



This open access document is posted as a preprint in the Beilstein Archives at <https://doi.org/10.3762/bxiv.2024.6.v1> and is considered to be an early communication for feedback before peer review. Before citing this document, please check if a final, peer-reviewed version has been published.

This document is not formatted, has not undergone copyediting or typesetting, and may contain errors, unsubstantiated scientific claims or preliminary data.

Preprint Title Monitoring carbohydrate 3D structure quality with the *Privateer* database

Authors Jordan S. Dialpuri, Haroldas Bagdonas, Lucy C. Schofield, Phuong Thao Pham, Lou Holland and Jon Agirre

Publication Date 31 Jan. 2024

Article Type Full Research Paper

ORCID® IDs Jordan S. Dialpuri - <https://orcid.org/0000-0002-6205-2661>;
Haroldas Bagdonas - <https://orcid.org/0000-0001-5028-4847>; Lucy C. Schofield - <https://orcid.org/0009-0001-2069-878X>; Phuong Thao Pham - <https://orcid.org/0000-0002-6205-1298>; Lou Holland - <https://orcid.org/0000-0002-3867-1833>; Jon Agirre - <https://orcid.org/0000-0002-1086-0253>



License and Terms: This document is copyright 2024 the Author(s); licensee Beilstein-Institut.

This is an open access work under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>). Please note that the reuse, redistribution and reproduction in particular requires that the author(s) and source are credited and that individual graphics may be subject to special legal provisions.

The license is subject to the Beilstein Archives terms and conditions: <https://www.beilstein-archives.org/xiv/terms>.

The definitive version of this work can be found at <https://doi.org/10.3762/bxiv.2024.6.v1>

Monitoring carbohydrate 3D structure quality with the *Privateer* database

Jordan S. Dialpuri¹, Haroldas Bagdonas¹, Lucy C. Schofield¹, Phuong Thao Pham¹, Lou Holland¹ and Jon Agirre^{1,*}

¹ York Structural Biology Laboratory, Department of Chemistry, University of York, UK.

*Correspondence: jon.agirre@york.ac.uk

Abstract

The remediation of the carbohydrate data of the Protein Data Bank (PDB) has brought numerous enhancements to the findability and interpretability of deposited glycan structures, yet crucial quality indicators are either missing or hard to find on the PDB pages. Without a way to access wider glyco-chemical context, problematic structures may be taken as fact by keen but inexperienced scientists. The *Privateer* software is a validation and analysis tool that provides access to a number of metrics and links to external experimental resources, allowing users to evaluate structures using carbohydrate-specific methods. Here we present the *Privateer* database, a free resource that aims to complement the growing glycan content of the PDB.

Keywords

database, validation, polysaccharides, carbohydrates, *N*-glycosylation, *N*-glycans, website

Introduction

Carbohydrate modelling is an important but often cumbersome stage in the macromolecular X-ray structure solution workflow. Modelling glycoproteins accurately is pivotal in understanding the complex biochemical interactions which affect the physiological function of cells [1]. Any mechanistic analyses done with finely grained approaches such as QM/MM [2] rely heavily on the correctness of the starting coordinates. Despite this, carbohydrate models often contain modelling inconsistencies that cannot easily be attributed to known biochemical principles [3]. These inconsistencies cannot solely be attributed to model-building inexperience, as carbohydrate model-building is an inherently difficult task, which in the past has been plagued with software related problems: from incorrect libraries to incomplete support [4]. Carbohydrates are mobile, highly branched additions to the comparatively rigid protein framework; in macromolecular crystallography, this causes heterogeneity throughout the crystal lattice, and therefore poorly resolved density regions, whereas in electron cryo-microscopy different conformations and compositions are averaged out during image classification and volume reconstruction [5].

Owing to these difficulties, it is not uncommon to find problematic carbohydrate structures in the PDB: from the initial works of Lütteke, Frank and von der Lieth [6,7], who identified numerous issues affecting nomenclature and linkages (estimated to affect 30% of the structures at the time), to the reports of surprising – or indeed glyco-chemically impossible – linkages in a glycoprotein by Crispin and collaborators [8], and more recently the realisation

that high-energy ring conformations, a rare event in six-membered pyranosides, were present in ~15% of the *N*-glycan components of glycoproteins in the PDB [3]. Many of these findings originated the development of new resources, including services and databases [9–13], and standalone software [14–18]. Among these, the *Privateer* software package has been a key tool for glycoprotein and protein-carbohydrate complex validation: *Privateer* analyses the conformational plausibility of each sugar model [3], checks that structures match the nomenclature used for deposition in the PDB [14], compares glycan compositions to known structures as reported by glycomics (e.g. GlyConnect [19]) and glyco-informatics (e.g. GlyTouCan [20]) databases and repositories [15], and checks how close the overall conformation of *N*-glycans comes to that of validated deposited structures [16].

The PDB-REDO [21] database is a separate resource, albeit linked to the PDB in that the entries that compound PDB-REDO are those original PDB crystallographic entries that included experimental data (i.e. reflection intensities or amplitudes): each entry includes a re-refined, sometimes even re-built to some extent, copy of the original model. These newer versions are produced with state of the art methods, many of which were probably not available at the time of deposition, hence the quality of the models is expected to improve. Because the methodology included in PDB-REDO had been affected by the lack of automatic support that plagued general purpose crystallographic model building and refinement software [4], carbohydrate-specific methods have been gradually introduced over the years [22,23].

Whilst *Privateer* has been a *staple* tool in carbohydrate validation, the results of *Privateer* have not been collated in such a way that allows for easy judgement of carbohydrate model quality in the Protein Data Bank (PDB). Providing users with metrics which allow them to make chemically sound conclusions about the model is an important facility, especially for novice users. To allow this to happen readily on PDB [24] distribution sites, we present the *Privateer* database, a freely available, up-to-date collection of validation information for both the PDB [24] and PDB-REDO [21] archives.

Materials and methods

The *Privateer* software package [14] was used to compute metrics and statistics for each entry in the PDB [24], or PDB-REDO [21]. For each structure in the PDB, the carbohydrate-containing chains are first identified before being validated using the suite of validation tools available within *Privateer*. Using the Python bindings available within the latest versions of *Privateer* a validation report can be generated for each carbohydrate in the molecular model. This report is put out in JSON format for easy consumption by web-based database frontends. The initial report generation was completed in parallel over 64 CPU cores in around 5 hours. After the initial surveys through the PDB and PDB-REDO, this process only needs to be completed when new molecular models are deposited into the PDB, which occurs weekly. Although compiling validation reports for only new structures would be more efficient, this would fail to encompass changes in structures in historical entries, therefore the *Privateer* database is recompiled weekly.

The database is hosted on GitHub, which receives any updates to the reports after recompilation. The database is separated into PDB and PDB-REDO sections, which are in turn structured in the same format as the PDB archive, separated into folders by the middle

two characters of the PDB four-letter code. For convenience, the presentation of the database is hosted alongside the *Privateer Web App* [25]; the database part can be accessed at <https://privateer.york.ac.uk/database> or by navigating to the database icon on the top right of the screen. The website is dynamic and compatible with desktop and laptop computers, plus tablets and smartphones.

Results and Discussion

Format of the Validation Report

The JSON file deposited for each PDB entry follows a consistent format, as shown in Figure 1. At the top level, the file contains metadata about the validation report. This metadata provides the date that the validation report was generated as well as the availability of experimental data. It is helpful to have this information easily accessible as *Privateer* cannot calculate the real space correlation coefficient without experimental data, therefore programmatic access to further validation metrics could be streamlined, knowing the information is not present.

Also at the top level of the validation report is the beginning of the carbohydrate information, listed as 'glycans' in the JSON format. Within this 'glycan' scope, information is segmented into glycan type, i.e. 'n-glycan', 'o-glycan', 's-glycan', 'c-glycan' and 'ligand'. Each of these glycan types contains an array of individual glycans of that type and the format of the data inside each of these glycan types is identical.

The data contained in each glycan entry is shown in Table 1. Each entry contains information about the protein chain attachment, the number of sugars in the glycan, the WURCS2.0 code [26], the standard nomenclature for glycan SVG and an array of sugar entries. The validation data calculated by *Privateer* for each sugar entry is shown in Table 2.

Table 1 - Data contained within each glycan entry

Key	Example	Type
proteinResidueType	ASN	String
proteinResidueId	61	String
proteinResidueSeqnum	61	Number
proteinChainId	A	String
rootSugarChainId	C	String
numberOfSugars	7	Number
wurcs	WURCS=2.0/3,7,6/...	String
snfg	<svg> ... </svg>	String
sugars	See Table 2	Array

Table 2 - Data contained within each sugar entry

Key	Example	Type
sugarID	NAG-D-1	String
q	0.54	Number
phi	303.44	Number
theta	6.45	Number
rsc	0.922	Number
detectedType	beta-D-aldopyranose	String
conformation	4c1	String
bFactor	22.367	Number
mFo	0.421	Number
diagnostic	yes	String

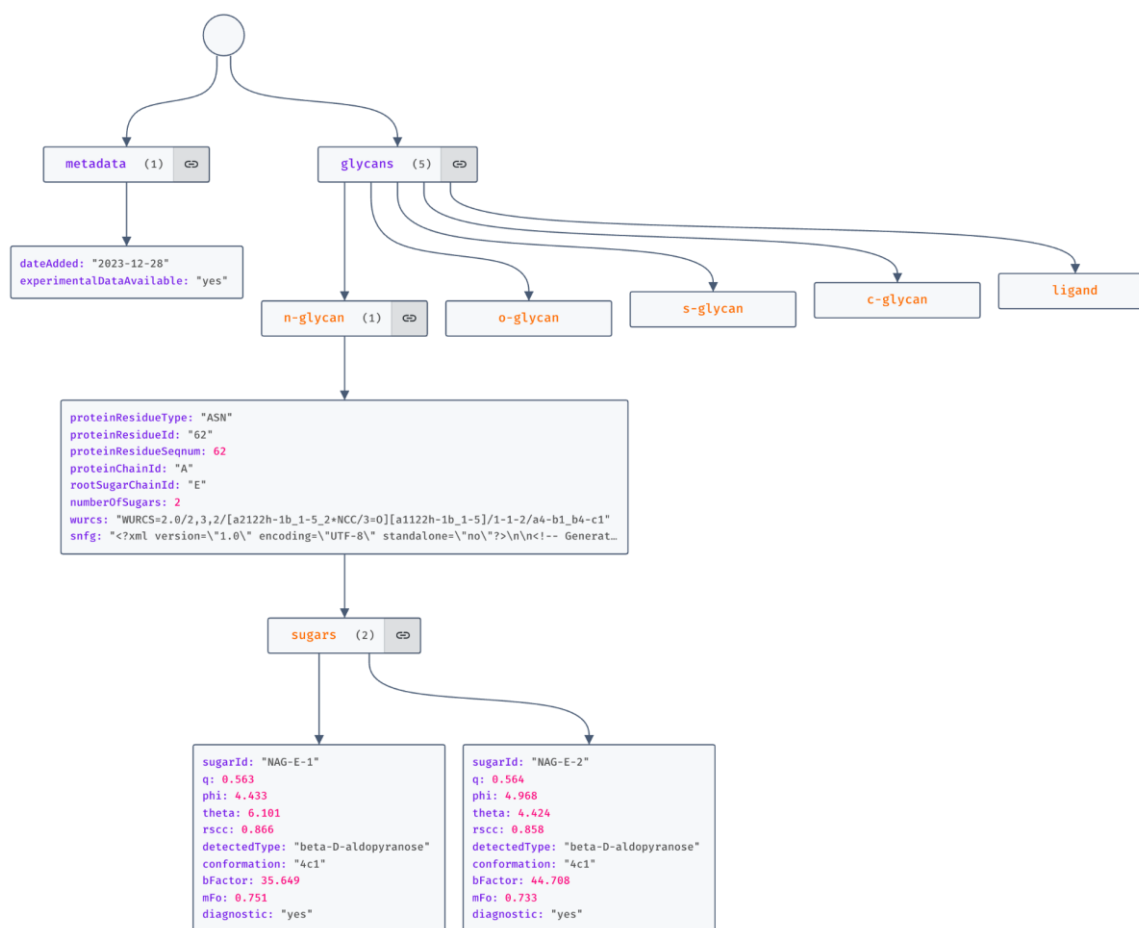


Figure 1 - Format of validation report in JSON format. At the top level of the tree, the report contains metadata about itself, such as the date the entry was added to the database and if experimental data is available. Also at the top level of the tree is the glycan information, separated into glycan type. Each glycan also contains a list of sugars, with a range of validation information. Tree visualisation was created with jsoncrack.com.

Visualising a Validation Report

While the database is available on GitHub for programmatic access, viewing a validation report entry in plaintext can be difficult, time-consuming and would certainly be a poor experience for the end user. To improve the utility of this database, we have provided a visualisation of the information contained within the validation report for both PDB and PDB-REDO databases, which is available alongside the *Privateer Web App* [25] - <https://privateer.york.ac.uk/database>.

The first section of this visual report displays a global outlook on the validity of the model through two graphs. The first graph shows the conformational landscape for the pyranose sugars. For a sugar model to be deemed valid, the ring must be in the 4C_1 chair conformation. This can be measured through the Cremer-Pople parameters θ and ψ [27], theta angles of $0^\circ < \theta < 360^\circ$ indicate that the sugar may be in a higher energy confirmation, and therefore caution should be placed on any conclusions drawn from the sugars molecular model. Also in the first section of the visual validation report is a plot of B-factor (temperature factor) versus the real space correlation coefficient (RSCC). A well-refined, well-built model would expect to have a B factor which increases somewhat linearly as the RSCC decreases. Over-refined models may deviate from this trend and would be trivial to identify.

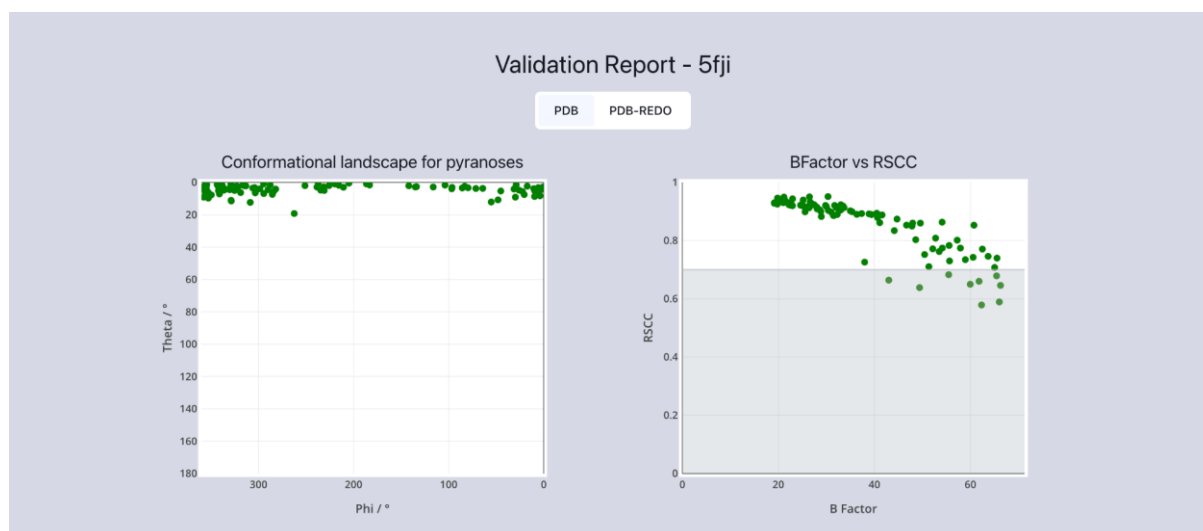


Figure 2 - Left: Graphical representation of the conformational landscape of pyranose sugars, a well-modelled sugar would be expected to be in the lowest energy conformation and have a theta angle close to 0° . Right: Real space correlation coefficient plotted against B-factor, which enables the refinement of the sugars to be assessed. A slight negative correlation would be expected for a well-refined model.

The validation report also displays a table representing two-dimensional descriptions of each glycan in the model. Each row in the table represents a unique glycan and includes the chain identifier, standard Symbol Nomenclature for Glycans (SNFG [28]) visualisation and copyable WURCS [26] identifier. The SNFG displayed for each glycan paints a picture of how well built the glycan model is, as the metrics and validity conclusions calculated by *Privateer* are

embedded within each shape and linkage of the diagram. For example, a shape with an orange highlight indicates something is abnormal about the ring's conformation, puckering or monosaccharide nomenclature [29]. Similarly, a linkage with an orange highlight indicates that the torsion angles between the linkages are unexpected and require further inspection [16].

In addition to the SNFG, also displayed for each table entry is a copyable WURCS link, which encodes the complete glycan format in a linear code. The decision to present this information as a copyable link, as opposed to as plaintext is due to the inherent difficulty and unlikeliness for a human to read and understand the WURCS code. It is much more likely that the WURCS code would be copied and searched for in a glycomics database, hence we provide that functionality in a streamlined way.



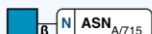





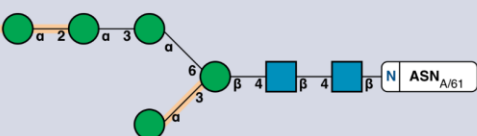

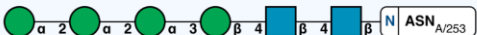

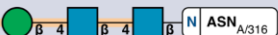

2D Glycan Descriptions		
Chain	SNFG	WURCS
A		ASN_{A/543} 
A		ASN_{A/715} 
B		ASN_{B/543} 
B		ASN_{B/715} 
C		ASN_{A/61} 
D		ASN_{A/253} 
E		ASN_{A/316} 

Figure 3 - Table of two-dimensional Symbol Nomenclature for Glycan (SNFG) visualisations, which can allow for easy oversight of the validity of a particular glycan. Sugars which have issues identified by *Privateer* are highlighted in yellow, and linkages which have unusual torsion angles are also highlighted in yellow. The WURCS codes for each glycan are also available to copy to the clipboard.

The final section of the validation report includes all of the validation metrics calculated by *Privateer* and, most importantly, the diagnostic provided by *Privateer*. A 'yes' diagnostic indicates the conformation is correct for the glycosylation type (e.g. ⁴C₁ for GlcNAc in an N-

glycan, 1C_4 for mannose in a C-glycan), has the correct anomer, and has an acceptable fit to density. This diagnostic indicates that the sugar is valid, whereas, a diagnostic of 'check' indicates that *Privateer* has detected a potential inconsistency affecting ring conformation, which requires manual inspection. Finally, a 'no' diagnostic indicates that the sugar needs a more detailed manual inspection to correct any conformational issues, anomeric issues, or fitting issues.

Detailed monosaccharide validation data

Sugar ID	Q	Phi	Theta	RSCC	B Factor	Detected Type	mFo	Type	Diagnostic
NAG-A-1601	0.54	285.52	7.40	0.65	59.97	beta-D-aldopyranose	0.51	n-glycan	yes
NAG-A-1801	0.55	294.49	6.92	0.68	55.48	beta-D-aldopyranose	0.39	n-glycan	yes
NAG-B-1601	0.54	262.44	19.22	0.58	62.32	beta-D-aldopyranose	0.59	n-glycan	yes
NAG-B-1801	0.56	313.27	2.19	0.66	61.80	beta-D-aldopyranose	0.13	n-glycan	yes
NAG-C-1	0.55	9.63	5.07	0.95	19.84	beta-D-aldopyranose	0.45	n-glycan	yes
NAG-C-2	0.57	356.29	5.94	0.94	22.96	beta-D-aldopyranose	0.78	n-glycan	yes
BMA-C-3	0.58	231.57	2.75	0.92	33.47	beta-D-aldopyranose	0.80	n-glycan	yes
MAN-C-4	0.56	304.07	3.21	0.88	40.66	alpha-D-aldopyranose	0.47	n-glycan	yes
MAN-C-5	0.57	133.82	2.64	0.77	52.16	alpha-D-aldopyranose	0.61	n-glycan	yes
MAN-C-6	0.57	329.17	1.39	0.71	51.35	alpha-D-aldopyranose	0.45	n-glycan	yes
MAN-C-7	0.57	358.97	3.78	0.73	58.94	alpha-D-aldopyranose	0.70	n-glycan	yes
NAG-D-1	0.55	303.45	6.45	0.92	22.37	beta-D-aldopyranose	0.42	n-glycan	yes
NAG-D-2	0.58	48.06	10.80	0.91	25.77	beta-D-aldopyranose	0.37	n-glycan	yes
BMA-D-3	0.57	22.33	5.61	0.89	32.27	beta-D-aldopyranose	0.51	n-glycan	yes
MAN-D-4	0.57	225.80	1.95	0.91	33.69	alpha-D-aldopyranose	0.70	n-glycan	yes
MAN-D-5	0.57	238.50	2.86	0.85	47.83	alpha-D-aldopyranose	0.45	n-glycan	yes
MAN-D-6	0.57	327.47	2.18	0.64	49.42	alpha-D-aldopyranose	0.16	n-glycan	yes
NAG-E-1	0.55	31.25	4.10	0.93	19.18	beta-D-aldopyranose	0.11	n-glycan	yes
NAG-E-2	0.55	339.00	4.19	0.89	31.46	beta-D-aldopyranose	0.14	n-glycan	yes
BMA-E-3	0.57	294.67	0.96	0.68	65.45	beta-D-aldopyranose	0.16	n-glycan	yes

Figure 4 - Table of validation data for each sugar residue within deposited PDB mode available in the visual validation report. The table contains all validation metrics calculated by *Privateer* including the Cremer-Pople puckering parameters, correlation coefficient and importantly *Privateer* diagnostic, which can be used to identify the validity of each sugar.

Trends in the *Privateer* Database

Using the *Privateer* database global statistics throughout the PDB and PDB-REDO can be calculated with ease. Observing deposition trends in the PDB is often interesting as it can provide insight into the kinds of structures that are experimentally obtainable over time. With the *Privateer* database, trends in glycosylation deposition in the PDB over time can be measured, as shown in Figure 5. Importantly, as the *Privateer* database is completely

recompiled every week, these trends remain consistent with the PDB. To allow for easy and up-to-date observation for anyone, compiled statistics are freely available alongside the *Privateer Web App* - <https://privateer.york.ac.uk/statistics>.

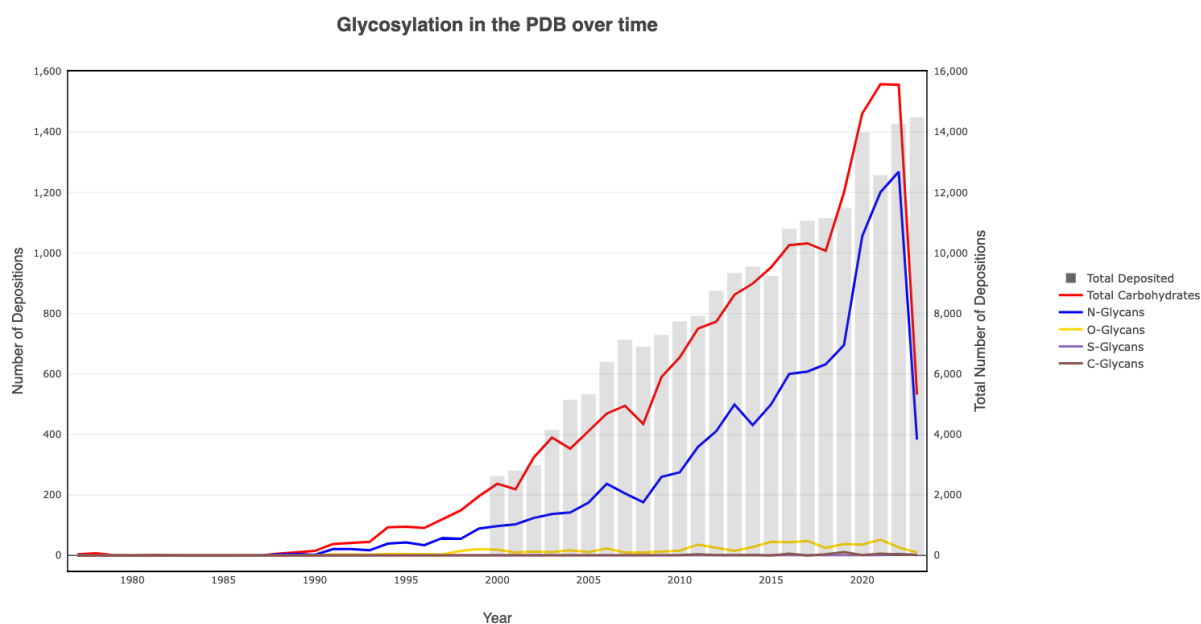


Figure 5 - Plot showing trends in deposition in the PDB over time from 1975 to the present. Blue bars show the total deposited models into the PDB for all structural determination methods. Lines show glycosylation in the PDB over time, split into *N*-glycans, *O*-glycans, *S*-glycans and *C*-glycans.

While simply looking at glycosylation over time using the *Privateer* database is possible, the validation reports calculated by *Privateer* contain a whole host of other interesting pieces of information. In an analogous way to looking at glycosylation over time, the type and validity of carbohydrates in the PDB can also be observed over time. The statistics page available alongside the *Privateer Web App* contains up-to-date plots of validation and conformational errors over time and resolution.

Conclusions

In conclusion, the new *Privateer* database encompasses the carbohydrate validation capabilities of *Privateer* in an easily accessible pre-prepared form. The database contains all validation metrics calculated by *Privateer* as well as highlighted SNFG diagrams in SVG format for easy third-party web use. Statistics are automatically computed weekly and are available alongside the database both on GitHub and the interactive web page.

ORCID® iDs

Jordan S. Dialpuri - <https://orcid.org/0000-0002-6205-2661>
Haroldas Bagdonas - <https://orcid.org/0000-0001-5028-4847>
Lucy C. Schofield - <https://orcid.org/0009-0001-2069-878X>
Phuong Thao Pham - <https://orcid.org/0000-0002-6205-1298>
Lou A. Holland - <https://orcid.org/0000-0002-3867-1833>
Jon Agirre - <https://orcid.org/0000-0002-1086-0253>

Availability and reproducibility

All source code is publicly available on GitHub (<https://github.com/glycojones/privateer> and <https://github.com/Dialpuri/PrivateerDatabase>). The *Privateer* database is available at <https://privateer.york.ac.uk/database> and calculated statistics are available at <https://privateer.york.ac.uk/statistics>. Both pages will remain automatically updated with respect to the source code on GitHub.

Acknowledgements

We are grateful to the University of York IT Services and Darren Miller in particular for accommodating our needs and offering timely and excellent technical support. Lastly, we should like to acknowledge and highlight the contributions of Thomas Lütteke, Martin Frank and the late Willy von der Lieth, pioneers of carbohydrate structure validation, whose research informed some of the methods showcased in the *Privateer* database.

Funding

Jordan Dialpuri is funded by the Biotechnology and Biological Sciences Research Council (BBSRC; grant No. BB/T0072221). Haroldas Bagdonas is funded by The Royal Society (grant No. RGF/R1/181006). Lucy Schofield is funded by STFC/CCP4 PhD studentship agreement 4462290 (York) / S2 2024 012 (STFC) awarded to Jon Agirre. Phuong Thao Pham is a self-funded PhD student. Lou Holland is funded by The Royal Society (URF\R\221006). Jon Agirre is a Royal Society University Research Fellow (awards UF160039 and URF\R\221006).

References

- (1) Brockhausen, I.; Schutzbach, J.; Kuhns, W. *Acta Anat.* **1998**, *161*, 36–78
- (2) Calvelo, M.; Males, A.; Alteen, M. G.; Willems, L. I.; Vocado, D. J.; Davies, G. J.; Rovira, C. *ACS Catal.* **2023**, *13*, 13672–13678
- (3) Agirre, J.; Davies, G.; Wilson, K.; Cowtan, K. *Nat. Chem. Biol.* **2015**, *11*, 303
- (4) Agirre, J. *Acta Crystallogr. D Struct. Biol.* **2017**, *73*, 171–186
- (5) Atanasova, M.; Bagdonas, H.; Agirre, J. *Curr. Opin. Struct. Biol.* **2020**, *62*, 70–78
- (6) Lütteke, T.; Frank, M.; von der Lieth, C.-W. *Nucleic Acids Res.* **2005**, *33*, D242–D246
- (7) Lütteke, T.; Frank, M.; von der Lieth, C.-W. *Carbohydr. Res.* **2004**, *339*, 1015–1020
- (8) Crispin, M.; Stuart, D. I.; Jones, E. Y. *Nat. Struct. Mol. Biol.* **2007**, *14*, 354; discussion 354–355
- (9) Frank, M.; Lütteke, T.; von der Lieth, C.-W. *Nucleic Acids Res.* **2007**, *35*, 287–290
- (10) von der Lieth, C.-W.; Freire, A. A.; Blank, D.; Campbell, M. P.; Ceroni, A.; Damerell, D. R.; Dell, A.; Dwek, R. A.; Ernst, B.; Fogh, R.; Frank, M.; Geyer, H.; Geyer, R.; Harrison, M. J.; Henrick, K.; Herget, S.; Hull, W. E.; Ionides, J.; Joshi, H. J.; Kamerling, J. P.; LeeFlang, B. R.; Lütteke, T.; Lundborg, M.; Maass, K.; Merry, A.; Ranzinger, R.; Rosen, J.; Royle, L.; Rudd, P. M.; Schloissnig, S.; Stenutz, R.; Vranken, W. F.; Widmalm, G.; Haslam, S. M. *Glycobiology* **2011**, *21*, 493–502

- (11) Lütteke, T.; Bohne-Lang, A.; Loss, A.; Goetz, T.; Frank, M.; von der Lieth, C.-W. *Glycobiology* **2006**, *16*, 71R – 81R
- (12) Toukach, P. V.; Egorova, K. S. *Nucleic Acids Res.* **2016**, *44*, D1229–D1236
- (13) Böhm, M.; Bohne-Lang, A.; Frank, M.; Loss, A.; Rojas-Macias, M. A.; Lütteke, T. *Nucleic Acids Res.* **2019**, *47*, D1195–D1201
- (14) Agirre, J.; Iglesias-Fernández, J.; Rovira, C.; Davies, G. J.; Wilson, K. S.; Cowtan, K. D. *Nat. Struct. Mol. Biol.* **2015**, *22*, 833–834
- (15) Bagdonas, H.; Ungar, D.; Agirre, J. *Beilstein J. Org. Chem.* **2020**, *16*, 2523–2533
- (16) Dialpuri, J. S.; Bagdonas, H.; Atanasova, M.; Schofield, L. C.; Hekkelman, M. L.; Joosten, R. P.; Agirre, J. *Acta Crystallogr D Struct Biol* **2023**, *79*, 462–472
- (17) Emsley, P.; Crispin, M. *Acta Crystallogr D Struct Biol* **2018**, *74*, 256–263
- (18) Atanasova, M.; Nicholls, R. A.; Joosten, R. P.; Agirre, J. *Acta Crystallogr D Struct Biol* **2022**, *78*, 455–465
- (19) Alocci, D.; Mariethoz, J.; Gastaldello, A.; Gasteiger, E.; Karlsson, N. G.; Kolarich, D.; Packer, N. H.; Lisacek, F. *J. Proteome Res.* **2019**, *18*, 664–677
- (20) Fujita, A.; Aoki, N. P.; Shinmachi, D.; Matsubara, M.; Tsuchiya, S.; Shiota, M.; Ono, T.; Yamada, I.; Aoki-Kinoshita, K. F. *Nucleic Acids Res.* **2021**, *49*, D1529–D1533
- (21) Joosten, R. P.; Long, F.; Murshudov, G. N.; Perrakis, A. *IUCrJ* **2014**, *1*, 213–220
- (22) van Beusekom, B.; Lütteke, T.; Joosten, R. P. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2018**, *74*, 463–472
- (23) van Beusekom, B.; Wezel, N.; Hekkelman, M. L.; Perrakis, A.; Emsley, P.; Joosten, R. P. *Acta Crystallogr D Struct Biol* **2019**, *75*, 416–425
- (24) Berman, H.; Henrick, K.; Nakamura, H.; Markley, J. L. *Nucleic Acids Res.* **2007**, *35*, D301–D303
- (25) Dialpuri, J. S.; Bagdonas, H.; Schofield, L. C.; Pham, P. T.; Holland, L.; Bond, P. S.; Sánchez Rodríguez, F.; McNicholas, S. J.; Agirre, J. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2024**, *80*. doi:10.1107/S2053230X24000359
- (26) Matsubara, M.; Aoki-Kinoshita, K. F.; Aoki, N. P.; Yamada, I.; Narimatsu, H. *J. Chem. Inf. Model.* **2017**, *57*, 632–637
- (27) Cremer, D.; Pople, J. A. **2002**. doi:10.1021/ja00839a011
- (28) Neelamegham, S.; Aoki-Kinoshita, K.; Bolton, E.; Frank, M.; Lisacek, F.; Lütteke, T.; O’Boyle, N.; Packer, N. H.; Stanley, P.; Toukach, P.; Varki, A.; Woods, R. J.; SNFG Discussion Group. *Glycobiology* **2019**, *29*, 620–624
- (29) Agirre, J.; Davies, G. J.; Wilson, K. S.; Cowtan, K. D. *Curr. Opin. Struct. Biol.* **2017**, *44*, 39–47