Article:
Holland, Lou, Dialpuri, Jordan S, Schofield, Lucy C et al. (2026) Privateer for UCSF ChimeraX: Carbohydrate structure validation and visualization. Protein science : a publication of the Protein Society. e70716. ISSN: 1469-896X

https://doi.org/10.1002/pro.70716

Privateer for UCSF ChimeraX: Carbohydrate Structure Validation and Visualisation

Lou Holland[1], Jordan S. Dialpuri[2], Lucy C. Schofield[1], Phuong Thao Pham[1], Shuai Wang[1], William Cook[1], Tristan I. Croll[3] and Jon Agirre[1,*]

[1] York Structural Biology Laboratory, Department of Chemistry, University of York, York YO10 5DD, UK.
[2]MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, UK.
[3]Institute of Medicine, Altos Labs Inc., Cambridge, CB21 6GP, UK.
[*]Correspondence: jon.agirre@york.ac.uk

Abstract

Modelling 3D carbohydrate structures, either as ligands or covalently bonded to proteins, still presents a number of challenges. As such, carbohydrate structure validation is a key step in the modelling process. The *Privateer* software package has long provided methods to aid in this validation, considering unlikely ring conformations of the sugars, stereochemistry, comparison to known glycans, and unlikely linkage torsions. However, uptake of these tools depends on the pipeline used for structure solution. It is therefore important to make glycan validation available in as many model building programs as possible. To this end, we have created a *Privateer* bundle for ChimeraX. This bundle provides intuitive access to key validation metrics for carbohydrates present in a structure model within the molecular visualisation program while working on your structure solution, as well as providing the ability to visualise monosaccharides in the structure using 3D glycan symbols to aid in understanding of interactions. The *Privateer* bundle for ChimeraX can be installed from ChimeraX's toolshed: https://cxtoolshed.rbvi.ucsf.edu/apps/chimeraxprivateer.

Keywords

Carbohydrates, validation, polysaccharides, oligosaccharides, glycans, ligands, cryo-EM, X-ray crystallography

Introduction

Protein glycosylation is a type of protein modification where an oligosaccharide moiety is covalently linked to an amino acid. The most common type of protein glycosylation is *N*-glycosylation, where the glycosidic link is formed to the side-chain nitrogen atom of an Asparagine residue within the consensus sequence Asn–X–Ser/Thr (where X is any amino acid except proline). Next most common is *O*-glycosylation, with *C*-glycans, and *S*-glycans being far less frequent (Schofield et al. 2024; Varki et al. 2022). These glycans can play a key role in determining protein structure and function (Schofield et al. 2024).

The number of structures deposited in the Protein Data Bank (PDB) containing carbohydrates, such as *N*-glycans, increases every year (Dialpuri, Bagdonas, Schofield, Pham, Holland, and Agirre 2024). However, despite their prevalence in the PDB, the accurate modelling and refinement of these carbohydrates still present a number of challenges. Their variety of possible ring conformations, branched structure, and anomeric

forms introduce additional complexity to the modelling process. Their flexible nature, prone to heterogeneity, can often result in a decrease in the quality of experimental density resolution. This can lead to distortion of ring conformations during the refinement process, as well as glycosidic linkages becoming skewed (Agirre et al. 2015; Crispin et al. 2007; Atanasova et al. 2020).

As such, glycan validation is an important step in the model building pipeline. A number of tools have been developed over the years to aid in this process, including online databases (Alocci et al. 2019; Fujita et al. 2021; Frank et al. 2007; Böhm et al. 2019), analysis and validation software (Agirre et al. 2015; Dialpuri et al. 2023), restraint dictionaries (Atanasova et al. 2022), and model building tools (Emsley and Crispin 2018). The *Privateer* software is one such tool. Initially introduced to address common issues with high-energy ring-conformations (Agirre et al. 2015), later additions included the ability to check glycans against those found in glycomics databases (Bagdonas et al. 2020) and validation of glycosidic linkage stereochemistry and torsion angles (Dialpuri et al. 2023). *Privateer* is currently available through the CCP4 software suite (Agirre et al. 2023), as a web app (Dialpuri, Bagdonas, Schofield, Pham, Holland, Bond, et al. 2024), and integrated into the molecular graphics web app *Moorhen (*https://moorhen.org). *Privateer* validation reports for structures deposited in the PDB can also be accessed through the *Privateer* Database (Dialpuri, Bagdonas, Schofield, Pham, Holland, and Agirre 2024).

However, depending on the software packages that structural biologists use, carbohydrate validation may not be an easily integrated step in the model building process. This can lead to more modelling errors slipping through into the PDB. It is therefore important to make carbohydrate validation software as easy to access as possible for all, including integrating it into other software suites. To this end, we have developed a version of *Privateer* within ChimeraX (Meng et al. 2023).

ChimeraX is a molecular visualisation program including many tools for structure building, analysis, and validation. It is favoured by many who work with cryo electron microscopy (cryo-EM) as it contains many useful tools relating to processing and using cryo-EM maps. It also provides easy routes to collaboration and software integration through its “Toolshed”, allowing developers to build their own bundles which users can then install within ChimeraX. However, at present, there are no tools specific to carbohydrate validation. The *Privateer* bundle for ChimeraX aims to fill this gap.

Results and Discussion

The *Privateer* bundle for ChimeraX was designed to integrate into ChimeraX’s existing user interface, providing new methods for carbohydrate analysis, validation, and visualisation. It can be installed from within ChimeraX by clicking “More tools…” in the “Tools” menu, or it can be downloaded from the ChimeraX toolshed and installed with the “toolshed install” command.

Once installed, the *Privateer* tool can be opened from the “Tools” menu under “Validation”. Doing so opens a tool window, from which the user can run the various functions of *Privateer* for ChimeraX. This tool window can be seen in **Fig. 1**. From this tool window, the user can perform a number of functions, such as opening a user guide which explains how to use

Privateer within ChimeraX, running *Privateer* to produce an interactive validation report (with the option for this to update automatically as the model is edited), and drawing the 3D equivalent of 2D Symbol Nomenclature For Glycans (SNFGs) as 3D glycan symbols on the model in ChimeraX's view (with the option for these symbols to update automatically as the model is edited, as well as with the option for the depth of the symbols to change depending on how zoomed out the user is). An example of what these 3D glycan symbols look like when fully zoomed out (both with and without the variable depth) is shown in **Fig. 2**. Examples of what these 3D glycan symbols look like when zoomed in can be seen in **Figs. 6**, **8**, **9**, and **10**. The 3D shapes were chosen to be analogous to the 2D ones (rather than fully symmetrical such as cubes or spheres) so that the plane of the sugar could be clearly observed from the symbols.

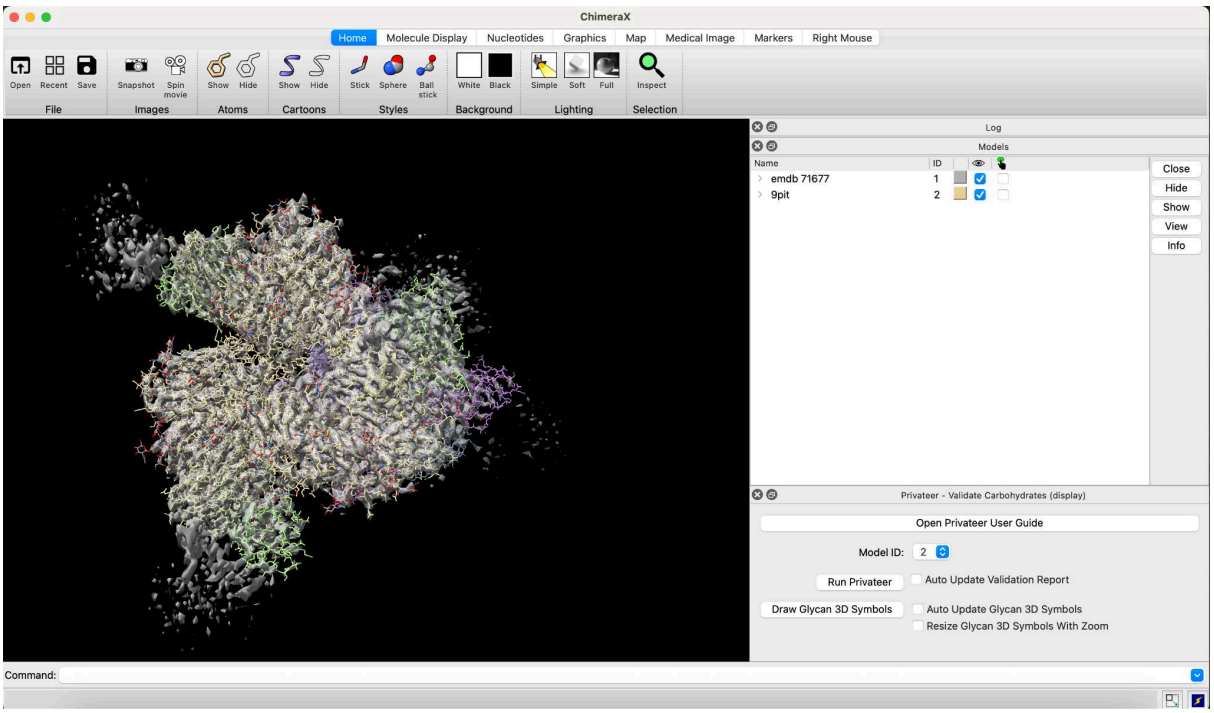


Figure 1. The *Privateer* tool panel, open within ChimeraX. The large button at the top opens a user guide which describes the functions of each of the buttons and tick boxes as well as giving an overview of how *Privateer* works. The drop down menu allows the user to select which model they wish to run *Privateer* for or draw 3D glycan symbols for, according to its model ID within ChimeraX. The "Run Privateer" button runs *Privateer* on the selected structure and produces an interactive validation report for it. Ticking the box "Auto Update Validation Report" will make this validation report update as the structure is edited. Unticking this box means that the validation report will remain unchanged with any edits. The "Draw Glycan 3D Symbols" button will draw a 3D representation of the 2D SNFGs on the structure. Ticking the box "Auto Update Glycan 3D Symbols will make these 3D representations update as the structure is edited. Unticking this box means that the 3D symbols will remain unchanged with any edits. Ticking the box "Resize Glycan 3D Symbols With Zoom" will alter the depth of the 3D symbols depending on how zoomed in or zoomed out the user is, allowing them to remain clearly visible when zoomed out fully, but allowing clear visualisation of the plain of the sugar when zoomed in. Unticking this box will result in the 3D symbols maintaining a constant depth no matter how zoomed in or zoomed out the user is.

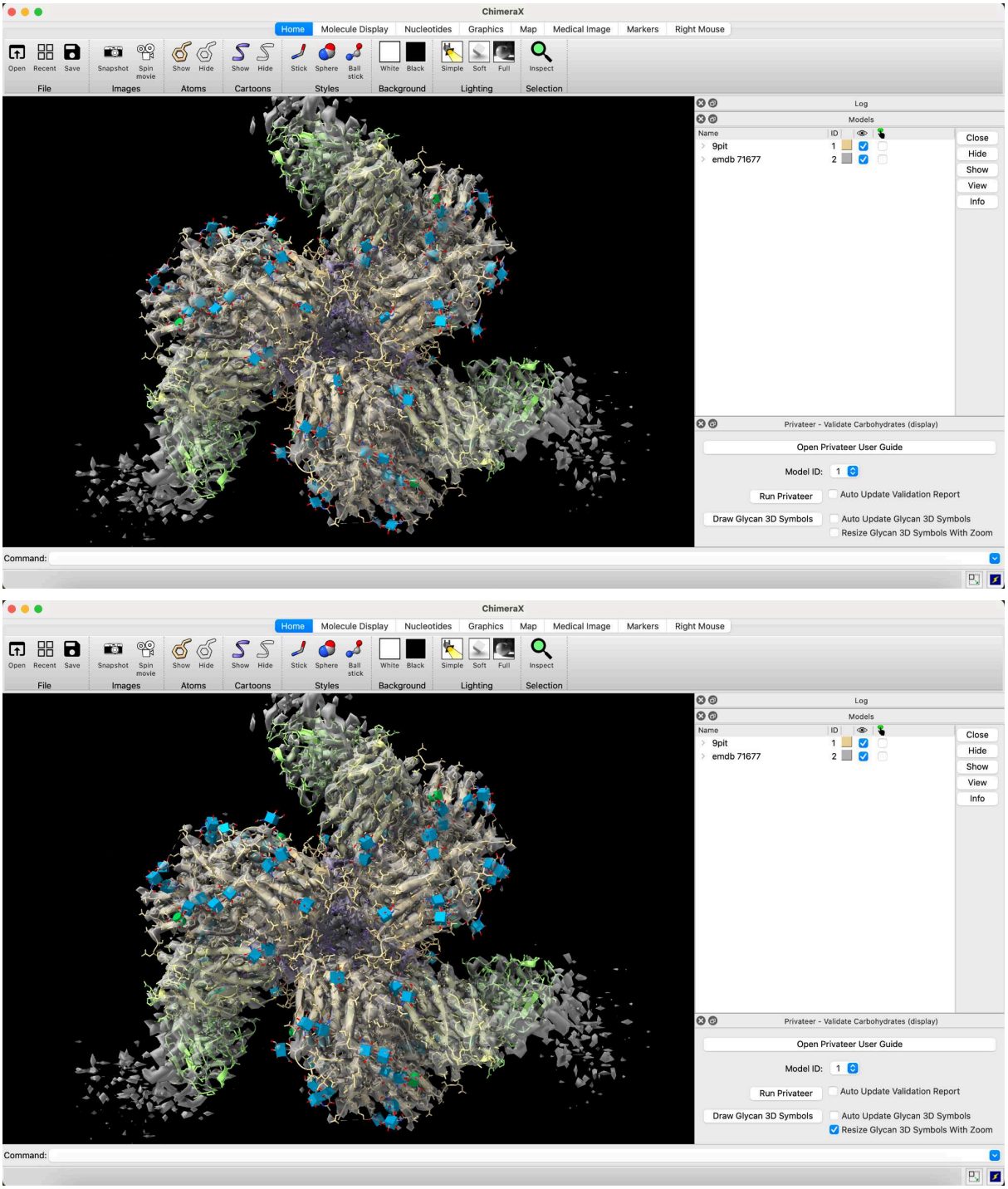


Figure 2. The 3D glycan symbols that *Privateer* draws in ChimeraX for structure with EMDB entry number 71677(Bader et al. 2026). The colours and shapes match those used in SNFGs. Top panel shows what these look like when fully zoomed out without the option "Resize Glycan 3D Symbols with Zoom" ticked, while the bottom panel shows what it looks like with this option ticked, where the depth of the 3D glycan symbols is increased in proportion to how zoomed out the ChimeraX view is in order to help the symbols stand out more.

The validation report produced by *Privateer* in ChimeraX displays in a child tool window. The format of the report matches the typical *Privateer* report used within the CCP4 suite and by the *Privateer* Web App for consistency (Dialpuri, Bagdonas, Schofield, Pham, Holland, Bond, et al. 2024). An example report for an envelope glycoprotein in complex with a transmembrane protein (EMDB entry 71677, (Bader et al. 2026)) is shown in **Fig. 3**. If the

user runs *Privateer* multiple times to produce multiple validation reports, it is only the most recent validation report for a given model that is "active" so will update with any changes to the model if that option is selected. Older validation reports for the same model become "inactive" and will no longer update regardless of the status of the tickbox, allowing the user to maintain a copy of the original validation report while another auto-updates for comparison after any changes to the structure. Which validations reports are inactive is indicated in the tab title within the validation report window.

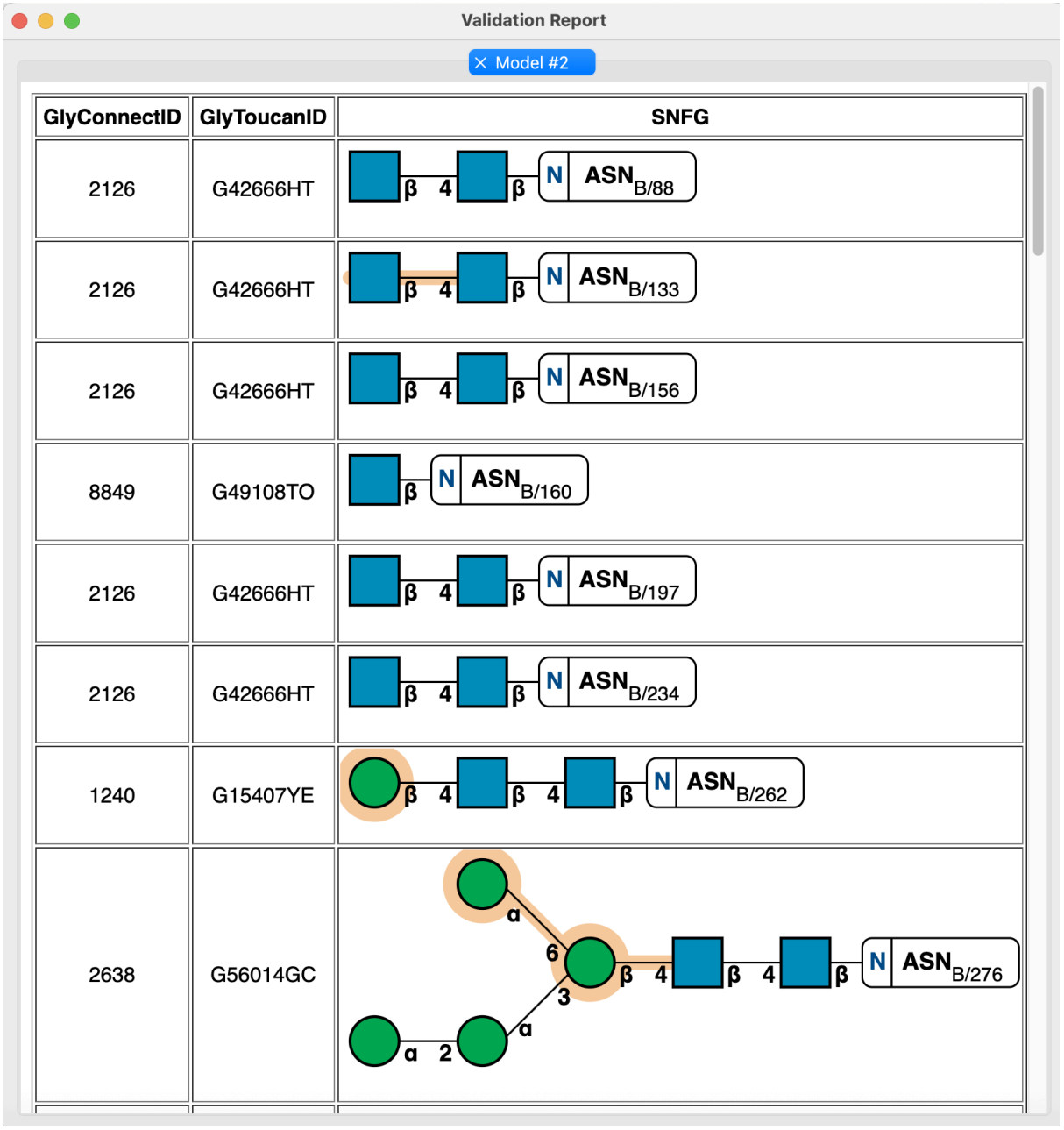


Figure 3. A section of a validation report as produced by *Privateer* within ChimeraX for structure with EMDB entry number 71677 (Bader et al. 2026). The first two columns contain glycan IDs as identified in online databases GlyConnect and GlyToucan. If the glycan is not found in either of these databases, this is stated in place of the ID, potentially indicating an issue. The SNFG gives a visual representation of the glycans present in the structure. The chain and residue ID of the glycan roots are provided for easy location, and linkage information is annotated on. Highlighted portions indicate potential issues. If a sugar is highlighted, this could be due to an anomer mismatch or a high energy conformation. A highlighted linkage indicates that the torsion angles of that linkage are outliers compared to *Privateer*'s torsion database.

The first two columns of the validation report, as shown in **Fig. 3**, contain the GlyConnect ID (Alocci et al. 2019) and the GlyToucan ID (Fujita et al. 2021) of the glycans identified in the structure, if those glycans exist in either of those glycomics databases. If the glycans are not found in the glycomics databases, the phrase “Not Found” will appear in place of the ID. The third column contains interactive SNFG representations of the glycans present in the structure. The chain ID and residue ID of the glycan root are annotated on the SNFG, as is linkage classification. Highlighted elements indicate potential issues. More information can be gleaned by hovering the cursor over different elements of the SNFG. Hovering over a sugar will display a tooltip which tells the user the sugar type, conformation, and B-factor, alongside noting any potential issues. Hovering over a link displays a tooltip which tells the user the linkage type and torsion angles. Hovering over the highlighted portion of a link displays a tooltip with the linkage Z-score, indicating how much of an outlier it is compared to the data in the *Privateer* torsion database (Dialpuri et al. 2023). In addition to simply hovering over elements in the SNFG, the user can also click on sugars to centre ChimeraX’s view on that sugar, or on a linkage to display a plot of the torsion angles of that linkage overlaid with data for that linkage from *Privateer*’s torsion database, if there are enough data points. An example of such a torsion plot tool window is shown in **Fig. 4**. This can be used to help see just how much of an outlier a particular linkage is or isn’t, in conjunction with examining the structure itself to determine if there are any interactions that might be causing the linkage to be an outlier, rather than a modelling or refinement error. For more precise information, the user can hover their cursor over any point on the plot to see the exact torsion angles at that point.

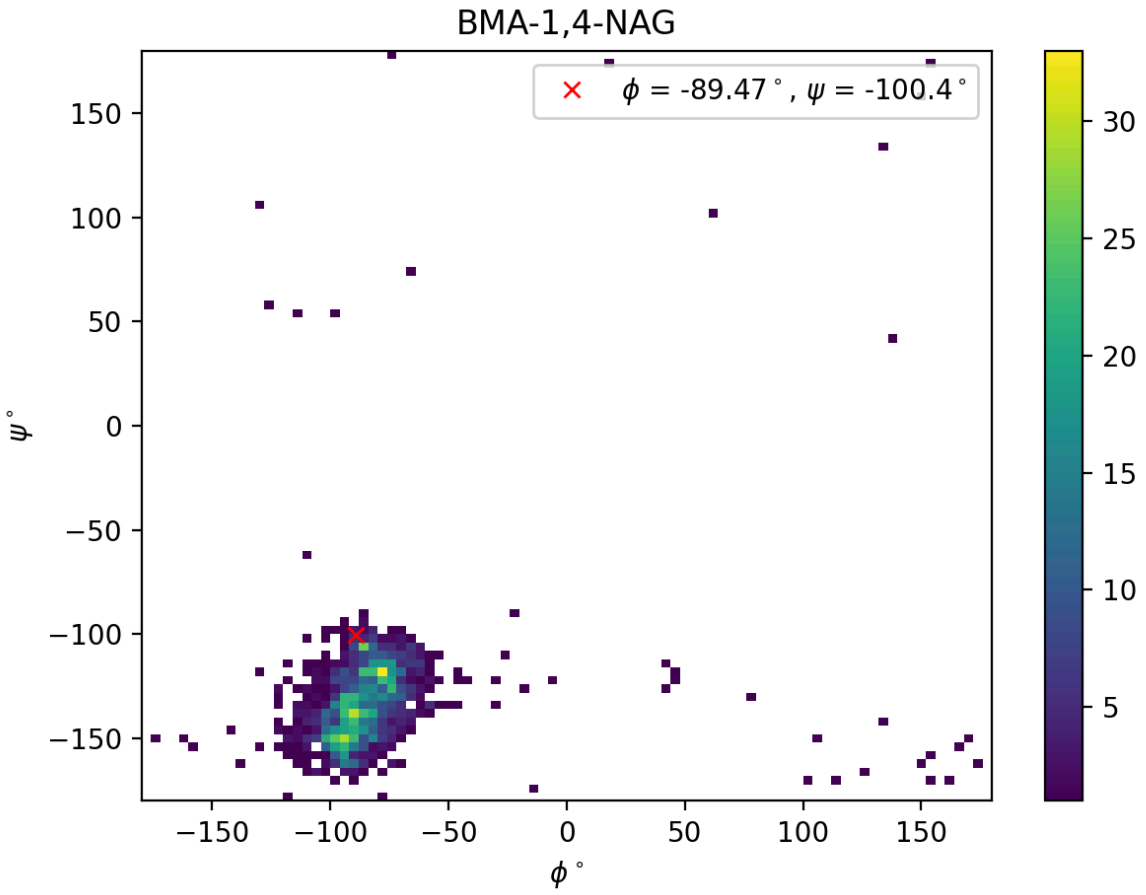


Figure 4. A torsion plot as produced by *Privateer* within ChimeraX for the BMA-1,4-NAG linkage in the B/276/Asn

linked *N*-glycan in the structure with EMDB entry number 71677 (Bader et al. 2026). BMA and NAG are the relevant three letter codes for the sugars involved in the linkage (beta-D-mannose and beta-N-acetylglucosamine). The numbers 1 and 4 refer to the atom numbers within the sugar ring which form the linkage. The torsion angles of the linkage in the structure are shown by a red cross, the specific values of which are given in the legend. Data from the *Privateer* torsion database is displayed as a heat map to indicate frequency.

Case Study: Validating Carbohydrates in a Structure

In order to demonstrate a possible use case of the *Privateer* bundle for ChimeraX, we present here an example using an envelope glycoprotein in complex with a transmembrane protein, the structure of which was determined using cryo-EM with a resolution of 3.1 Å. The EMDB entry number of this structure is 71677 (Bader et al. 2026), and the PDB accession code is 9PIT (Bader et al. 2026).

The structure can be imported into ChimeraX using the “Fetch by ID” function after selecting “EMDB & fit PDBs” as the database to fetch from. Once this is done, clicking the “Run Privateer” button in the *Privateer* tool produces a validation report for the structure, summarising all carbohydrates present within the structure. A section of this validation report is already shown in **Fig. 3**.

Scrolling through the validation report, it is easy to identify any potential issues, as they are highlighted in orange in the SNFG. It should be noted that an element highlighted in orange is not necessarily wrong, but does warrant further inspection. For example, nearby interactions could legitimately skew a linkage, resulting in it becoming an outlier from the dataset in the *Privateer* torsion database. However, sugars in high-energy conformations are highly unlikely.

This structure contains fifty-one glycans. For the purposes of this case study, we will focus on the three-sugar *N*-glycan linked to B/262/Asn and the five-sugar *N*-glycan linked to B/276/Asn (the two bottom rows of the section of the validation report in **Fig. 3**). The full validation report for the structure is shown in the supplementary information. Hovering the mouse over highlighted elements in the SNFGs for these two glycans provides further insight into potential issues, as demonstrated in **Fig. 5**.

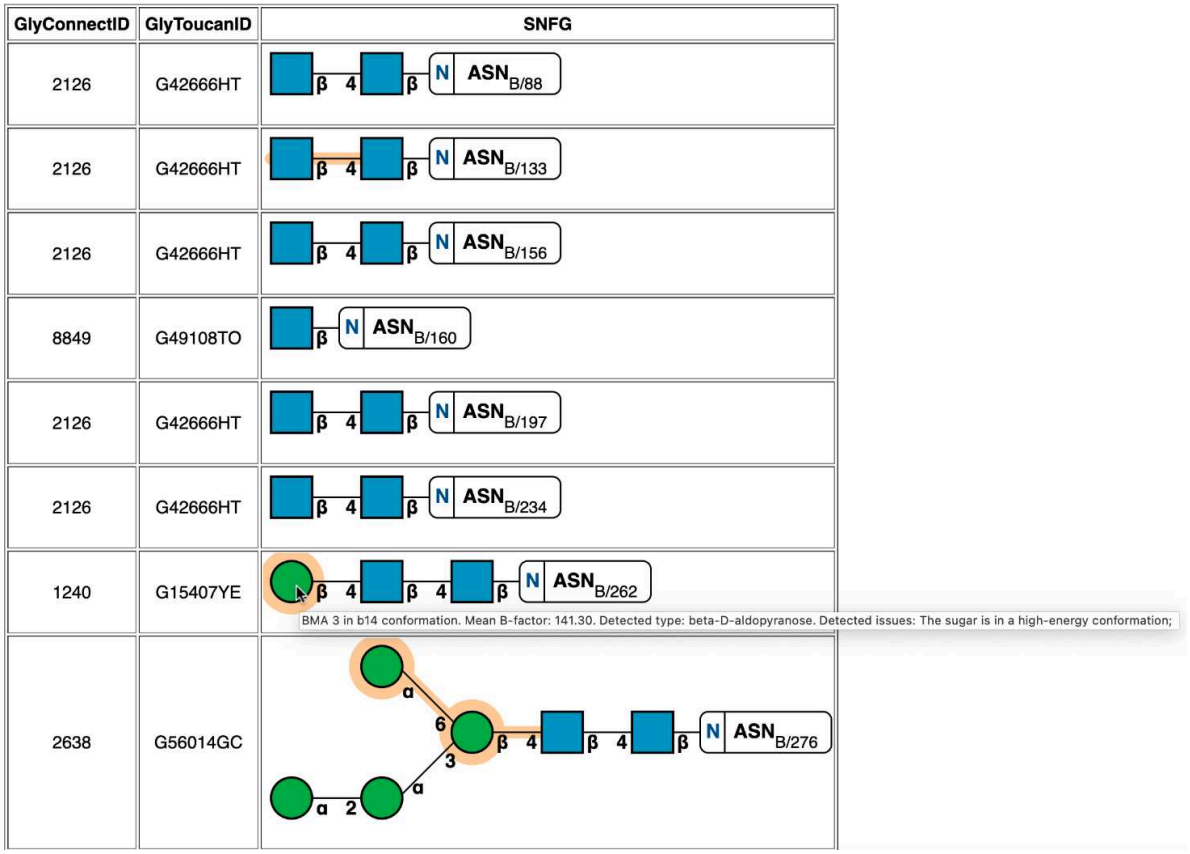


Figure 5. A section of the validation report as produced by *Privateer* within ChimeraX for the structure with EMDB entry number 71677 (Bader et al. 2026), focusing on the *N*-glycan linked to B/262/Asn. Hovering the mouse over one of the sugars highlighted as having a potential problem is displaying a tooltip with further information. This tooltip states that the sugar is in a high energy conformation.

The tooltips displayed by hovering the mouse over these highlighted elements reveals a potential conformation issue in the terminal sugar of the B/262/Asn linked *N*-glycan, where the *β*1,4-linked mannose is in a boat ($B_{1,4}$) conformation. For additional information on map model fit, the ChimeraX Q-score tool can be utilised, yielding the result that this terminal sugar has a Q-score of 0.597 (Pintilie et al. 2020). Clicking on this terminal sugar centres ChimeraX's view on it, allowing further inspection of this glycan, alongside the cryo-EM map map in ChimeraX. In doing so it is possible to see that there is potentially additional unmodelled density, suggesting that this glycan could be extended by the presence of an additional sugar. This is shown in **Fig. 6**, where the glycan in question is shown in yellow. The terminal sugar can be seen to be in a high energy conformation, and a large blob of unmodelled potential extends beyond it, with the map level set at 0.28.

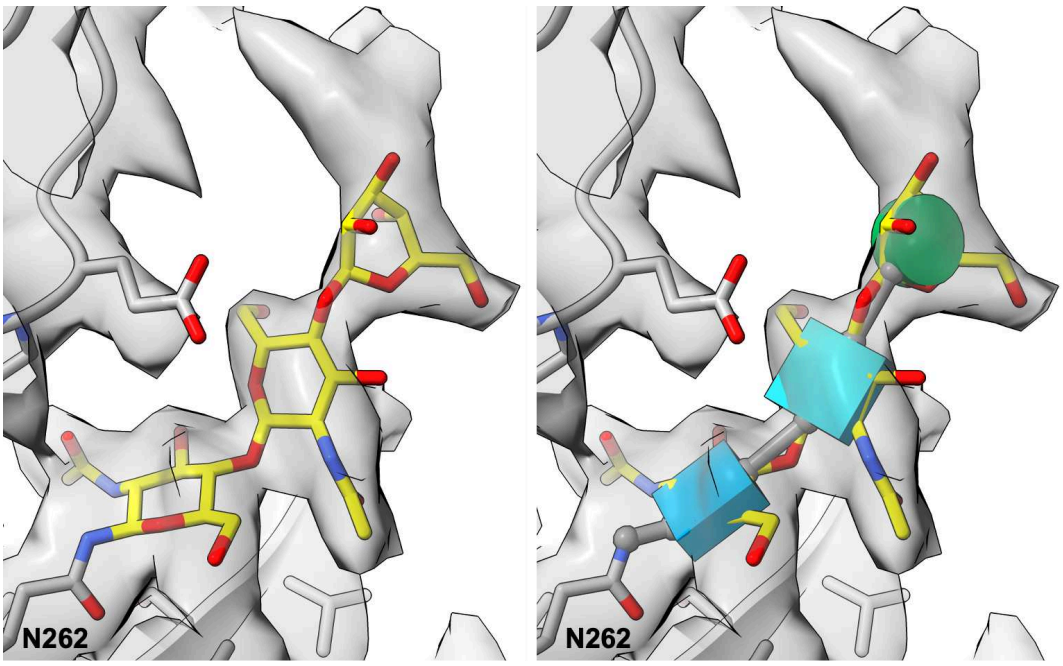


Figure 6. The *N*-glycan linked to B/262/Asn in structure with EMDB entry number 71677(Bader et al. 2026). The map level is set to 0.28 and is shown as a translucent surface; any blobs < 8 Å were hidden. The glycan molecules are depicted in yellow. The left panel shows just the atoms and bonds while the right panel also shows *Privateer*'s 3D glycan symbols for easy comparison to the SNFG in the validation report. The terminal sugar with a potential conformation issue can be seen to be sticking out of the cryo-EM map a little, and there is additional unmodelled potential continuously attached to that for the glycan.

For the B/276/Asn linked *N*-glycan, there are two highlighted linkages and sugars, as can be seen in **Figs. 3** and **5**. The *β*1,4-linked mannose is in a half-chair ($^{1}H_{2}$), and the *α*1,6-linked mannose is in an envelope (^{3}E), both of which are unlikely high-energy conformations. The Q-scores of these two sugars (as calculated using ChimeraX's Q-score tool) are 0.597 and 0.508 respectively. Hovering over the *β*1,4-linkage displays a tooltip with the linkage Z-score of -1.27, as shown in **Fig 7**, where anything less than -1 is considered an outlier (Dialpuri et al. 2023). Doing the same for the *α*1,6-linkage reveals a Z-score of -1.17. The torsion plot for the BMA-1,4-NAG linkage (where BMA is the three letter code for beta-D-mannose and NAG is the three letter code for beta-N-acetylglucosamine) within this glycan can be viewed by clicking on the linkage, and is shown in **Fig. 4**. From this, it can be seen that the linkage torsion angles are relatively close to the expected ones, though are on the edge of the distribution. Doing the same for the MAN-1,6-BMA linkage (where MAN and BMA are the three letter codes for D-mannose in the alpha and beta anomeric forms respectively) reveals that it is not particularly close to the peak of the distribution for torsion angles in this linkage within the *Privateer* torsion database. Again, we can click on one of the sugars in this glycan in the SNFG to centre ChimeraX's view on that sugar for further inspection alongside the cryo-EM map, as shown in **Fig. 8**. From this, it appears as if the incorrect glycan conformations are skewing the linkages which are also highlighted, so fixing one issue will likely fix the other.

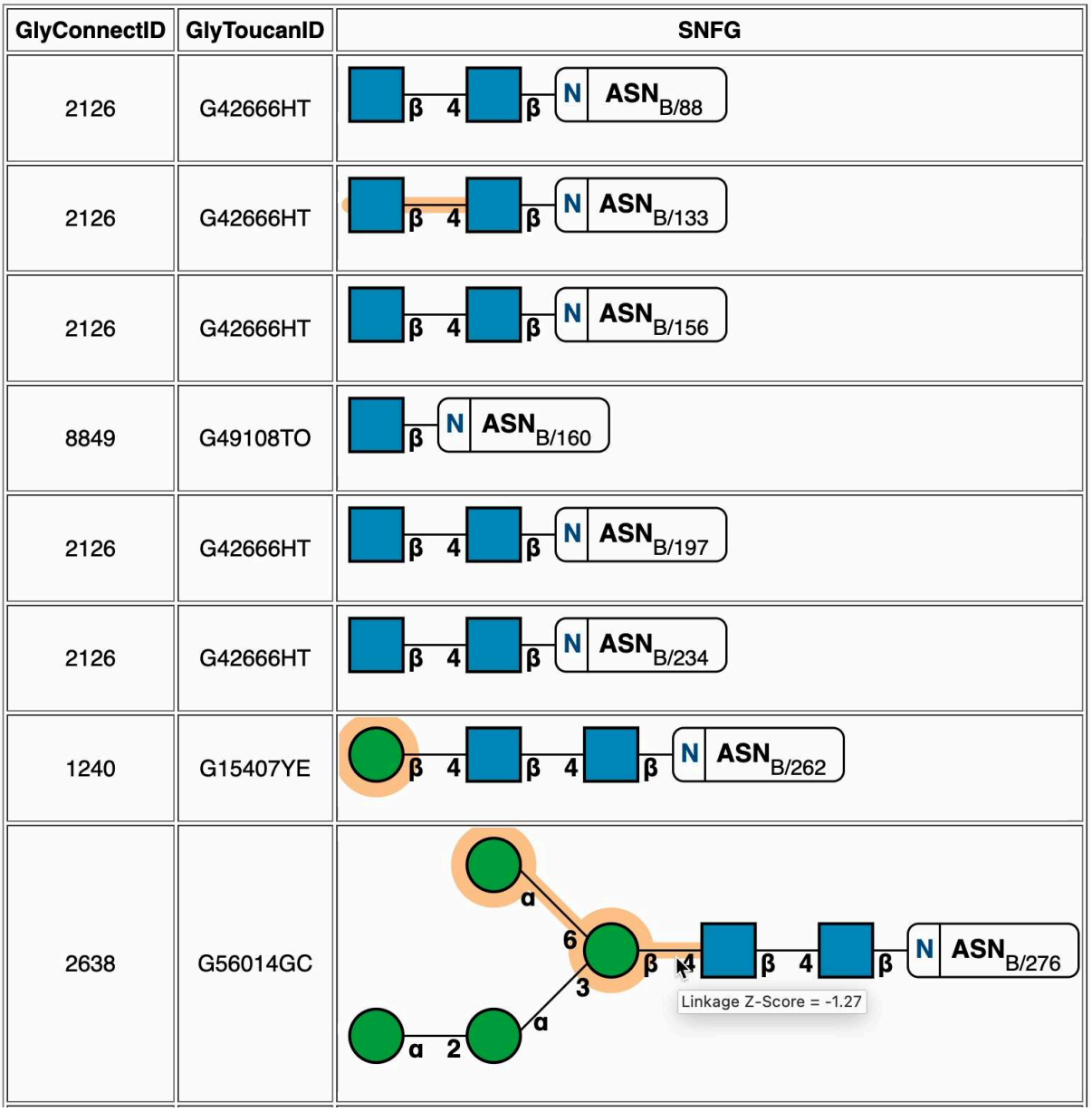


Figure 7. A section of the validation report as produced by *Privateer* within ChimeraX for the structure with EMDB entry number 71677(Bader et al. 2026), focusing on the *N*-glycan linked to B/276/Asn. Hovering the mouse over one of the linkages highlighted as having a potential problem is displaying a tooltip with further information. This tooltip states the Z-score of the torsion angles for that linkage as a measure of how much of an outlier they are.

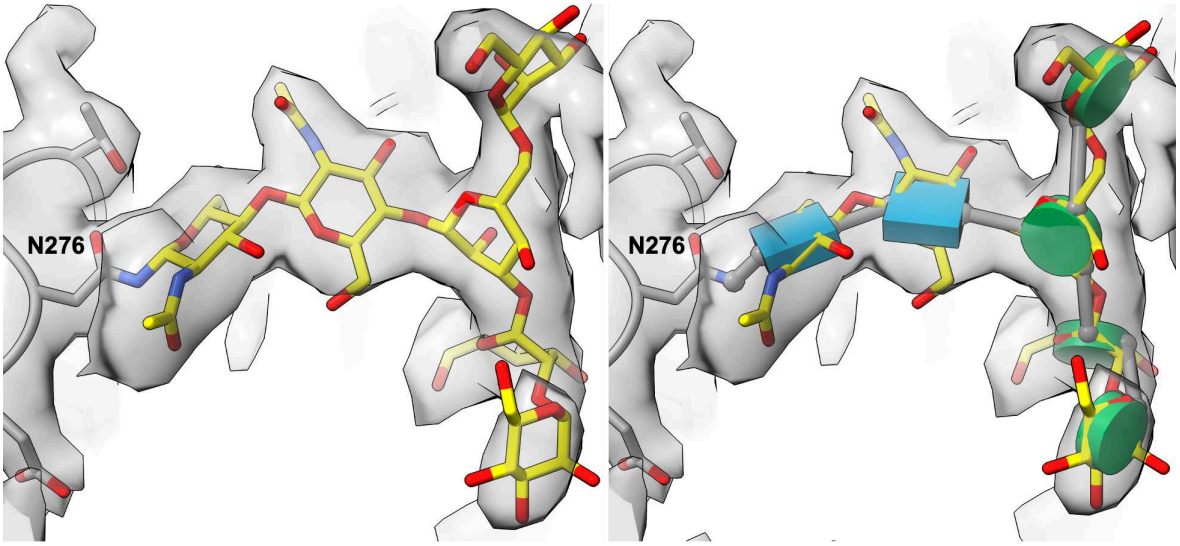


Figure 8. The *N*-glycan linked to B/276/Asn in the structure with EMDB entry number 71677(Bader et al. 2026). The map level is set to 0.28 and is shown as a translucent surface; any blobs < 8 Å were hidden The glycan molecules are depicted in yellow. The left panel shows just the atoms and bonds while the right panel also shows *Privateer*'s 3D glycan symbols for easy comparison to the SNFG in the validation report. The BMA sugar in a high energy conformation is shown in the middle, to the right of the figure, and the terminal MAN sugar in a high energy conformation is shown to the top right. The two linkages with outlying torsion angles come off of the top of the BMA sugar.

In order to improve the modelling of these glycans, we used a combination of ISOLDE within ChimeraX (Croll 2018) and Servalcat (Yamashita et al. 2021) for refinement along with manual refinement in Coot (Emsley and Cowtan 2004) owing to its carbohydrate specific methods. While the *Privateer* tool in ChimeraX provides methods for carbohydrate validation and visualisation, further methods for carbohydrate modelling would be advantageous. To fix conformation issues, the skewed sugars were deleted within Coot (Emsley et al. 2010) and replaced with sugars in the correct anomeric form and conformation – alternatively, this could have been done in ISOLDE, which conveniently restrains pyranose ring conformation to their lowest energy. In the case of the terminal sugar in the *N*-glycan linked to B/262/Asn, this meant deleting the *β*1,4-linked mannose in a boat ($B_{1,4}$) and replacing it with a *β*1,4-linked mannose in a chair ($^4C_1$). An additional *α*1,3-linked mannose residue in a chair ($^4C_1$) was then added to the end of this glycan to fill the unmodelled blob in the map. A similar procedure was followed for fixing the conformations of sugars in the B/276/Asn linked *N*-glycan, which in turn resolved the potential linkage issues. In order to verify that these, and other linkage torsion outliers in the structure, were resolved, we used Coot to overlay the glycans with reference glycans which had the optimal glycosidic linkages. The refinement carried out in Servalcat consisted of twenty cycles of rigid body refinement, including hydrogens.

The *N*-glycan linked to B/262/Asn following this refinement is shown in **Fig. 9**. The additional sugar added to the end of the glycan can be seen to be a terminal *α*1,3-linked mannose residue which fits tolerably within the density. For a more quantitative measure of map model fit, the ChimeraX Q-score tool can be used, revealing a Q-score of 0.357. Prior to these changes, the Q-score of the terminal BMA sugar in a high energy conformation was calculated as 0.597. After fixing the conformation and extending the glycan, the Q-score of this sugar improved marginally to 0.608. It is not always expected that map model fit will improve after fixing conformation issues, as the refinement program is essentially further

constraining the sugar rather than allowing it to skew into whatever unlikely conformation best fits the density.

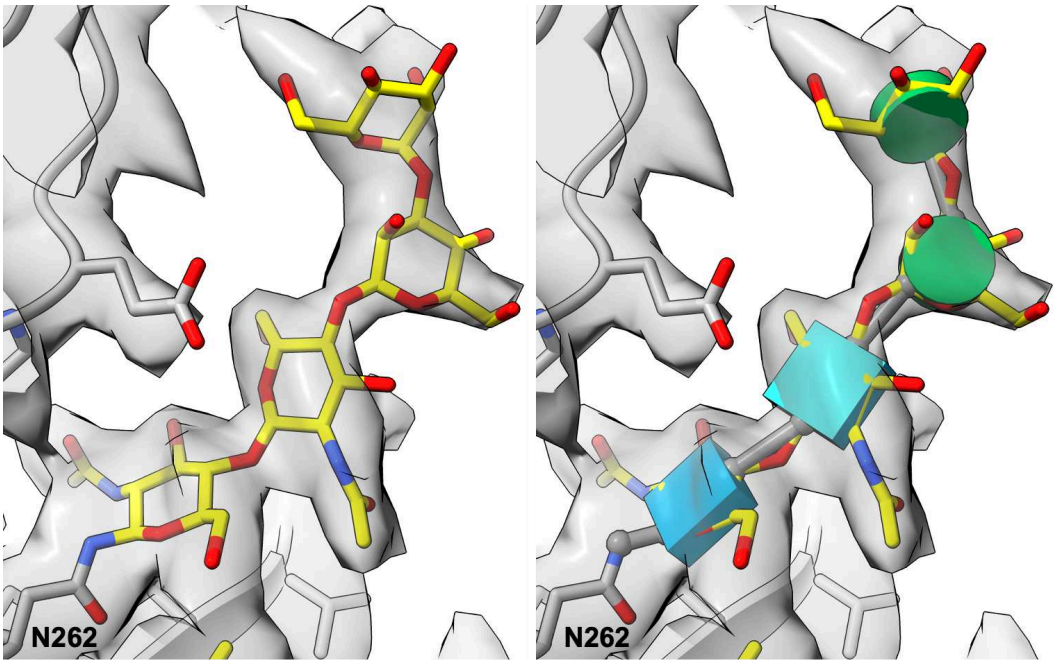


Figure 9. The *N*-glycan linked to B/262/Asn in the structure with EMDB entry number 71677(Bader et al. 2026) after fixing the conformation issue and extending the glycan. The map level is set to 0.28 and is shown as a translucent surface; any blobs < 8 Å were hidden. The glycan molecules are depicted in yellow. The left panel shows just the atoms and bonds while the right panel shows the same view but with *Privateer*'s 3D glycan symbols drawn on for easy comparison to the SNFG in the validation report.

The *N*-glycan linked to B/276/Asn in the structure following our structure improvements is shown in **Fig. 10**. With the conformations of the sugars fixed, the linkage torsions are no longer skewed, and they still appear to fit well within the map. This conclusion is supported by improved Q-scores of the two sugars for which the conformation was fixed from 0.597 to 0.69 for the *β*1,4-linked mannose and from 0.508 to 0.673 for the terminal *α*1,6-linked mannose.

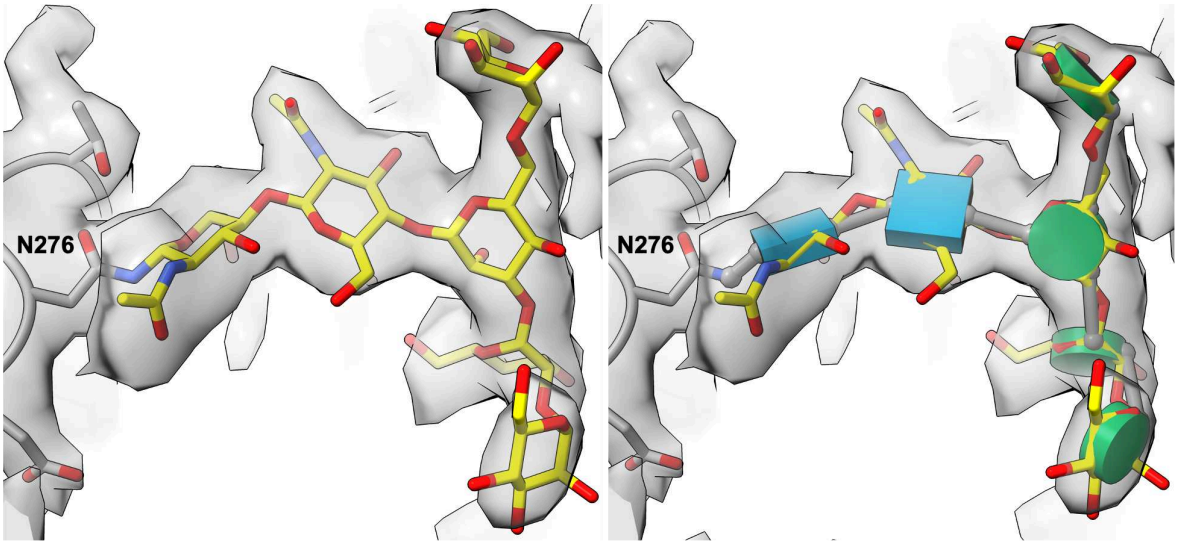


Figure 10. The *N*-glycan linked to B/276/Asn in the structure with EMDB entry number 71677(Bader et al. 2026) after fixing the conformation issue which in turn removed the linkage outliers. The map level is set to 0.28 and is shown as a translucent surface; any blobs < 8 Å were hidden. The glycan molecules are depicted in yellow. The left panel shows just the atoms and bonds while the right panel shows the same view but with *Privateer*'s 3D glycan symbols drawn on for easy comparison to the SNFG in the validation report.

Equivalent changes were made to the other chains in the structure, along with fixing the linkage issues present in the Asn-311, Asn-392, and Asn-488 linked *N*-glycans present in chains B, E, and F. We also extended the Asn-611 and Asn-618 linked *N*-glycans in chains A, C and D as we noticed potential unmodelled residues in the map during our refinement. The same section of the validation report (as shown in **Figs. 3**, **5**, and **7** for the improved structure can be seen in **Fig. 11**. The full validation report for both before and after the changes can be seen in the supplementary information. From this, it can be seen that there are no longer any highlighted sugars, meaning that no sugars are in unlikely high energy conformations, and there are no longer any highlighted linkages, meaning that all linkage torsion angles fall within expected ranges according to *Privateer*'s torsion database.

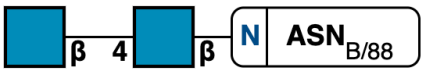
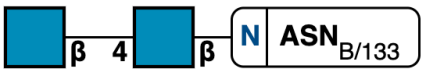
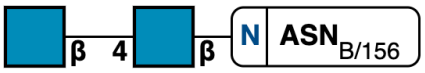
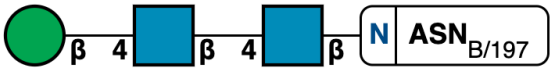
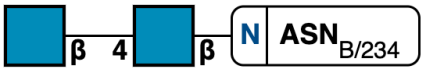
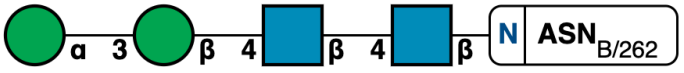
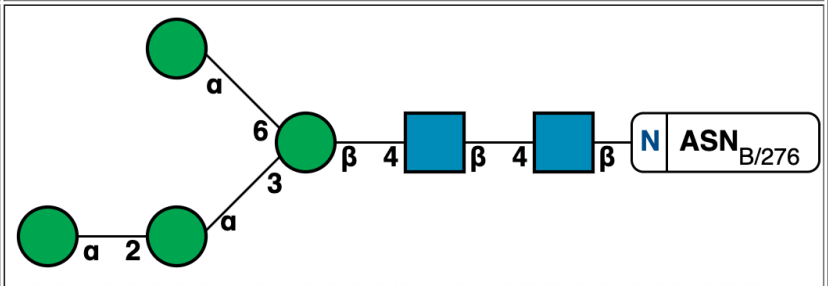
GlyConnectID	GlyToucanID	SNFG
2126	G42666HT	
2126	G42666HT	
2126	G42666HT	
8849	G49108TO	
1240	G15407YE	
2126	G42666HT	
545	G81315DD	
2638	G56014GC	

Figure 11. A validation report as produced by *Privateer* within ChimeraX for structure with EMDB entry number 71677(Bader et al. 2026) following remodelling and refinement of the glycans present in the structure. The first two columns contain glycan IDs as identified in online databases GlyConnect and GlyToucan. If the glycan is not found in either of these databases, this is stated in place of the ID, potentially indicating an issue. The SNFG gives a visual representation of the glycans present in the structure. The chain and residue ID of the glycan roots are provided for easy location, and linkage information is annotated on. There are no longer any highlighted portions indicating potential issues.

In addition to using the *Privateer* tool within ChimeraX to validate the glycans in the structure before and after our changes, we also used ChimeraX's Q-score tool to check the map to model fit of all the sugars in the structure. A plot of these Q-scores before and after our changes can be seen in **Fig. 12**. In this plot, each point represents the average Q-score of a single sugar within the structure, with different symbols representing sugars where the conformation was fixed, new sugars that were added, sugars that were part of glycans where a linkage torsion outlier was fixed, or sugars that were part of glycans where new sugars were added. In cases where a new sugar was added, which was not present in the original structure, the original Q-score was set to zero. The grey dashed line represents the case where a Q-score has not changed. Sugars above this line have an improved map to model fit after our changes, whereas sugars below the line have a worse map to model fit. It is often expected that fixing sugar conformations might worsen metrics such as Q-score, adding additional restraints instead of letting the sugar skew into whatever unlikely conformation it maximises metrics used by refinement. Similarly, adding sugars to a glycan can worsen the Q-score of other sugars present in that glycan where two sugars are fit into a blob of density which previously contained only a single sugar, pushing that single sugar out of the centre of the blob. It can be seen from the plot that the majority of sugars have improved Q-scores after our changes.

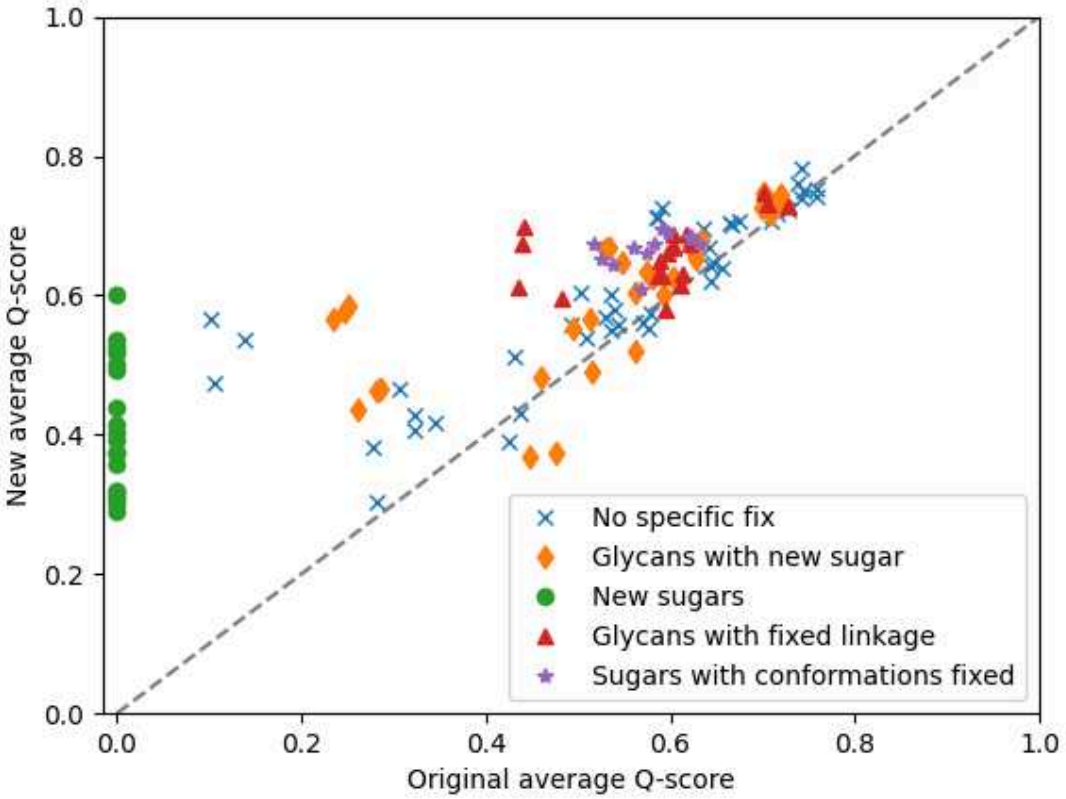


Figure 12. The average Q-score of sugars in the structure EMDB entry number 71677(Bader et al. 2026) both before (x-axis) and after (y-axis) changes based on *Privateer* for ChimeraX's validation report. The Q-scores were calculated using the Q-score tool within ChimeraX, where average Q-score indicates the average over the whole residue as opposed to of a single atom within the residue. Each point represents a sugar within the structure. Where a sugar wasn't present in one of the versions of the structure, its Q-score was set to zero. Different symbols are used to represent sugars that were either specifically fixed themselves (due to conformation issues) or which were part of glycans where specific changes were made such as adding new sugars or flipping sugars to resolve linkage torsion outliers. Sugars above the grey dashed line had a better map to model fit after our changes, while the inverse is true for sugars below this line.

Conclusions and Future Work

ChimeraX is a popular molecular visualisation tool for structural biologists, particularly those working with cryo-EM data, however, until recently, it has lacked carbohydrate specific methods. The *Privateer* bundle provides users the ability to validate carbohydrates in their structures, checking conformations, anomers, linkage torsion angles, as well as comparing to trusted glycomics databases, all from within ChimeraX without the need to install additional software, or upload files. It also provides further visualisation options for glycans in a structure, allowing easier viewing of potentially important interactions.

However, the *Privateer* bundle does not at present provide carbohydrate building or modelling tools. These additional features could prove extremely useful to those wishing to accurately build structures which contain carbohydrates.

Methods

Validation and Visualisation

The *Privateer* bundle for ChimeraX utilises the same underlying methods and code for carbohydrate validation. This means that it checks carbohydrate ring conformation using Cremer-Pople puckering coordinates (Cremer and Pople 1975), linkage stereochemistry, and glycan nomenclature to identify potential issues in carbohydrates present within a model (Agirre et al. 2015).

The bundle also includes later additions to the *Privateer* software such as the ability to search for glycans identified in the model in online glycomics databases GlyConnect and GlyToucan (Bagdonas et al. 2020). Analysis of the linkage torsion angles within the *N*-glycans (Dialpuri et al. 2023) and *C*-glycans (Holland et al. 2025) is also included. This allows for comparison of the torsion angles present in the modeled glycan to *Privateer*'s torsion database in order to identify possible outliers.

In addition to this validation, the *Privateer* bundle for ChimeraX also allows for improved visualisation of the carbohydrates present in a model by displaying 3D glycan symbols over the model displayed in ChimeraX. The symbols are based on the Symbol Nomenclature For Glycans (SNFGs), with the 2D shapes simply being extended into the third dimension — sugars represented by squares in SNFG are represented by cuboids, circles become cylinders, etc. — rather than using fully symmetrical 3D shapes such as cubes and spheres. This allows the plane of the symbol to match the plane of the sugar ring to aid in

understanding of the orientation of the sugars relative to each other, the protein, or anything they may be interacting with. The position of the symbols as drawn on the model are centered on the center of the sugar ring, as calculated by *Privateer* as the average of the coordinates of all atoms that form the sugar ring. The symbols are then aligned to the normal of the sugar ring mean plane.

Implementation

The *Privateer* bundle for ChimeraX is based off of the original *Privateer* source code, the repository for which can be found on GitHub (https://github.com/glycojones/privateer.git). *Privateer* primarily uses C++ for key computations which are speed-critical, with some of these functions made accessible via higher-level functions written in Python. The Python functions inherit from the original C++ functions, wrapped using PyBind 11.

The *Privateer* bundle for ChimeraX utilises the same C++ functions, but with new Python bindings (also wrapped using PyBind11) to provide the core functionality required. This includes analysing sugar conformations (Agirre et al. 2015), linkage stereochemistry and torsion angles (Dialpuri et al. 2023), as well as checking detected glycans against glycomics databases (Bagdonas et al. 2020) GlyConnect (Alocci et al. 2019) and GlyToucan (Fujita et al. 2021). The other C++ library dependencies are inherited from ChimeraX Clipper (https://cxtoolshed.rbvi.ucsf.edu/apps/chimeraxclipper), upon which the *Privateer* bundle is dependent. The user interface builds off of existing ChimeraX Python classes, utilising Qt (https://www.qt.io/). This user interface is designed to replicate the *Privateer* web app as much as possible for the sake of consistency and usability. The bundle is then compiled into a Python Wheel through ChimeraX.

Due to significant changes in the compilation method, as well as numerous additional Python scripts for the user interface, the *Privateer* bundle for ChimeraX will be maintained as a separate branch on the *Privateer* GitHub repository named “chimeraX_bundle”.

Availability and reproducibility

The full source code implemented and used here is available at https://github.com/glycojones/privateer, the Python Wheel can be downloaded from ChimeraX’s toolshed https://cxtoolshed.rbvi.ucsf.edu/apps/chimeraxprivateer, or installed directly from within ChimeraX. The data used are publicly available via the PDB and EMDB, and the edited version of PDB entry 9PIT and EMDB entry 71677 can be found at https://doi.org/10.5281/zenodo.19910199.

Acknowledgements

The ChimeraX team provided much helpful insight via the ChimeraX developers form.

Funding

Lou Holland is funded by The Royal Society (URF\R\221006). Jordan Dialpuri is funded by the Biotechnology and Biological Sciences Research Council (BBSRC; grant No. BB/T0072221). 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. Shuai Wang is funded by the Chemistry Wild Fund PhD Studentship of the University of York. Jon Agirre is a Royal Society University Research Fellow (awards UF160039 and URF\R\221006) and is partly funded by the BBSRC (grant no. BB/Y00888X/1).

References

Agirre, Jon, Mihaela Atanasova, Haroldas Bagdonas, et al. 2023. “The CCP4 Suite: Integrative Software for Macromolecular Crystallography.” *Acta Crystallographica. Section D, Structural Biology* 79 (Pt 6): 449–461.

Agirre, Jon, Javier Iglesias-Fernández, Carme Rovira, Gideon J. Davies, Keith S. Wilson, and Kevin D. Cowtan. 2015. “Privateer: Software for the Conformational Validation of Carbohydrate Structures.” *Nature Structural & Molecular Biology* 22 (11): 833–834.

Alocci, Davide, Julien Mariethoz, Alessandra Gastaldello, et al. 2019. “GlyConnect: Glycoproteomics Goes Visual, Interactive, and Analytical.” *Journal of Proteome Research* 18 (2): 664–677.

Atanasova, Mihaela, Haroldas Bagdonas, and Jon Agirre. 2020. “Structural Glycobiology in the Age of Electron Cryo-Microscopy.” *Current Opinion in Structural Biology* 62 (June): 70–78.

Atanasova, Mihaela, Robert A. Nicholls, Robbie P. Joosten, and Jon Agirre. 2022. “Updated Restraint Dictionaries for Carbohydrates in the Pyranose Form.” *Acta Crystallographica. Section D, Structural Biology* 78 (Pt 4): 455–465.

Bader, D. L. V., G. Ozorowski, and A. B. Ward. 2026. “HIV-1 bnAb 1-23 in Complex with BG505 MD39 SOSIP and RM19R: 9pit.” In *Worldwide Protein Data Bank*. Worldwide Protein Data Bank, March 4. https://doi.org/10.2210/pdb9pit/pdb.

Bagdonas, Haroldas, Daniel Ungar, and Jon Agirre. 2020. “Leveraging Glycomics Data in Glycoprotein 3D Structure Validation with Privateer.” In *Beilstein Archives*. July 20. https://doi.org/10.3762/bxiv.2020.83.v1.

Böhm, Michael, Andreas Bohne-Lang, Martin Frank, Alexander Loss, Miguel A. Rojas-Macias, and Thomas Lütteke. 2019. “Glycosciences.DB: An Annotated Data Collection Linking Glycomics and Proteomics Data (2018 Update).” *Nucleic Acids Research* 47 (D1): D1195–D1201.

Cremer, D., and J. A. Pople. 1975. “General Definition of Ring Puckering Coordinates.” *Journal of the American Chemical Society* 97 (6): 1354–1358.

Crispin, Max, David I. Stuart, and E. Yvonne Jones. 2007. “Building Meaningful Models of Glycoproteins.” *Nature Structural & Molecular Biology* 14 (5): 354; discussion 354–5.

Croll, Tristan Ian. 2018. “ISOLDE: A Physically Realistic Environment for Model Building into Low-Resolution Electron-Density Maps.” *Acta Crystallographica. Section D, Structural Biology* 74 (Pt 6): 519–530.

Dialpuri, Jordan S., Haroldas Bagdonas, Mihaela Atanasova, et al. 2023. “Analysis and Validation of Overall N-Glycan Conformation in Privateer.” *Acta Crystallographica. Section D, Structural Biology* 79 (Pt 6): 462–472.

Dialpuri, Jordan S., Haroldas Bagdonas, Lucy C. Schofield, Phuong Thao Pham, Lou Holland, Paul S. Bond, et al. 2024. “Online Carbohydrate 3D Structure Validation with the Privateer Web App.” *Acta Crystallographica. Section F, Structural Biology Communications* 80 (Pt 2): 30–35.

Dialpuri, Jordan S., Haroldas Bagdonas, Lucy C. Schofield, Phuong Thao Pham, Lou Holland, and Jon Agirre. 2024. “Monitoring Carbohydrate 3D Structure Quality with the Privateer Database.” *Beilstein Journal of Organic Chemistry* 20 (April): 931–939.

Emsley, Paul, and Kevin Cowtan. 2004. “Coot: Model-Building Tools for Molecular Graphics.” *Acta Crystallographica. Section D, Biological Crystallography* 60 (Pt 12 Pt 1): 2126–2132.

Emsley, Paul, and Max Crispin. 2018. “Structural Analysis of Glycoproteins: Building N-Linked Glycans with Coot.” *Acta Crystallographica. Section D, Structural Biology* 74 (Pt 4): 256–263.

Emsley, P., B. Lohkamp, W. G. Scott, and K. Cowtan. 2010. “Features and Development of Coot.” *Acta Crystallographica. Section D, Biological Crystallography* 66 (Pt 4): 486–501.

Frank, M., T. Lütteke, and C-W von der Lieth. 2007. “GlycoMapsDB: A Database of the Accessible Conformational Space of Glycosidic Linkages.” *Nucleic Acids Research* 35 (Database issue): 287–290.

Fujita, Akihiro, Nobuyuki P. Aoki, Daisuke Shinmachi, et al. 2021. “The International Glycan Repository GlyTouCan Version 3.0.” *Nucleic Acids Research* 49 (D1): D1529–D1533.

Holland, Lou, Phuong Thao Pham, Haroldas Bagdonas, Jordan S. Dialpuri, Lucy C. Schofield, and Jon Agirre. 2025. “Methods for Detecting, Building, and Improving Tryptophan Mannosylation in Glycoprotein Structures.” *Protein Science* 34 (2): e70025.

Meng, Elaine C., Thomas D. Goddard, Eric F. Pettersen, et al. 2023. “UCSF ChimeraX: Tools for Structure Building and Analysis.” *Protein Science* 32 (11): e4792.

Pintilie, Grigore, Kaiming Zhang, Zhaoming Su, Shanshan Li, Michael F. Schmid, and Wah Chiu. 2020. “Measurement of Atom Resolvability in Cryo-EM Maps with Q-Scores.” *Nature Methods* 17 (3): 328–334.

Schofield, Lucy C., Jordan S. Dialpuri, Garib N. Murshudov, and Jon Agirre. 2024. “Post-Translational Modifications in the Protein Data Bank.” *Acta Crystallographica. Section D, Structural Biology* 80 (Pt 9): 647–660.

Varki, Ajit, Richard D. Cummings, Jeffrey D. Esko, et al., eds. 2022. *Essentials of Glycobiology, Fourth Edition*. 4th ed. Cold Spring Harbor Laboratory Press.

Yamashita, Keitaro, Colin M. Palmer, Tom Burnley, and Garib N. Murshudov. 2021. “Cryo-EM Single-Particle Structure Refinement and Map Calculation Using Servalcat.” *Acta Crystallographica. Section D, Structural Biology* 77 (Pt 10): 1282–1291.

Supplementary Information

Here, we include the full validation report for the structure with EMDB entry number 71677 and PDB accession code 9PIT prior to our changes (**Figs. 1A – 3A**) and after our changes (**Figs. 4A – 5A**). Each is split over three figures to fit onto a page while still maintaining readability.

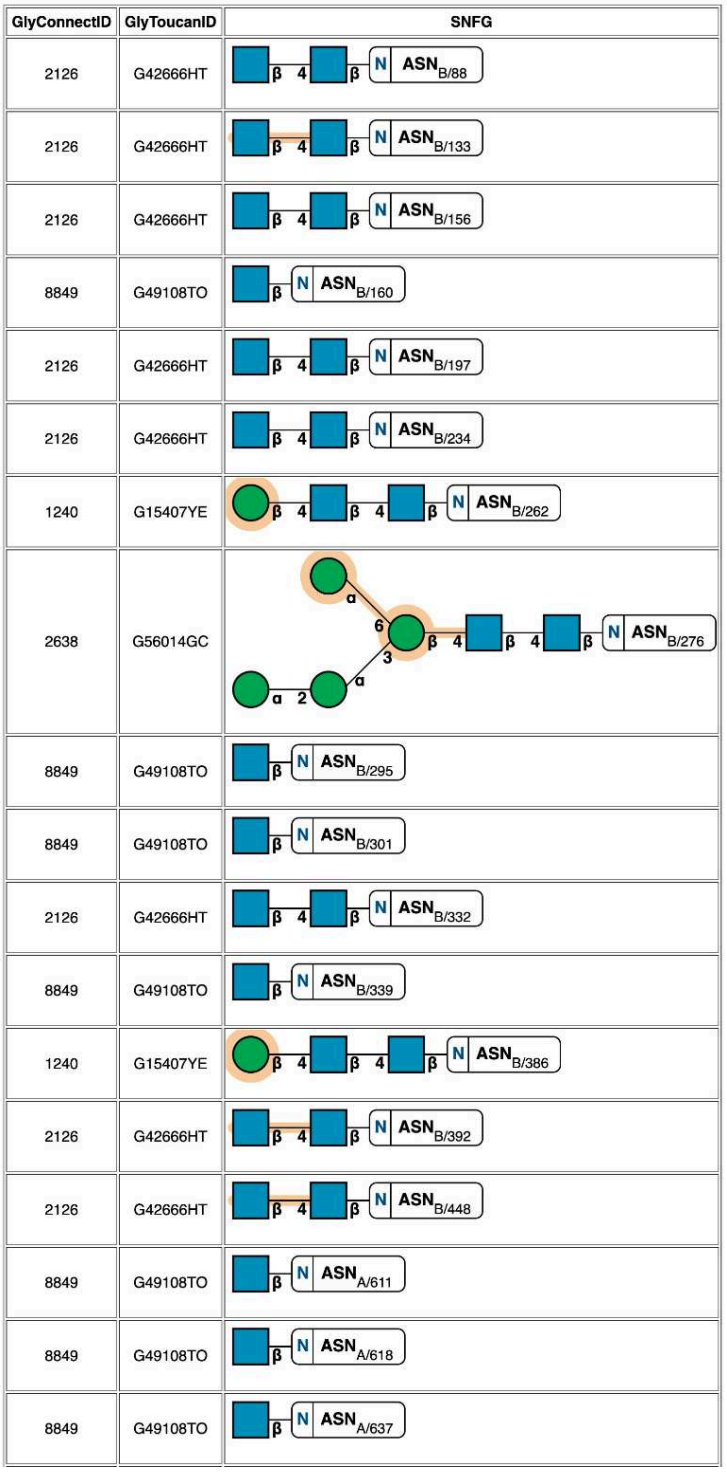


Figure 1A. Section 1 of validation report as produced by *Privateer* within ChimeraX for structure with EMDB entry number 71677(Bader et al. 2026). The first two columns contain glycan IDs as identified in online databases GlyConnect and GlyToucan. The third column contains the SNFG. The chain and residue ID of the glycan roots are provided for easy location, and linkage information is annotated on. Highlighted elements indicate potential issues.

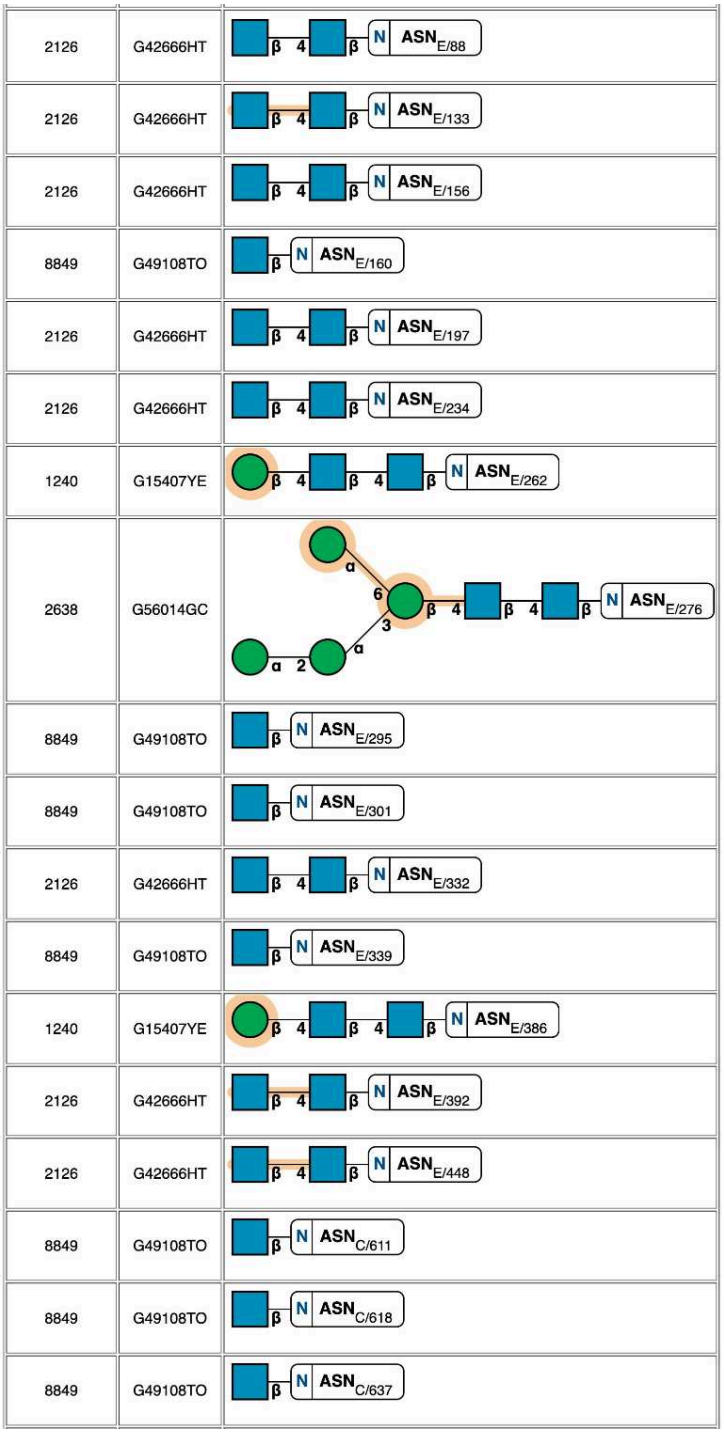


Figure 2A. Section 2 of validation report as produced by *Privateer* within ChimeraX for structure with EMDB entry number 71677(Bader et al. 2026). The first two columns contain glycan IDs as identified in online databases GlyConnect and GlyToucan. The third column contains the SNFG. The chain and residue ID of the glycan roots are provided for easy location, and linkage information is annotated on. Highlighted elements indicate potential issues.

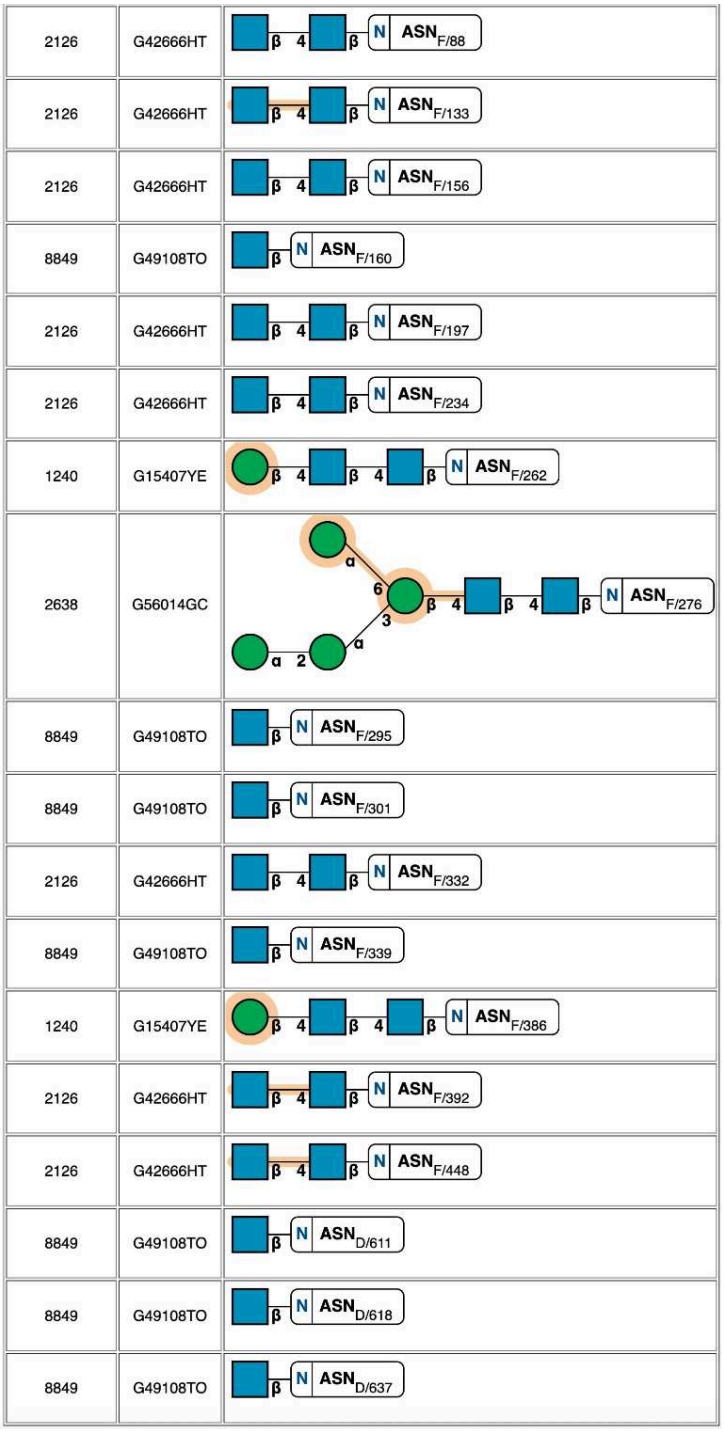


Figure 3A. Section 3 of validation report as produced by *Privateer* within ChimeraX for structure with EMDB entry number 71677(Bader et al. 2026). The first two columns contain glycan IDs as identified in online databases GlyConnect and GlyToucan. The third column contains the SNFG. The chain and residue ID of the glycan roots are provided for easy location, and linkage information is annotated on. Highlighted elements indicate potential issues.

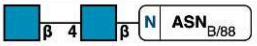
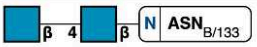
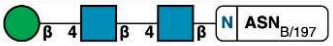
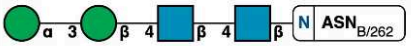
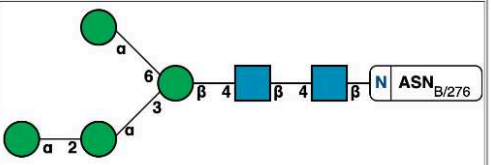
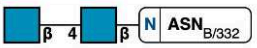
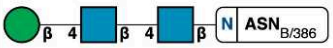
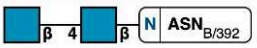
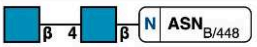
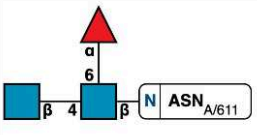
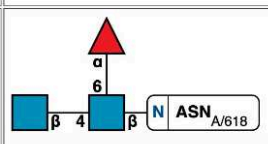
GlyConnectID	GlyToucanID	SNFG
2126	G42666HT	
2126	G42666HT	
2126	G42666HT	
8849	G49108TO	
1240	G15407YE	
2126	G42666HT	
545	G81315DD	
2638	G56014GC	
8849	G49108TO	
8849	G49108TO	
2126	G42666HT	
8849	G49108TO	
1240	G15407YE	
2126	G42666HT	
2126	G42666HT	
54	G21290RB	
54	G21290RB	

Figure 4A. Section 1 of validation report as produced by *Privateer* within ChimeraX for structure with EMDB entry number 71677(Bader et al. 2026) following our changes. The first two columns contain glycan IDs as identified in online databases GlyConnect and GlyToucan. The third column contains the SNFG. The chain and residue ID of the glycan roots are provided for easy location, and linkage information is annotated on. Highlighted elements indicate potential issues.

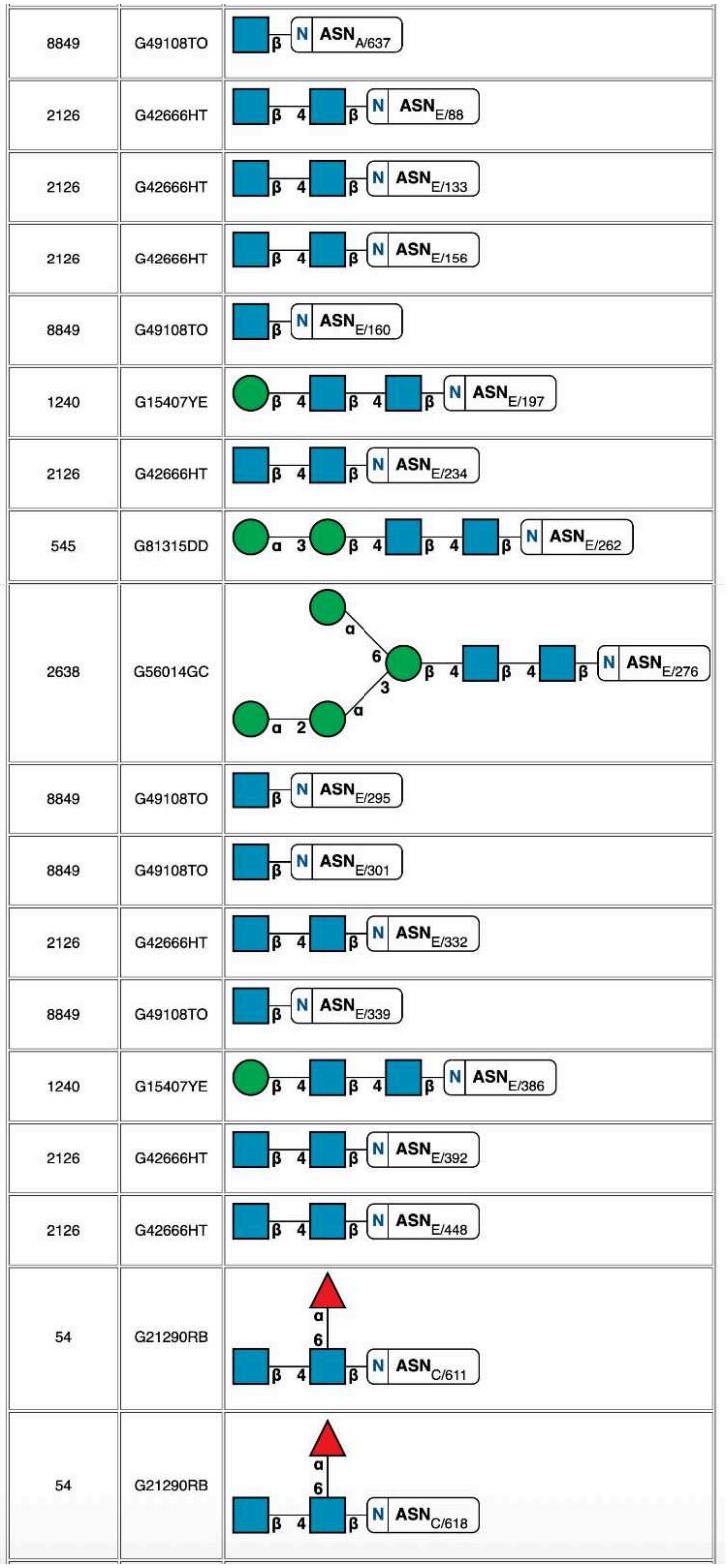


Figure 5A. Section 2 of validation report as produced by *Privateer* within ChimeraX for structure with EMDB entry number 71677(Bader et al. 2026) following our changes. The first two columns contain glycan IDs as identified in online databases GlyConnect and GlyToucan. The third column contains the SNFG. The chain and residue ID of the glycan roots are provided for easy location, and linkage information is annotated on. Highlighted elements indicate potential issues.

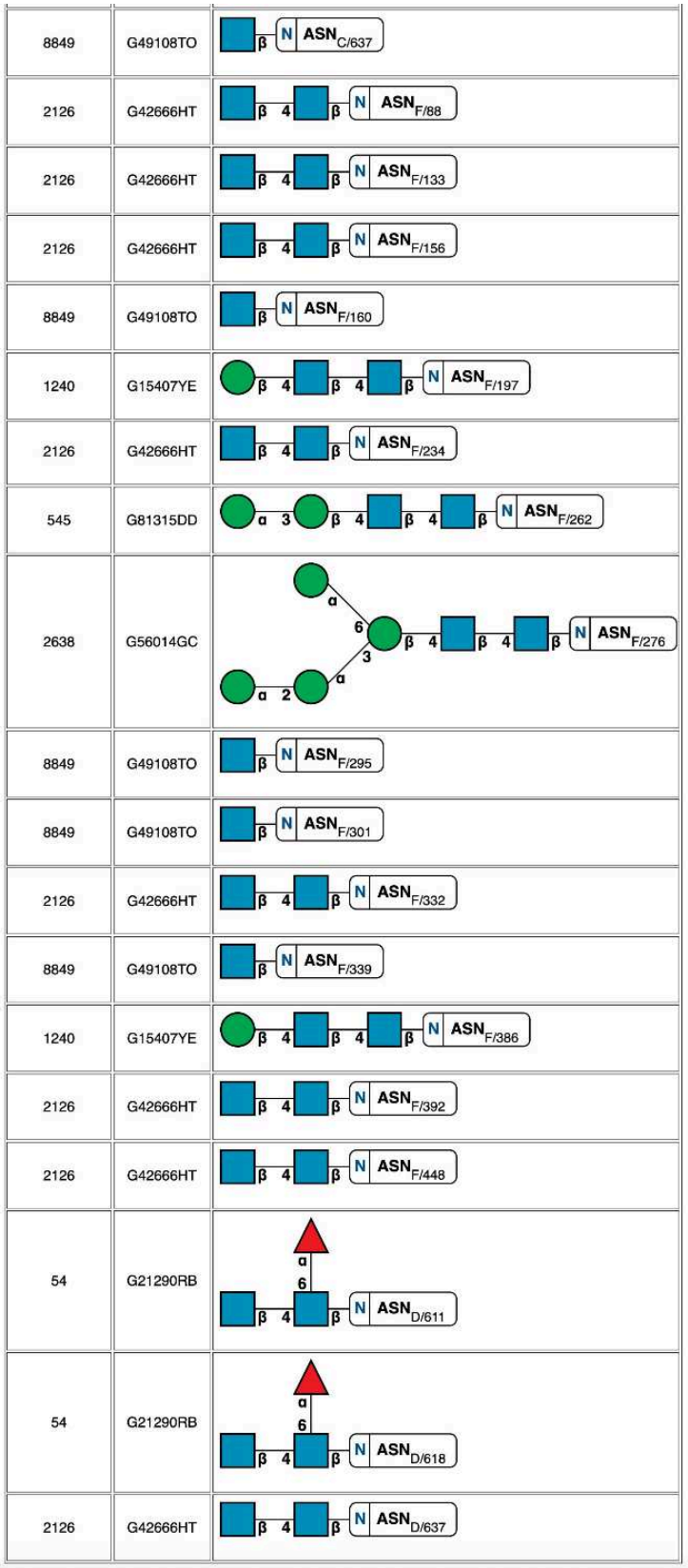


Figure 6A. Section 3 of validation report as produced by *Privateer* within ChimeraX for structure with EMDB entry number 71677(Bader et al. 2026) following our changes. The first two columns contain glycan IDs as identified in online databases GlyConnect and GlyToucan. The third column contains the SNFG. The chain and residue ID of the glycan roots are provided for easy location, and linkage information is annotated on. Highlighted elements indicate potential issues.