. A. *J. Med. Chem.* **2006,** *49,* 2689–2702. doi:10.1021/jm050968b
51. Lazrek, H. B.; Taourirte, M.; Oulih, T.; Kabbaj, Y.; Barascut, J. L.; Imbach, J. L.; Almasoudi, N. A.; Pfleiderer, W. *Nucleosides Nucleotides* **1997,** *16,* 1073–1077. doi:10.1080/07328319708006135
52. Horwitz, J. P.; Chua, J.; Noel, M. *J. Org. Chem.* **1964,** *29,* 2076–2078. doi:10.1021/jo01030a546
53. Epp, J. B.; Widlanski, T. S. *J. Org. Chem.* **1999,** *64,* 293–295. doi:10.1021/jo981316g
54. Wu, T.; Nauwelaerts, K.; Van Aerschot, A.; Froeyen, M.; Lescrinier, E.; Herdewijn, P. *J. Org. Chem.* **2006,** *71,* 5423–5431. doi:10.1021/jo052194c
55. O-Yang, C.; Wu, H. Y.; Fraser-Smith, E. B.; Walker, K. A. M. *Tetrahedron Lett.* **1992,** *33,* 37–40. doi:10.1016/s0040-4039(00)77667-2
56. Jones, G. H.; Taniguchi, M.; Tegg, D.; Moffatt, J. G. *J. Org. Chem.* **1979,** *44,* 1309–1317. doi:10.1021/jo01322a025
57. Fecher, R.; Boswell, K. H.; Wittick, J. J.; Shen, T. Y. *J. Am. Chem. Soc.* **1970,** *92,* 1400–1402. doi:10.1021/ja00708a051
58. Fecher, R.; Boswell, K. H.; Wittick, J. J.; Shen, T. Y. *Carbohydr. Res.* **1970,** *13,* 105–111. doi:10.1016/s0008-6215(00)84900-9
59. Kašnar, B.; Škarić, V.; Klaić, B.; Žinić, M. *Tetrahedron Lett.* **1993,** *34,* 4997–5000. doi:10.1016/s0040-4039(00)74067-6
60. Montgomery, J. A.; Hewson, K. In *Synthetic Procedures in Nucleic Acid Chemistry;* Zorbach, W. W.; Tipson, R. S., Eds.; Interscience Publishers, John Wiley & Sons Ltd.: London–New York, 1968; Vol. 1, pp 180–182.
61. Baker, J. J.; Mellish, P.; Riddle, C.; Somerville, A. R.; Tittensor, J. R. *J. Med. Chem.* **1974,** *17,* 764–766. doi:10.1021/jm00253a027
62. Wang, G.; Middleton, P. J. *Tetrahedron Lett.* **1996,** *37,* 2739–2742. doi:10.1016/0040-4039(96)00418-2
63. Sharma, P. K.; Mikkelsen, B. H.; Christensen, M. S.; Nielsen, K. E.; Kirchhoff, C.; Pedersen, S. L.; Sørensen, A. M.; Østergaard, K.; Petersen, M.; Nielsen, P. *Org. Biomol. Chem.* **2006,** *4,* 2433–2445. doi:10.1039/b603830a
64. Borsting, P.; Nielsen, K. E.; Nielsen, P. *Org. Biomol. Chem.* **2005,** *3,* 2183–2190. doi:10.1039/b502720a
65. Andersen, C.; Sharma, P. K.; Christensen, M. S.; Pedersen, N. S.; Nielsen, P. *Nucleic Acids Symp. Ser.* **2008,** *52,* 275–276. doi:10.1093/nass/nrn139
66. Fensholdt, J.; Wengel, J. *Acta Chem. Scand.* **1996,** *50,* 1157–1163.

67. Kumar, S.; Steffansen, S. I.; Albæk, N.; Nielsen, P. *Tetrahedron* **2014,** *70,* 583–589. doi:10.1016/j.tet.2013.12.013
68. Albæk, N.; Petersen, M.; Nielsen, P. *J. Org. Chem.* **2006,** *71,* 7731–7740. doi:10.1021/jo061225g
69. Umemoto, T.; Wengel, J.; Madsen, A. S. *Org. Biomol. Chem.* **2009,** *7,* 1793–1797. doi:10.1039/b901028a
70. Enderlin, G.; Nielsen, P. *J. Org. Chem.* **2008,** *73,* 6891–6894. doi:10.1021/jo801081t
71. Kumar, P.; Sharma, P. K.; Hansen, J.; Jedinak, L.; Reslow-Jacobsen, C.; Hornum, M.; Nielsen, P. *Bioorg. Med. Chem.* **2016,** *24,* 742–749. doi:10.1016/j.bmc.2015.12.043
72. Kumar, P.; Sorinas, A. F.; Nielsen, L. J.; Slot, M.; Skytte, K.; Nielsen, A. S.; Jensen, M. D.; Sharma, P. K.; Vester, B.; Petersen, M.; Nielsen, P. *J. Org. Chem.* **2014,** *79,* 8020–8030. doi:10.1021/jo501151w
73. Dalager, M.; Andersen, N. K.; Kumar, P.; Nielsen, P.; Sharma, P. K. *Org. Biomol. Chem.* **2015,** *13,* 7040–7049. doi:10.1039/c5ob00872g
74. Jacobsen, M. F.; Andersen, C. S.; Knudsen, M. M.; Gothelf, K. V. *Org. Lett.* **2007,** *9,* 2851–2854. doi:10.1021/ol071006x
75. Jacobsen, M. F.; Knudsen, M. M.; Gothelf, K. V. *J. Org. Chem.* **2006,** *71,* 9183–9190. doi:10.1021/jo061694i
76. Kumar, P.; Østergaard, M. E.; Hrdlicka, P. J. *Curr. Protoc. Nucleic Acid Chem.* **2011,** 4.43.1–4.43.22. doi:10.1002/0471142700.nc0443s44
77. Robins, M. J.; Barr, P. J. *J. Org. Chem.* **1983,** *48,* 1854–1862. doi:10.1021/jo00159a012
78. Graham, D.; Parkinson, J. A.; Brown, T. *J. Chem. Soc., Perkin Trans. 1* **1998,** 1131–1138. doi:10.1039/a707031d
79. Hurley, D. J.; Seaman, S. E.; Mazura, J. C.; Tor, Y. *Org. Lett.* **2002,** *4,* 2305–2308. doi:10.1021/ol026043x
80. Le, B. T.; Hornum, M.; Sharma, P. K.; Nielsen, P.; Veedu, R. N. *RSC Adv.* **2017,** *7,* 54542–54545. doi:10.1039/c7ra10964d
81. Østergaard, M. E.; Kumar, P.; Nichols, J.; Watt, A.; Sharma, P. K.; Nielsen, P.; Seth, P. P. *Nucleic Acid Ther.* **2015,** *25,* 266–274. doi:10.1089/nat.2015.0547
82. Goubet, A.; Chardon, A.; Kumar, P.; Sharma, P. K.; Veedu, R. N. *Bioorg. Med. Chem. Lett.* **2013,** *23,* 761–763. doi:10.1016/j.bmcl.2012.11.096
83. Knapp, D. M.; Gillis, E. P.; Burke, M. D. *J. Am. Chem. Soc.* **2009,** *131,* 6961–6963. doi:10.1021/ja901416p
84. Browne, D. L.; Harrity, J. P. A. *Tetrahedron* **2010,** *66,* 553–568. doi:10.1016/j.tet.2009.10.085
85. Ferrier, R. J.; Tyler, P. C. *J. Chem. Soc., Perkin Trans. 1* **1980,** 2767–2773. doi:10.1039/p19800002767
86. Verma, V.; Maikhuri, V. K.; Khatri, V.; Singh, A.; Prasad, A. K. *Synth. Commun.* **2021,** *51,* 446–452. doi:10.1080/00397911.2020.1836224
87. Khatri, V.; Kumar, A.; Singh, B.; Malhotra, S.; Prasad, A. K. *J. Org. Chem.* **2015,** *80,* 11169–11174. doi:10.1021/acs.joc.5b01933
88. Elhalabi, J.; Rice, K. G. *Nucleosides, Nucleotides Nucleic Acids* **2004,** *23,* 195–205. doi:10.1081/ncn-120027828
89. Christensen, M. S.; Bond, A. D.; Nielsen, P. *Org. Biomol. Chem.* **2008,** *6,* 81–91. doi:10.1039/b713888a
90. Daves, G. D., Jr. *Acta Pharm. Suec.* **1987,** *24,* 275–288.
91. El Ashry, E. S. H.; Rashed, N.; Shobier, A. H. *Pharmazie* **2000,** *55,* 251–262.
92. El Ashry, E. S. H.; Rashed, N.; Shobier, A. H. *Pharmazie* **2000,** *55,* 331–348.
93. El Ashry, E. S. H.; Rashed, N.; Shobier, A. H. *Pharmazie* **2000,** *56,* 403–415.
94. Joshi, K. C.; Jain, S. K.; Jain, A. K. *Curr. Sci.* **1982,** *51,* 346–348.
95. Omar, A.-M. M. E.; Eshba, N. H.; Aboushleib, H. M. *J. Heterocycl. Chem.* **1986,** *23,* 1731–1735. doi:10.1002/jhet.5570230626
96. Holla, B. S.; Udupa, K. V. *J. Indian Chem. Soc.* **1988,** *65,* 524–525.
97. Abdel-Latif, F. F.; Shaker, R. M.; Mahgoub, S. A.; Badr, M. Z. A. *J. Heterocycl. Chem.* **1989,** *26,* 769–772. doi:10.1002/jhet.5570260348
98. Shaban, M. A. E.; Nasr, A. Z.; Morgaan, A. E. A. *Farmaco* **1999,** *54,* 800–809. doi:10.1016/s0014-827x(99)00107-x
99. Kaminsky, D. Method for Producing Antihypertensive Activity. U.S. Patent US3,752,891, Aug 14, 1973.
100. Monge, A.; Palop, J. A.; Ramirez, C.; Fernandez-Alvarez, E. *Acta Farm. Bonaerense* **1987,** *8,* 157–162.
101. Tomchin, A. B.; Ignat'eva, M. A.; Masyuta, G. F. *Khim.-Farm. Zh.* **1972,** *6,* 23.
102. Eshaba, N. H.; Salama, H. M.; Labouta, I. M.; Omar, A.-M. M. E. *Pharmazie* **1999,** *54,* 580–587.
103. Yousef, T.; Tawil, G. G. *Pharmazie* **1980,** *35,* 698–701.
104. Amara, S.; Tchouar, N.; Belaidi, S. *Int. Lett. Chem., Phys. Astron.* **2015,** *61,* 1–11. doi:10.18052/www.scipress.com/ilcpa.61.1
105. Paolo, F.; Daniele, S.; Franca, S.; Noemi, P. *Farmaco* **1996,** *51,* 659–664.
106. Gupta, R.; Sudan, S.; Kachroo, P. L. *Indian J. Chem., Sect. B* **1996,** *35,* 718–720.
107. Gupta, R.; Sudan, S.; Kachroo, P. L. *J. Indian Chem. Soc.* **1996,** *73,* 625–626.
108. Sagredou, S.; Dalezis, P.; Nikoleousakos, N.; Nikolaou, M.; Voura, M.; Almpanakis, K.; Panayiotidis, M. I.; Sarli, V.; Trafalis, D. T. *OncoTargets Ther.* **2020,** *13,* 7369–7386. doi:10.2147/ott.s254856
109. Bujji, S.; Edigi, P. K.; Subhashini, N. J. P. *J. Heterocycl. Chem.* **2020,** *57,* 3318–3325. doi:10.1002/jhet.4047
110. Holla, B. S.; Shivanada, M. K.; Akberali, P. M.; Baliga, S. S. *Farmaco* **1996,** *51,* 785–792.
111. Wu, T.; Froeyen, M.; Schepers, G.; Mullens, K.; Rozenski, J.; Busson, R.; Van Aerschot, A.; Herdewijn, P. *Org. Lett.* **2004,** *6,* 51–54. doi:10.1021/ol0360647
112. Ashry, E. S. H. E.; Rashed, N.; Abdel-Rahman, A.; Awad, L. F.; Rasheed, H. A. *Nucleosides, Nucleotides Nucleic Acids* **2006,** *25,* 925–939. doi:10.1080/15257770600793919
113. El Ashry, E. S. H.; Abdel Rahman, A. A. H.; Rashed, N.; Rasheed, H. A. *Pharmazie* **1999,** *54,* 893–897.
114. Rashed, N.; Hamid, H. A.; Ramadan, E. S.; El Ashry, E. S. H. *Nucleosides Nucleotides* **1998,** *17,* 1373–1384. doi:10.1080/07328319808003476
115. Hamed, A.; Abo-Amaym, E. R.; El Ashry, E. S. H. *Nucleosides Nucleotides* **1998,** *17,* 1385–1407. doi:10.1080/07328319808003477

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Chemical approaches to discover the full potential of peptide nucleic acids in biomedical applications

Nikita Brodyagin[‡1,§], Martins Katkevics[‡2,§], Venubabu Kotikam[‡1,§], Christopher A. Ryan[‡1,§] and Eriks Rozners[*1]

Review

Open Access

Address:
[1]Department of Chemistry, Binghamton University, The State University of New York, Binghamton, New York 13902, United States and [2]Latvian Institute of Organic Synthesis, Aizkraukles 21, Riga, LV-1006, Latvia

Email:
Eriks Rozners[*] - erozners@binghamton.edu

* Corresponding author ‡ Equal contributors
§ The authors are listed in alphabetical order.

Keywords:
antisense; chemical modifications; diagnostics; peptide nucleic acid; PNA

Beilstein J. Org. Chem. **2021,** *17,* 1641–1688.
https://doi.org/10.3762/bjoc.17.116

Received: 01 April 2021
Accepted: 28 June 2021
Published: 19 July 2021

This article is part of the thematic issue "Celebrating the role of chemistry in the success of oligonucleotides as therapeutics".

Guest Editors: P. Kumar and T. Brown

Abstract

Peptide nucleic acid (PNA) is arguably one of the most successful DNA mimics, despite a most dramatic departure from the native structure of DNA. The present review summarizes 30 years of research on PNA's chemistry, optimization of structure and function, applications as probes and diagnostics, and attempts to develop new PNA therapeutics. The discussion starts with a brief review of PNA's binding modes and structural features, followed by the most impactful chemical modifications, PNA enabled assays and diagnostics, and discussion of the current state of development of PNA therapeutics. While many modifications have improved on PNA's binding affinity and specificity, solubility and other biophysical properties, the original PNA is still most frequently used in diagnostic and other in vitro applications. Development of therapeutics and other in vivo applications of PNA has notably lagged behind and is still limited by insufficient bioavailability and difficulties with tissue specific delivery. Relatively high doses are required to overcome poor cellular uptake and endosomal entrapment, which increases the risk of toxicity. These limitations remain unsolved problems waiting for innovative chemistry and biology to unlock the full potential of PNA in biomedical applications.

Introduction

Peptide nucleic acid (PNA) is a DNA mimic where the sugar–phosphate backbone of DNA is replaced with a neutral and achiral pseudopeptide backbone (Figure 1) [1]. PNA retains the natural DNA nucleobases that are connected to the amide-linked backbone through additional amide linkages. PNA was originally designed as a DNA mimic to improve the properties of triplex-forming oligonucleotides [1,2]. Two key considerations were elimination of electrostatic repulsion (neutral back-

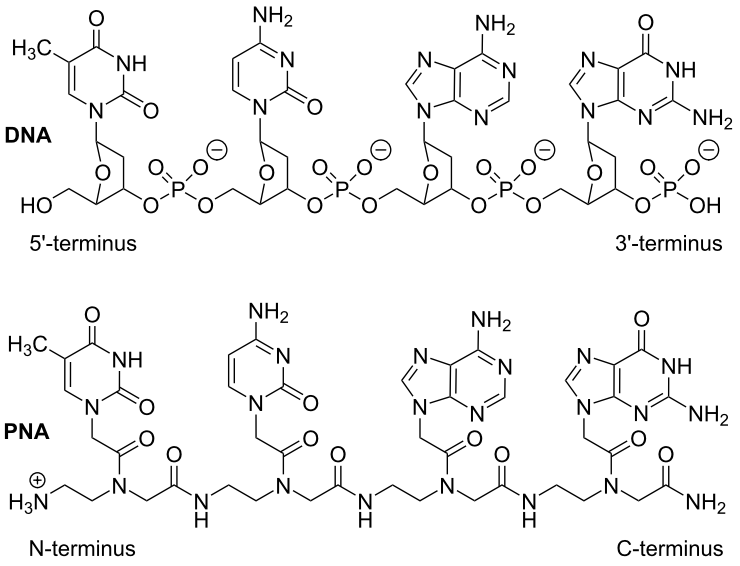


Figure 1: Structure of DNA and PNA.

bone) and synthetic accessibility (simple to make achiral amide linkages) [3]. The design was guided by a simple computer model where the natural sugar-phosphodiester backbone of the Hoogsteen strand of a T•A–T DNA triplex was replaced by an achiral and neutral pseudopeptide backbone having the same number of atoms [2,3]. It is remarkable that this simple design resulted in a nucleic acid analogue that had the right degree of flexibility and favorable conformational properties, enforced by the rotational preferences around amide linkages, to form strong and sequence specific complexes with natural DNA and RNA [3]. As will be discussed below, despite extensive studies [4-6], relatively few modifications have improved this simple original design.

Since its inception, PNA has become an extremely useful research tool and enabling component of many assays and diagnostics [4,7-9]. On the other hand, development of PNA based therapeutics has notably lagged behind other nucleic acid technologies [10,11]. In the present review, we summarize the remarkable journey of PNA from the initial design, through many chemical modifications and various applications, to the current state of the field. We also seek insights into the key question of why PNA, despite its impressive biophysical properties, has still not entered clinical trials.

The most significant difference between PNA and the natural nucleic acids is the lack of negative charge on PNA's backbone. Electrostatic repulsion of the negatively charged phosphates dominates the conformational properties and structure of nucleic acids. In contrast to proteins that prefer to fold in compact structures, DNA and RNA inherently prefer extended conformations that minimize the electrostatic repulsion. The maintenance and function of long double-stranded DNA (dsDNA) is achieved through complex mechanisms involving histones and other proteins. Large non-coding RNAs (e.g., ribosomes) manage electrostatic repulsion using positively charged RNA-binding proteins and cations (e.g., magnesium ions), and achieve remarkably complex folded structures. Nevertheless, the electrostatic repulsion is the main force that disfavors folding and association of nucleic acids. With this consideration in mind, neutral PNA was expected to have superior binding to negatively charged nucleic acids due to the lack of electrostatic repulsion [1-3].

As will be reviewed below, because of its robust metabolic stability and high affinity and sequence specificity, PNA has become a vital component of many research assays and diagnostics [4]. Nevertheless, PNA has not been without shortcomings and vulnerabilities. Limited water solubility, especially for purine rich sequences, was noted in early studies. To improve water solubility and decrease aggregation, typical PNA designs place a lysine at the C-terminus (Figure 1) introducing a second positive charge in addition to the charge at the N-terminus of PNA [1]. Even with the additional lysine, the solubility of PNA decreases as the polymer length increases. PNA solubility in the HEPES buffer at pH 7.3 and 37 °C is estimated to be in the 0.1–0.5 mM range [12,13]. The hydrophobic nature and lack of electrostatic repulsion of the PNA backbone favors folding in compact structures and aggregation in concentrated solutions [13].

Other bottlenecks for in vivo applications of PNA have been poor cellular uptake and unfavorable pharmacokinetics [14-16]. Unmodified PNAs are not taken up by eukaryotic cells in vitro and are cleared rapidly (within 10–30 min in mice) through the kidneys after administration to animals by either intravenous or intraperitoneal injection [16]. In another study, PNA elimination half-life in rats was ≈17 minutes and ≈90% of PNA was recovered unchanged in the urine 24 h after administration [17].

To address these problems, many research groups have worked on chemical modifications to the backbone and nucleobases of PNA, as well as conjugating PNA to other biomolecules (e.g., cell-penetrating peptides) [4]. The present review summarizes the most significant efforts and achievements in optimizing various aspects of PNA applications. We start with a brief review of PNA's binding modes and structural features, continue to the most impactful chemical modifications, PNA enabled assays and diagnostics, and finish with discussion of the current state of development of PNA therapeutics. The common theme that emerges is that despite extensive studies reviewed below, PNA still needs innovative chemistry to break through in clinic and other in vivo applications.

Review

PNA binding modes to DNA and RNA

PNA was originally designed with an expectation to improve the binding properties of negatively charged triplex-forming oligonucleotides to dsDNA [1,2]. The parallel PNA-dsDNA triplex, where the N-terminus of PNA aligns with the 5′-terminus of a polypurine strand of DNA (Figure 2A), is a binding mode that is particularly sensitive to electrostatic repulsion as three negatively charged strands are brought in proximity. PNA was also found to bind single-stranded DNA and RNA (ssDNA and ssRNA) in an antiparallel fashion (the C-terminus of PNA aligning with the 5′-terminus of ssDNA) with affinity and sequence selectivity significantly higher than that of the natural oligonucleotides [18,19]. The thermal stabilities of duplexes involving PNA usually follow the order PNA–PNA > PNA–RNA > PNA–DNA [20,21]. Hybridization of PNA with complementary nucleic acids is enthalpy driven, involving large favorable gains in enthalpy compensated by significant unfavorable entropy, as typically observed for nucleic acid complexes [22]. The binding is highly sequence specific as one Watson–Crick base pair mismatch can drop the melting temperature of the complex with PNA by 8–20 °C making PNA an excellent nucleic acid analogue for development of probes and diagnostics. This strong and selective binding has made PNA a key component of assays and diagnostics that depend on Watson–Crick hydrogen bonding to natural nucleic acids. An unexpected discovery of early studies was that the triplex-forming PNAs built of pyrimidine monomers formed a 2:1 PNA–DNA–PNA strand-invasion triplex instead of the expected 1:1 PNA–dsDNA triplex (c.f., Figure 2A and 2B) [1,23]. This unprecedented binding mode was enabled by PNA's unique ability to displace the pyrimidine-rich strand of dsDNA as the so-called P-loop, which was clearly facilitated by the neutral backbone [1].

Later studies showed that there was a delicate balance between the two binding modes. The strand invasion (Figure 2B) was favored at low ionic strength and high PNA concentration, and

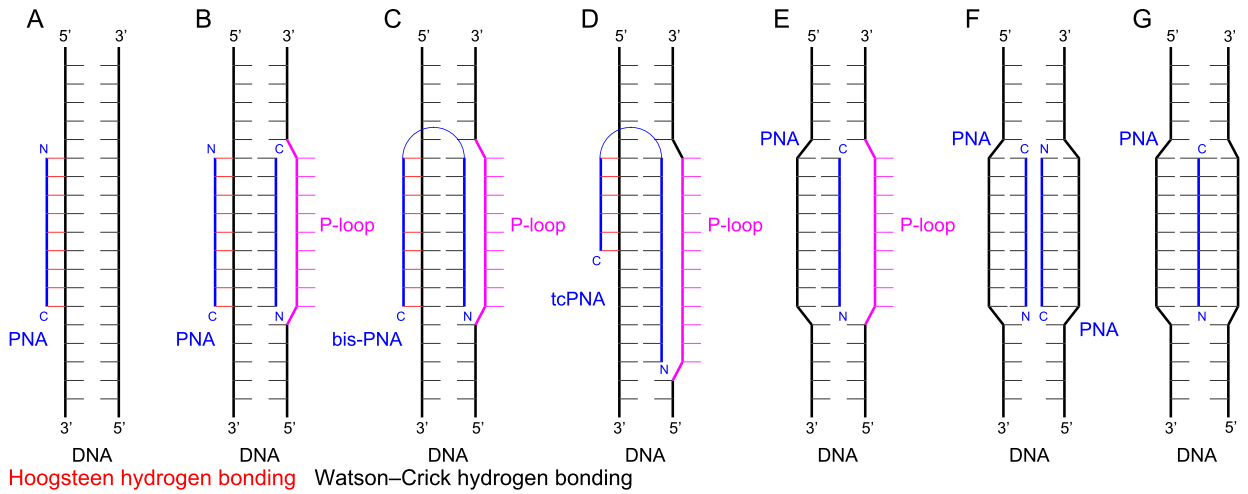


Figure 2: PNA binding modes: (A) PNA–dsDNA 1:1 triplex; (B) PNA–DNA–PNA strand-invasion triplex; (C) the Hoogsteen and Watson–Crick parts are linked together in a bis-PNA; (D) shortening the Hoogsteen part and extending the Watson–Crick part of the bis-PNA creates a tail-clamp PNA (tcPNA); (E) and (F) single and double invasion using only Watson–Crick hydrogen bonding; (G) Janus-wedge triple helix.

required longer reaction times [24]. In contrast, physiological ionic strength inhibited strand invasion and shifted the binding mode towards the major groove Hoogsteen triple helix (Figure 2A) [24]. The binding mode was also affected by PNA's sequence with thymine-rich PNAs generally preferring invasion complexes and cytosine-rich PNAs generally preferring triple helix formation [25]. Overall, while PNA formed stronger triple helices with dsDNA than negatively charged oligonucleotides, the stability of the triplexes was still lower than that of the Watson–Crick PNA–DNA and PNA–RNA duplexes and required a tract of at least 15 consecutive purines for chemically-modified triplex-forming PNA to achieve low nanomolar binding [26]. Triple-helical binding of PNA to dsRNA was not explored until 2010 when Rozners and co-workers showed that PNAs as short as hexamers formed strong and sequence specific triplexes at pH 5.5 [27]. Later studies using nucleobase-modified PNA (vide infra) confirmed that PNA had >10-fold higher affinity for dsRNA than for the same sequence of dsDNA [28-31].

While parallel PNA–DNA and PNA–RNA triple helices formed by PNAs built of C and T monomers are well documented (as reviewed above), the antiparallel triplexes formed by PNAs built of G and T or G and A monomers have not been reported. It is conceivable, that the limited solubility and tendency to aggregate prevent such binding modes involving purine-rich PNAs, as discussed in a recent review [32]. However, it is also possible that this is an underexplored PNA binding mode. G-rich PNAs do not form stable G-quadruplexes [33], which suggests that with innovative chemistry, it may be possible to explore G-rich PNAs for antiparallel triplexes.

The strand invasion complex contains two PNA molecules binding the purine-rich strand of DNA. While one PNA strand forms an antiparallel Watson–Crick duplex, the other strand forms a parallel Hoogsteen triplex, which brings the N- and C-ends of the two strands in proximity (Figure 2B). An innovative design links the two ends together with an ethylene glycol linker (Figure 2C), which reduced the unfavorable loss of entropy by converting the binding event from a trimolecular to a bimolecular process [34-36]. The new bis-PNAs (Figure 2C) showed about two orders of magnitude stronger binding (lower EC_{50}) to ssDNA targets compared to the trimolecular formation of the PNA–DNA–PNA triplex [35]. However, the need for polypurine tracts remained a limitation of bis-PNAs. A further development that extended the sequence scope that can be targeted by bis-PNAs was to shorten the Hoogsteen part and extend the Watson–Crick part of the bis-PNA by creating a tail-clamp PNA (tcPNA, Figure 2D) [37]. Tail-clamp PNAs are currently at the forefront of PNA therapeutic development (vide infra).

Single or double invasion of dsDNA (Figure 2E and 2F, respectively) using only Watson–Crick base pairing at mixed sequences that do not have polypurine tracts is also possible, but requires chemical modifications to alter the binding properties of PNAs. These binding modes further illustrate the diversity of molecular recognition that can be achieved with PNAs. Taken together, the early discoveries that revealed the remarkable nucleic acid binding properties of PNA boosted enthusiasm about PNA's potential as an antisense and antigene therapeutic agent [38].

Structures of PNA complexes

Early NMR structural studies suggested that PNA formed heteroduplexes with DNA [39] and RNA [40] that resembled the B- and A-form conformations of natural nucleic acids. The PNA–RNA duplex adopted a conformation very close to the standard A-form helix [40]. In contrast, the PNA–DNA duplex adopted an intermediate structure where positioning of the base pairs was A-like, while the backbone curvature, sugar conformation (C2′-endo), base pair inclination, and helical rise resembled B-DNA [39].

The first X-ray crystal structure of a PNA–DNA–PNA triplex revealed a previously unknown helix with a wide diameter of ≈26 Å (compared to 20 Å for A-form duplex) and a wide and deep major groove (Figure 3), given the name "P-form helix" by the study authors [41]. Despite the much larger displacement of the bases from the helix axis, the base stacking in the P-form helix resembles that of an A-form DNA duplex. The sugars of a DNA strand adopt C3′-endo conformations with an average interphosphate distance of ≈6 Å, which is similar to A-type DNA and RNA, and allows the O1P oxygen from each DNA phosphate to form a hydrogen bond to the amide proton of each residue of the PNA backbone of the Hoogsteen strand [41]. More recent structural work by Rozners and co-workers confirmed that the PNA–dsRNA triplex had similar structural features [42]. The hydrogen bonding between PNA and RNA backbones is most likely the reason behind the >10-fold higher stability PNA–dsRNA triplexes [28-31] (compared to PNA–dsDNA) that favor structures having the ideal interphosphate distance of ≈6 Å. In contrast, the interphosphate distances in B-form structures (preferred by DNA) would be ≈7 Å. Most likely, PNA–dsDNA triplexes must pay an energy penalty by compromising between different stabilizing interactions that favor either B-like or A-like structures, which results in overall lower stability than the PNA-dsRNA triplexes where the stabilizing interactions are better aligned.

The crystal structure of a self-complementary PNA–PNA duplex was very similar to the P-form helix showing a wide helix (28 Å diameter) with a very large pitch of ≈18 base pairs

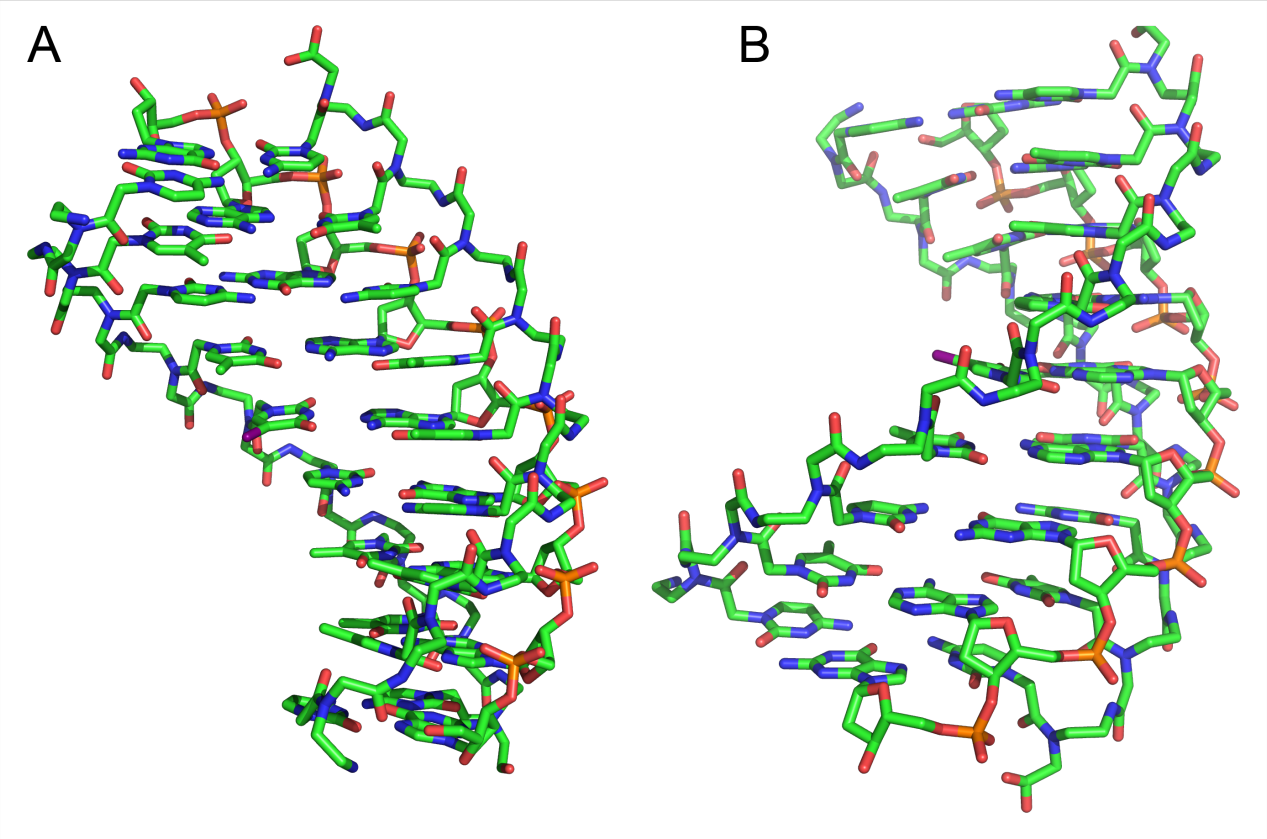


Figure 3: Structure of P-form PNA–DNA–PNA triplex from reference [41]. (A) view in the major groove and (B) view in the minor groove.

per turn, compared to 10 and 11 base pairs per turn for DNA and RNA, respectively, and a nucleobase stacking pattern similar to that of the A-form RNA [43]. Another crystal structure of a partially self-complementary PNA–PNA duplex revealed PNA's ability to combine the P-form Watson–Crick duplex with higher order structural features, such as reversed Hoogsteen base pairing, interstrand intercalation, triplex formation, and backbone chirality shifts [44]. A similar P-form helix having a wide and deep major groove and a shallow and narrow minor groove was also observed for an NMR solution structure of a self-complementary PNA–PNA duplex [45]. Taken together, these results confirmed that, while PNA was able to adopt to the conformations of DNA and RNA to some extent, the P-form was the naturally preferred helical conformation of PNA.

PNA backbone modifications

PNA design was originally assisted by simple computer modeling that replaced the phosphodiester backbone of DNA with pseudopeptide linkages having the same number of atoms and linking bonds [2]. Not surprisingly, backbone modification has been a major focus of follow up attempts to improve the original PNA design. Early studies showed that maintaining proper distances (number of bonds) along the backbone and between the backbone and nucleobases of PNA was critical for effective nucleic acid binding as extension of either by additional methylene groups strongly decreased the binding affinity of PNA to either single- or double-stranded nucleic acids [46-48]. Furthermore, replacing amide linkages connecting the PNA's backbone and the nucleobase with a tertiary amine also destabilized PNA complexes with complementary DNA [49]. The majority of the following studies focused on adding substituents to the original backbone for conformational control and improving PNA's biophysical properties.

Conformationally constrained backbones

Nielsen and co-workers [50] were the first to test restricting PNA backbone conformation by locking the backbone in a fused cyclohexane ring of either *S*,*S* or *R*,*R* configuration (chPNA, Figure 4). Both *S*,*S* or *R*,*R* chPNAs formed weaker complexes with complementary DNA and RNA than unmodified PNA [50]. Later, Kumar, Ganesh and co-workers [51-54] reported that either *S,R*- or *R,S*-modified chPNA had lower affinity for complementary DNA and RNA as well. The decreased binding affinity of chPNAs was most likely due to unfavorable dihedral angles for proper organization of PNA's backbone. In contrast, Appella and co-workers found that restricting the backbone's conformation with the fused *S,S*-cyclopentane

ring increased the binding affinity of cpPNA (Figure 4) for complementary DNA and RNA compared to the unmodified PNA [55,56]. Govindaraju, Kumar and Ganesh [57,58] reported that isolated *S*,*R*- and *R*,*S*-cyclopentane modifications had variable effects on PNA binding affinity depending on their location (C-terminus, middle, or N-terminus) in PNA, while fully *S*,*R*- and *R*,*S*-modified cpPNAs were binding stronger to complementary DNA and RNA than the unmodified PNA. The *R*,*S*-modified cpPNAs appeared to be somewhat stronger binders than the *S*,*R*-modified counterparts [57,58]. Interestingly, PNAs having constrained backbones, including modifications that lowered affinity, were more sequence selective (less tolerant to mismatches) than unmodified PNA, which is important for development of diagnostics and therapeutics.

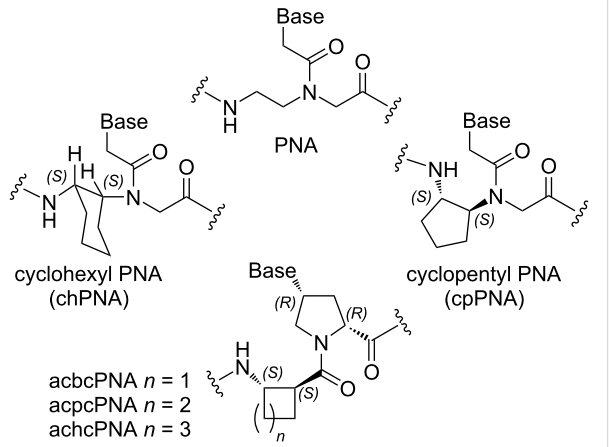


Figure 4: Structures of backbone-modified PNA.

Recently, more detailed biophysical and structural studies on *S,S*-cpPNA by Appella and co-workers [59,60] show that the *S,S*-configuration of cyclopentane modification enforces dihedral angles of PNA backbone favorable for binding to complementary DNA. PNA binding affinity and sequence selectivity increase with increasing number of *S,S*-cyclopentane modifications allowing rational fine tuning of the complex stability. The recently published crystal structure of a duplex between completely modified *S,S*-cpPNA and a complementary DNA strand reveals preorganization of PNA backbone into a right handed-helix favorable for DNA binding [60]. At the time of this writing, binding of *S,S*-cpPNA to complementary RNA remains less well explored; however, other constrained backbone-modified PNAs reviewed above have shown stronger binding to RNA over DNA. *S,S*-cpPNA may be expected to follow this trend and, at this time, appears to be the most promising conformationally constrained PNA analogue.

Vilaivan and co-workers developed pyrrolidinyl PNA based on an α/β-dipeptide backbone that is one atom longer than the canonical PNA and contains two amide bonds and two cyclic moieties in one monomer (Figure 4) [61]. Cyclobutane-derived acbcPNA and cyclopentane-derived acpcPNA formed stable duplexes with matching DNA and RNA, while cyclohexane-derived achcPNA did not form complexes with either DNA or RNA, which was explained by unfavorable torsional angles and conformational rigidity of the cyclohexane backbone [62]. Interestingly and in contrast to other backbone-constrained PNAs, pyrrolidinyl α/β-dipeptide PNA formed PNA–DNA complexes having higher thermal stability compared to PNA–RNA complexes [63,64]. Most likely, the one atom longer PNA backbone, which is rigidified and preorganized by cyclic moieties, may align better with the B-form DNA helix rather than with the A-form RNA helix. While pyrrolidinyl α/β-dipeptide PNAs formed stable antiparallel duplexes with DNA and RNA with high mismatch intolerance, due to constrained nature, two pyrrolidinyl α/β-dipeptide PNAs had low ability to self-hybridize [62,65]. This property makes pyrrolidinyl α/β-dipeptide PNA especially suitable for double duplex invasion of dsDNA [66]. In general, pyrrolidinyl α/β-dipeptide PNA is another promising modification along with cyclopentane constrained PNAs studied by Kumar, Ganesh, and Appella.

PNA modified at alpha and gamma positions of the backbone

α-Modified PNA: Adding substituents to the *N*-(2-aminoethyl)glycine backbone has been an obvious starting point for PNA modification. Nielsen and co-workers were the first to replace the glycine residues in PNA backbone with various chiral amino acids [67,68]. Most of these α-modified PNA monomers (Figure 5) slightly reduced PNA binding affinity, with D-amino acids being somewhat better accommodated in the backbone than L-amino acids and D-Lys being the only α-backbone modification that slightly increased PNA's binding affinity to complementary DNA (but not RNA) [67]. Circular dichroism studies showed that the D-Lys modification induced a right-handed helical conformation favorable for DNA binding while the L-Lys modification induced a left-handed helical conformation that disfavored PNA binding to DNA [69]. Interestingly, a crystal structure of PNA having three α-D-Lys modifications in the middle [70] resembled the P-form helices formed by PNA–PNA and PNA–DNA–PNA more than the PNA–DNA structure [39].

Ly and co-workers synthesized α-modified PNAs derived from L-arginine (α-GPNA, Figure 5) and showed that the positively charged guanidinium group increased the stability of PNA duplexes with complementary DNA and RNA, without compromising the sequence selectivity, and improved the cellular uptake of PNA [71]. The same group later demonstrated that GPNA derived from D-arginine formed more stable duplexes

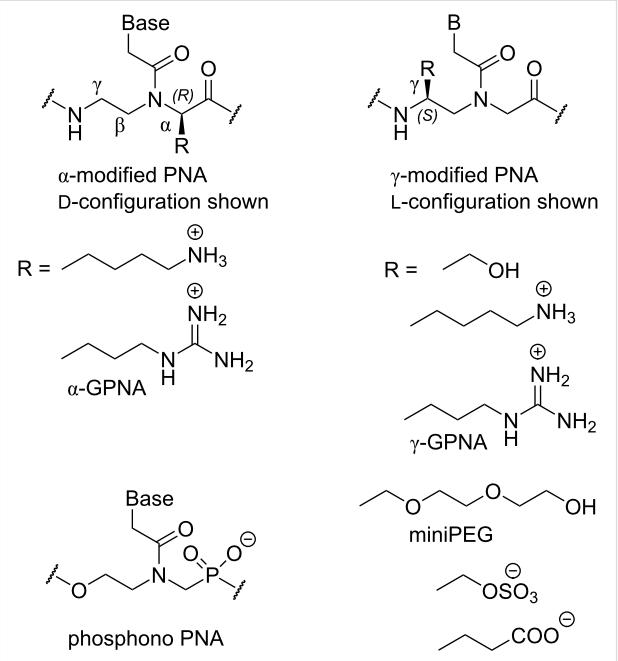


Figure 5: Structures of PNA having α- and γ-substituted backbones.

with RNA and was readily taken up by both human somatic and embryonic stem cells [72]. GPNA targeting the transcriptional start-site of the human E-cadherin gene had potent and sequence-specific antisense activity and was less toxic to the cells than the PNA–polyarginine conjugate [73]. Interestingly, the α-arginine modification in either L- or D-configuration destabilized PNA–dsRNA triplexes [74].

γ-Modified PNA: Later studies focused on introducing substituents in the ethylenediamine moiety of the PNA backbone. Ly and co-workers showed that introduction of simple substituents, such as methyl (derived from L-alanine) or hydroxymethyl (derived from L-serine) at the γ-position (Figure 5) preorganized the PNA backbone in a right-handed helical structure favorable for stronger binding to complementary DNA and RNA [75]. The NMR structure showed that γ-methyl-PNA folded in a P-form helix similar to that observed for non-modified PNA but having more resemblance to A-form [76]. The γ-methyl-PNA helix was slightly more unwound and had a smaller twist angle than the P-helix of unmodified PNA. In a crystal structure, γ-methyl-PNA–DNA heteroduplex also adopted a P-form helix, with greater resemblance to A-form than B-form DNA, accommodating 15 base pairs per turn [77]. Dynamic transitions between different binding modes of γ-hydroxymethyl-modified triplex-forming PNAs have been also explored [78].

Englund and Appella showed that PNA containing γ-modifications derived from L-lysine formed stronger duplexes with DNA and RNA, while γ-modifications derived from D-lysine decreased the stability of duplexes [79,80]. Ly and co-workers showed that γ-modified PNA derived from L-arginine (γ-GPNA, Figure 5) were preorganized into a right-handed helix, which improved their binding to complementary DNA and RNA while retaining sequence selectivity [81]. As expected, the guanidine modifications greatly improved cellular uptake of γ-GPNA. Others have also investigated positively charged α- and γ-modifications of the PNA backbone, and most of them showed promising hybridization properties and improved cellular uptake [82-86]. Very recent work has used α- and γ-positions of the PNA backbone to attach additional nucleobases, which enable these "double face" PNAs to form higher order double and triple helical structures [87,88].

Ly and co-workers followed up on the promising conformational properties of γ-hydroxymethyl PNA by extending the side chain into a miniPEG modification (Figure 5). In addition to retaining the superior nucleic acid binding (due to preorganization of PNA's backbone) miniPEG greatly improves aqueous solubility of PNA without causing any cytotoxicity [89]. Because of the superior binding properties, miniPEG-modified PNAs can invade any sequence of dsDNA using only Watson–Crick base pairing to recognize the target [89]. As will be discussed later in this review, PNAs having guanidine (γ-GPNA) and miniPEG γ-modifications are currently among the most promising PNA derivatives explored in medicinal chemistry and preclinical studies.

Anionic PNA: Anionic functionalities have been introduced in PNA to improve water solubility and better mimic DNA/RNA structure. One of the early studies was on chimeras of PNA and phosphono-PNA (Figure 5) that improved water solubility and in some cases resulted in stronger hybridization with complementary DNA and RNA [90]. The phosphono-PNAs retained the stability against nucleases. In another study, conjugation with glutamine phosphonate or lysine bis-phosphonate amino acid derivatives introduced up to twelve negative charges (phosphonate moieties) into PNAs [91]. The negative charges allowed cationic lipid-mediated delivery of PNAs to HeLa cells achieving sub-nanomolar antisense activity [91]. More recent studies introduced sulphate and carboxylate groups at the γ-position of PNA backbone (Figure 5) but neither modification showed promising hybridization profiles or improved cellular uptake [92,93].

Modified nucleobases in PNA

PNA nucleobases for Hoogsteen recognition of guanine: As discussed in the Introduction, PNA was originally designed with the idea that the neutral backbone would improve binding properties of triplex-forming oligonucleotides. However, elec-

trostatic repulsion is not the only weakness of triple helical recognition of nucleic acids. The binding affinity and sequence selectivity of triplex-forming oligonucleotides derives from thymine recognition of A–T (or A–U in RNA) base pairs (T•A–T or T•A–U triplet) and protonated cytosine recognition of G–C base pairs (C^+•G–C triplet) via Hoogsteen hydrogen-bonding (Figure 6) [94]. A significant bottleneck for triple helix formation is the requirement for cytosine protonation to form the natural C^+•G–C triplet. Because of the low pK_a of cytosine (≈4.5), formation of the C^+•G–C triplet is unfavorable at physiological pH, which severely destabilizes the parallel triple helices and limits their applications in biological systems.

Two obvious strategies to solve this problem are to modify the cytosine heterocycle to either 1) increase the pK_a or 2) create neutral analogues of protonated cytosine. In the latter strategy, Ono et. al. introduced pseudoisocytosine (J, Figure 6) in triplex-forming oligonucleotides, alleviating the problem of unfavorable cytosine protonation [95,96]. Nielsen and co-workers replaced Cs with Js in the Hoogsteen strand of their original design of bis-PNAs in 1995 [34]. While J demonstrated weaker binding than C at pH 5, J enabled formation of relatively stable triplexes at physiological pH of 7.4. Later, the same research group reported that 1,8-naphthyridin-2,7-(1,8*H*)-dione (K, Figure 6), a bicyclic mimic of protonated cytosine, afforded stronger binding to G–C base pairs compared to J, most likely due to the increased surface area of the bicyclic nucleobases that enabled better π-stacking [97]. Despite the superior binding properties, the original report on K has not been followed up with more detailed studies and J remains the current gold standard for triple-helical recognition of G–C base pairs in PNA.

However, more recent studies show that J can be further optimized. Chen and co-workers reported that substitution of oxygen-4 of J with sulfur improved the Hoogsteen binding properties of 4-thiopseudoisocytosine (L, Figure 6) [98]. UV thermal melting and gel electrophoresis studies showed that L formed more stable L•G–C triplets than J when binding to dsRNA, which was suggested to be a combined effect of improved van der Waals contacts, base stacking, hydrogen bonding, and reduced dehydration energy [98]. Replacement of three Js with Ls increased the binding affinity of a PNA 8-mer ≈4-fold [98]. In addition, the sulfur modification removed the undesired ability of J to form a Watson–Crick base pair with G in single-stranded nucleic acids. This is important for avoiding off-target binding to single-stranded RNA and DNA in biological systems. L appears to be a promising improvement of J as a neutral analogue of protonated C for Hoogsteen recognition of G–C base pairs.

An alternative strategy that increases the basicity of cytosine through chemical modifications was pioneered by Povsic and Dervan who showed that addition of a 5-methyl substituent increased the stability of $^{Me}C^+$•G–C triplet apparently through a subtle modulation of the pK_a and better π-stacking [99]. Several other research groups have further increased the pK_a value by removing electronegative substituents from C arriving at derivatives of 2-aminopyridine (M, Figure 6) as more basic nucleo-

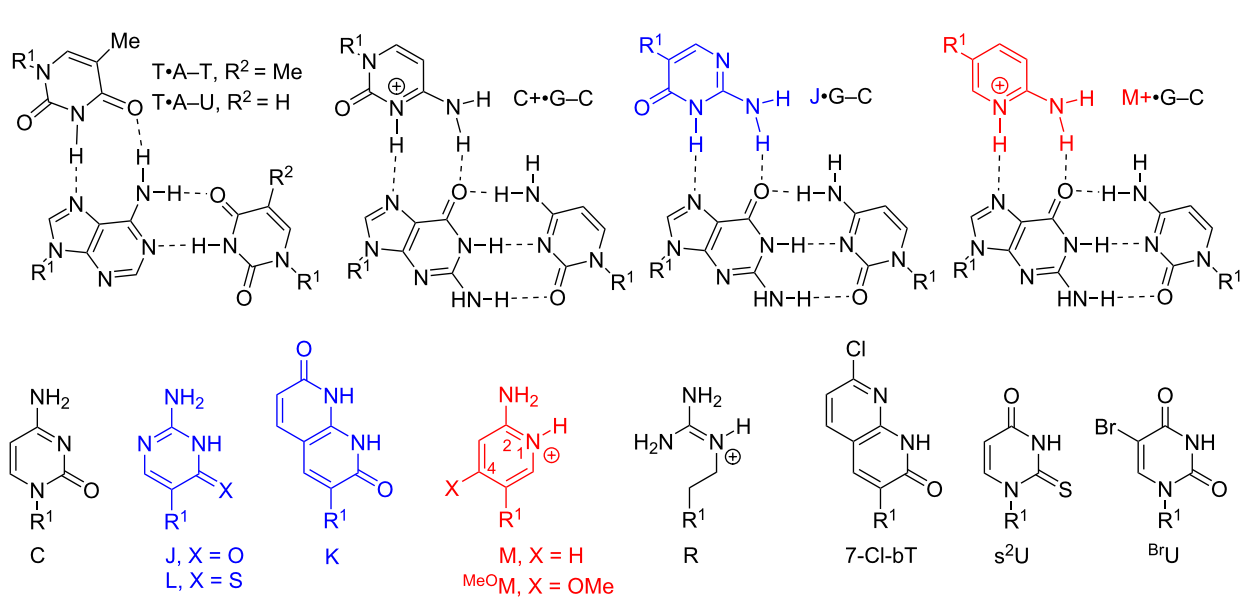


Figure 6: Structures of modified nucleobases in PNA to improve Hoogsteen hydrogen bonding to guanine and adenine. R^1 denotes DNA, RNA, or PNA backbones.

bases that improve binding of triplex-forming oligonucleotides at neutral pH [100-102]. Rozners and co-workers were the first to introduce M in triplex-forming PNAs targeting dsRNA [28]. Having a pK_a of ≈6.7, M is partially protonated at physiological pH 7.4, which facilitates fast binding and formation of strong triplex [28,30,31]. While all Hoogsteen triplets in Figure 6 are stabilized by two hydrogen bonds, because of the positive charge, M forms a significantly more stable M^+•G–C triplet compared to either J•G–C or T•A–U in dsRNA [28,30]. In a recent study, replacement of six Js with Ms increased the binding affinity of a PNA 9-mer ≈100-fold [31]. Preliminary results suggest that PNA–dsDNA triplexes follow similar trends. Similar to L, M does not form a Watson–Crick base pair with G or any other natural nucleobase, which is important for avoiding off-target effects of triple-helical recognition in biological systems. M is unique among cationic RNA binding compounds, perhaps, because the protonation event is coupled with the Hoogsteen hydrogen bond formation. As a result, the partially protonated M strengthens the triple helix without compromising the sequence specificity of recognition [28,30,31].

As discussed above, guanidine groups have been attractive modifications because of their potential to improve cellular uptake of PNA. Interestingly, simple guanidine (R, Figure 6) as a single nucleobase replacement appeared to form a strong and selective R•G–C triplet; however, two consecutive R modifications destabilized the PNA–dsRNA triplex, most likely due to reduced ability of R to π-stack [103]. As expected, fluorescence microscopy showed improved cellular uptake of the cationic guanidinium-modified PNAs [103].

PNA nucleobases for Hoogsteen recognition of adenine: Because the T•A–T triplets are reasonably stable under physiological conditions, development of novel nucleobases for Hoogsteen recognition of A has attracted less attention than the problem of C protonation discussed above. Similar to K, 7-Cl-bT, a substituted naphthyridine derivative (Figure 6), forms stronger Watson–Crick base pairs and Hoogsteen triplets with A, most likely because of enhanced stacking of the bicyclic π-system [104,105]. However, beyond the original studies, 7-Cl-bT has not been further explored for either duplex or triplex stabilization. Similar to L, substitution of thymine with 2-thiouracil (s^2U) or 5-halouracils (e.g., ^{Br}U, Figure 6) strengthens the Hoogsteen recognition of A. The stabilization provided by these nucleobases is most likely due to improved π-stacking, which may be sensitive to sequence context that needs to be further studied [106,107]. MacKay and co-workers designed an extended nucleobase based on isoorotic acid (Io4, Figure 7) to recognize the entire Hoogsteen face of the A–U base pair [108]. Io4 formed about a two-fold stronger triplet with the A–U base pair with good sequence selectivity. A PNA containing four consecutive Io4 nucleobases showed stronger binding to the complementary dsRNA than PNA containing four Ts suggesting that Io4 may be a promising alternative to T where stronger binding is desired [108].

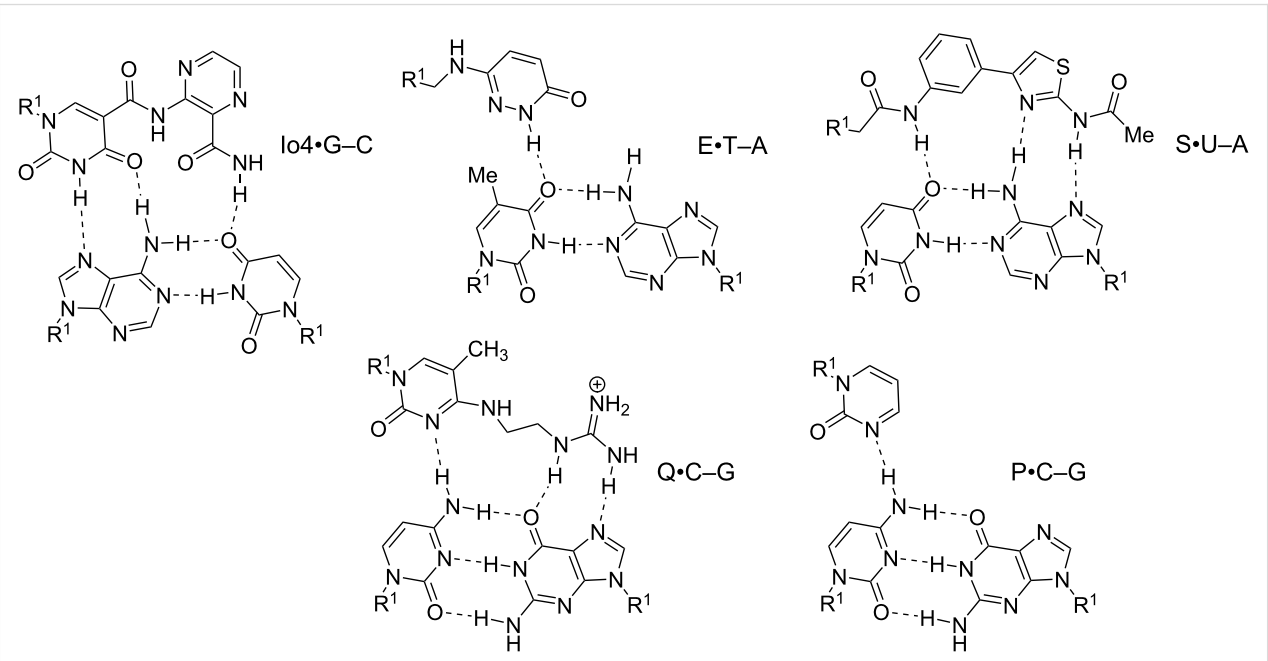


Figure 7: Proposed hydrogen bonding schemes for modified PNA nucleobases designed to recognize pyrimidines or the entire Hoogsteen face of the Watson–Crick base pairs. R^1 denotes DNA, RNA, or PNA backbones.

PNA nucleobases for Hoogsteen recognition of pyrimidines: Triple helix formation, especially using tailored oligonucleotide analogues as PNA, could be a general and sequence specific approach for molecular recognition of dsDNA and dsRNA. However, the triple helical recognition has a severe sequence limitation – the requirement of polypurine tracts in target nucleic acids. Natural triple helices allow only T•A–T (or U•A–U) and C^+•G–C triplets stabilized by two Hoogsteen hydrogen bonds (Figure 6) [94]. Analogous recognition of pyrimidines in hypothetical X•T–A or X•C–G triplets is complicated by two problems: 1) pyrimidines present only one hydrogen bond acceptor (C=O in T or U) or donor (-NH_2 in C) in the major groove, and 2) the six-membered pyrimidine ring extends further out in the major groove than the five-membered ring of purines causing a steric clash with the incoming third nucleobase. In other words, the Hoogsteen face of Watson–Crick base pairs in the major groove is not isomorphous providing more space and better hydrogen bonding options for purines than for pyrimidines. Despite significant efforts by nucleic acid chemists, a universal solution to triple helical pyrimidine recognition is still missing [94,109].

Nielsen and co-workers introduced 3-oxo-2,3-dihydropyridazine (E, Figure 7), a synthetic nucleobase designed to form a single hydrogen bond with U in PNA–DNA–PNA clamps [110]. Their design connected E to the PNA backbone with a linker two atoms longer than in standard PNA, which was suggested to circumvent the 5-methyl group of thymine and enable hydrogen bonding to the 4-oxo group [110]. More recent work [111] has questioned the originally proposed hydrogen bonding scheme shown in Figure 7. In fact, all of the hydrogen bonding schemes in Figure 7, while reasonable, are proposed. They are not necessarily confirmed by structural studies. E was later used for recognition of U in PNA–dsRNA triple helices [112]. Most recent study from Sugimoto, Rozners, and co-workers showed that triplex formation by E- and M-modified PNAs was able to inhibit maturation of pri-microRNA hairpins in SH-SY5Y cells [113].

An alternative approach to pyrimidine recognition has been to develop extended nucleobases that bind the entire Hoogsteen face of a Watson–Crick base pair and take advantage of the hydrogen bonding options on the purine base as well. An extended nucleobase S (Figure 7) originally developed for triplex-forming oligonucleotides [114,115], was introduced in PNAs targeting U interruptions in polypurine tracts of dsRNA triplexes [111]. However, in PNA, S showed limited sequence specificity binding strongly to either U–A or C–G base pairs [111]. The low mismatch discrimination suggests that S may have binding modes other than the hydrogen bonding depicted in Figure 7, for example, intercalation as has been previously observed for other similar unnatural nucleobases in DNA [116]. At the time of writing, E remains the most commonly used PNA nucleobase for recognition of U–A base pairs in dsDNA and RNA [117].

Several heterocyclic nucleobases in triplex-forming oligonucleotides have been designed to form a single hydrogen bond with the exocyclic -NH_2 of cytosine [94,109]. Rozners and co-workers [112] followed up on original work by Leumann [118] and showed that pyrimidin-2-one (P, Figure 7) could selectively recognize C–G, albeit with lower binding affinity than that of the standard Hoogsteen triplets. Despite the lower affinity, P-modified PNA formed a sequence specific triplex with a hairpin structure in the 5'-UTR of an mRNA, which inhibited ribosome assembly and suppressed mRNA translation in vitro and in cells [119]. This study was the first demonstration of the biological effect of binding of M- and P-modified PNAs to dsRNA in live cells. Recent work from our labs [120] systematically surveyed simple nitrogen heterocycles and found that the 3-pyridazinyl nucleobase formed significantly more stable triplets with C–G than other heterocycles, including P. Several groups have explored extended PNA nucleobases for recognition of C–G base pairs [121,122]. Chen and co-workers followed up on original work by Seidman [123] and showed that Q (Figure 7) in PNAs targeting dsRNA, recognized C–G base pairs with good selectivity. However, the stability of the Q•C–G triplet was reduced compared to T•A–U (≈8-fold) or L•G–C (≈24-fold) triplets [122]. Thus, an optimal solution for recognition of the C–G base pair in dsDNA and dsRNA remains elusive.

While the modified nucleobases reviewed above have given promising results, they typically lack either the binding affinity or selectivity of the natural triplets. This is especially true when the task is to recognize several pyrimidines, not just a single interruption of longer polypurine tract. Therefore, the search for new and better nucleobases for triple-helical recognition of any sequence of dsDNA or dsRNA remains an important goal and active area of research.

Nucleobases improving Watson–Crick recognition of PNA: We previously noted that PNA forms duplexes with complementary DNA and RNA that are more stable than duplexes involving only natural nucleic acids. Nevertheless, nucleobase modifications can further improve the remarkable binding properties of PNAs. One of the most promising nucleobases for improving Watson–Crick binding is G-clamp (Figure 8), the phenoxazine-derived tricyclic analogue of cytosine [124]. The improvements in affinity provided by the G-clamp are likely a combined effect of superior π-stacking of the rigid and planar aromatic system, electrostatic attraction of the positively

charged amine, and additional Hoogsteen hydrogen bonding [125]. Inserting just one G-clamp nucleobase into a PNA sequence increased the duplex melting temperature with complementary DNA or RNA by 13–20 °C while maintaining good mismatch discrimination [126].

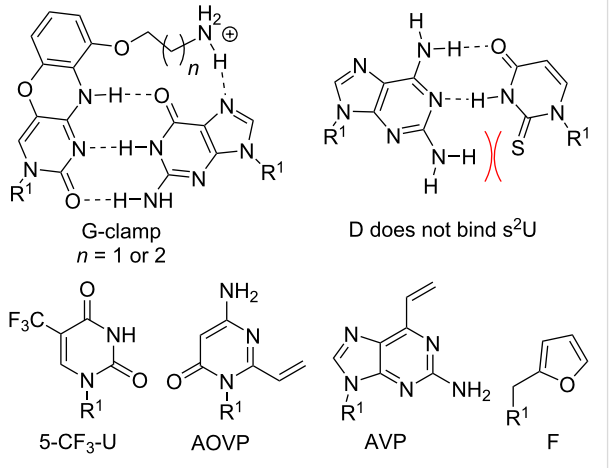


Figure 8: Modified nucleobases to modulate Watson–Crick base pairing and chemically reactive crosslinking PNA nucleobases. R^1 denotes DNA, RNA, or PNA backbones.

Ganesh and co-workers found that substitution of the 5-position in uracil with fluorine or trifluoromethyl improved PNA binding affinity for complementary DNA and RNA [127]. Moreover, fluorination increased the cellular uptake of PNAs [127]. Fluorinated uracil derivatives are also useful probes for studying different binding modes of PNA using ^{19}F NMR [128].

PNA nucleobases for double duplex invasion: Double duplex invasion (Figure 2F) critically depends on the ability of two PNAs to recognize each strand of dsDNA while not forming an unproductive PNA–PNA complex. Because the two DNA strands that are invaded are complementary, the two PNA strands have inherent complementarity as well. An elegant solution to this problem has been to use 2,6-diaminopurine (D) instead of adenosine and 2-thiouridine (s^2U) instead of uridine as modified nucleobases in PNAs designed for double duplex invasion [129,130]. D and s^2U form more stable Watson–Crick base pairs with T and A, respectively, than the natural A–T, but do not cross-bind in a D–s^2U pair because of a steric clash between the 2-amino group of D and 2-thiocarbonyl group of s^2U [129,130]. A recent report described an improved synthesis of s^2U and s^2T, which will help future applications of this currently somewhat underexplored technology [131].

Chemically reactive crosslinking PNA nucleobases: PNA has become a highly useful probe for detection of nucleic acids. Not surprisingly, chemists have developed reactive nucleobases for covalently crosslinking PNA and nucleic acid targets. 4-Amino-6-oxo-2-vinylpyrimidine (AOVP, Figure 8), a chemically reactive mimic of cytosine, exhibited selective crosslinking reactivity with thymine in DNA when incorporated at the terminal position of a PNA probe [132]. Interestingly, the activity of the crosslinking reaction was lower in RNA. Because AOVP functional groups do not match well any Watson–Crick base pairing scheme, AOVP lowered the stability of PNA duplexes with complementary DNA and RNA [132]. Similarly, vinyl-modified purine (AVP) effectively crosslinked with thymine in DNA and with uracil in RNA. The crosslinking resulted in inhibition of Dicer processing of microRNA precursors in vitro [133].

Furan (F, Figure 8) as a reactive nucleobase mimic was well accommodated in a duplex with DNA without decreasing its thermal stability [134]. Upon oxidation of the furan ring, F-modified PNAs reacted preferentially with cytosine and adenine and irreversibly crosslinked with ssDNA and dsDNA [134]. Covalent crosslinking of PNA with DNA or RNA upon hybridization is potentially highly useful for diagnostics and other applications as more stringent washing could be applied after hybridization with the complementary nucleic acid.

Janus-wedge PNA triple helix: McLaughlin and co-workers described a novel Janus-wedge triple helix (Figure 2G) where the wedge nucleobases (W_1 and W_2, Figure 9) of an incoming third PNA strand insert between two natural nucleobases hydrogen bonding with the Watson–Crick faces of the two DNA target strands from the major groove side [135,136]. This approach showed best results when invading DNA having consecutive C–T mismatches (C–W_1–T triplet, Figure 9). W_2 effectively bonded with the G–C base pair (G–W_2–C triplet), but recognition of the A–T base pair (A–W_1–T triplet) was significantly weaker and the Janus-wedge PNA was not able to invade a fully matched DNA duplex [136]. Bong and co-workers used melamine as a Janus-wedge nucleobase (K^M, Figure 9) to organize two identical strands of oligothymidine DNA tracts (or oligouridine RNA tracts) on a peptide template to form peptide−DNA(RNA) triplex structures [137]. This approach was applied to induce RNA–RNA kissing loop dimerization and RNA–protein binding [138].

Ly and co-workers developed Janus-wedge nucleobases that invade both dsDNA and dsRNA Watson–Crick base pairs from the minor groove side. At the time of writing, three Janus nucleobases, E, F, and I (Figure 9) have been reported for recognition of C–G, G–C, and A–A base pairs, respectively [139,140]. While still in relatively early stages of development, the Janus-wedge triplex has already shown intriguing potential

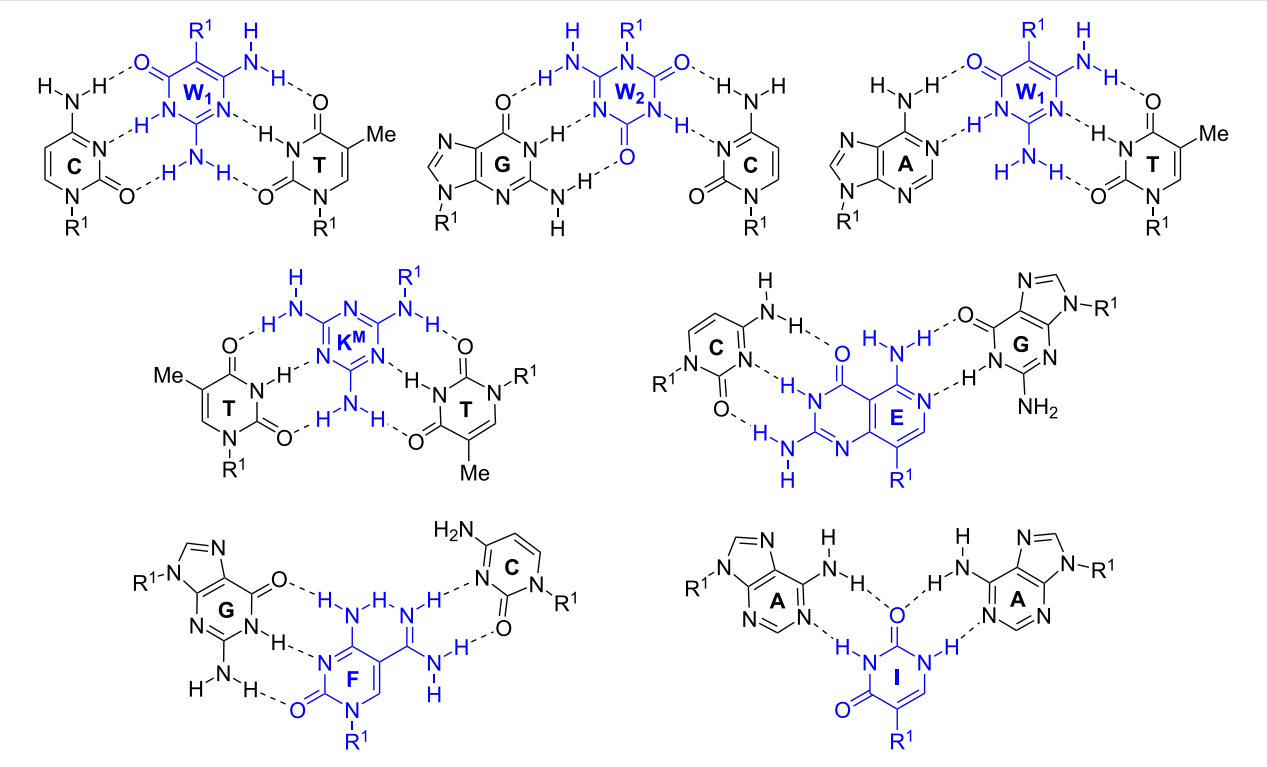


Figure 9: Examples of triplets formed by Janus-wedge PNA nucleobases (blue). R^1 denotes DNA, RNA, or PNA backbones.

as a diagnostic or therapeutic approach for Huntington's or related genetic diseases [139].

Fluorescent nucleobases in PNA: Because PNA has become a key component of many assays and diagnostics, development of fluorescent nucleobases as labels for PNA has attracted considerable attention. 2-Aminopurine (Figure 10), a fluorescent structural isomer of adenine [141], was one of the first fluorescent nucleobases used in PNA [142]. Melting of a duplex formed by 2-aminopurine-modified PNA and complementary DNA increased the fluorescence signal, which had likely been quenched by adjacent nucleobases in the duplex [142]. Interestingly, quenching was also observed in a single stranded PNA alone, which diminished the applicability of 2-aminopurine in PNA probes. Hudson and co-workers developed several fluorescent PNA nucleobases derived from phenylpyrrolocytosine [143-145]. One of the most promising analogues, mmguaPhpC (Figure 10), formed a stronger base pair with G than the native C–G pair which was followed by a 30–70% decrease of emission intensity (dependent on the sequence context) upon hybridization with complementary DNA and RNA [145]. Another analogue, 5,6-BenzopC (Figure 10) had high quantum yield and superior base pairing properties, but its fluorescence was completely quenched upon hybridization with DNA and RNA [146]. Inspired by these findings, Cheruiyot and Rozners attempted to design fluorescent analogues of 2-aminopyridine; PhEthM (Figure 10) gave the best binding and fluorescence properties, but was strongly quenched upon formation of PNA–dsRNA triplex [147]. In general, quenching of PNA fluo-

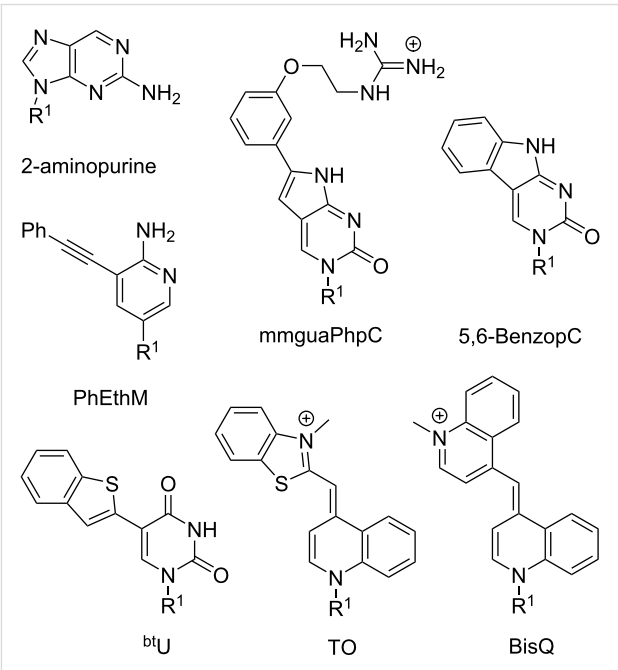


Figure 10: Examples of fluorescent PNA nucleobases. R^1 denotes DNA, RNA, or PNA backbones.

rescence upon binding to target DNA or RNA is less useful than the increase in signal intensity.

Chen and co-workers found that 5-benzothiopheneuracil (^{bt}U, Figure 10) modified PNAs increased the fluorescence upon binding to dsRNA, acting as light-up triplex-forming PNA probes [148]. This was the first report of a modified natural nucleobase that did not quench the fluorescence upon hybridization [148].

Köhler and Seitz introduced thiazole orange (TO, Figure 10), an intercalator dye originally designed for DNA [149], as a forced intercalation (FIT) probe in PNA. Because of rotation around the methine bond connecting thiazole and quinoline, TO fluorescence is almost completely quenched in ssPNA, but increases significantly upon hybridization to the complementary DNA [150]. The intercalation of TO in PNA–DNA duplex restricts rotation around the methine bond enforcing planarity of the two TO's aromatic system, which leads to fluorescence increase [151,152]. TO can be considered as a "universal base" due to its ability to pair equally well with each of the four natural DNA nucleobases [150]. Later, Nishizawa and co-workers used TO-modified triplex-forming PNAs as fluorescent probes sensitive to adjacent mismatched base pairs in dsRNA [153,154]. Replacement of thiazole in TO with another quinoline gives bis-quinoline (BisQ, Figure 10), a red-shifted PNA nucleobase analogous to TO [155]. Although binding of BisQ with all four natural DNA nucleobases has not been explored in detail, BisQ-modified FIT PNAs showed promising fluorescent enhancements and an ability to detect mismatches in live cells [155]. Overall, the TO- and BisQ-modified FIT PNAs are currently among the most promising fluorescent PNA probes.

While promising, the studies discussed in this section leave plenty of room for designing better fluorophores, especially, red-shifted dyes with stronger fluorescence enhancement. Future design of novel PNA nucleobases that enhance the fluorescence signal while selectively hybridizing to natural nucleobases will be highly beneficial for in vitro and in vivo probes and diagnostics.

Covalent PNA conjugates for delivery in cells and animal models

Delivery and uptake of oligonucleotides to target tissues and cells is one of the greatest challenges for development of nucleic acid detection probes and therapeutics [14]. This problem is especially critical for in vivo applications of PNA because unmodified PNA, despite being charge neutral, does not readily cross cellular membranes [16,156-158]. Not surprisingly, the first demonstration of PNA-mediated suppression of gene expression by Babiss and co-workers used nuclear micro injection [38]. Another common method for PNA delivery has been electroporation [119,159,160]. Looking forward, conjugation of PNA with various delivery enhancing compounds, most notably cell-penetrating peptides (CPP) that deliver the conjugates mainly through endocytosis (Figure 11) has become one of the most promising approaches to improving cellular uptake of PNA [161,162]. However, the uptake of most PNA–CPP conjugates is limited by endosomal entrapment. While the uptake can be improved either by increasing the concentration of PNA–CPP conjugates or by using endosomolytic compounds (for example, chloroquine or calcium ions) this leads to toxicity that is not viable for in vivo applications [163]. Inefficient and incomplete release from endosomes remains an unsolved problem for PNA–CPP conjugates [164]. In this section we review the initial approaches and some of the most promising and foundational studies undertaken in addressing the cellular delivery issue using the covalent conjugation of PNA to delivery enhancing compounds.

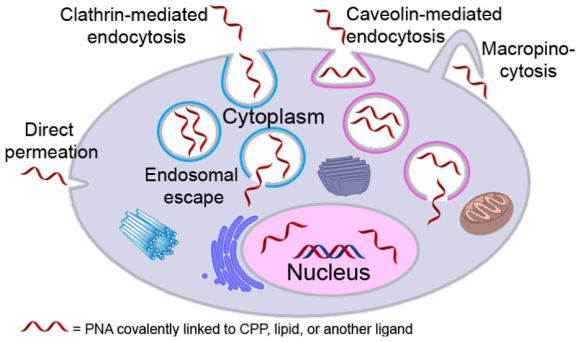


Figure 11: Endosomal entrapment and escape pathways of PNA and PNA conjugates.

Cell-penetrating peptides derived from natural proteins: The initial success of PNA delivery involved PNA conjugates taken up by receptor-mediated endocytosis. Pardridge and co-workers successfully demonstrated in vivo delivery and blood-brain barrier crossing of PNAs by intravenous administration of PNA conjugated to OX26 murine monoclonal antibody to the rat transferrin receptor [165]. The limitation of this strategy was complexity of the construct and lack of clear evidence for the cellular uptake. The first report of using the PNA–peptide conjugate approach involved the conjugation of PNA to (D)-insulin-like growth factor 1 peptide (IGF1) that enabled the delivery to cells expressing the IGF1 receptor [166]. Later developments adopted CPPs derived from natural proteins (Figure 12A), such as penetratin (16-amino acid peptide from the third helix of the Antennapedia homeodomain) [167], Tat (14-amino acid peptide from HIV-1 TAT protein) [168],

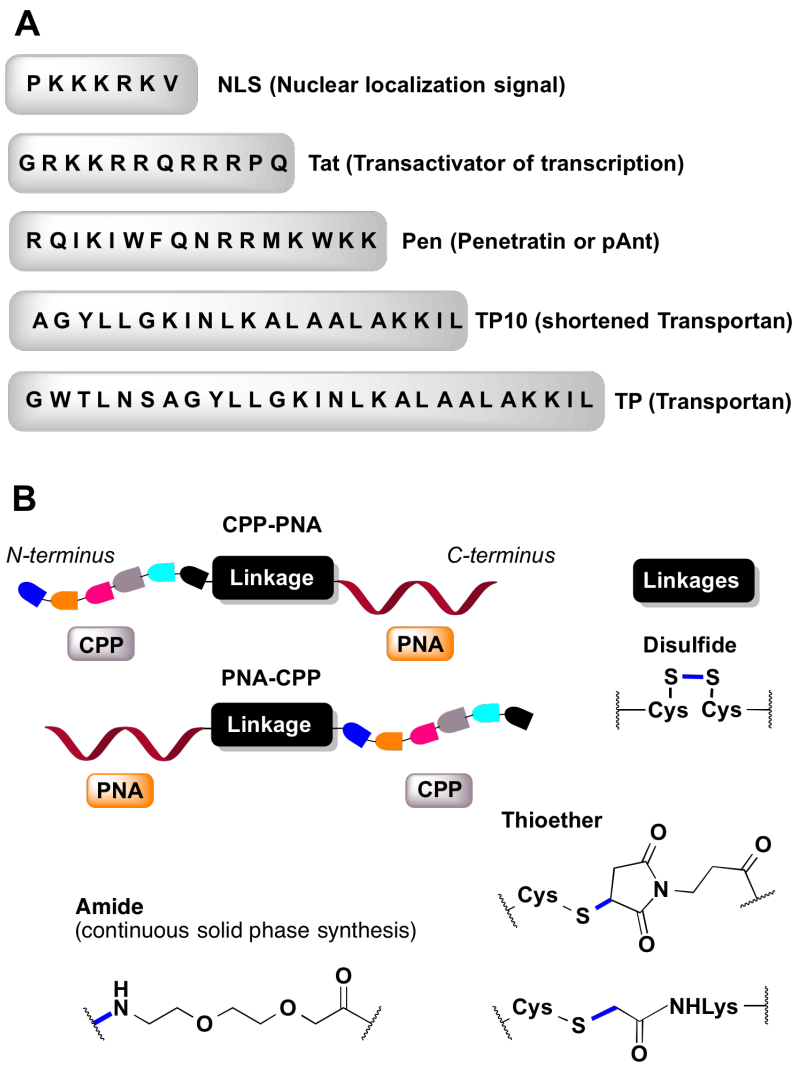


Figure 12: (A) representative cell-penetrating peptides (CPPs), (B) conjugation designs and linker chemistries.

and transportan (chimeric 27-amino acid peptide derived from galanin and mastoparan) [169].

Corey and co-workers were the first to demonstrate that conjugation of an 11-mer PNA to penetratin peptide enabled uptake of the conjugate in DU145 cancer cells as analyzed by fluorescence-activated cell sorting (FACS). However, the conjugate did not inhibit the targeted human telomerase in cells [170]. Langel and co-workers conjugated antisense PNA targeting mRNA of galanin receptor type 1 (GalR1) through disulfide linkages to transportan and penetratin peptides. The PNA–peptide conjugates were effectively internalized in human Bowes melanoma cells and in vivo in rats [171]. Transportan peptide localized the PNA in membranous structures of cells, while the penetratin conjugate preferred nuclear localization. The conjugates inhibited ^{125}I-galanin binding in Bowes cells with 91% efficiency of PNA–penetratin (3 μM) and 83% of PNA–transportan (1.5 μM), which compared favorably with 5% efficiency of antisense DNA (10 μM) and 37% of phosphorothioate-modified antisense DNA (12 μM). In rats, intrathecally administered PNA–penetratin conjugate (3 × 10 μL of 150 μM) caused a 40% decrease in ^{125}I-galanin binding in spinal cord sections compared to rats treated with the saline control. The PNA–peptide conjugates showed no toxicity in these studies [171].

Boffa and co-workers conjugated antigene PNA to a nuclear localization signal (NLS) peptide (PKKKRKV, Figure 12A), and showed that the PNA–NLS conjugates localized predominantly in the nucleus rather than in the cytoplasm of Burkitt's lymphoma cell lines (BRG, BJAB, HBL2) [172]. The opposite trend was observed for unmodified PNA or PNA conjugated to a scrambled-NLS peptide (KKVKPKR). UV melting studies showed that the conjugation of basic NLS peptide to PNA did

not influence the binding ability for the complementary DNA. In BRG cells at 10 μM concentration, PNA–NLS targeting *c-myc* gene reduced its expression by 75% compared to controls having scrambled PNA or peptide sequence, or unmodified PNA [172].

Peschke and co-workers conjugated a dual peptide construct built of penetratin (for cytosolic delivery) and NLS (for nuclear delivery) at the N-terminus of PNA and demonstrated efficient delivery and distribution of the conjugate (100 nM) in the nucleus of DU 145 or R3327-AT1 prostate tumor cells [173]. Importantly, the efficient delivery of PNA to the nucleus was achieved only when the penetratin and NLS peptides were connected by a cleavable disulfide linkage (Figure 12B). PNA conjugates with penetratin only or dual peptide with a non-cleavable linker localized mostly in the cytosol with very little nuclear delivery. Confocal imaging studies of a fluorescently labeled dual peptide–PNA conjugate revealed initial cytosolic delivery, followed by cleavage of the disulfide linkage in cytosol and nuclear uptake of NLS–PNA. The ability to achieve delivery and diffused nuclear localization of PNA using only 100 nM concentration of the dual peptide conjugate was a significant achievement; however, this study did not demonstrate antisense or other biological effects of the PNA–penetratin conjugate [173].

Nielsen and co-workers compared the cellular uptake of unmodified PNA with α-backbone-modified PNA derived from lysine (T_{Lys}-PNA, Figure 5), CPP (Tat or Penetratin, Figure 12A) alone, and PNA–CPP conjugates in HeLa (cervical carcinoma), SK-BR-3 (breast carcinoma) and IMR-90 (fetal lung fibroblast) monolayer cells, as well as in H9 (lymphoid) and U937 (monocytic) suspension cells [174]. At 2.0 μM concentration, T_{Lys}–PNA and PNA–CPP were readily taken up by the three monolayer cell lines but were confined exclusively to the cytosolic vesicular compartments. T_{Lys}–PNA and PNA–CPP showed very weak membrane staining in H9 cells and no uptake in U937 cells. The vesicular uptake was time, temperature and concentration dependent indicating an endocytic pathway (Figure 11). PNA alone and CPPs alone were not taken up in cells under the experimental conditions used in this study. It was also noted that depending on the cell type, the PNA–CPP conjugates were cytotoxic above 5–10 μM [174].

Gait and co-workers studied the effect of different CPPs and linkers (Figure 13) on activity of PNA conjugates targeting the apical stem-loop of TAR at the 5′-end of HIV-1 RNA [175]. In this study, the inhibition of HIV-1 Tat-mediated *trans*-activation in HeLa cells was monitored using an integrated double-luciferase reporter system [175]. PNAs conjugated through a stable amide linker to various CPPs (Figure 12B) showed no inhibitory activity at 2.5 μM while cell viability remained >95%. Co-administration with 100 μM chloroquine showed significant to weak inhibitory activity for Tat–PNA, TP–PNA, TP10–PNA, NLS–PNA–Tat, PNA–TP10, and Tat–PNA–NLS (Figure 12). However, no inhibition activity was recovered for NLS–PNA, PNA–NLS, and K_8–PNA–K. Some conjugates having cleavable linkers, such as, Tat–S–S–PNA, Pen–S–S–PNA, and R_9F_2–S–S–PNA showed no inhibitory activity at 2.5 μM either with or without 100 μM of chloroquine. Three conjugates having cleavable linkers, R_6-penetratin–S–S–PNA, TP–S–S–PNA and TP(int)–S–S–PNA showed significant levels of inhibitory activity at 2.5 μM, which was further increased in the presence of 100 μM chloroquine, while maintaining sequence-specificity. Overall, the poor activity of most of the CPP–PNA conjugates in the nucleus was attributed to the poor escape from endosomes or other membrane-bound compartments [175].

Cao and co-workers conjugated a PNA targeting the direct repeats of hepatitis B virus (HBV) to Tat peptide using 1,4-addition of C-terminal cysteine thiol on Tat to N-terminal maleimide on PNA [176]. The resulting Tat-PNA conjugate showed excellent in vitro and in vivo antiviral properties. In HepG2.2.15 cells, the Tat–PNA conjugate blocked expression of HBV DNA, RNA and proteins (HBeAg, HBsAg, HBV core, x protein, reverse transcriptase) indicating multiple modes of action, in contrast to the single mode of reverse transcriptase inhibition by the clinically approved drug lamivudine. The Tat–PNA conjugate was not toxic at 100 μM in multiple cell lines from hepatocytes and erythrocytes. Intravenous injection of the Tat–PNA conjugate at 50 mg/kg in mice did not cause acute toxicity or immune response as judged by levels of IgG and IgM measured by ELISA. The Tat–PNA conjugate suppressed HBV DNA concentration in serum of mice infected with HBV as measured by quantitative real time PCR (qRT-PCR) to 1.4×10^4 copies/mL, which compared favorably with 1.2×10^4 copies/mL in lamivudine treated mice and was lower than 6.9×10^4 copies/mL in untreated mice. In mouse liver tissues, HBV core-protein-positive hepatocytes were reduced to 1.7% compared to 4.5% in untreated mice. In addition, very low levels of viral antigens (HBeAg and HBsAg) were observed in the blood of mice treated with the Tat–PNA conjugate [176]. These results suggested that targeting of direct repeats of HBV using PNA–CPP conjugates might be explored as a potential therapeutic strategy against HBV.

Engelman and co-workers discovered that a 36-residue polypeptide derived from transmembrane helix C of bacteriorhodopsin spontaneously inserts into the lipid bilayer under slightly acidic conditions [177]. Follow-up studies developed a pH-low insertion peptide (pHLIP) that translocates imperme-

able drug molecules specifically across the membranes of cells with low surface pH ≈ 6 (Figure 13) [178-180]. Peptides of the pHLIP family typically contain a transmembrane peptide sequence, which is essential for interactions with the lipid bilayer of cells, and short flanking sequences at the C- and N-terminus that promote membrane insertion and peptide solubility [178,180].

Slack and co-workers conjugated a 23-mer PNA targeting miRNA-155 to the C-terminus of pHLIP through a cleavable disulfide linkage. In A549 and DLBCL tumor cell lines, enhanced delivery of pHLIP–PNA was observed at the slightly acidic extracellular pH of tumor cells [181]. Intravenous administration of the pHLIP-PNA conjugate (2 mg/kg) in two mouse models, *mir-155*^{LSLtTA} subcutaneous flank model and *mir-155*^{LSLtTA} model of lymphoma was studied [181]. The systemically administered pHLIP-PNA accumulated in the enlarged lymph nodes of transgenic *mir-155*^{LSLtTA} mice. Significant reduction in the tumor growth was achieved in the flank tumor model. The survival time of 11 days for pHLIP–PNA treated mice compared favorably with 7 days for mice treated with commercial locked nucleic acid (LNA) anti-miR at 17–40-fold higher concentrations than pHLIP–PNA. The pHLIP–PNA conjugate not only delayed the tumor growth but also suppressed the metastatic spread of neoplastic lymphocytes to other organs with no clinical signs of distress, toxicity or renal damage [181].

Glazer and co-workers conjugated pHLIP via a disulfide linkage to antisense γ-miniPEG-modified PNA (Figure 5) targeting nonenzymatic-NHEJ factor Ku80 mRNA [182]. In human lung adenocarcinoma (A549) cells at pH 6.2, this pHLIP–PNA conjugate showed ≈45% reduction of Ku80; no activity observed at pH 7.8. Systemic delivery of the pHLIP–PNA conjugate (5 mg/kg) in mice bearing DLD1-BRCA2KO human colon cancer xenograft reduced the Ku80 expression by ≈40%. Similar partial suppression was observed in EMT6 tumors as well. No significant toxicity or immune response was noted in mice treated with the pHLIP–PNA conjugate and, unlike with many anticancer therapeutics, no bone marrow toxicity was observed [182].

Pentelute and co-workers achieved efficient cytosolic delivery of PNA using the two nontoxic components of the anthrax toxin, the protective antigen (PA) and the N-terminal domain of lethal factor (LF_N) [183]. The antisense PNA was conjugated to the C-terminus of LF_N through sortase-mediated ligation. The advantage of LF_N/PA mediated delivery was demonstrated by the 100- to 1000-fold higher antisense activity at nanomolar concentrations (250 nM LF_N–PNA and 50 nM PA protein) in cancer cell lines compared to PNA alone or Tat–PNA conjugates (no activity up to 5 μM). The robustness of LF_N/PA delivery system was demonstrated by delivering PNAs across a panel of nine cancer cell lines from breast and blood lineages. The PNA–LF_N conjugate (100 nM) in the presence of PA protein (50 nM) caused a significant decrease in the viability of BT549 and HCC1954 breast cancer cells (50%) and Toledo and HUT 78 blood cancer cells (80%). Neither the length nor the sequence of PNA affected the translocation efficiency using the LF_N/PA delivery system; however, neutralizing antibodies produced by the immune system remained a critical challenge for this delivery system [183].

Synthetic cell-penetrating peptides: Kole and co-workers compared PNAs conjugated to one, two, and four lysines (PNA–K, PNA–K_2, and PNA–K_4) with negatively charged

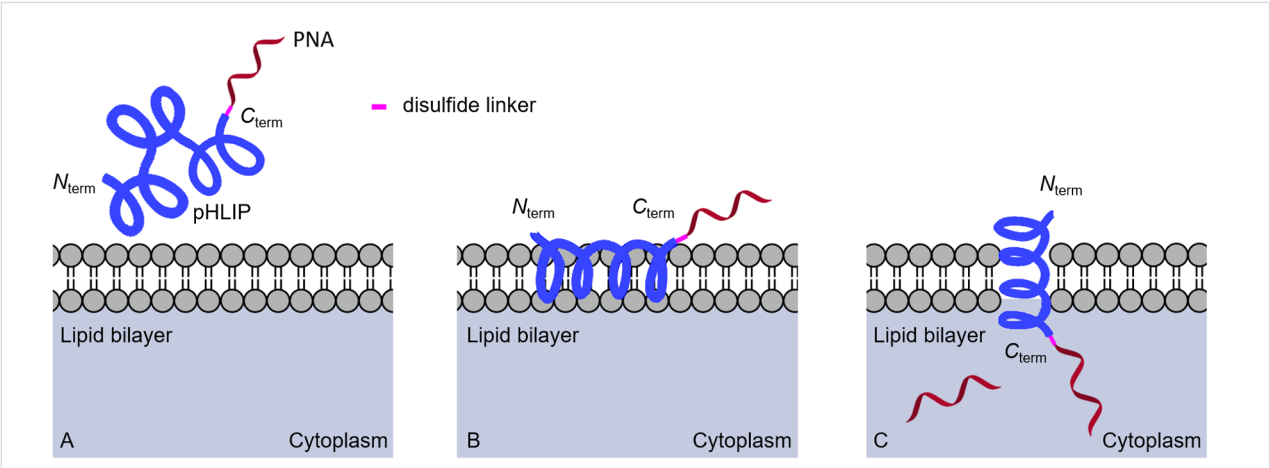


Figure 13: Proposed delivery mode by pHLIP-PNA conjugates (A) the transmembrane section of pHLIP interacting with lipid bilayer, (B) low surface pH leads to partial protonation of negative residues triggering interfacial helix formation and deeper partitioning into lipid bilayer, and (C) the transmembrane helix formation and release of PNA into cytosol by disulfide cleavage.

2′-*O*-alkyl oligonucleotide derivatives and neutral morpholino phosphorodiamidates (PMOs) in HeLa cells [184]. Passive uptake studies by FACS showed that PNA–K, PNA–K_2, PNA–K_4, and PMOs crossed the cellular membrane and gained access to the nucleus more readily than the anionic oligonucleotide analogues. In a splicing correction assay, increasing the number of lysines in the series PNA–K, PNA–K_2, and PNA–K_4 correlated with increased splicing modulation activity with EC_{50} of 4.7, 3.3, and 2.1 μM, respectively. The uptake mechanism was similar to that of PNA–penetratin conjugates. MTT assay showed no toxicity associated with PNA–K_4 even at 10 μM. In the clinically relevant β-thalassemia model, in the absence of transfection reagents, the correct splicing of IVS2-654 human β-globin pre-mRNA was four-fold higher with PNA–K_4 compared to PMO as measured by qRT-PCR [184].

Kole and co-workers also compared antisense activity of PNA–K_4, PMO and 2′-O-methoxyethyl phosphorothioate (2′-O-MOE-PS) oligonucleotides in EGFP-654 transgenic mice [185]. In this model, antisense activity restores correct splicing and expression of enhanced green fluorescence protein (EGFP) providing an easy readout of in vivo activity. Systemically injected 2′-O-MOE-PS and PNA–K_4 oligomers showed sequence-specific antisense activity in cardiac muscle, cortex of kidney, liver hepatocytes, lung and small intestine, while PMOs had weak or moderate activity in all these tissues and PNA–K was completely inactive. PNA–K_4 was the most effective antisense in all the tissues except small intestine where 2′-O-MOE-PS was more effective [185]. No antisense activity was observed in brain, skin and stomach with any of the oligomers.

Follow up studies by Corey [186,187] and Gait [164,175,188] and co-workers demonstrated that PNAs conjugated to short oligolysine peptides (four to eight residues) were efficiently taken up in cancer cell lines. Later studies demonstrated delivery and antisense activity of PNA–K_8 and K–PNA–K_3 conjugates in mice [189,190]. The cellular uptake of these simple conjugates was further optimized by addition of a terminal thiol group (cysteine in C–K–PNA–K_3) [191].

Corey and co-workers compared PNA–$(AAKK)_4$, PNA–NLS, and unmodified PNA delivered by complementary DNA/lipid co-transfectant [192]. They found that PNA–$(AAKK)_4$ and PNA–NLS were taken up in cultured cells but required higher PNA concentration to achieve the same uptake as that of DNA/lipid-mediated PNA delivery. In the absence of DNA/lipid co-transfectant, unmodified PNA and NLS–PNA did not inhibit expression of the human caveolin 1 (hCav-1) gene, while PNA–$(AAKK)_4$ reduced the expression of hCav-1 with IC_{50} 2 μM.

Wright and co-workers enhanced the antisense activity of the PNA–K_8 conjugate in the presence of PA protein (the protective antigen from anthrax) in CHO and HeLa cells [193]. Interestingly, reducing the lysine tail at the C-terminus to four in PNA–K_4 reduced the antisense activity ≈2-fold. Reducing the lysine tail further from four to two residues completely eliminated the antisense activity, highlighting the importance of lysine conjugation at the C-terminus of PNA. Administration of PNA–K_8 (300 nM) and PA protein (2 × 300 ng/mL) corrected the β-globin splice defect in cultured erythroid precursor cells from a patient with β-thalassemia, while no correction was observed with PNA–K_8 alone, highlighting the role of PA protein in delivering the PNA into cells [193].

Nielsen and co-workers demonstrated the antibacterial properties of PNAs by targeting 23S rRNA using unmodified bis-PNA, which inhibited the growth of the AS19 strain of *E. coli* that had a compromised and permeable cell membrane [194]. However, no growth inhibition was observed in case of the membrane intact K12 strain of *E. coli* [194]. In a later study by Good and Nielsen, conjugation of an antisense PNA targeting the *lacZ* gene in *E. coli* to a synthetic antibacterial peptide $(KFF)_3K$ [195] composed of cationic lysine and hydrophobic phenylalanine, inhibited growth of *E. coli* K12, with a minimal inhibitory concentration of 3.0 μM, while free peptide and unmodified PNAs showed no activity [196]. A $(KFF)_3K$–bis-PNA conjugate targeting mRNA of acyl carrier protein (*acp*P) at 2.0 μM concentration reduced the colony forming units (CFU) from 10^5 per mL to zero in three hours. Most importantly, the $(KFF)_3K$–bis-PNA conjugate at 2.0 μM fully cured the *E. coli* infection in *E. coli* K12 infected HeLa cells without harming the host HeLa cells [196].

Gait and co-workers developed a series of CPPs called PNA internalization peptides (Pip, Figure 14) by combining and optimizing the amino acid sequences of $(RXR)_4$, previously developed for delivery of charge-neutral PMOs [197], and penetratin CPPs [198]. The uptake of Pip–PNA conjugates followed the pathway of clathrin-dependent endocytosis, as previously established for Tat–PNA and $(RXR)_4$–PMO conjugates [199]. In HeLa pLuc705 cells, the Pip1–PNA conjugate showed higher splice correction activity (EC_{50} = 0.5 μM) than R_6Pen–PNA (EC_{50} = 1.0 μM) or $(RXR)_4$–PNA (EC_{50} = 3–4 μM) conjugates, but was fully cleaved within 1 hour in 20% mouse serum. Pip1 was further optimized into two serum-stabilized peptides, Pip2a and Pip2b (both differ by a single amino acid at position 11, underlined in Figure 14). In cultured *mdx* mouse myotubes, Pip2a–PNA and Pip2b–PNA conjugates targeting the exon 23 mutation in the dystrophin gene induced significant exon skipping at 1 and 2 μM, while maintaining the cell viability above 80% at concentrations up to 5 μM. The Pip1–PNA and

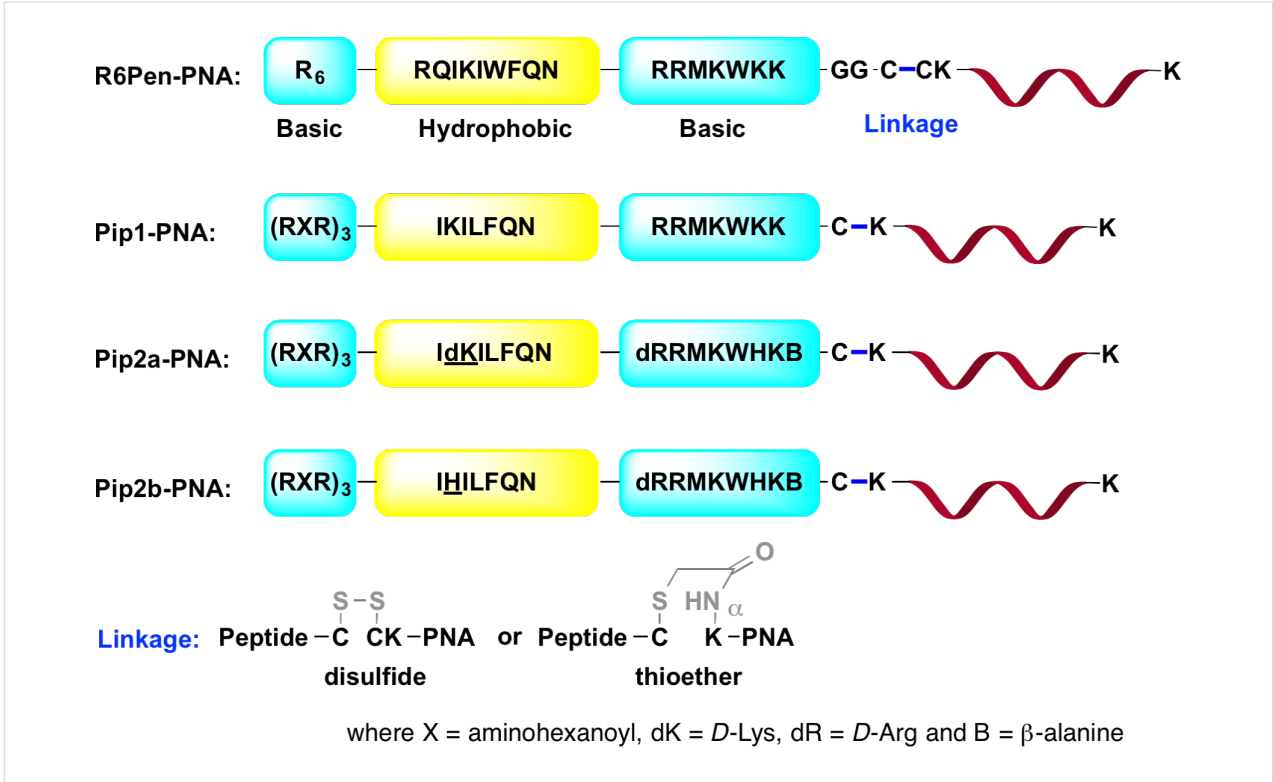


Figure 14: Structures of modified penetratin CPP conjugates with PNA linked through either disulfide (for study in HeLa pLuc705 cells) or thioether bonds (for study in cultured *mdx* mouse myotubes or mouse model).

$(RXR)_4$–PNA conjugates induced a small amount of exon skipping at 2 μM. In a mouse model of Duchenne muscular dystrophy (DMD), a single dose of 5 μg of Pip2a–PNA and Pip2b–PNA conjugates showed a significant increase in the dystrophin-positive myofibers [198].

Gambari and co-workers conjugated PNA with the well-established octaarginine CPP [200] and used R_8–PNA in Glioma cells to inhibit microRNA-221 (miRNA-221), which down regulates the expression of $p27^{Kip1}$ among several other genes [201]. Surface plasmon resonance confirmed that conjugation with the highly cationic R_8 peptide did not compromise sequence specificity of the R_8–PNA conjugate. FACS and confocal microscopy showed high levels of uptake of the R_8–PNA conjugate at 2.0 μM compared to unmodified PNA in U251, U373, and T98G Glioma cells. Strong miRNA-221 inhibitory effects were observed at 2 μM with the R_8–PNA conjugate while no inhibitory effects were observed with an unmodified PNA or R_8–PNA conjugate having mutated PNA. Moreover, the R_8–PNA conjugate did not inhibit the closely related miRNA-210 and -222, members of the same family as miRNA-221 [201].

Searching for a general membrane transporter for therapeutic agents, Pei and co-workers discovered that cyclic peptides were ≈20-fold more efficient for cytosolic delivery in HeLa cells compared to common CPPs, such as, Tat and R_9 [202]. Yavin and co-workers adopted this strategy and synthesized a PNA conjugate with a cyclic peptide C_9–PNA (Figure 15) [203]. After incubation at 500 nM for 3 h, C_9–PNA showed significant uptake in U87MG cells as judged by live cell fluorescence microscopy and FACS analysis, compared to less efficient uptake of K_4–PNA under the same conditions. In U87MG cells, which are difficult to transfect, at 500 nM concentration C_9–PNA and K_4–PNA reduced the miRNA-155 levels by ≈80 and 65%, respectively [203].

Lipid-based delivery of PNA: Murphy and co-workers conjugated a lipophilic phosphonium cation (TPP, Figure 16A) through a thioether linkage to a PNA targeting a point mutation in mitochondrial DNA [204]. The TPP–PNA conjugates (1.0 μM) were efficiently taken up in mitochondria of cultured human cells, myoblasts and fibroblasts, driven by the inner membrane potential across the lipid bilayer of mitochondria as evidenced by microscopic images. Mitochondrial localization of the TPP–PNA conjugate was noticed after 4 h but, surprisingly, did not show inhibition of target mitochondrial DNA replication [204]. Patino and co-workers conjugated a PNA targeting the TAR region of HIV RNA to TPP cation through a combination of carbamate and disulfide linkages (Figure 16B) [205].

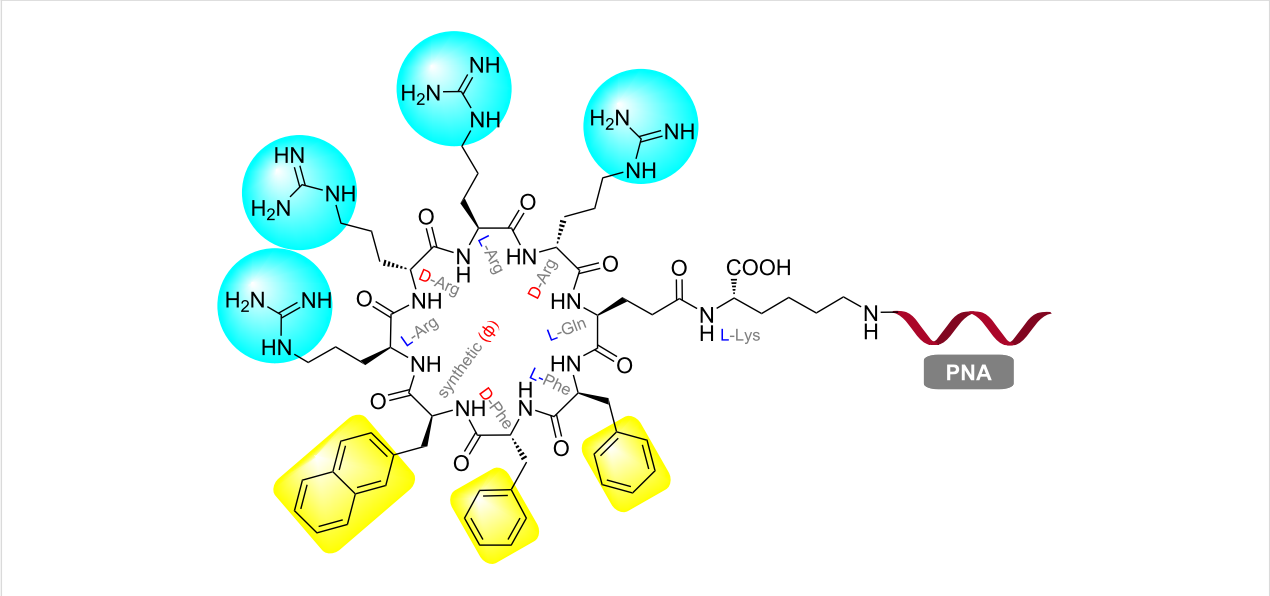


Figure 15: Chemical structure of C_9–PNA, a stable amphipathic (cyclic-peptide)–PNA conjugate.

The linker was stable in media containing 10% fetal calf serum for 48 h but was easily cleaved by glutathione treatment. FACS analysis showed 43% uptake of fluorescently labeled TPP–PNA conjugates in CEM cells in 6 h. The TPP–PNA conjugate inhibited replication of pseudotyped HIV-1 virions in CEM cells with IC_{50} 1.0 μM, while unmodified PNA was inactive. The TPP–PNA conjugate was not toxic at 2 μM [205].

Taylor and co-workers evaluated the splice correcting activity of PNA–R_9 with additional conjugation of lipids and phospholipids at the N-terminus, such as, L–PNA–R_9, P–PNA–R_9, LP–PNA–R_9, and LSS–PNA–R_9 (Figure 16C) [206]. In HeLa pLuc705 cells, L–PNA–R_9, LP–PNA–R_9, and LSS–PNA–R_9 showed similar bioactivity in the 1–3 μM range while PNA–R_9 and P–PNA–R_9 showed very little activity. The activity increased in the presence of 100 μM chloroquine suggesting that endosomal entrapment was limiting the efficiency [206]. A disadvantage of these lipid constructs was significantly higher toxicity compared to PNA and PNA–R_9. The LC_{50} values for LSS–PNA–R_9, L–PNA–R_9, and LP–PNA–R_9 were 3 (most toxic), 6, and 11 μM [206].

Nielsen and co-workers conjugated cholesterol or cholic acid at the N-terminus of PNA (Figure 17) targeting a cryptic splice site in pre-mRNA in HeLa pLuc 705 cell line [207]. The conjugates were inactive in the splice correction assay when administered alone in up to 1 μM concentration. In contrast, both cholesterol and cholic acid PNA conjugates exhibited nanomolar antisense activity (EC_{50} = 25 nM, as measured by qRT-PCR) when delivered in the presence of lipofectamine2000, which was several-fold higher than the activity of PNA delivered by the DNA/lipid co-transfectant strategy [207].

PNA delivery using receptor-specific ligands: Corey and co-workers conjugated eight lactose moieties at the N-terminus of PNA targeting human telomerase and demonstrated cell-specific uptake of the Lac_8–PNA conjugate in HepG2 cells that expresses surface bound asialoglycoprotein receptor (ASGPR). The addition of eight lactose moieties did not interfere with PNA's binding to the target. In HepG2 cells, the Lac_8–PNA conjugate linked through a cleavable disulfide bond was more active in inhibiting cellular telomerase (IC_{50} = 6 μM) than the conjugate linked by a stable amide bond (IC_{50} = 20 μM) [208]. However, the activity was still 50-fold lower compared to PNA delivered by the DNA/lipid co-transfectant strategy [209]. The Lac_8–PNA conjugate having mismatched PNA or PNA conjugated to eight maltose moieties showed no activity at 20 μM [208].

Biessen and co-workers conjugated an antisense PNA targeting the human microsomal triglyceride transfer protein (MTP) to a bivalent $(GalNAc)_2K$ ligand (Figure 18), which has nanomolar affinity for the ASGPR [210], for receptor-mediated delivery of PNAs in hepatic cells [211]. In HepG2 cells, the antisense $(GalNAc)_2K$–PNA at 100 nM concentration reduced the target huMTP mRNA levels by 35–40%, whereas no reduction was observed for scrambled PNA glycoconjugate and unmodified PNA [211]. A radiolabeled [^{125}I]-$(GalNAc)_2K$–PNA accumulated in parenchymal liver cells after intravenous injection in larger amounts than unmodified PNA (46% vs 3%). However, [^{125}I]-$(GalNAc)_2K$–PNA was rapidly cleared from the blood-

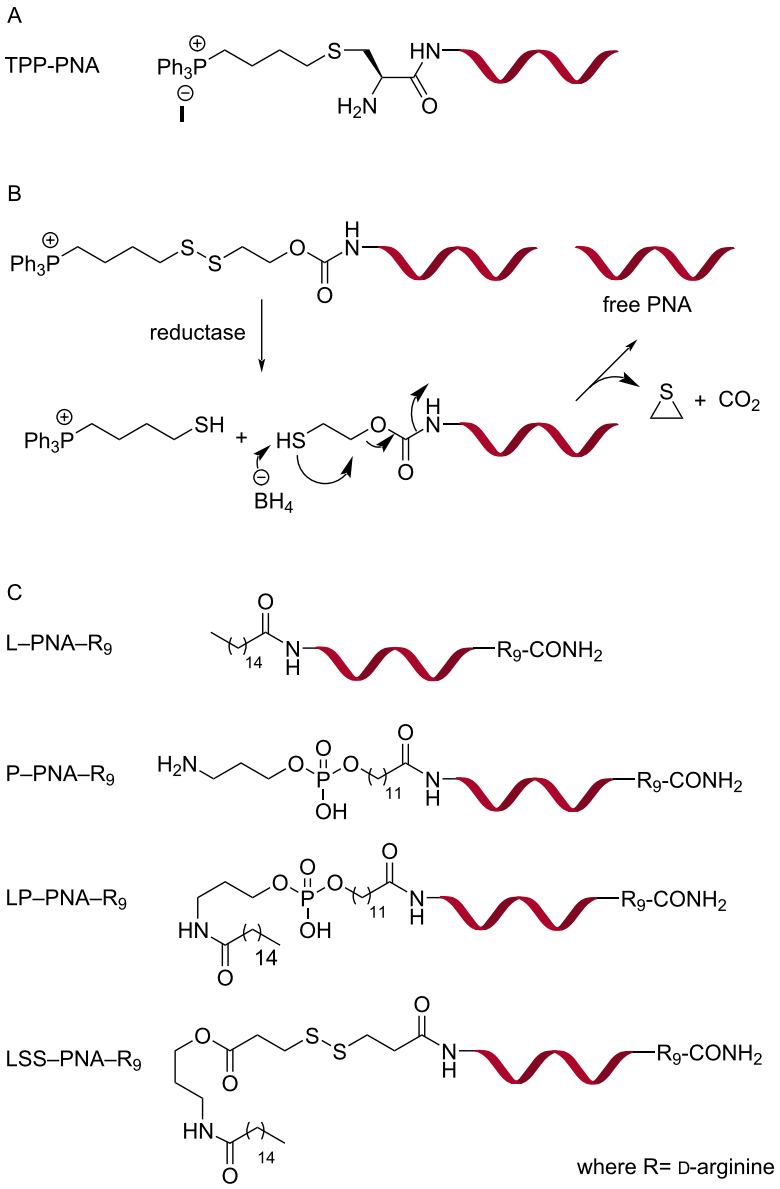


Figure 16: Structures of PNA conjugates with a lipophilic triphenylphosphonium cation (TPP–PNA) through (A) thioether and (B) cleavable disulfide linkage; (C) PNA–R_9 conjugates with lipids, phospholipids and cleavable lipids.

stream with a plasma half-life of 0.38 ± 0.04 min [211]. In another study, $(GalNAc)_2K$–PNA reduced MTP expression in mouse parenchymal liver cells by 70% [212].

Ganesh and co-workers, inspired by the recent success of siRNA-GalNAc platform [213-215], conjugated PNA to a trimeric GalNAc ligands, $(GalNAc)_3$ and $(T\text{-}\gamma\text{-}GalNAc)_3$ for receptor-mediated delivery to hepatocytes [216]. The triantennary $(GalNAc)_3$–PNA conjugate (Figure 18B) at 4.0 μM specifically internalized in HepG2 cells that express ASGPR on their cell membrane, but not in Hek293 cells which lack ASGPR. Interestingly, the architecture of GalNAc conjugation to the PNA influenced the delivery. The trivalent $(T\text{-}\gamma\text{-}GalNAc)_3$ having sequentially appended GalNAc units connected through the γ-carbons of the three T monomers (Figure 18C) showed 13-fold better uptake compared to a branched triantennary $(GalNAc)_3$ unit (Figure 18B) (39% vs 3%) [216]. The GalNAc–PNA conjugates showed no cytotoxicity at 4.0 μM over 12 h; however, no in vitro antisense activity was studied [216].

As mammalian cells are incapable of synthesizing vitamin B_{12}, they have developed a well-established dietary uptake mechanism. Recently, the unique pathway of vitamin B_{12} absorption

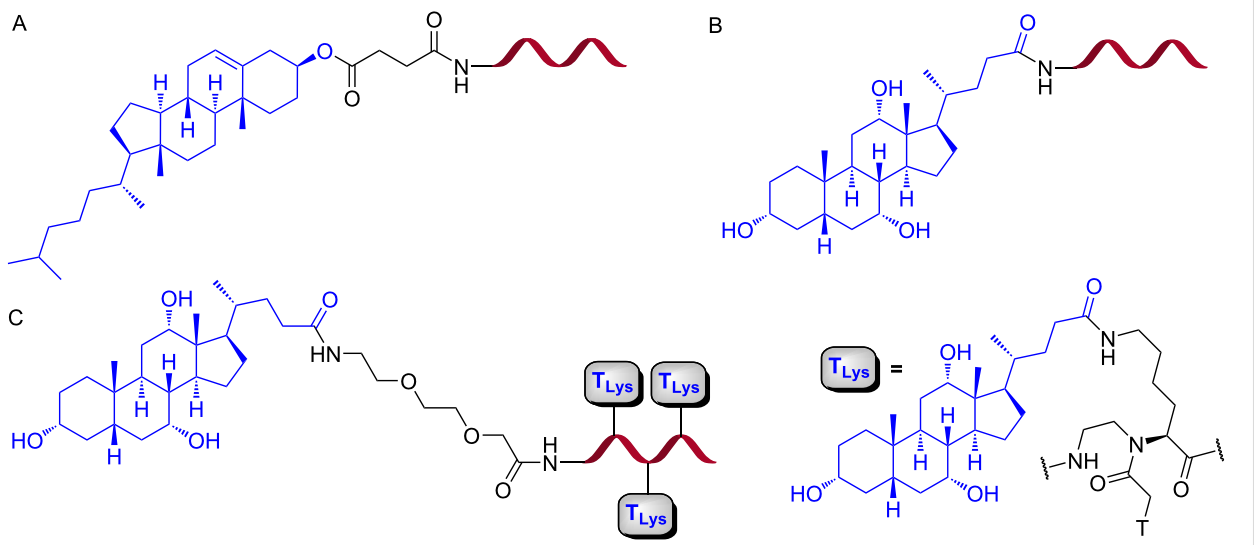


Figure 17: Structures of (A) chloesteryl–PNA, (B) cholate–PNA and (C) cholate–PNA(cholate)$_3$.

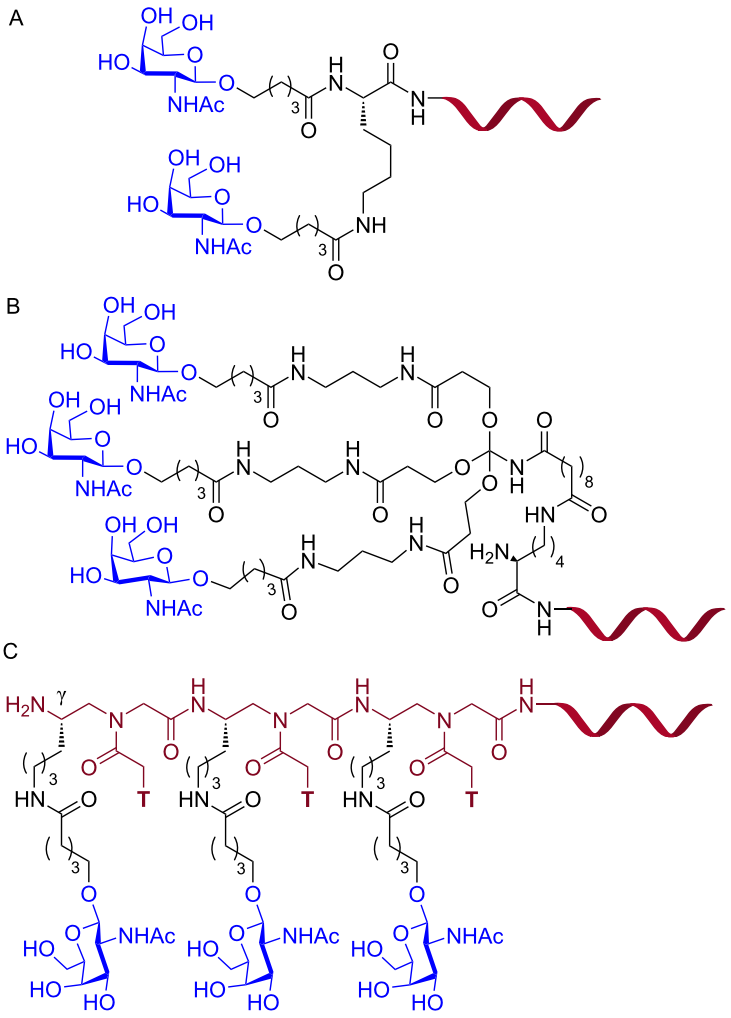


Figure 18: Structures of PNA–GalNAc conjugates (A) (GalNAc)$_2$K, (B) triantennary (GalNAc)$_3$, and (C) trivalent (T-γ-GalNAc)$_3$.

was used to deliver potential drug candidates, such as peptides and proteins, into the cells [217,218]. Gryko, Trylska and co-workers developed a synthetic strategy to covalently conjugate vitamin B_{12} (functionalized at the 5′-position of the ribose sugar) and PNA through a cleavable disulfide linkage (Figure 19) [219]. The same group synthesized a series of vitamin PNA–B_{12} conjugates with cleavable and non-cleavable linkers as well as various spacer length between PNA and B_{12}. All conjugates were stable in bacterial Davis minimal broth and fetal bovine serum [220].

In *E. coli*, the PNA–B_{12} conjugates showed concentration dependent inhibition of *mrfp1* gene expressing a reporter red fluorescent protein, which was in contrast to the PNA–$(KFF)_3K$ conjugate that had constant activity of 70% over the 1–16 μM concentration range [220]. In *E. coli* the conjugates having the longest linker, PNA–$(CH_2)_{12}$–B_{12} and the shortest PNA showed slightly better uptake than PNA–$(KFF)_3K$ conjugate, while the opposite was observed in *S. typhimurium*. The PNA–B_{12} conjugate with a cleavable linker was the least effective in *E. coli*, whereas in *S. typhimurium* all PNA–B_{12} conjugates were equally effective. The activity differences in two bacterial cell lines highlighted the interplay between different bacterial cell walls and B_{12} in the membrane transport system [220]. Although the antisense effect of PNA–B_{12} and PNA–$(KFF)_3K$ conjugates was clearly demonstrated in the bacterial cells, it should be noted that both carriers reduced the binding affinity of PNA for the complementary RNA in cell-free systems [220].

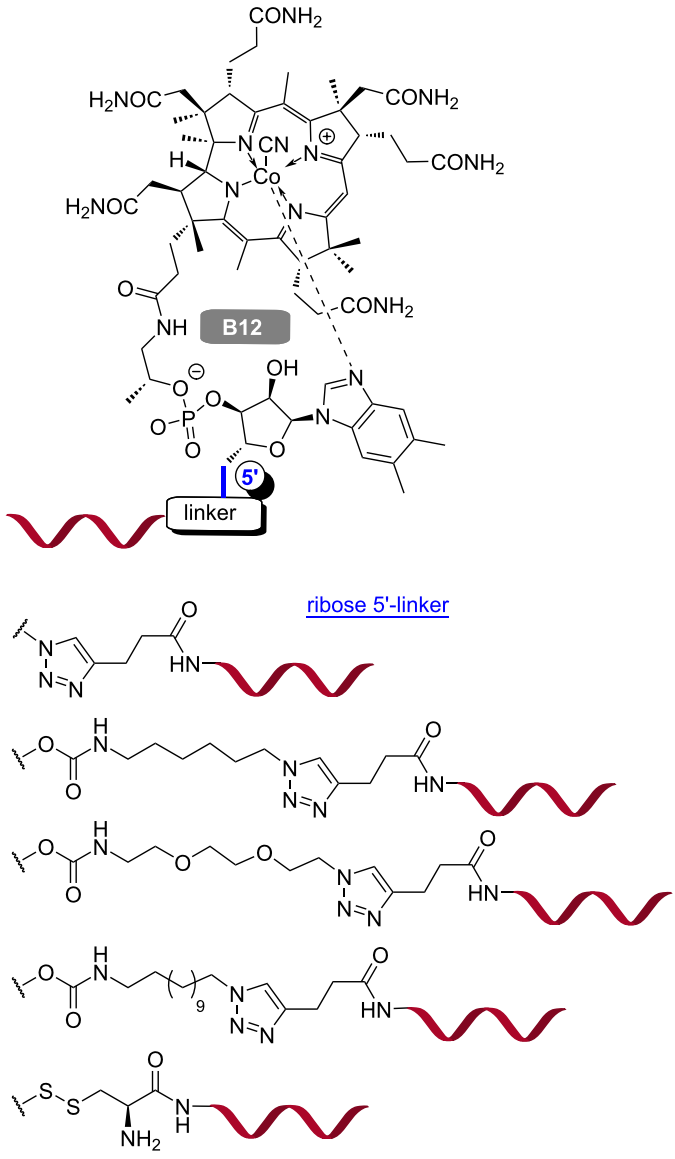


Figure 19: Vitamin B_{12}–PNA conjugates with different linkages.

In follow up studies, Gryko, Trylska and co-workers observed a similar antibacterial activity of PNA–B_{12} and PNA–$(KFF)_3K$ conjugates targeting the essential *acpP* gene in *E. coli* at 5 μM [221]. However, the bacterial growth inhibition activity of the PNA–B_{12} conjugates was media dependent in contrary to the PNA–$(KFF)_3K$ conjugates. Mueller Hinton broth (MHB) is a nutritionally rich medium where the receptors of vitamin B_{12} uptake on *E. coli* cell wall might have saturated, resulting in no antibacterial activity of the PNA–B_{12} conjugates compared to complete bacterial growth inhibition by the PNA–$(KFF)_3K$ conjugates. Changing the medium from MHB to Scarlet and Turner medium restored the bacterial growth inhibition activity of the PNA–B_{12} conjugates [221]. Most recently, Pienko, Trylska and co-workers reported that both B_{12} and B_{12}-conjugates enter *E. coli* via the same route, a TonB-dependent unidirectional delivery through a recognition by the outer-membrane bound BtuB (vitamin B_{12}-specific) receptor [222].

PNA conjugates with cationic carbohydrate ligands: Decout, Pandey and co-workers conjugated PNA with neamine (rings I and II of neomycin B, Figure 20A) [223,224]. The PNA–neamine conjugate showed improved water solubility and antiviral activity in CEM cells infected with HIV-1 carrying a reporter gene (IC_{50} = 1.0 μM). Interestingly, the PNA–neamine (Figure 20B) conjugate cleaved the target RNA sequence specifically [224]. In a later study, a PNA–neosamine (ring II of neomycin B, Figure 20C) conjugated through an amide linkage at the N-terminus of a PNA targeting HIV-1 TAR RNA performed even better than the PNA–neamine conjugate [225]. In CEM (T-lymphocytes) cells, 100% cellular uptake in the cytosol and nucleus of the PNA–neosamine conjugate at 0.3 μM was observed compared to 30% uptake of the PNA–neamine conjugate at 2 μM concentration [224,225].

The mechanism of uptake was studied in the Huh7.5 cells which have larger cytoplasmic space than the CEM cells. Unlike the delivery of PNAs using Tat and poly-arginine conjugates where a majority of the PNA conjugates were sequestered in endosome–lysosome compartments, the cellular distribution of PNA–neosamine conjugates was not affected by chloroquine co-treatment suggesting the absence of endosomal entrapment. No cytotoxicity was observed for the PNA–neosamine conjugates in the 0.2 to 1.0 μM range [225]. In CEM cells transfected with a reporter plasmid construct (pHIV-1 LTR-Luc), the PNA–neosamine conjugate at 0.5 μM and 1 μM inhibited Tat-mediated transactivation of HIV-1-LTR by 64 and 75%, respectively. The PNA–neosamine conjugates inhibited HIV-1 transcription in CEM cells infected with pseudo typed HIV-1 particles carrying a luciferase reporter with IC_{50} = 0.8 μM, without inducing cellular toxicity. Even at the concentrations as high as 100 and 500 μM, the PNA–neosamine conjugates had no negative effect on the cellular proliferation [225].

Kierzek, Chen, Prabhakaran and co-workers conjugated a triplex-forming PNA targeting the dsRNA panhandle structure of influenza virus with neamine [226]. As observed previously [224], the cellular distribution of the PNA–neamine conjugate in MDCK cells (Madin–Darby canine kidney) was homoge-

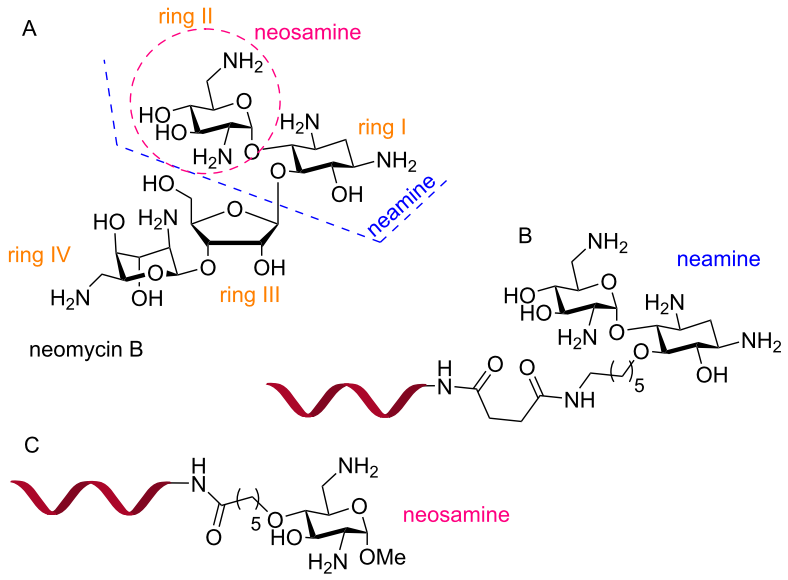


Figure 20: Structures of (A) neomycin B, (B) PNA–neamine conjugate, and (C) PNA–neosamine conjugate.

neous, including nuclei and mitochondria [226]. The PNA–neamine conjugate showed significant inhibition of viral RNA replication ($IC_{50} \approx 3$ μM as measured by qRT-PCR) compared to a lack of inhibition with unmodified PNA. In another study, Chen and co-workers demonstrated that delivery of an antisense PNA–neamine conjugate in HEK 293T cells enabled splicing modulation comparable to the activity of the same antisense PNA delivered using a commercial X-tremeGENE 9 Transfection Reagent (both at 20 μM) [227].

Despite extensive research reviewed above, delivery of PNA is still an unsolved problem. Most of the PNA delivery systems have average EC/IC_{50} values in the range of ≈1–5 μM, with only a few reports of nanomolar activity. Increased cytotoxicity has been a limiting factor for most cationic peptides. Tat and $(KKF)_3K$ peptides are among the most common PNA-delivery reagents, most likely due to the balance between their ability to penetrate the membranes of various cell lines and synthetic accessibility. Common linkers to conjugate CPP and PNA are cleavable disulfides and stable amides, thioethers, or carbamates; the selection of linker becomes important based on the application, tissue/cell line, and mechanism of uptake of the CPP involved. Endosomal or vesicular entrapment and poor release remain as major reasons for the frequently observed micromolar activity of PNA conjugates. Therefore, the development and optimization of new non-endocytic delivery systems such as pHLIP, neosamine, vitamin B12, etc. or new peptides such as cyclic CPPs, etc. capable of efficient endosomal release might help realizing the full potential of PNAs for therapeutic and biotechnology applications. In conclusion, cellular uptake and in vivo delivery of PNA remains an area of active research where future developments hold promise for significant breakthroughs.

PNA probes for research and diagnostic applications

PNA's development has largely been guided by the biophysical limitations of PNA itself, specifically, its reduced solubility compared to native nucleic acids and poor cellular uptake. Early experiments were highly promising suggesting PNA binds nucleic acid targets with significantly higher affinity than analogous DNA/RNA probes. However, these experiments represented the potential of PNA under idealized conditions to bind to target compounds. Synthetic modifications, as previously discussed, have been extensively explored to translate these binding properties to applications in live cells, tissues, and living organisms where conditions are far from ideal. Many in vitro applications, where cellular uptake is not a concern and conditions are well controlled, use minimally modified PNA as the gold standard. Employing PNA in cells or tissues is more challenging, as the matrix becomes increasingly complex, more extensively modified PNAs are required to facilitate solubility and cellular uptake while maintaining selectivity. As a result, PNA has been found to have many applications as a research and diagnostic tool both in the lab and in the clinic [7-9], while advancement of PNA therapeutics, especially when compared to other nucleic acid derivatives [10,11], has notably lagged behind. To better understand the potential of PNA-based technologies, we will examine selected research and diagnostic applications highlighting the versatility of PNA as well as key limitations that hinder the extension of these technologies to therapeutic applications.

PNA-mediated PCR clamping

The high specificity of PNA for target oligonucleotides in vitro was immediately identified as useful for PCR applications. One such approach, termed PNA-mediated PCR clamping, allows for selective PCR amplification of low population target sequences by suppressing the amplification of more abundant targets through PNA/DNA duplex formation. In the first report of PNA clamping (Figure 21A) the authors used 10- to 20-mer PNAs to suppress amplification of a plasmid DNA by clamping its primer [228]. A control plasmid without the target primer sequence showed no inhibition of PCR by the PNA clamps, demonstrating the sequence selectivity of amplification suppression. Clamping at or near the primer binding site was generally more effective while binding further from the primer gave effective clamping in two of the three PNAs tested. A single mismatch sufficiently destabilized the PNA–DNA duplex to allow for primer binding and selective amplification. This approach also worked using homopyrimidine PNAs that formed a PNA/DNA 2:1 triplex. This approach was extended to the detection of Ras proto-oncogene mutations [229]. A 15-mer PNA targeting codons 12 and 13 of wild type Ki-ras suppressed its PCR amplification. Mismatches between the PNA and mutant Ki-ras sequences resulted in lower stability allowing for 23-mer DNA primers to displace PNA turning on PCR amplification of the mutant sequences.

The ability to discriminate single-nucleotide polymorphism (SNP) in mixed populations makes PNA clamping especially useful in cancer detection [7]. Targeting epidermal growth-factor receptor (EGFR) mutations in non-small cell lung cancer revealed genetic heterogeneity in different lung cancer cell lines [230]. EGFR mutations can impact responsiveness to anti-cancer drugs, such as gefitinib. Clamping was done using 14- to 18-mer PNAs along with LNA molecular beacons (Figure 21B) to track the total amplification of different mutant subtypes. The mutated sequences were identified in the presence of 100 to 1,000-fold background of the wild-type EGFR. In total, 30 cell lines were screened by this method with 19 of those containing an EGFR mutation.

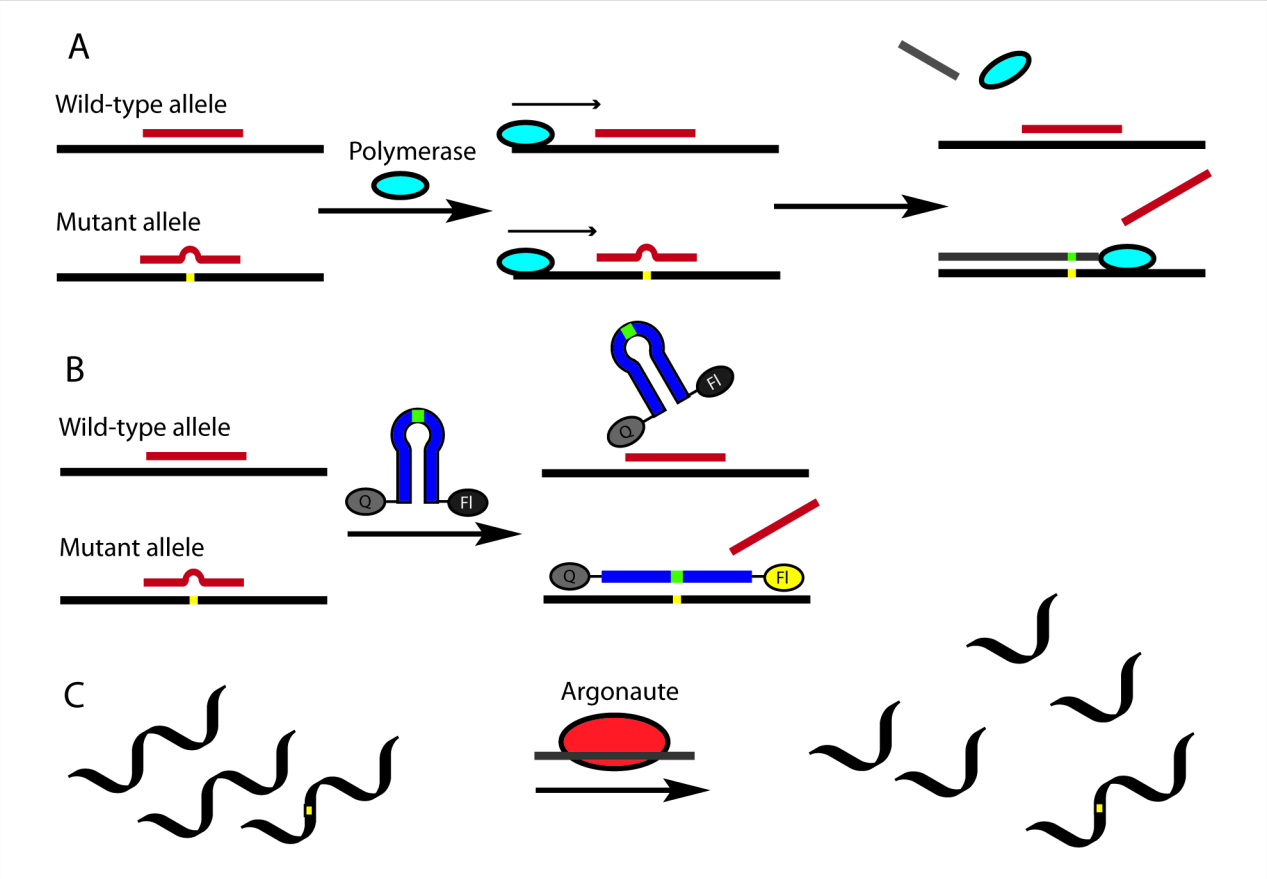


Figure 21: PNA clamp (red) binding to target DNA containing a mixture of sequences (A) PNA binds with higher affinity to the perfectly matched wild-type sequence while binding to the mutant containing as few as one mismatch is weaker. Once elongation begins, the perfectly matched complex stalls the polymerase inhibiting elongation while the mismatched complex dissociates allowing for elongation to continue; (B) LNA probes (blue) can also out compete PNA/DNA complexes mismatched allowing for sequence selective detection of mutant alleles; (C) NAVIGATER uses DNA-guided Argonaute to selective degrade wild-type oligos to enrich the mutant population increasing the sensitivity of PCR clamping.

This approach was later integrated into a clinical application focusing on identifying mutations that make non-small cell lung cancer more susceptible to gefitinib [231]. A total of 132 patient biopsied tissue samples were analyzed at the Saitama Medical University Hospital with 34% being positive for mutations. A total of 29 exon 19 deletions and 16 exon 21 point mutations were detected by PNA clamping, all of which were confirmed by sequencing. The PNA probes displayed excellent sensitivity and selectivity, even for a mutant present at 1% with no false positives. Mutations in EGFR can also be detected in circulating free DNA from plasma [232]. Analyzing plasma samples is less invasive to patients making it an attractive alternative to biopsy sampling. Plasma samples from 60 patients were analyzed using PNA-mediated PCR clamping for mutations in exons 19 and 21 of EGFR. Of the 60 patients, 66.7% tested positive for EGFR mutations in the targeted exon. Of these, 70% were in-frame deletions in exon 19 and 30% were a specific arginine to leucine mutation in exon 21. Detection of mutants present in <1% in plasma samples, such as the T790M, remained a challenge. Sensitivity of PNA-mediated PCR clamping was recently improved by including DNA-guided Argonaute from *Thermus thermophilus* (*Tt*Ago) in an approach called NAVIGATER (Figure 21C) [233]. The DNA guide in *Tt*Ago corresponds to the wild-type allele for various genes (*KRAS, EGFR,* and *BRAF*). Prior to PNA-mediated PCR clamping, *Tt*Ago enriches either circulating free DNA or mRNA in mutant alleles by cleaving wild-type alleles complementary to the DNA guide. Sensitivity of PNA-mediated PCR clamping to mutations increased roughly 10-fold through this enrichment.

PNA-mediated PCR clamping directly applies PNAs high binding affinity and selectivity to silence an enzymatic process. Hybridization of PNA probes targets wild-type sequences to suppress their amplification with excellent selectivity and sensitivity blocking amplification based on a single nucleotide difference. While this is certainly impressive with clear implications in antisense and antigene applications PCR is an in vitro application that bypasses cellular uptake, which remains a significant roadblock to effective application of PNA in vivo. As the

PCR application is in vitro, PNA already displays sufficiently high affinity and selectivity and therefore requires minimal improvements. Instead, most improvements in PCR technology have come from improved sampling methods either from a clinical standpoint (i.e., circulating free DNA detection) or from a biochemical standpoint (i.e., enrichment of low population species via NAVIGATER). Regardless, the application of PNA in PCR demonstrates both its selectivity and specificity as well as PNAs ability to impact enzymatic processes as a result of its strong binding.

Rolling-circle amplification

PNA can also be used to liberate a target sequence from dsDNA using bis-PNAs openers to invade the double helix generating a P-loop ssDNA structure (Figure 2C). The liberated ssDNA can then serve as a hybridization platform allowing for oligonucleotide capture, topological labeling, or sequence-specific detection [234,235]. The ssDNA platform can also hybridize with so-called padlock probes to generate circularized oligonucleotides for rolling-circle amplification (RCA, Figure 22). After hybridization to the P-loop, the termini of padlock probes are fused by a DNA ligase generating an earring structure that acts as a primer for DNA polymerase resulting in the synthesis of long, repeating ssDNA which can then be detected. The PNA-mediated approach to RCA was first applied in topological labeling of dsDNA corresponding to the HIV-1 *nef* gene [236]. Two different DNA targets were used to determine the impact of topological constrain on RCA. One target was composed of a linear dsDNA fragment while the other was circularized forming a closed dumb bell structure. RCA proceeded smoothly despite of the geometric constraints of the dumb bell structure. While the kinetics of RCA were slower for the P-loops than for free ssDNA, signal generation still occurred quickly taking less than 90 minutes to reach its maximum.

PNA-mediated RCA has displayed a high level of sensitivity making it amenable to diagnostic applications. Detection of single-copy genomic DNA in *E. coli*, *B. subtilis* and *S. mutans* was accomplished using 8-mer bis-PNA openers and fluorescent probes targeting the ssDNA of the resulting P-loop [237]. A total of eight target sites were tested, all of which resulted in species-specific detection based on their unique P-loop sequences. This approach can be extended to targeting human chromosomal and mitochondrial DNA [238]. Multiple labels could be introduced by simultaneously targeting common 7-mer polypurine sequences flanking unique ≈20 nucleotide sequences. Chromosome specific padlock probes bound to each site specifically. Padlock probes also contained a shared sequence which was targeted by a sequence-specific fluorescent label allowing for visualization of multiple chromosomes with a single fluorescent probe. Chromosome specific labeling occurred for all targets with the main limitation being imaging sequences on sister chromatids with signals being distinguishable in only ≈30% of cases.

PNA can also be used as a capture probe in the design of microarrays for detection of genetic mutations. Recently, detection of mutations in *EGFR* was accomplished using RCA of the ssDNA of *EGFR* [239]. PNA complementary to the conserved 3′-end of the *EGFR* gene was covalently linked to the microarray through the N-terminus. As the target was ssDNA, no openers were required for padlock hybridization. Detection employed fluorescently labeled probes with graphene oxide acting as a quencher to increase sensitivity, which will be discussed in more detail in the coming sections. Selective detection of the mutant *EGFR* over the wild type was achieved using a species-specific padlock probe. Clear bands were observed down to 1 pM of the target sequence and was specific for the targeted mutation with the wild type generating no signal.

In RCA, the strong binding of PNA enables localized disruption of nucleic acid structure through invasion. The formation of P-loops is potentially useful for antigene and gene-editing technology with one major limitation. For simple PNA, low salt

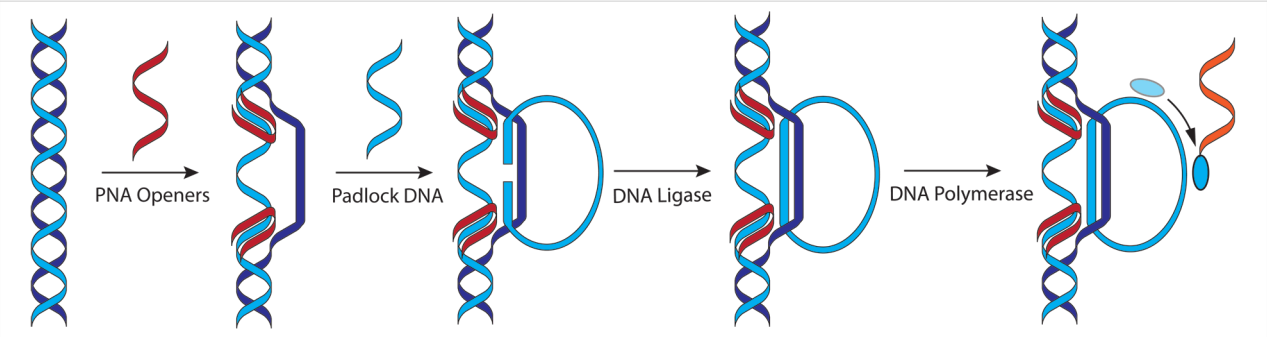


Figure 22: Rolling circle amplification using PNA openers (red) to invade a dsDNA target forming a P-loop. A padlock DNA probe (blue) can bind to the DNA liberated by the PNA openers. Ligase circularizes the padlock DNA resulting in an earring complex which acts as a primer for DNA polymerase. The resulting rolling circle amplification product (orange) can then be isolated or detected in solution.

concentrations are required for invasion to occur that differs significantly from physiological conditions. Under physiological salt concentration, little invasion occurs as the dsDNA is stabilized making P-loop formation difficult. This is in part intrinsic to double-stranded oligonucleotide systems but can be partially remedied using more advanced PNA modifications. For example, replacement of pseudoisocytosine (J) in the triplex-forming portion of the clamp by 2-aminopyridine (M), which displays a higher binding affinity, may improve overall clamping efficiency [31]. The use of M as a partially cationic nucleobase may help counteract the stabilizing effect of salts on dsDNA affording potentially easier invasions.

The in vitro applications discussed above illustrate the power of PNA technology. The strong binding of PNA allows for suppression of enzymatic processes, such as PCR, and enable localized disruption of nucleic acid structure as demonstrated in RCA. PNA-mediate PCR clamping has been particularly impactful in diagnostic applications because of its efficacy and ease of application. The principle of disrupting either enzymatic processes or nucleic acid structure also has significant implications for PNA as a therapeutic. However, the biophysical limitations of PNA in cellulo and in vivo (i.e., low solubility, poor cellular uptake, etc.) have made the transition to antisense and antigene applications challenging.

Detection of DNA and mRNA

Imaging oligonucleotides using PNA is also widespread for in situ, in vitro, and in cellulo applications. The high binding affinity, sequence selectivity, chemical and enzymatic stability, and convenient functionalization makes PNA attractive for oligonucleotide sensing applications, such as fluorescence in situ hybridization (FISH) [8,240]. Fluorescent labeling of PNA is often operationally simple involving conjugation of dyes to the amino terminus, terminally attached amino acid residues, or functional groups of backbone-modified PNAs [241]. PNA-FISH was used to detect immunoglobulin kappa light chain mRNA in paraffin sections of fixed cells from tonsils using antibody-based signal amplification [242]. The immunoglobulin kappa light chain is one of the more abundant mRNAs in tonsil cells, making it an attractive first target. FITC-labeled PNA or DNA probes complementary to the target mRNA were hybridized in the sections of fixed cells. An anti-FITC antibody containing an alkaline phosphatase was then conjugated to the PNA/RNA duplex. After washing, treatment with 5-bromo-4-chloro-3-indolyl phosphate and nitro blue tetrazolium generated the observed signal through phosphatase-mediated enzymatic redox reaction. A similar amplification-based approach was used to detect HIV-1 in the cells of two AIDS patients in 2001 [243]. An N-terminally labeled FITC-PNA probe was designed to hybridize to the HIV protease gene. A horseradish peroxidase labeled anti-FITC antibody was then used to label the PNA. Next biotinylated tyramine reacts with the peroxidase, which is, in turn, labeled with horseradish peroxidase conjugated streptavidin. The cycle is repeated with the last step utilizing an Alexa Fluor 488 labeled streptavidin resulting in multiple Alexa Fluor 488 labels per hybridized PNA complex. Labeling occurred predominately in the nucleus, but some cytosolic labeling was also observed, possibly due to the presence of either HIV-1 DNA or RNA in the cytoplasm. Signal amplification is critical in generating a sufficiently bright enough signal for detection. Enzymatic signal amplification can be effective, but has limited applicability, as it often involves cumbersome antibodies and multiple rounds of amplification to generate a detectable response.

Fluorogenic PNA helps address this limitation through the design of fluorescent systems which are somehow quenched in the absence of the complementary target sequence [241]. Several fluorogenic designs exist with molecular beacons being identified early as a means of increasing the sensitivity of PNA probes [244,245]. Due to sequence complementarity at the beacon termini, these probes form a hairpin structure in the absence of a complementary nucleic acid target referred to as a closed state. In the closed state, a fluorophore (Fl, Figure 23) and quencher (Q) are in proximity resulting in quenching of the fluorescence signal. Two different designs were reported in 1998. Lizardi and co-workers included 7-amino-4-methyl-3-coumarinylacetic acid (AMCA, FL) and 4-((4-(dimethylamino)phenyl)azo)benzoic acid (DABCYL, Q) modified T monomers in the last two AT/TA base pairs of their DNA/PNA chimera beacon (Figure 23A) [244]. Hybridization to the target sequence resulted in linearization of the PNA/DNA chimera probe enhancing fluorescence [244]. Schuster and co-workers replaced nucleobases with aminoacridine (Fl) and anthraquinone (Q) at proximal base pair positions in the middle of a PNA hairpin stem (Figure 23B) [245]. Titration experiments confirmed a 1:1 ratio between the probe and complementary dsDNA hairpins indicating PNA and DNA hairpins both open to form a PNA/DNA duplex.

Soon after, it was discovered that the stem portion of the design could be eliminated as PNA aggregation favored stacking interactions that quenched fluorescence in so-called stemless beacons (Figure 23C). Stemless PNA beacons binding either fully complementary or single-mismatched 16-mer ssDNA gave enhancement of the fluorescence signal [246]. An N-terminal cystine residue was modified with 5-((2-aminoethyl)amino)naphthalene-1-sulfonic acid (EDANS) which serves as the fluorophore while DABCYL-modified adenine acted as the quencher. The position of DABCYL impacted fluorescence enhancement with modification closer to the

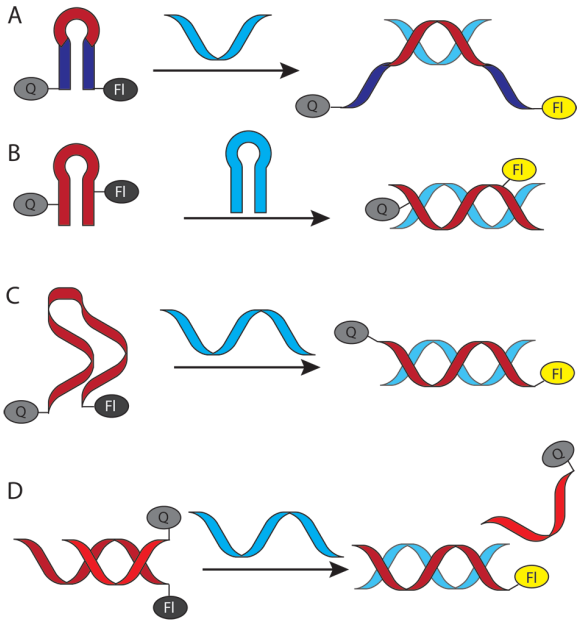


Figure 23: Molecular beacons containing generic fluorophores (Fl) and quenchers (Q) recognizing a complementary oligonucleotide. (A) PNA/DNA chimeras [244] (PNA in red, DNA in blue) and (B) PNA [245] with self-complementary stems were originally used to ensure close proximity of the fluorophore and quencher; (C) stemless beacons [246] lack partially self-complementary sequences instead relying on PNA aggregation to keep the fluorophore and quencher in proximity; (D) two complementary PNAs can also be used to ensure the proximity in dsPNA beacons.

C-terminus giving ≈6-fold enhancement while modification closer to the middle of the sequence giving ≈4-fold enhancement. Stemless PNA molecular beacons were superior to stemmed PNA and stemless DNA molecular beacons when targeting ssDNA and dsDNA [247]. A fluorescein/DABCYL FRET pair attached to the termini of an 11-mer PNA displayed a rapid fluorescence response to ssDNA targets that was independent of salt concentration. The stemless DNA beacon also had a rapid response, but PNA had higher signal-to-noise ratio of ≈10. To target dsDNA, PNA openers were employed to generate a P-loop which acts as the hybridization platform for the PNA molecular beacon. Selectivity was modest with a matched-to-mismatched signal ratio of 1.5 at 25 °C, which increased to 20 at 46 °C.

Another prominent approach in fluorogenic PNA probe design uses thiazole orange (TO) [151,248] or other cyanine dyes. These fluorophores display fluorescence enhancement on binding and intercalation in DNA, which eliminates non-radiative collisional quenching with solvent (Figure 24A). Early designs involved N-terminal labeling of PNA through a flexible linker allowing the dye to intercalate when the PNA probe was hybridized to a target oligonucleotide (Figure 24B) [151]. A 10-mer duplex forming PNA with a 10-atom linker displayed the greatest fluorescence enhancement of 45-fold (Φ_{free} = 0.0015, Φ_{bound} = 0.068). Homopyrimidine sequences had generally lower quantum yields (Φ_{bound} = 0.04–0.07) than mixed sequences (Φ_{bound} = 0.06–0.14). Ground state quantum yields varied significantly, likely because of different π-stacking interactions in the unhybridized probe. Kubista and co-workers applied a TO-PNA probe designed to detect a 1098 bp fragment of the *gusA* reporter gene [248]. A 10-mer polypyrimidine PNA using a 5-carbon linker to the quinoline ring of TO was designed to anneal at 67 °C, between the primer annealing temperature (54 °C) and the elongation temperature (74 °C), so the probe would not interfere with PCR amplification. This method displayed an excellent linear response over a large copy number range (R^2 = 0.999, 300–10^9 copies).

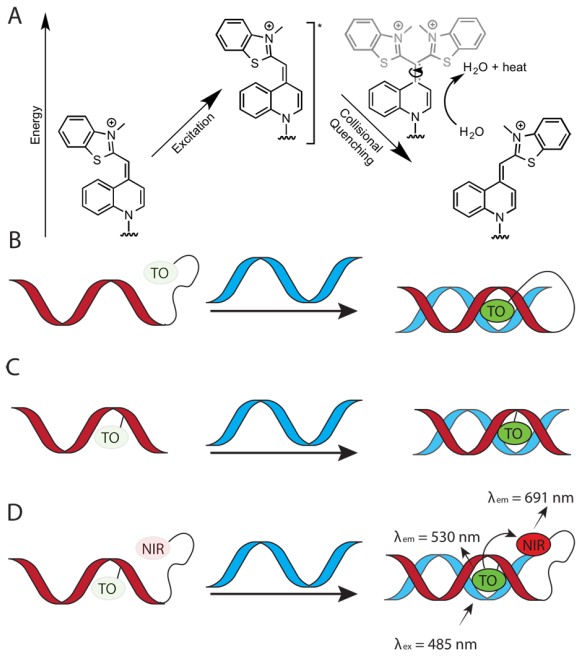


Figure 24: (A) Light-up fluorophores such as thiazole orange display fluorescence enhancement upon binding to a target oligo. In the free, single-stranded state, thiazole orange has a low fluorescence quantum yield as a result of collisional quenching with solvent upon excitation. (B) Thiazole orange can be tethered to PNA either at the terminus [151] or (C) through modified base pairs [150]. Modifying PNA at a nucleobase position with thiazole orange, typically referred to as forced intercalation (FIT) probes also results in sequence specific fluorescence enhancement. (D) FIT probes can be coupled in a FRET system with NIR-667 dye [249].

PNA probes having TO attached through a terminal linker showed promising light-up properties but exhibited significant signal variability depending on the sequence context. A more reliable fluorescence signal was achieved using a modified PNA monomer with TO serving as a nucleobase surrogate (Figure 24C), originally synthesized in 1999 by Seitz and co-workers [250]. While the TO nucleobase decreased PNA's

binding affinity compared to the fully complementary PNA/DNA duplex, the decrease was relatively minor (ΔT_m ≈1–3 °C) and showed little sequence dependency (± 1 °C when TO was paired opposite A, T, C, or G) [150]. Stacking interactions of TO helped stabilize PNA–DNA duplexes while simultaneously enhancing the fluorescence signal. The fluorescence response of TO was sensitive to the opposing nucleobase with fluorescence enhancement decreasing in the order of T > G > C > A.

Seitz and co-workers explored detection of single nucleotide polymorphisms using PNAs modified with the TO nucleobase [251]. To optimize these FIT-probes, attachment of TO through the quinoline or benzothiazole ring using linkers of various lengths (n = 1, 2, or 5) was tested in a 12- and 13-mer PNA against complementary 12- or 13-mer ssDNA. The FIT-PNA probe with the shortest linker attached to the quinoline ring had the highest sensitivity to mismatched base pairs adjacent to the TO nucleobase. Differences in melting temperatures ranged from 8 to 15 °C depending on the sequence as well as the position and identity of the mismatch. Fluorescence enhancement was 11–19-fold for fully-matched sequences while mismatched sequences only showed a 4–8-fold increase. Increasing the temperature increased mismatch discrimination.

FRET-based systems (Figure 24D) can use a single PNA containing intercalating nucleobases such as TO and a terminally tethered chromophore. Normally, this would be problematic as the background FRET signal may be high. When TO is used as a FRET donor, this is not the case, as fluorescence from TO is minimal in the unhybridized probe. Initial reports used NIR-667 conjugated to a lysine residue to serve as a FRET acceptor for TO [252]. Later studies significantly expanded the list of FRET pairs involving TO [253]. The presence of complementary DNA resulted in a 7- to 28-fold increase in TO fluorescence and a 15- to 89-fold increase in NIR-667 fluorescence. Several probes displayed significant enhancement in both TO fluorescence and FRET. One example using indotricarbocyanine (ITCC) separated by 10 nucleotides from TO gave a 452-fold enhancement in TO fluorescence on binding. Another example using NIR-664 separated from TO by 10 atoms displayed a 254-fold enhancement in FRET signal.

Intron splicing of mRNA can be monitored using two labeled PNAs in a similar FRET-based detection method. This method employed two PNAs using TO and Alexa-594 to target the *RPS14A* gene mRNA [249]. In the unspliced pre-mRNA, these two PNAs are separated by >300 nucleotides. Upon splicing, this distance is shortened to 6 to 12 nucleotides, increasing FRET efficiency (Figure 25A). Using TO as a FRET donor can be exceptionally useful as the FRET signal is dependent not only on distance but also on intercalation of TO. Similarly, Artero and co-workers used Cy3/Cy5 labeled PNAs as FRET-based probes to visualize lymphocyte antigen 6 pre-mRNA which has two isoforms resulting from mRNA splicing in HeLa cells [254]. One N-terminal Cy5-labeled PNA was used as the acceptor with different Cy3-labeled PNAs acting as donors. The Cy3-labeled PNAs targeting either the spliced form or unspliced form of the mRNA displayed the expected FRET response on mRNA splicing.

Templated fluorogenic reactions use similar principles as FRET-based probes in that two probes with terminal labels are designed to hybridize in proximity on a target strand. Unlike simple FRET, a chemical reaction occurs due to the proximity of the labels, which produces a fluorescent signal [255]. Early efforts used Staudinger reaction to liberate amino groups from azido-modified fluorophores such as azidocoumarin or azidorhodamine (Figure 25B) [256,257]. The coumarin-based templated reaction used a C-terminal 7-azidocoumarin label and an N-terminal triphenylphosphine modification [256]. The reaction gave excellent fluorescence turn-on using two 8-mer PNA probes targeting an 18-mer ssDNA target. As little as 1% of the matched template generated a fluorescence response. In the presence of 20% template, single mismatches were easily discriminated with <5% conversion after ≈40 minutes compared to >30% conversion in the same time frame for the matched sequence. A 10-fold increase in fluorescence was observed using a catalytic amount of matched template DNA after only 15 minutes. The method was extended to visualize mRNA of O-6-methylguanine-DNA methyltransferase (MGMT) in HEK293 cells using cell-permeable GPNAs (Figure 5) having azidorhodamine and tris(2-carboxyethyl)phosphine (TCEP) modifications [257]. Incubation times were relatively short, less than 90 minutes, with the templated reaction showing similar fluorescence enhancement and mismatch discrimination as the azidocoumarin system.

Detection of mRNA in cellulo is more restrictive than detection in vitro requiring careful consideration of probe biophysics. FIT-PNA probes using TO, oxazole yellow (YO), and benzothiazole orange (BO) are ideal in this application as the fluorophore is relatively small and cationic limiting any negative impact on solubility. FIT-PNA probes enabled multichannel detection of influenza mRNA in MDCK cells [258]. In this study, 14-mer PNAs with TO outperformed PNAs with YO and BO in photophysical response at 25 °C displaying a 16-fold enhancement compared to 3.4-fold for YO and BO. At 60 °C all dyes performed admirably giving 34-, 15-, and 27-fold enhancements, respectively. Two probes, TO-FIT probe for neuraminidase (NA) and BO-FIT probe for matrix protein 1 (M1) were used for qRT-PCR as well as in-cell imaging. Streptolysin O facilitated the uptake of FIT PNAs into MDCK cells,

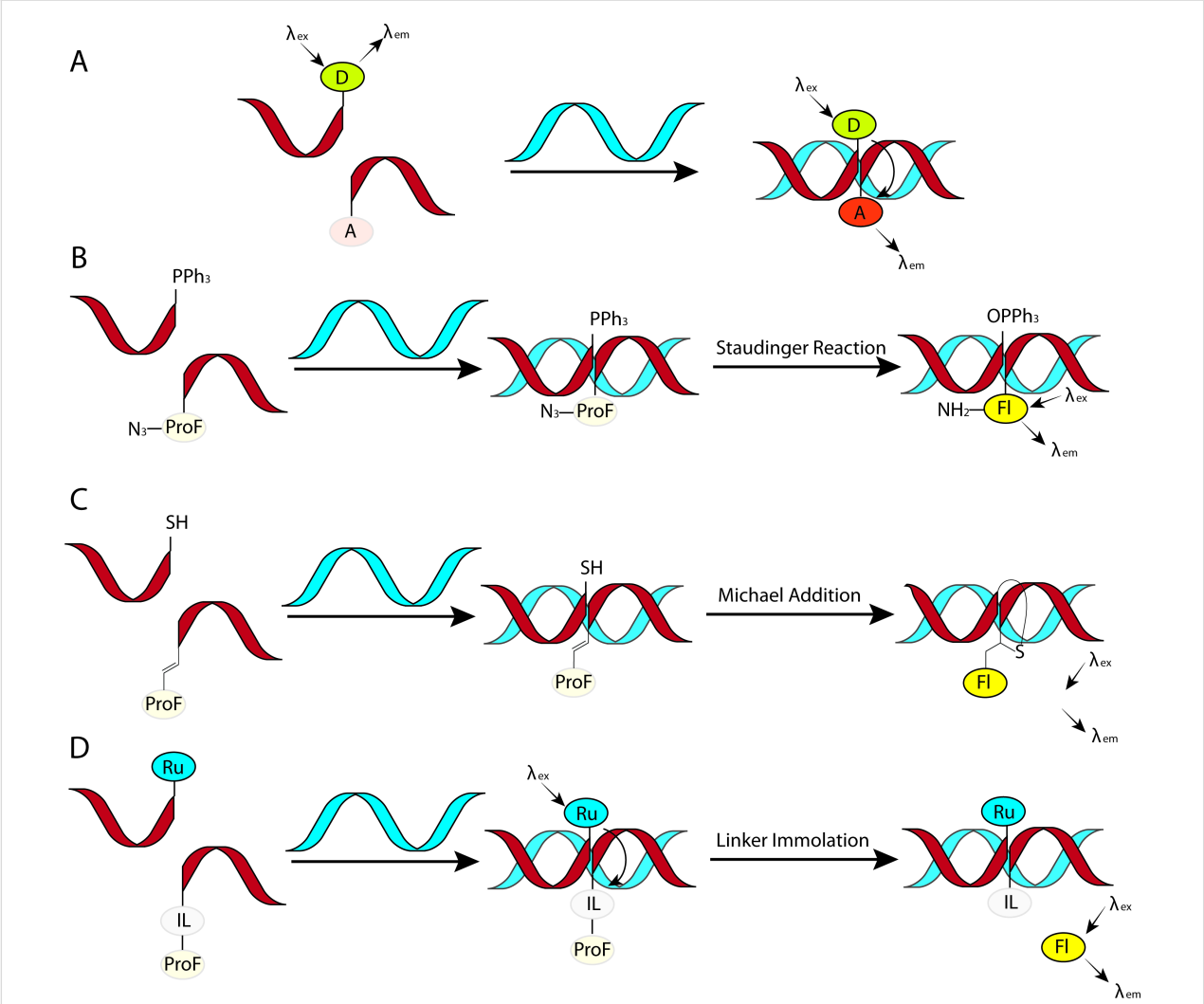


Figure 25: Templated fluorogenic detection of oligonucleotides using two PNAs. (A) Templated FRET depends on hybridization of PNAs to adjacent positions on the target sequence to bring the donor and acceptor in proximity. Templated reactions such as (B) Staudinger reaction or (C) conjugate addition of thiols can be used to turn on fluorescence of a caged pro-fluorophore. (D) Photochemical templated reactions target an immolative linker which both tethers and quenches a pro-fluorescent molecule.

which were then infected with influenza A. Fluorescence from the two probes developed at different time points with the TO-FIT probe for neuraminidase generating signal at two hours post infection. The signal was initially localized in compartments identified as nucleoli and spread into the cytosol over time. Control cells generated no signal suggesting the localization was a result of mRNA distribution and not PNA compartmentalization. The BO-FIT probe for M1 generated diffuse signal throughout the cell starting at five hours post infection.

Fluorescence reporters exhibiting red-shifted emissions are generally desirable because autofluorescence is reduced and the lower energy light required for excitation is less damaging to cells. Bisquinoline (BisQ) is a cyanine dye similar to TO with red-shifted emission (λ_{em} ≈ 610 nm BisQ, λ_{em} ≈ 500 nm TO) [155]. BisQ FIT-PNAs targeting the mutated *KRAS* oncogene DNA or mRNA had exceptional brightness (quantum yields Φ_{bound} = 0.22–0.26) and showed selective fluorescence from Panc-1 cells (*KRAS* mutant) but not HT-29 or Bxpc-3 cells (*KRAS* wild type) [155]. The sequence context for BisQ fluorescence response has been examined thoroughly to help in the design of BisQ FIT-PNAs [259].

BisQ-modified PNAs have been used to detect KRT20 mRNA, which is over expressed in colorectal cancer, in live cancer cells [260]. An 18-mer BisQ-modified PNA displayed 20-fold fluorescence enhancement in the presence of the target sequence while a scrambled PNA sequence containing BisQ gave no fluorescence response. The red emission from BisQ allows for superior detection of mRNA in tumors as red light scatters less

and can penetrate deeper in tissue. Using two-photon microscopy allows for excitation of BisQ using longer wavelengths of light further improving tissue imaging. Spraying tumors with a solution of BisQ FIT-PNA targeting KRT20 visualized the mRNA to a depth of 240-micron in tumor tissue.

Detection of genes and mRNA has driven a large number of innovations in PNA technology, specifically in detection methods. Moving from parafilm sections to in-cell imaging showed that PNA can progress from in situ applications to more complex biological systems. The shift of detection to fluorogenic designs significantly improved the technology for diagnostic and research applications. Specifically, the development of FIT-PNAs has greatly improved the consistency of fluorescence responses using simple modifications without compromising sensitivity. FIT-PNAs have also been applied to in imaging in cells and tissues representing a significant step forward in this technology. Templated reactions have also shown promise as a sensitive method of detection with excellent selectivity which could be adapted for potential sequence selective payload delivery. The main limitation in extending these types of systems to therapeutics is the 1:1 ratio of PNA to target oligonucleotide. While this is acceptable for many modern sensing applications, the non-catalytic nature of silencing puts PNA/DNA or PNA/RNA duplex-based systems at a distinct disadvantage to enzymatic methods such as CRISPR/Cas9 or short interfering RNA (siRNA).

Detection of ribosomal RNA

The stronger affinity of PNA for RNA compared to DNA as well as the abundance of ribosomal RNA (rRNA) in cells makes rRNA targeted PNA a powerful diagnostic tool. Initial reports targeting rRNA focused on *Mycobacterium tuberculosis* complex (MTC) which is a genetically related group of bacteria responsible for tuberculosis. Both the 16S and 23S ribosomal subunits of several mycobacteria were screened to find partial sequence alignments specific to two members of MTC: *M. tuberculosis* and *M. bovis*. N-terminal FITC-conjugated 15-mer PNAs at 25–100 nM were shown to selectively target rRNA sequences specific to MTC complex or other mycobacteria [261]. After this initial report, both bacterial [262] and fungal infections [263] were identified from blood culture tubes using rRNA targeting PNA probes. Both publications used 15-mer FITC-conjugated PNAs targeting either the 16S rRNA of *Staphylococcus aureus* or the 26S rRNA of *Candida albicans* in clinically relevant samples. A total of 48 clinical isolates of *S. aureus* produced only one false-positive for *Stomatococcus*. Testing of 87 clinical blood culture specimens gave a 97% true positive rate and a 100% true negative rate [262]. For *C. albicans*, this technique had 100% sensitivity and specificity in samples of 148 clinical isolates and 33 real yeast-positive clinical blood cultures [263]. Both tests were fast and accessible, taking only 2.5 hours to obtain a potential diagnosis using techniques common in microbiology labs.

Raskin and co-workers imaged rRNA in fixed *E. coli* cells using a 16-mer stemless PNA molecular beacon with C-terminal DABCYL and N-terminal FITC labels [264]. This work compared the PNA beacon with an analogous 24-mer stem-containing DNA beacon. The DNA probe at 50 nM showed a linear response of fluorescence intensity depending on concentration of extracted target rRNA down to 12.5 nM of target rRNA, while the PNA probe's linear response extended down to 0.39 nM. Both probes showed selective staining of rRNA in *E. coli* and *M. acetivorans*, but the PNA beacon was 3-fold brighter than the DNA-based probe. Signal intensity increased sharply during the first 15 minutes while reaching its peak at one hour while the DNA probe required several hours to generate a fluorescence response.

The exceptional sequence specificity of PNA along with high sensitivity and short time of analysis in imaging rRNA from blood cultures led to early development of commercial kits for PNA testing. Specifically, identifying *Candida* fungi has become increasingly important in determining course of treatment as different species of *Candida* respond differently to common antifungal drugs such as fluconazole. A multi-institute study comparing the *Candida* PNA FISH assay from AdvanDx with other routine tests showed that PNA FISH improved accuracy in microbe identification [265]. Similar results were obtained for PNA-FISH detection of different Gram-positive *Staphylococci* where accurate detection of *S. aureus* significantly improved the outcome for patients in intensive care [266]. In some instances, mutations in the rRNA sequence of bacteria or fungi may be associated with phenotypic changes such as antibiotic resistance. PNA-FISH is capable of identifying these mutant strains by directly targeting the rRNA mutation as demonstrated by the identification of clarithromycin-resistance in *Helicobacter pylori* [267]. Several different point mutations in the peptidyltransferase region in domain V of the 23S rRNA gene associated with the clarithromycin resistance were identified using 15-mer PNAs. These PNAs were labeled at the N-terminus with either Alexa Fluor 488 for the mutant rRNA or with Alexa Fluor 594 for the wild type rRNA. The PNAs were specific and sensitive to their target mutants and discriminated resistant and susceptible strains because of a single mismatch in the middle of the 15-mer PNA sequence.

Recent PNA probes targeting rRNA for clinical applications expand the scope of testing, improve the signal-to-noise, and reduce time of analysis. *Candida* QuickFISH BC from AdvanDx improves on their PNA-FISH kit for *C. albicans*

[268]. Specific labeling for *C. albicans, C. galbrata, or C. parapsilosis* is done in multiplex using species-specific PNAs with different fluorescent labels. Quencher probes are then used to eliminate fluorescence from unhybridized PNA. Overall, the sensitivity was 99.7% and the specificity was 98.0% for the three strains of *Candida* targeted in this study. The time of analysis for this approach is only 30 minutes affording a fast and accurate diagnosis of multiple strains of *Candida* in one test.

Recently, a single-cell-based microfluidic detection of Gram-negative bacterial pathogens used molecular beacon PNAs targeting rRNA [269]. Two beacon designs, dsPNA beacons and stemless ssPNA beacons (Figure 23D), were compared. Cell lysates were incubated with PNA beacons at 25 nM to 200 nM followed by a quencher DNA sequence to eliminate fluorescence from any unhybridized probe. Probes were tested on four bacterial strains: *E. coli* (UPEC), *P. aeruginosa* (Pa127), *P. mirabilis* (Pm159), and *K. pneumoniae* (Kp128). The first two served as positive controls while the last two served as negative controls. Of these conditions, the dsPNA beacon at 25 nM had the highest signal-to-noise ratio and was species specific for *E. coli* and *P. aeruginosa*. Single cell experiments in 7 pL droplets using microfluidics confirmed results observed in bulk fluid analysis. The experiment aimed to seed 10% of droplets with bacterial cells. The dsPNA beacon resulted in 8% of droplets displaying fluorescence after 30 minutes compared to 1% of the droplets treated with the ssPNA probe suggesting faster hybridization of the dsPNA probe. Signal from bacteria-containing droplets compared to empty droplets was higher for dsPNA probes (≈3.4) to the ssPNA probe (≈2.2), suggesting that in no-wash applications, dsPNA beacons are superior to stemless molecular beacons in high-throughput diagnostics.

Diagnosis of bacterial and fungal infections is exceptionally accurate using PNA-based probes. The strong binding of PNA and the abundance of target rRNA has led to the development of commercial kits for disease identification. The simplicity and accuracy of these diagnostics has resulted in wide-spread adoption of this technique in clinical settings. While most PNA applications in rRNA sensing are limited to in vitro experiments, the strong binding of PNA to this critical component of cellular machinery make rRNA-targeting PNA therapeutics an attractive approach to treating microbial infections. The abundance of rRNA in cells, similarly to mRNA discussed previously, would likely be limiting to this technology as PNA binding and inactivating the rRNA would be non-catalytic and limited by the cellular uptake of PNA. However, development of therapeutic technology based on rRNA targeting with PNA may help supplement the physician's toolkit as bacterial resistance to traditional antibiotics increases over time.

Detection of microRNAs

High binding affinity is critical for detection of microRNAs (miRNAs) because of their generally low copy number in cells and short sequence length (18–22 nucleotides). With miRNAs identified as increasingly prominent players in regulating gene expression, detection and quantification of these species is critical to deepening our understanding of miRNAs relation to disease. PNA-based fluorescence and electrochemical sensors of miRNAs have seen increasing use in a number of applications as highlighted in a recent review [9]. Early attempts at miRNA detection mirrored those of mRNA, using in situ enzymatic amplification to generate an optical signal [270]. Electrochemical detection using PNA for miRNA have also been explored with early reports using silicon nanowires [271]. As PNA lacks an intrinsic charge, this approach is well developed displaying excellent sensitivity.

Photochemically-induced templated reactions involving a $[Ru(bpy)_2phen]^{2+}$ catalyst have attracted significant attention due to ease of spatiotemporal control. In this two PNA templated system, one probe has an N-terminal rhodamine attached via an azide-caged immolative carbamate with the second probe containing a C-terminal $[Ru(bpy)_2phen]^{2+}$ group which can be excited with 455 nm light [272]. In the presence of a reducing agent, such as sodium ascorbate or NADPH, and the template sequence, excitation of $[Ru(bpy)_2phen]^{2+}$ results in azide reduction, which uncages rhodamine and generates a fluorescence signal. This process results in signal amplification as uncaged PNA dissociates and is replaced by another PNA still bearing the caged fluorophore. Backbone-modified (γ-CH_2-OH, Figure 5) PNAs displayed the fastest reaction times and were sensitive to single mismatches when targeting ssDNA containing the sequences for either miRNA-21 or -31. Increasing the distance between the probes on the target sequence slightly decreased the efficiency of reaction, but ≈50% conversion was reported after 90 minutes, even when the PNAs were separated by 12 nucleotides. Templated reactions in BT474 cells and HeLa cells targeting miRNA-21 and -31 selectively showed fluorescence signal when using perfectly matched PNA, while a single mismatch in one of the two probes resulted in no observable fluorescence.

A FRET-based detection method using fluorescently labeled PNA along with nano graphene oxide referred to as PANGO have also been used to detect miRNAs [273]. Graphene oxide facilitates cellular uptake of PNA [273] while also quenching fluorescence via π-stacking [274]. This approach was used to target miRNA-21, -125b, and -96 with carboxy fluorescein (FAM), 6-carboxy-X-rhodamine (ROX), and Cy5 N-terminally-labeled PNAs. In all cases, a steady increase in fluorescence was observed up to 1,000 nM with a 1 pM detection limit. This

approach could be multiplexed for miRNA detection in complex samples as no cross-reactivity was observed between the miRNAs and probes. The method did not show significant toxicity with a >90% viability in four cancer cell lines at ≤200 μg/mL of PANGO complex. While impressive, this approach lacks the signal amplification of templated reactions. RCA of miRNA synthesizing ssDNA with tandem repeats can be used along with PANGO complexes to increase detection sensitivity [275]. As discussed previously, RCA generates long, repeating ssDNA using a circularized padlock DNA probe complementary to the target oligonucleotide. In this case, the target miRNA-21, overexpressed in lung cancer patients, was normalized against miRNA-16. In the presence of graphene oxide, fluorescence of unbound PNA was completely quenched. The limit of detection was 0.4 pM for isolated miRNA and 0.7 pM when tested using total cellular RNA from A549 lung cancer cells. The method also worked in multiplex detection of miRNA-21, -31, and, -155 using three different FITC-, ATTO550-, and Cy5-labeled PNAs in a multi-well plate.

Fluorogenic coumarins can be uncaged using thiols through Michael additions that disrupt conjugation to a quencher. This was used to detect miRNA-132, -141, and -375 using PNA having C-terminal *n*-butyl thiol groups and PNAs having N-terminal styrene-quenched coumarin 334 [276]. The presence of DNA corresponding to the matched miRNA target gave a 15-fold increase in fluorescence intensity using two 7-mer PNA strands. Mismatches in the middle of the PNA probes significantly (>50%) decreased the fluorescence while mismatches close to the end of probes resulted in a modest fluorescence reduction.

Lateral flow devices using PNAs have been developed for miRNA sensing. These devices often employ a similar design using a streptavidin-labeled lane in the middle of a strip of nitrocellulose paper which binds to a so-called PNA anchor through an N-terminal biotin label (Figure 26). Detection is then achieved through ligation to a separate fluorescently labeled PNA or through a templated fluorogenic reaction which forms a covalent linkage between the PNA probes [277,278]. Native chemical ligation is a well-established reaction involving a cysteine-mediated reaction of thioester to generate a peptide bond. A seleno-variant of this reaction was used in a lateral flow device and demonstrated a 10-fold faster reaction rate than the sulfur-based reaction (Figure 26A) [278]. This reaction was used to ligate two 9-mer PNAs with one containing a FITC-label. The limit of naked eye detection was <0.1 nM based on titration experiments of ligated PNA product. This method was then used to detect miRNA-31 and -21 in lysates from HeLa, MCF-7, and HEK293-T cells. HeLa cell lysates were positive for miRNA-31 and negative for miRNA-21, while MCF-7 cells gave the opposite result. HEK293-T cells were used as negative controls and, as expected, displayed no labeling. Another lateral flow device was developed by Ladame and co-workers using two 7-mer PNAs to detect miRNA-150-5p, which is a biomarker for preterm birth [277]. The two PNAs were connected through a templated Michael addition (Figure 26B). The detection limit was 9 nM with a linear correlation between signal intensity and target concentration between 10–200 nM. Plasma extracts from 18 patients tested using the lateral flow strip generated a statistically greater fluorescence signal (p value = 0.0006) from eight patients who delivered preterm than from the ten who delivered at term [277].

Targeting double-stranded pre-miRNA hairpins is also an effective sensing strategy as demonstrated by Winssinger and co-workers using two triplex-forming PNAs modified with $[Ru(bpy)_2phen]^{2+}$ and coumarin attached via an immolative pyridinium linker [117]. Sequence context proved to be important in maintaining selectivity for dsRNA pre-miRNA-31 hairpin, where longer PNAs (an 11-mer and 13-mer) showed some off-target fluorescence in the presence of ssRNA from the cleaved pre-miRNA, while shorter sequences (two 9-mer PNAs) were selective for dsRNA. A detectible fluorescence response was observed after 30 minutes in the presence of 12.5 nM pre-miRNA-31 [117]. Signal enhancement using this approach was as high as 20-fold.

Detection of miRNA is exceptionally important in the study of genetic diseases such as cancer. A number of miRNA biomarkers for disease and injury have been established and the ability to detect and quantify miRNAs with increasing sensitivity and precision will undoubtedly expand this list. Despite their relatively low abundance in cells, detection of biomarker miRNAs using PNA has developed rapidly as a viable diagnostic tool due to PNA's strong affinity for RNA. PNA-based detection of miRNA has even been applied to potential consumer-friendly products, such as lateral flow devices. While many current applications focus on processed miRNA, targeting pre-miRNA is also a viable diagnostic approach. Developing methods for detection of both miRNA and pre-miRNA using PNA can help with understanding the role of miRNA in cells. Targeting these species has potential therapeutic implications as well, as the PNA–RNA complex may prevent processing of pre-miRNA or loading of miRNA into the RISC complex for mRNA silencing. The role of miRNA in coordinating cellular function through fine-tuning mRNA levels in cells makes it an attractive potential therapeutic target. Unlike mRNA or rRNA, the low copy number of miRNA and their broad effects means PNA-based silencing or attenuation of miRNA may have a strong impact provided the affinity of PNA is high enough. Hence, exploring

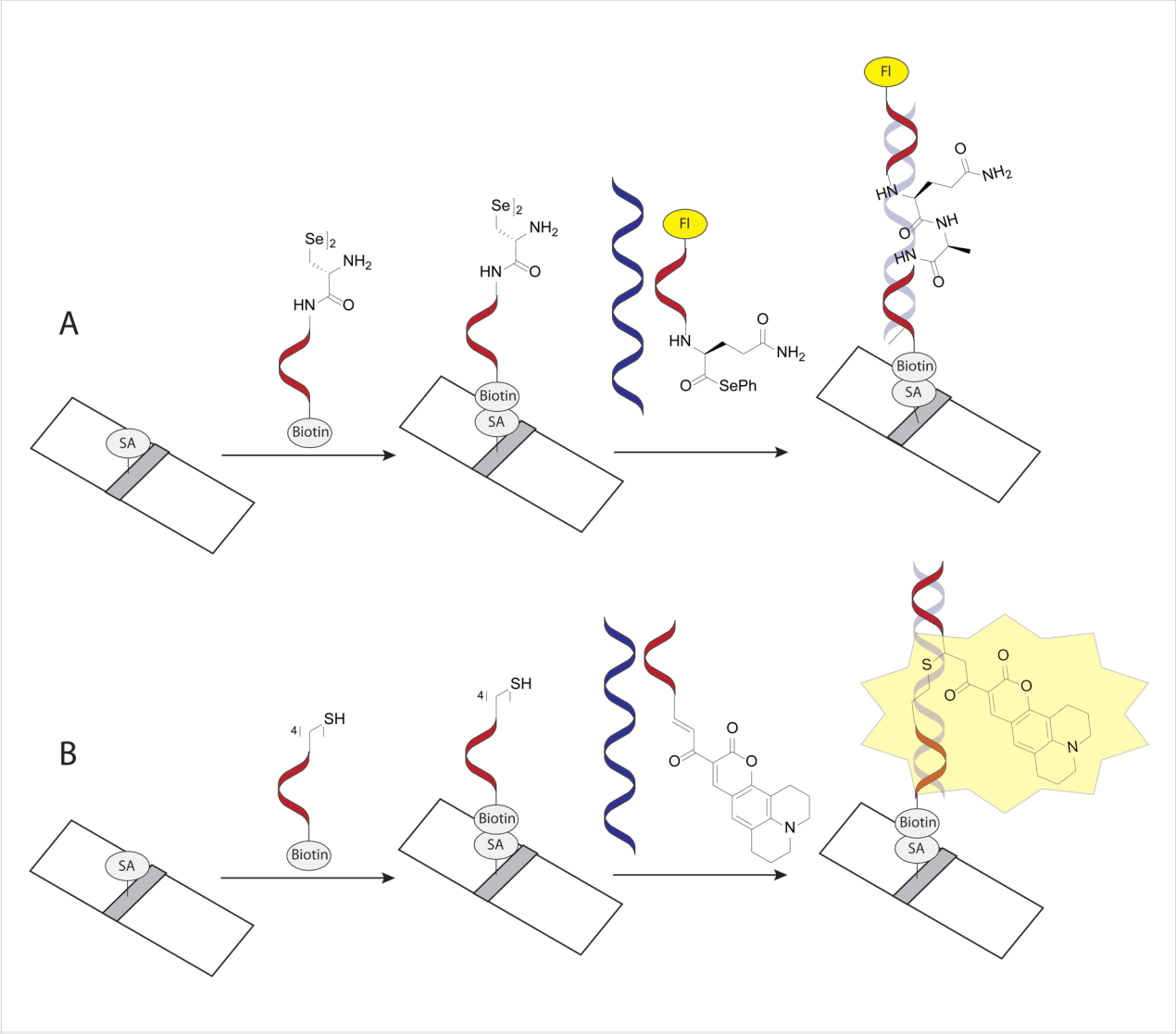


Figure 26: Lateral flow devices use a streptavidin labeled strip on nitrocellulose paper to anchor a capture PNA (red). The target oligonucleotide (blue) and the detection PNA probe (red) are then allowed to run the length of the strip. If the target is present, it will act as a hybridization scaffold bringing the two PNAs in proximity. This allows for either simple ligation (A) or fluorogenic ligation (B) which generates an optical signal allowing for detection of the target.

anti-miRNA applications could be a fruitful area of research for PNA therapeutics.

Protein sensing

While PNA is typically designed to target nucleic acids, it can also be used to sense proteins. Hairpin peptide beacons function similarly to regular molecular beacons: they utilize a protein-targeting peptide sequence flanked on either end by short complementary strands of PNA to form a closed hairpin structure. Protein binding to the peptide sequence unfolds the structure giving fluorescence enhancement. Several proteins that bind short peptide substrates were targeted using this approach [279-281]. Src kinase is an important signaling protein that interacts with other proteins through its SH2 domain, which binds phosphorylated tyrosine residues on target proteins. A short peptide sequence from a known Src-SH2 binding protein containing phosphotyrosine served as the protein binding site for the probe. Two complementary 4-mer PNA sequences, terminated with pyrene-modified lysine residues that induced excimer/monomer fluorescence, closed the harpin. At 1 μM of peptide–PNA conjugate in the presence of 16 μM SH2-protein, a 10-fold increase in fluorescence was observed. Renin, an aspartic acid protease, was also targeted using a short renin peptide-inhibitor sequence and 4-mer PNA stems modified with NIR644 and DABCYL. At 100 nM of peptide–PNA conjugate in the presence of 120 nM renin, an 8-fold increase in NIR644 fluorescence was observed. Only a two-fold increase in the presence of 600 nM renin was observed from an analogous

unstructured peptide probe. Another probe used a portion of HIV protein p17 and two PNA base pairs to form the closed structure quenching C-terminal BODIPY with N-terminal tryptophan [280]. This probe was then used to quantify anti-HIV antibodies that bind to the target peptide sequence in the probe. Unlike the previous report, PNA modification decreased the affinity of this peptide for its target from $K_d \approx 200$ pM to 4 nM. However, a three-fold fluorescence enhancement and good emissivity allowed detection of anti-HIV antibodies down to 300 pM.

A similar approach was used recently to detect protein S100B, a known biomarker for brain trauma [281]. As protein expression is low in trauma victims and absent in healthy patients, a low nanomolar affinity is necessary for effective probe design. The peptide TRTK-12 served as the protein-sensitive portion of the probe as the peptide itself has a $K_d \approx 0.3$ μM for the protein. Because S100B has two peptide binding sites, the best probe design contained two TRTK peptides connected through a peptide linker which also contained two G–C PNA base pairs. Using two TRTK peptides increased the probe affinity by two orders of magnitude ($K_d \approx 3$ nM). Detection was achieved using lysine residues modified with either Alexa 488 or DABCYL in proximity with two G–C PNA base pairs to assure proximal fluorescence quenching. At 2 nM, the beacon generated 5-fold fluorescence enhancement in the presence of 80 nM S100B. Removing the PNA base pairs limited this enhancement to <1.5-fold [281].

While detection of proteins using PNA is relatively uncommon, the ease with which PNA monomers can be linked to peptides has been exploited in improving cellular penetration of PNAs for years. Using PNA base pairs to structure biologically relevant peptides therefore requires minimal adaptation of established procedures in peptide synthesis. The main strength of this approach is also its weakness, as PNA base pairs form strong interactions which help maintain the peptide in the closed state, but also hinder opening of the structure in the presence of the target protein. In spite of this, peptide beacons are useful in the detection of proteins with strong binding affinities for short target peptides. Similarly to PNA–peptide conjugates previously discussed, the combination of PNA with peptides and other biomolecules may lead to new or expanded applications of PNA both as research and diagnostic tools as well as potential therapeutics.

Preclinical studies and attempts at therapeutic development

Demonstration of antisense and antigene potential

The potential of PNA for antisense and antigene applications was recognized almost immediately after its invention. Babiss and co-workers demonstrated that 10- to 20-mer PNAs could terminate both transcription and reverse transcription in vitro [38]. Nuclear microinjection of 15-mer or 20-mer PNA targeting SV40 T antigen mRNA reduced expression of the SV40 T antigen in 40% and 50% of injected cells, respectively. Similar results were obtained by Buchardt and co-workers two years later, showing that 10-mer PNAs arrested transcription of DNA under the control of T3 and T7 promoters [282]. Antisense properties of PNAs were explored more thoroughly in 1996 [283]. Both duplex and triplex formation with PNA could arrest translation of chloramphenicol acetyltransferase (CAT) mRNA both in vitro and in cell lysate. Inhibition using duplex forming PNA was limited to sequences immediately to the 5′ of the AUG start codon while targeting the coding region had little effect. Triplex formation using either two PNAs, bis-PNAs, or tail-clamp PNAs could arrest translation by binding either at the start codon or within the coding region of the mRNA. Corey and co-workers further explored PNA's ability to inhibit translation by using 27 PNAs to target 18 different sites in a luciferase mRNA [284]. Duplex-forming PNAs targeting the terminus of the 5' UTR were found to be very effective (80%) in inhibiting translation of luciferase in COS-7 cells. Targeting other regions, including the start codon, was less effective. Sequence length also played an important role with 15- to 18-mer sequences giving <20% residual luciferase activity while a 10-mer sequence resulted in >85% residual activity.

PNA properties related to pharmacology

The high affinity and sequence specificity of PNA for natural nucleic acids [18,19] inspired multiple attempts to develop therapeutic approaches, such as antisense, antigene, and even more complex gene editing technologies [285]. PNAs form Watson–Crick duplexes with complementary DNA and RNA having significantly higher thermal stability and sequence selectivity (mismatch intolerance) than the natural DNA/DNA or DNA/RNA duplexes [3,286]. These favorable binding properties are critical for potency and selectivity of on-target pharmacological activity. Moreover, because of the entirely unnatural backbone, PNA is remarkably resistant to degradation by nucleases or proteases. In biological systems, PNA showed no significant degradation under conditions that completely cleaved various peptide substrates [287]. PNAs were stable in human and animal serums and eukaryotic cellular extracts under conditions where unmodified oligonucleotides had a half-life of only a few minutes [288]. Collectively, early studies clearly showed that PNA had impressive nucleic acid recognition potential and sufficient biostability for therapeutic applications. The most advanced examples of preclinical studies and cases were PNAs were tested in animal models in are summarized in Table 1 and discussed in more detailed below.

Table 1: Examples of advanced studies attempting therapeutic applications.

Disease or disorder	Mode of action	Target	Carrier, construct	Test model	Observed effect	Refs.
HIV	antisense	viral genome transactivation response element (TAR) in the 5′ long terminal repeat (LTR)	PNA– transportan (Figure 12)	HIV-1-infected H9 cells	inhibition of HIV-1 production	[289]
HIV	antisense	viral genome transactivation response element (TAR) in the 5′ long terminal repeat (LTR)	PNA–neamine (Figure 20B)	CEM cells infected with pseudo-typed HIV-1 S1 strain	inhibition of viral replication, cleavage of TAR RNA	[223]
hepatitis B virus (HBV)	antisense	terminal direct repeat (DR) sequence of pgRNA and mRNAs encoding HBV e antigen (HBeAg), core protein, x protein (HBx), and reverse transcriptase (RT)	PNA–Tat (Figure 12)	HepG2.2.15 cells and mouse model of acute hepatitis B	significant inhibitory effects against HBV replication in vitro and in vivo	[176]
malaria	antisense	PfSec13 mRNA essential for parasite proliferation in human erythrocytes	PNA–K_8	parasites modified with luciferase reporter gene	dose dependent inhibition of parasite proliferation	[290]
duchenne muscular dystrophy (DMD)	antisense	spliceosomal skipping of exon 23 to restore correct dystrophin gene translation	PNA PNA–Tat PNA–MSP[a] PNA–AAV6 PNA–AAV8	Mdx mouse model of muscular dystrophy	dystrophin expression in a dose-dependent manner in the injected muscle	[291,292]
DMD	antisense	spliceosomal skipping of exon 23 to restore correct dystrophin gene translation	unmodified PNA	Mdx mouse model of muscular dystrophy	dystrophin expression at high dose 50–100 mg/kg	[293]
DMD	antisense	spliceosomal skipping of exon 23 to restore correct dystrophin gene translation	PNA-BEPO[b]	Mdx mouse model of muscular dystrophy	low levels of exon skipping and dystrophin expression	[294]
DMD	antigene	exon 10 of the dystrophin gene	unmodified PNA	Mdx mouse model of muscular dystrophy	3% of gene repair was observed in the injected muscle	[295]
DMD	antigene	exon 10 of the dystrophin gene	unmodified PNA	muscle satellite stem cells form mdx mice, transplanted after PNA treatment into injured mdx mice	increase in the number of dystrophin-positive fibers detected after six months following transplantation in muscle	[296]
thalassemia	gene editing	β-thalassemia-associated splicing mutation at IVS2-654	γ-miniPEG tcPNA–K_3, donor DNA, PLGA–NPs[c]	β-globin/GFP transgenic mice	editing of the defective gene with low levels of off-target modifications	[297]
lymphoma	antigene	enhancer Eμ DNA sequence that controls c-*myc* oncogene over-expression	PNA–NLS	Burkitt's lymphoma (BL) cells and human BL lymphoma cells introduced in mice	decreased tumor size	[298,299]

Table 1: Examples of advanced studies attempting therapeutic applications. (continued)

multiple myeloma	antigene	transcription start site of RAD51 gene	PNA–NLS	SCID mice with implanted rabbit bone segments	sensitized multiple myeloma cells to melphalan treatment	[300]
tumor (lymphoma, leukemia)	antisense	miRNA-155 overexpressed in many cancers	PNA–penetratin, PLGA–NPs	nude mice with subcutaneously injected tumors	reduced tumor growth and miRNA-155 levels	[301]
tumor (lymphoma)	antisense	miRNA-155 overexpressed in many cancers	K_3–PNA–K, R_3–PNA–R, PLGA–NPs	NSG mice with injected subcutaneously tumors (U2932 lymphoma cells)	reduced tumor growth and miRNA-155 levels	[302]
tumor (HeLa cells)	antisense	miRNA-210 upregulated in response to hypoxia in various cancer cells and almost all solid tumors	γ-miniPEG PNA, PLGA–NPs	athymic nude mice with injected subcutaneously tumors (HeLa cells)	reduced tumor growth	[303]
tumor (lymphoma)	antisense	miR-155 overexpressed in many cancers	PNA–pHLIP	Tet-Off-based mouse model expressing miRNA-155 Nude mice subcutaneously implanted with neoplastic B cells	inhibition of miRNA-155 in vivo, delayed tumor growth, suppressed metastatic spread	[181]
prostate cancer	antisense	miRNA-21 frequently upregulated in solid tumors	PNA–Tat (49–57)	murine prostate cancer model with metastatic bone tumors	reduced tumor growth and metastasis	[304]
antibacterial	antisense	*acpP* gene encoding the ACP protein responsible for fatty acid biosynthesis	$(KFF)_3$K–PNA	intraperitoneal mouse model of *E. coli* infection	reduced levels of bacteria	[305]
antibacterial	antisense	*rpo*A gene encoding RNA polymerase α subunit	PNA–KFF, PNA–ANT, PNA–Tat, PNA–PXR, PNA–RFR	*C. elegans* infected with *Listeria monocytogenes*	complete bacterial clearance with PNA–RXR at 32 μM.	[306]

[a]Sequences of cell penetrating peptides: MSP – ASSLNIASSL; AAV6 – TVAVNLQSSSTDPATGDVHVM; AAV8 – IVADNLQQQNTAPQIGTVNSQ; NLS – PKKKRKV (nuclear localization signal), pHLIP – AAEQNPIYWARYADWLFTTPLLLLDLALLVDADEGT(CNPys)G; TAT – YGRKKRRQRRRP, GRKKRRQRRRPGC, RRRQRRKKR or GRKKKRRQRRRYK; KFF – KFFKFFKFFK, ANT – RQIKIWFQNRRMKWKK; RXR – RXRRXRRXRRXRXB, RFR – RFRRFRRFRRFRXB. [b]A long-acting, injectable in situ depot forming technology based on diblock and triblock poly(ethylene glycol)-polyesters solubilized in a biocompatible solvent; [c]poly(lactic-*co*-glycolic acid) nanoparticle.

Inhibition of HIV and hepatitis B virus

Transactivation response (TAR) element from 5' nontranslated region of HIV-1 viral genome together with transactivator protein are essential for the initiation of viral replication [307]. Pandey and co-workers reported that anti-TAR PNA conjugated with transportan CPP inhibited transactivation of the HIV-1 LTR, resulting in decreased production of HIV-1 virions by chronically infected H9 cells [289]. Latter studies found that the mechanism of cellular uptake of the PNA–transportan conjugate was neither receptor-dependent nor endocytosis [308]. The PNA conjugate permeated the virus envelope and inactivated HIV-1 virions in the plasma prior to their entry into cells; hence, these conjugates could be envisioned as potential prophylactic agents to block HIV-1 infection following accidental exposure to the virus. In another study, the same sequence of PNA was conjugated to neamine (Figure 20B) which improved PNA solubility and cellular uptake. The PNA–neamine conjugate also enabled cleavage of target RNA thus enhancing HIV inhibition [223].

The terminal direct repeat (DR) sequence of hepatitis B virus (HBV) pre-genomic RNA plays an important role in the synthesis of the HBV genome. As discussed above, the PNA–Tat conjugate antisense targeting DR effectively inhibited HBV replication in vitro and in vivo, with potency similar to clinically used antiretroviral drug lamivudine [176]. This study suggested that PNA–Tat has potential for treatment of HBV infections.

Malaria

The PNA–K_8 conjugate was explored as an inhibitor of malaria-causing protozoan *Plasmodium falciparum* [290,309]. To reach the target RNA in parasites at their intracellular blood stage, PNA should cross four membranes: the erythrocyte membrane, the parasitophorous vacuole, the parasite's plasma membrane, and the parasite's nuclear envelope. The PNA–K_8 antisense effect was more pronounced when the conjugate was added in the trophozoite stage and 4.8 μM of the anti-Sec13 PNA–K_8 conjugate downregulated PfSec13 expression by approximately 75% [290].

Duchenne muscular dystrophy

Duchenne muscular dystrophy (DMD) is an X-linked genetic disorder and the most common form of muscular dystrophy caused by mutations in the dystrophin gene that lead to essential shortage of the functional protein. Respiratory or cardiac failure caused by DMD usually become fatal before the end of the third decade of life. Antisense oligonucleotides have been shown to induce specific exon skipping and restore the correct reading frame and expression of functional dystrophin [294].

Wood and co-workers found that unmodified PNA and various PNA–peptide conjugates, including Tat, muscle-specific peptide (MSP), and adenoassociated virus functional domains AAV6 and AAV8 induced exon skipping and dystrophin expression in a dose-dependent manner after intramuscular injection in mdx mice [291]. Interestingly, this study observed no significant difference in potency of unmodified PNA and its peptide conjugates, which was attributed to peptides selected in this study not being sufficiently efficient in transfecting specifically in muscle and escaping endosomal entrapment. Examination of morphology of muscle cells treated with unmodified PNA or PNA–peptide conjugates by hematoxylin and eosin staining did not show local muscle toxicity [291].

Yin and co-workers explored the potential of PNA (various lengths from 20- to 30-mer) to induce exon skipping and expression of dystrophin by systemically administering PNA to mdx mice through weekly intravenous injections of 50–100 mg/kg, over the course of three weeks [293]. Enlarged number of dystrophin-positive fibers was observed in several tissues: abdominal muscle, gastrocnemius, and quadriceps, but not in the heart. The longest PNA (30-mer) caused more significant increase of dystrophin expression in tibialis anterior muscles than other shorter versions. However, an acidity-related toxicity was observed for PNA 30-mer, which may be related to difficulties in purification and solubilization of longer PNAs [293].

Brolin and Nielsen investigated the effect of in situ forming depot technology (BEPO, PEG-PLA biodegradable polymer) and PNA-oligonucleotide formulation in systemic administration of a 20-mer splice switching antisense PNA through intravenous and subcutaneous routes in the mdx mice [294]. Intravenous administration resulted in fast renal/bladder excretion of the PNA (half-life ≈20 min) while subcutaneous administration led to a 2–3 times slower excretion. However, due to biphasic kinetics, release of 50% of the PNA dose from BEPO–PNA formulation takes approximately 10 days. Overall, the PNA–BEPO administration did not significantly improve antisense activity [294].

Brolin and Nielsen observed lower dystrophin expression than that reported by Yin and co-workers [293]. Interestingly, Brolin and Nielsen observed PNA precipitation when the pH of PNA administration solution was adjusted above 4 at 1 mM concentration (required for dosing at 50 mg/kg), suggesting that acidity required to solubilize longer PNAs may have caused toxicity observed by Yin and co-workers [293]. Collectively, the PNA antisense agents targeting muscles, even in the case of compromised muscle fibers in muscular dystrophy, still need major improvements to become therapeutically relevant, regardless of the administration route and long-acting depot or heteroduplex formulation [294].

Thalassemia

Thalassemia is an inherited blood illness characterized by decreased hemoglobin production. As a monogenic disorder, β-thalassemia has been a focus of gene therapy efforts. Most notably, significant progress in gene editing of hematopoietic stem cells has been reported by Glazer's team in collaboration with other groups [297]. Glazer and co-workers have been studying triplex-forming bis-PNAs as gene mutagenesis and editing tools for more than two decades [285,310,311]. Recently, addition of γ-miniPEG modification (Figure 5) to tail-clamp PNAs (tcPNA, Figure 2) increased the gene editing frequency of up to 6.9% in a thalassemic mouse model [297]. The gene editing construct included γ-miniPEG modified-tcPNA, conjugated with three lysines at each termini, and donor DNA, formulated in poly(lactic-*co*-glycolic acid) nanoparticles (PLGA-NPs), and was used together with stimulation of the stem cell factor (SCF)/c-Kit pathway. The use of γ-miniPEG modified-tcPNA gave almost double gene editing than unmodified tcPNA, presumably due to enhanced strand invasion and DNA binding because of the helical pre-organization enforced by the γ-miniPEG modification [89].

PLGA-NPs were previously used for systemic delivery of FDA-approved drugs and effectively delivered PNA/donor DNA combinations into primary human and mouse hematopoietic

cells with essentially no toxicity [301,312,313]. For in vivo studies, PNAs and donor DNAs, at a molar ratio of 2:1, were incorporated into PLGA-NPs and administrated by intravenous injection while SCF was administrated intraperitoneally 3 h before PLGA-NP injections. Importantly, the overall off-target modification frequency in γ-miniPEG modified-tcPNA treated thalassemic mice was 0.0032%, which was 1,218-fold lower than the frequency of β-globin gene editing [297]. In addition, minimal immune or inflammatory responses were observed in this study according to cytokine array analyses. The combination of nanoparticle delivery, γ-miniPEG modified-tcPNA, and SCF treatment can be basis for a minimally invasive cure for genetic disorders that can be achieved simlpy and safely by intravenous and intraperitoneal administration [297]. About 4% frequency of gene editing in total bone marrow cells achieved in the thalassemic mice was adequate to achieve a clear improvement in phenotype. Higher editing frequencies have been achieved in cell culture carrying the same thalassemia-associated β-globin mutation using TALENS (33%) and CRISPR/Cas9 (12–16%) [314,315]. However, direct comparison of PNA with TALENS or CRISPR/Cas9 is not possible because the studies used different cell lines and data analysis methodologies.

Anticancer PNAs

PNAs have been explored as antigene and antisense agents against various types of cancer (Table 1). Boffa and co-workers reported that a PNA–NLS conjugate (18-mer) complementary to intronic Eμ enhancer DNA sequences, inhibited the expression of the c-myc oncogene under the Eμ enhancer control in Burkitt's lymphoma (BL) cells and human BL lymphoma cells introduced in mice [298,299]. After injection in mice, PNA reached the maximum concentration in the tumor in 90 minutes, with less accumulation in kidney, liver, spleen, heart, and brain. PNA was present in tumors for at least 600 minutes at a concentration that effectively inhibited BL cell growth in culture [299]. Short-term or long-term toxic effects were not observed. The tumor volume started to plateau after eight injections of PNA–NLS. Necrosis (about 8% of the neoplastic cell) was observed in histology of the tumor of PNA-treated mice [299]. Reis and co-workers reported that PNA–NLS targeting the transcription start site of RAD51, protein that mediated recombinational DNA repair and is overexpressed in multiple myeloma, sensitized multiple myeloma cells to melphalan treatment [300]. Melphalan is chemotherapy medication used to treat multiple myeloma, ovarian cancer, melanoma, and AL amyloidosis.

MiRNAs have been a well-established target for antisense anticancer approaches [316,317]. Fabani, Vigorito, and co-workers reported that antisense PNA conjugated with three lysines ($K–PNA–K_3$) completely abolished the expression of miRNA-155 induced by intraperitoneal lipopolysaccharide (LPS) injection after dosed systemically at 50 mg/kg for two days [190]. Slack and co-workers reported that antisense PNA decreased miRNA-155 expression and tumor growth when injected as PLGA–NP formulations in nude mice carrying tumor cells from NesCre8 [301]. The PLGA–NPs were modified by penetratin that is attached to the NP surface via a PEG linker. A single local intratumor injection of PNA–PLGA-NP at 1.5 mg/kg reduced tumor increase from 10-fold to 2-fold, while two intravenous injections (1.5 mg/kg) reduced tumor increase by ≈50% relative to control tumors. These decreases in tumor growth correlated with a decreased number of miRNA-155 per tumor cell.

Bahal and co-workers studied short PNAs (8-mers) targeting the seed region of miRNA-155 in NSG mice carrying tumors induced by subcutaneous injection of U2932 lymphoma cells [302]. PNA conjugates with lysine and arginine, $K_3–PNA–K$ and $R_3–PNA–R$, were formulated with PLGA–NPs and delivered by tail vein injection. The PNA 8-mer showed similar and even better efficacy in reducing the tumor growth compared to full length PNA 23-mer; PNAs without additional amino acids did not bind to miRNA-155 and arginine conjugates were slightly better than lysine conjugates. The authors did not observe any signs of immune response or toxicity in histology of liver, kidney and other organs [302].

Glazer and co-workers showed that γ-miniPEG modified-PNA antisense to miRNA-210, an oncogenic miRNA that helps tumor cells to survive and proliferate under hypoxic conditions, significantly delayed growth of a human tumor xenograft when administered by intratumoral injection in mice using PLGA–NPs [303]. The γ-miniPEG modified-PNA was significantly more active than unmodified PNAs. However, intravenous administration of the PLGA–NPs/PNA was not effective in preventing the tumor growth. In another study, Slack and co-workers showed that antisense PNA–pHLIP conjugate (Figure 13) showed significant survival advantages in nude mice subcutaneously implanted with neoplastic B cells compared with a commercially available locked nucleic acid antimiR optimized for in vivo miRNA-155 silencing [181]. The PNA–pHLIP conjugate delayed tumor growth and suppressed the metastatic spread of neoplastic lymphocytes to other organs, without causing toxicity in healthy mice.

Youn and co-workers compared antisense PNA and locked nucleic acids (LNA) conjugated with a shorter version (amino acids 49–57) of the Tat peptide and targeting miRNA-21 murine prostate cancer model [304]. This study found that PNA

conjugates showed better stability and efficacy than LNA conjugates with 86% and 25% reduction in the tumor volume, respectively, after intravenous injection at 200 nM in the mouse model of metastatic bone tumors.

Antibacterial PNAs

Antisense PNAs have been extensively studied as potential antibacterial agents. The scope and limitation of these studies have been recently reviewed [318], therefore only examples where in vivo data were presented are included in Table 1. Cell-penetrating peptides (CPP) are the most commonly used ligands for delivery of PNA to bacteria; however, this approach is not universally applicable because the CPP-mediated transport across bacterial cell membrane may be dependent on the specific strain of bacteria. Tan and co-workers showed that PNA conjugated with the $(KFF)_3K$ peptide inhibited bacterial growth in vivo in BALB/c mice infected with SM101 or K12 strains of *E. coli* [305]. The antisense PNAs targeting the acpP gene that encodes protein ACP responsible to fatty acid biosynthesis were more effective against the SM101 strain, which has a defective outer membrane and hence is easier penetrated by PNA conjugates.

Seleem and co-workers compared conjugation to five different CPPs for delivery of PNA antisense to *rpo*A gene encoding RNA polymerase α subunit, which also causes suppressive effects on other essential bacterial genes and virulence factors [306]. In murine macrophage cells infected with *Listeria monocytogenes*, $(RXR)_4XB$–PNA conjugate was the most effective, with significant reduction at 2 μM and complete clearance of intracellular *Listeria* at 8 μM. Tat–PNA and $(RFR)_4XB$–PNA conjugates also showed significant activity at 2 μM. In *C. elegans* infected with *L. monocytogenes*, the $(RXR)_4XB$–PNA conjugate achieved complete bacterial clearance at 32 μM. Collectively, the in vitro and in vivo results suggested that $(RXR)_4XB$ followed by Tat and $(RFR)_4XB$ were the best CPPs for delivery of the anti-*rpo*A PNA to cells infected with *L. monocytogene* [306].

While PNA–CPP conjugates have shown promising anticancer and antibacterial activity in cell cultures and in vivo, they are not without drawbacks and vulnerabilities. Activity of PNA–CPP conjugates can drastically decrease in the presence of blood serum [319] and typically require excessively high (10 to 50 mg/kg) and repeated dosing to achieve therapeutic effect in vivo [16,157]. In addition, some CPPs are larger than their PNA cargo, increasing the complexity of the therapeutic system. Despite extensive studies, primary literature lacks reports on comprehensive and conclusive studies on long-term toxicity and possible innate and adaptive immune responses [320]. In summary, while many attempts at therapeutic development have given promising preliminary results, PNAs have still not entered clinical trials [14].

Conclusion

The year 2021 marks the 30th anniversary of the original PNA publication [1]. The remarkable biophysical properties of the first neutral DNA mimic, especially the high binding affinity and sequence selectivity for complementary native nucleic acids, were recognized immediately. However, the limitations imposed by poor solubility and inefficient crossing of cellular membranes quickly became obvious. Over these 30 years, extensive research focused on either direct chemical modification or conjugation of PNA with various ligands to address the limitations and improve the biophysical and biological properties of PNA. The present review covers only selected examples of an enormous body of these studies, but aims to present a comprehensive picture of the versatility of PNA.

It is fascinating to think that, while many chemical modifications of both backbone and nucleobases have been reported, relatively few provide significant improvements on the original design. Among the backbone modifications, pre-organizing of PNA in a right-handed helix favorable for DNA binding either by cyclopentane or γ-substituents has shown the most promise. Work towards the development of nucleobase modifications continues to address the limitations of triple helical recognition of dsDNA and dsRNA. Cellular uptake remains an unsolved problem, and both backbone and nucleobase modifications may deliver future advances. In this context, the 2-aminopyridine (M) nucleobase has afforded interesting preliminary results by enhancing both molecular recognition of dsRNA and cellular uptake of triplex-forming PNAs. In applications where solubility and cell permeability are not the limiting factors, such as PCR or FISH, PNA is widely used due to its exceptional binding strength and specificity.

Cell-penetrating peptides have been extensively explored as delivery-enhancing ligands. While many of the conjugates have shown promising in vitro and even in vivo activity, PNA-based therapeutic candidates have not yet entered clinical trials. It appears that the key remaining bottleneck is the necessity for high doses of PNA conjugates, to overcome the problem of endosomal entrapment, and associated toxicity. In other words, the chemical modifications that have succeeded in addressing the problems of cellular uptake, biodistribution, and tissue delivery of PNA have also increased the toxicity of the conjugates beyond acceptable therapeutic windows. Nevertheless, both academic and industrial research groups continue creative research into new chemical modifications and PNA–ligand conjugations. The optimism remains high, that with the right

combination of innovative chemistry and biology, the full potential of PNA in biomedical applications will be discovered in the near future.

Acknowledgements

The authors thank James A. MacKay and Emily Harding for critical reading of the manuscript.

Funding

The PNA research in Rozners lab has been supported by National Institutes of Health (R35 GM130207) and National Science Foundation (CHE-1708761).

ORCID® iDs

Martins Katkevics - https://orcid.org/0000-0003-1547-0075
Venubabu Kotikam - https://orcid.org/0000-0001-6670-1713
Eriks Rozners - https://orcid.org/0000-0001-7649-0040

References

1. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991,** *254,* 1497–1500. doi:10.1126/science.1962210
2. Buchardt, O.; Egholm, M.; Berg, R. H.; Nielsen, P. E. *Trends Biotechnol.* **1993,** *11,* 384–386. doi:10.1016/0167-7799(93)90097-s
3. Nielsen, P. E.; Egholm, M.; Buchardt, O. *Bioconjugate Chem.* **1994,** *5,* 3–7. doi:10.1021/bc00025a001
4. Saarbach, J.; Sabale, P. M.; Winssinger, N. *Curr. Opin. Chem. Biol.* **2019,** *52,* 112–124. doi:10.1016/j.cbpa.2019.06.006
5. Corradini, R.; Sforza, S.; Tedeschi, T.; Totsingan, F.; Marchelli, R. *Curr. Top. Med. Chem.* **2007,** *7,* 681–694. doi:10.2174/156802607780487759
6. Gupta, A.; Mishra, A.; Puri, N. *J. Biotechnol.* **2017,** *259,* 148–159. doi:10.1016/j.jbiotec.2017.07.026
7. Fouz, M. F.; Appella, D. H. *Molecules* **2020,** *25,* 786. doi:10.3390/molecules25040786
8. Frickmann, H.; Zautner, A. E.; Moter, A.; Kikhney, J.; Hagen, R. M.; Stender, H.; Poppert, S. *Crit. Rev. Microbiol.* **2017,** *43,* 263–293. doi:10.3109/1040841x.2016.1169990
9. Cadoni, E.; Manicardi, A.; Madder, A. *Molecules* **2020,** *25,* 1296. doi:10.3390/molecules25061296
10. Shen, X.; Corey, D. R. *Nucleic Acids Res.* **2018,** *46,* 1584–1600. doi:10.1093/nar/gkx1239
11. Setten, R. L.; Rossi, J. J.; Han, S.-p. *Nat. Rev. Drug Discovery* **2019,** *18,* 421–446. doi:10.1038/s41573-019-0017-4
12. Noble, S. A.; Bonham, M. A.; Bisi, J. E.; Bruckenstein, D. A.; Brown, P. H.; Brown, S. C.; Cadilla, R.; Gaul, M. D.; Hanvey, J. C.; Fred Hassman, C.; Josey, J. A.; Luzzio, M. J.; Myers, P. M.; Pipe, A. J.; Ricca, D. J.; Su, C. W.; Stevenson, C. L.; Thomson, S. A.; Wiethe, R. W.; Babiss, L. E. *Drug Dev. Res.* **1995,** *34,* 184–195. doi:10.1002/ddr.430340208
13. Lundin, K. E.; Good, L.; Strömberg, R.; Gräslund, A.; Smith, C. I. E. Biological Activity and Biotechnological Aspects of Peptide Nucleic Acid. *Advances in Genetics;* Academic Press, 2006; Vol. 56, pp 1–51. doi:10.1016/s0065-2660(06)56001-8
14. Roberts, T. C.; Langer, R.; Wood, M. J. A. *Nat. Rev. Drug Discovery* **2020,** *19,* 673–694. doi:10.1038/s41573-020-0075-7
15. Nielsen, P. E. *Curr. Opin. Mol. Ther.* **2010,** *12,* 184–191.
16. Nielsen, P. E. *Q. Rev. Biophys.* **2005,** *38,* 345–350. doi:10.1017/s0033583506004148
17. McMahon, B. M.; Mays, D.; Lipsky, J.; Stewart, J. A.; Fauq, A.; Richelson, E. *Antisense Nucleic Acid Drug Dev.* **2002,** *12,* 65–70. doi:10.1089/108729002760070803
18. Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992,** *114,* 1895–1897. doi:10.1021/ja00031a062
19. Egholm, M.; Nielsen, P. E.; Buchardt, O.; Berg, R. H. *J. Am. Chem. Soc.* **1992,** *114,* 9677–9678. doi:10.1021/ja00050a068
20. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature* **1993,** *365,* 566–568. doi:10.1038/365566a0
21. Wittung, P.; Nielsen, P. E.; Buchardt, O.; Egholm, M.; Nordén, B. *Nature* **1994,** *368,* 561–563. doi:10.1038/368561a0
22. Tomac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P. E.; Nordén, B.; Gräslund, A. *J. Am. Chem. Soc.* **1996,** *118,* 5544–5552. doi:10.1021/ja960495l
23. Cherny, D. Y.; Belotserkovskii, B. P.; Frank-Kamenetskii, M. D.; Egholm, M.; Buchardt, O.; Berg, R. H.; Nielsen, P. E. *Proc. Natl. Acad. Sci. U. S. A.* **1993,** *90,* 1667–1670. doi:10.1073/pnas.90.5.1667
24. Bentin, T.; Hansen, G. I.; Nielsen, P. E. *Nucleic Acids Res.* **2006,** *34,* 5790–5799. doi:10.1093/nar/gkl736
25. Wittung, P.; Nielsen, P.; Nordén, B. *Biochemistry* **1997,** *36,* 7973–7979. doi:10.1021/bi963136b
26. Hansen, M. E.; Bentin, T.; Nielsen, P. E. *Nucleic Acids Res.* **2009,** *37,* 4498–4507. doi:10.1093/nar/gkp437
27. Li, M.; Zengeya, T.; Rozners, E. *J. Am. Chem. Soc.* **2010,** *132,* 8676–8681. doi:10.1021/ja101384k
28. Zengeya, T.; Gupta, P.; Rozners, E. *Angew. Chem., Int. Ed.* **2012,** *51,* 12593–12596. doi:10.1002/anie.201207925
29. Muse, O.; Zengeya, T.; Mwaura, J.; Hnedzko, D.; McGee, D. W.; Grewer, C. T.; Rozners, E. *ACS Chem. Biol.* **2013,** *8,* 1683–1686. doi:10.1021/cb400144x
30. Hnedzko, D.; McGee, D. W.; Karamitas, Y. A.; Rozners, E. *RNA* **2017,** *23,* 58–69. doi:10.1261/rna.058362.116
31. Ryan, C. A.; Brodyagin, N.; Lok, J.; Rozners, E. *Biochemistry* **2021,** *60,* 1919–1925. doi:10.1021/acs.biochem.1c00275
32. Doluca, O.; Withers, J. M.; Filichev, V. V. *Chem. Rev.* **2013,** *113,* 3044–3083. doi:10.1021/cr300225q
33. Krishnan-Ghosh, Y.; Stephens, E.; Balasubramanian, S. *J. Am. Chem. Soc.* **2004,** *126,* 5944–5945. doi:10.1021/ja031508f
34. Egholm, M.; Christensen, L.; Deuholm, K. L.; Buchardt, O.; Coull, J.; Nielsen, P. E. *Nucleic Acids Res.* **1995,** *23,* 217–222. doi:10.1093/nar/23.2.217
35. Griffith, M. C.; Risen, L. M.; Greig, M. J.; Lesnik, E. A.; Sprankle, K. G.; Griffey, R. H.; Kiely, J. S.; Freier, S. M. *J. Am. Chem. Soc.* **1995,** *117,* 831–832. doi:10.1021/ja00107a033
36. Kuhn, H.; Demidov, V. V.; Frank-Kamenetskii, M. D.; Nielsen, P. E. *Nucleic Acids Res.* **1998,** *26,* 582–587. doi:10.1093/nar/26.2.582
37. Bentin, T.; Larsen, H. J.; Nielsen, P. E. *Biochemistry* **2003,** *42,* 13987–13995. doi:10.1021/bi0351918
38. Hanvey, J. C.; Peffer, N. J.; Bisi, J. E.; Thomson, S. A.; Cadilla, R.; Josey, J. A.; Ricca, D. J.; Hassman, C. F.; Bonham, M. A.; Au, K. G. *Science* **1992,** *258,* 1481–1485. doi:10.1126/science.1279811
39. Eriksson, M.; Nielsen, P. E. *Nat. Struct. Mol. Biol.* **1996,** *3,* 410–413. doi:10.1038/nsb0596-410
40. Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. *Science* **1994,** *265,* 777–780. doi:10.1126/science.7519361

41. Betts, L.; Josey, J. A.; Veal, J. M.; Jordan, S. R. *Science* **1995,** *270,* 1838–1841. doi:10.1126/science.270.5243.1838
42. Kotikam, V.; Kennedy, S. D.; MacKay, J. A.; Rozners, E. *Chem. – Eur. J.* **2019,** *25,* 4367–4372. doi:10.1002/chem.201806293
43. Rasmussen, H.; Kastrup, J. S.; Nielsen, J. N.; Nielsen, J. M.; Nielsen, P. E. *Nat. Struct. Mol. Biol.* **1997,** *4,* 98–101. doi:10.1038/nsb0297-98
44. Petersson, B.; Nielsen, B. B.; Rasmussen, H.; Larsen, I. K.; Gajhede, M.; Nielsen, P. E.; Kastrup, J. S. *J. Am. Chem. Soc.* **2005,** *127,* 1424–1430. doi:10.1021/ja0458726
45. He, W.; Hatcher, E.; Balaeff, A.; Beratan, D. N.; Gil, R. R.; Madrid, M.; Achim, C. *J. Am. Chem. Soc.* **2008,** *130,* 13264–13273. doi:10.1021/ja800652h
46. Hyrup, B.; Egholm, M.; Rolland, M.; Nielsen, P. E.; Berg, R. H.; Buchardt, O. *J. Chem. Soc., Chem. Commun.* **1993,** 518–519. doi:10.1039/c39930000518
47. Hyrup, B.; Egholm, M.; Nielsen, P. E.; Wittung, P.; Norden, B.; Buchardt, O. *J. Am. Chem. Soc.* **1994,** *116,* 7964–7970. doi:10.1021/ja00097a002
48. Kumar, V.; Brodyagin, N.; Rozners, E. *ChemBioChem* **2020,** *21,* 3410–3416. doi:10.1002/cbic.202000432
49. Hyrup, B.; Egholm, M.; Buchardt, O.; Nielsen, P. E. *Bioorg. Med. Chem. Lett.* **1996,** *6,* 1083–1088. doi:10.1016/0960-894x(96)00176-x
50. Lagriffoule, P.; Eriksson, M.; Jensen, K. K.; Nielsen, P. E.; Wittung, P.; Nordén, B.; Buchardt, O. *Chem. – Eur. J.* **1997,** *3,* 912–919. doi:10.1002/chem.19970030613
51. Govindaraju, T.; Gonnade, R. G.; Bhadbhade, M. M.; Kumar, V. A.; Ganesh, K. N. *Org. Lett.* **2003,** *5,* 3013–3016. doi:10.1021/ol034933m
52. Govindaraju, T.; Kumar, V. A.; Ganesh, K. N. *J. Org. Chem.* **2004,** *69,* 1858–1865. doi:10.1021/jo035747x
53. Govindaraju, T.; Kumar, V. A.; Ganesh, K. N. *J. Am. Chem. Soc.* **2005,** *127,* 4144–4145. doi:10.1021/ja044142v
54. Kumar, V. A.; Ganesh, K. N. *Acc. Chem. Res.* **2005,** *38,* 404–412. doi:10.1021/ar030277e
55. Myers, M. C.; Witschi, M. A.; Larionova, N. V.; Franck, J. M.; Haynes, R. D.; Hara, T.; Grajkowski, A.; Appella, D. H. *Org. Lett.* **2003,** *5,* 2695–2698. doi:10.1021/ol0348811
56. Pokorski, J. K.; Witschi, M. A.; Purnell, B. L.; Appella, D. H. *J. Am. Chem. Soc.* **2004,** *126,* 15067–15073. doi:10.1021/ja046280q
57. Govindaraju, T.; Kumar, V. A.; Ganesh, K. N. *Chem. Commun.* **2004,** 860–861. doi:10.1039/b317000d
58. Govindaraju, T.; Kumar, V. A.; Ganesh, K. N. *J. Org. Chem.* **2004,** *69,* 5725–5734. doi:10.1021/jo049442+
59. Zheng, H.; Saha, M.; Appella, D. H. *Org. Lett.* **2018,** *20,* 7637–7640. doi:10.1021/acs.orglett.8b03374
60. Zheng, H.; Botos, I.; Clausse, V.; Nikolayevskiy, H.; Rastede, E. E.; Fouz, M. F.; Mazur, S. J.; Appella, D. H. *Nucleic Acids Res.* **2021,** *49,* 713–725. doi:10.1093/nar/gkaa1249
61. Vilaivan, T.; Suparpprom, C.; Harnyuttanakorn, P.; Lowe, G. *Tetrahedron Lett.* **2001,** *42,* 5533–5536. doi:10.1016/s0040-4039(01)01020-6
62. Mansawat, W.; Vilaivan, C.; Balázs, Á.; Aitken, D. J.; Vilaivan, T. *Org. Lett.* **2012,** *14,* 1440–1443. doi:10.1021/ol300190u
63. Vilaivan, T.; Lowe, G. *J. Am. Chem. Soc.* **2002,** *124,* 9326–9327. doi:10.1021/ja012543u
64. Vilaivan, T.; Srisuwannaket, C. *Org. Lett.* **2006,** *8,* 1897–1900. doi:10.1021/ol060448q
65. Vilaivan, T. *Acc. Chem. Res.* **2015,** *48,* 1645–1656. doi:10.1021/acs.accounts.5b00080
66. Muangkaew, P.; Vilaivan, T. *Bioorg. Med. Chem. Lett.* **2020,** *30,* 127064. doi:10.1016/j.bmcl.2020.127064
67. Nielsen, P. E.; Haaima, G.; Lohse, A.; Buchardt, O. *Angew. Chem., Int. Ed. Engl.* **1996,** *35,* 1939–1942. doi:10.1002/anie.199619391
68. Püschl, A.; Sforza, S.; Haaima, G.; Dahl, O.; Nielsen, P. E. *Tetrahedron Lett.* **1998,** *39,* 4707–4710. doi:10.1016/s0040-4039(98)00862-4
69. Sforza, S.; Haaima, G.; Marchelli, R.; Nielsen, P. E. *Eur. J. Org. Chem.* **1999,** 197–204. doi:10.1002/(sici)1099-0690(199901)1999:1<197::aid-ejoc197>3.0.co;2-n
70. Menchise, V.; De Simone, G.; Tedeschi, T.; Corradini, R.; Sforza, S.; Marchelli, R.; Capasso, D.; Saviano, M.; Pedone, C. *Proc. Natl. Acad. Sci. U. S. A.* **2003,** *100,* 12021–12026. doi:10.1073/pnas.2034746100
71. Zhou, P.; Wang, M.; Du, L.; Fisher, G. W.; Waggoner, A.; Ly, D. H. *J. Am. Chem. Soc.* **2003,** *125,* 6878–6879. doi:10.1021/ja029665m
72. Dragulescu-Andrasi, A.; Zhou, P.; He, G.; Ly, D. H. *Chem. Commun.* **2005,** 244–246. doi:10.1039/b412522c
73. Dragulescu-Andrasi, A.; Rapireddy, S.; He, G.; Bhattacharya, B.; Hyldig-Nielsen, J. J.; Zon, G.; Ly, D. H. *J. Am. Chem. Soc.* **2006,** *128,* 16104–16112. doi:10.1021/ja063383v
74. Gupta, P.; Muse, O.; Rozners, E. *Biochemistry* **2012,** *51,* 63–73. doi:10.1021/bi201570a
75. Dragulescu-Andrasi, A.; Rapireddy, S.; Frezza, B. M.; Gayathri, C.; Gil, R. R.; Ly, D. H. *J. Am. Chem. Soc.* **2006,** *128,* 10258–10267. doi:10.1021/ja0625576
76. He, W.; Crawford, M. J.; Rapireddy, S.; Madrid, M.; Gil, R. R.; Ly, D. H.; Achim, C. *Mol. BioSyst.* **2010,** *6,* 1619–1629. doi:10.1039/c002254c
77. Yeh, J. I.; Shivachev, B.; Rapireddy, S.; Crawford, M. J.; Gil, R. R.; Du, S.; Madrid, M.; Ly, D. H. *J. Am. Chem. Soc.* **2010,** *132,* 10717–10727. doi:10.1021/ja907225d
78. Tähtinen, V.; Granqvist, L.; Murtola, M.; Strömberg, R.; Virta, P. *Chem. – Eur. J.* **2017,** *23,* 7113–7124. doi:10.1002/chem.201700601
79. Englund, E. A.; Appella, D. H. *Angew. Chem., Int. Ed.* **2007,** *46,* 1414–1418. doi:10.1002/anie.200603483
80. Englund, E. A.; Appella, D. H. *Org. Lett.* **2005,** *7,* 3465–3467. doi:10.1021/ol051143z
81. Sahu, B.; Chenna, V.; Lathrop, K. L.; Thomas, S. M.; Zon, G.; Livak, K. J.; Ly, D. H. *J. Org. Chem.* **2009,** *74,* 1509–1516. doi:10.1021/jo802211n
82. Manicardi, A.; Fabbri, E.; Tedeschi, T.; Sforza, S.; Bianchi, N.; Brognara, E.; Gambari, R.; Marchelli, R.; Corradini, R. *ChemBioChem* **2012,** *13,* 1327–1337. doi:10.1002/cbic.201100745
83. Mitra, R.; Ganesh, K. N. *Chem. Commun.* **2011,** *47,* 1198–1200. doi:10.1039/c0cc03988h
84. Mitra, R.; Ganesh, K. N. *J. Org. Chem.* **2012,** *77,* 5696–5704. doi:10.1021/jo300860f
85. Kumar, P.; Jain, D. R. *Tetrahedron* **2015,** *71,* 3378–3384. doi:10.1016/j.tet.2015.03.093
86. Tähtinen, V.; Verhassel, A.; Tuomela, J.; Virta, P. *ChemBioChem* **2019,** *20,* 3041–3051. doi:10.1002/cbic.201900393
87. Bhingardeve, P.; Madhanagopal, B. R.; Ganesh, K. N. *J. Org. Chem.* **2020,** *85,* 13680–13693. doi:10.1021/acs.joc.0c01853
88. Gupta, M. K.; Madhanagopal, B. R.; Ganesh, K. N. *J. Org. Chem.* **2021,** *86,* 414–428. doi:10.1021/acs.joc.0c02158
89. Bahal, R.; Sahu, B.; Rapireddy, S.; Lee, C.-M.; Ly, D. H. *ChemBioChem* **2012,** *13,* 56–60. doi:10.1002/cbic.201100646

90. Efimov, V. A.; Choob, M. V.; Buryakova, A. A.; Kalinkina, A. L.; Chakhmakhcheva, O. G. *Nucleic Acids Res.* **1998,** *26,* 566–575. doi:10.1093/nar/26.2.566
91. Shiraishi, T.; Hamzavi, R.; Nielsen, P. E. *Nucleic Acids Res.* **2008,** *36,* 4424–4432. doi:10.1093/nar/gkn401
92. Avitabile, C.; Moggio, L.; Malgieri, G.; Capasso, D.; Di Gaetano, S.; Saviano, M.; Pedone, C.; Romanelli, A. *PLoS One* **2012,** *7,* e35774. doi:10.1371/journal.pone.0035774
93. Kirillova, Y.; Boyarskaya, N.; Dezhenkov, A.; Tankevich, M.; Prokhorov, I.; Varizhuk, A.; Eremin, S.; Esipov, D.; Smirnov, I.; Pozmogova, G. *PLoS One* **2015,** *10,* e0140468. doi:10.1371/journal.pone.0140468
94. Fox, K. R.; Brown, T.; Rusling, D. A. DNA Recognition by Parallel Triplex Formation. In *Chemical Biology No. 7: DNA-targeting Molecules as Therapeutic Agents;* Waring, M. J., Ed.; Royal Society of Chemistry: Cambridge. UK, 2018; Vol. 7, pp 1–32. doi:10.1039/9781788012928-00001
95. Ono, A.; Ts'o, P. O. P.; Kan, L. S. *J. Am. Chem. Soc.* **1991,** *113,* 4032–4033. doi:10.1021/ja00010a077
96. Ono, A.; Ts'o, P. O. P.; Kan, L. S. *J. Org. Chem.* **1992,** *57,* 3225–3230. doi:10.1021/jo00037a048
97. Christensen, C.; Eldrup, A. B.; Haaima, G.; Nielsen, P. E. *Bioorg. Med. Chem. Lett.* **2002,** *12,* 3121–3124. doi:10.1016/s0960-894x(02)00658-3
98. Devi, G.; Yuan, Z.; Lu, Y.; Zhao, Y.; Chen, G. *Nucleic Acids Res.* **2014,** *42,* 4008–4018. doi:10.1093/nar/gkt1367
99. Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989,** *111,* 3059–3061. doi:10.1021/ja00190a047
100. Hildbrand, S.; Leumann, C. *Angew. Chem., Int. Ed. Engl.* **1996,** *35,* 1968–1970. doi:10.1002/anie.199619681
101. Bates, P. J.; Laughton, C. A.; Jenkins, T. C.; Capaldi, D. C.; Roselt, P. D.; Reese, C. B.; Neidle, S. *Nucleic Acids Res.* **1996,** *24,* 4176–4184. doi:10.1093/nar/24.21.4176
102. Cassidy, S. A.; Slickers, P.; Trent, J. O.; Capaldi, D. C.; Roselt, P. D.; Reese, C. B.; Neidle, S.; Fox, K. R. *Nucleic Acids Res.* **1997,** *25,* 4891–4898. doi:10.1093/nar/25.24.4891
103. Krishna, M. S.; Wang, Z.; Zheng, L.; Bowry, J.; Ong, A. A. L.; Mu, Y.; Prabakaran, M.; Chen, G. *Biochemistry* **2019,** *58,* 3777–3788. doi:10.1021/acs.biochem.9b00608
104. Eldrup, A. B.; Nielsen, B. B.; Haaima, G.; Rasmussen, H.; Kastrup, J. S.; Christensen, C.; Nielsen, P. E. *Eur. J. Org. Chem.* **2001,** 1781–1790. doi:10.1002/1099-0690(200105)2001:9<1781::aid-ejoc1781>3.0.co;2-k
105. Eldrup, A. B.; Christensen, C.; Haaima, G.; Nielsen, P. E. *J. Am. Chem. Soc.* **2002,** *124,* 3254–3262. doi:10.1021/ja0117027
106. Patil, K. M.; Toh, D.-F. K.; Yuan, Z.; Meng, Z.; Shu, Z.; Zhang, H.; Ong, A. A. L.; Krishna, M. S.; Lu, L.; Lu, Y.; Chen, G. *Nucleic Acids Res.* **2018,** *46,* 7506–7521. doi:10.1093/nar/gky631
107. Ong, A. A. L.; Toh, D.-F. K.; Krishna, M. S.; Patil, K. M.; Okamura, K.; Chen, G. *Biochemistry* **2019,** *58,* 3444–3453. doi:10.1021/acs.biochem.9b00521
108. Brodyagin, N.; Maryniak, A. L.; Kumpina, I.; Talbott, J. M.; Katkevics, M.; Rozners, E.; MacKay, J. A. *Chem. – Eur. J.* **2021,** *27,* 4332–4335. doi:10.1002/chem.202005401
109. Fox, K. R.; Brown, T. *Q. Rev. Biophys.* **2005,** *38,* 311–320. doi:10.1017/s0033583506004197
110. Eldrup, A. B.; Dahl, O.; Nielsen, P. E. *J. Am. Chem. Soc.* **1997,** *119,* 11116–11117. doi:10.1021/ja9717424
111. Ong, A. A. L.; Toh, D.-F. K.; Patil, K. M.; Meng, Z.; Yuan, Z.; Krishna, M. S.; Devi, G.; Haruehanroengra, P.; Lu, Y.; Xia, K.; Okamura, K.; Sheng, J.; Chen, G. *Biochemistry* **2019,** *58,* 1319–1331. doi:10.1021/acs.biochem.8b01313
112. Gupta, P.; Zengeya, T.; Rozners, E. *Chem. Commun.* **2011,** *47,* 11125–11127. doi:10.1039/c1cc14706d
113. Endoh, T.; Brodyagin, N.; Hnedzko, D.; Sugimoto, N.; Rozners, E. *ACS Chem. Biol.* **2021,** *16*. doi:10.1021/acschembio.1c00133
114. Guianvarc'h, D.; Benhida, R.; Fourrey, J.-L.; Maurisse, R.; Sun, J.-S. *Chem. Commun.* **2001,** 1814–1815. doi:10.1039/b103743a
115. Rusling, D. A.; Powers, V. E. C.; Ranasinghe, R. T.; Wang, Y.; Osborne, S. D.; Brown, T.; Fox, K. R. *Nucleic Acids Res.* **2005,** *33,* 3025–3032. doi:10.1093/nar/gki625
116. Koshlap, K. M.; Gillespie, P.; Dervan, P. B.; Feigon, J. *J. Am. Chem. Soc.* **1993,** *115,* 7908–7909. doi:10.1021/ja00070a059
117. Kim, K. T.; Chang, D.; Winssinger, N. *Helv. Chim. Acta* **2018,** *101,* e1700295. doi:10.1002/hlca.201700295
118. Prévot-Halter, I.; Leumann, C. J. *Bioorg. Med. Chem. Lett.* **1999,** *9,* 2657–2660. doi:10.1016/s0960-894x(99)00451-5
119. Endoh, T.; Hnedzko, D.; Rozners, E.; Sugimoto, N. *Angew. Chem., Int. Ed.* **2016,** *55,* 899–903. doi:10.1002/anie.201505938
120. Brodyagin, N.; Kumpina, I.; Applegate, J.; Katkevics, M.; Rozners, E. *ACS Chem. Biol.* **2021,** *16,* 872–881. doi:10.1021/acschembio.1c00044
121. Kumpina, I.; Brodyagin, N.; MacKay, J. A.; Kennedy, S. D.; Katkevics, M.; Rozners, E. *J. Org. Chem.* **2019,** *84,* 13276–13298. doi:10.1021/acs.joc.9b01133
122. Toh, D.-F. K.; Devi, G.; Patil, K. M.; Qu, Q.; Maraswami, M.; Xiao, Y.; Loh, T. P.; Zhao, Y.; Chen, G. *Nucleic Acids Res.* **2016,** *44,* 9071–9082. doi:10.1093/nar/gkw778
123. Semenyuk, A.; Darian, E.; Liu, J.; Majumdar, A.; Cuenoud, B.; Miller, P. S.; MacKerell, A. D., Jr.; Seidman, M. M. *Biochemistry* **2010,** *49,* 7867–7878. doi:10.1021/bi100797z
124. Lin, K.-Y.; Matteucci, M. D. *J. Am. Chem. Soc.* **1998,** *120,* 8531–8532. doi:10.1021/ja981286z
125. Wilds, C. J.; Maier, M. A.; Tereshko, V.; Manoharan, M.; Egli, M. *Angew. Chem., Int. Ed.* **2002,** *41,* 115–117. doi:10.1002/1521-3773(20020104)41:1<115::aid-anie115>3.0.co;2-r
126. Rajeev, K. G.; Maier, M. A.; Lesnik, E. A.; Manoharan, M. *Org. Lett.* **2002,** *4,* 4395–4398. doi:10.1021/ol027026a
127. Ellipilli, S.; Palvai, S.; Ganesh, K. N. *J. Org. Chem.* **2016,** *81,* 6364–6373. doi:10.1021/acs.joc.6b01009
128. Kiviniemi, A.; Murtola, M.; Ingman, P.; Virta, P. *J. Org. Chem.* **2013,** *78,* 5153–5159. doi:10.1021/jo400014y
129. Haaima, G.; Hansen, H. F.; Christensen, L.; Dahl, O.; Nielsen, P. E. *Nucleic Acids Res.* **1997,** *25,* 4639–4643. doi:10.1093/nar/25.22.4639
130. Lohse, J.; Dahl, O.; Nielsen, P. E. *Proc. Natl. Acad. Sci. U. S. A.* **1999,** *96,* 11804–11808. doi:10.1073/pnas.96.21.11804
131. Hudson, R. H. E.; Heidari, A.; Martin-Chan, T.; Park, G.; Wisner, J. A. *J. Org. Chem.* **2019,** *84,* 13252–13261. doi:10.1021/acs.joc.9b00821
132. Akisawa, T.; Ishizawa, Y.; Nagatsugi, F. *Molecules* **2015,** *20,* 4708–4719. doi:10.3390/molecules20034708
133. Akisawa, T.; Yamada, K.; Nagatsugi, F. *Bioorg. Med. Chem. Lett.* **2016,** *26,* 5902–5906. doi:10.1016/j.bmcl.2016.11.002
134. Manicardi, A.; Gyssels, E.; Corradini, R.; Madder, A. *Chem. Commun.* **2016,** *52,* 6930–6933. doi:10.1039/c6cc02062c
135. Chen, D.; Meena; Sharma, S. K.; McLaughlin, L. W. *J. Am. Chem. Soc.* **2004,** *126,* 70–71. doi:10.1021/ja038081x

136. Chen, H.; Meena; McLaughlin, L. W. *J. Am. Chem. Soc.* **2008,** *130,* 13190–13191. doi:10.1021/ja804607v
137. Zeng, Y.; Pratumyot, Y.; Piao, X.; Bong, D. *J. Am. Chem. Soc.* **2012,** *134,* 832–835. doi:10.1021/ja2099326
138. Miao, S.; Liang, Y.; Marathe, I.; Mao, J.; DeSantis, C.; Bong, D. *J. Am. Chem. Soc.* **2019,** *141,* 9365–9372. doi:10.1021/jacs.9b03435
139. Thadke, S. A.; Perera, J. D. R.; Hridya, V. M.; Bhatt, K.; Shaikh, A. Y.; Hsieh, W.-C.; Chen, M.; Gayathri, C.; Gil, R. R.; Rule, G. S.; Mukherjee, A.; Thornton, C. A.; Ly, D. H. *Biochemistry* **2018,** *57,* 2094–2108. doi:10.1021/acs.biochem.8b00062
140. Thadke, S. A.; Hridya, V. M.; Perera, J. D. R.; Gil, R. R.; Mukherjee, A.; Ly, D. H. *Commun. Chem.* **2018,** *1,* 79. doi:10.1038/s42004-018-0080-5
141. Ward, D. C.; Reich, E.; Stryer, L. *J. Biol. Chem.* **1969,** *244,* 1228–1237. doi:10.1016/s0021-9258(18)91833-8
142. Gangamani, B. P.; Kumar, V. A. *Chem. Commun.* **1997,** 1913–1914. doi:10.1039/a705539k
143. Hudson, R. H. E.; Viirre, R. D.; McCourt, N.; Tse, J. *Nucleosides, Nucleotides Nucleic Acids* **2003,** *22,* 1029–1033. doi:10.1081/ncn-120022728
144. Wojciechowski, F.; Hudson, R. H. E. *J. Am. Chem. Soc.* **2008,** *130,* 12574–12575. doi:10.1021/ja804233g
145. Wojciechowski, F.; Hudson, R. H. E. *Org. Lett.* **2009,** *11,* 4878–4881. doi:10.1021/ol9019474
146. Suchý, M.; Hudson, R. H. E. *J. Org. Chem.* **2014,** *79,* 3336–3347. doi:10.1021/jo402873e
147. Cheruiyot, S. K.; Rozners, E. *ChemBioChem* **2016,** *17,* 1558–1562. doi:10.1002/cbic.201600182
148. Krishna, M. S.; Toh, D.-F. K.; Meng, Z.; Ong, A. A. L.; Wang, Z.; Lu, Y.; Xia, K.; Prabakaran, M.; Chen, G. *Anal. Chem. (Washington, DC, U. S.)* **2019,** *91,* 5331–5338. doi:10.1021/acs.analchem.9b00280
149. Ishiguro, T.; Saitoh, J.; Yawata, H.; Otsuka, M.; Inoue, T.; Sugiura, Y. *Nucleic Acids Res.* **1996,** *24,* 4992–4997. doi:10.1093/nar/24.24.4992
150. Köhler, O.; Seitz, O. *Chem. Commun.* **2003,** 2938–2939. doi:10.1039/b308299g
151. Svanvik, N.; Westman, G.; Wang, D.; Kubista, M. *Anal. Biochem.* **2000,** *281,* 26–35. doi:10.1006/abio.2000.4534
152. Svanvik, N.; Nygren, J.; Westman, G.; Kubista, M. *J. Am. Chem. Soc.* **2001,** *123,* 803–809. doi:10.1021/ja002294u
153. Sato, T.; Sato, Y.; Nishizawa, S. *J. Am. Chem. Soc.* **2016,** *138,* 9397–9400. doi:10.1021/jacs.6b05554
154. Sato, T.; Sato, Y.; Nishizawa, S. *Chem. – Eur. J.* **2017,** *23,* 4079–4088. doi:10.1002/chem.201604676
155. Kolevzon, N.; Hashoul, D.; Naik, S.; Rubinstein, A.; Yavin, E. *Chem. Commun.* **2016,** *52,* 2405–2407. doi:10.1039/c5cc07502e
156. Koppelhus, U.; Nielsen, P. E. *Adv. Drug Delivery Rev.* **2003,** *55,* 267–280. doi:10.1016/s0169-409x(02)00182-5
157. Quijano, E.; Bahal, R.; Ricciardi, A.; Saltzman, W. M.; Glazer, P. M. *Yale J. Biol. Med.* **2017,** *90,* 583–598.
158. Montazersaheb, S.; Hejazi, M. S.; Nozad Charoudeh, H. *Adv. Pharm. Bull.* **2018,** *8,* 551–563. doi:10.15171/apb.2018.064
159. Karras, J. G.; Maier, M. A.; Lu, T.; Watt, A.; Manoharan, M. *Biochemistry* **2001,** *40,* 7853–7859. doi:10.1021/bi010263l
160. Siwkowski, A. M.; Malik, L.; Esau, C. C.; Maier, M. A.; Wancewicz, E. V.; Albertshofer, K.; Monia, B. P.; Bennett, C. F.; Eldrup, A. B. *Nucleic Acids Res.* **2004,** *32,* 2695–2706. doi:10.1093/nar/gkh584
161. Shiraishi, T.; Nielsen, P. E. Cellular bioavailability of peptide nucleic acids (PNAs) conjugated to cell penetrating peptides. In *Delivery Technologies for Biopharmaceuticals: Peptides, Proteins, Nucleic Acids and Vaccines;* Jorgensen, L.; Mørck Nielsen, H., Eds.; John Wiley & Sons: Chichester, UK, 2009; pp 305–338. doi:10.1002/9780470688397.ch16
162. Mäe, M.; El Andaloussi, S.; Lehto, T.; Langel, Ü. *Expert Opin. Drug Delivery* **2009,** *6,* 1195–1205. doi:10.1517/17425240903213688
163. Shiraishi, T.; Nielsen, P. E. *Nat. Protoc.* **2006,** *1,* 633–636. doi:10.1038/nprot.2006.92
164. Abes, S.; Turner, J. J.; Ivanova, G. D.; Owen, D.; Williams, D.; Arzumanov, A.; Clair, P.; Gait, M. J.; Lebleu, B. *Nucleic Acids Res.* **2007,** *35,* 4495–4502. doi:10.1093/nar/gkm418
165. Pardridge, W. M.; Boado, R. J.; Kang, Y. S. *Proc. Natl. Acad. Sci. U. S. A.* **1995,** *92,* 5592–5596. doi:10.1073/pnas.92.12.5592
166. Basu, S.; Wickstrom, E. *Bioconjugate Chem.* **1997,** *8,* 481–488. doi:10.1021/bc9700650
167. Derossi, D.; Joliot, A. H.; Chassaing, G.; Prochiantz, A. *J. Biol. Chem.* **1994,** *269,* 10444–10450. doi:10.1016/s0021-9258(17)34080-2
168. Fawell, S.; Seery, J.; Daikh, Y.; Moore, C.; Chen, L. L.; Pepinsky, B.; Barsoum, J. *Proc. Natl. Acad. Sci. U. S. A.* **1994,** *91,* 664–668. doi:10.1073/pnas.91.2.664
169. Pooga, M.; Hällbrink, M.; Zorko, M.; Langel, Ü. *FASEB J.* **1998,** *12,* 67–77. doi:10.1096/fsb2fasebj.12.1.67
170. Simmons, C. G.; Pitts, A. E.; Mayfield, L. D.; Shay, J. W.; Corey, D. R. *Bioorg. Med. Chem. Lett.* **1997,** *7,* 3001–3006. doi:10.1016/s0960-894x(97)10136-6
171. Pooga, M.; Soomets, U.; Hällbrink, M.; Valkna, A.; Saar, K.; Rezaei, K.; Kahl, U.; Hao, J.-X.; Xu, X.-J.; Wiesenfeld-Hallin, Z.; Hökfelt, T.; Bartfai, T.; Langel, Ü. *Nat. Biotechnol.* **1998,** *16,* 857–861. doi:10.1038/nbt0998-857
172. Cutrona, G.; Carpaneto, E. M.; Ulivi, M.; Roncella, S.; Landt, O.; Ferrarini, M.; Boffa, L. C. *Nat. Biotechnol.* **2000,** *18,* 300–303. doi:10.1038/73745
173. Braun, K.; Peschke, P.; Pipkorn, R.; Lampel, S.; Wachsmuth, M.; Waldeck, W.; Friedrich, E.; Debus, J. *J. Mol. Biol.* **2002,** *318,* 237–243. doi:10.1016/s0022-2836(02)00031-1
174. Koppelhus, U.; Awasthi, S. K.; Zachar, V.; Holst, H. U.; Ebbesen, P.; Nielsen, P. E. *Antisense Nucleic Acid Drug Dev.* **2002,** *12,* 51–63. doi:10.1089/108729002760070795
175. Turner, J. J.; Ivanova, G. D.; Verbeure, B.; Williams, D.; Arzumanov, A. A.; Abes, S.; Lebleu, B.; Gait, M. J. *Nucleic Acids Res.* **2005,** *33,* 6837–6849. doi:10.1093/nar/gki991
176. Zeng, Z.; Han, S.; Hong, W.; Lang, Y.; Li, F.; Liu, Y.; Li, Z.; Wu, Y.; Li, W.; Zhang, X.; Cao, Z. *Mol. Ther.–Nucleic Acids* **2016,** *5,* e295. doi:10.1038/mtna.2016.11
177. Hunt, J. F.; Rath, P.; Rothschild, K. J.; Engelman, D. M. *Biochemistry* **1997,** *36,* 15177–15192. doi:10.1021/bi970147b
178. Andreev, O. A.; Karabadzhak, A. G.; Weerakkody, D.; Andreev, G. O.; Engelman, D. M.; Reshetnyak, Y. K. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107,* 4081–4086. doi:10.1073/pnas.0914330107
179. An, M.; Wijesinghe, D.; Andreev, O. A.; Reshetnyak, Y. K.; Engelman, D. M. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107,* 20246–20250. doi:10.1073/pnas.1014403107
180. Andreev, O. A.; Engelman, D. M.; Reshetnyak, Y. K. *Front. Physiol.* **2014,** *5,* 97. doi:10.3389/fphys.2014.00097

181. Cheng, C. J.; Bahal, R.; Babar, I. A.; Pincus, Z.; Barrera, F.; Liu, C.; Svoronos, A.; Braddock, D. T.; Glazer, P. M.; Engelman, D. M.; Saltzman, W. M.; Slack, F. J. *Nature* **2015,** *518,* 107–110. doi:10.1038/nature13905
182. Kaplan, A. R.; Pham, H.; Liu, Y.; Oyaghire, S.; Bahal, R.; Engelman, D. M.; Glazer, P. M. *Mol. Cancer Res.* **2020,** *18,* 873–882. doi:10.1158/1541-7786.mcr-19-0661
183. Lu, Z.; Paolella, B. R.; Truex, N. L.; Loftis, A. R.; Liao, X.; Rabideau, A. E.; Brown, M. S.; Busanovich, J.; Beroukhim, R.; Pentelute, B. L. *ACS Chem. Biol.* **2020,** *15,* 1358–1369. doi:10.1021/acschembio.9b01027
184. Sazani, P.; Kang, S.-H.; Maier, M. A.; Wei, C.; Dillman, J.; Summerton, J.; Manoharan, M.; Kole, R. *Nucleic Acids Res.* **2001,** *29,* 3965–3974. doi:10.1093/nar/29.19.3965
185. Sazani, P.; Gemignani, F.; Kang, S.-H.; Maier, M. A.; Manoharan, M.; Persmark, M.; Bortner, D.; Kole, R. *Nat. Biotechnol.* **2002,** *20,* 1228–1233. doi:10.1038/nbt759
186. Hu, J.; Matsui, M.; Gagnon, K. T.; Schwartz, J. C.; Gabillet, S.; Arar, K.; Wu, J.; Bezprozvanny, I.; Corey, D. R. *Nat. Biotechnol.* **2009,** *27,* 478–484. doi:10.1038/nbt.1539
187. Hu, J.; Corey, D. R. *Biochemistry* **2007,** *46,* 7581–7589. doi:10.1021/bi700230a
188. Fabani, M. M.; Gait, M. J. *RNA* **2008,** *14,* 336–346. doi:10.1261/rna.844108
189. Wancewicz, E. V.; Maier, M. A.; Siwkowski, A. M.; Albertshofer, K.; Winger, T. M.; Berdeja, A.; Gaus, H.; Vickers, T. A.; Bennett, C. F.; Monia, B. P.; Griffey, R. H.; Nulf, C. J.; Hu, J.; Corey, D. R.; Swayze, E. E.; Kinberger, G. A. *J. Med. Chem.* **2010,** *53,* 3919–3926. doi:10.1021/jm901489k
190. Fabani, M. M.; Abreu-Goodger, C.; Williams, D.; Lyons, P. A.; Torres, A. G.; Smith, K. G. C.; Enright, A. J.; Gait, M. J.; Vigorito, E. *Nucleic Acids Res.* **2010,** *38,* 4466–4475. doi:10.1093/nar/gkq160
191. Torres, A. G.; Fabani, M. M.; Vigorito, E.; Williams, D.; Al-Obaidi, N.; Wojciechowski, F.; Hudson, R. H. E.; Seitz, O.; Gait, M. J. *Nucleic Acids Res.* **2012,** *40,* 2152–2167. doi:10.1093/nar/gkr885
192. Kaihatsu, K.; Huffman, K. E.; Corey, D. R. *Biochemistry* **2004,** *43,* 14340–14347. doi:10.1021/bi048519l
193. Wright, D. G.; Zhang, Y.; Murphy, J. R. *Biochem. Biophys. Res. Commun.* **2008,** *376,* 200–205. doi:10.1016/j.bbrc.2008.08.124
194. Good, L.; Nielsen, P. E. *Proc. Natl. Acad. Sci. U. S. A.* **1998,** *95,* 2073–2076. doi:10.1073/pnas.95.5.2073
195. Vaara, M.; Porro, M. *Antimicrob. Agents Chemother.* **1996,** *40,* 1801–1805. doi:10.1128/aac.40.8.1801
196. Good, L.; Awasthi, S. K.; Dryselius, R.; Larsson, O.; Nielsen, P. E. *Nat. Biotechnol.* **2001,** *19,* 360–364. doi:10.1038/86753
197. Abes, S.; Moulton, H. M.; Clair, P.; Prevot, P.; Youngblood, D. S.; Wu, R. P.; Iversen, P. L.; Lebleu, B. *J. Controlled Release* **2006,** *116,* 304–313. doi:10.1016/j.jconrel.2006.09.011
198. Ivanova, G. D.; Arzumanov, A.; Abes, R.; Yin, H.; Wood, M. J. A.; Lebleu, B.; Gait, M. J. *Nucleic Acids Res.* **2008,** *36,* 6418–6428. doi:10.1093/nar/gkn671
199. Richard, J. P.; Melikov, K.; Brooks, H.; Prevot, P.; Lebleu, B.; Chernomordik, L. V. *J. Biol. Chem.* **2005,** *280,* 15300–15306. doi:10.1074/jbc.m401604200
200. Stanzl, E. G.; Trantow, B. M.; Vargas, J. R.; Wender, P. A. *Acc. Chem. Res.* **2013,** *46,* 2944–2954. doi:10.1021/ar4000554
201. Brognara, E.; Fabbri, E.; Bazzoli, E.; Montagner, G.; Ghimenton, C.; Eccher, A.; Cantù, C.; Manicardi, A.; Bianchi, N.; Finotti, A.; Breveglieri, G.; Borgatti, M.; Corradini, R.; Bezzerri, V.; Cabrini, G.; Gambari, R. *J. Neuro-Oncol.* **2014,** *118,* 19–28. doi:10.1007/s11060-014-1405-6
202. Qian, Z.; Martyna, A.; Hard, R. L.; Wang, J.; Appiah-Kubi, G.; Coss, C.; Phelps, M. A.; Rossman, J. S.; Pei, D. *Biochemistry* **2016,** *55,* 2601–2612. doi:10.1021/acs.biochem.6b00226
203. Soudah, T.; Khawaled, S.; Aqeilan, R. I.; Yavin, E. *ACS Omega* **2019,** *4,* 13954–13961. doi:10.1021/acsomega.9b01697
204. Muratovska, A.; Lightowlers, R. N.; Taylor, R. W.; Turnbull, D. M.; Smith, R. A. J.; Wilce, J. A.; Martin, S. W.; Murphy, M. P. *Nucleic Acids Res.* **2001,** *29,* 1852–1863. doi:10.1093/nar/29.9.1852
205. Mehiri, M.; Upert, G.; Tripathi, S.; Di Giorgio, A.; Condom, R.; Pandey, V. N.; Patino, N. *Oligonucleotides* **2008,** *18,* 245–256. doi:10.1089/oli.2008.0126
206. Shen, G.; Fang, H.; Song, Y.; Bielska, A. A.; Wang, Z.; Taylor, J.-S. A. *Bioconjugate Chem.* **2009,** *20,* 1729–1736. doi:10.1021/bc900048y
207. Shiraishi, T.; Nielsen, P. E. *Bioconjugate Chem.* **2012,** *23,* 196–202. doi:10.1021/bc200460t
208. Zhang, X.; Simmons, C. G.; Corey, D. R. *Bioorg. Med. Chem. Lett.* **2001,** *11,* 1269–1272. doi:10.1016/s0960-894x(01)00198-6
209. Herbert, B.-S.; Pitts, A. E.; Baker, S. I.; Hamilton, S. E.; Wright, W. E.; Shay, J. W.; Corey, D. R. *Proc. Natl. Acad. Sci. U. S. A.* **1999,** *96,* 14276–14281. doi:10.1073/pnas.96.25.14276
210. Valentijn, A. R. P. M.; van der Marel, G. A.; Sliedregt, L. A. J. M.; van Berkel, T. J. C.; Biessen, E. A. L.; van Boom, J. H. *Tetrahedron* **1997,** *53,* 759–770. doi:10.1016/s0040-4020(96)01018-6
211. Biessen, E. A. L.; Sliedregt-Bol, K.; 'T Hoen, P. A. C.; Prince, P.; Van der Bilt, E.; Valentijn, A. R. P. M.; Meeuwenoord, N. J.; Princen, H.; Bijsterbosch, M. K.; Van der Marel, G. A.; Van Boom, J. H.; Van Berkel, T. J. C. *Bioconjugate Chem.* **2002,** *13,* 295–302. doi:10.1021/bc015550g
212. van Rossenberg, S. M. W.; Sliedregt-Bol, K. M.; Prince, P.; van Berkel, T. J. C.; van Boom, J. H.; van der Marel, G. A.; Biessen, E. A. L. *Bioconjugate Chem.* **2003,** *14,* 1077–1082. doi:10.1021/bc0340417
213. Nair, J. K.; Willoughby, J. L. S.; Chan, A.; Charisse, K.; Alam, M. R.; Wang, Q.; Hoekstra, M.; Kandasamy, P.; Kel'in, A. V.; Milstein, S.; Taneja, N.; O'Shea, J.; Shaikh, S.; Zhang, L.; van der Sluis, R. J.; Jung, M. E.; Akinc, A.; Hutabarat, R.; Kuchimanchi, S.; Fitzgerald, K.; Zimmermann, T.; van Berkel, T. J. C.; Maier, M. A.; Rajeev, K. G.; Manoharan, M. *J. Am. Chem. Soc.* **2014,** *136,* 16958–16961. doi:10.1021/ja505986a
214. Prakash, T. P.; Graham, M. J.; Yu, J.; Carty, R.; Low, A.; Chappell, A.; Schmidt, K.; Zhao, C.; Aghajan, M.; Murray, H. F.; Riney, S.; Booten, S. L.; Murray, S. F.; Gaus, H.; Crosby, J.; Lima, W. F.; Guo, S.; Monia, B. P.; Swayze, E. E.; Seth, P. P. *Nucleic Acids Res.* **2014,** *42,* 8796–8807. doi:10.1093/nar/gku531
215. Scott, L. J. *Drugs* **2020,** *80,* 335–339. doi:10.1007/s40265-020-01269-0
216. Bhingardeve, P.; Madhanagopal, B. R.; Naick, H.; Jain, P.; Manoharan, M.; Ganesh, K. *J. Org. Chem.* **2020,** *85,* 8812–8824. doi:10.1021/acs.joc.0c00601
217. Petrus, A. K.; Fairchild, T. J.; Doyle, R. P. *Angew. Chem., Int. Ed.* **2009,** *48,* 1022–1028. doi:10.1002/anie.200800865
218. Clardy, S. M.; Allis, D. G.; Fairchild, T. J.; Doyle, R. P. *Expert Opin. Drug Delivery* **2011,** *8,* 127–140. doi:10.1517/17425247.2011.539200

219. Wierzba, A.; Wojciechowska, M.; Trylska, J.; Gryko, D. *Bioconjugate Chem.* **2016,** *27,* 189–197. doi:10.1021/acs.bioconjchem.5b00599
220. Równicki, M.; Wojciechowska, M.; Wierzba, A. J.; Czarnecki, J.; Bartosik, D.; Gryko, D.; Trylska, J. *Sci. Rep.* **2017,** *7,* 7644. doi:10.1038/s41598-017-08032-8
221. Równicki, M.; Dąbrowska, Z.; Wojciechowska, M.; Wierzba, A. J.; Maximova, K.; Gryko, D.; Trylska, J. *ACS Omega* **2019,** *4,* 819–824. doi:10.1021/acsomega.8b03139
222. Pieńko, T.; Czarnecki, J.; Równicki, M.; Wojciechowska, M.; Wierzba, A. J.; Gryko, D.; Bartosik, D.; Trylska, J. *Biophys. J.* **2021,** *120,* 725–737. doi:10.1016/j.bpj.2021.01.004
223. Riguet, E.; Tripathi, S.; Chaubey, B.; Désiré, J.; Pandey, V. N.; Décout, J.-L. *J. Med. Chem.* **2004,** *47,* 4806–4809. doi:10.1021/jm049642d
224. Chaubey, B.; Tripathi, S.; Désiré, J.; Baussanne, I.; Décout, J.-L.; Pandey, V. N. *Oligonucleotides* **2007,** *17,* 302–313. doi:10.1089/oli.2007.0085
225. Das, I.; Désiré, J.; Manvar, D.; Baussanne, I.; Pandey, V. N.; Décout, J.-L. *J. Med. Chem.* **2012,** *55,* 6021–6032. doi:10.1021/jm300253q
226. Kesy, J.; Patil, K. M.; Kumar, S. R.; Shu, Z.; Yong, H. Y.; Zimmermann, L.; Ong, A. A. L.; Toh, D.-F. K.; Krishna, M. S.; Yang, L.; Decout, J.-L.; Luo, D.; Prabakaran, M.; Chen, G.; Kierzek, E. *Bioconjugate Chem.* **2019,** *30,* 931–943. doi:10.1021/acs.bioconjchem.9b00039
227. Ong, A. A. L.; Tan, J.; Bhadra, M.; Dezanet, C.; Patil, K. M.; Chong, M. S.; Kierzek, R.; Decout, J.-L.; Roca, X.; Chen, G. *Molecules* **2019,** *24,* 3020. doi:10.3390/molecules24163020
228. Ørum, H.; Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O.; Stanley, C. *Nucleic Acids Res.* **1993,** *21,* 5332–5336. doi:10.1093/nar/21.23.5332
229. Thiede, C.; Bayerdörffer, E.; Blaszczyk, R.; Wittig, B.; Neubauer, A. *Nucleic Acids Res.* **1996,** *24,* 983–984. doi:10.1093/nar/24.5.983
230. Nagai, Y.; Miyazawa, H.; Huqun; Tanaka, T.; Udagawa, K.; Kato, M.; Fukuyama, S.; Yokote, A.; Kobayashi, K.; Kanazawa, M.; Hagiwara, K. *Cancer Res.* **2005,** *65,* 7276–7282. doi:10.1158/0008-5472.can-05-0331
231. Tanaka, T.; Nagai, Y.; Miyazawa, H.; Koyama, N.; Matsuoka, S.; Sutani, A.; Udagawa, K.; Murayama, Y.; Nagata, M.; Shimizu, Y.; Ikebuchi, K.; Kanazawa, M.; Kobayashi, K.; Hagiwara, K. *Cancer Sci.* **2007,** *98,* 246–252. doi:10.1111/j.1349-7006.2006.00377.x
232. Kim, H.-R.; Lee, S. Y.; Hyun, D.-S.; Lee, M. K.; Lee, H.-K.; Choi, C.-M.; Yang, S.-H.; Kim, Y.-C.; Lee, Y. C.; Kim, S. Y.; Jang, S. H.; Lee, J. C.; Lee, K. Y. *J. Exp. Clin. Cancer Res.* **2013,** *32,* 50. doi:10.1186/1756-9966-32-50
233. Song, J.; Hegge, J. W.; Mauk, M. G.; Chen, J.; Till, J. E.; Bhagwat, N.; Azink, L. T.; Peng, J.; Sen, M.; Mays, J.; Carpenter, E. L.; van der Oost, J.; Bau, H. H. *Nucleic Acids Res.* **2020,** *48,* e19. doi:10.1093/nar/gkz1165
234. Demidov, V. V. *Expert Rev. Mol. Diagn.* **2001,** *1,* 343–351. doi:10.1586/14737159.1.3.343
235. Bukanov, N. O.; Demidov, V. V.; Nielsen, P. E.; Frank-Kamenetskii, M. D. *Proc. Natl. Acad. Sci. U. S. A.* **1998,** *95,* 5516–5520. doi:10.1073/pnas.95.10.5516
236. Kuhn, H.; Demidov, V. V.; Frank-Kamenetskii, M. D. *Nucleic Acids Res.* **2002,** *30,* 574–580. doi:10.1093/nar/30.2.574
237. Smolina, I.; Lee, C.; Frank-Kamenetskii, M. *Appl. Environ. Microbiol.* **2007,** *73,* 2324–2328. doi:10.1128/aem.02038-06
238. Yaroslavsky, A. I.; Smolina, I. V. *Chem. Biol.* **2013,** *20,* 445–453. doi:10.1016/j.chembiol.2013.02.012
239. Xu, X.; Xing, S.; Xu, M.; Fu, P.; Gao, T.; Zhang, X.; Zhao, Y.; Zhao, C. *RSC Adv.* **2019,** *9,* 38298–38308. doi:10.1039/c9ra06758b
240. Lansdorp, P. M.; Verwoerd, N. P.; van de Rijke, F. M.; Dragowska, V.; Little, M.-T.; Dirks, R. W.; Raap, A. K.; Tanke, H. J. *Hum. Mol. Genet.* **1996,** *5,* 685–691. doi:10.1093/hmg/5.5.685
241. Vilaivan, T. *Beilstein J. Org. Chem.* **2018,** *14,* 253–281. doi:10.3762/bjoc.14.17
242. Thisted, M. J. T.; Pluzek, K.-J.; Petersen, K. H.; Hyldigg-Nielsen, J.-J.; Godtfredsen, S. E. *Cell Visions* **1996,** *3,* 358–363.
243. Murakami, T.; Hagiwara, T.; Yamamoto, K.; Hattori, J.; Kasami, M.; Utsumi, M.; Kaneda, T. *J. Pathol.* **2001,** *194,* 130–135. doi:10.1002/path.843
244. Ortiz, E.; Estrada, G.; Lizardi, P. M. *Mol. Cell. Probes* **1998,** *12,* 219–226. doi:10.1006/mcpr.1998.0175
245. Armitage, B.; Ly, D.; Koch, T.; Frydenlund, H.; Ørum, H.; Schuster, G. B. *Biochemistry* **1998,** *37,* 9417–9425. doi:10.1021/bi9729458
246. Seitz, O. *Angew. Chem., Int. Ed.* **2000,** *39,* 3249–3252. doi:10.1002/1521-3773(20000915)39:18<3249::aid-anie3249>3.0.co;2-m
247. Kuhn, H.; Demidov, V. V.; Coull, J. M.; Fiandaca, M. J.; Gildea, B. D.; Frank-Kamenetskii, M. D. *J. Am. Chem. Soc.* **2002,** *124,* 1097–1103. doi:10.1021/ja0041324
248. Svanvik, N.; Ståhlberg, A.; Sehlstedt, U.; Sjöback, R.; Kubista, M. *Anal. Biochem.* **2000,** *287,* 179–182. doi:10.1006/abio.2000.4824
249. Robertson, K. L.; Yu, L.; Armitage, B. A.; Lopez, A. J.; Peteanu, L. A. *Biochemistry* **2006,** *45,* 6066–6074. doi:10.1021/bi052050s
250. Seitz, O.; Bergmann, F.; Heindl, D. *Angew. Chem., Int. Ed.* **1999,** *38,* 2203–2206. doi:10.1002/(sici)1521-3773(19990802)38:15<2203::aid-anie2203>3.0.co;2-2
251. Köhler, O.; Jarikote, D. V.; Seitz, O. *ChemBioChem* **2005,** *6,* 69–77. doi:10.1002/cbic.200400260
252. Socher, E.; Bethge, L.; Knoll, A.; Jungnick, N.; Herrmann, A.; Seitz, O. *Angew. Chem., Int. Ed.* **2008,** *47,* 9555–9559. doi:10.1002/anie.200803549
253. Socher, E.; Knoll, A.; Seitz, O. *Org. Biomol. Chem.* **2012,** *10,* 7363–7371. doi:10.1039/c2ob25925g
254. Blanco, A. M.; Rausell, L.; Aguado, B.; Perez-Alonso, M.; Artero, R. *Nucleic Acids Res.* **2009,** *37,* e116. doi:10.1093/nar/gkp551
255. Koppitz, M.; Nielsen, P. E.; Orgel, L. E. *J. Am. Chem. Soc.* **1998,** *120,* 4563–4569. doi:10.1021/ja974190y
256. Pianowski, Z. L.; Winssinger, N. *Chem. Commun.* **2007,** 3820–3822. doi:10.1039/b709611a
257. Pianowski, Z.; Gorska, K.; Oswald, L.; Merten, C. A.; Winssinger, N. *J. Am. Chem. Soc.* **2009,** *131,* 6492–6497. doi:10.1021/ja809656k
258. Kummer, S.; Knoll, A.; Socher, E.; Bethge, L.; Herrmann, A.; Seitz, O. *Bioconjugate Chem.* **2012,** *23,* 2051–2060. doi:10.1021/bc300249f
259. Peled, I.; Yavin, E. *ACS Omega* **2018,** *3,* 3813–3818. doi:10.1021/acsomega.8b00184
260. Hashoul, D.; Shapira, R.; Falchenko, M.; Tepper, O.; Paviov, V.; Nissan, A.; Yavin, E. *Biosens. Bioelectron.* **2019,** *137,* 271–278. doi:10.1016/j.bios.2019.04.056
261. Stender, H.; Mollerup, T. A.; Lund, K.; Petersen, K. H.; Hongmanee, P.; Godtfredsen, S. E. *Int. J. Tuberc. Lung Dis.* **1999,** *3,* 830–837.

262. Oliveira, K.; Procop, G. W.; Wilson, D.; Coull, J.; Stender, H. *J. Clin. Microbiol.* **2002,** *40,* 247–251. doi:10.1128/jcm.40.1.247-251.2002
263. Rigby, S.; Procop, G. W.; Haase, G.; Wilson, D.; Hall, G.; Kurtzman, C.; Oliveira, K.; Von Oy, S.; Hyldig-Nielsen, J. J.; Coull, J.; Stender, H. *J. Clin. Microbiol.* **2002,** *40,* 2182–2186. doi:10.1128/jcm.40.6.2182-2186.2002
264. Xi, C.; Balberg, M.; Boppart, S. A.; Raskin, L. *Appl. Environ. Microbiol.* **2003,** *69,* 5673–5678. doi:10.1128/aem.69.9.5673-5678.2003
265. Wilson, D. A.; Joyce, M. J.; Hall, L. S.; Reller, L. B.; Roberts, G. D.; Hall, G. S.; Alexander, B. D.; Procop, G. W. *J. Clin. Microbiol.* **2005,** *43,* 2909–2912. doi:10.1128/jcm.43.6.2909-2912.2005
266. Ly, T.; Gulia, J.; Pyrgos, V.; Waga, M.; Shoham, S. *Ther. Clin. Risk Manage.* **2008,** *4,* 637–640. doi:10.2147/tcrm.s2838
267. Cerqueira, L.; Fernandes, R. M.; Ferreira, R. M.; Carneiro, F.; Dinis-Ribeiro, M.; Figueiredo, C.; Keevil, C. W.; Azevedo, N. F.; Vieira, M. J. *BMC Microbiol.* **2011,** *11,* 101. doi:10.1186/1471-2180-11-101
268. Abdelhamed, A. M.; Zhang, S. X.; Watkins, T.; Morgan, M. A.; Wu, F.; Buckner, R. J.; Fuller, D. D.; Davis, T. E.; Salimnia, H.; Fairfax, M. R.; Lephart, P. R.; Poulter, M. D.; Regi, S. B.; Jacobs, M. R. *J. Clin. Microbiol.* **2015,** *53,* 1672–1676. doi:10.1128/jcm.00549-15
269. Mach, K. E.; Kaushik, A. M.; Hsieh, K.; Wong, P. K.; Wang, T.-H.; Liao, J. C. *Analyst* **2019,** *144,* 1565–1574. doi:10.1039/c8an02194e
270. Su, X.; Teh, H. F.; Lieu, X.; Gao, Z. *Anal. Chem. (Washington, DC, U. S.)* **2007,** *79,* 7192–7197. doi:10.1021/ac0709403
271. Zhang, G.-J.; Chua, J. H.; Chee, R.-E.; Agarwal, A.; Wong, S. M. *Biosens. Bioelectron.* **2009,** *24,* 2504–2508. doi:10.1016/j.bios.2008.12.035
272. Sadhu, K. K.; Winssinger, N. *Chem. – Eur. J.* **2013,** *19,* 8182–8189. doi:10.1002/chem.201300060
273. Ryoo, S.-R.; Lee, J.; Yeo, J.; Na, H.-K.; Kim, Y.-K.; Jang, H.; Lee, J. H.; Han, S. W.; Lee, Y.; Kim, V. N.; Min, D.-H. *ACS Nano* **2013,** *7,* 5882–5891. doi:10.1021/nn401183s
274. Kotikam, V.; Fernandes, M.; Kumar, V. A. *Phys. Chem. Chem. Phys.* **2012,** *14,* 15003–15006. doi:10.1039/c2cp42770b
275. Hong, C.; Baek, A.; Hah, S. S.; Jung, W.; Kim, D.-E. *Anal. Chem. (Washington, DC, U. S.)* **2016,** *88,* 2999–3003. doi:10.1021/acs.analchem.6b00046
276. Metcalf, G. A. D.; Shibakawa, A.; Patel, H.; Sita-Lumsden, A.; Zivi, A.; Rama, N.; Bevan, C. L.; Ladame, S. *Anal. Chem. (Washington, DC, U. S.)* **2016,** *88,* 8091–8098. doi:10.1021/acs.analchem.6b01594
277. Pavagada, S.; Channon, R. B.; Chang, J. Y. H.; Kim, S. H.; MacIntyre, D.; Bennett, P. R.; Terzidou, V.; Ladame, S. *Chem. Commun.* **2019,** *55,* 12451–12454. doi:10.1039/c9cc05607f
278. Sayers, J.; Payne, R. J.; Winssinger, N. *Chem. Sci.* **2018,** *9,* 896–903. doi:10.1039/c7sc02736b
279. Thurley, S.; Röglin, L.; Seitz, O. *J. Am. Chem. Soc.* **2007,** *129,* 12693–12695. doi:10.1021/ja075487r
280. Oh, K. J.; Cash, K. J.; Lubin, A. A.; Plaxco, K. W. *Chem. Commun.* **2007,** 4869–4871. doi:10.1039/b709776j
281. Dhar, A.; Ahmed, I.; Mallick, S.; Roy, S. *ChemBioChem* **2020,** *21,* 2121–2125. doi:10.1002/cbic.202000097
282. Nielsen, P. E.; Egholm, M.; Buchardt, O. *Gene* **1994,** *149,* 139–145. doi:10.1016/0378-1119(94)90422-7
283. Knudsen, H.; Nielsen, P. E. *Nucleic Acids Res.* **1996,** *24,* 494–500. doi:10.1093/nar/24.3.494
284. Doyle, D. F.; Braasch, D. A.; Janowski, B. A.; Corey, D. R. *Biochemistry* **2001,** *40,* 53–64. doi:10.1021/bi0020630
285. Economos, N. G.; Oyaghire, S.; Quijano, E.; Ricciardi, A. S.; Saltzman, W. M.; Glazer, P. M. *Molecules* **2020,** *25,* 735. doi:10.3390/molecules25030735
286. Hyrup, B.; Nielsen, P. E. *Bioorg. Med. Chem.* **1996,** *4,* 5–23. doi:10.1016/0968-0896(95)00171-9
287. Demidov, V. V.; Potaman, V. N.; Frank-Kamenetskil, M. D.; Egholm, M.; Buchard, O.; Sönnichsen, S. H.; Nielsen, P. E. *Biochem. Pharmacol.* **1994,** *48,* 1310–1313. doi:10.1016/0006-2952(94)90171-6
288. Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W. *Angew. Chem., Int. Ed.* **1998,** *37,* 2796–2823. doi:10.1002/(sici)1521-3773(19981102)37:20<2796::aid-anie2796>3.0.co;2-k
289. Kaushik, N.; Basu, A.; Palumbo, P.; Myers, R. L.; Pandey, V. N. *J. Virol.* **2002,** *76,* 3881–3891. doi:10.1128/jvi.76.8.3881-3891.2002
290. Kolevzon, N.; Nasereddin, A.; Naik, S.; Yavin, E.; Dzikowski, R. *PLoS One* **2014,** *9,* e86802. doi:10.1371/journal.pone.0086802
291. Yin, H.; Lu, Q.; Wood, M. *Mol. Ther.* **2008,** *16,* 38–45. doi:10.1038/sj.mt.6300329
292. Yin, H.; Betts, C.; Saleh, A. F.; Ivanova, G. D.; Lee, H.; Seow, Y.; Kim, D.; Gait, M. J.; Wood, M. J. A. *Mol. Ther.* **2010,** *18,* 819–827. doi:10.1038/mt.2009.310
293. Gao, X.; Shen, X.; Dong, X.; Ran, N.; Han, G.; Cao, L.; Gu, B.; Yin, H. *Mol. Ther.–Nucleic Acids* **2015,** *4,* e255. doi:10.1038/mtna.2015.27
294. Brolin, C.; Lim, E. W. K.; Grizot, S.; Olsen, C. H.; Yavari, N.; Krag, T. O.; Nielsen, P. E. *Nucleic Acid Ther.* **2021,** *31,* 208–219. doi:10.1089/nat.2020.0856
295. Kayali, R.; Bury, F.; Ballard, M.; Bertoni, C. *Hum. Mol. Genet.* **2010,** *19,* 3266–3281. doi:10.1093/hmg/ddq235
296. Nik-Ahd, F.; Bertoni, C. *Stem Cells (Durham, NC, U. S.)* **2014,** *32,* 1817–1830. doi:10.1002/stem.1668
297. Bahal, R.; Ali McNeer, N.; Quijano, E.; Liu, Y.; Sulkowski, P.; Turchick, A.; Lu, Y.-C.; Bhunia, D. C.; Manna, A.; Greiner, D. L.; Brehm, M. A.; Cheng, C. J.; López-Giráldez, F.; Ricciardi, A.; Beloor, J.; Krause, D. S.; Kumar, P.; Gallagher, P. G.; Braddock, D. T.; Mark Saltzman, W.; Ly, D. H.; Glazer, P. M. *Nat. Commun.* **2016,** *7,* 13304. doi:10.1038/ncomms13304
298. Boffa, L. C.; Cutrona, G.; Cilli, M.; Mariani, M. R.; Matis, S.; Pastorino, M.; Damonte, G.; Millo, E.; Roncella, S.; Ferrarini, M. *Oligonucleotides* **2005,** *15,* 85–93. doi:10.1089/oli.2005.15.85
299. Boffa, L. C.; Cutrona, G.; Cilli, M.; Matis, S.; Damonte, G.; Mariani, M. R.; Millo, E.; Moroni, M.; Roncella, S.; Fedeli, F.; Ferrarini, M. *Cancer Gene Ther.* **2007,** *14,* 220–226. doi:10.1038/sj.cgt.7701002
300. Alagpulinsa, D. A.; Yaccoby, S.; Ayyadevara, S.; Shmookler Reis, R. J. *Cancer Biol. Ther.* **2015,** *16,* 976–986. doi:10.1080/15384047.2015.1040951
301. Babar, I. A.; Cheng, C. J.; Booth, C. J.; Liang, X.; Weidhaas, J. B.; Saltzman, W. M.; Slack, F. J. *Proc. Natl. Acad. Sci. U. S. A.* **2012,** *109,* E1695–E1704. doi:10.1073/pnas.1201516109
302. Malik, S.; Lim, J.; Slack, F. J.; Braddock, D. T.; Bahal, R. *J. Controlled Release* **2020,** *327,* 406–419. doi:10.1016/j.jconrel.2020.08.026
303. Gupta, A.; Quijano, E.; Liu, Y.; Bahal, R.; Scanlon, S. E.; Song, E.; Hsieh, W.-C.; Braddock, D. E.; Ly, D. H.; Saltzman, W. M.; Glazer, P. M. *Mol. Ther.–Nucleic Acids* **2017,** *9,* 111–119. doi:10.1016/j.omtn.2017.09.001

304. Kim, K.; Kim, H. H.; Lee, C.-H.; Kim, S.; Cheon, G. J.; Kang, K. W.; Chung, J.-K.; Youn, H. *Biochem. Biophys. Res. Commun.* **2020,** *529,* 707–713. doi:10.1016/j.bbrc.2020.05.215
305. Tan, X.-X.; Actor, J. K.; Chen, Y. *Antimicrob. Agents Chemother.* **2005,** *49,* 3203–3207. doi:10.1128/aac.49.8.3203-3207.2005
306. Abushahba, M. F. N.; Mohammad, H.; Thangamani, S.; Hussein, A. A. A.; Seleem, M. N. *Sci. Rep.* **2016,** *6,* 20832. doi:10.1038/srep20832
307. Feng, S.; Holland, E. C. *Nature* **1988,** *334,* 165–167. doi:10.1038/334165a0
308. Chaubey, B.; Tripathi, S.; Ganguly, S.; Harris, D.; Casale, R. A.; Pandey, V. N. *Virology* **2005,** *331,* 418–428. doi:10.1016/j.virol.2004.10.032
309. Amit-Avraham, I.; Pozner, G.; Eshar, S.; Fastman, Y.; Kolevzon, N.; Yavin, E.; Dzikowski, R. *Proc. Natl. Acad. Sci. U. S. A.* **2015,** *112,* E982–E991. doi:10.1073/pnas.1420855112
310. Faruqi, A. F.; Egholm, M.; Glazer, P. M. *Proc. Natl. Acad. Sci. U. S. A.* **1998,** *95,* 1398–1403. doi:10.1073/pnas.95.4.1398
311. Chin, J. Y.; Kuan, J. Y.; Lonkar, P. S.; Krause, D. S.; Seidman, M. M.; Peterson, K. R.; Nielsen, P. E.; Kole, R.; Glazer, P. M. *Proc. Natl. Acad. Sci. U. S. A.* **2008,** *105,* 13514–13519. doi:10.1073/pnas.0711793105
312. McNeer, N. A.; Chin, J. Y.; Schleifman, E. B.; Fields, R. J.; Glazer, P. M.; Saltzman, W. M. *Mol. Ther.* **2011,** *19,* 172–180. doi:10.1038/mt.2010.200
313. Oyaghire, S. N.; Quijano, E.; Piotrowski-Daspit, A. S.; Saltzman, W. M.; Glazer, P. M. Poly(Lactic-*co*-Glycolic Acid) Nanoparticle Delivery of Peptide Nucleic Acids In Vivo. In *Methods in Molecular Biology;* Nielsen, P. E., Ed.; Methods in Molecular Biology, Vol. 2105; Humana Press: New York, NY, USA, 2020; pp 261–281. doi:10.1007/978-1-0716-0243-0_17
314. Xu, P.; Tong, Y.; Liu, X.-z.; Wang, T.-t.; Cheng, L.; Wang, B.-y.; Lv, X.; Huang, Y.; Liu, D.-p. *Sci. Rep.* **2015,** *5,* 12065. doi:10.1038/srep12065
315. Song, B.; Fan, Y.; He, W.; Zhu, D.; Niu, X.; Wang, D.; Ou, Z.; Luo, M.; Sun, X. *Stem Cells Dev.* **2015,** *24,* 1053–1065. doi:10.1089/scd.2014.0347
316. Li, Z.; Rana, T. M. *Nat. Rev. Drug Discovery* **2014,** *13,* 622–638. doi:10.1038/nrd4359
317. Cha, W.; Fan, R.; Miao, Y.; Zhou, Y.; Qin, C.; Shan, X.; Wan, X.; Cui, T. *Med. Chem. Commun.* **2018,** *9,* 396–408. doi:10.1039/c7md00285h
318. Wojciechowska, M.; Równicki, M.; Mieczkowski, A.; Miszkiewicz, J.; Trylska, J. *Molecules* **2020,** *25,* 559. doi:10.3390/molecules25030559
319. Good, L.; Sandberg, R.; Larsson, O.; Nielsen, P. E.; Wahlestedt, C. *Microbiology (London, U. K.)* **2000,** *146,* 2665–2670. doi:10.1099/00221287-146-10-2665
320. Gambari, R. *Expert Opin. Ther. Pat.* **2014,** *24,* 267–294. doi:10.1517/13543776.2014.863874

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Cationic oligonucleotide derivatives and conjugates: A favorable approach for enhanced DNA and RNA targeting oligonucleotides

Mathias B. Danielsen and Jesper Wengel*

Review

Open Access

Address:
Biomolecular Nanoscale Engineering Center, Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark

Email:
Jesper Wengel* - jwe@sdu.dk

* Corresponding author

Keywords:
antisense oligonucleotides; backbone modifications; cations; nucleobase modifications; sugar modifications

Beilstein J. Org. Chem. **2021,** *17,* 1828–1848.
https://doi.org/10.3762/bjoc.17.125

Received: 06 April 2021
Accepted: 14 July 2021
Published: 29 July 2021

This article is part of the thematic issue "Celebrating the role of chemistry in the success of oligonucleotides as therapeutics".

Guest Editors: P. Kumar and T. Brown

Abstract

Antisense oligonucleotides (ASOs) have the ability of binding to endogenous nucleic acid targets, thereby inhibiting the gene expression. Although ASOs have great potential in the treatment of many diseases, the search for favorable toxicity profiles and distribution has been challenging and consequently impeded the widespread use of ASOs as conventional medicine. One strategy that has been employed to optimize the delivery profile of ASOs, is the functionalization of ASOs with cationic amine groups, either by direct conjugation onto the sugar, nucleobase or internucleotide linkage. The introduction of these positively charged groups has improved properties like nuclease resistance, increased binding to the nucleic acid target and improved cell uptake for oligonucleotides (ONs) and ASOs. The modifications highlighted in this review are some of the most prevalent cationic amine groups which have been attached as single modifications onto ONs/ASOs. The review has been separated into three sections, nucleobase, sugar and backbone modifications, highlighting what impact the cationic amine groups have on the ONs/ASOs physiochemical and biological properties. Finally, a concluding section has been added, summarizing the important knowledge from the three chapters, and examining the future design for ASOs.

Introduction

Antisense oligonucleotides (ASOs) are single-stranded (ss) oligomers composed of typically 10–25 nucleotides linked by negatively charged phosphorus-based linkages. ASOs have the distinctive ability to bind endogenous nucleic acid targets in a sequence-specific manner, thereby inhibiting gene expression and offering opportunities for the treatment of a broad range of diseases. As ASOs interact with their RNA (or DNA) targets through complementary Watson–Crick base-pairing, the sequence options of ASO lead compounds can be rationalized based on a knowledge of the endogenetic gene sequence to

be targeted, thus further offering a potential against targets which are considered undruggable using conventional small-molecule drugs. The key features of ASOs further enable them to be transformed into personalized medicines, eventually even targeting patient-specific sequences and very rare diseases [1].

ASOs can mediate gene silencing via different mechanisms of action. ASOs that induce RNase H degradation of the endogenous RNA target generally are of the gapmer-design class (Figure 1), where a central segment of at least five DNA nucleotides termed the 'gap' is flanked by modified nucleotides that promote target binding and protection against exonucleolytic degradation [2]. Another class are the steric block ASOs that bind to the target with high affinity without inducing RNase H mediated degradation. Such ASOs are usually in part ('mixmers'), or in full, composed of nucleotides that structurally are incompatible with RNase H activity [3]. A limited number of ASOs has been approved by different agencies as medicines for the treatment of various diseases, such as Fomivirsen (1998, withdrawn), Mipomersen (2013), Eteplirsen (2016), Nusinersen (2016), Inotersen (2018), Golodirsen (2019), Volanesorsen (2019), Viltolarsen (2020), and Casimersen (2021) [3-5].

In practice, the design of ASOs that mediate efficient gene silencing without side-effects has turned out to be challenging. These side-effects have been shown to emerge due to off-target interactions [6-11], toxicities [12] or saturation of RNA-processing pathways [13]. Additionally, the delivery of ASOs to the target tissues or organs is a major hurdle that needs to be addressed before ASOs can find more widespread use [3,9-11,14]. A lack of efficient delivery of ASOs can be caused by various reasons such as degradation [15,16], insufficient endosomal escape [17], glomerular filtration [18], or binding to one or more proteins [19,20]. Notably, optimization of the therapeutic window of ASOs is closely related to improved delivery, and a variety of chemical strategies have been investigated in this context, such as ASO–lipid conjugates for improved endosomal escape [21], ASO–triantennary *N*-acetylgalactosamine (GalNAc) conjugates for improved liver targeting [22,23] and ASO–glucagon-like peptide-1 (GLP1) conjugation for improved delivery to pancreatic β-cells [24].

Previously, Menzi et al. reviewed the impact of cationic modifications and conjugations for ONs and siRNAs biophysical and biological activities until 2015 [25]. In this review, we focus on important monomeric cationic modifications for ASOs, including locked nucleic acid (LNA) monomers, and their synthesis. Such modifications have been achieved either by direct conjugation to the nucleobase, the sugar or the backbone of nucleotide monomers of such ASOs. In addition, a table design showing which modification has duplex stabilizing properties, as well as improved nuclease resistance and cell activity, has been chosen for optimal visual presentation.

This approach has spurred considerable interest since the introduction of positively charged groups results in ASOs with an overall reduced negative charge compared to the corresponding ASO without such groups. As the large number of negative charges of an ASO, i.e., $n-1$, if n is the number of nucleotides in an ASO with phosphodiester (PO) or phosphorothioate (PS) linkages, is assumed to contribute to the limited membrane permeability of ASOs. A reduction in the net negative charge may have beneficial delivery-related effects in addition to other possible effects such as improved resistance towards nuclease degradation or increased binding to the negatively charged RNA complements, the latter as a result of reduced electrostatic repulsion. In the following sections, a series of ASO-type oligonucleotides (ONs) which have been chemically modified with positively charged groups will be described, and their properties highlighted.

Review

ONs containing amine-group conjugates and nucleotide derivatives

Many parameters can affect the physicochemical properties of ASOs that have been modified with amine groups, where factors such as length and shape of the amine moiety, the

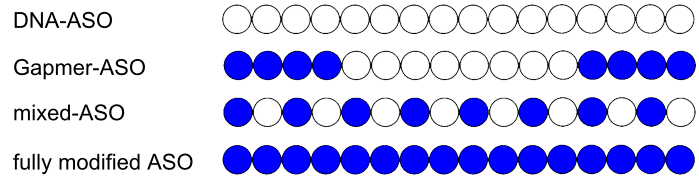


Figure 1: A schematic representation of 16-mer ASOs in different designs. White circles represent unmodified DNA monomers; blue circles represent nucleotide modifications. The gapmer-ASO shown is an example of a so-called 4-8-4 gapmer. The patterns of modified and unmodified nucleotides may vary and only examples are shown. Also, the phosphodiester linkages between the nucleotide monomers may be modified, and here phosphorothioate linkages are often used.

attachment site of the amine group (to the sugar, backbone, or nucleobase) and the position within the ON/ASO, i.e., the 5'-terminus, 3'-terminus, or in the center. The modifications chosen for inclusion into this review involve some of the more commonly used amine groups that have been attached as a single modification either on the nucleobase, sugar, or internucleotide linkage. When structurally depicted in this paper, modifications are shown with amine functionalities in their neutral, i.e., non-protonated form.

Regarding synthetic strategies, both the use of amine group functionalized phosphoramidites, i.e., functionalized monomeric building blocks, as well as conjugation with amine groups after completion of the ON assembly, are being discussed. Amino acids and cationic modifications that replace the core structure of the nucleobase, sugar, or the internucleotide linkage have been excluded.

Cationic amine-functionalized group substitutions at nucleobases

One strategy that has attracted a lot of interest is the attachment of cationic (poly)amine groups via the nucleobase on ASOs, thereby improving the RNA-binding affinity [26]. This strategy can be employed either on the nucleoside level, which requires many different nucleotide building blocks to be synthesized or via the so-called post-synthetic modification strategy of ONs. The latter strategy can be divided into the conjugation of amine groups onto ONs still attached to the solid support, or onto ONs in solution after cleavage from the solid support. The C-5 position of the pyrimidine ring has in general been the most used attachment point since it is not involved in hydrogen bonding and is facing the major groove upon duplex formation [27].

An illustrative approach that allowed the exploration of the 5-position of the pyrimidine ring as attachment site involved the use of 5-methoxycarbonylmethyl-2'-deoxyuridine and 5-trifluoroethoxycarbonyl-2'-deoxycytidine building blocks in the ON synthesis [28-30]. The corresponding modified ONs could be converted in a versatile manner to oligomers carrying the desired amine-functionalized groups at the 5-position on the pyrimidine nucleobase [28-30]. Similarly, the more reactive 5-cyanomethoxycarbonylmethyl-2'-deoxyuridine monomer has been used [27]. The reactivity of the above-mentioned chemical groups has enabled the attachment of various amine-functionalized groups onto the 5-position of pyrimidines via both the modified monomeric building blocks and post-synthetic ON chemistry [27,28,31-41]. The structures of some of the uridine derivatives are shown in Table 1. In general, the attachment of amine-functionalized groups on the pyrimidine C-5 position positively affects both the thermal stability and the nuclease resistance of the resulting amine-modified ONs. A tris-aminated derivative group, when conjugated to the C-5 position on 2'-deoxyuridine (Table 1B, **12**), improved antisense activity while reducing toxicity [39].

In addition, a 15-mer PS-ASO, modified with the C-5 tris-aminated 2'-deoxyuridine **12**, improved anti-HIV activity and reduced cytotoxicity relative to the unmodified PS-ASO [43]. It is important to notice that C-5 modifications, besides resulting in improved nuclease stability and providing improved target hybridization, allow for activation of RNase H. This was demonstrated by Matsuda and co-workers when they incorporated modification **2** or **3** into a 17-mer ON containing a stretch of four DNA nucleotides in the middle, flanked by the modifications in a ´mixmer´ design, which is important for designing gapmer ASOs [31].

Table 1: Amine-functionalized groups on the nucleobase.[a]

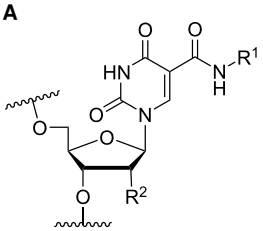
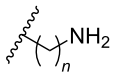
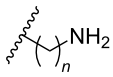
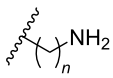
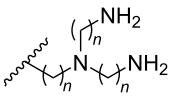
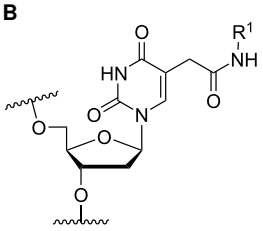
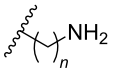
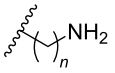
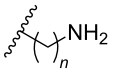
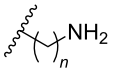
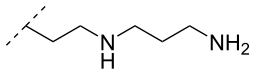
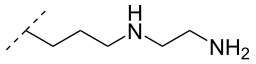
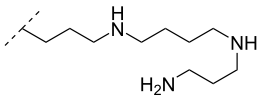
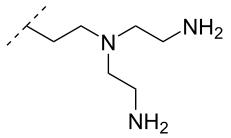
base modifications	R^1		n R^2	Ref.	thermo-stability	nuclease resistance	activity in cell
A 		**1**	n = 2 R^2 = H	[28,30]	+	+	n.d.
		**2**	n = 6 R^2 = H	[28,30-32]	+	+	n.d.
		**3**	n = 6 R^2 = OMe	[31,33]	+	+	n.d.
		**4**	n = 2 or 3	[37]	+	+	n.d.

Table 1: Amine-functionalized groups on the nucleobase.[a] (continued)

B 		**5**	*n* = 2	[34,35]	+	+	n.d.
		**6**	*n* = 5	[27]	n.d.	n.d.	n.d.
		**7**	*n* = 6	[34-36]	+	+	n.d.
		**8**	*n* = 7	[27]	n.d.	n.d.	n.d.
		**9**		[35,38]	+	n.d.	n.d.
		**10**		[38]	+	n.d.	n.d.
		**11**		[40,41]	+	n.d.	n.d.
		**12**		[34,39,42,43]	+	+	X

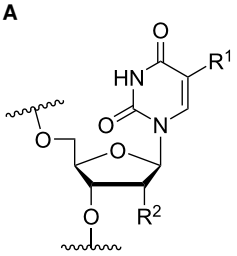
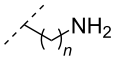
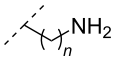
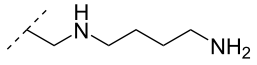
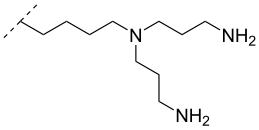
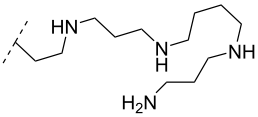
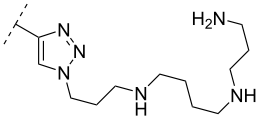
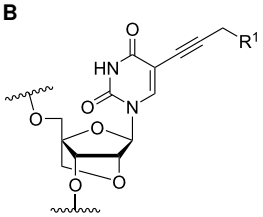
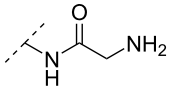
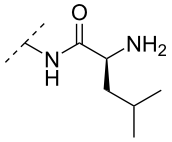
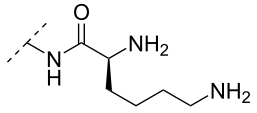
[a]A '+' sign has been added when the modified ON/ASO showed improved thermal stability (T_m) either towards ssDNA, ssRNA, or double-stranded (ds) DNA, and when the nuclease stability for the modified ON/ASO demonstrated improved stability, all compared to the effects mediated by control DNA or RNA strands. The 'X' sign has been added when the modified ASO demonstrated equal or better gene inhibitory activity in cells relative to the ASO control, while n.d. = not determined.

Another well-established method for C-5 pyrimidine modification involves the Sonogashira cross-coupling reaction between an alkyne group and a 5-iodo-modified nucleobase/nucleoside followed, if desired, by reduction [44] to give a more flexible group, or the alkyne group can be retained, depending on the modification needed [45-47]. This method has been extensively used to study various modifications, and some of them can be seen in Table 2 (A and B) [44,48-54]. Interestingly, when ONs were modified with C-5 amino acid-functionalized LNA nucleotides **20**–**22**, significant increases in the melting temperatures (T_m) were measured with up to 14 °C for modification **22** towards complementary RNA, relative to the unmodified DNA-ON. Furthermore, this was 5.5 °C higher than the ON modified with LNA. It is important to mention that the positioning of the modification in the 9-mer strand had a significant impact on the stability of the corresponding duplex with its RNA complement. All three modifications **20**–**22** showed better hybridization properties than LNA-thymidine; however, only modification **22** gave significant increases in T_m relative to modification **19** used as control. This finding was ascribed to both the extended π-conjugation of the alkynyl-functionalized nucleobase and stabilizing electrostatic interactions [54]. Positioning the modifications near the 3'-terminus increased the resistance toward 3'-exonuclease degradation relative to both the unmodified and the LNA-modified ONs [54].

Although a more simple modification regarding the chemical composition, the C5-aminopropynyl-functionalized LNA **19** has shown good duplex-stabilizing properties with up to 13 °C per modification towards RNA [47] while conferring also high triplex stability [55]. A further investigation demonstrated that C5-aminopropynyl-functionalized LNA, after being introduced into so-called bisLNAs (triplex-forming ONs (TFOs) linked to a Watson–Crick interacting ON, both targeting the same ssDNA) exhibited the ability to invade double-stranded DNA (dsDNA) targets in vitro [56].

The functionalization with aminoalkyl variants onto the nucleobase is not limited to the C-5 position on the pyrimidine base. Another important site is the C-4 position and/or a combination of both as deposed in Table 3 (A and B), with a selection of amine modifications attached. One of these is the cytosine ana-

Table 2: Amine-functionalized groups on the nucleobase.[a]

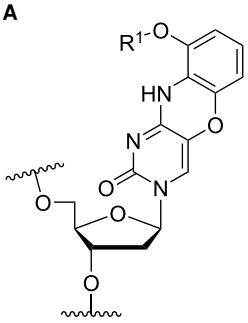
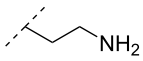
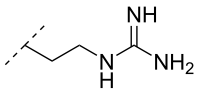
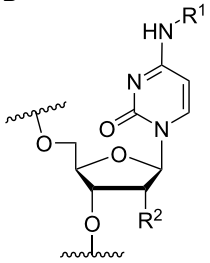
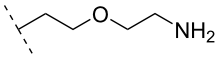
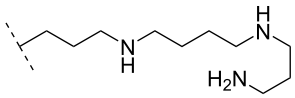
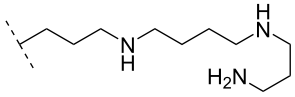
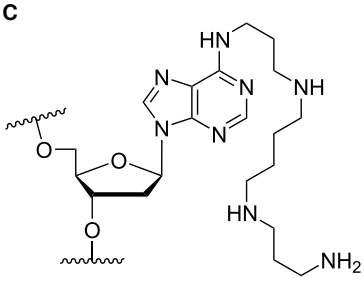
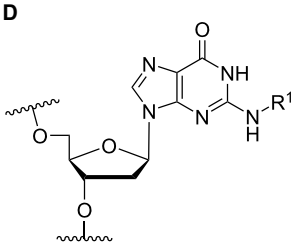
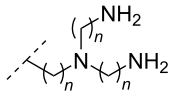
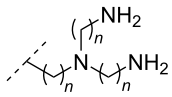
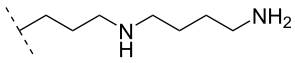
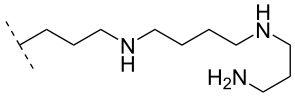
base modifications	R^1		n/R^2	ref.	thermo-stability	nuclease resistance	activity in cell
A 		**13**	n = 3, R^2 = H	[48-50]	+	n.d.	n.d.
		**14**	n = 6, R^2 = H	[44]	+	n.d.	n.d.
		**15**	R^2 = H	[51]	+	+	n.d.
		**16**	R^2 = H	[52]	+	+	n.d.
		**17**	R^2 = OH	[53]	n.d.	+	n.d.
		**18**	R^2 = OMe	[62]	+	n.d.	n.d.
B 		**19**		[47,55,56]	+	+	n.d.
		**20**		[54]	+	+	n.d.
		**21**		[54]	+	+	n.d.
		**22**		[54]	+	+	n.d.

[a]A '**+**' sign has been added when the modified ON/ASO showed improved thermal stability (T_m) either towards ssDNA, ssRNA or dsDNA, and when the nuclease stability for the modified ON/ASO demonstrated improved stability, all compared to the effects mediated by control DNA or RNA strands. The '**X**' sign has been added when the modified ASO demonstrated equal or better gene inhibitory activity in cells relative to the ASO control, while n.d. = not determined.

logue termed G-clamp (modification **23**) which increased the T_m of a DNA duplex by 18 °C when incorporated centrally into a decamer ON (Table 3A) [57]. The G-clamp modification was later observed to have antisense inhibition activity involving RNase H cleavage with a single incorporation into a PS-ON [58]. Afterwards, a guanidino-G-clamp (modification **24**) was synthesized to increase the number of hydrogen bonds that could be established between the modified nucleobase and the corresponding guanidine, which resulted in an increase in T_m of 16 °C, i.e., in the same range as obtained with the original G-clamp (Table 3A) [59].

Generally, conversions of nucleoside phosphoramidite synthons have been explored only rarely. However, the commercially available 3'-phosphoramidite derivative of 5'-*O*-dimethoxytrityl-2'-*O*-methyluridine could be converted into an N^4-triazole-modified 2'-OMe-cytidine phosphoramidite [60]. This concept was later used to prepare spermine-functionalized

Table 3: Amine-functionalized groups on the nucleobase.[a]

base modification	R^1		n/R^2	ref.	thermo-stability	nuclease resistance	activity in cell
A 		**23**		[57,58]	+	n.d.	**X**
		**24**		[59]	+	n.d.	n.d.
B 		**25**	R^2 = H	[57]	+	n.d.	n.d.
		**26**	R^2 = OMe	[61]	+	n.d.	n.d.
		**27**	R^2 = O-MOE	[61]	+	n.d.	n.d.
C 		**28**		[63-65]	n.d.	n.d.	n.d.
D 		**29**	n = 2	[70,71]	+	+	n.d.
		**30**	n = 3	[70,71]	+	+	n.d.
		**31**		[68]	+	n.d.	n.d.
		**32**		[64-69]	+	n.d.	n.d.

[a]A '**+**' sign has been added when the modified ON/ASO showed improved thermal stability (T_m) either towards ssDNA, ssRNA or dsDNA, and when the nuclease stability for the modified ON/ASO demonstrated improved stability, all compared to the effects mediated by control DNA or RNA strands. The '**X**' sign has been added when the modified ASO demonstrated equal or better gene inhibitory activity in cells relative to the ASO control, while n.d. = not determined.

2'-OMe and 2'-*O*-((2-methoxy)ethyl) (MOE)cytidine phosphoramidites as building blocks for incorporation into PS-ONs (Table 3B, **26**, **27**). In this study, two modifications of the N^4-spermine-modified 2'-*O*-MOE-cytidine monomer **26** was incorporated into a 12-mer PS-ON centrally and at the 5'-end, resulting in significant increases in T_m (by more than 16 °C) towards complementary RNA [61]. In general, the study showed that the modifications had a positive effect on T_m for the formed ON/RNA duplex for all ONs substituted with cytidine monomers **26** or **27**.

These findings were also in agreement with an earlier study where the conjugation of a spermine-substituted triazole group at the C-5 position of a 2'-OMe-uridine monomer, positioning the group in the major groove of the formed duplex, had a significant stabilizing effect on RNA binding (Table 2A, **18**) [62]. Another method for introducing spermine at the N^4 position involved the reaction of spermine and 4-*N*-*p*-toluenesulfonyl-5'-*O*-dimethoxytrityl-2'-deoxycytidine followed by ON synthesis [63].

A less investigated strategy is the anchoring of amine functionalities onto purines. Beginning with adenine, the N^6 position has been the most explored attachment point. For example, a 2'-deoxy-N^6-triazole-substituted adenosine monomer (9-(5'-*O*-dimethoxytrityl-2'-deoxy-β-D-erythro-pentofuranosyl)-N^6-(1,2,4-triazol-4-yl)adenine) was reacted with spermine to yield 5'-*O*-dimethoxytrityl-N^6-(4,9,13-triazatridecane-1-yl)-2'-deoxyadenosine (**28**) and subsequently incorporated into an ON (Table 3C) [63-65].

A different approach must be taken when conjugating amine moieties onto guanine. Interestingly, in a post-oligo synthetic modification approach, a 2-fluoro-6-*p*-nitrophenylethyl-2'-deoxyinosine-3'-phosphoramidite monomer was incorporated twice into an 11-mer ON, whereupon spermine was attached to the 2-position of the purine simultaneously with cleavage from the solid support [66]. It was found that the modified ON gave an improved duplex stability relative to the unmodified ON by 15 °C at 150 mM NaCl [66]. Later, improved syntheses of the phosphoramidite derivatives of guanine analogues have been developed [67]. Thus, many studies have been carried out for C2-spermine modified 2'-deoxyguanosine (Table 3D), and in general, when the above modification is incorporated on the C2 position, a positive and cooperative stability-enhancing effect is observed for the duplex formation between the modified ON and the targeted complementary strands [66,68,69]. Additionally, when a shorter group is introduced (spermidine, **31**, Table 3D) into a 12-mer ON, at the 5-position and at the 5'-end, a similar increase in duplex stability was observed, i.e., +22 °C for modification **31** and +21 °C for modification **32** towards complementary DNA (150 mM NaCl) [68].

It has been demonstrated that C-2 modified guanidine analogues containing nor-spermidine (**30**) and the shorter diethylenetriamine (**29**) (Table 3D) could be synthesized via the C2-fluoro modified monomer. This resulted in ONs with slightly higher T_m (approximately 3 °C for **29** and **30**) when having one incorporation relative to their unmodified versions. Additionally, the modified ONs exhibited enhanced nuclease resistance [70,71]. A recent review has recently been published with a more extensive coverage of the post-synthetic ON functionalizations [72].

Cationic amine-functionalized group moieties attached to the sugar scaffold

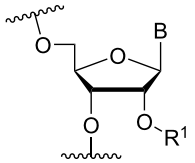
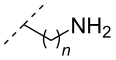
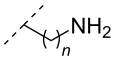
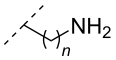
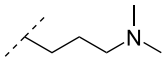
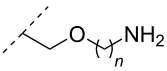
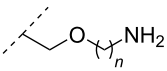
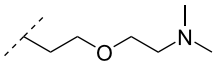
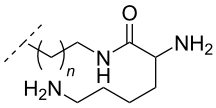
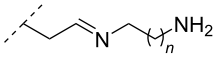
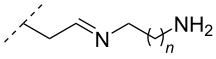
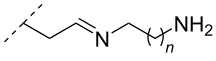
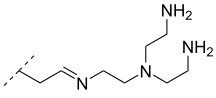
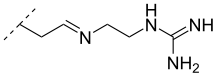
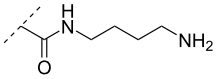
The sugar moiety of ONs has been extensively studied with respect to the significance of structure and configurations of substituents, and the resulting conformations of the furanose ring, on the properties of ONs (including ASOs). The great variation at which substituents can be positioned has led to the investigation of the impact of cationic amine-functionalized groups on the biophysical properties of the resulting ONs.

One position of the sugar moiety that has been explored in detail is the 2'-position. Modifications at this site have resulted in highly therapeutically relevant monomers like 2'-OMe- and 2'-*O*-MOE-RNA [3]. Additionally, the attachment of amine-functionalized groups at the 2'-position can be readily performed and allows for the amine-functionalized group to be positioned into the minor groove of the ON duplex [73].

Aminoalkyl functionalization of the furanose sugar moiety is mainly achieved through two major pathways. The first is the synthesis of cationic nucleotide derivatives, while the second proceeds via post-synthetic chemistry on a susceptible ON. The latter strategy, in theory, allows for a more versatile approach in testing the effects of different amine-functionalized groups.

Beginning with the 2'-*O*-alkylated RNA nucleotides (Table 4) it has been shown that the introduction of an aminopropyl group via 2'-*O*-alkylation (modification **34**) leads to moderately improved hybridization properties for the modified ON against its RNA complement and improved nuclease resistance. Additionally, a 20-mer PS-ASO with nine incorporations of modification **34** near the 3'-end was introduced into A549 cells via the electroporation method to induce c-*raf* mRNA and protein knockdown. Improved activity of the modified PS-ASO relative to the unmodified PS-ASO was observed: a ten-fold higher concentration of the control PS-ASO was needed to obtain a similar knockdown effect relative to the modified PS-ASO [74].

Table 4: Amine-functionalized groups attached to the sugar scaffold.[a]

Sugar modifications	R^1		*n*	ref.	thermo-stability	nuclease resistance	activity in cell
		**33**	*n* = 2	[77]	n.d.	n.d.	**X**
		**34**	*n* = 3	[74,77,78,81]	+	+	**X**
		**35**	*n* = 6	[75]	n.d.	+	n.d.
		**36**		[78]	+	+	n.d.
		**37**	*n* = 2	[73,77]	n.d.	n.d.	**X**
		**38**	*n* = 3	[73,77]	n.d.	n.d.	**X**
		**39**		[79]	+	+	n.d.
		**40**	*n* = 5	[76,82,83]	+	+	**X**
		**41**	*n* = 1	[80]	+	n.d.	n.d.
		**42**	*n* = 2	[80]	+	n.d.	n.d.
		**43**	*n* = 3	[80]	+	n.d.	n.d.
		**44**		[80]	+	n.d.	n.d.
		**45**		[80]	+	n.d.	n.d.
		**46**		[84]	n.d.	+	n.d.

[a]A '**+**' sign has been added when the modified ON/ASO showed improved T_m either towards ssDNA, ssRNA or dsDNA and when the nuclease stability for the modified ON/ASO demonstrated improved stability compared to the DNA or RNA control. The '**X**' sign has been added when the modified ASO demonstrated equal or better activity in cells relative to the ASO control, while n.d. = not determined.

Extending to aminohexyl (monomer **35**) resulted in a small decrease in duplex stability relative to the native ON, whereas the high nuclease resistance was maintained [75]. The attachment of a lysine onto the aminohexyl residue resulted in a lysyl-aminohexyl group (monomer **40**) which displayed a gradual increase in T_m upon incorporation of up to three modifications, and also improved the resistance against nuclease degradation relative to the native ON. Additionally, compared to wild-type ONs or siRNA, ASOs carrying three modifications resulted in an equal or higher downregulation of ICAM-1 expression [76].

A large study including 2'-aminoethyl RNA (monomer **33**) showed that two incorporations at the 3'-end of a 22-mer antisense strand of a siRNA had better eGFP gene silencing activity compared to the control siRNA with a single 2'-OMe RNA monomer near the 5'-end [77]. Conversion of the primary

amine on the aminopropyl modification into a tertiary amine (monomer **36**) and insertion four times in a 16-mer ON (two in the middle and one near either end) resulted in modified ONs having a high binding affinity towards RNA relative to the affinity of the DNA-control. Afterwards, a modified 19-mer ON carrying four copies of modification **36** near the 3'-end demonstrated high nuclease stability, as also observed with modification **34** [78].

In an attempt to introduce high yielding phosphoramidite building blocks suitable for automated ON synthesis, 2′-*O*-aminoethoxymethyl and 2′-*O*-aminopropoxymethyl nucleotides were developed. This method introduced the primary amine functionality through an azide reduction [73]. The corresponding monomers **37** and **38** improved the silencing activity of a siRNA when incorporated into the passenger strand (in the eGFP assay mentioned above). However, a decrease in the silencing activity was observed when incorporated into the guide strand [77]. To design a 2'-*O*-MOE cationic analogue, the 2'-*O*-(2-(2-(*N*,*N*-dimethylamino)ethoxy)ethyl) monomer **39** has been prepared and shown to moderately enhance RNA affinity and induce high nuclease resistance, similar to that of modification **34** [79].

Optimizing the triplex stability of complexes formed between modified TFOs and their dsDNA target is an important direction of research. This has been explored utilizing the reactivity between primary amines and the aldehyde moiety of a 2'-*O*-(2-oxoethyl)uridine nucleotide, incorporated centrally in an 11-mer TFO, to form a Schiff base (monomers **41**–**45**) [80]. All aminoalkylated moieties improved the triplex stability. Notably, a significant improvement in T_m was observed when the functionalizing groups were changed from ethylenediamine to either trimethylenediamine (monomer **42**) or putrescine (monomer **43**), demonstrating that the length between the cationic amino group and the sugar scaffold is important for the thermal stability effects. The best stabilization was obtained for the 2-(aminoethyl)guanidine monomer **45** and tris(2-aminoethyl)amine (monomer **44**) variants [80].

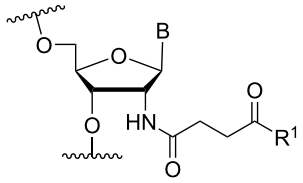
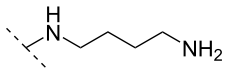
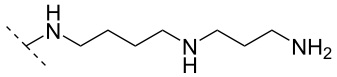
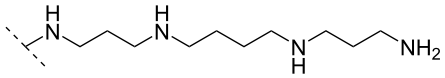
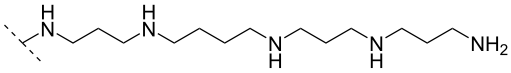
Another chemical group utilized for the 2'-modification is 2'-*O*-carbamoyl [85-87]. However, it has proven difficult to stabilize the duplex formed between the modified ON and targeted DNA/RNA with cationic aminoalkylated groups [85], which is thought to be due to the close contact between the carbonyl of the 2'-*O*-carbamoyl substituent and the O-2 on the nucleobase [88]. To circumvent this issue, the 2'-*O*-(*N*-(4-aminobutylcarbamoyl))uridine monomer **46** has been synthesized [84]. When incorporated centrally in a 2'-OMe-RNA ON, a significant stabilization relative to that of 2′-*O*-carbamoyluridine was observed against the RNA target. However, relative to the control ON carrying pure 2'-OMe RNA modified ONs, monomer **46** had an affinity-lowering effect [84]. A nuclease stability assay showed a clear improvement relative to the control and the simple 2′-*O*-carbamoyluridine modification [84].

A different approach to introduce aminoalkyl groups at the 2'-position was achieved via a benzyl protected 2'-succinylamido-2'-deoxyuridine building block attached either to a solid support or incorporated using conventional phosphoramidite chemistry. Previously such a method had been used to attaching different moieties onto an ON still bound to a solid support [89,90]. Putrescine (**47**), spermidine (**48**), spermine (**49**) and a synthetic pentaamine (**50**) were attached to the 2'-position (Table 5).

Interestingly the resulting nucleotides were found to adopt a conformation which was, with respect to duplex stability, tolerated at terminal positions but not at internal positions [91]. The modified PS-ASOs were transfected into human-607B melanoma cells, and after 48 h, the B-cell lymphoma 2 (bcl-2) protein levels were examined. All PS-ASOs gave improved downregulation relative to the control (scrambled sequence), but only PS-ASOs carrying the nucleotides modified with spermine (i.e., monomer **49**) or pentaamine (i.e., monomer **50**) induced improved downregulation of gene expression relative to the downregulation of the reference ASO (Oblimersen) [91].

In an entirely different approach, the 2'-amino group of amino-LNA-thymine (amino-LNA-T) has been explored as an attachment point for various cationic groups. As one example, amino acids such as glycine, lysine, and proline in different combinations have been attached to the 2'-amino group with substantial success regarding duplex stability [92]. The 2'-amino-LNA scaffold has further been modified with amine-functionalized groups at the nucleoside level creating different nucleotide building blocks for ON synthesis, or at the ON level via post-ON synthesis conjugation chemistry. The first method was used to attach 1-piperazinepropionic acid through an amide coupling onto 2'-amino-LNA. This monomer (**51**) induced high binding affinity towards complementary targets upon incorporation into a 9-mer ON. In DNA, an increase of 7.0 °C and 17.5 °C for one and three incorporations, respectively, was observed and in RNA, an increase of 9.0 °C and 24.5 °C for one and three incorporations, respectively, was observed relative to the DNA 9-mer control strand. Additionally, a high nuclease resistance compared to the DNA control was observed [93]. In a follow-up study, after being introduced into bisLNAs, the piperazino-modified 2'-amino-LNA-T nucleotide was compatible with invasion into dsDNA targets in vitro [56]. In further studies utilizing the same amide coupling, nor-spermidine with different group lengths (**52** and **53**), a glycol-amine-functionalized

Table 5: Amine-functionalized groups attached to the sugar scaffold.[a]

sugar modifications	R^1		ref.	thermo-stability	nuclease resistance	acitivty in cell
		**47**	[91]	n.d.	n.d.	**X**
		**48**	[91]	n.d.	n.d.	**X**
		**49**	[91]	+	n.d.	**X**
		**50**	[91]	+	n.d.	**X**

[a]A '+' sign has been added when the modified ON/ASO showed improved T_m either towards ssDNA, ssRNA or dsDNA and when the nuclease stability for the modified ON/ASO demonstrated improved stability compared to the DNA or RNA control. The '**X**' sign has been added when the modified ASO demonstrated equal or better activity in cells relative to the ASO control, while n.d. = not determined.

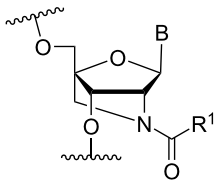
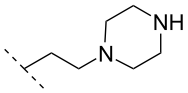
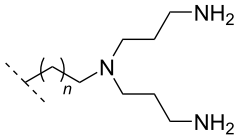
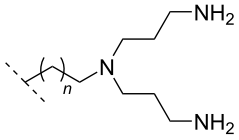
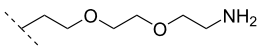
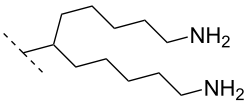
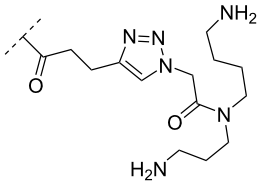
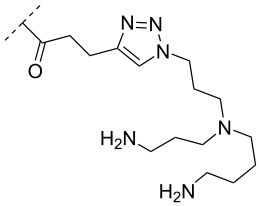
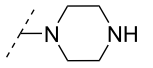
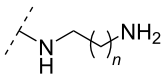
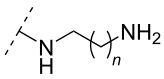
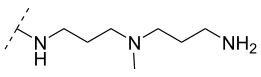
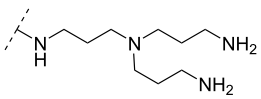
group (**54**), and a bis-C6-amine-functionalized group (**55**) were attached to 2'-amino-LNA-T. All these modifications demonstrated high duplex stabilizing capabilities combined with high nuclease resistance [94,95]. Additionally, the nor-spermidine and amino-glycol modified 2'-amino-LNA-T when incorporated into TFOs all induced excellent triplex stability at pH 7.0 [94] (Table 6).

To circumvent the laborious work related to the monomers described above, post-ON conjugation via click-chemistry was utilized to attach two different spermidine analogues, carrying either two (**56**) or three (**57**) positive charges, onto the 2'-amino-LNA scaffolds. This was demonstrated to be a successful design as both monomers showed very high duplex stabilizing properties towards RNA (T_m +10.0 °C) and DNA (T_m +8.5 °C). Additionally, nuclease resistance was shown to be high [95] which was in agreement with other 2'-aminoalkylated-LNA monomers [94]. The triplex stability of the monomers was determined, and the spermidine variant carrying three cationic charges (**57**) had the highest triplex stabilizing effect [95]. Interestingly, at biologically relevant pH (7.0), two incorporations of the nor-spermidine variant (**53**) [94], and the triazole-linker variant carrying three cationic charges (**57**) [95] stabilized the formed triplex by 28.0 °C and 30.5 °C, respectively [94,95]. These cationic 2'-amine-functionalized LNA modifications (**51**–**57**) all gave high binding affinity towards RNA with excellent nuclease resistance, making them ideal ASO modifications as only a limited number of modifications is needed for a substantial effect.

Recently, a versatile method of post-ON synthesis conjugation, different from the click chemistry method, has been applied to 2'-amino-LNA. A class of 2'-urea-LNA analogues (**58–62**) has been prepared by reacting various amine-functionalized groups with a 2'-*N*-pentafluorophenoxycarbonyl-2'-amino-LNA monomer already incorporated into an ON. All modifications improved hybridization towards both DNA and RNA complements when compared to natural DNA nucleotides [96].

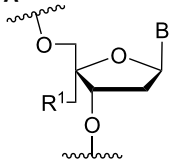
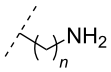
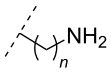
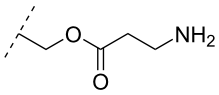
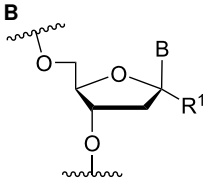
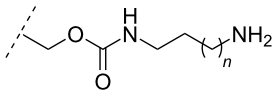
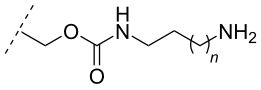
The attachment of an aminoalkyl-group to the 4'-position (Table 7A) offers an advantage since this site is in close proximity to the backbone, which potentially allows the basic amino group via a relative short linking group to engage electrostatically with the acidic phosphodiester moiety [97]. This was initially explored for ONs containing 4′-*C*-(aminomethyl)thymidine (monomer **63**) [98,99]. Later, the amine-functionalized group was expanded into the monomer 4′-*C*-(2-((*N*-(2-aminoethyl)carbamoyl)oxy)ethyl)thymidine (not shown) that was shown to display improved resistance against endo- and exonuclease cleavage relative to the DNA control ON [100]. This was followed by 4′-*C*-amidoethyl- (**64**) and 4′-*C*-amidopropylthymidine (**65**) derivatives which continued the trend of good nuclease stability. In general, all the 4'-substituted nucleotides (**63–66**) mentioned here stabilized duplexes formed with DNA complements with modification **64** giving the best stabilization (up to 5.7 °C for four incorporations in an 18-mer ON relative to the control), whereas all gave similar and/or decreased thermal stability against RNA relative to the natural control ON [97]. It is noteworthy to mention that all monomers

Table 6: Amine-functionalized groups attached to the sugar scaffold.[a]

sugar modifications	R^1		n/R^2	ref.	thermostability	nuclease resistance	activity in cell
		**51**		[93]	+	+	n.d.
		**52**	n = 1	[94]	+	+	n.d.
		**53**	n = 2	[94]	+	+	n.d.
		**54**		[94]	+	+	n.d.
		**55**		[95]	+	+	n.d.
		**56**		[95]	+	+	n.d.
		**57**		[95]	+	+	n.d.
		**58**		[96]	+	n.d.	n.d.
		**59**	n = 1	[96]	+	n.d.	n.d.
		**60**	n = 2	[96]	+	n.d.	n.d.
		**61**		[96]	+	n.d.	n.d.
		**62**		[96]	+	n.d.	n.d.

[a]A '**+**' sign has been added when the modified ON/ASO showed improved T_m either towards ssDNA, ssRNA or dsDNA and when the nuclease stability for the modified ON/ASO demonstrated improved stability compared to the DNA or RNA control. The '**X**' sign has been added when the modified ASO demonstrated equal or better activity in cells relative to the ASO control, while n.d. = not determined.

Table 7: Amine-functionalized groups attached to the sugar scaffold.[a]

sugar modifications	R^1		n/R^2	ref.	thermo-stability	nuclease stability	activity in cell
A 		**63**		[97-99]	+	+	n.d.
		**64**	n = 1	[97]	+	+	n.d.
		**65**	n = 2	[97]	+	+	n.d.
		**66**		[97]	+	+	n.d.
B 		**67**	n = 2	[101]	+	+	n.d.
		**68**	n = 4	[101]	n.d.	n.d.	n.d.

[a]A '**+**' sign has been added when the modified ON/ASO showed improved T_m either towards ssDNA, ssRNA or dsDNA and when the nuclease stability for the modified ON/ASO demonstrated improved stability compared to the DNA or RNA control. The '**X**' sign has been added when the modified ASO demonstrated equal or better activity in cells relative to the ASO control, while n.d. = not determined.

allowed the design of ASOs that were substrates for RNase H [97].

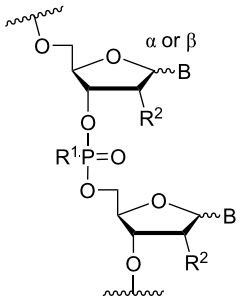
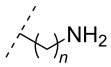
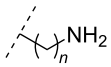
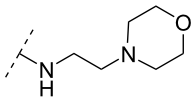
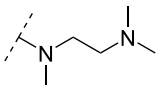
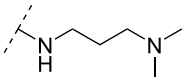
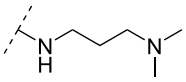
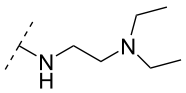
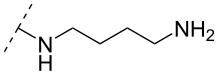
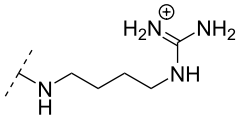
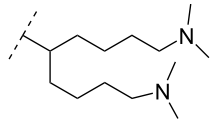
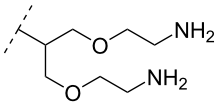
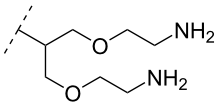
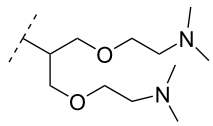
The 1'-position on the furanose ring has been studied to a lesser degree regarding the functionalization by amine-containing moieties. This site allows for the substitution to be positioned towards the minor groove. To develop a new way of attaching various functional groups onto the ONs without disturbing duplex formation, Matsuda and co-workers developed some 2'-deoxyuridine analogues carrying aminoalkyl groups at the 1'-position (Table 7B). These were intended to be used as an attachment point, and the 1'-aminobutane variant of the 2'-deoxyuridine analogue (monomer **67**) was also tested for the duplex forming capabilities. Here it was observed that this modification positioned at the 5'-end in a poly-T ON stabilized the duplex (+3.0 °C relative to the control ON) whereas a central insertion of modification **67** or **68** had a neutral or slightly negative effect on the T_m value relative to the control ON [101].

Cationic amine-functionalized group functionalities as internucleoside linkage

Although previous research has highlighted the relevance of the phosphodiester-linked backbone in the overall function of nucleic acids [102,103], many researchers have still sought to change the properties of ONs, i.e., enzymatic stability, hybridization, biodistribution, and cell-uptake, via the introduction of non-natural internucleoside linkages. The most well-known modification is the phosphorothioate-linked backbone, which is known to enhance not only nuclease resistance but also protein interactions compared to the phosphodiester backbone [104]. In an effort to reduce the overall negative charge of the backbone, also a large number of different artificially linked backbone ONs has been synthesized, and a selected number of these can be seen in Table 8 and Table 9. The latter approach of modifying the internucleotide linkage is unique as it may reduce in part or in full the negative charge of ONs, including ASOs.

One strategy that has been employed is the aminoalkyl phosphoramidate linkage (Table 8). Pioneering work was done by Letsinger and co-workers in 1986, who reported the synthesis of a 2'-deoxyadenosyl dinucleotide linked via an aminoethyl phosphoramidate linkage that was positively charged under neutral to acidic conditions [105]. A subsequent work based on these findings resulted in the synthesis of short cationic DNA ONs linked via *N*-alkylated phosphoramidate linkages (Table 8) [106]. In contrast to the dinucleotide which was synthesized by solution phase chemistry [105], the modified ONs were synthesized on solid-support employing H-phosphonate chemistry, followed by the oxidative coupling with the appropriate diamines to give the desired *N*-ethyl-2-morpholino- (monomer

Table 8: Amine-functionalized groups as internucleoside linkage.[a]

backbone modifications	R1		α/β	n/R2	ref.	thermo-stability	nuclease resistance	activity in cell
		**69**	β	*n* = 1 R2 = H	[113]	+	+	n.d.
		**70**	β	*n* = 2 R2 = H	[114]	+	+	n.d.
		**71**	α/β	R2 = H	[106]	+	+	n.d.
		**72**	α/β	R2 = H	[106]	+	+	n.d.
		**73**	α	R2 = H	[107,109,110,112]	+	n.d.	**X**
		**74**	α/β	R2 = OMe	[107]	+	n.d.	n.d.
		**75**	α	R2 = H	[108,111,116]	+	+	**X**
		**76**	α	R2 = H	[112]	+	n.d.	n.d.
		**77**	α	R2 = H	[112]	+	n.d.	n.d.
		**78**	β	R2 = H	[115]	n.d.	n.d.	n.d.
		**79**	β	R2 = H	[115]	+	n.d.	n.d.
		**80**	β	R2 = OMe	[115]	n.d.	n.d.	n.d.
		**81**	β	R2 = OMe	[115]	+	+	**X**

[a]A '**+**' sign has been added when the modified ON/ASO showed improved thermal stability (T_m) either towards ssDNA, ssRNA or dsDNA, and when the nuclease stability for the modified ON/ASO demonstrated improved stability, all compared to the effects mediated by control DNA or RNA strands. The '**X**' sign has been added when the modified ASO demonstrated equal or better gene inhibitory activity in cells relative to the ASO control, while n.d. = no determined.

71) and *N*-methyl-2-(dimethylamino)ethyl (monomer **72**) phosphoramidate linkages [106]. How these *N*-alkylated phosphoramidate-linked ONs interacted with the complementary DNA was highly dependent on the ionic strength and the pH of the relevant medium. An inverse effect between hybridization stability and salt concentration was observed for cationic ONs when compared to their anionic counterparts. The study demonstrated a decrease in hybridization for the phosphoramidate-modified ONs towards complementary DNA when a high salt concentration (1.0 M NaCl) was used, caused by electrostatic shielding mediated by the salt ions [106]. Additionally, this class of ONs showed high resistance towards nuclease-catalyzed degradation [106].

Subsequently, two variants of this phosphoramidate linker strategy were synthesized, one being the *N,N*-(dimethylaminopropyl)phosphoramidate linkage (monomers **73** and **74**) (DMAP) [107], and the other the *N,N*-diethyl-ethylenediamine phosphoramidate linkage (**75**) (DEED) [108]. Stereo-uniform ONs (either R_p or S_p) containing the DMAP modification were synthesized via dinucleotide derivatives obtained by a phosphitylation reaction followed by oxidative amidate coupling to create an epimeric mixture of the dinucleotide phosphoramidate-linked derivatives with subsequent separation of the two stereoisomers. These were then incorporated into the desired ONs after O3'-desilylation and phosphitylation of the dimers [107]. The authors found that for each of the sequences investigated in the study, one phosphoramidate stereoisomer induced improved hybridization towards targeted DNA, while the other stereoisomer induced lower affinity, all relative to their corresponding phosphodiester control ON [107].

Vasseur, Debart and co-workers introduced the DMAP modification into α-configured ONs [109,110]. These zwitterionic or cationic α-ONs hybridized with high affinity to their complementary DNA and RNA targets, while a significant impairment in hybridization was found when mismatches were introduced [109,110]. This was especially seen in case of a 12-mer α-ON containing 11 modifications of the DMAP linkage, giving the ON an overall net charge of +11, resulting in a T_m increase of 27.0 °C for the ON/DNA duplex and 10.4 °C for the ON/RNA duplex [110]. A fully DMAP-modified 18-mer α-ASO was incubated with HePG2 hepatoma cells, inhibiting firefly luciferase activity in a dose-depending manner in a whole cell assay, while the scrambled control showed no effect [110]. Interestingly, this effect was observed without any transfection agents. However, the isosequential 2'-OMe ASO and the methoxyethylphosphoramidate (PNHME) ASO (with a neutral backbone) showed no activity when they were incubated without a transfection agent [110]. The β-configured ON variants resulted in a decrease in T_m towards their complementary RNA and DNA targets, which was ascribed to increased steric hindrance [109], although an improved triplex stability for the 2'-OMe RNA phosphoramidate variant was observed (**75**).

The cationic phosphoramidate variant termed DEED was originally tested for its triplex-forming capabilities. The authors found that ONs containing the DEED modification were more capable at forming triplexes under conditions that approximated the magnesium, pH, and potassium levels found in vivo [108]. A later study conducted by Weeks and co-workers reported that a TFO modified with the DEED modification could efficiently inhibit the expression of plasmid DNA injected into *Xenopus oocytes* [111]. The study demonstrated that a sufficiently long mismatch-free DNA target needed to be present for the modified TFO to work effectively, thus demonstrating the significance of sequence-specific binding. It was however important that the TFO and plasmid were mixed prior to injection, to get near-complete inhibition of gene expression. Only partial inhibition could be observed, if TFOs were injected before the plasmid, and no inhibition could be observed when the plasmid was injected first. This indicated that a competition between the cationic TFOs and the histones for DNA binding had a large impact [111].

The library of phosphoramidate variants was expanded when the aminobutyl phosphoramidate and the guanidinobutyl phosphoramidate were synthesized [112]. A facile post-synthetic method was successfully employed to convert amine functionalities attached to ONs into guanidinium tethers. Both, the aminobutyl (**76**) and guanidinobutyl (**77**) modifications were introduced into α-ONs which elicited significant stabilization of the formed ON/DNA and ON/RNA duplexes. Interestingly, towards RNA complements, these modifications resulted in more pronounced increases in T_m, relative to the DMAP modification mentioned above, with the guanidino phosphoramidate modification providing a 14.0 °C increase in T_m relative to the DMAP modification for fully modified ONs [112]. Furthermore, a noticeable increase in T_m was observed for all fully modified α-TFOs irrespectively of modification (DMAP, aminobutyl or guanidinobutyl phosphoramidate) relative to the unmodified β-TFO control [112]. A cell uptake assay was conducted between two 12-mer poly-T ASOs, one being an α-ASO with a fully modified guanidinobutyl phosphoramidate backbone and the other a PS-β-ASO control, both fluorescein-labelled at the 5'-end. The study found that the guanidinium modification increased the cellular uptake. However, ASOs carrying the novel guanidinium modification were mainly localized in the cytoplasm, indicating that the ASOs are taken up by endocytosis but are retained in part in the endocytic vesicles [112].

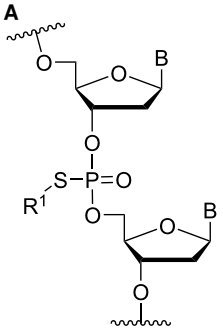
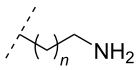
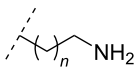
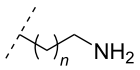
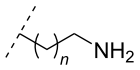
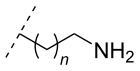
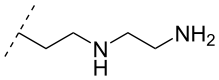
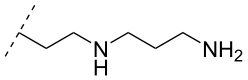
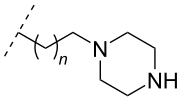
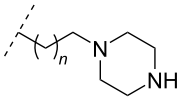
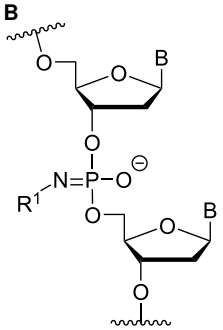
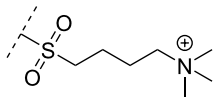
Utilizing the phosphorus atom as an attachment point for cationic aminoalkyl groups has been employed with slight variations. Fathi, Cook and co-workers successfully introduced aminomethyl phosphonate [113] (**69**) and aminoethyl phosphonate [114] (**70**) linkages (Table 8). The introduction of the stereo-pure aminomethyl phosphonate linkage was achieved by preparing the appropriate stereo-pure (R_p or S_p) thymidine dinucleotide linked through the 3'-5' oxygen atoms modified with the phthalimidomethyl phosphonate linkage [113]. Later, a halogenated phthalimide protection was employed for the synthesis of the amidoethyl phosphonate variant [114]. The desired dinucleotides were phosphitylated and incorporated into 13-mer poly-T ONs [113,114]. For both modifications, the ONs modified with the R_p-isomer formed more stable duplexes with DNA and RNA complements relative to the control ON. The ONs carrying the S_p-isomer had a destabilizing effect. When tested for their nuclease resistance, an increase in stability was observed relative to their unmodified ON [113,114]. Interestingly, a difference in hydrolysis rate was noticed: the aminomethyl (**69**) modification was readily hydrolysed at pH 7 whereas the aminoethyl (**70**) modification was completely stable under the same conditions [114]. A preliminary cell uptake assay was conducted for a net-neutral ASO carrying the aminoethyl phosphonate linkage. This showed that under appropriate conditions (1 μM and at 37 °C) the net-neutral ASO had improved uptake relative to the anionic PO-ASO, demonstrating a concentration-dependent uptake [114].

A new class of internucleotide linkages has recently been introduced, termed branched charge-neutralizing sleeves (BCNSs). These cationic internucleotide linkages were synthesized through conventional phosphoramidate chemistry with a slight variation. In contrast to the standard method, bis(diisopropylamino)chlorophosphine was first reacted with either of the three diaminoalcohol groups, before subsequent phosphitylation with the DMT-protected nucleosides [115]. After incorporation into ONs, conversion to monomers **78–81** was accomplished as shown in Table 8. Although monomer **78** showed good yield in both phosphoramidite synthesis and coupling efficiency on the synthesizer, significant loss of the hydrocarbon-linked group was observed during the alkaline deprotection conditions [115]. After insertion into a 19-mer DNA-ON, modification **79** did not show any significant increase in T_m. This trend was changed when the modification **81** was inserted into a 21-mer 2'-OMe RNA-ON and hybridized to a complementary 2'-OMe RNA complement. Relative to the unmodified 2'-OMe RNA sequence, the modified ON carrying six insertions of monomer **81** gave an increase in T_m of 11 °C. The serum stability of modification **81** was evaluated in PBS/bovine serum at 37 °C for 8 h. Here, a significant improvement over the unmodified 2'-OMe RNA was observed for modification **81** [115]. The two BCNSs used for monomers **79–81** (1,3-bis(2-(amino)ethoxy)-2-propyl and bis(2-(dimethylamino)ethoxy)-2-propyl) were tested for their cell uptake properties. The BCNS carrying the dimethylamino groups had a higher effect on the cellular uptake for 5'-FAM-labelled ASOs relative to the oligomers carrying BCNS being primary amines. However, both modifications demonstrated an improved cellular uptake relative to the unmodified 5'-end FAM-labelled 2'-OMe ASO [115].

Another strategy to introduce cationic aminoalkylated moieties onto phosphorus atoms of the ON backbone involves aminoalkylated phosphorothioate linkages. In general, modifications of the backbone have been used in the context of the phosphodiester linkage, while only a few examples can be found for the PS linkage [117-121]. Rahman, Obika and co-workers investigated the properties of ONs modified with the aminoalkyl–PS linkage [117], based on the post-synthetic alkylation protocol developed earlier by Chen and Gothelf reacting PS-ONs with 2-bromoethylammonium bromide in HEPES (4-(2-hydroxyethyl)-1-piperazine-ethanesulfonic acid) buffer (dimethylformamide/H_2O 1:9) at 45 °C [118]. Initially, a 12-mer sequence containing the nucleobases guanine, thymine, and cytosine was tested by incorporating the earlier reported aminoethyl–PS linkage [118] (modification **82**). However, cleavage products formed by guanine alkylation prompted a switch to a 12-mer sequence containing only cytosine and thymine [117].

Conjugation of the desired aminoalkyl moieties with the stereopure PS-ON (R_p or S_p) gave the aminoalkyl–PS linkages shown in Table 9A, in yields between 24–55%. Extensive work has been carried out for the synthesis of stereochemically pure PS-ONs and one of these methods employs the proline-derived bicyclic oxazaphospholidine monomer [122] which was later used in the scalable synthetic process of therapeutic stereopure PS-ASOs [123]. When the modified ONs were hybridized with their complementary DNA strands it was clear that the R_p stereoisomer of the aminoalkylated PS linkage improved the stability of the formed duplexes, while the S_p stereoisomer destabilized the formed duplexes, all relative to their PS-ON controls. Against the complementary RNA target, the R_p stereoisomer either had a similar or slightly lower T_m than the control ON, while the S_p stereoisomeric linkages destabilized the formed aminoalkylated PS-ON/RNA duplex significantly. When tested for their triplex-forming properties, a similar pattern was observed. The R_p stereoisomeric linkages gave an improved stabilization of up to +6 °C, with the non-cyclic diamine (monomers **87** and **88**) stabilizing duplex formation the most, while all S_p stereoisomeric linkages gave lower T_m than their control [117]. The nuclease stability of the aminoethyl (**82**) and aminohexyl (**86**) PS linkages was tested relative to the

Table 9: Amine-functionalized groups as internucleoside linkage.[a]

backbone modifications	R^1		n/R^2	ref.	thermo-stability	nuclease resistance	activity in cell
A 		**82**	$n = 1$	[117,118]	+	+	n.d.
		**83**	$n = 2$	[117]	+	n.d.	n.d.
		**84**	$n = 3$	[117]	+	n.d.	n.d.
		**85**	$n = 4$	[117]	+	n.d.	n.d.
		**86**	$n = 5$	[117]	+	+	n.d.
		**87**		[117]	+	n.d.	n.d.
		**88**		[117]	+	n.d.	n.d.
		**89**	$n = 1$	[117]	+	n.d.	n.d.
		**90**	$n = 2$	[117]	+	n.d.	n.d.
B 		**91**		[124]	+	+	n.d.

[a]A '+' sign has been added when the modified ON/ASO showed improved thermal stability (T_m) either towards ssDNA, ssRNA or dsDNA, and when the nuclease stability for the modified ON/ASO demonstrated improved stability, all compared to the effects mediated by control DNA or RNA strands. The '**X**' sign has been added when the modified ASO demonstrated equal or better gene inhibitory activity in cells relative to the ASO control, while n.d. = no determined.

PS-ON control. Interestingly, both modifications showed an improved resistance relative to the control ON; however, the S_p-aminoethyl (**82**) PS-linkage was slightly better than the S_p-aminohexyl (**86**) linkage while the opposite was true for the R_p stereoisomeric linkages [117]. It is noteworthy to mention that the 2-(3-aminopropyl)aminoethyl (**88**) PS linkage was too heat-labile for nuclease tests [117].

Filichev and co-workers have recently synthesized a variant of the phosphoramidate internucleotide linkage (Table 9B) and have obtained some preliminary results. Modification **91** was introduced via Staudinger reaction during solid-phase DNA synthesis, circumventing some of the laborious work involved in the synthesis of some internucleotide linkages. The authors demonstrated that this new modification could enhance the

stability of the ON/RNA duplex, although no significant change was observed for the ON /DNA duplex (100 mM NaCl). Additionally, the dependency of the position of the modification had a large effect, since changing the position, for a single modification from the 5'-end to the middle and then to the 3'-end, resulted in different T_m increases of 7 °C, 1 °C, and 12 °C, respectively, for the ON/RNA duplex, relative to the DNA control [124]. At pH 5.0, the internucleotide linkage could stabilize a triplex between 10–11 °C when the modification was positioned in the middle, near the 3'-end or as a double-modified TFO near the 5'- and 3'-ends, although this high stability diminished when three or four modifications were inserted [124]. At pH 6.0 no significant stabilization was observed; however, a good nuclease stability was observed for the ON carrying four insertions of modification **91** [124].

The concept of introducing cationic backbone linkages extends beyond conjugating cationic aminoalkyl groups to the phosphorus atom which, however, is beyond the scope of this summary. For a more in-depth account of cationic backbone modifications, we direct the reader to a recent review by Meng and Ducho [125].

Conclusion

Cationic amine- and polyamine-conjugates/derivatives have the potential to improve the properties of ASOs for defined applications. Many different aspects need to be considered when optimizing an ASO for a defined application. This includes the position of a modification as well as the chemical composition of the cationic group, but also what scaffold the cationic group is to be attached to, i.e., is it a 2'-amino-LNA nucleotide or the common DNA/RNA nucleotides. This difference is reflected when evaluating groups attached to the base and sugar moieties. When the cationic group is connected via the nucleobase, thus allowing the group to be positioned in the major groove of the duplex, an increased target binding affinity is usually seen relative to modifications on the furanose ring (minor groove). However, when the 2'-amino-LNA scaffolds is used, a very high duplex stability is seen irrespectively of the target (DNA or RNA), as a synergistic effect between the locked sugar conformation and the cationic moiety emerges. Additionally, sugar modifications tend to bring higher nuclease resistance compared with nucleobase modifications. However, a drawback is the lack of RNase H activation for most of the sugar-modified derivatives. This is contrary to the nucleobase-modified variants that are in general well tolerated RNase H substrates. Another important aspect is the overall net-charge and the charge density of the ON. As the sugar and nucleobase substitutions compensate for the anionic charge carried by the PO and PS-backbone, a more densely charged ON is created which can result in synthetic difficulties. This can be circumvented by using internucleotide linkage modifications which generate net neutral or net positive ASOs, which usually have a high resistance towards nuclease degradation. Additionally, the internucleotide modifications result in ASOs that can stabilize the ASO/RNA duplex to a relatively high degree, although the R_p stereoisomer is generally the preferred isomer for improved ASO/RNA stabilization. These considerations relate to the overall design of the ASOs, since the gapmer design allows for sugar-modified nucleotides to be used on the flanks, while for a mixmer- or a fully modified ASO a more diverse composition might be used.

Another consideration is the use of more densely modified nucleotides, i.e., carrying both modified base and sugar moieties. Although not as extensively explored, this strategy was employed by Fox and co-workers who elaborated on a 2'-deoxyuridine variant carrying a 1-propargylamino group (Figure 2A) that had already demonstrated enhanced stability for the T·AT triplet [126,127]. This monomer was further modified to a bis-modified uridine analogue, i.e., a nucleotide containing both the 1-propargylamino group on the 5-position and an aminoethoxy moiety at the 2'-position (Figure 2B), thus creating a monomer that could further stabilize the triplex [128]. In theory, this strategy allows for fewer modifications to be used in designing ASOs with high RNA target affinities.

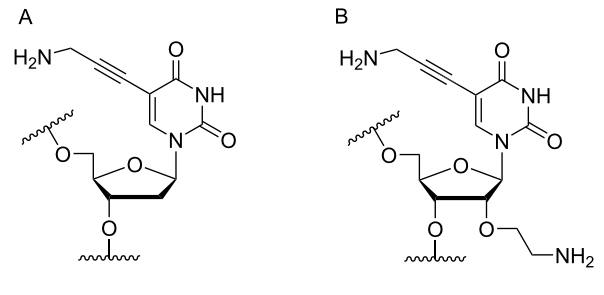


Figure 2: Structures of 5-(1-propargylamino)-2'-deoxyuridine (A) and 2'-aminoethoxy-5-propargylaminouridine (B).

Funding

The VILLUM FONDEN is thanked for funding the Biomolecular Nanoscale Engineering Center (BioNEC), a VILLUM center of excellence, grant number VKR18333.

ORCID® iDs

Mathias B. Danielsen - https://orcid.org/0000-0001-9862-3967
Jesper Wengel - https://orcid.org/0000-0001-9835-1009

References

1. Kim, J.; Hu, C.; Moufawad El Achkar, C.; Black, L. E.; Douville, J.; Larson, A.; Pendergast, M. K.; Goldkind, S. F.; Lee, E. A.; Kuniholm, A.; Soucy, A.; Vaze, J.; Belur, N. R.; Fredriksen, K.; Stojkovska, I.; Tsytsykova, A.; Armant, M.; DiDonato, R. L.; Choi, J.; Cornelissen, L.; Pereira, L. M.; Augustine, E. F.; Genetti, C. A.; Dies, K.; Barton, B.; Williams, L.; Goodlett, B. D.; Riley, B. L.; Pasternak, A.; Berry, E. R.; Pflock, K. A.; Chu, S.; Reed, C.; Tyndall, K.; Agrawal, P. B.; Beggs, A. H.; Grant, P. E.; Urion, D. K.; Snyder, R. O.; Waisbren, S. E.; Poduri, A.; Park, P. J.; Patterson, A.; Biffi, A.; Mazzulli, J. R.; Bodamer, O.; Berde, C. B.; Yu, T. W. *N. Engl. J. Med.* **2019,** *381,* 1644–1652. doi:10.1056/nejmoa1813279
2. Monia, B. P.; Lesnik, E. A.; Gonzalez, C.; Lima, W. F.; McGee, D.; Guinosso, C. J.; Kawasaki, A. M.; Cook, P. D.; Freier, S. M. *J. Biol. Chem.* **1993,** *268,* 14514–14522. doi:10.1016/s0021-9258(19)85268-7
3. Roberts, T. C.; Langer, R.; Wood, M. J. A. *Nat. Rev. Drug Discovery* **2020,** *19,* 673–694. doi:10.1038/s41573-020-0075-7
4. Stein, C. A.; Castanotto, D. *Mol. Ther.* **2017,** *25,* 1069–1075. doi:10.1016/j.ymthe.2017.03.023
5. Xiong, H.; Veedu, R. N.; Diermeier, S. D. *Int. J. Mol. Sci.* **2021,** *22,* 3295. doi:10.3390/ijms22073295
6. Hagedorn, P. H.; Pontoppidan, M.; Bisgaard, T. S.; Berrera, M.; Dieckmann, A.; Ebeling, M.; Møller, M. R.; Hudlebusch, H.; Jensen, M. L.; Hansen, H. F.; Koch, T.; Lindow, M. *Nucleic Acids Res.* **2018,** *46,* 5366–5380. doi:10.1093/nar/gky397
7. Scharner, J.; Ma, W. K.; Zhang, Q.; Lin, K.-T.; Rigo, F.; Bennett, C. F.; Krainer, A. R. *Nucleic Acids Res.* **2020,** *48,* 802–816. doi:10.1093/nar/gkz1132
8. Yoshida, T.; Naito, Y.; Yasuhara, H.; Sasaki, K.; Kawaji, H.; Kawai, J.; Naito, M.; Okuda, H.; Obika, S.; Inoue, T. *Genes Cells* **2019,** *24,* 827–835. doi:10.1111/gtc.12730
9. Jackson, A. L.; Bartz, S. R.; Schelter, J.; Kobayashi, S. V.; Burchard, J.; Mao, M.; Li, B.; Cavet, G.; Linsley, P. S. *Nat. Biotechnol.* **2003,** *21,* 635–637. doi:10.1038/nbt831
10. Jackson, A. L.; Burchard, J.; Schelter, J.; Chau, B. N.; Cleary, M.; Lim, L.; Linsley, P. S. *RNA* **2006,** *12,* 1179–1187. doi:10.1261/rna.25706
11. Scacheri, P. C.; Rozenblatt-Rosen, O.; Caplen, N. J.; Wolfsberg, T. G.; Umayam, L.; Lee, J. C.; Hughes, C. M.; Shanmugam, K. S.; Bhattacharjee, A.; Meyerson, M.; Collins, F. S. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** *101,* 1892–1897. doi:10.1073/pnas.0308698100
12. Burdick, A. D.; Sciabola, S.; Mantena, S. R.; Hollingshead, B. D.; Stanton, R.; Warneke, J. A.; Zeng, M.; Martsen, E.; Medvedev, A.; Makarov, S. S.; Reed, L. A.; Davis, J. W., II; Whiteley, L. O. *Nucleic Acids Res.* **2014,** *42,* 4882–4891. doi:10.1093/nar/gku142
13. Grimm, D.; Streetz, K. L.; Jopling, C. L.; Storm, T. A.; Pandey, K.; Davis, C. R.; Marion, P.; Salazar, F.; Kay, M. A. *Nature* **2006,** *441,* 537–541. doi:10.1038/nature04791
14. Persengiev, S. P.; Zhu, X.; Green, M. R. *RNA* **2004,** *10,* 12–18. doi:10.1261/rna5160904
15. Kaczmarkiewicz, A.; Nuckowski, Ł.; Studzińska, S.; Buszewski, B. *Crit. Rev. Anal. Chem.* **2019,** *49,* 256–270. doi:10.1080/10408347.2018.1517034
16. Studzińska, S. *Talanta* **2018,** *176,* 329–343. doi:10.1016/j.talanta.2017.08.025
17. Du Rietz, H.; Hedlund, H.; Wilhelmson, S.; Nordenfelt, P.; Wittrup, A. *Nat. Commun.* **2020,** *11,* 1809. doi:10.1038/s41467-020-15300-1
18. Yu, R. Z.; Geary, R. S.; Siwkowski, A.; Levin, A. A. Modified Antisense Oligonucleotides in Animals and Man. In *Antisense drug technology: principles, strategies, and applications;* Crooke, S. T., Ed.; CRC Press: Boca Raton, FL, USA, 2008; pp 305–326.
19. Shen, W.; De Hoyos, C. L.; Migawa, M. T.; Vickers, T. A.; Sun, H.; Low, A.; Bell, T. A., III; Rahdar, M.; Mukhopadhyay, S.; Hart, C. E.; Bell, M.; Riney, S.; Murray, S. F.; Greenlee, S.; Crooke, R. M.; Liang, X.-h.; Seth, P. P.; Crooke, S. T. *Nat. Biotechnol.* **2019,** *37,* 640–650. doi:10.1038/s41587-019-0106-2
20. Migawa, M. T.; Shen, W.; Wan, W. B.; Vasquez, G.; Oestergaard, M. E.; Low, A.; De Hoyos, C. L.; Gupta, R.; Murray, S.; Tanowitz, M.; Bell, M.; Nichols, J. G.; Gaus, H.; Liang, X.-h.; Swayze, E. E.; Crooke, S. T.; Seth, P. P. *Nucleic Acids Res.* **2019,** *47,* 5465–5479. doi:10.1093/nar/gkz247
21. Wang, S.; Allen, N.; Prakash, T. P.; Liang, X.-h.; Crooke, S. T. *Nucleic Acid Ther.* **2019,** *29,* 245–255. doi:10.1089/nat.2019.0794
22. Prakash, T. P.; Graham, M. J.; Yu, J.; Carty, R.; Low, A.; Chappell, A.; Schmidt, K.; Zhao, C.; Aghajan, M.; Murray, H. F.; Riney, S.; Booten, S. L.; Murray, S. F.; Gaus, H.; Crosby, J.; Lima, W. F.; Guo, S.; Monia, B. P.; Swayze, E. E.; Seth, P. P. *Nucleic Acids Res.* **2014,** *42,* 8796–8807. doi:10.1093/nar/gku531
23. Tanowitz, M.; Hettrick, L.; Revenko, A.; Kinberger, G. A.; Prakash, T. P.; Seth, P. P. *Nucleic Acids Res.* **2017,** *45,* 12388–12400. doi:10.1093/nar/gkx960
24. Ämmälä, C.; Drury, W. J., III; Knerr, L.; Ahlstedt, I.; Stillemark-Billton, P.; Wennberg-Huldt, C.; Andersson, E.-M.; Valeur, E.; Jansson-Löfmark, R.; Janzén, D.; Sundström, L.; Meuller, J.; Claesson, J.; Andersson, P.; Johansson, C.; Lee, R. G.; Prakash, T. P.; Seth, P. P.; Monia, B. P.; Andersson, S. *Sci. Adv.* **2018,** *4,* eaat3386. doi:10.1126/sciadv.aat3386
25. Menzi, M.; Lightfoot, H. L.; Hall, J. *Future Med. Chem.* **2015,** *7,* 1733–1749. doi:10.4155/fmc.15.90
26. Herdewijn, P. *Antisense Nucleic Acid Drug Dev.* **2000,** *10,* 297–310. doi:10.1089/108729000421475
27. Kohgo, S.; Shinozuka, K.; Ozaki, H.; Sawai, H. *Tetrahedron Lett.* **1998,** *39,* 4067–4070. doi:10.1016/s0040-4039(98)00660-1
28. Ono, A.; Haginoya, N.; Kiyokawa, M.; Minakawa, N.; Matsuda, A. *Bioorg. Med. Chem. Lett.* **1994,** *4,* 361–366. doi:10.1016/s0960-894x(01)80145-1
29. Nomura, Y.; Haginoya, N.; Ueno, Y.; Matsuda, A. *Bioorg. Med. Chem. Lett.* **1996,** *6,* 2811–2816. doi:10.1016/s0960-894x(96)00519-7
30. Haginoya, N.; Ono, A.; Nomura, Y.; Ueno, Y.; Matsuda, A. *Bioconjugate Chem.* **1997,** *8,* 271–280. doi:10.1021/bc970021r
31. Ito, T.; Ueno, Y.; Komatsu, Y.; Matsuda, A. *Nucleic Acids Res.* **2003,** *31,* 2514–2523. doi:10.1093/nar/gkg374
32. Ueno, Y.; Kumagai, I.; Haginoya, N.; Matsuda, A. *Nucleic Acids Res.* **1997,** *25,* 3777–3782. doi:10.1093/nar/25.19.3777
33. Juan, E. C. M.; Kondo, J.; Kurihara, T.; Ito, T.; Ueno, Y.; Matsuda, A.; Takenaka, A. *Nucleic Acids Res.* **2007,** *35,* 1969–1977. doi:10.1093/nar/gkl821
34. Ozaki, H.; Nakamura, A.; Arai, M.; Endo, M.; Sawai, H. *Bull. Chem. Soc. Jpn.* **1995,** *68,* 1981–1987. doi:10.1246/bcsj.68.1981
35. Shinozuka, K.; Umeda, A.; Aoki, T.; Sawai, H. *Nucleosides, Nucleotides Nucleic Acids* **1998,** *17,* 291–300. doi:10.1080/07328319808005177
36. Sawai, H.; Nakamura, A.; Sekiguchi, S.; Yumoto, K.; Endoh, M.; Ozaki, H. *J. Chem. Soc., Chem. Commun.* **1994,** 1997–1998. doi:10.1039/c39940001997

37. Ueno, Y.; Mikawa, M.; Matsuda, A. *Bioconjugate Chem.* **1998,** *9,* 33–39. doi:10.1021/bc9701508
38. Ozaki, H.; Mine, M.; Shinozuka, K.; Sawai, H. *Nucleosides, Nucleotides Nucleic Acids* **2004,** *23,* 339–346. doi:10.1081/ncn-120027903
39. Matsukura, M.; Okamoto, T.; Miike, T.; Sawai, H.; Shinozuka, K. *Biochem. Biophys. Res. Commun.* **2002,** *293,* 1341–1347. doi:10.1016/s0006-291x(02)00383-2
40. Moriguchi, T.; Sakai, H.; Suzuki, H.; Shinozuka, K. *Chem. Pharm. Bull.* **2008,** *56,* 1259–1263. doi:10.1248/cpb.56.1259
41. Sakai, H.; Moriguchi, T.; Suzuki, H.; Matsukura, M.; Shinozuka, K. *Nucleic Acids Symp. Ser.* **2001,** *1,* 127–128. doi:10.1093/nass/1.1.127
42. Masud, M. M.; Masuda, T.; Inoue, Y.; Kuwahara, M.; Sawai, H.; Ozaki, H. *Bioorg. Med. Chem. Lett.* **2011,** *21,* 715–717. doi:10.1016/j.bmcl.2010.11.125
43. Shinozuka, K.; Matsukura, M.; Okamoto, T.; Sawai, H. *Nucleosides Nucleotides* **1998,** *17,* 2081–2084. doi:10.1080/07328319808004749
44. Hashimoto, H.; Nelson, M. G.; Switzer, C. *J. Am. Chem. Soc.* **1993,** *115,* 7128–7134. doi:10.1021/ja00069a009
45. Someya, T.; Ando, A.; Kimoto, M.; Hirao, I. *Nucleic Acids Res.* **2015,** *43,* 6665–6676. doi:10.1093/nar/gkv638
46. Bag, S. S.; Jana, S.; Kasula, M. Sonogashira Cross-Coupling: Alkyne-Modified Nucleosides and Their Applications. In *Palladium-Catalyzed Modification of Nucleosides, Nucleotides and Oligonucleotides;* Kapdi, A. R.; Maiti, D.; Sanghvi, Y. S., Eds.; Elsevier, 2018; pp 75–146. doi:10.1016/b978-0-12-811292-2.00004-0
47. Kumar, P.; Østergaard, M. E.; Baral, B.; Anderson, B. A.; Guenther, D. C.; Kaura, M.; Raible, D. J.; Sharma, P. K.; Hrdlicka, P. J. *J. Org. Chem.* **2014,** *79,* 5047–5061. doi:10.1021/jo500614a
48. Heystek, L. E.; Zhou, H.-q.; Dande, P.; Gold, B. *J. Am. Chem. Soc.* **1998,** *120,* 12165–12166. doi:10.1021/ja982039y
49. Strauss, J. K.; Prakash, T. P.; Roberts, C.; Switzer, C.; James Maher, L., III. *Chem. Biol.* **1996,** *3,* 671–678. doi:10.1016/s1074-5521(96)90135-0
50. Soto, A. M.; Kankia, B. I.; Dande, P.; Gold, B.; Marky, L. A. *Nucleic Acids Res.* **2001,** *29,* 3638–3645. doi:10.1093/nar/29.17.3638
51. Takeda, T.; Ikeda, K.; Mizuno, Y.; Ueda, T. *Chem. Pharm. Bull.* **1987,** *35,* 3558–3567. doi:10.1248/cpb.35.3558
52. Nara, H.; Ono, A.; Matsuda, A. *Bioconjugate Chem.* **1995,** *6,* 54–61. doi:10.1021/bc00031a005
53. Marsh, A. J.; Williams, D. M.; Grasby, J. A. *Org. Biomol. Chem.* **2004,** *2,* 2103–2112. doi:10.1039/b404150j
54. Guenther, D. C.; Kumar, P.; Anderson, B. A.; Hrdlicka, P. J. *Chem. Commun.* **2014,** *50,* 9007–9009. doi:10.1039/c4cc03623a
55. Sau, S. P.; Kumar, P.; Anderson, B. A.; Østergaard, M. E.; Deobald, L.; Paszczynski, A.; Sharma, P. K.; Hrdlicka, P. J. *Chem. Commun.* **2009,** 6756–6758. doi:10.1039/b917312a
56. Geny, S.; Moreno, P. M. D.; Krzywkowski, T.; Gissberg, O.; Andersen, N. K.; Isse, A. J.; El-Madani, A. M.; Lou, C.; Pabon, Y. V.; Anderson, B. A.; Zaghloul, E. M.; Zain, R.; Hrdlicka, P. J.; Jørgensen, P. T.; Nilsson, M.; Lundin, K. E.; Pedersen, E. B.; Wengel, J.; Smith, C. I. E. *Nucleic Acids Res.* **2016,** *44,* 2007–2019. doi:10.1093/nar/gkw021
57. Lin, K.-Y.; Matteucci, M. D. *J. Am. Chem. Soc.* **1998,** *120,* 8531–8532. doi:10.1021/ja981286z
58. Flanagan, W. M.; Wolf, J. J.; Olson, P.; Grant, D.; Lin, K.-Y.; Wagner, R. W.; Matteucci, M. D. *Proc. Natl. Acad. Sci. U. S. A.* **1999,** *96,* 3513–3518. doi:10.1073/pnas.96.7.3513
59. Wilds, C. J.; Maier, M. A.; Tereshko, V.; Manoharan, M.; Egli, M. *Angew. Chem., Int. Ed.* **2002,** *41,* 115–117. doi:10.1002/1521-3773(20020104)41:1<115::aid-anie115>3.0.co;2-r
60. Webb, T. R.; Matteucci, M. D. *Nucleic Acids Res.* **1986,** *14,* 7661–7674. doi:10.1093/nar/14.19.7661
61. Decuypere, E.; Lepikhina, A.; Halloy, F.; Hall, J. *Helv. Chim. Acta* **2019,** *102,* e1900222. doi:10.1002/hlca.201900222
62. Menzi, M.; Wild, B.; Pradère, U.; Malinowska, A. L.; Brunschweiger, A.; Lightfoot, H. L.; Hall, J. *Chem. – Eur. J.* **2017,** *23,* 14221–14230. doi:10.1002/chem.201701670
63. Markiewicz, W. T.; Godzina, P.; Markiewicz, M.; Astriab, A. *Nucleosides Nucleotides* **1998,** *17,* 1871–1880. doi:10.1080/07328319808004725
64. Markiewicz, W. T.; Godzina, P.; Markiewicz, M. *Nucleosides Nucleotides* **1999,** *18,* 1449–1454. doi:10.1080/07328319908044749
65. Godzina, P.; Adrych-Rozek, K.; Markiewicz, W. T. *Nucleosides Nucleotides* **1999,** *18,* 2397–2414. doi:10.1080/07328319908044615
66. Schmid, N.; Behr, J.-P. *Tetrahedron Lett.* **1995,** *36,* 1447–1450. doi:10.1016/0040-4039(95)00027-a
67. Adib, A.; Potier, P. F.; Doronina, S.; Huc, I.; Behr, J.-P. *Tetrahedron Lett.* **1997,** *38,* 2989–2992. doi:10.1016/s0040-4039(97)00540-6
68. Diaz, A. R.; Eritja, R.; Garcia, R. G. *Nucleosides Nucleotides* **1997,** *16,* 2035–2051. doi:10.1080/07328319708002554
69. Potier, P.; Adib, A.; Kochkine, A.; Huc, I.; Behr, J.-P. *Nucleosides Nucleotides* **1999,** *18,* 1467–1468. doi:10.1080/07328319908044754
70. Shinozuka, K.; Onodera, M.; Ikeda, H.; Sawai, H. *Nucleic Acids Symp. Ser.* **1999,** *42,* 229–230. doi:10.1093/nass/42.1.229
71. Shinozuka, K.; Onodera, M.; Ikeda, H.; Sawai, H. *Chem. Lett.* **2002,** *31,* 200–201. doi:10.1246/cl.2002.200
72. Bartosik, K.; Debiec, K.; Czarnecka, A.; Sochacka, E.; Leszczynska, G. *Molecules* **2020,** *25,* 3344. doi:10.3390/molecules25153344
73. Bobkov, G. V.; Mikhailov, S. N.; Van Aerschot, A.; Herdewijn, P. *Tetrahedron* **2008,** *64,* 6238–6251. doi:10.1016/j.tet.2008.04.110
74. Griffey, R. H.; Monia, B. P.; Cummins, L. L.; Freier, S.; Greig, M. J.; Guinosso, C. J.; Lesnik, E.; Manalili, S. M.; Mohan, V.; Owens, S.; Ross, B. R.; Sasmor, H.; Wancewicz, E.; Weiler, K.; Wheeler, P. D.; Cook, P. D. *J. Med. Chem.* **1996,** *39,* 5100–5109. doi:10.1021/jm950937o
75. Noe, C. R.; Winkler, J.; Urban, E.; Gilbert, M.; Haberhauer, G.; Brunar, H. *Nucleosides, Nucleotides Nucleic Acids* **2005,** *24,* 1167–1185. doi:10.1081/ncn-200067400
76. Winkler, J.; Gilbert, M.; Kocourková, A.; Stessl, M.; Noe, C. R. *ChemMedChem* **2008,** *3,* 102–110. doi:10.1002/cmdc.200700169
77. Bramsen, J. B.; Laursen, M. B.; Nielsen, A. F.; Hansen, T. B.; Bus, C.; Langkjær, N.; Babu, B. R.; Højland, T.; Abramov, M.; Van Aerschot, A.; Odadzic, D.; Smicius, R.; Haas, J.; Andree, C.; Barman, J.; Wenska, M.; Srivastava, P.; Zhou, C.; Honcharenko, D.; Hess, S.; Müller, E.; Bobkov, G. V.; Mikhailov, S. N.; Fava, E.; Meyer, T. F.; Chattopadhyaya, J.; Zerial, M.; Engels, J. W.; Herdewijn, P.; Wengel, J.; Kjems, J. *Nucleic Acids Res.* **2009,** *37,* 2867–2881. doi:10.1093/nar/gkp106
78. Prakash, T. P.; Manoharan, M.; Fraser, A. S.; Kawasaki, A. M.; Lesnik, E. A.; Owens, S. R. *Tetrahedron Lett.* **2000,** *41,* 4855–4859. doi:10.1016/s0040-4039(00)00703-6

79. Prhavc, M.; Prakash, T. P.; Minasov, G.; Cook, P. D.; Egli, M.; Manoharan, M. *Org. Lett.* **2003,** *5,* 2017–2020. doi:10.1021/ol0340991
80. Azéma, L.; Bathany, K.; Rayner, B. *ChemBioChem* **2010,** *11,* 2513–2516. doi:10.1002/cbic.201000538
81. Teplova, M.; Wallace, S. T.; Tereshko, V.; Minasov, G.; Symons, A. M.; Cook, P. D.; Manoharan, M.; Egli, M. *Proc. Natl. Acad. Sci. U. S. A.* **1999,** *96,* 14240–14245. doi:10.1073/pnas.96.25.14240
82. Winkler, J.; Noe, C. R. *Nucleosides, Nucleotides Nucleic Acids* **2007,** *26,* 939–942. doi:10.1080/15257770701507978
83. Winkler, J.; Giessrigl, B.; Novak, C.; Urban, E.; Noe, C. R. *Monatsh. Chem.* **2010,** *141,* 809–815. doi:10.1007/s00706-010-0318-0
84. Seio, K.; Tokugawa, M.; Kanamori, T.; Tsunoda, H.; Ohkubo, A.; Sekine, M. *Bioorg. Med. Chem. Lett.* **2012,** *22,* 2470–2473. doi:10.1016/j.bmcl.2012.02.012
85. Korshun, V. A.; Stetsenko, D. A.; Gait, M. J. *J. Chem. Soc., Perkin Trans. 1* **2002,** 1092–1104. doi:10.1039/b111434b
86. Misra, A.; Dwivedi, P.; Shahid, M. *Russ. J. Bioorg. Chem.* **2009,** *35,* 62–67. doi:10.1134/s1068162009010087
87. Dioubankova, N. N.; Malakhov, A. D.; Shenkarev, Z. O.; Korshun, V. A. *Tetrahedron* **2004,** *60,* 4617–4626. doi:10.1016/j.tet.2004.03.076
88. Prhavc, M.; Lesnik, E. A.; Mohan, V.; Manoharan, M. *Tetrahedron Lett.* **2001,** *42,* 8777–8780. doi:10.1016/s0040-4039(01)01908-6
89. Winkler, J.; Urban, E.; Losert, D.; Wacheck, V.; Pehamberger, H.; Noe, C. R. *Nucleic Acids Res.* **2004,** *32,* 710–718. doi:10.1093/nar/gkh229
90. Winkler, J.; Urban, E.; Noe, C. R. *Bioconjugate Chem.* **2005,** *16,* 1038–1044. doi:10.1021/bc049729d
91. Winkler, J.; Saadat, K.; Díaz-Gavilán, M.; Urban, E.; Noe, C. R. *Eur. J. Med. Chem.* **2009,** *44,* 670–677. doi:10.1016/j.ejmech.2008.05.012
92. Johannsen, M. W.; Crispino, L.; Wamberg, M. C.; Kalra, N.; Wengel, J. *Org. Biomol. Chem.* **2011,** *9,* 243–252. doi:10.1039/c0ob00532k
93. Lou, C.; Vester, B.; Wengel, J. *Chem. Commun.* **2015,** *51,* 4024–4027. doi:10.1039/c5cc00322a
94. Lou, C.; Samuelsen, S. V.; Christensen, N. J.; Vester, B.; Wengel, J. *Bioconjugate Chem.* **2017,** *28,* 1214–1220. doi:10.1021/acs.bioconjchem.7b00061
95. Danielsen, M. B.; Christensen, N. J.; Jørgensen, P. T.; Jensen, K. J.; Wengel, J.; Lou, C. *Chem. – Eur. J.* **2021,** *27,* 1416–1422. doi:10.1002/chem.202004495
96. Yamashita, S.; Nishida, K.; Osawa, T.; Nakanishi, A.; Ito, Y.; Hari, Y. *Molecules* **2020,** *25,* 346. doi:10.3390/molecules25020346
97. Kanazaki, M.; Ueno, Y.; Shuto, S.; Matsuda, A. *J. Am. Chem. Soc.* **2000,** *122,* 2422–2432. doi:10.1021/ja9934706
98. Wang, G.; Seifert, W. E. *Tetrahedron Lett.* **1996,** *37,* 6515–6518. doi:10.1016/0040-4039(96)01432-3
99. Wang, G.; Middleton, P. J.; Lin, C.; Pietrzkowski, Z. *Bioorg. Med. Chem. Lett.* **1999,** *9,* 885–890. doi:10.1016/s0960-894x(99)00103-1
100. Ueno, Y.; Nagasawa, Y.; Sugimoto, I.; Kojima, N.; Kanazaki, M.; Shuto, S.; Matsuda, A. *J. Org. Chem.* **1998,** *63,* 1660–1667. doi:10.1021/jo9720492
101. Ono, A.; Dan, A.; Matsuda, A. *Bioconjugate Chem.* **1993,** *4,* 499–508. doi:10.1021/bc00024a012
102. Westheimer, F. H. *Science* **1987,** *235,* 1173–1178. doi:10.1126/science.2434996
103. Benner, S. A. *Acc. Chem. Res.* **2004,** *37,* 784–797. doi:10.1021/ar040004z
104. Crooke, S. T.; Vickers, T. A.; Liang, X.-h. *Nucleic Acids Res.* **2020,** *48,* 5235–5253. doi:10.1093/nar/gkaa299
105. Letsinger, R. L.; Bach, S. A.; Eadie, J. S. *Nucleic Acids Res.* **1986,** *14,* 3487–3499. doi:10.1093/nar/14.8.3487
106. Letsinger, R. L.; Singman, C. N.; Histand, G.; Salunkhe, M. *J. Am. Chem. Soc.* **1988,** *110,* 4470–4471. doi:10.1021/ja00221a089
107. Chaturvedi, S.; Horn, T.; Letsinger, R. L. *Nucleic Acids Res.* **1996,** *24,* 2318–2323. doi:10.1093/nar/24.12.2318
108. Dagle, J. M.; Weeks, D. L. *Nucleic Acids Res.* **1996,** *24,* 2143–2149. doi:10.1093/nar/24.11.2143
109. Laurent, A.; Naval, M.; Debart, F.; Vasseur, J.-J.; Rayner, B. *Nucleic Acids Res.* **1999,** *27,* 4151–4159. doi:10.1093/nar/27.21.4151
110. Michel, T.; Martinand-Mari, C.; Debart, F.; Lebleu, B.; Robbins, I.; Vasseur, J.-J. *Nucleic Acids Res.* **2003,** *31,* 5282–5290. doi:10.1093/nar/gkg733
111. Bailey, C. P.; Weeks, D. L.; Dagle, J. M. *Nucleic Acids Res.* **1998,** *26,* 4860–4867. doi:10.1093/nar/26.21.4860
112. Deglane, G.; Abes, S.; Michel, T.; Prévot, P.; Vives, E.; Debart, F.; Barvik, I.; Lebleu, B.; Vasseur, J.-J. *ChemBioChem* **2006,** *7,* 684–692. doi:10.1002/cbic.200500433
113. Fathi, R.; Huang, Q.; Syi, J. L.; Delaney, W.; Cook, A. F. *Bioconjugate Chem.* **1994,** *5,* 47–57. doi:10.1021/bc00025a007
114. Fathi, R.; Huang, Q.; Coppola, G.; Delaney, W.; Teasdale, R.; Krieg, A. M.; Cook, A. F. *Nucleic Acids Res.* **1994,** *22,* 5416–5424. doi:10.1093/nar/22.24.5416
115. Yanachkov, I.; Zavizion, B.; Metelev, V.; Stevens, L. J.; Tabatadze, Y.; Yanachkova, M.; Wright, G.; Krichevsky, A. M.; Tabatadze, D. R. *Org. Biomol. Chem.* **2017,** *15,* 1363–1380. doi:10.1039/c6ob02576e
116. Bailey, C.; Weeks, D. L. *Nucleic Acids Res.* **2000,** *28,* 1154–1161. doi:10.1093/nar/28.5.1154
117. Rahman, S. M. A.; Baba, T.; Kodama, T.; Islam, M. A.; Obika, S. *Bioorg. Med. Chem.* **2012,** *20,* 4098–4102. doi:10.1016/j.bmc.2012.05.009
118. Chen, M.; Gothelf, K. V. *Org. Biomol. Chem.* **2008,** *6,* 908–911. doi:10.1039/b717664c
119. Fidanza, J. A.; Ozaki, H.; McLaughlin, L. W. *J. Am. Chem. Soc.* **1992,** *114,* 5509–5517. doi:10.1021/ja00040a004
120. Lee, J. H.; Wong, N. Y.; Tan, L. H.; Wang, Z.; Lu, Y. *J. Am. Chem. Soc.* **2010,** *132,* 8906–8908. doi:10.1021/ja103739f
121. Vekhoff, P.; Halby, L.; Oussedik, K.; Dallavalle, S.; Merlini, L.; Mahieu, C.; Lansiaux, A.; Bailly, C.; Boutorine, A.; Pisano, C.; Giannini, G.; Alloatti, D.; Arimondo, P. B. *Bioconjugate Chem.* **2009,** *20,* 666–672. doi:10.1021/bc800494y
122. Iwamoto, N.; Oka, N.; Sato, T.; Wada, T. *Angew. Chem., Int. Ed.* **2009,** *48,* 496–499. doi:10.1002/anie.200804408
123. Iwamoto, N.; Butler, D. C. D.; Svrzikapa, N.; Mohapatra, S.; Zlatev, I.; Sah, D. W. Y.; Meena; Standley, S. M.; Lu, G.; Apponi, L. H.; Frank-Kamenetsky, M.; Zhang, J. J.; Vargeese, C.; Verdine, G. L. *Nat. Biotechnol.* **2017,** *35,* 845–851. doi:10.1038/nbt.3948
124. Su, Y.; Bayarjargal, M.; Hale, T. K.; Filichev, V. V. *Beilstein J. Org. Chem.* **2021,** *17,* 749–761. doi:10.3762/bjoc.17.65
125. Meng, M.; Ducho, C. *Beilstein J. Org. Chem.* **2018,** *14,* 1293–1308. doi:10.3762/bjoc.14.111
126. Gowers, D. M.; Bijapur, J.; Brown, T.; Fox, K. R. *Biochemistry* **1999,** *38,* 13747–13758. doi:10.1021/bi9911637

127. Bijapur, J.; Bergqvist, S.; Brown, T.; Keppler, M. D.; Fox, K. R. *Nucleic Acids Res.* **1999,** *27,* 1802–1809. doi:10.1093/nar/27.8.1802
128. Sollogoub, M.; Darby, R. A. J.; Cuenoud, B.; Brown, T.; Fox, K. R. *Biochemistry* **2002,** *41,* 7224–7231. doi:10.1021/bi020164n

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Synthetic strategies toward 1,3-oxathiolane nucleoside analogues

Umesh P. Aher[1], Dhananjai Srivastava[1], Girij P. Singh[1] and Jayashree B. S[*2]

Review

Open Access

Address:
[1]Chemical Research Department, Lupin Research Park, Lupin Limited, 46A/47A, Village Nande, Taluka Mulshi, Pune-412115, Maharashtra, India and [2]Department of Pharmaceutical Chemistry, Manipal College of Pharmaceutical Sciences, Manipal Academy of Higher Education, Manipal-576104, Karnataka, India

Email:
Jayashree B. S[*] - bs.jayashree@manipal.edu

* Corresponding author

Keywords:
chiral auxiliaries; enzymes; Lewis acids; N-glycosylation; 1,3-oxathiolane sugar and nucleosides; separation of racemic nucleosides; stereoselectivity

Beilstein J. Org. Chem. **2021,** *17,* 2680–2715.
https://doi.org/10.3762/bjoc.17.182

Received: 21 June 2021
Accepted: 14 October 2021
Published: 04 November 2021

This article is part of the thematic issue "Celebrating the role of chemistry in the success of oligonucleotides as therapeutics".

Guest Editors: P. Kumar and T. Brown

Abstract

Sugar-modified nucleosides have gained considerable attention in the scientific community, either for use as molecular probes or as therapeutic agents. When the methylene group of the ribose ring is replaced with a sulfur atom at the 3'-position, these compounds have proved to be structurally potent nucleoside analogues, and the best example is BCH-189. The majority of methods traditionally involves the chemical modification of nucleoside structures. It requires the creation of artificial sugars, which is accompanied by coupling nucleobases via N-glycosylation. However, over the last three decades, efforts were made for the synthesis of 1,3-oxathiolane nucleosides by selective N-glycosylation of carbohydrate precursors at C-1, and this approach has emerged as a strong alternative that allows simple modification. This review aims to provide a comprehensive overview on the reported methods in the literature to access 1,3-oxathiolane nucleosides. The first focus of this review is the construction of the 1,3-oxathiolane ring from different starting materials. The second focus involves the coupling of the 1,3-oxathiolane ring with different nucleobases in a way that only one isomer is produced in a stereoselective manner via N-glycosylation. An emphasis has been placed on the C–N-glycosidic bond constructed during the formation of the nucleoside analogue. The third focus is on the separation of enantiomers of 1,3-oxathiolane nucleosides via resolution methods. The chemical as well as enzymatic procedures are reviewed and segregated in this review for effective synthesis of 1,3-oxathiolane nucleoside analogues.

Introduction

Among all the biomolecules in an organism, nucleic acids, namely DNA and RNA, have the unique role of storing the genetic code – the nucleotide sequence that specifies the amino acid sequence of proteins that is essential for life on Earth. These molecules play a significant role in replication, transmission, and transcription of genetic material in life forms [1].

Structural analogues similar to the naturally occurring 2'-deoxynucleosides and ribonucleosides, the DNA and RNA building blocks, respectively, are expected to mimic their counterparts during DNA or RNA synthesis, a biological role that is crucial for cellular reproduction [2]. Most of the drugs that are incorporated in the viral DNA upon phosphorylation in vivo block the DNA polymerase enzyme. However, DNA polymerase recognizes 2',3'-dideoxynucleosides as substrates, which are incorporated into the growing DNA strand. However, the absence of a 3'-hydroxy group prevents further strand elongation. The anticancer and antiviral activity of 2',3'-dideoxynucleosides is mainly based on inhibition of DNA synthesis, either through the chain termination process or by competitive inhibition [3,4]. These compounds are the structural analogues of the naturally occurring 2'-deoxynucleosides, the building blocks of DNA.

The World Health Organization (WHO) newsroom announced the primary statistics that HIV and cancer remain a significant global public health issue, having claimed over 47.6 million lives so far [5,6]. The statistics confirm that 1 in 6 deaths happening globally are due to cancer [5]. In 2019, 690,000 people died from HIV-related causes worldwide and by the end of 2019, around 38 million people were living with HIV. From these, 1.7 million people were newly diagnosed [6]. Nucleoside analogues have been in clinical use for almost 50 years and have been the mainstay of treating patients with cancer and viral infections [7,8]. The 2',3'-dideoxynucleoside analogues, such as AZT (zidovudine), ddI (didanosine), ddC (zalcitabine), and d4T (stavudine), are modified examples of the natural nucleosides with β-D-configuration in the carbohydrate part. These molecules are known to have a common HIV transcriptase inhibition mechanism, in which cytoplasmic enzymes progressively phosphorylate the analogues to 5'-triphosphates. This then competes with the naturally occurring nucleoside triphosphate substrate to bind to cellular DNA polymerase and viral reverse transcriptase [9]. The effectiveness of nucleoside analogues depends on the ability to replicate naturally occurring nucleosides, interfering with viral as well as cellular enzymes and hampering essential metabolism processes of nucleic acid components. Therefore, it was assumed until recently that effective inhibition of the metabolic enzyme is only possible by D-nucleoside analogues, which have the stereochemistry of natural nucleosides. This was proved to be untrue when the antiviral activity of 1,3-oxathiolane nucleosides with L-configuration was discovered, and this led to the approval of 3TC (lamivudine, (−)-BCH-189, **1**) as an antiviral drug and, among many others, to the use of FTC (emtricitabine, **2**) and L-FMAU (clevudine). L-Nucleosides can have a comparable and often greater antiviral efficacy than the D-counterparts, with more favorable toxicological profiles and a greater stability [10]. A variety of nucleoside analogues as possible antiviral agents has appeared, possessing the unusual β-L-configuration. Work has been motivated by the fact that, while retaining strong antiviral and/or antibacterial activity, L-nucleosides are typically endowed with lower host toxicity [11,12]. The antiviral activity and cytotoxicity in MT-4 cells showed that racemic (±)-BCH-189 (**1c**) possesses lower anti-HIV activity (ID_{50} = 0.37–1.31 μM) than AZT (D-nucleoside, ID_{50} = 0.0048–0.0217 μM). However, (±)-BCH-189 (**1c**) appeared to be a more effective antiviral agent than AZT in PBM cells and U937 cells [13]. The BCH-189 core structure bears two stereocenters, and hence four stereoisomers are possible. The individual stereoisomers were also evaluated against HIV-1 activity in PBM cells and based on this study, it was found that out of four stereoisomers, the β-configured L-(−)-enantiomer **1** (EC_{50} = 0.02 μM) is more potent in primary human lymphocytes than the β-configured D-(+)-enantiomer **1a** (EC_{50} = 0.2 μM) in CEM cells [14]. Similarly, the 5-fluoro-substituted analogue of cytidine, i.e., β-configured L-(−)-enantiomer **2**, exhibits potential antiviral activity against HIV-1 (EC_{50} = 0.009 μM) in CEM cells. However, the corresponding D-(+)-enantiomer is less active against HIV-1 (EC_{50} = 0.84 μM) [15]. The fusion of an appropriate sugar element, carbacycle, or heterocyclic equivalent with an activated base results in the corresponding analogues of D- and L-configured nucleosides and other unnatural nucleoside derivatives [16-19]. Therefore, further demand for various effective chemical syntheses of these nucleoside analogues is rapidly growing.

The FDA has approved modified nucleoside analogues such as zidovudine, didanosine, zalcitabine, stavudine, lamivudine (**1**), and abacavir (a carbanucleoside) for treating HIV infection, along with protease and nonnucleoside reverse transcriptase inhibitors (NNRTIs). Phosphorylation of 1,3-oxathiolane nucleosides, such as 3TC (**1**) and FTC (**2**), occurs in vivo to compete with natural deoxynucleotides for incorporation into (viral) DNA. Chain elongation via reverse transcriptase is thus inhibited. This class of drugs is referred to as nucleoside reverse transcriptase inhibitors (NRTIs). In NRTIs, 3TC (**1**) possesses chemical and biological properties, a sulfur atom instead of C-3', and an unnatural L-configured sugar [20]. The presence of oxygen as a second heteroatom in the sugar ring was also found to result in anti-HIV activity in 1,3-dioxolane nucleosides [21]. A good example for the preparation of 2',3'-dideoxy-3'-oxacytidine in a stereospecific manner was reported by Chu et al. [22]. Choi et al. [23] produced the 5-fluoro-substituted analogue of a 1,3-oxathiolane nucleoside as a racemic mixture, and the enantiomers were separated using pig liver esterase (PLE) enzyme, which resulted in 5'-butyroyl ester derivatives. They further explained the higher antiviral activity and lower toxicity of the unnatural L-(−)-enantiomer over the D-(−)-enantiomer.

The enantiomers of natural nucleosides are known to have a greater biological activity since they possess structural and configurational similarity to naturally occurring counterparts. In turn, for oxathiolane nucleoside analogues, it was noticed that unnatural (−)-enantiomers have higher anti-HIV activity and lower toxicity in comparison to natural (+)-enantiomers. The activation of such analogues was established to occur preferentially by the enzymes (kinases) or by the target enzymes (polymerases), which may be responsible for such differences [24].

Initial results point at a conventional mechanism of action. Therein, the investigation of the cellular metabolism predicts triphosphate formation of the compounds by phosphorylation, and the resulting nucleotide is a selective inhibitor of the HIV-1 reverse transcriptase [25]. Nucleosides with sulfur atom-containing heterocyclic sugar rings at the 3'-position are important pharmaceutical substances. Two well-known important molecules in this category are lamivudine (**1**) and emtricitabine (**2**), as shown in Figure 1.

It was found that there is a remarkable reduction in deaths related to HIV/AIDS in the United States due to usage of combination drug therapies [26]. In these combination therapies, L-(−)-2',3'-dideoxy-3'-thiacytidine (**1**) is one of the standard

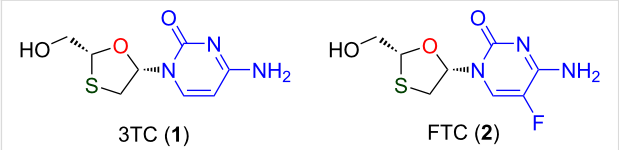


Figure 1: Representative modified 1,3-oxathiolane nucleoside analogues.

components, having an enhanced pharmacological profile over AZT and other dideoxynucleotide inhibitors [26-28]. 3TC (**1**) has a β-L-oxathiolane ring structure, instead of the ribose ring in the canonical nucleosides, and studies have shown that the triphosphate of **1** (i.e., 3TCTP) inhibits reverse transcriptase due to DNA chain termination [26,29,30]. In comparison to some of the other NRTIs that are hardly effective inhibitors of HIV-1 reverse transcriptase, 3TC (**1**) acts as a good substitute with lower toxicity. This could be because it is an unfavorable substrate for mitochondrial DNA polymerases [26,27,30]. The drug triphosphate interferes with HIV reverse transcriptase by competing with natural nucleotides for incorporation into the growing HIV DNA chain. The result if the triphosphate is taken up is the termination of the chain elongation because the drug lacks the 3'-hydroxy group on the deoxyribose ring that is necessary for the sugar–phosphate linking as shown in Figure 2.

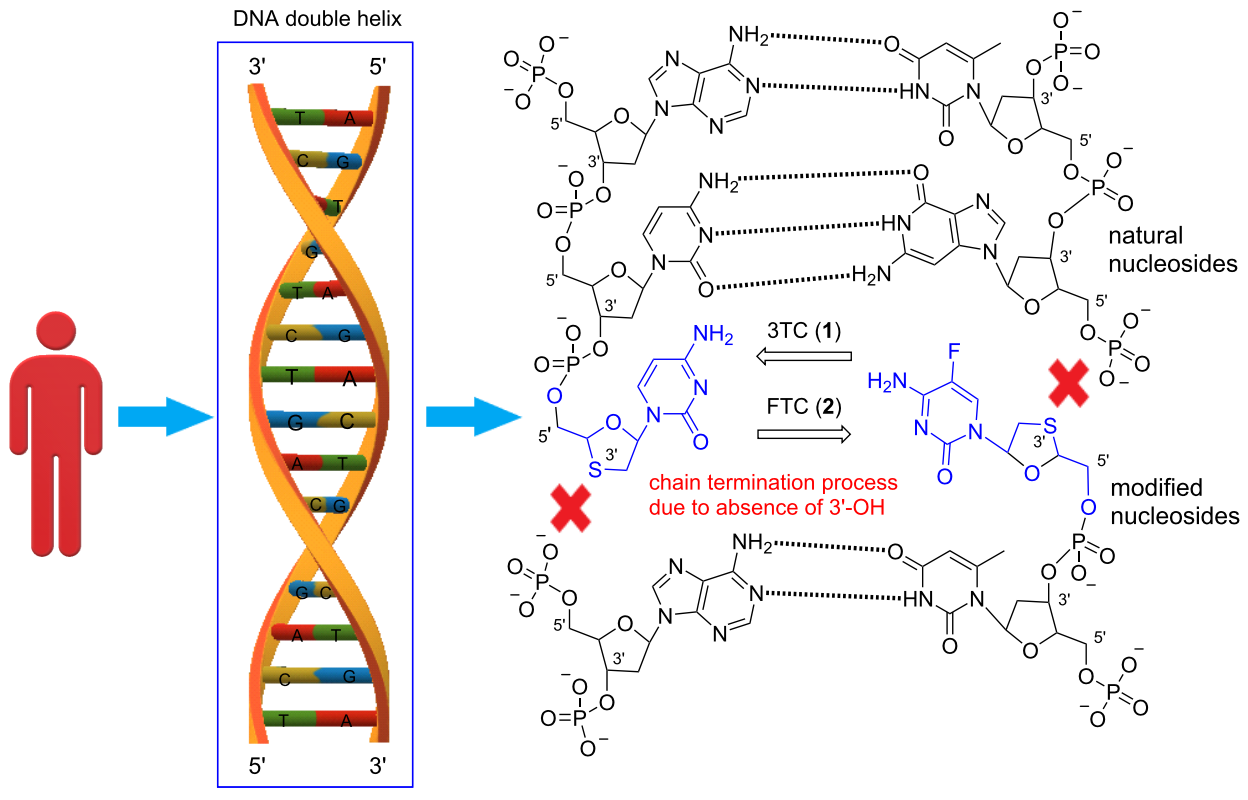


Figure 2: Mechanism of antiviral action of 1,3-oxathiolane nucleosides, 3TC (**1**) and FTC (**2**), as chain terminators.

The chemical approaches that were broadly used in the past to access these compounds are separated into two main groups: i) those that modify intact nucleosides by modifying the sugar, nucleobase, or both and ii) those that modify the sugar and introduce a nucleobase to a suitable position of the sugar. Since there is more than one chiral center in the structure of these nucleosides, the possibility of stereoisomer formation exists. In most cases, only one stereoisomer is found to be potent and the remaining, undesired isomers are significantly more toxic. Thus, it remains crucial for chemists to establish synthetic approaches toward single desired isomers. The methodologies for modified nucleosides are also known as linear approach and convergent approach [3]. We recognized that there are three major obstacles that have to be cleared: i) efficient preparation of the oxathiolane sugar ring, ii) a stereoselective N-glycosylation process that is compatible with an enantiomerically pure substrate, and iii) separation of enantiomers by chemical or enzymatic resolution methods. This review summarizes the methods used to synthesize 1,3-oxathiolane nucleosides. Many methods provide the formation of a diastereomeric mixture or a racemate of the resultant nucleosides [31]. However, the enantiomers of chiral drugs have indistinguishable chemical and physical properties in an achiral environment. One enantiomer may exhibit a more diverse pharmacological and chemical behavior than the other enantiomer in a chiral environment [32]. Additionally, on the grounds that living systems are, in a sense, themselves chiral, each of the enantiomers of a chiral drug can perform very differently in vivo. Therefore, there is a requirement to synthesize enantiomerically pure nucleosides that are free from undesired isomers.

Over the past three decades, several research groups have been working on devising novel methods for installing glycosidic linkages during the synthesis of modified nucleosides. For 1,3-oxathiolane nucleosides, to achieve β-selective glycosylation, a certain key intermediate was employed in the earlier studies, from 1989 to 2013. The several significant studies have been thoroughly reviewed in 2003 by Chu et al. [33]. The book contains a thorough section on the biological importance and synthesis of oxathiolane nucleosides. Herein, we tried to explore recent developments in comparison to previously reported methods to access 1,3-oxathiolane nucleosides. Similarly, a book chapter by D'alonzo and Guaragna published in 2013 summarizes the synthesis and biological applications of these important analogues [34]. However, a brief account is presented in this section for the sake of continuity.

While targeting to discover antiviral agents [35-37], particularly the class of dideoxynucleotides, it is essential to investigate possible fundamental alteration of the furanose ring and the practical and convenient synthesis of these analogues. These investigations are needed to improve the logic while depicting comparison with the established series of nucleoside analogues. In chiral synthesis, it is often important to establish the ratio of enantiomers before focusing on the isolation of a specific enantiomer. Therefore, having a good overview on enzymatic and chemical resolution methods, for example for the resolution of the oxathiolane nucleoside herein, is beneficial [24].

Review

Construction of the 1,3-oxathiolane sugar ring

The 1,3-oxathiolane ring structure has been known for a long time. However, in recent years, that ring has been utilized in place of the sugar ring in nucleoside analogues. The enantiomerically pure 1,3-oxathiolane core has been an important building block in precursors that result in a defined stereochemistry of the resultant nucleoside product after N-glycosylation. Dynamic kinetic resolution (DKR) is a processes that interconverts a racemic mixture into a single enantiomer via an in situ stereoinversion, and it was implemented in some of the examples described herein. Therefore, for the construction of the 1,3-oxathiolane sugar ring, an extensive number of efficient and environmentally friendly chemical and enzymatic approaches has been established (Figure 3).

Chemical approaches

Modified sugar rings containing a sulfur heteroatom at C-3' are found in medicinal chemistry. The reaction between oxygen-containing substrates (such as aldehydes or acetals) and sulfur sources (such as thiols or sulfenyl compounds) is one of the most important methods to give the 1,3-oxathiolane sugar ring. Herein, the research on 1,3-oxathiolane sugar ring formation strategies, mainly starting from oxygen- and sulfur-containing substrates, is summarized.

In 1989, Belleau and co-workers [38] produced the first oxathiolane nucleoside as a racemic mixture, popularly known as (±)-BCH-189 (**1c**). The key oxathiolane **4**, a precursor of the corresponding nucleoside, was obtained as a 1:1 mixture of anomers (60%) from benzoyloxyacetaldehyde (**3a**) and 2-mercapto-substituted dimethyl acetal **3na**. The reaction was performed in toluene in the presence of *p*-toluenesulfonic acid (*p*-TSA) catalyst at reflux (Scheme 1).

Sadayoshi and co-workers [39] developed the synthesis of 1,3-oxathiolane derivative **8** (Scheme 2). The protected glycolic aldehyde **3b** was isolated after ozonolysis of alkene **3ra**. The reaction between an aldehyde **3b** and 2-mercaptoacetic acid (**3o**) was carried out at reflux temperature in toluene to synthesize the 1,3-oxathiolane lactone **6** via intermediate **5** after elimi-

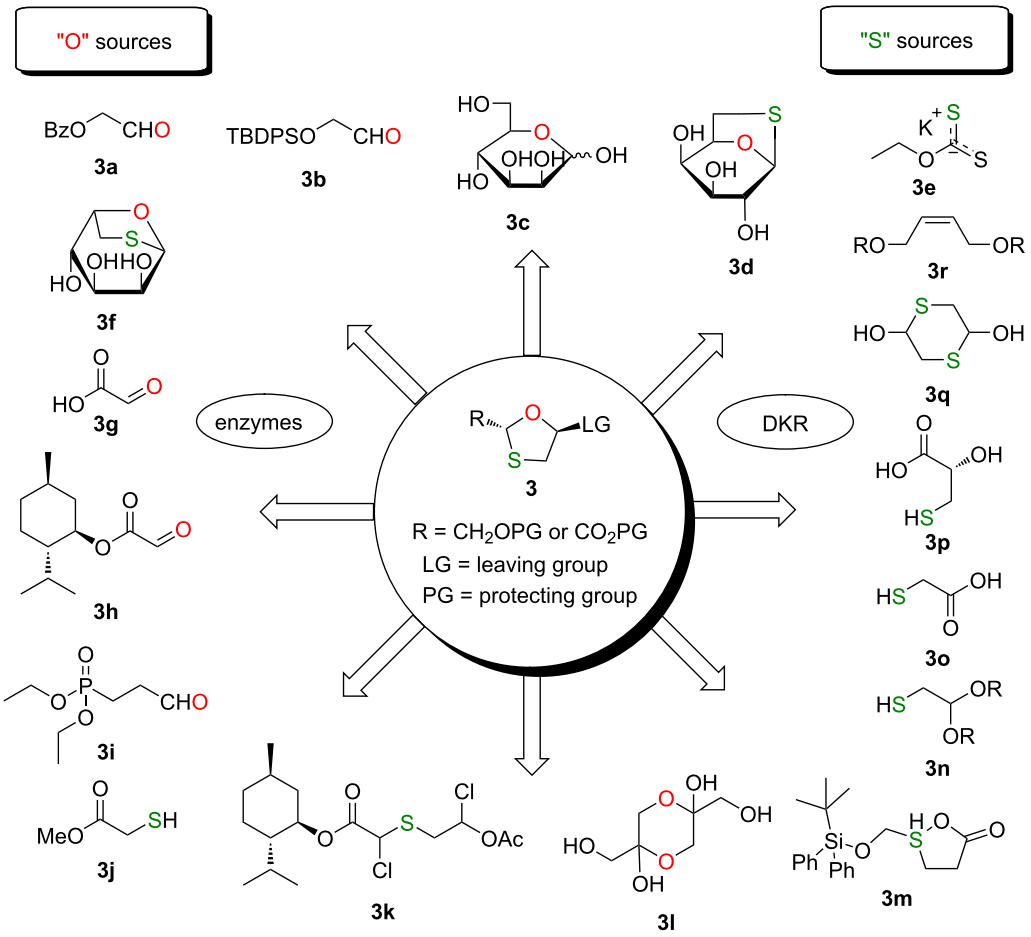
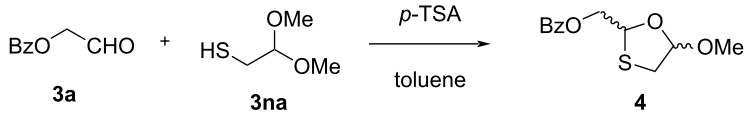
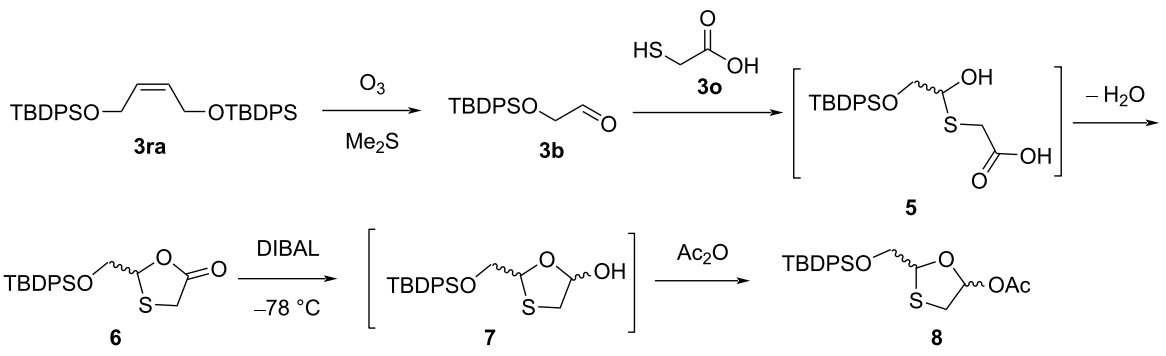


Figure 3: Synthetic strategies for the construction of the 1,3-oxathiolane sugar ring.



Scheme 1: Synthesis of **4** from benzoyloxyacetaldehyde (**3a**) and 2-mercapto-substituted dimethyl acetal **3na**.

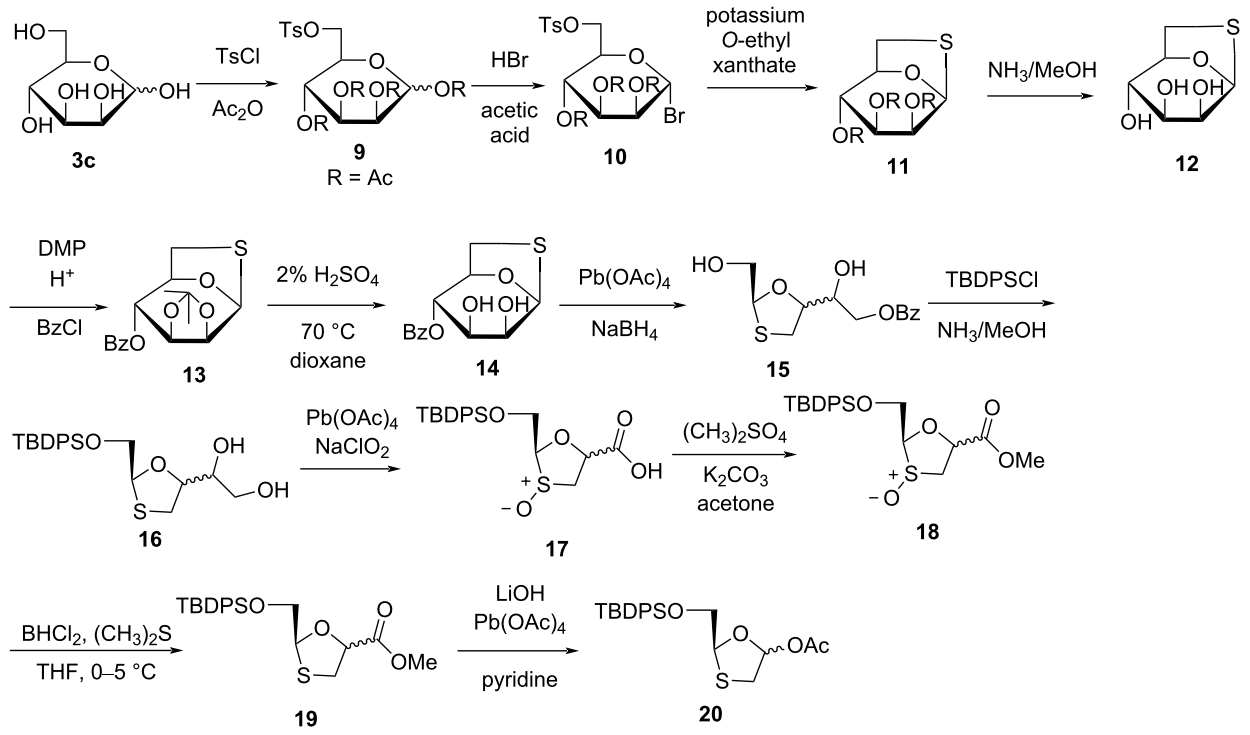


Scheme 2: Synthesis of **8** from protected glycolic aldehyde **3b** and 2-mercaptoacetic acid (**3o**).

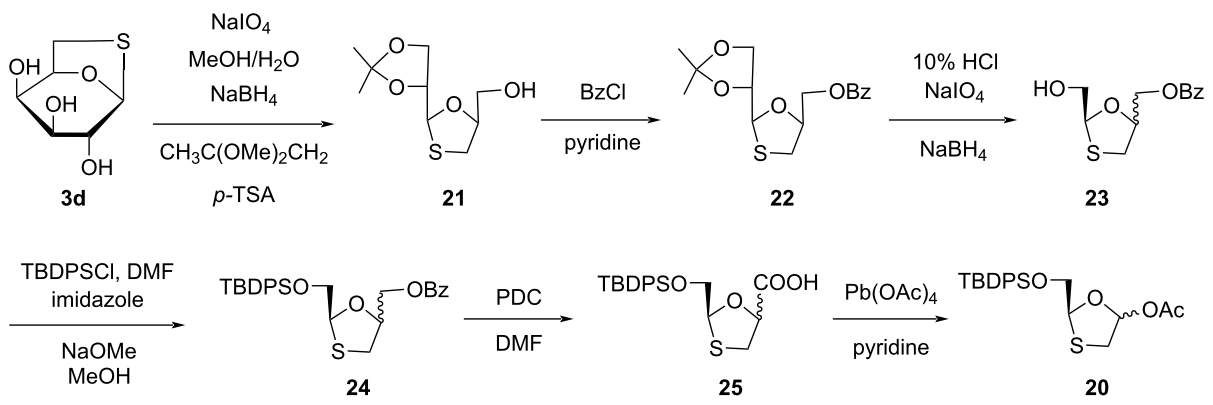
nation of a water molecule. This was further reduced with diisobutylaluminum hydride (DIBAL) in toluene at −78 °C or by lithium tri-*tert*-butoxyaluminum hydride in THF at 0 °C to obtain lactol **7**, which was subsequently acetylated with acetic anhydride to afford **8** as a 2:1 mixture of anomers.

Chu and co-workers [40] applied a novel strategy for the synthesis of enantiomerically pure (+)-BCH-189 (**1a**) using D-mannose (**3c**) as a starting material (Scheme 3). 1,2,3,4-Tetraacetyl-D-mannose derivative **9** was prepared from D-mannose (**3c**) by protecting the primary alcohol with a tosyl group, followed by protection of the four hydroxy groups by acetylation. Further, bromo-substituted sugar compound **10** was obtained by a bromination reaction of the anomeric acetyl group. 1,6-Thioanhydro-β-mannose derivative **11** was obtained by cyclization with 3 equivalents of potassium *O*-ethyl xanthate. It was then treated with a methanolic ammonia solution to give triol compound **12**. The protection of the *cis*-2,3-vicinal hydroxy groups of **12** with an isopropylidene, followed by benzoylation, gave compound **13**. Using 2% aqueous sulfuric acid, the isopropylidene group of **13** was selectively deprotected at 70 °C in dioxane to obtain diol **14**. This diol was further cleaved using lead tetraacetate, and further reduction with sodium borohydride produced compound **15**. The 5'-hydroxy group of **15** was then treated with *tert*-butyldiphenylsilyl chloride (TBDPSCl) for selective protection. The compound was further debenzoylated by ammonolysis, which gave compound **16**. Compound **16** underwent oxidative cleavage using lead tetraacetate, and the intermediate aldehyde was oxidized to the carboxylic acid using sodium chlorite, which afforded acid derivative **17**. This was obtained as a mixture of *endo*- and *exo*-sulfoxides. Esterification of **17** was carried out with dimethyl sulfate to give methyl ester **18**, which was further reduced using dichloroborane and dimethyl sulfide to provide sulfide **19** in 80% yield in THF as solvent. Hydrolysis of compound **19** provided the corresponding carboxylic acid, and further oxidative decarboxylation with lead tetraacetate and pyridine provided oxathiolane **20**.

Chu and co-workers [41,42] further established a more proficient system for the synthesis of (+)-BCH-189 (**1a**) from 1,6-thioanhydro-D-galactose (**3d**, Scheme 4). Sodium periodate was used for oxidative cleavage of *cis*-diol **3d**. The subsequent aldehyde was then converted to a vicinal diol by reduction with sodium borohydride. Further, it was protected by 2,2-dimethoxypropane to give the 1,3-oxathiolane derivative **21**. The benzoylated compound **22** was obtained by reaction of benzoyl chloride in pyridine to protect the hydroxy group, which results in a high yield. The isopropylidene group was selectively deprotected using 10% HCl, followed by oxidative breakage of the carbon–carbon bond of the resulting diol using sodium periodate. Further reduction of the aldehyde into a primary



Scheme 3: Synthesis of **20** from D-mannose (**3c**).



Scheme 4: Synthesis of **20** from 1,6-thioanhydro-D-galactose (**3d**).

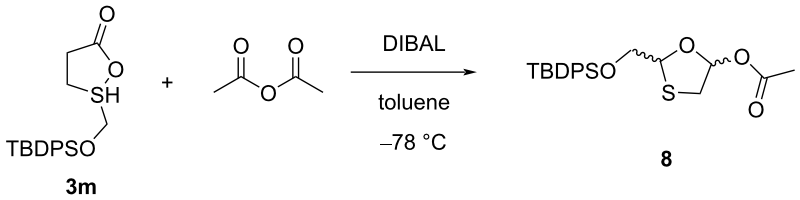
alcohol with sodium borohydride affords compound **23**. The protection of the hydroxy group of compound **23** was carried out by TBDPSCl in the presence of imidazole and *N*,*N*-dimethylformamide (DMF) as solvent, and deprotection of the benzoyl group by ammonolysis provides silylated compound **24**. Reaction of **24** with pyridinium dichromate (PDC) in DMF solvent afforded the acid derivative **25**. This derivative was converted to the key intermediate **20** by oxidative decarboxylation [33].

Han et al. [43] developed a method for the novel oxathiolane intermediate 2-(*tert*-butyldiphenylsilyloxy)methyl-5-acetoxy-1,3-oxathiolane (**8**) from 2-(*tert*-butyldiphenylsilyloxy)methyl-5-oxo-1,2-oxathiolane (**3m**, Scheme 5). Compound **3m** was dissolved in toluene and cooled to −78 °C. Further, a DIBAL solution was added slowly while maintaining the reaction temperature below −70 °C. The reaction mixture was further treated with acetic anhydride at room temperature. After workup by adding water and diethyl ether, the reaction mass was filtered and distilled until a residue was obtained. The colorless liquid compound **8** was obtained in 64% yield (as 6:1 mixture of anomers) after flash chromatography with 20% ethyl acetate in hexanes.

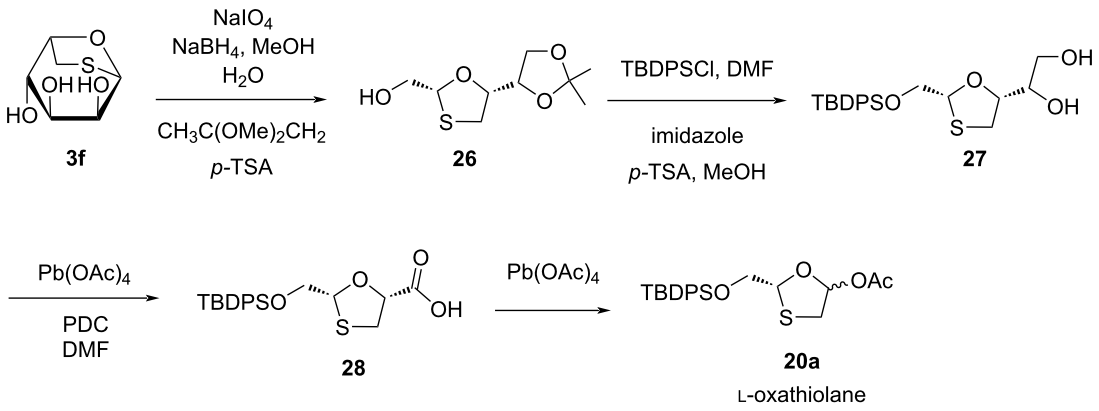
Chu and colleagues [44] constructed a synthetic approach to access (−)-BCH-189 (**1**) from L-gulose derivative **3f** (Scheme 6). Compound **26** was obtained by oxidation, reduction, and protection of the primary hydroxy group from **3f**. Further, lead tetraacetate directly cleaved diol **27** at room temperature, and oxidation with a mild oxidizing agent, PDC, provided **28**. Using the reaction of lead tetraacetate with **28** via oxidative decarboxylation afforded oxathiolane acetate derivative **20a**.

The synthesis of a 1,3-oxathiolane precursor required for the preparation of 3TC (**1**) in four steps was reported by Humber et al. [45]. They started with a coupling reaction of (+)-thiolactic acid **3p** and 2-benzoyloxyacetaldehyde (**3a**) using boron trifluoride etherate. A diastereomeric mixture of oxathiolane acids **29** and **30** was prepared in a 1:2 ratio in good yield (Scheme 7). Further separation of the diastereomers by silica gel column chromatography and reaction with lead tetraacetate provided the key oxathiolane derivative **31**.

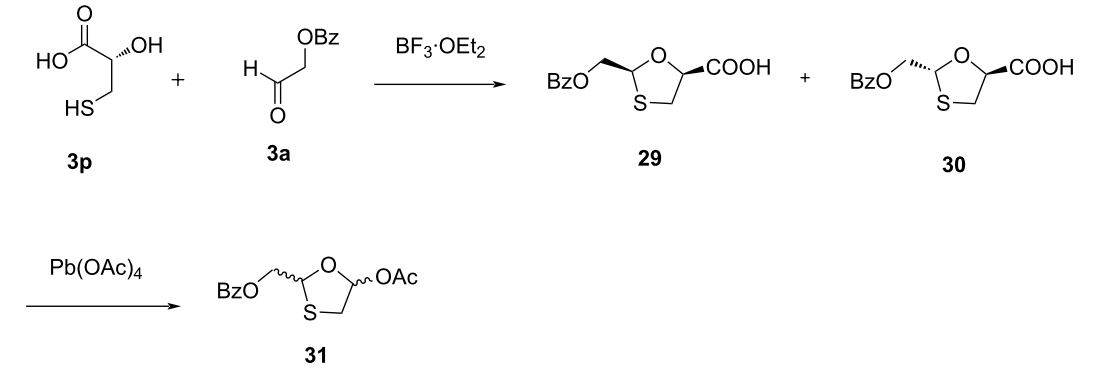
In 1995, Jin et al. [46] carried out the reaction of 1,4-dithiane-2,5-diol (**3q**) with glyoxylic acid (**3g**) hydrate at reflux temperature in *tert*-butyl methyl ether, which provided the hydroxyoxathiolane **32**. Further, acetylation of the hydroxyoxathiolane in the presence of methanesulfonic acid gave a 1:2 mixture of the *trans*-diastereomer **33** and the *cis*-diastereomer **34**. The esterification using L-menthol as a chiral auxiliary resulted in a diastereomeric mixture, which was successfully recrystallized to obtain the enantiomerically pure L-menthyl ester **35a** (Scheme 8).



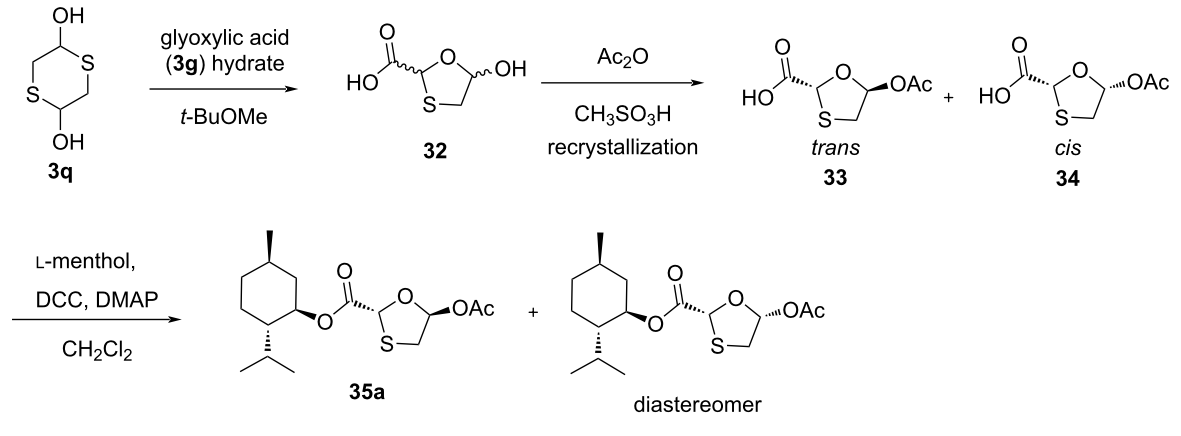
Scheme 5: Synthesis of **8** from 2-(*tert*-butyldiphenylsilyloxy)methyl-5-oxo-1,2-oxathiolane (**3m**).



Scheme 6: Synthesis of **20a** from L-gulose derivative **3f**.



Scheme 7: Synthesis of **31** from (+)-thiolactic acid **3p** and 2-benzoyloxyacetaldehyde (**3a**).

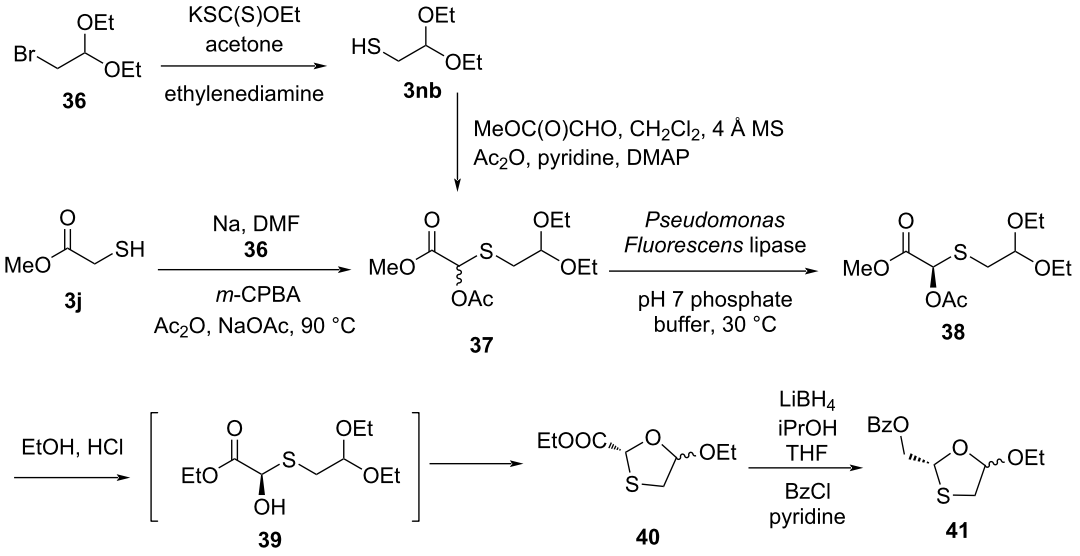


Scheme 8: Synthesis of **35a** from 1,4-dithiane-2,5-diol (**3q**) and glyoxylic acid (**3g**) hydrate.

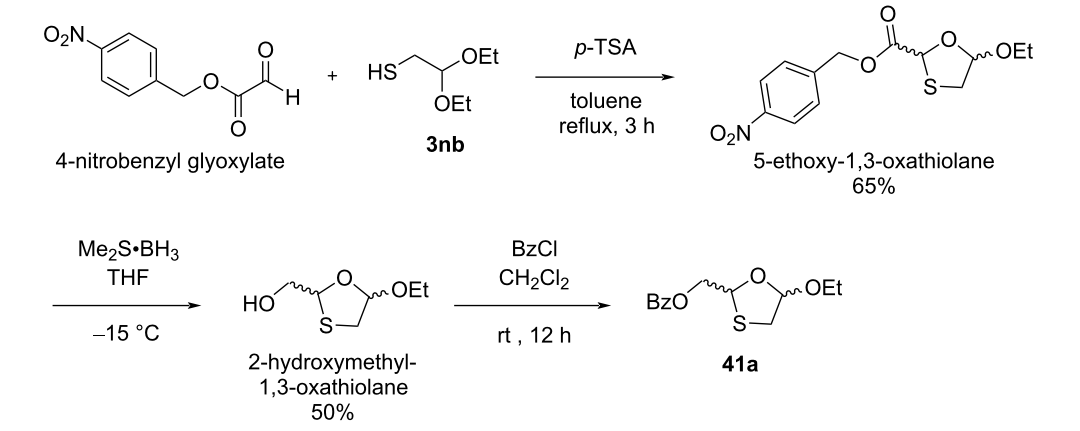
Milton et al. [47] synthesized the key intermediate **38** by two synthetic routes. The first route involves a reaction of bromoacetaldehyde diethyl acetal (**36**) with a xanthate ester, followed by treatment of ethylenediamine, which afforded the thiol compound **3nb**. Further treatment of the thiol **3nb** with methyl glyoxylate in dichloromethane solvent along with molecular sieves (4 Å), followed by in situ acetylation using Ac_2O, pyridine, and catalytic 4-(*N*,*N*-dimethylamino)pyridine

(DMAP) provided compound **37**. The second route involves condensation of the sodium salt of methyl 2-mercaptoacetate (**3j**) with bromoacetaldehyde diethyl acetal (**36**) in DMF solvent and further oxidation of the sulfide using *m*-CPBA, followed by Pummerer rearrangement using Ac_2O and sodium acetate at 90 °C, which provides compound **37** (Scheme 9). α-Acetoxy sulfide intermediate **37** was resolved using a lipase in *t*-BuOMe, resulting in a high enantiomeric excess. They used an enzymatic resolution of an acetoxy sulfide with a *Pseudomonas fluorescens* lipase to obtain compound **38**. Reaction of chiral acetoxy sulfide **38** with HCl in dry ethanol induced acetate removal by transesterification to give the hemithioacetal **39**, which cyclized to the oxathiolane **40** in situ with minor isomerization. The reduction of the ester group with $LiAlH_4$, followed by benzoylation using benzoyl chloride and pyridine gave 1,3-oxathiolane derivative **41**.

Kraus and Attardo [48] developed new strategies for the synthesis of a new 2,5-substituted 1,3-oxathiolane intermediate (Scheme 10). The approach involved the cyclocondensation reaction of anhydrous 4-nitrobenzyl glyoxylate with mercaptoacetaldehyde diethyl acetal (**3nb**) at reflux temperature in toluene solvent. This led to the formation of a 5-ethoxy-1,3-oxathiolane derivative. Further, reduction of the ester functionality with borane dimethyl sulfide at −15 °C afforded the corresponding 2-(hydroxymethyl)-1,3-oxathiolane in 50% yield. It was further treated with benzoyl chloride in the presence of triethylamine (TEA), which provided the desired compound **41a** as 1:1 mixture of *cis*- and *trans*-isomers.



Scheme 9: Synthetic routes toward **41** through Pummerer reaction from methyl 2-mercaptoacetate (**3j**) and bromoacetaldehyde diethyl acetal (**36**).



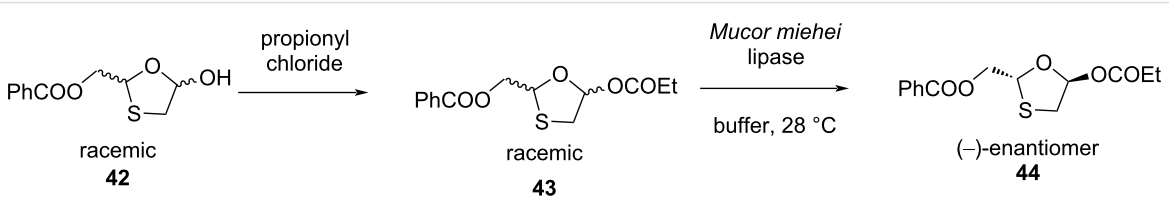
Scheme 10: Strategy for the synthesis of 2,5-substituted 1,3-oxathiolane **41a** using 4-nitrobenzyl glyoxylate and mercaptoacetaldehyde diethyl acetal (**3nb**).

In 1995, Cousins and co-workers [49] investigated enzymatic methods to resolve an oxathiolane intermediate (Scheme 11). Racemic intermediate **42** was converted into **43** with propionyl chloride protection. The procedure provides the enzymatic resolution of oxathiolane propionate derivative **43** by using *Mucor miehei* lipase, which affords (−)-enantiomer **44** as residual substrate. This enantioenriched precursor was useful to obtain the pure corresponding nucleoside analogue.

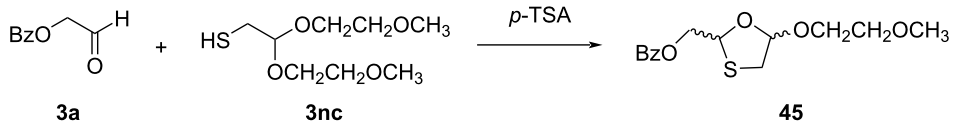
Faury and co-workers [50] synthesized the tetrazole analogues of 1,3-oxathiolane nucleosides to show the antiviral activity in comparison to ribavirin. The condensation reaction between benzoyloxyacetaldehyde (**3a**) and 2-mercaptoacetaldehyde bis(2-methoxyethyl) acetal (**3nc**) in the presence of *p*-TSA as catalyst afforded the intermediate 2-benzoyloxymethyl-5-(2-methoxyethyloxy)-1,3-oxathiolane **45** (Scheme 12).

Kraus [51] developed the cyclocondensation of 2-mercaptoacetaldehyde bis(2-methoxyethyl) acetal (**3nc**) with diethyl 3-phosphonoaldehyde (**3i**) to provide the novel oxathiolane intermediate **46** (Scheme 13). The reaction was carried out in the presence of *p*-TSA at reflux temperature in toluene solvent.

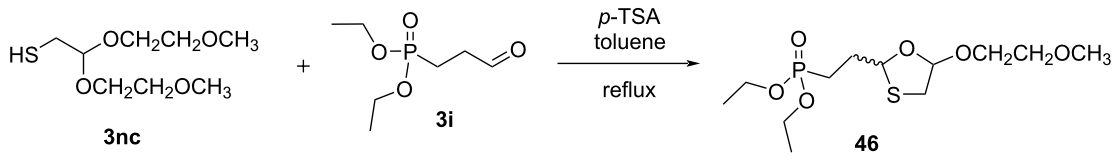
The synthesis and antiviral evaluation of 4'-(hydroxymethyl)oxathiolane nucleosides was reported by Chao and Nair [52]. The synthetic approach used 1,3-dihydroxyacetone dimer **3l** (Scheme 14). This dimer, upon acetylation using acetic anhydride in pyridine, produced compound **47**. Further, cyclocondensation of **47** with 2-mercaptoacetaldehyde diethyl acetal (**3nb**) in the presence of *p*-TSA in benzene solvent afforded 1,3-oxathiolane intermediate (±)-2,2-bis(acetoxymethyl)-5-ethoxy-1,3-thioxalane (**48**).



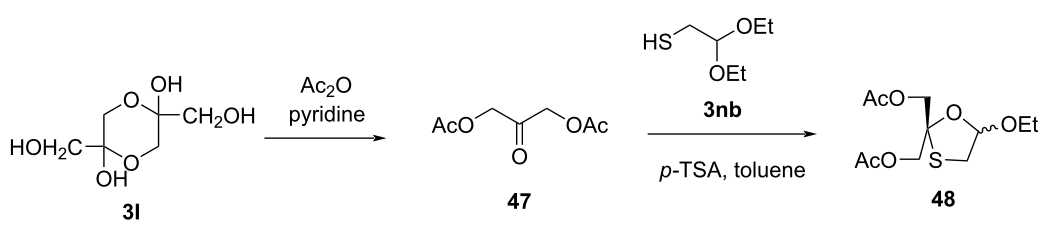
Scheme 11: Synthesis of **44** by a resolution method using *Mucor miehei* lipase.



Scheme 12: Synthesis of **45** from benzoyloxyacetaldehyde (**3a**) and 2-mercaptoacetaldehyde bis(2-methoxyethyl) acetal (**3nc**).



Scheme 13: Synthesis of **46** from 2-mercaptoacetaldehyde bis(2-methoxyethyl) acetal (**3nc**) and diethyl 3-phosphonoaldehyde **3i**.



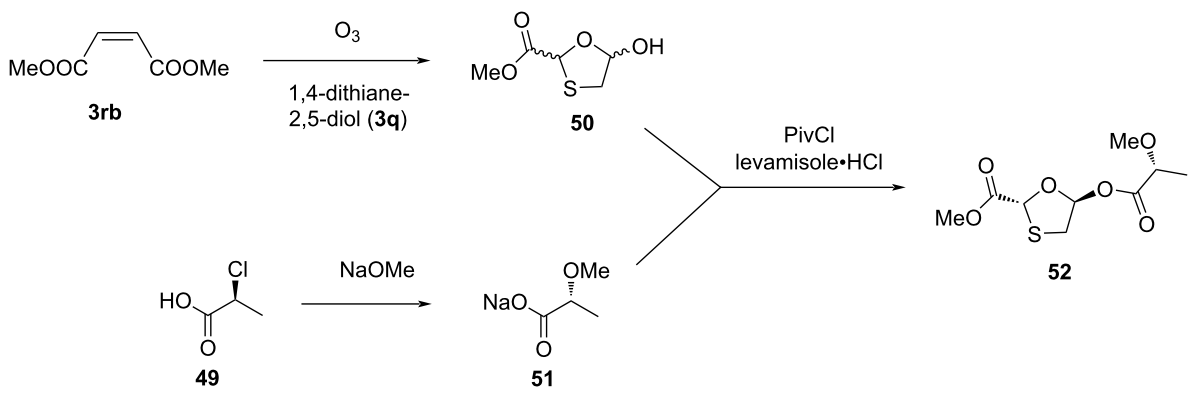
Scheme 14: Synthesis of **48** from 1,3-dihydroxyacetone dimer **3l**.

More recently, an approach developed by Snead et al. [53] at the Medicines for All Institute used lactic acid derivatives to test the impact of a chiral auxiliary on N-glycosylation. Compound **50** was synthesized by ozonolysis of alkene **3rb**, followed by reaction of aldehyde (generated in situ from alkene) with 1,4-dithiane-2,5-diol (**3q**). The use of lactic acid derivatives provided both enantiomers of oxathiolane precursors. The use of an (*S*)-lactic acid derivative resulted in the formation of an oxathiolane precursor with the opposite configuration of the desired one, which eventually led to the opposite enantiomer of 3TC (**1a**). Therefore, the authors changed the procedure and used the (*R*)-lactic acid derivative **51** to facilitate the formation of 3TC (**1**). In this procedure, the oxathiolane **50** was acylated using the lactic acid derivative sodium (*R*)-2-methoxypropanoate (**51**), which provided the derivative **52** (Scheme 15). Compound **51** was obtained by reaction of (*S*)-2-chloropropanoic acid (**49**) with sodium methoxide. Further, selective recrystallization in an appropriate solvent (toluene/hexanes) resulted in a single isomer (50:1 dr) in solution. The oxathiolane derivative **52** has the opposite configuration of that required for 3TC (**1**) synthesis. This acylation reaction was accomplished using pivaloyl chloride in the presence of levamisole, which gave an improved overall yield of **52** of up to 67%. In this approach, the required stereochemistry of the thioacetal compound was created, so that the coupling with a nucleobase in a further step determines the stereochemistry of the attaching nucleobase at the anomeric center, which is governed by an anchimeric effect. Thus, the method determines the configuration of proximal as well as remote stereocenters in a single step, and both enantiomers of the β-nucleoside were accessed from affordable starting materials.

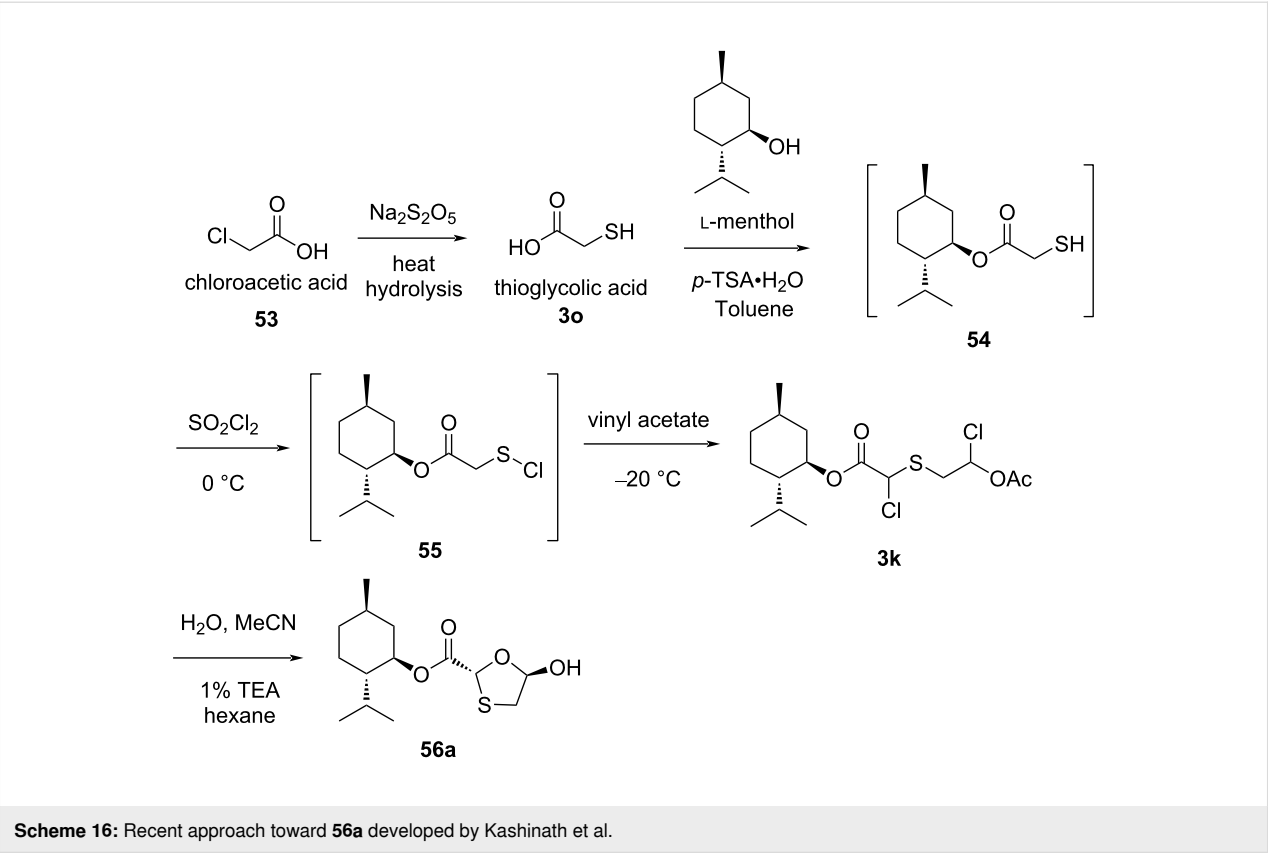
Kashinath and co-workers [54] also identified an innovative route to access an oxathiolane intermediate, which was further used for the synthesis of lamivudine (**1**) as well as emtricitabine (**2**). They presented an efficient reaction path that utilized commonly available and inexpensive starting materials. Sulfenyl chloride chemistry was used to synthesize the oxathiolane precursor **56a** from acyclic precursors. The method used chloroacetic acid (**53**), vinyl acetate, sodium thiosulfate, and water to construct the oxathiolane moiety. The use of sulfenyl chloride provided a new method to access such oxathiolanes (Scheme 16). Thioglycolic acid (**3o**), upon reaction with L-menthol, afforded the relevant thiol-substituted esters **54**, which further reacted with sulfuryl chloride to give compound **55**. The reaction of compound **55** with vinyl acetate constructed a sulfur–carbon bond and produced **3k**. The sulfuryl chloride reagent simultaneously allowed for chlorination at the α-position of the ester. The dichloro-substituted intermediate **3k** was further cyclized to produce the oxathiolane **56a** by reaction with water in the presence of acetonitrile as solvent. The focus of this novel route was to access basic reagents that are useful for the synthesis of 3TC (**1**) and FTC (**2**).

One of the methods of choice for the industrial manufacturing of lamivudine (**1**) follows the procedure suggested by Whitehead et al. [55]. This procedure involves the use of compound **56a**, where an L-menthyl moiety as chiral auxiliary is connected to an enantiomerically pure oxathiolane-based lactol. This is a necessary requirement to produce the desired stereochemistry in the product. It was extensively reported that the β-selectivity could be due to the formation of an oxonium ion, which is stabilized through anchimeric assistance of the L-menthyl ester function. The method requires highly effective crystallization-induced DKR to achieve an efficient synthesis of enantiomerically pure oxathiolane-based lactol **56a** from L-menthyl glyoxylate (**3h**) monohydrate and 1,4-dithiane-2,5-diol (**3q**, Scheme 17).

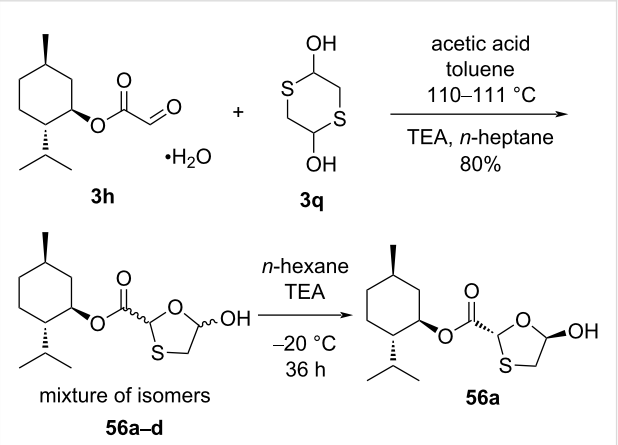
The investigation proved that the base TEA was capable of effecting the equilibration at C-2 but advantageous for the crystallization process. A number of bases was also evaluated by this research group: pyridine gave only a small amount of inter-



Scheme 15: Approach toward **52** from protected alkene **3rb** and lactic acid derivative **51** developed by Snead et al.



Scheme 16: Recent approach toward **56a** developed by Kashinath et al.



Scheme 17: Synthesis of **56a** from L-menthyl glyoxylate (**3h**) hydrate by DKR.

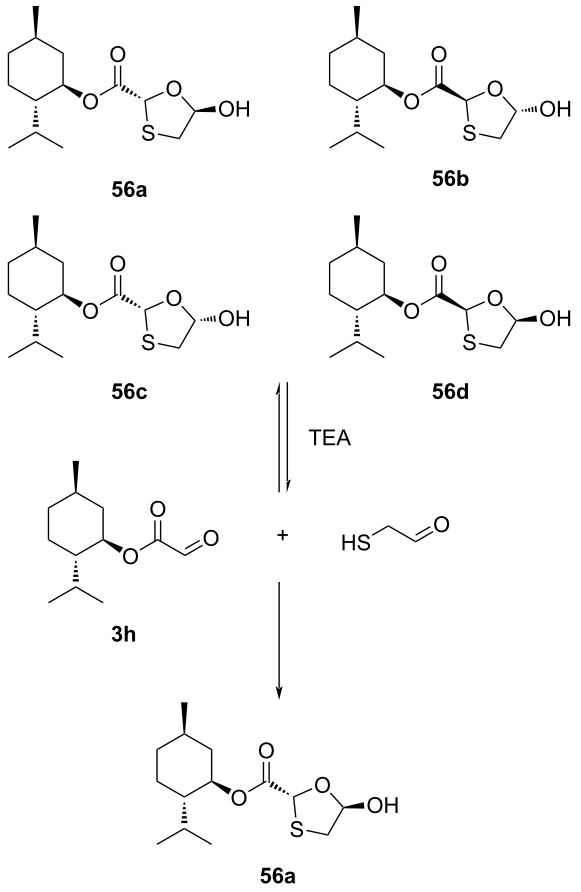
conversion, whereas TEA caused rapid interconversion. Furthermore, it was discovered that instant interconversion and crystallization of **56a** in 80% yield (Scheme 18) was possible through a mechanism that required the addition of a catalytic amount of TEA.

A method was established utilizing a Vorbrüggen reaction [55,56] of 5-acetoxyoxathiolane **35a**, which is an enantiomerically pure compound that can be used for the synthesis of lamivudine (**1**), as summarized in the next section. Access to crystalline 5-acetoxyoxathiolane **35a** was accomplished either by selective crystallization in the presence of the remaining diastereoisomers, although in only 16% yield, or by a classical resolution method using the norephedrine salt **58** (Scheme 19). The other diastereomer **59** remained dissolved in the mother liquor. The treatment of the norephedrine salt **58** with 5 M HCl afforded the enantiopure acid **60**, which was further converted to the desired 1,3-oxathiolane-substituted L-menthyl ester **35a**.

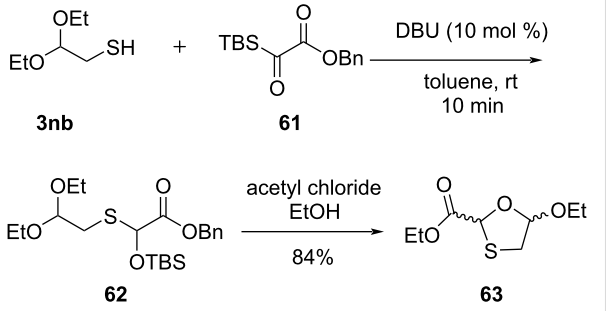
The synthetic use of [1,2]-Brook rearrangement for the synthesis of lamivudine (**1**) and the opposite enantiomer **1a** was demonstrated by Han et al. [57]. They carried out the [1,2]-Brook rearrangement of silyl glyoxylate **61** using thiol **3nb** as the nucleophile. Under optimized conditions, the reaction of the key intermediate **62** with acetyl chloride in ethanol results in the formation of the 1,3-oxathiolane species **63** (Scheme 20).

Enzymatic approaches

1,3-Oxathiolanes have shown broad biological activities, including the most important intermediates in the synthesis of the pharmaceuticals lamivudine (**1**) and emtricitabine (**2**), which have been approved as drugs to treat HIV infection [58] as well as human chronic hepatitis B [59]. Using asymmetric synthesis



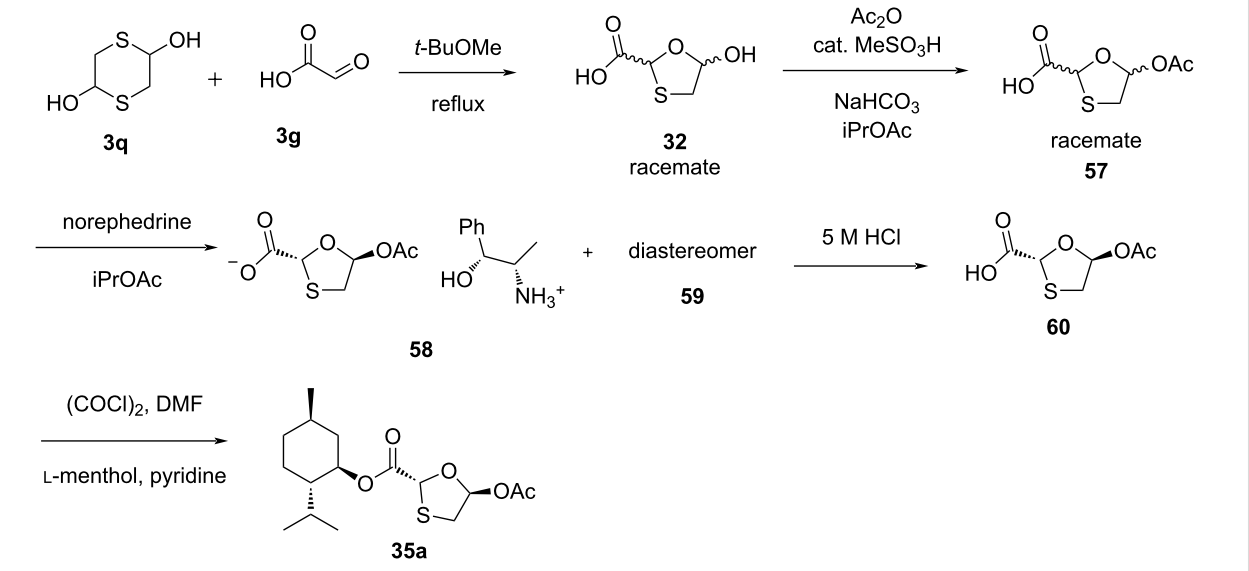
Scheme 18: Possible mechanism with catalytic TEA for rapid interconversion of isomers.



Scheme 20: Synthesis of **63** via [1,2]-Brook rearrangement from silyl glyoxylate **61** and thiol **3nb**.

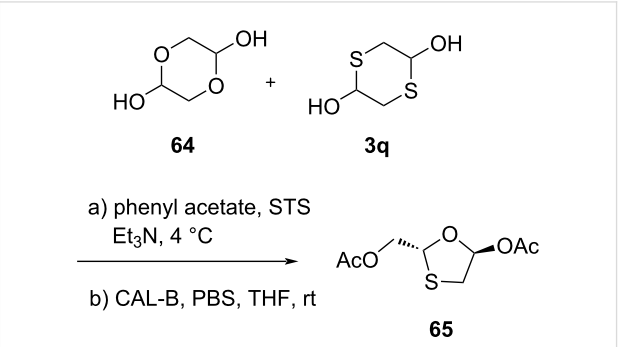
or resolution with appropriate enzymes to prepare these enantiopure 1,3-oxathiolanes has gained extensive attention due to the good stereoselectivity, high efficiency, mild reaction conditions, and eco-friendliness.

Ren and colleagues [59] recently reported the preparation of an enantiopure 1,3-oxathiolane **65** utilizing a multienzymatic cascade protocol (Scheme 21). The combined use of surfactant-treated Subtilisin Carlsberg (STS) and *Candida antarctica* lipase B (CAL-B) resulted in the 1,3-oxathiolane ring in THF and phosphate-buffered saline (PBS). The reaction used **64** and **3q** as starting materials and was stereocontrolled efficiently, providing an enantiomeric excess of about >99%. The subsequent N-glycosylation further provided enantiopure lamivudine (**1**). Hu et al. [60] explained that chiral HPLC and nuclear Overhauser effect (NOE) NMR spectroscopy are useful tools to



Scheme 19: Synthesis of **35a** by a classical resolution method through norephedrine salt **58** formation.

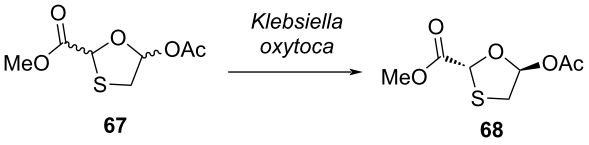
monitor and control the chirality when utilizing a modified 1,3-oxathiolane intermediate **65** obtained via enzyme-catalyzed selective hydrolysis.



Scheme 21: Combined use of STS and CAL-B as catalysts to synthesize an enantiopure oxathiolane precursor **65**.

Hu et al. [61] established a green catalyst, STS, for the asymmetric synthesis of lamivudine (**1**). Specifically, this approach used enzyme optimization techniques to efficiently synthesize highly enantiopure nucleoside analogues. The group found that the stereochemistry of the target molecules was selectively obtained using different enzymes. Importantly, the stereochemistry of the 1,3-oxathiolane intermediates **65** and **66** could be controlled well (Scheme 22). The glycolaldehyde dimer **64** and 1,4-dithiane-2,5-diol (**3q**) were reacted in the presence of TEA and the acyl donor phenyl acetate. The presence of CAL-B allowed the formation of the intermediate **66** and ultimately the corresponding nucleoside **1a** in a protocol by Vorbrüggen et al. In turn, using STS, this valuable asymmetric synthesis provided the intermediate **65**, which led to lamivudine (**1**).

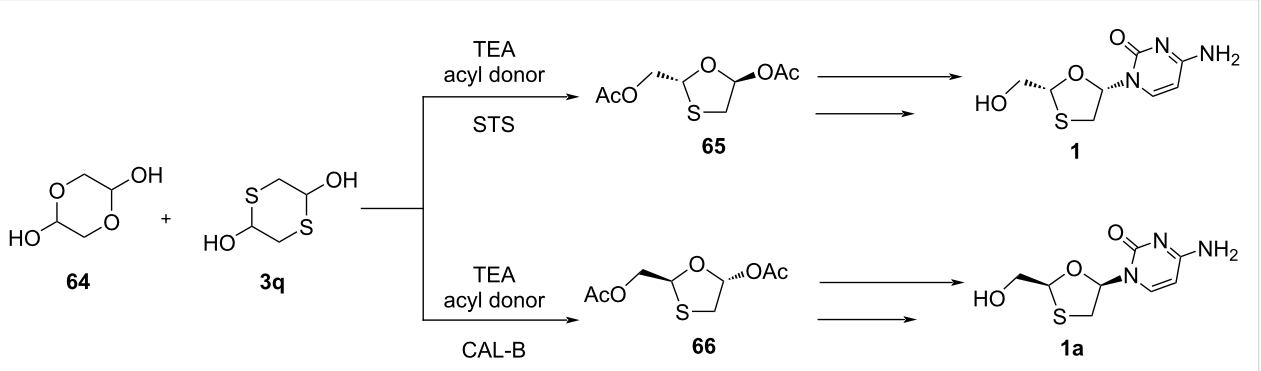
Recently, Chen at al. [62] reported the isolation of the strain *Klebsiella oxytoca* from soil by a target-oriented process, and it was utilized as a catalyst for the asymmetric preparation of a chiral intermediate of the antiviral agent lamivudine (**1**, Scheme 23). Further, the reaction conditions were optimized, and a series of factors was explored, including pH value, concentration, temperature, as well as the presence of metal ions and surfactant. Exceptionally, the end product was obtained in 99.9% ee by using whole-cell *Klebsiella oxytoca* catalysis and enantioselective resolution of the racemic mixture at 30 °C, pH 7.0, a substrate concentration of 1.5 g/L, and no additives. As compared to nearly all of the lipase-catalyzed methods to produce the chiral oxathiolane precursor **68** of lamivudine (**1**) from a mixture of isomers, i.e., **67**, the reaction occurred in a single-phase aqueous system, which may be considered a green chemistry approach.



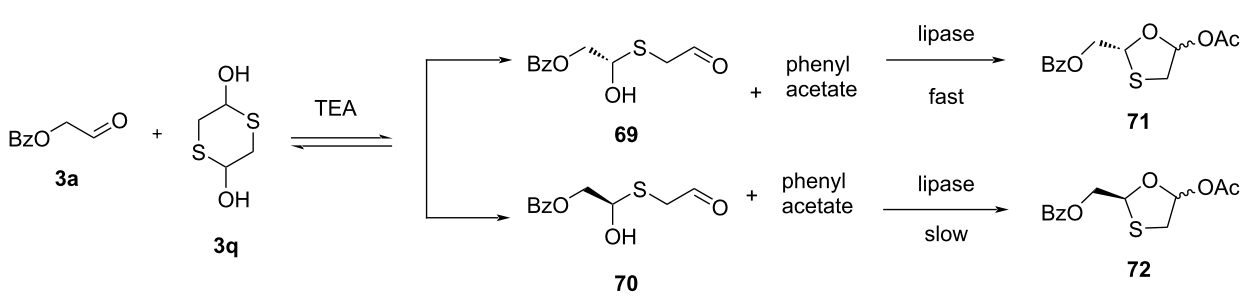
Scheme 23: Synthesis of **68** by using *Klebsiella oxytoca*.

Recently, Zhang and co-workers [63] developed a one-pot enzymatic synthesis of enantiopure 1,3-oxathiolane with *Trichosporon laibachii* lipase and a kinetic resolution. The synthesis of enantiopure ((*R*)-5-acetoxy-1,3-oxathiolan-2-yl)methyl benzoate (**71**) was carried out from the substrates **3a**, 1,4-dithiane-2,5-diol (**3q**), and phenyl acetate via dynamic covalent kinetic resolution. This was a one-pot process that reached 96.5% ee through the combination of the reversible hemithioacetal transformation and the enantioselective lactonization catalyzed by the immobilized lipase from *Trichosporon laibachii* (Scheme 24). As a result, the desired stereochemistry of 1,3-oxathiolane precursors **71** and **72** was achieved.

In 2014, Zhang et al. [64] reported an optimized asymmetric synthesis of 1,3-oxathiolan-5-ones **77** and **78** via dynamic cova-



Scheme 22: Synthesis of **1** and **1a** from glycolaldehyde dimer **64** and 1,4-dithiane-2,5-diol (**3q**) using STS and CAL-B, respectively.



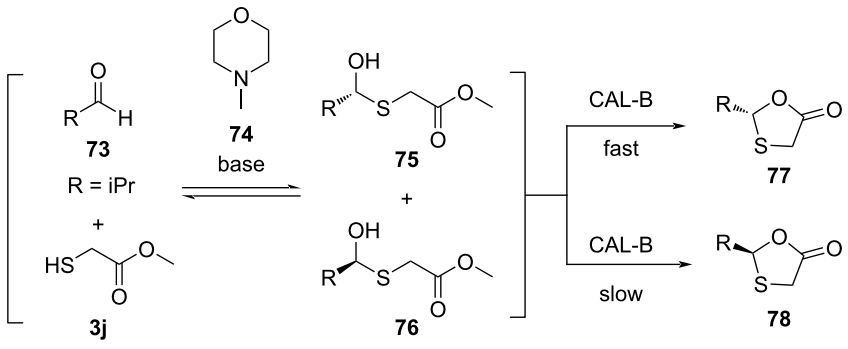
Scheme 24: Synthesis of **71** and **72** using *Trichosporon taibachii* lipase and kinetic resolution.

lent kinetic resolution using hemithioacetal chemistry coupled with a lipase-catalyzed cyclization (Scheme 25). Methyl thioglycolate (**3j**) was used in the reaction with aldehyde **73**. These acted as hemithioacetal substrate and acyl donor, respectively. CAL-B was further utilized for the subsequent intramolecular cyclization of hemithioacetal intermediates **75** and **76**. Screening of base additives showed that good results could be obtained by addition of 4-methylmorpholine (**74**). Enantioselectivity for a wide range of substrates was achieved in good yield with rigorous optimization of the reaction conditions by utilization of wild-type CAL-B.

Synthetic N-glycosylation strategies for glycosidic C–N bond formation in 1,3-oxathiolane nucleosides

This section will discuss the methods for constructing glycosidic C–N bonds in 1,3-oxathiolane nucleosides. The chemistry detailed in this section will concentrate on building N-nucleosides. There have been several excellent reviews on the construction of nucleosides over the past decades [33,34,65]. Accordingly, this section begins with an introduction on important classical approaches and older yet creative methods to provide the reader with a historical context. For comparison, this will be followed by a discussion of more modern techniques, including chiral auxiliaries for neighboring group participation and transition metal-catalyzed reactions, along with recent new promoter-dependent advances. It is generally agreed that the stereochemical outcome of glycosylation can be affected by multiple factors [66-69], which include i) structure and conformation of the glycosyl substrates, ii) glycosylation reagents or promoters, iii) the solvent, iv) presence of a participating or chiral auxiliary protecting group, v) the presence of a conformationally locked protecting group, vi) the presence of a glycosyl acceptor tethering group, and/or vii) the presence of an exogenous nucleophilic additive.

The distinction between α- and β-glycosidic bonds depends on the relative orientation of the anomeric carbon atom and the stereocenter furthest from position C-1 in the sugar. For example, when the nucleobase at C-1 is oriented *cis* to the hydroxymethyl group of the sugar at C-4, it is a β-glycosidic bond, whereas if it is orientated *trans*, it is referred to as α-glycosidic bond [70]. The exact character of the glycosidic bond in the structure defines the physicochemical properties and biological role of the molecule [71]. There have been numerous efforts to synthesize these nucleoside analogues in order to achieve the desired stereoselectivity during β-selective glycosidic bond formation. The general pathway for glycosidic bond formation (Figure 4) shows that the glycoside donor moiety has to be activated using an appropriate activator to form an oxonium ion.



Scheme 25: Synthesis of 1,3-oxathiolan-5-ones **77** and **78** via dynamic covalent kinetic resolution.

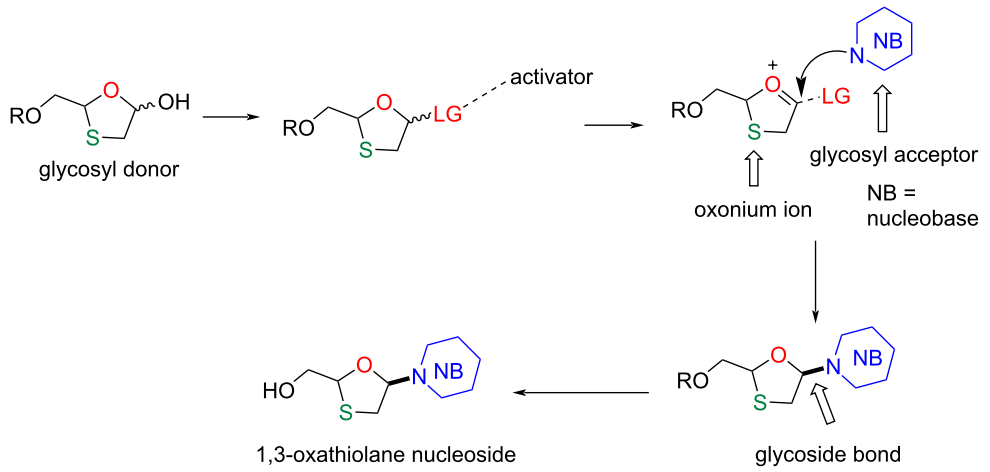


Figure 4: Pathway for glycosidic bond formation.

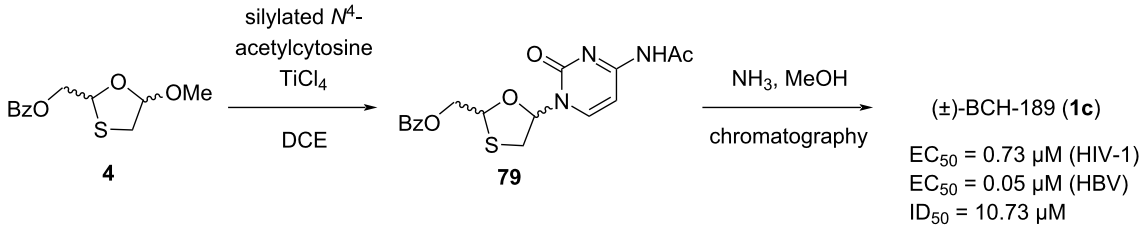
The attack of a nucleobase (glycosyl acceptor) may occur on either side of the oxonium ion, which can result in two anomers, i.e., an α- and a β-anomer. The factors affecting such stereocontrolled glycoside bond formations are also discussed in this review.

The preparation of the racemate **1c** was reported by Belleau et al. in 1989 (Scheme 26) [38]. The method involved the coupling of oxathiolane derivative **4** with silylated cytosine, which afforded **79** as a mixture of the *cis*- and *trans*-nucleosides. The process used N-protected cytosine and further chromatographic separation. The deprotection with a methanolic ammonia solution provided racemic (±)-BCH-189 (**1c**). In vitro studies of (±)-BCH-189 (**1c**) showed potent anti-HIV-1 activity. The EC_{50} value of (±)-BCH-189 (**1c**) was reported to be in the range of 0.37–1.31 μM (mean 0.73 μM), and the compound was effective against HIV-1 in MT-4 cells [13].

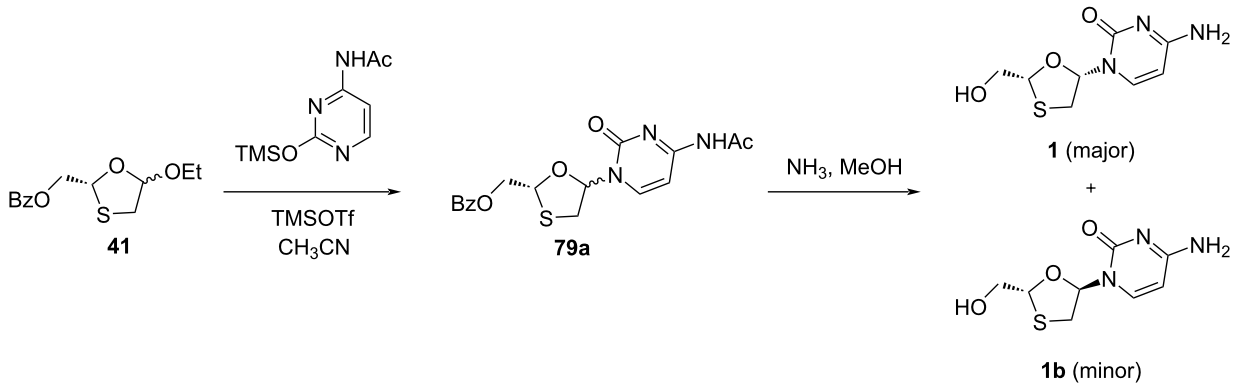
Enantioselective enzymatic synthesis of 3TC (**1**) was also reported by Milton et al. [47], who isolated oxathiolane precursor **41**, as discussed earlier, by enzymatic resolution of an acetoxy sulfide by a *Pseudomonas fluorescens* lipase. Using this pure precursor **41**, the synthesis of 3TC (**1**) was accomplished by N-glycosylation with silylated base using trimethylsilyl trifluoromethanesulfonate (TMSOTf) catalyst to obtain nucleoside derivative **79a**, followed by deprotection using methanolic ammonia (Scheme 27).

Cousins et al. [49] carried out the coupling of enantiomerically enriched oxathiolane propionate **44** with silylated cytosine in the presence of the Lewis acid trimethylsilyl iodide (TMSI), which gave a *cis*/*trans* ratio of 1.3:1 for the nucleoside intermediate **79a**. Further, the nucleoside intermediate **79a** was deprotected using a type of basic resin. This gave the *cis*-diastereomer 3TC (**1**), which was purified by chiral HPLC, resulting in an ee value of 70% (Scheme 28).

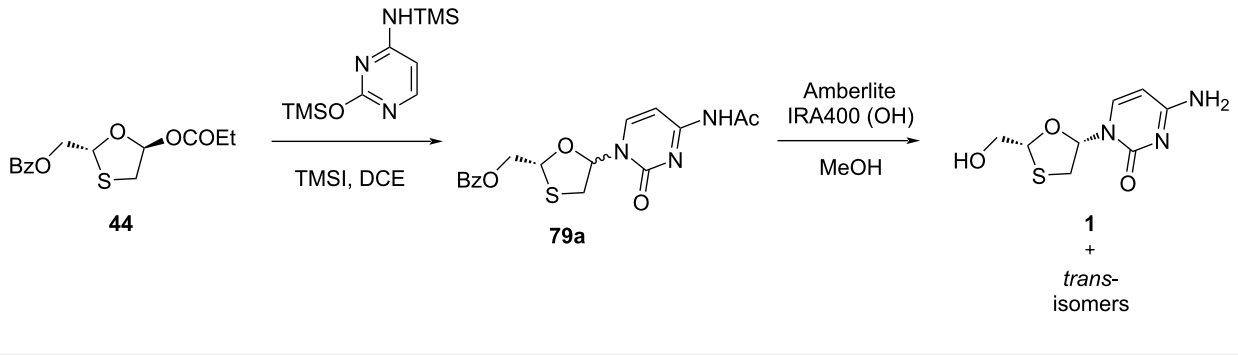
Further developments in the effective enantiopure synthesis of lamivudine (**1**) were achieved by many scientists. The synthesis of 1,3-oxathiolane nucleosides utilizing stereoselective coupling of a nucleobase with the oxathiolane sugar intermediate via in situ chelation was reported by Liotta and co-workers (Scheme 29) [72]. Appropriate Lewis acids form a complex with the oxathiolane intermediates via in situ chelation. The ex-



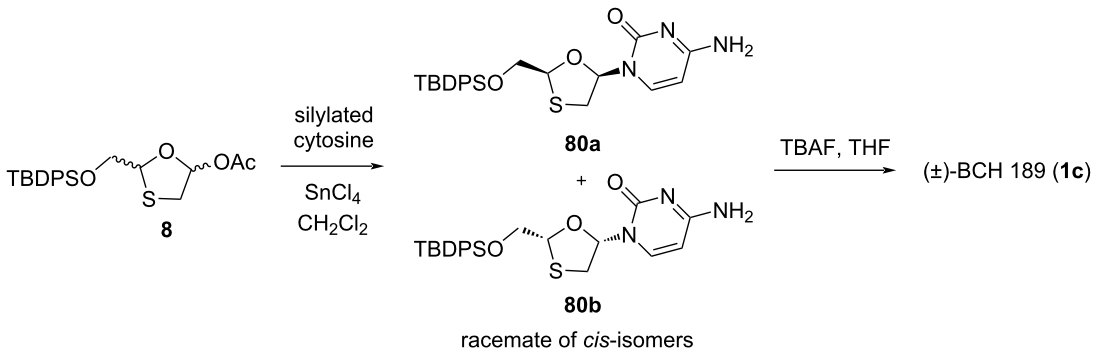
Scheme 26: First synthesis of (±)-BCH-189 (**1c**) by Belleau et al.



Scheme 27: Enantioselective synthesis of 3TC (**1**).



Scheme 28: Synthesis of *cis*-diastereomer 3TC (**1**) from oxathiolane propionate **44**.



Scheme 29: Synthesis of (±)-BCH-189 (**1c**) via $SnCl_4$-mediated N-glycosylation of **8**.

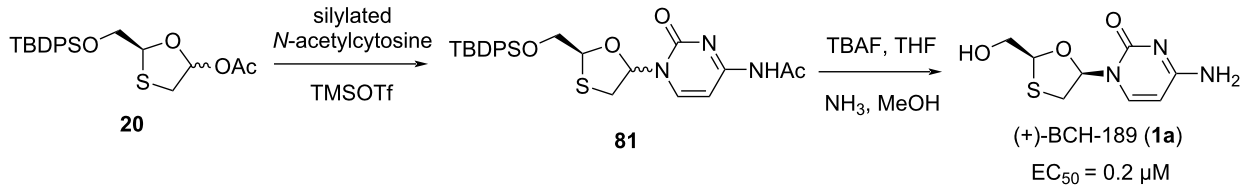
clusive formation of the β-anomer was observed upon coupling of the anomer mixture **8** with silylated cytosine using stannic chloride (about 2 equiv in CH_2Cl_2 at room temperature). This stereoselective outcome could have been due to an in situ chelation process. The level of selectivity was determined by HPLC to be >300:1 in favor of the β-configured *cis*-isomers (racemic mixture of **80a** and **80b**) [30]. Further, the desilylation using tetrabutylammonium fluoride (TBAF) gave racemic (±)-BCH-189 (**1c**).

Chu et al. [40] described coupling of crude 1,3-oxathiolane precursor **20** with silylated acetylcytosine utilizing TMSOTf as a

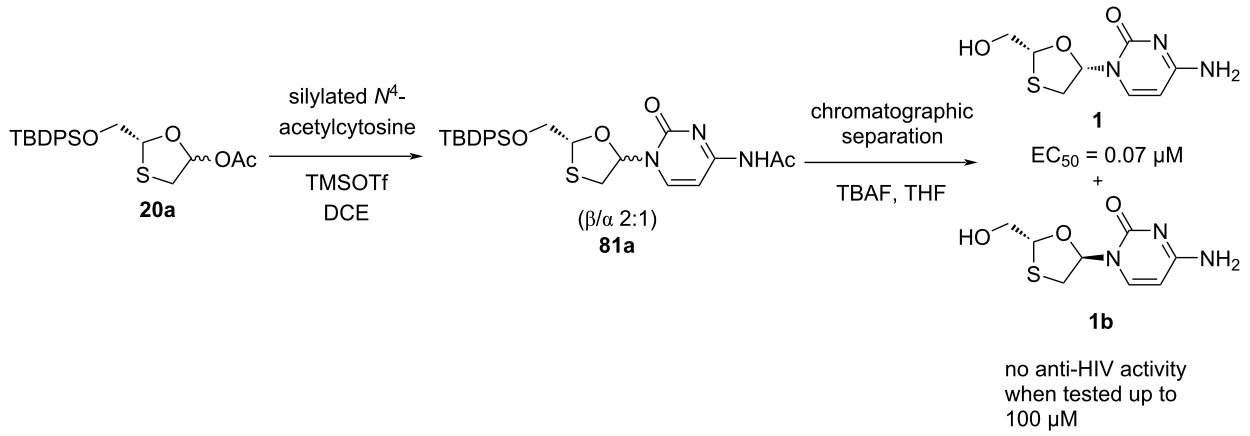
Lewis acid, which gave a mixture of α- and β-anomers (1:2 ratio) of **81** (Scheme 30). The mixture of anomers was further separated by silica gel column chromatography. (+)-BCH-189 (**1a**) and the α-anomer were produced individually by further deacetylation using methanolic ammonia and desilylation with TBAF [33]. (+)-BCH-189 (**1a**) was found to be less active against HIV-1 (EC_{50} = 0.2 μM in CEM cells) than (−)-BCH-189 (**1**, EC_{50} = 0.07 μM in CEM cells) [14].

Chu and co-workers [44] reported a synthetic procedure to access (−)-BCH-189 (**1**). Compound **20a** was synthesized from L-gulose derivative **3f** (Scheme 6). The glycosylation reaction of oxathiolane intermediate **20a** with silylated N^4-acetylcytosine in dichloroethane using TMSOTf as a catalyst gave **81a** as a β/α 2:1 mixture (Scheme 31). Separation by chromatography and deprotection with TBAF in THF afforded the (−)-isomer 3TC (**1**, EC50 = 0.07 μM in CEM cells) and the *trans*-isomer **1b**. The *trans*-substituted (+)-isomer **1b** did not shown any activity when it was tested (up to 100 μM). Further investigation showed that using stannic chloride instead of TMSOTf for the N-glycosylation procedure afforded a racemic mixture. This could be due to the opening as well as closing of the oxathiolane ring under the reaction conditions.

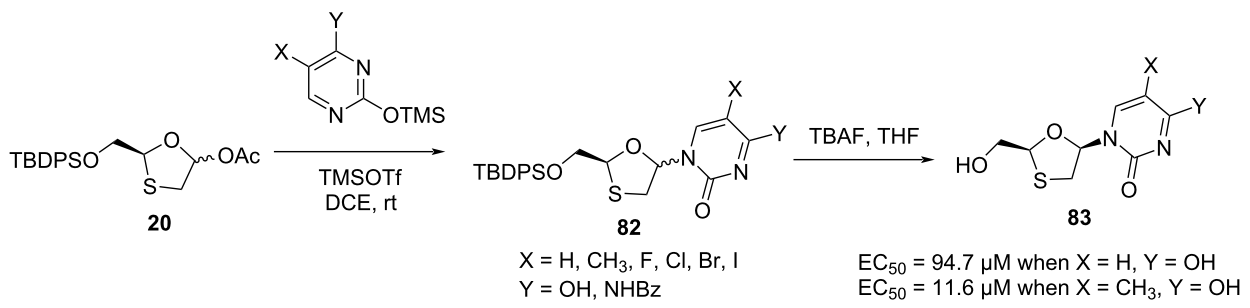
Optically pure β-D- and α-D-configured 1,3-oxathiolane pyrimidine and 1,3-oxathiolane purine nucleosides with natural nucleoside configuration were synthesized by Jeong et al. (Scheme 32 and Scheme 33) [41,42]. The purpose of this was



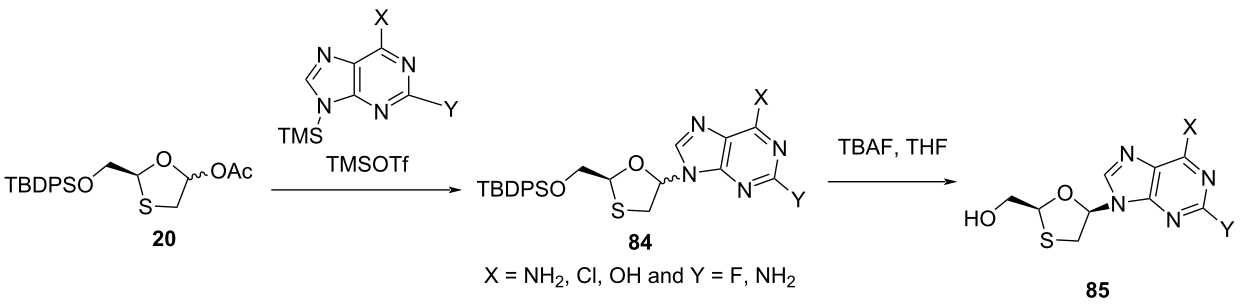
Scheme 30: Synthesis of (+)-BCH-189 (**1a**) via TMSOTf-mediated N-glycosylation of **20**.



Scheme 31: Synthesis of 3TC (**1**) from oxathiolane precursor **20a**.



Scheme 32: Synthesis of **83** via N-glycosylation of **20** with pyrimidine bases.



Scheme 33: Synthesis of **85** via N-glycosylation of **20** with purine bases.

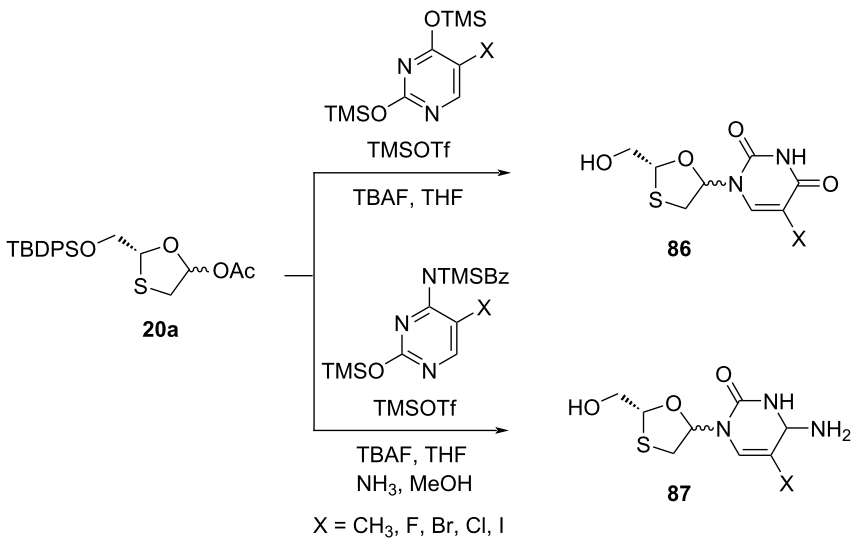
the investigation of the structure–activity relationships as anti-HIV-1 agents. The oxathiolane intermediate **20**, produced from D-mannitol, was further condensed with a range of pyrimidine and purine nucleobases via N-glycosylation. The anti-HIV activity of the nucleosides **83** was quantified by EC_{50} values of 94.7 μM and 11.6 μM when X = H or CH_3 and Y = OH, respectively [33]. The α-anomers were also isolated by chromatographic separation methods.

To study the structure–activity relationships of various nucleobase derivatives, oxathiolane acetate **20a** was further condensed with various pyrimidines (Scheme 34) and purines (Scheme 35), as reported by Jeong et al. [73]. When X = F, the cytosine derivative **87**, among all of these nucleosides having anti-HIV activity, was found to be the most potent. The pyrimidine analogues **90** and **91** were also found to be active against HIV-1, with EC_{50} = 0.28 μM and 2.8 μM, respectively [33].

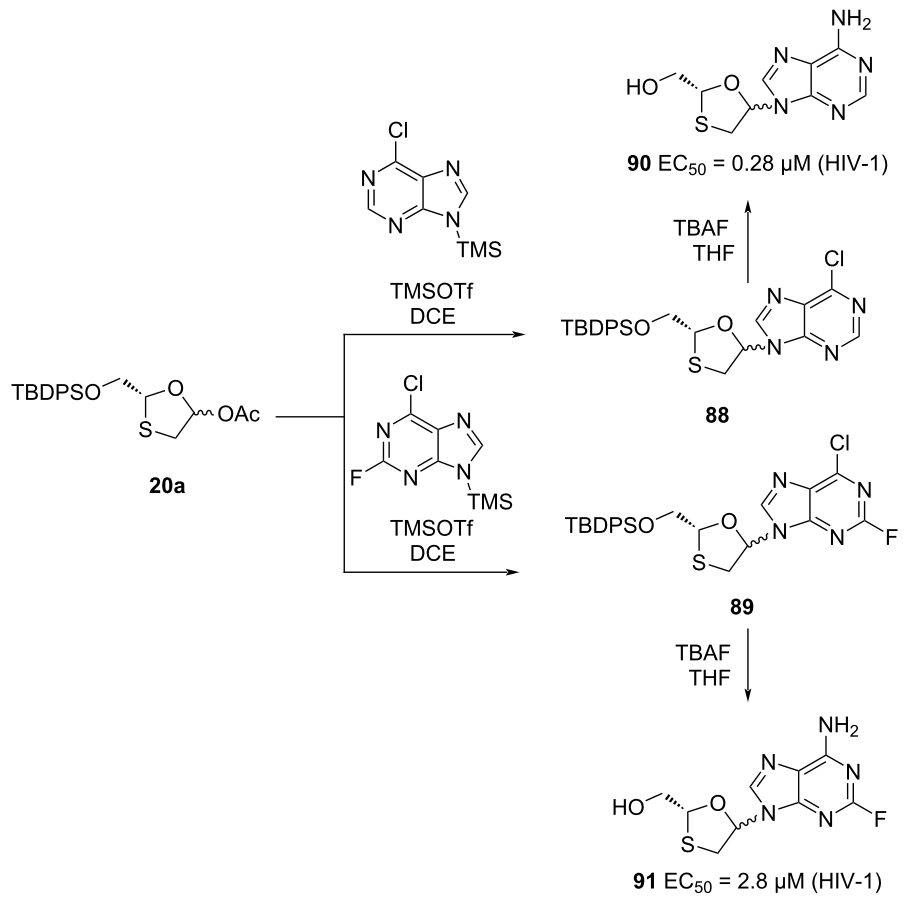
In 1992, Humber et al. [45] established a method for glycosylation of benzoylated oxathiolane **31** with silylated cytosine in the presence of trimethylsilyl iodide (TMSI) as a catalyst, which afforded nucleoside **92** as a β/α 1.3:1 mixture. Furthermore, anomeric mixture separation and deprotection using Amberlite IRA 400(OH) afforded 3TC (**1**) and the α-anomer **1b** (Scheme 36).

In 1995, a novel route was developed by Jin and co-workers [46], which utilized a Vorbrüggen N-glycosylation of the enantiomerically pure 5-acetoxyoxathiolane **35a** with presilylated cytosine as the key convergent step. This N-glycosylation reaction required the Lewis acid TMSI in a significant quantity to produce the desired cytidine **1** (Scheme 37).

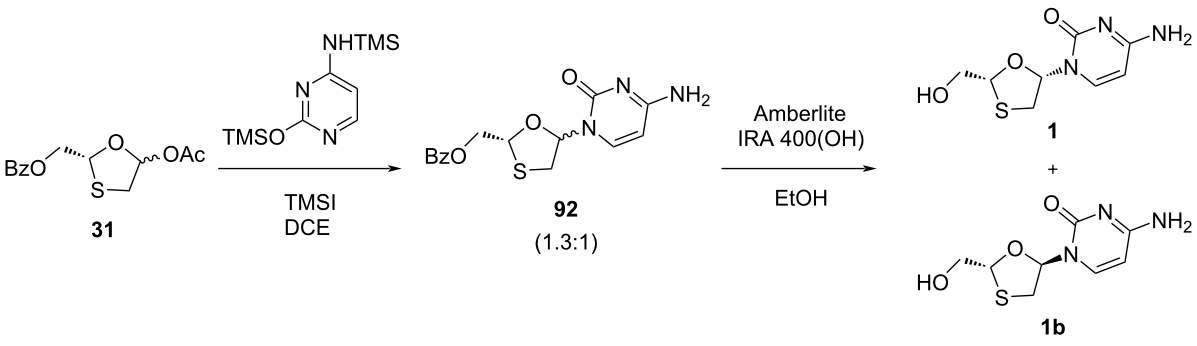
As shown in a plausible mechanism in Scheme 37, it is assumed that 5-α-iodooxathiolane **III** is formed stereoselectively by reac-



Scheme 34: Synthesis of **86** and **87** via N-glycosylation using TMSOTf and pyrimidines.



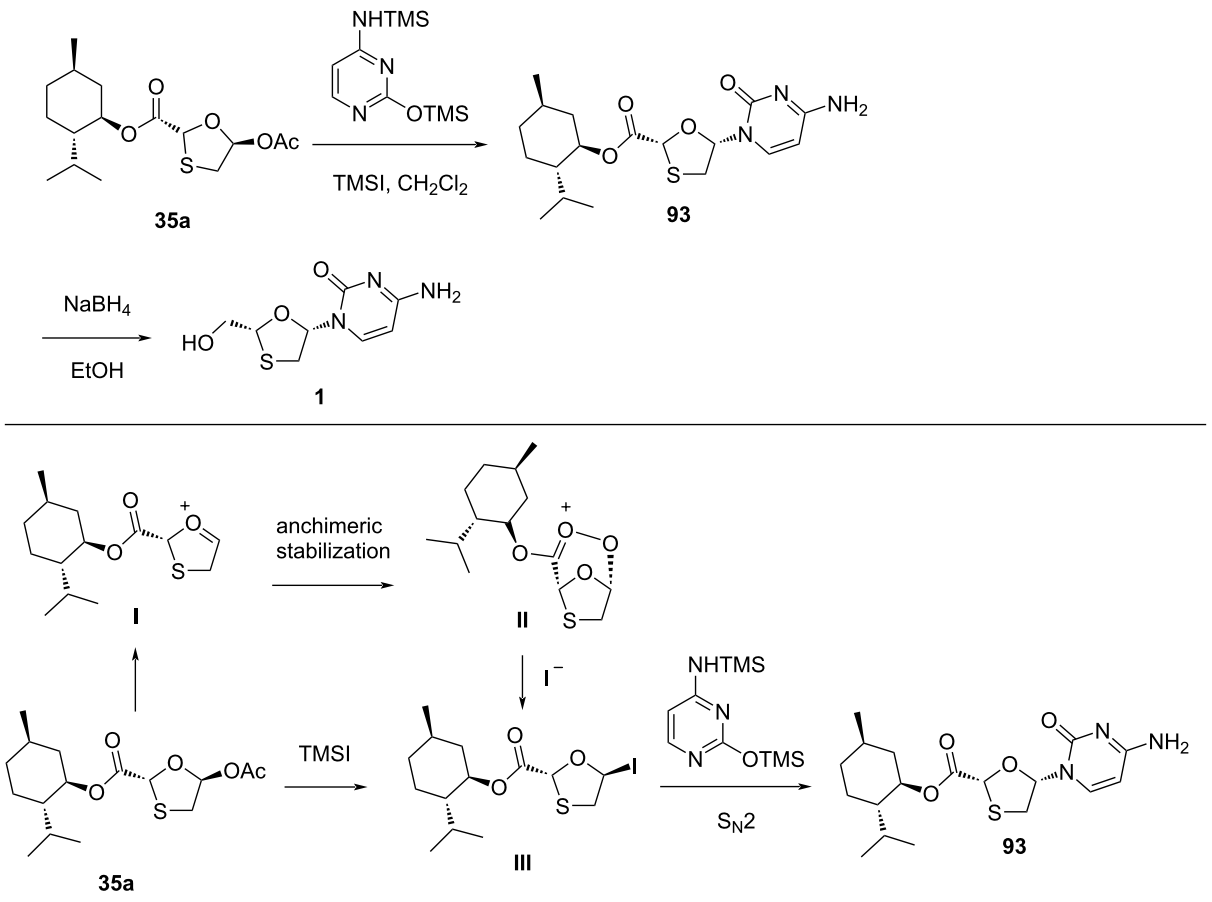
Scheme 35: Synthesis of **90** and **91** via N-glycosylation using TMSOTf and purines.



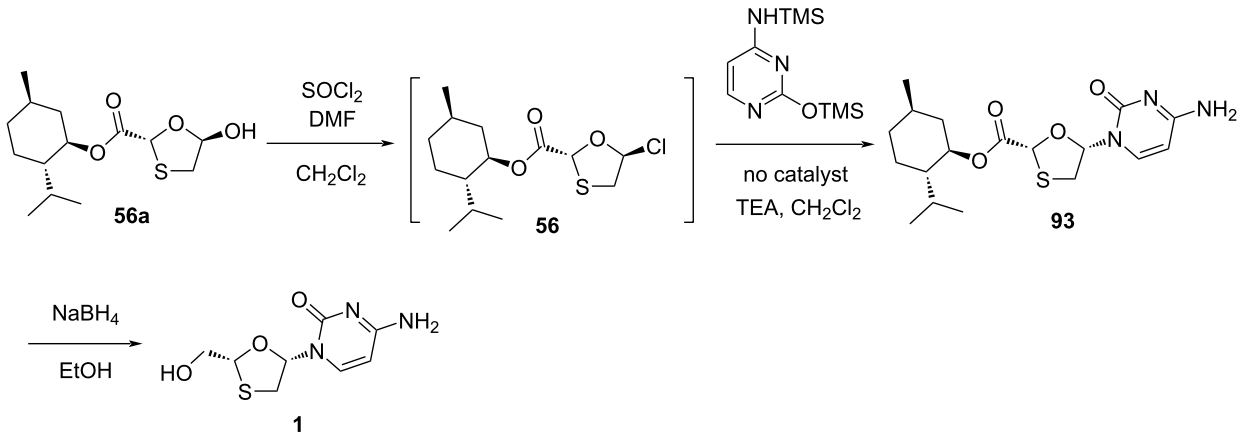
Scheme 36: Synthesis of 3TC (**1**) via TMSI-mediated N-glycosylation.

tion of the oxonium ion **I**, generated in situ by reaction of 5-acetoxyoxathiolane **35a** with TMSI, which stabilizes a C-2 ester substituent via anchimeric assistance (see **II**). The postulated mechanism shows that the iodide ion attacks on the stable oxonium ion to provide an anomer, which further reacts with the presilylated nucleobase in an S_N2 manner and predominantly affords the β-cytidine adduct.

DKR overcomes the drawback of classical resolution since it is theoretically possible to obtain 100% yield of the desired isomer [74]. 5-Hydroxyoxathiolane intermediate **56a** was isolated in a DKR procedure by Whitehead and co-workers (Scheme 38) [55]. Further, 5-chlorooxathiolane **56** was isolated from chlorination reaction of 5-hydroxyoxathiolane **56a** using thionyl chloride in presence of catalytic DMF and dichloro-



Scheme 37: Stereoselective N-glycosylation for the synthesis of **1** by anchimeric assistance of a chiral auxiliary.



Scheme 38: Whitehead and co-workers' approach for the synthesis of **1** via direct N-glycosylation without an activator.

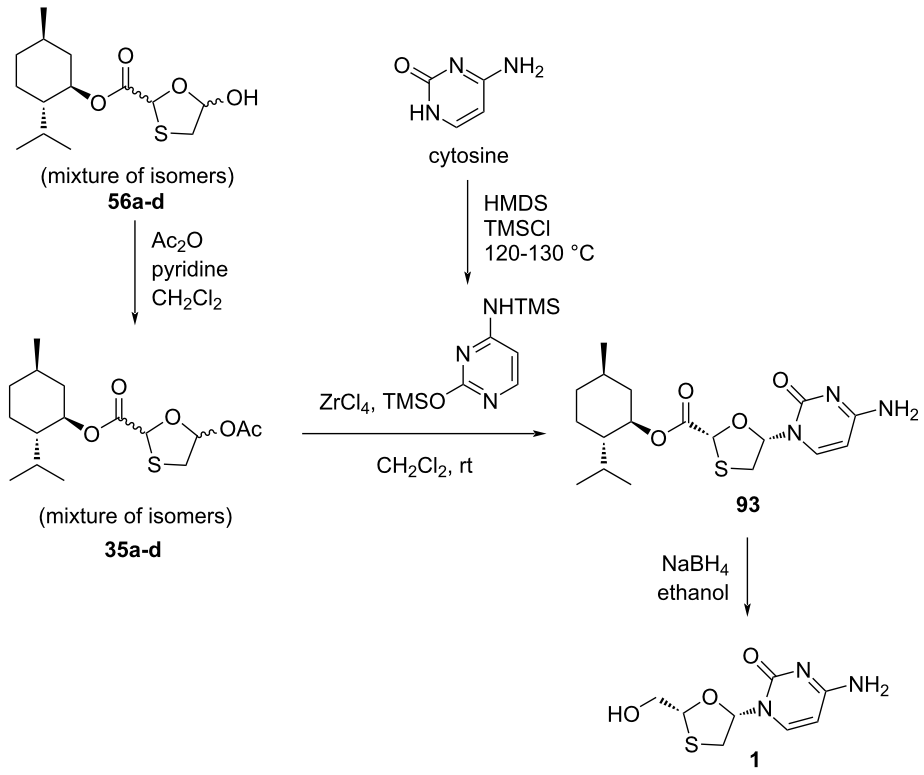
methane solvent. This further reacted directly with the presilylated cytosine without any promoter or additive and gave nucleoside **93** in a selective manner (β/α 10:1). The ester group of nucleoside derivative **93** was further reduced with sodium borohydride in ethanol, which gave lamivudine (**1**). An efficient and enantioselective synthesis of lamivudine (**1**) was de-

veloped, which utilizes a highly effective DKR as the key step for obtaining pure substrate. The synthesis of **56a** via DKR was discussed earlier in this review (Scheme 17).

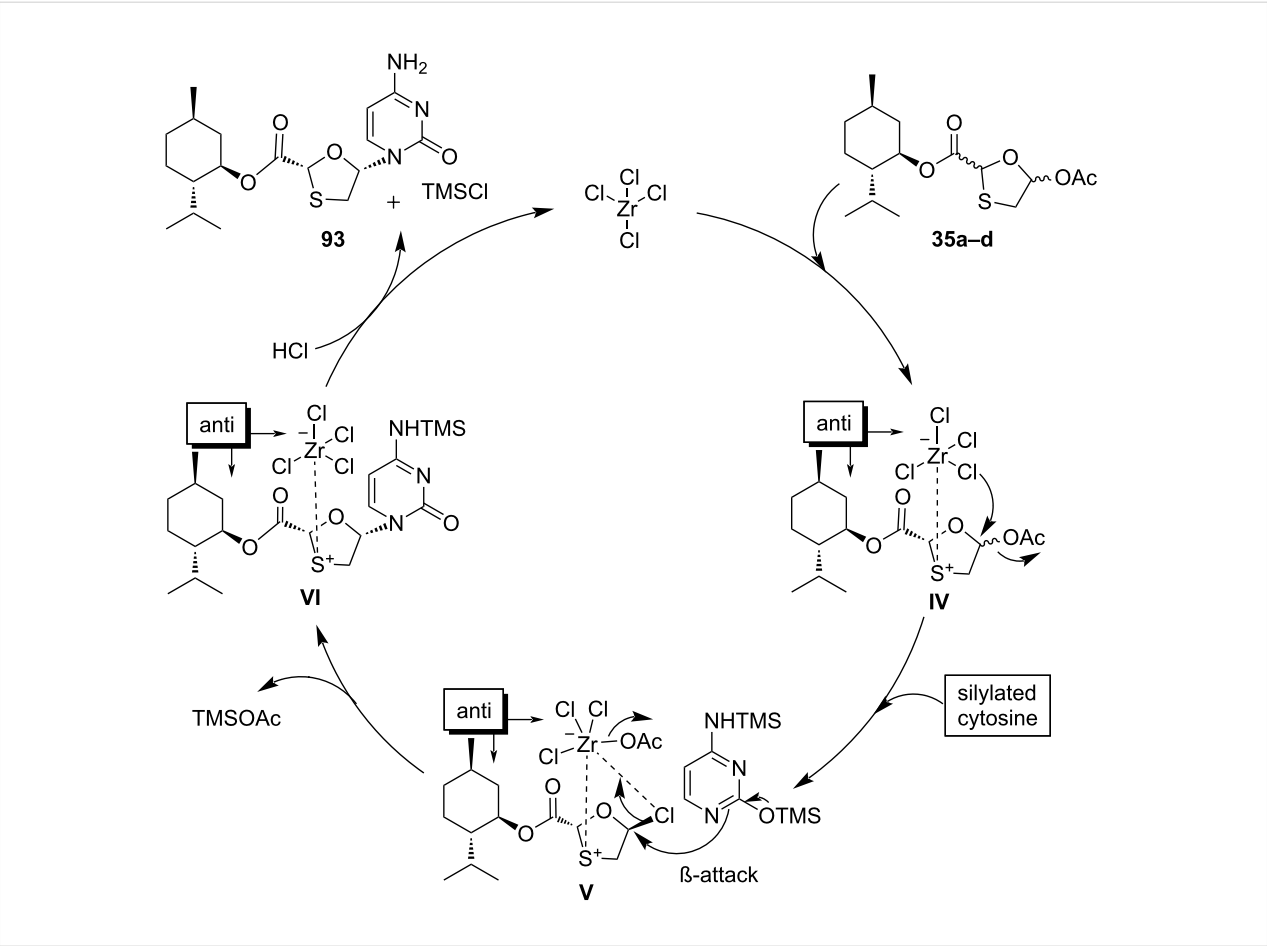
Recently, we have developed [75] an effective method for selective glycosylation using 0.5 equiv of $ZrCl_4$ via the activation of oxathiolane acetates **35a**–**d**. The reaction was complete after a reduced reaction time and suitable for large-scale production with good yield at ambient temperature (Scheme 39). The usefulness of this method was that even without isolation of enantiomerically pure oxathiolane substrate, the facile stereoselective glycosylation took place and was improved compared to previously reported methods. The oxathiolane acetates **35a**–**d** were used in situ and stereoselectively led to a single nucleoside isomer **93**. After preparation of nucleoside ester intermediate **93**, lamivudine (**1**) was obtained with reducing agent sodium borohydride.

The plausible reaction mechanism was also described for this selective N-glycosylation methodology (Scheme 40). A previously reported [72] plausible mechanism involving the use of $SnCl_4$ was considered while proposing the mechanism when using $ZrCl_4$ catalyst for the stereoselective N-glycosylation. We hypothesize that because of the Lewis acid character, $ZrCl_4$ could most likely form a precomplex with the sulfur atom of the oxathiolane ring, as in **IV**. The presence of the chiral L-menthyl ester auxiliary function assists the complexation with $ZrCl_4$ in a specific orientation and could minimize the destabilization through 1,2-steric interactions. Therefore, the selectivity could herein be accomplished by means of anchimeric assistance by the L-menthyl ester. Additionally, formation of intermediate **V** probably occurred due to the attack of one chloride ion on the anomeric carbon atom while maintaining α-configuration and simultaneous elimination of an acyl group as illustrated in intermediate **IV**. Further, attack of silylated cytosine on α-chloro-substituted derivative **V** in an S_N2 reaction results in the formation of a glycosidic C–N bond in the β-configured nucleoside intermediate **VI**. In the last step, addition of HCl easily deprotects the TMS group of intermediate **VI** and affords compound **93** through simultaneous removal of $ZrCl_4$. This approach of in situ precomplexing disallows the α-face attack of silylated cytosine.

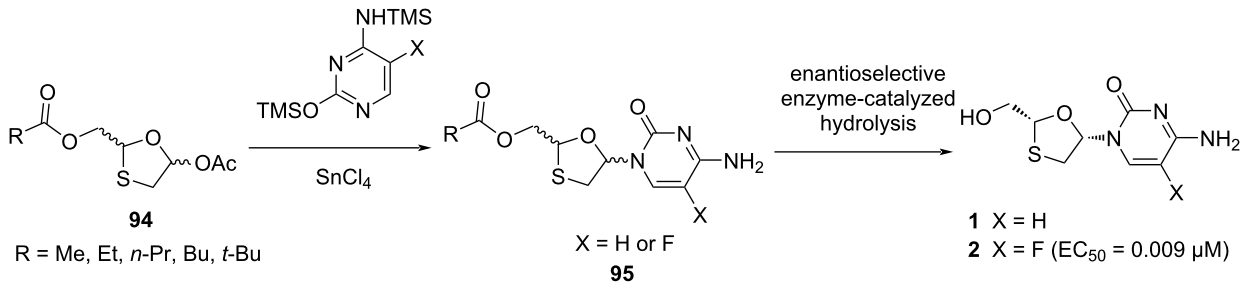
Liotta and co-workers [76] established an enzyme-catalyzed hydrolysis of protected racemic nucleosides to synthesize the enantiomerically pure oxathiolane nucleoside analogues **1** and **2** (Scheme 41). The protected racemic nucleoside derivatives **95** were synthesized by tin-mediated N-glycosylation of the corresponding acetate precursor **94** with silylated cytosine or 5-fluorocytosine. Further, hydrolysis of the



Scheme 39: $ZrCl_4$-mediated stereoselective N-glycosylation.



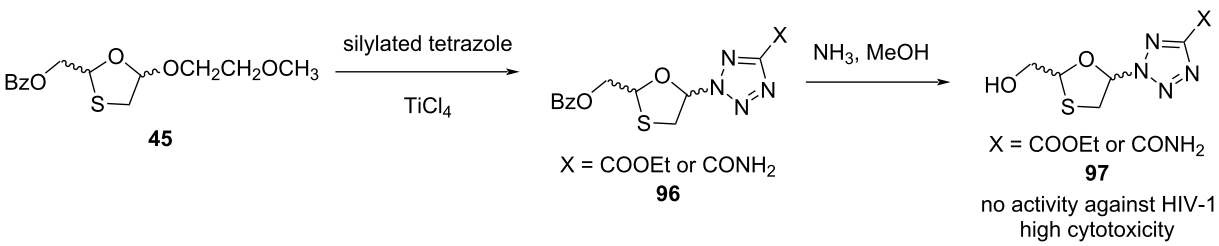
Scheme 40: Plausible reaction mechanism for stereoselective N-glycosylation using $ZrCl_4$.



Scheme 41: Synthesis of enantiomerically pure oxathiolane nucleosides **1** and **2**.

5'-*O*-acetyl group was evaluated with respect to reactivity and enantioselectivity utilizing several enzymes. They found that the butyrate ester derivative was hydrolyzed with a higher rate than the 5'-*O*-acetate derivative during the synthesis of L-(−)-2',3'-dideoxy-5-fluoro-3'-thiacytidine (**2**). However, hydrolysis was observed to occur with a comparable rate to that of the 5'-*O*-valerate and 5'-*O*-propionate esters. Additionally, the rate of hydrolysis for the ester derivatives of FTC (**2**) was significantly higher than for the corresponding 3'-thiacytidine derivatives.

The tetrazole analogues of 1,3-oxathiolane nucleosides were synthesized by Faury et al. [50]. N-Glycosylation of silylated tetrazole with 1,3-oxathiolane precursor **45** in the presence of titanium tetrachloride or TMSOTf, followed by deprotection in methanolic ammonia gave the final nucleoside **97** (Scheme 42).

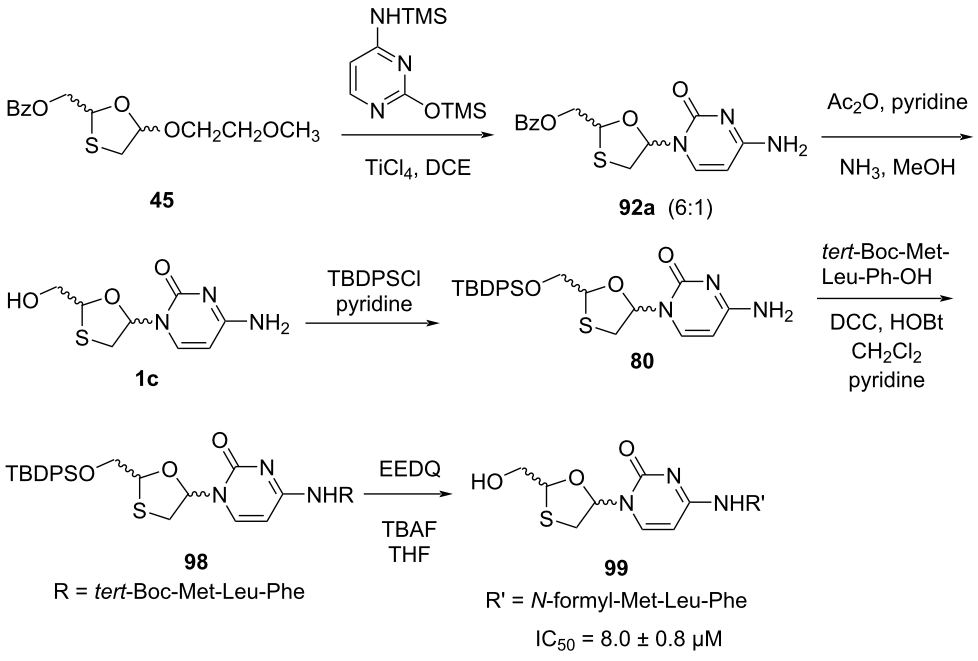


Scheme 42: Synthesis of tetrazole analogues of 1,3-oxathiolane nucleosides **97**.

Unfortunately, the introduction of a tetrazole ring to the oxathiolane moiety did not result in any anti-HIV activity and higher cytotoxicity.

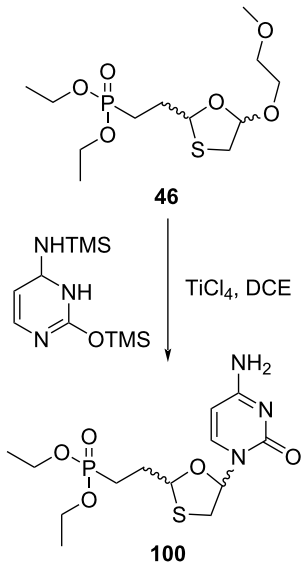
The synthesis of N^4-substituted analogue **99** of 2',3'-dideoxy-3'-thiacytosine was discovered by Camplo et al. (Scheme 43) [77]. The prodrug was devised for targeting specific receptors on the leukocytes membrane. The crucial N-glycosylation reaction between 1,3-oxathiolane precursor **45** and silylated cytosine was carried out using $TiCl_4$ as a catalyst. The N-acylation of compound **92a** was performed for flash chromatography, and further ammonolysis in methanol affords compound **1c**. The silylation of **1c** with TBDPSCl was carried out, and then coupling reaction with *tert*-Boc-Met-Leu-Phe-OH in the presence of DCC and HOBt provided compound **98**. The *tert*-Boc protecting group was further removed in formic acid, and the resulting nucleoside peptide was formylated using 2-ethoxy-1-ethoxycarbonyl-1,2-dihydroquinoline (EEDQ) as formylating reagent. Finally, the *tert*-butyldiphenylsilyl (TBDPS) group was desilylated using TBAF in THF solvent, which gave compound **99**. For compound **99**, the IC_{50} value of HIV-I cytopathogenicity in MT-4 cells was 8.0 μM at a concentration nontoxic to the host cells.

In 1993, Kraus [51] developed the phosphonate analogue **100** of 3'-thia-2',3'-dideoxycytidine. The Lewis acid-mediated N-glycosylation reaction of the phosphonate analogue **46** of an oxathiolane precursor with an appropriate nucleobase afforded the phosphonate analogue **100** (Scheme 44). To obtain both the α- and β-anomers for biological assessment, $TiCl_4$ was used as a Lewis acid in the glycosylation procedure in place of $SnCl_4$. Separation of the α- and β-anomers was carried out after N^4-acetylation by using acetic anhydride in DMF. The pure isomers were isolated in 80% yield in a 1:1 ratio. The phos-



Scheme 43: Synthetic approach toward **99** from 1,3-oxathiolane **45** by Camplo et al.

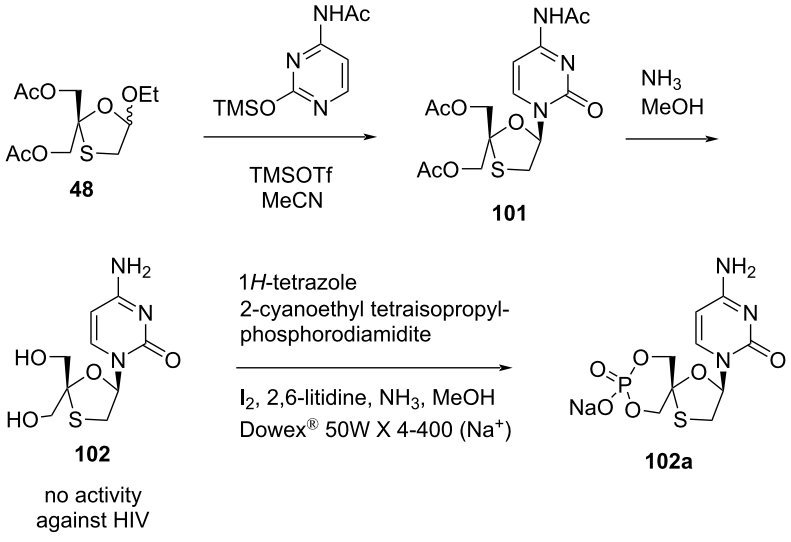
phonate nucleosides were isolated by hydrolysis of the phosphonic acid ethyl ester, followed by treatment with methanolic ammonia. Anti-HIV assessment of these analogues demonstrated that the α-anomer was not active, while the β-anomer was less potent than the parent compound (±)-BCH-189 (**1c**). This could be because the phosphorylated modified analogue **100** was not a proper substrate for nucleotide kinases.



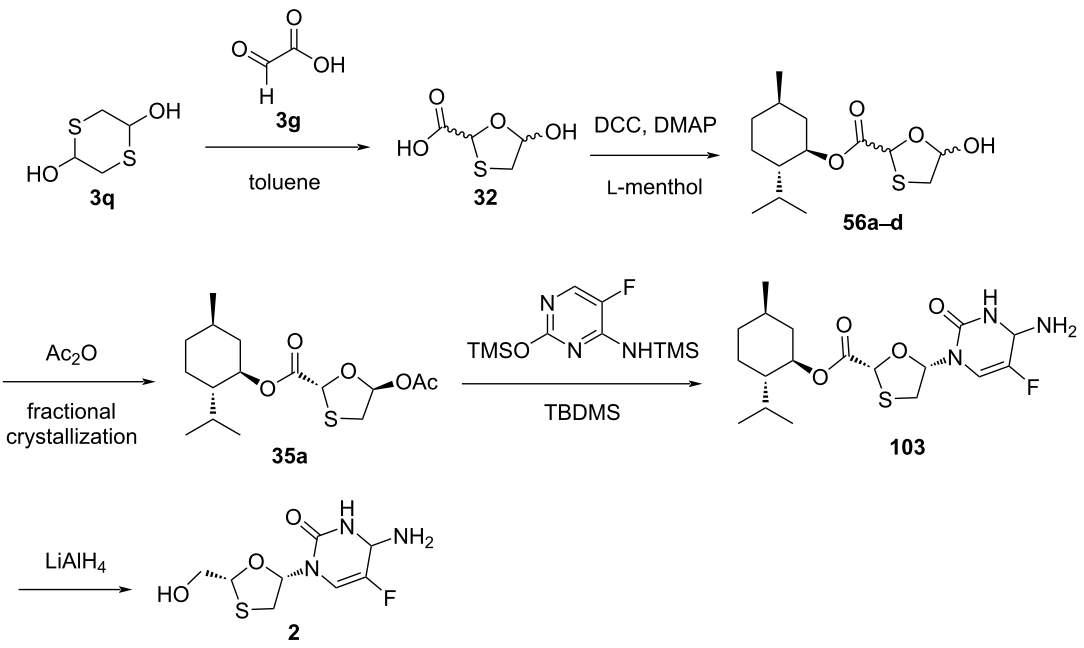
Scheme 44: Synthesis of **100** from oxathiolane phosphonate analogue **46**.

1,3-Oxathiolane derivative **48** was glycosylated directly with persilylated *N*-acetylcytosine to provide nucleoside **101**, which gave nucleoside **102** after deprotection (Scheme 45). Thymine was also used instead of *N*-acetylcytosine, which gave the corresponding thymine-based nucleoside derivative. These were also converted to the spirocyclic monophosphate nucleosides **102a**, but none of the synthesized compounds showed anti-HIV activity. This study was performed by Chao and Nair in 1997 [52]. The procedure synthesized a racemic 4'-hydroxymethylated 2',3'-dideoxy-3'-thianucleoside analogue starting from compound **48** via N-glycosylation with silylated nucleobase in the presence of Lewis acid in acetonitrile solvent. Further deacetylation was carried out in methanolic ammonia to afford nucleoside **102**. Cyclic thianucleoside monophosphate **102a** was synthesized when nucleoside **102** was treated with 2-cyanoethyl tetraisopropylphosphorodiamidite in the presence of 1*H*-tetrazole in DMF, followed by oxidation with iodine and deprotection with methanolic ammonia.

Mansour et al. [78] described a highly diastereoselective processes for producing *cis*-nucleoside analogues and derivatives in high optical purity. The oxathiolane derivative **32** was synthesized by the reaction of 1,4-dithiane-2,5-diol (**3q**) with glyoxylic acid (**3g**). Further, L-menthol as a chiral auxiliary was introduced using DCC and DMAP, which gave *cis*- and *trans*-esters **56a**–**d** as a diastereomeric mixture (Scheme 46). The glycosylation reaction of **35a** with presilylated 5-fluorocytosine, followed by ester group reduction of **103** using $LiAlH_4$, provided emtricitabine (**2**). The procedure illustrates the advantages of generating nucleosides of which the configuration can easily be controlled by the selection of the appropriate starting material.

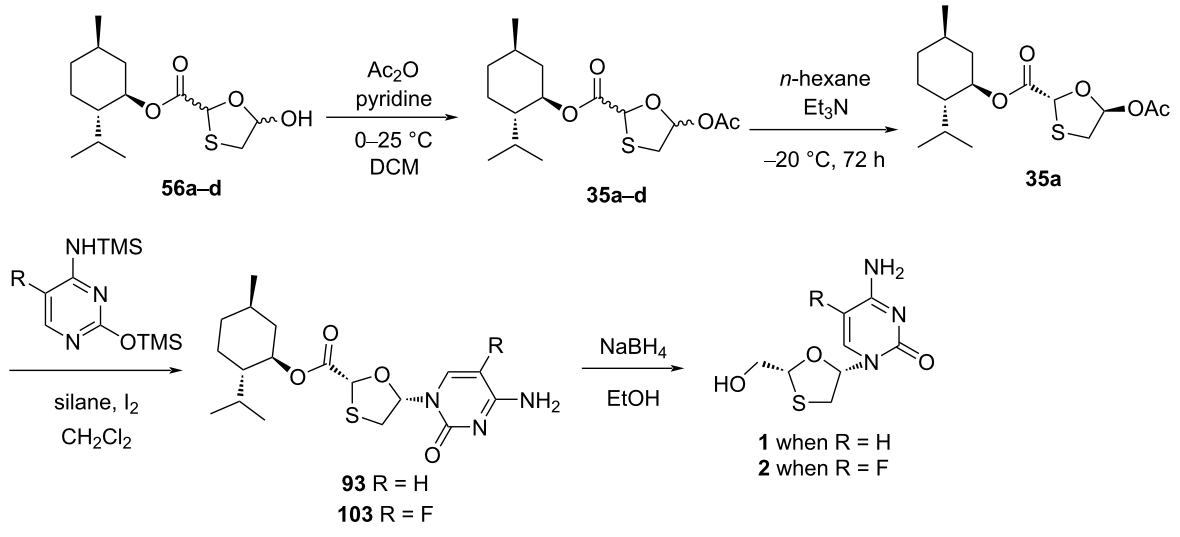


Scheme 45: Synthetic approach toward **102** and the corresponding cyclic thianucleoside monophosphate **102a** by Chao and Nair.



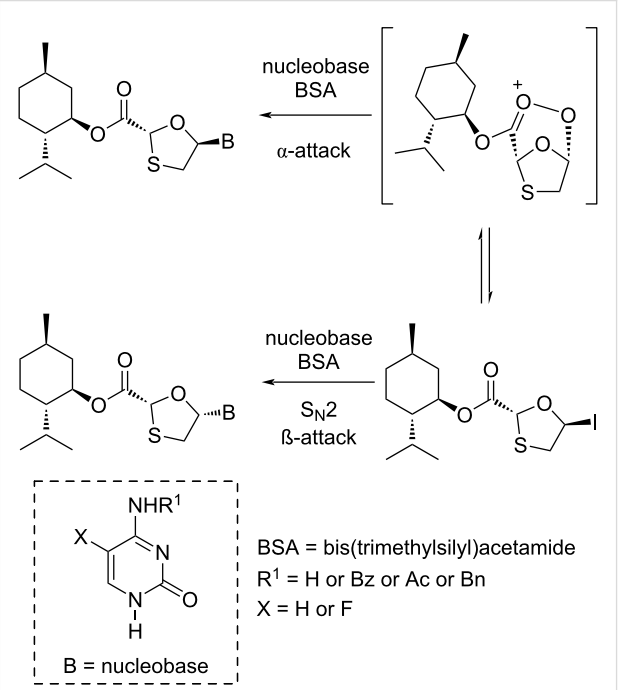
Scheme 46: Synthesis of emtricitabine (**2**) from 1,4-dithiane-2,5-diol (**3q**) and glyoxylic acid (**3g**).

The silanes Et_3SiH and PMHS, respectively, were used along with I_2 as novel N-glycosylation reagents for the synthesis of 3TC (**1**) and FTC (**2**), as reported by Caso et al. [79]. These systems were developed to promote the substrate N-glycosylation. The enantiopure 1,3-oxathiolane acetate **35a** was isolated from *n*-hexane and TEA at −20 °C after stirring for about 72 h. This intermediate was further reacted with a silylated cytosine derivative via N-glycosylation, which afforded the nucleoside analogues **93** and **103**, respectively. Stereoselectivity was achieved in the reactions, and the stereochemical outcome of the reaction was influenced by the nature of the protecting group at position N^4 of 5-fluorocytosine (Scheme 47). This method was reasonably considered as an effective alternative to the available procedures because of the use of inexpensive and more stable reagents. An important role in determining the stereochemical outcome was played by the N^4-protecting group of 5-fluorocytosine, presumably based on the capacity to increase the soft character of the nucleobase. A possible mechanism was also provided (Scheme 48), in which the chiral auxiliary L-menthol assists the selective β-nucleoside formation. The



Scheme 47: Synthesis of **1** and **2**, respectively, from **56a–d** using iodine-mediated N-glycosylation.

ester functionality of the nucleoside derivatives **93** and **103** was easily converted to a primary hydroxy group upon reduction with sodium borohydride in ethanol, which gave **1** and **2**, respectively.

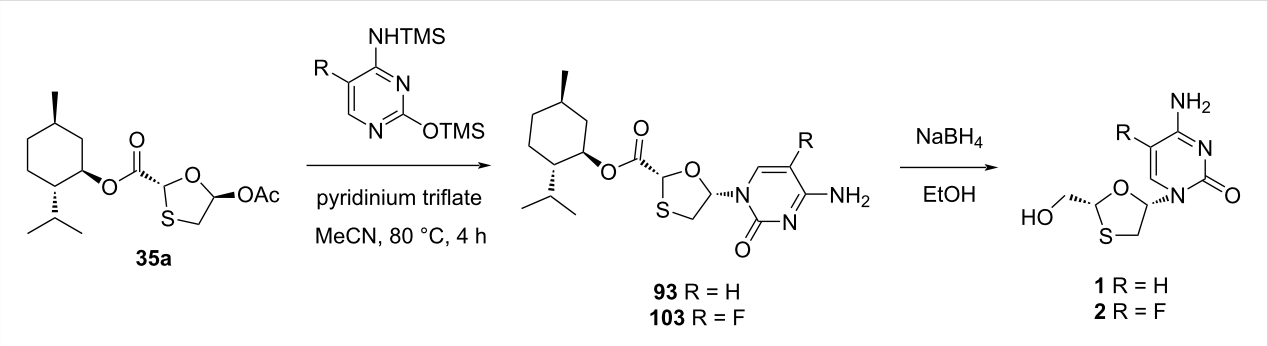


Scheme 48: Plausible mechanism for silane- and I_2-mediated N-glycosylation.

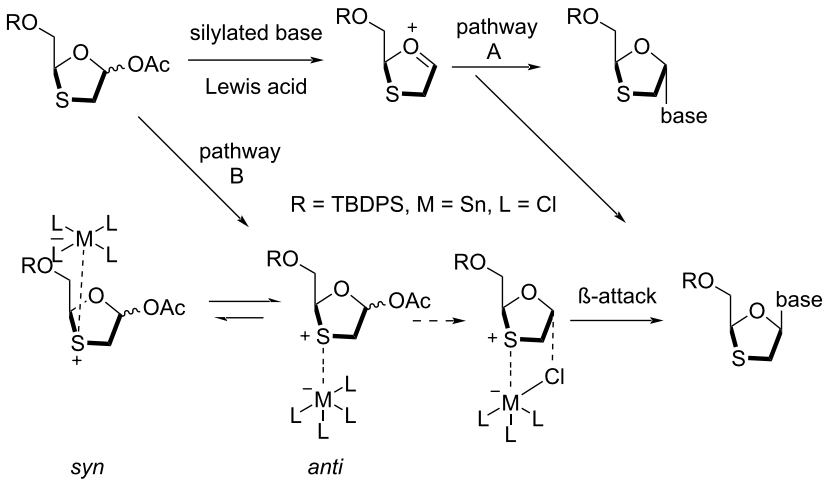
Mandala and Watts [80] reported the first use of pyridinium triflate as a novel N-glycosylation reagent for the synthesis of the antiviral drugs lamivudine (**1**) and emtricitabine (**2**, Scheme 49). The key 5-acetoxyoxathiolane intermediate **35a** was prepared in high yield by a catalyst- and solvent-free method within a minimum reaction time. Further, a greener procedure by using sodium bicarbonate as the base for the acetylation reaction instead of pyridine was implemented to prepare **35a**.

In the 1990s, Liotta, Choi, and co-authors [18,72] reported a highly stereoselective N-glycosylation reaction that was controlled via in situ chelation of the oxathiolane moiety and an appropriate Lewis acid (Scheme 50). The exclusive formation of the β-anomer of a precursor of lamivudine (**1**) was achieved by the use of stannic(IV) chloride in dichloromethane solvent at ambient temperature. This way, the stereochemistry in the N-glycosylation reaction is predictable. The stereoselectivity in the N-glycosylation reaction could be organized based on a preferential interaction between the sulfur heteroatom and an appropriate Lewis acid. The use of the Lewis acids TMSOTf and TMSI generates an oxonium ion, which reacts further following pathway A in Scheme 50, and hence no stereocontrol was found in the resultant product. But when the Lewis acid precomplexed the sulfur heteroatom of the ring, selectivity in a diastereofacial manner could be achieved (i.e., pathway B) by complexation *anti* to the protected hydroxymethyl substituent. This complexation may restrict the orientation of the attack of the nucleobase moiety to the α-face. The metal that provides a chloride ligand to the α-face of the oxonium ion could possibly undergo S_N2 attack, and hence the formation of the β-N-nucleoside resulted.

Barral and co-workers [81] synthesized cyclic 2′,3′-dideoxynucleoside compounds in which a 3-hydroxy-2-methylpyridin-4-one species was used as the nucleobase. The synthesized nucleosides **108** contain sugar moieties similar to the oxathiolane nucleosides, namely 3TC (**1**). The heterocyclic base 3-benzyloxy-2-methylpyridin-4-one (**107**) was silylated using HMDS in the presence of catalytic ammonium sulfate. The reaction further involved conventional N-glycosylation with the oxathiolane precursor **8** in 1,2-dichloroethane using TMSOTf as a catalyst. As oxathiolane precursor **8** was sterically impurely obtained from racemic thialactone **104** after reduction with DIBAL and subsequent acetylation, **105** was formed as a mixture of racemic *cis*-nucleosides and racemic *trans*-nucleosides after N-glycosylation. The TBDPS group of these nucleosides was further deprotected using TBAF in THF. Since the Pd cata-



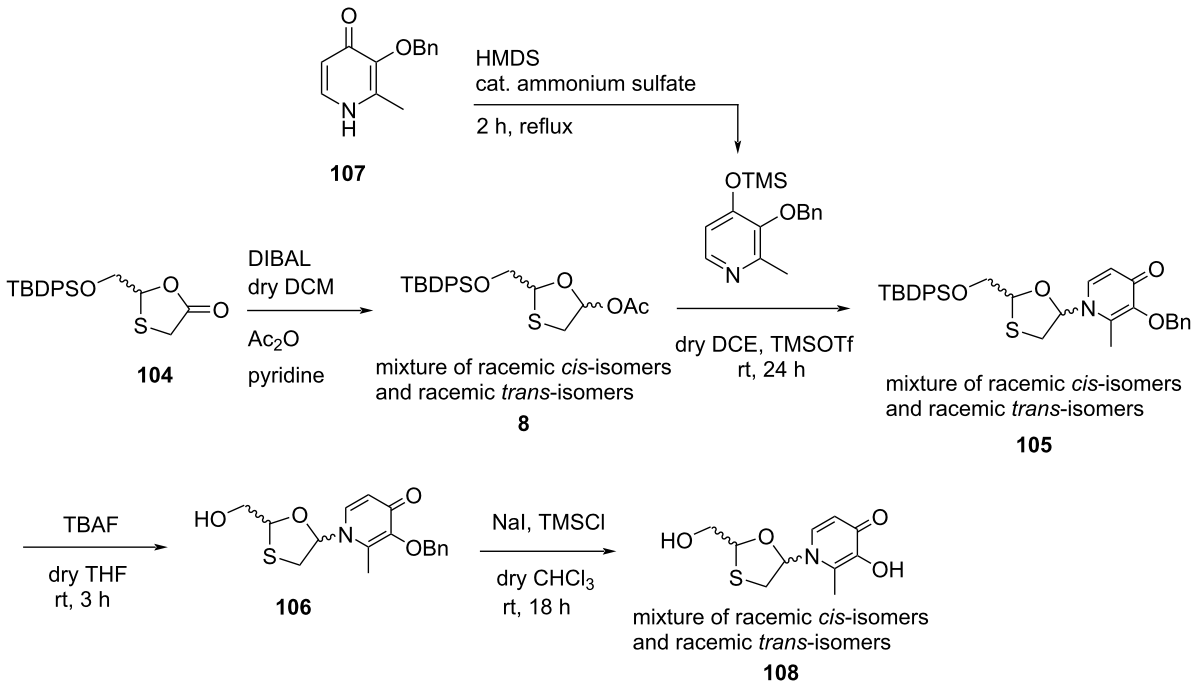
Scheme 49: Pyridinium triflate-mediated N-glycosylation of **35a**.



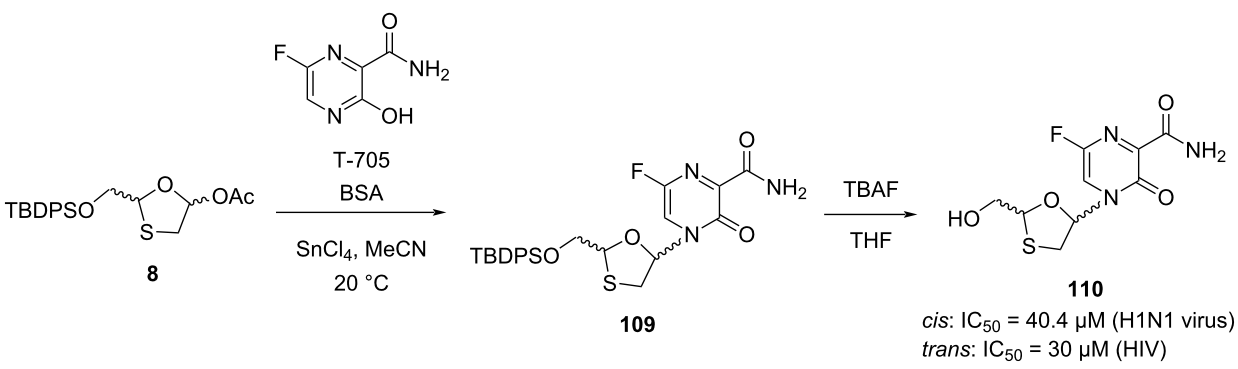
Scheme 50: Possible pathway for stereoselective N-glycosylation via in situ chelation with a metal ligand.

lyst is poisoned due to the sulfur present in the oxathiolane ring, further debenzylation of the nucleobase was achieved by using in situ-generated trimethylsilyl iodide, which gave final **108** as a mixture of racemic *cis*-nucleoside and racemic *trans*-nucleoside (Scheme 51).

A novel class of 1,3-oxathiolane nucleoside derivatives of favipiravir (T-705) was synthesized and investigated recently by Han et al. Some of the analogues were found to have good anti-HIV and anti-H1N1 activity [43]. The N-glycosylation reaction of 1,3-oxathiolane derivative **8** with a novel nucleobase, which is known as T-705, was carried out. Firstly, silylation of nucleobase T-705 was performed in a BSA and acetonitrile mixture. The N-glycosylation was accomplished using $SnCl_4$ catalyst (Scheme 52), providing nucleoside **109**, which was further converted to the nucleoside **110** using TBAF in THF. The nucleosides analogue **110** formed as a mixture of *cis*-(45%) and *trans*-isomers (50%). Interestingly, the *cis*-isomer



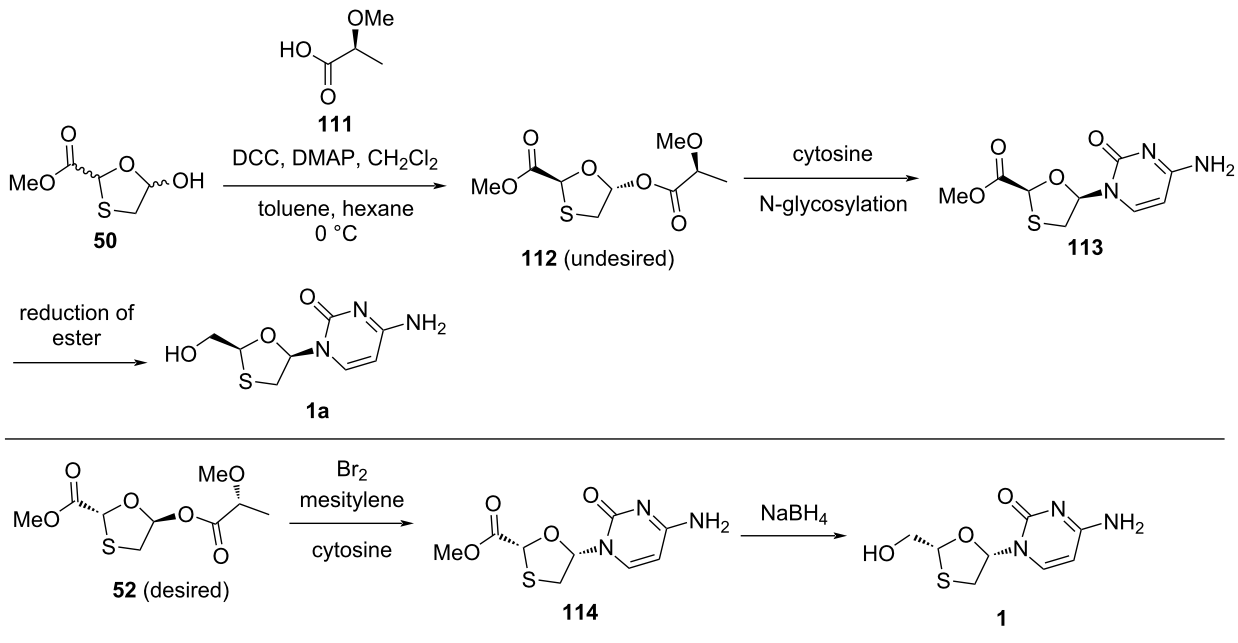
Scheme 51: Synthesis of novel 1,3-oxathiolane nucleoside **108** from oxathiolane precursor **8** and 3-benzyloxy-2-methylpyridin-4-one (**107**).



Scheme 52: Synthesis of **110** using T-705 as a nucleobase and 1,3-oxathiolane derivative **8** via N-glycosylation.

showed activity against the H1N1 influenza virus (IC_{50} = 40.4 μM), while the *trans*-isomer showed weak activity against HIV (IC_{50} = 30 μM).

Snead and co-workers [53] recently developed a new approach for stereoselective nucleoside synthesis that enables a cost-effective approach to lamivudine (**1**, Scheme 53). The synthesis of lamivudine (**1**) was established by employing a method that defines the stereochemistry at the oxathiolane moiety. For this, a commercially available lactic acid derivate **111** served a dual purpose, namely the activation of the anomeric center for N-glycosylation and the transfer of the stereochemical information to the substrate. The enantiomers of the lactic acid derivative **111** are available and used to access the β-enantiomer. The research group also discovered that an asymmetric leaving group was useful for acylation in a selective manner by directing the absolute stereochemistry of the resultant nucleoside, and it provides reliable access to either enantiomer. The acylation of 1,3-oxathiolane **50** with (*S*)-lactic acid derivative **111** and further crystallization in toluene/hexane at 0 °C provided a single isomer **112**. However, compound **112** did not have the desired configuration. Therefore, while using the compound **112**, the undesired 3TC-derived enantiomer **1a** may end up as a final product via N-glycosylation, followed by reduction of the ester group to the primary hydroxy group. Considering this proof of concept, the research group used the other isomer **52** to access the desired configuration of 3TC (**1**). The synthesis was a high-yielding linear four-step sequence that made use of inexpensive raw materials. Also, the use of low-molecular-weight intermediates efficiently increased the material throughput,



Scheme 53: Synthesis of **1** using an asymmetric leaving group and N-glycosylation with bromine and mesitylene.

setting the stage for reduced costs of goods derived from 3TC (**1**). For the N-glycosylation reaction, bromine and mesitylene reagents were used, which generated HBr, and hence acylated oxathiolane **52** was quantitatively transformed in situ to the brominated analogue, which acted as an active precursor to the nucleoside. Then, coupling with the nucleobase cytosine resulted in the formation of nucleoside **114** in good yield. Further, removal of the ester group of **114** using sodium borohydride afforded 3TC (**1**).

Efforts for the separation of racemic mixtures of 1,3-oxathiolane nucleosides

Biological activities of nucleosides generally reside in a single enantiomer, and enzymes are often used for the resolution of racemic nucleosides [82]. To understand which of the enantiomers of a nucleoside has potential antiviral activity, scientist have separated the enantiomers with a variety of methods, such as chiral HPLC as well as enzymatic and chemical methods for the determination of the anti-HIV activity and cytotoxicity in vitro. Coates et al. [83] made efforts for the separation of enantiomers of racemic (±)-BCH-189 (**1c**) with a chiral HPLC method, and it was accomplished using a column known as Cyclobond I acetyl (acetylated β-cyclodextrin). This section reviews the enzymatic as well as the chemical methods used to separate a racemic mixture of 1,3-oxathiolane nucleosides.

Enzymatic methods

The use of enzymes for the resolution of racemic compounds is widespread, and enzymes have been used frequently in the synthesis of nucleosides. The synthesis of optically pure 3TC (**1**) by utilizing enzymatic resolution was also established by Mahmoudian et al. [84]. Cytidine deaminase from *Escherichia coli* deaminated only the D-form of 2'-deoxy-3'-thiacytidine, which converted **1c** to compound **115**, leaving the optically pure L-form 3TC (**1**) unreacted (Scheme 54). The cytidine deaminase EC 3.5.4.5 from *Escherichia coli* easily deaminated 2'-deoxy-3'-thiacytidine in an enantioselective manner and produced optically pure 3TC (**1**).

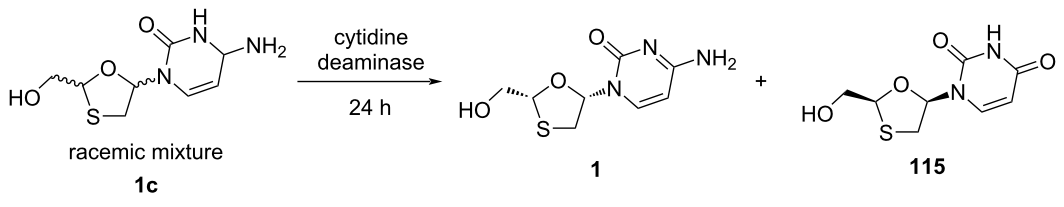
The enzymatic resolution of the monophosphate derivative **116** of (±)-*cis*-[2-(hydroxymethyl)-1,3-oxathiolan-5-yl]cytosine using the 5'-nucleotidase from *Crotalus atrox* venom allowed facile access to the individual enantiomers, which was reported by Storer et al. [24]. The racemic mixture **1c**, upon treatment with phosphorous oxychloride in the presence of trimethyl phosphate at a temperature 0 °C and further appropriate work-up, produced a racemic monophosphate as the ammonium salt **116**. Later, a solution of the racemic monophosphate **116** in an aqueous buffer at 37 °C was prepared from glycine and magnesium chloride upon treatment with 5'-nucleotidase (EC 3.1.3.5), which resulted in a two-component mixture. This was further separated by chromatography, which gave enantiomerically pure (+)-BCH-189 (**1a**) and the monophosphate **117** of (−)-BCH-189 (**1**). The product was then dephosphorylated by an alkaline phosphatase to afford (−)-BCH-189 (**1**, Scheme 55).

Liotta et al. [76] reported an approach for the highly enantioselective resolution to obtain emtricitabine (**2**) as well as related sulfur-containing nucleosides with enzyme catalysis, which uses a PLE-mediated hydrolysis procedure of butyrate ester derivative **118**. The use of the butyrate ester selectively separated the unreacted substrate **119** from the medium by an extraction procedure with chloroform. This process was developed to the synthesis of enantiomerically pure **2** in a gram quantity (Scheme 56).

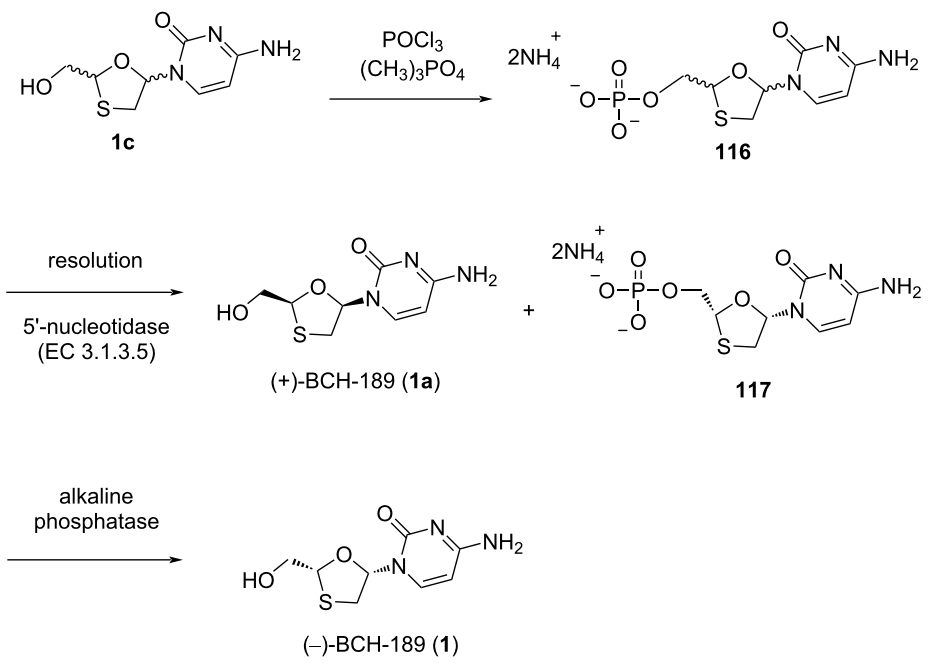
Chemical methods

The choice of a proper resolving agent and an appropriate crystallization solvent are the two determining factors for the successful resolution of enantiomers.

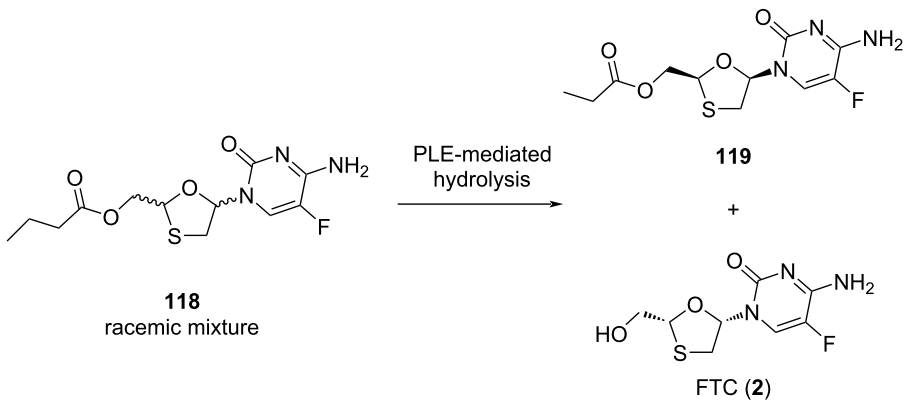
In 2002, Li et al. [85] described the chemical resolution of a racemic mixture of lamivudine (**1**) and **1a** using chiral resolving agents, such as (−)-camphanic acid chloride and (+)-menthyl chloroformate. Out of these two, (+)-menthyl chloroformate was used as a promising resolving agent to separate racemic (±)-BCH-189 (**1c**). The primary amine group was initially protected by acetylation using acetic anhydride in DMF. Further, the corresponding acetyl derivative **120** was reacted with (+)-menthyl chloroformate in the presence of pyridine, providing a mixture of the diastereomers **121** and **122**. The crystallization of these diastereomers in methanol at 0 °C afforded the



Scheme 54: Cytidine deaminase for enzymatic separation of **1c**.



Scheme 55: Enzymatic resolution of the monophosphate derivative **116** for the synthesis of (–)-BCH-189 (**1**) and (+)-BCH-189 (**1a**).

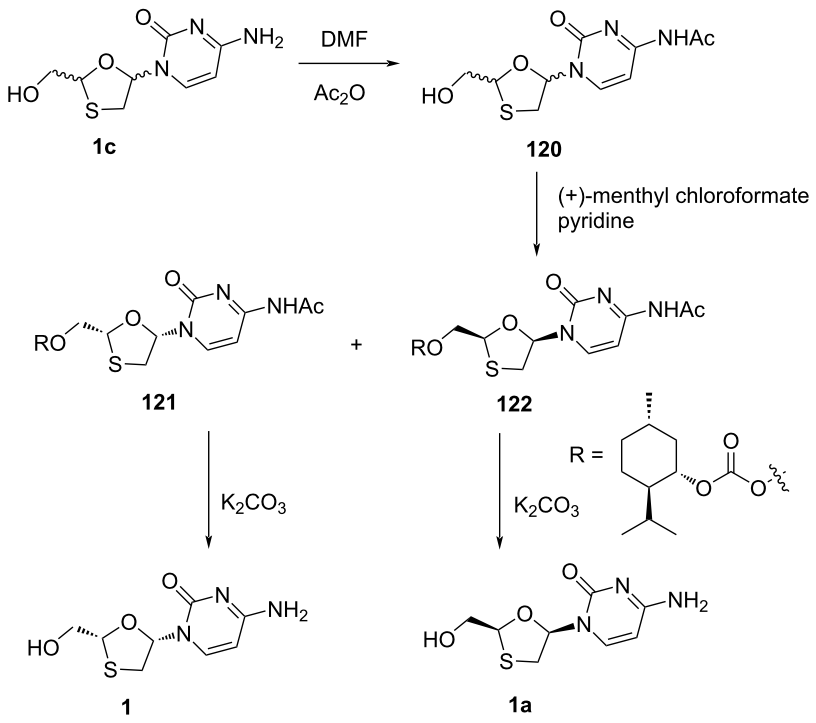


Scheme 56: Enantioselective resolution by PLE-mediated hydrolysis to obtain FTC (**2**).

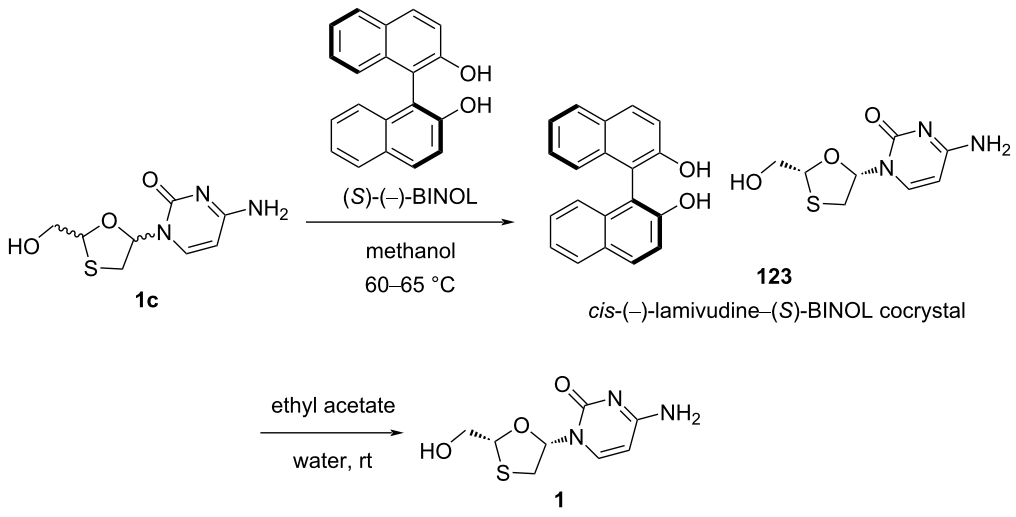
(−)-diastereomer, while the (+)-diastereomer was isolated by concentration, followed by recrystallization from mother liquor. Further, separate deprotection of the diastereomers with potassium carbonate gave the (−)-enantiomer lamivudine (**1**) and the opposite (+)-enantiomer **1a** (Scheme 57).

Through using chiral host compounds, such as dinaphthalenephenols (e.g., BINOL), diphenanthrenols, or tartaric acid derivatives, Deng and co-workers [86] reported the resolution of prazoles. The resolution approach resulted in the formation of a 1:1 of the complex, involving the chiral host and the desired enantiomer as a guest molecule, while the undesired enantiomer remained in solution. (*S*)-Omeprazole, a potent inhibitor of gastric acid secretion, has been isolated in pure form from a racemic mixture by using this chiral host–guest method involving (*S*)-(–)-BINOL.

In 2009, we demonstrated a chemical resolution process for racemic mixture **1c** consisting of lamivudine (**1**) and **1a** by forming cocrystal with (*S*)-(−)-BINOL (Scheme 58) [87]. Lamivudine (**1**) was obtained in high purity and more than 99.9% ee. Interestingly, it was found that the *cis*-(−)- and *trans*-(−)-enantiomers also formed cocrystals with (*S*)-(−)-BINOL, leaving behind the *cis*-(+)- and *trans*-(+)-isomers in the solu-



Scheme 57: (+)-Menthyl chloroformate as a resolving agent to separate a racemic mixture **120**.



Scheme 58: Separation of racemic mixture **1c** by cocrystal **123** formation with (*S*)-(−)-BINOL.

tion. It is worth to mention that the four stereoisomers based on the lamivudine core structure were also separated using this strategy. These isomers were further isolated and characterized. The racemic mixture **1c**, consisting of lamivudine (**1**) and **1a**, was mixed with (*S*)-(−)-BINOL in methanol at reflux temperature. The *cis*-(−)-lamivudine–(*S*)-BINOL complex **123** was isolated at room temperature by filtration. The *cis*-(−)-lamivudine–(*S*)-(−)-BINOL complex **123** was further dissolved in a water/ethyl acetate system where (*S*)-(−)-BINOL was extracted by ethyl acetate and lamivudine (**1**) remained in water.

Overall, over decades, numerous procedures have been utilized to synthesize 1,3-oxathiolane nucleosides. However, from an

industrial perspective, asymmetric procedures are more viable because they are more efficient with respect to the cost and atom economy. Attempts are currently being made to develop cost-effective, simpler, and atom-economic processes for these nucleoside analogues. Glycosylation reactions where the formation of the C–N bond to the anomeric center determines the stereochemistry of the resultant product are crucial. The use of enzymes for these syntheses have also been shown to be an alternative to existing chemical methods. However, the use of enzymes in industry is somewhat difficult to implement, but it is being developed nonetheless because of the current interest in sustainable chemistry. As a result, combined chemoenzymatic procedures can be recognized as a viable alternative to the conventional synthesis for such a type of modified nucleosides.

Conclusion

Invention and improvement of 1,3-oxathiolane nucleosides is indisputably one of the important success stories of recent research in nucleoside chemistry. In the last three decades since Belleau et al. [38] produced the first oxathiolane nucleoside as racemic (±)-BCH-189 (**1c**), significant contributions from different research groups to the stereoselectivity, structural modification, and reactivity of these analogues have led to applications as a variety of therapeutic agents. Due to the biological significance of these nucleosides, efficient synthetic routes with stereocontrol and high yield are in great demand. The heterogeneity of oxathiolane nucleoside structures, along with the unpredictable outcome of the glycosylation process, which depends on the reaction conditions, is influenced by the nature of the glycosyl accepter substrate, catalyst, type of leaving group, and protecting groups, makes it a more daunting task.

However, as we anticipate from the recent research discussed in this review, the efforts made to access such nucleoside analogues are inadequate. One promising area is the field of chiral auxiliaries or Lewis acid catalysis, where the strength of the coupling of the nucleobase with the sugar can allow for enhanced stabilization of oxonium ions, with the potential for stereoselectivity in N-glycosylation via in situ complexation or anchimeric assistance, a key step for nucleoside formation. Significant advances are also being made in fields where the use of enantiomerically pure oxathiolane precursors is more established. In this review, various methodologies to access 1,3-oxathiolane nucleoside analogues that are used to treat AIDS are reviewed. The efforts for the construction of many such nucleoside analogues require a selective glycosylation reaction, which is important to unite appropriate furanose sugar derivatives and nucleobases by forming C–N glycosidic bonds while maintaining selectivity.

In conclusion, we summarize that the desired stereoselectivity in 1,3-oxathiolane nucleoside synthesis was achieved by using i) asymmetric starting materials, ii) asymmetric leaving groups, iii) in situ chelation with appropriate activators or Lewis acids, iv) chiral-auxiliary-induced methods via anchimeric assistance, v) DKR methods, and vi) enzymatic or chemical resolution methods. This classification could be convenient for organic chemists and researchers working in the area of nucleoside research.

Regardless of all the global efforts, HIV/AIDS is still a chronic disease due to rapid mutations of the virus and more prevalent drug resistance. However, many antiviral drugs have been approved, particularly NRTIs, which in most cases can transform this disease from fatal to chronic. L-Nucleosides represent a specific class of drugs with a better activity and low toxicity due to their specific interaction with reverse transcriptase rather than human DNA polymerase. Within this class of drugs, 1,3-oxathiolane nucleosides, such as 3TC (**1**) and FTC (**2**), are still the most representative FDA-approved medicines, indicating a low cytotoxicity and a potent activity both in vitro and in vivo conditions. Their synthesis is still a major concern for researchers, who aim for high yield and good stereoselectivity as well as an inexpensive and safe production. Indeed, a modern synthetic approach is being sought that can avoid the utilization of unstable promoters in N-glycosylation protocols and attain a superior selectivity [88]. More specifically, there is still a need to accelerate new research for discovering anti-HIV agents that feature novel mechanisms of action and can work against drug resistance phenomena. The significance of more creative and easy procedures to access not only 1,3-oxathiolane nucleosides but various nucleoside molecules with desired structural modifications is still a major challenge for synthetic chemists [89,90]. On the grounds of this, synthetic nucleoside analogues have found an application in rational biomolecular designing as well as in medicinal chemistry [90]. These modified 1,3-oxathiolane nucleosides could be transformed into oligonucleotides to investigate the potential to act as antisense nucleosides. Accordingly, because some of these nucleosides have demonstrated anti-HIV, anti-HBV, and anti-H1V1 activity, they could also be thoroughly investigated for further biological activity. This review will benefit researchers in understanding the various processes for synthesizing 1,3-oxathiolane nucleosides as well as their involvement in the chain termination process in medicinal chemistry. The chemists researching modified nucleosides will be encouraged to advance the access to unexplored 1,3-oxathiolane nucleosides and the use of various nucleobases, such as purines, pyrimidines, or their derivatives. Further, it appears promising to develop the stereoselective chemistry of nucleoside analogues to evaluate the resulting products as potential anti-HIV and anticancer drugs.

Acknowledgements

This review is a recognition of the efforts of many prominent scientists working on this difficult subject, especially chemistry of sugar-modified nucleosides.

Funding

The authors would like to thank Lupin Research Park, Lupin Limited for generous funding. U. P. A. thanks Lupin Limited for the ASCENT Ph.D. program.

ORCID® iDs

Umesh P. Aher - https://orcid.org/0000-0002-4513-2766

References

1. Krasnova, L.; Wong, C.-H. *J. Am. Chem. Soc.* **2019,** *141,* 3735–3754. doi:10.1021/jacs.8b11005
2. De Clercq, E. *Curr. Opin. Microbiol.* **2005,** *8,* 552–560. doi:10.1016/j.mib.2005.08.010
3. Wilson, L. J.; Hager, M. W.; El-Kattan, Y. A.; Liotta, D. C. *Synthesis* **1995,** 1465–1479. doi:10.1055/s-1995-4142
4. De Clercq, E.; Neyts, J. Antiviral Agents Acting as DNA or RNA Chain Terminators. In *Antiviral Strategies;* Kräusslich, H. G.; Bartenschlager, R., Eds.; Handbook of Experimental Pharmacology, Vol. 189; Springer: Berlin, Heidelberg, 2009; pp 53–84. doi:10.1007/978-3-540-79086-0_3
5. Cancer. https://www.who.int/news-room/fact-sheets/detail/cancer (accessed June 21, 2021).
6. HIV/AIDS. https://www.who.int/news-room/fact-sheets/detail/hiv-aids (accessed June 21, 2021).
7. Jordheim, L. P.; Durantel, D.; Zoulim, F.; Dumontet, C. *Nat. Rev. Drug Discovery* **2013,** *12,* 447–464. doi:10.1038/nrd4010
8. De Clercq, E.; Li, G. *Clin. Microbiol. Rev.* **2016,** *29,* 695–747. doi:10.1128/cmr.00102-15
9. Albert, M.; De Souza, D.; Feiertag, P.; Hönig, H. *Org. Lett.* **2002,** *4,* 3251–3254. doi:10.1021/ol026460+
10. Gumina, G.; Song, G.-Y.; Chu, C. K. *FEMS Microbiol. Lett.* **2001,** *202,* 9–15. doi:10.1111/j.1574-6968.2001.tb10773.x
11. Nair, V.; Jahnke, T. S. *Antimicrob. Agents Chemother.* **1995,** *39,* 1017–1029. doi:10.1128/aac.39.5.1017
12. Wang, P.; Hong, J. H.; Cooperwood, J. S.; Chu, C. K. *Antiviral Res.* **1998,** *40,* 19–44. doi:10.1016/s0166-3542(98)00041-2
13. Soudeyns, H.; Yao, X. I.; Gao, Q.; Belleau, B.; Kraus, J. L.; Nguyen-Ba, N.; Spira, B.; Wainberg, M. A. *Antimicrob. Agents Chemother.* **1991,** *35,* 1386–1390. doi:10.1128/aac.35.7.1386
14. Schinazi, R. F.; Chu, C. K.; Peck, A.; McMillan, A.; Mathis, R.; Cannon, D.; Jeong, L. S.; Beach, J. W.; Choi, W. B.; Yeola, S. *Antimicrob. Agents Chemother.* **1992,** *36,* 672–676. doi:10.1128/aac.36.3.672
15. Choi, Y.; Choo, H.; Chong, Y.; Lee, S.; Olgen, S.; Schinazi, R. F.; Chu, C. K. *Org. Lett.* **2002,** *4,* 305–307. doi:10.1021/ol0171665
16. Chu, C. K.; Babu, J. R.; Beach, J. W.; Ahn, S. K.; Huang, H.; Jeong, L. S.; Lee, S. J. *J. Org. Chem.* **1990,** *55,* 1418–1420. doi:10.1021/jo00292a006
17. Bolon, P. J.; Wang, P.; Chu, C. K.; Gosselin, G.; Boudou, V.; Pierra, C.; Mathé, C.; Imbach, J.-L.; Faraj, A.; el Alaoui, A.; Sommadossi, J.-P.; Pai, S. B.; Zhu, Y.-L.; Lin, J.-S.; Cheng, Y.-C.; Schinazi, R. F. *Bioorg. Med. Chem. Lett.* **1996,** *6,* 1657–1662. doi:10.1016/0960-894x(96)00293-4
18. Wilson, L. J.; Liotta, D. *Tetrahedron Lett.* **1990,** *31,* 1815–1818. doi:10.1016/s0040-4039(00)98793-8
19. Chen, S.-H.; Li, X.; Li, J.; Niu, C.; Carmichael, E.; Doyle, T. W. *J. Org. Chem.* **1997,** *62,* 3449–3452. doi:10.1021/jo970177k
20. Lee, K.; Chu, C. K. *Antimicrob. Agents Chemother.* **2001,** *45,* 138–144. doi:10.1128/aac.45.1.138-144.2001
21. Blackburn, G. M.; Gait, M. J.; Loakes, D.; Williams, D. M., Eds. *Nucleic Acids in Chemistry and Biology;* Royal Society of Chemistry: Cambridge, United Kingdom, 2007. doi:10.1039/9781847555380
22. Kim, H. O.; Ahn, S. K.; Alves, A. J.; Beach, J. W.; Jeong, L. S.; Choi, B. G.; Van Roey, P.; Schinazi, R. F.; Chu, C. K. *J. Med. Chem.* **1992,** *35,* 1987–1995. doi:10.1021/jm00089a007
23. Choi, W.-B.; Yeola, S.; Liotta, D. C.; Schinazi, R. F.; Painter, G. R.; Davis, M.; St. Clair, M.; Furman, P. A. *Bioorg. Med. Chem. Lett.* **1993,** *3,* 693–696. doi:10.1016/s0960-894x(01)81256-7
24. Storer, R.; Clemens, I. R.; Lamont, B.; Noble, S. A.; Williamson, C.; Belleau, B. *Nucleosides Nucleotides* **1993,** *12,* 225–236. doi:10.1080/07328319308021208
25. Cihlar, T.; Ray, A. S. *Antiviral Res.* **2010,** *85,* 39–58. doi:10.1016/j.antiviral.2009.09.014
26. Sarafianos, S. G.; Das, K.; Clark, A. D., Jr.; Ding, J.; Boyer, P. L.; Hughes, S. H.; Arnold, E. *Proc. Natl. Acad. Sci. U. S. A.* **1999,** *96,* 10027–10032. doi:10.1073/pnas.96.18.10027
27. Furman, P. A.; Davis, M.; Liotta, D. C.; Paff, M.; Frick, L. W.; Nelson, D. J.; Dornsife, R. E.; Wurster, J. A.; Wilson, L. J.; Fyfe, J. A. *Antimicrob. Agents Chemother.* **1992,** *36,* 2686–2692. doi:10.1128/aac.36.12.2686
28. Kukhanova, M.; Liu, S.-H.; Mozzherin, D.; Lin, T.-S.; Chu, C. K.; Cheng, Y.-C. *J. Biol. Chem.* **1995,** *270,* 23055–23059. doi:10.1074/jbc.270.39.23055
29. Chen, C. H.; Vazquez-Padua, M.; Cheng, Y. C. *Mol. Pharmacol.* **1991,** *39,* 625–628.
30. Schinazi, R. F.; McMillan, A.; Cannon, D.; Mathis, R.; Lloyd, R. M.; Peck, A.; Sommadossi, J. P.; St Clair, M.; Wilson, J.; Furman, P. A. *Antimicrob. Agents Chemother.* **1992,** *36,* 2423–2431. doi:10.1128/aac.36.11.2423
31. Liotta, D. C.; Schinazi, R. F.; Choi, W. B. Method of resolution and antiviral activity of 1,3-oxathiolane nuclesoside enantiomers. U. S. Patent US6,703,396B1, March 9, 2004.
32. McConathy, J.; Owens, M. J. *Prim. Care Companion J. Clin. Psychiatry* **2003,** *5,* 70–73. doi:10.4088/pcc.v05n0202
33. Gumina, G.; Cooperwood, J. S.; Chu, C. K. Oxathiolane and Dioxolane Nucleosides: Synthesis and Antiviral Activity. *Antiviral Nucleoside: chiral synthesis and chemotherapy;* Elsevier: Amsterdam, Netherlands, 2003; pp 191–258. doi:10.1016/b978-044451319-9/50003-8
34. D'alonzo, D.; Guaragna, A. Stereoselective Methods in the Synthesis of Bioactive Oxathiolane and Dioxolane Nucleosides. *Chemical Synthesis of Nucleoside Analogues;* John Wiley & Sons: Hoboken, NJ, USA, 2013; pp 727–780. doi:10.1002/9781118498088.ch16
35. De Clercq, E. *J. Med. Chem.* **2019,** *62,* 7322–7339. doi:10.1021/acs.jmedchem.9b00175
36. De Clercq, E. *Nat. Rev. Drug Discovery* **2002,** *1,* 13–25. doi:10.1038/nrd703
37. De Clercq, E. *Annu. Rev. Pharmacol. Toxicol.* **2011,** *51,* 1–24. doi:10.1146/annurev-pharmtox-010510-100228

38. Belleau, B.; Dixit, D.; Nguyen-Ba, N.; Kraus, J. L. Abstract T.C.0.1. 5th International Conference on AIDS, Montreal, Canada, June 4–9, 1989; p 215.
39. Sadayoshi, S.; Shuichi, I.; Hiroshi, K.; Kozo, T.; Shinichi, M. *Bull. Chem. Soc. Jpn.* **1972,** *45,* 913–915.
40. Chu, C. K.; Beach, J. W.; Jeong, L. S.; Choi, B. G.; Comer, F. I.; Alves, A. J.; Schinazi, R. F. *J. Org. Chem.* **1991,** *56,* 6503–6505. doi:10.1021/jo00023a010
41. Jeong, L. S.; Alves, A. J.; Carrigan, S. W.; Kim, H. O.; Beach, J. W.; Chu, C. K. *Tetrahedron Lett.* **1992,** *33,* 595–598. doi:10.1016/s0040-4039(00)92319-0
42. Jeong, L. S.; Schinazi, R. F.; Beach, J. W.; Kim, H. O.; Shanmuganathan, K.; Nampalli, S.; Chun, M. W.; Chung, W. K.; Choi, B. G.; Chu, C. K. *J. Med. Chem.* **1993,** *36,* 2627–2638. doi:10.1021/jm00070a006
43. Han, M.; Zhao, X.; Wu, X.; Huang, W.; Li, X.; Yu, F. *Med. Chem.* **2018,** *14,* 595–603. doi:10.2174/1573406414666180112102225
44. Beach, J. W.; Jeong, L. S.; Alves, A. J.; Pohl, D.; Kim, H. O.; Chang, C. N.; Doong, S. L.; Schinazi, R. F.; Cheng, Y. C.; Chu, C. K. *J. Org. Chem.* **1992,** *57,* 2217–2219. doi:10.1021/jo00034a006
45. Humber, D. C.; Jones, M. F.; Payne, J. J.; Ramsay, M. V. J.; Zacharie, B.; Jin, H.; Siddiqui, A.; Evans, C. A.; Tse, H. L. A.; Mansour, T. S. *Tetrahedron Lett.* **1992,** *33,* 4625–4628. doi:10.1016/s0040-4039(00)61330-8
46. Jin, H.; Siddiqui, M. A.; Evans, C. A.; Tse, H. L. A.; Mansour, T. S.; Goodyear, M. D.; Ravenscroft, P.; Beels, C. D. *J. Org. Chem.* **1995,** *60,* 2621–2623. doi:10.1021/jo00113a050
47. Milton, J.; Brand, S.; Jones, M. F.; Rayner, C. M. *Tetrahedron Lett.* **1995,** *36,* 6961–6964. doi:10.1016/0040-4039(95)01380-z
48. Kraus, J.-L.; Attardo, G. *Synthesis* **1991,** 1046–1048. doi:10.1055/s-1991-26643
49. Cousins, R. P. C.; Mahmoudian, M.; Youds, P. M. *Tetrahedron: Asymmetry* **1995,** *6,* 393–396. doi:10.1016/0957-4166(95)00022-h
50. Faury, P.; Camplo, M.; Charvet, A.-S.; Chermann, J.-C.; Kraus, J.-L. *Nucleosides Nucleotides* **1992,** *11,* 1481–1488. doi:10.1080/07328319208021190
51. Kraus, J.-L. *Nucleosides Nucleotides* **1993,** *12,* 157–162. doi:10.1080/07328319308021202
52. Chao, Q.; Nair, V. *Bioorg. Med. Chem. Lett.* **1997,** *7,* 1199–2002. doi:10.1016/s0960-894x(97)00178-9
53. Snead, D. R.; McQuade, D. T.; Ahmad, S.; Krack, R.; Stringham, R. W.; Burns, J. M.; Abdiaj, I.; Gopalsamuthiram, V.; Nelson, R. C.; Gupton, B. F. *Org. Process Res. Dev.* **2020,** *24,* 1194–1198. doi:10.1021/acs.oprd.0c00083
54. Kashinath, K.; Snead, D. R.; Burns, J. M.; Stringham, R. W.; Gupton, B. F.; McQuade, D. T. *Org. Process Res. Dev.* **2020,** *24,* 2266–2270. doi:10.1021/acs.oprd.0c00145
55. Goodyear, M. D.; Hill, M. L.; West, J. P.; Whitehead, A. J. *Tetrahedron Lett.* **2005,** *46,* 8535–8538. doi:10.1016/j.tetlet.2005.10.002
56. Vorbrüggen, H.; Ruh-Pohlenz, C. *Org. React.* **1999,** *55,* 1–630. doi:10.1002/0471264180.or055.01
57. Han, M.-Y.; Yang, F.-Y.; Zhou, D.; Xu, Z. *Org. Biomol. Chem.* **2017,** *15,* 1418–1425. doi:10.1039/c7ob00005g
58. Pau, A. K.; George, J. M. *Infect. Dis. Clin. North Am.* **2014,** *28,* 371–402. doi:10.1016/j.idc.2014.06.001
59. Ren, Y.; Hu, L.; Ramström, O. *Mol. Catal.* **2019,** *468,* 52–56. doi:10.1016/j.mcat.2019.02.013
60. Hu, L.; Ren, Y.; Ramström, O. *J. Org. Chem.* **2015,** *80,* 8478–8481. doi:10.1021/acs.joc.5b01585
61. Hu, L.; Schaufelberger, F.; Zhang, Y.; Ramström, O. *Chem. Commun.* **2013,** *49,* 10376–10378. doi:10.1039/c3cc45551c
62. Chen, Y.; Zhang, X.; Zheng, G.; Gao, S. *Process Biochem. (Oxford, U. K.)* **2019,** *81,* 77–84. doi:10.1016/j.procbio.2019.03.025
63. Zhang, Y.; Sun, Y.; Tang, H.; Zhao, Q.; Ren, W.; Lv, K.; Yang, F.; Wang, F.; Liu, J. *Org. Process Res. Dev.* **2020,** *24,* 579–587. doi:10.1021/acs.oprd.0c00010
64. Zhang, Y.; Schaufelberger, F.; Sakulsombat, M.; Liu, C.; Ramström, O. *Tetrahedron* **2014,** *70,* 3826–3831. doi:10.1016/j.tet.2014.03.059
65. Romeo, G.; Chiacchio, U.; Corsaro, A.; Merino, P. *Chem. Rev.* **2010,** *110,* 3337–3370. doi:10.1021/cr800464r
66. Nigudkar, S. S.; Demchenko, A. V. *Chem. Sci.* **2015,** *6,* 2687–2704. doi:10.1039/c5sc00280j
67. Bennett, C. S. *Selective Glycosylations: Synthetic Methods and Catalysts;* Wiley-VCH: Weinheim, Germany, 2017. doi:10.1002/9783527696239
68. Crich, D. *Acc. Chem. Res.* **2010,** *43,* 1144–1153. doi:10.1021/ar100035r
69. Zhu, X.; Schmidt, R. R. *Angew. Chem., Int. Ed.* **2009,** *48,* 1900–1934. doi:10.1002/anie.200802036
70. Séquin, U. *Experientia* **1973,** *29,* 1059–1062. doi:10.1007/bf01946717
71. Miljković, M. Chemistry of the Glycosidic Bond. *Carbohydrates;* Springer: New York, NY, USA, 2010; pp 323–421. doi:10.1007/978-0-387-92265-2_12
72. Choi, W. B.; Wilson, L. J.; Yeola, S.; Liotta, D. C.; Schinazi, R. F. *J. Am. Chem. Soc.* **1991,** *113,* 9377–9379. doi:10.1021/ja00024a058
73. Jeong, L. S.; Schinazi, R. F.; Beach, J. W.; Kim, H. O.; Nampalli, S.; Shanmuganathan, K.; Alves, A. J.; McMillan, A.; Chu, C. K.; Mathis, R. *J. Med. Chem.* **1993,** *36,* 181–195. doi:10.1021/jm00054a001
74. Huerta, F. F.; Minidis, A. B. E.; Bäckvall, J.-E. *Chem. Soc. Rev.* **2001,** *30,* 321–331. doi:10.1039/b105464n
75. Aher, U. P.; Srivastava, D.; Jadhav, H. S.; Singh, G. P.; S., J. B.; Shenoy, G. G. *Org. Process Res. Dev.* **2020,** *24,* 387–397. doi:10.1021/acs.oprd.9b00414
76. Hoong, L. K.; Strange, L. E.; Liotta, D. C.; Koszalka, G. W.; Burns, C. L.; Schinazi, R. F. *J. Org. Chem.* **1992,** *57,* 5563–5565. doi:10.1021/jo00047a004
77. Camplo, M.; Faury, P.; Charvet, A.-S.; Lederer, F.; Chermann, J.-C.; Kraus, J.-L. *Nucleosides Nucleotides* **1993,** *12,* 631–641. doi:10.1080/07328319308019017
78. Mansour, T.; Jin, H.; Allan Tse, H. L.; Siddiqui, A. Processes for the diastereoselective synthesis of nucleoside analogues. U.S. Patent US5,696,254A, Dec 9, 1997.
79. Caso, M. F.; D'Alonzo, D.; D'Errico, S.; Palumbo, G.; Guaragna, A. *Org. Lett.* **2015,** *17,* 2626–2629. doi:10.1021/acs.orglett.5b00982
80. Mandala, D.; Watts, P. *ChemistrySelect* **2017,** *2,* 1102–1105. doi:10.1002/slct.201700052
81. Barral, K.; Hider, R. C.; Balzarini, J.; Neyts, J.; De Clercq, E.; Camplo, M. *Bioorg. Med. Chem. Lett.* **2003,** *13,* 4371–4374. doi:10.1016/j.bmcl.2003.09.033
82. Secrist, J. A., III; Montgomery, J. A.; Shealy, Y. F.; O'Dell, C. A.; Clayton, S. J. *J. Med. Chem.* **1987,** *30,* 746–749. doi:10.1021/jm00387a032
83. Coates, J. A.; Cammack, N.; Jenkinson, H. J.; Mutton, I. M.; Pearson, B. A.; Storer, R.; Cameron, J. M.; Penn, C. R. *Antimicrob. Agents Chemother.* **1992,** *36,* 202–205. doi:10.1128/aac.36.1.202

84. Mahmoudian, M.; Baines, B. S.; Drake, C. S.; Hale, R. S.; Jones, P.; Piercey, J. E.; Montgomery, D. S.; Purvis, I. J.; Storer, R.; Dawson, M. J.; Lawrence, G. C. *Enzyme Microb. Technol.* **1993,** *15,* 749–755. doi:10.1016/0141-0229(93)90005-m
85. Li, J.-z.; Gao, L.-x.; Ding, M.-x. *Synth. Commun.* **2002,** *32,* 2355–2359. doi:10.1081/scc-120006006
86. Deng, J.; Chi, Y.; Fu, F.; Cui, X.; Yu, K.; Zhu, J.; Jiang, Y. *Tetrahedron: Asymmetry* **2000,** *11,* 1729–1732. doi:10.1016/s0957-4166(00)00114-2
87. Roy, B. N.; Singh, G. P.; Srivastava, D.; Jadhav, H. S.; Saini, M. B.; Aher, U. P. *Org. Process Res. Dev.* **2009,** *13,* 450–455. doi:10.1021/op800228h
88. Caso, M. F. New synthetic routes to the stereoselective assembly and the elaboration of bioactive compounds. Ph.D. Thesis, Università degli Studi di Napoli Federico II, Napels, Italy, 2016.
89. Kapdi, A. R.; Sanghvi, Y. S. The Future of Drug Discovery: The Importance of Modified Nucleosides, Nucleotides, and Oligonucleotides. In *Palladium-Catalyzed Modification of Nucleosides, Nucleotides and Oligonucleotides;* Kapdi, A. R.; Maiti, D.; Sanghvi, Y. S., Eds.; Elsevier: Amsterdam, Netherlands, 2018; pp 1–18. doi:10.1016/b978-0-12-811292-2.00001-5
90. Verma, V.; Maity, J.; Maikhuri, V. K.; Sharma, R.; Ganguly, H. K.; Prasad, A. K. *Beilstein J. Org. Chem.* **2021,** *17,* 1392–1439. doi:10.3762/bjoc.17.98

License and Terms