
Article:
Agirre, Jon orcid.org/0000-0002-1086-0253, Iglesias-Fernández, Javier, Rovira, Carme et al. (3 more authors) (2015) Privateer:software for the conformational validation of carbohydrate structures. Nature Structural & Molecular Biology. pp. 833-834. ISSN 1545-9993

https://doi.org/10.1038/nsmb.3115

Privateer: software for the conformational validation of carbohydrate structures

Jon Agirre[1], Javier Iglesias-Fernández[2], Carme Rovira[2,3], Gideon J Davies[1], Keith S Wilson[1] & Kevin D Cowtan[1]

[1] York Structural Biology Laboratory, Department of Chemistry, The University of York, England.

[2] Departament de Química Orgànica and Institut de Química Teòrica i Computacional (IQTCUB), Universitat de Barcelona, Barcelona, Spain.

[3] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain.

e-mail: kevin.cowtan@york.ac.uk or jon.agirre@york.ac.uk

Nature Structural & Molecular Biology **22**, 833–834 (2015) doi:10.1038/nsmb.3115
Published online 04 November 2015

Carbohydrates, including O- and N-glycans attached to protein and lipid structures, are increasingly important in cellular biology. Crystallographic refinement of sugars is, however, very poorly performed with thousands of incorrect structures polluting the PDB[1,2]. While nomenclature validation became possible in the past decade with the introduction of tools such as *pdb-care*[3], inappropriate refinement protocols at resolutions lower than 1.6 Å can still force the correct sugar into a most improbable ring conformation, let alone distort one with chemical errors[1]. Higher-energy conformations are very infrequent in nature - even out of the question in N-glycans - and should always be backed by clear electron density, and otherwise treated as outliers.

We have developed the Privateer software package for the detection and prevention of conformational anomalies in cyclic carbohydrate structures. The software identifies incorrect regio- and stereo-chemistry and unlikely conformations, produces a visual checklist for rapid correction of the errors in real space, and ensures that conformational preferences are accounted for during subsequent rebuilding and refinement, achieved as follows.

Privateer holds a manually curated database of supported monosaccharides based on the *PDB Chemical Component Dictionary*, whose entries contain coordinates for an energy-minimized conformer that the PDB calculates upon new depositions using Corina (Molecular Networks) or Omega (OpenEye). For each of these supported sugars, the database holds the conformation as determined by the Cremer-Pople algorithm[4], as well as stereochemical and geometric information, to be used as reference upon validation with the software. Additionally, puckering amplitudes[4] of pyranoses are compared to those registered

during *ab initio* metadynamics simulations of the conformational free-energy landscape of cyclohexane[5].

A real space correlation coefficient against omit mFo-DFc electron density is reported as a quality-of-fit indicator. Bias-minimized map coefficients are exported automatically and can be subsequently used to assess identification of the sugars.

Privateer is able to feed results into Coot[6] through its Python or Scheme scripting interface, loading model and maps automatically and flagging issues up visually. Most of the conformational outliers detected at higher resolutions (< 1.6Å) are typically wrongly modelled from the start and may require rebuilding (Fig. 1, top panel). At lower resolutions, higher-energy conformations may appear as a consequence of under-parameterized refinement, in spite of starting from a correct input model. These can be corrected during both real and reciprocal space refinement (Fig. 1, bottom panel) provided that enough restraints are introduced to balance the parameter-to-observation ratio. A common way of introducing conformational preferences is by using harmonic torsion restraints, but these are equally satisfied by multiple conformations, *e.g.* a chair and a boat. To enforce a single conformation, Privateer produces CIF library files that contain monoperiodic torsion restraints on the ring bonds, computed directly from the minimal-energy conformer. These files, generated using the *CCP4 Storage, Retrieval and Search framework for small molecule data* (CCP4SRS), are compatible with all major refinement packages. In addition, torsion restraints can be turned on for the distorted sugars in Refmac5[7] using the keyword file generated by Privateer, making its integration into re-refinement pipelines such as PDB_REDO[8] straightforward. It is just such examples that emphasize the problems of modelling sugars where the density is poor. In both cases, the published structures have high energy conformations which are not justified by the density, while is clearly more appropriate to use the expected minimal energy conformation.

Privateer is distributed by CCP4[9]. A graphical user interface will be available in the upcoming CCP4 7.0 release, fully automating the validation, correction and re-refinement process, and providing access to interactive HTML5/Javascript graphical reports.

This publication should be considered the primary citation for Privateer.

Acknowledgments

This work was supported by the Biotechnology and Biological Sciences Research Council grant BB/K008153/1. The authors thank Eugene Krissinel (CCP4, UK) for valuable assistance in using CCP4SRS, and to the Barcelona Supercomputing Center (BSC).

Competing financial interests

The authors declare no competing financial interests.

References

[1] Agirre, J., *et al. Nat Chem Biol* **11(5)**, 303 (2015).
[2] Lütteke, T. *et al. Carbohydr. Res.* **339,** 1015-1020 (2004).
[3] Lütteke, T. & von der Lieth, C.W. *BMC bioinformatics.* **5,** 69 (2004).
[4] Cremer, D. & Pople, J.A. *J. Am. Chem. Soc.* **97**, 1354–1358 (1975).
[5] Iglesias-Fernandez, J., *et al. Chemical Science* **6(2)**, 1167-1177 (2015).
[6] Emsley, P. *et al. Acta Cryst.* **D66 (4)**, 486-501 (2010).
[7] Murshudov, G.N. *et al. Acta Cryst.* **D67(4),** 355-367 (2011).
[8] Joosten, R.P. *et al. Bioinformatics* **27**, 3392–3398 (2011).
[9] Winn, M.D. *et al. Acta Cryst.* **D67(4)**, 235–242 (2011).
[10]Brunger, A.T. *Nature Protocols* **2**, 2728-2733 (2007).

Figure 1

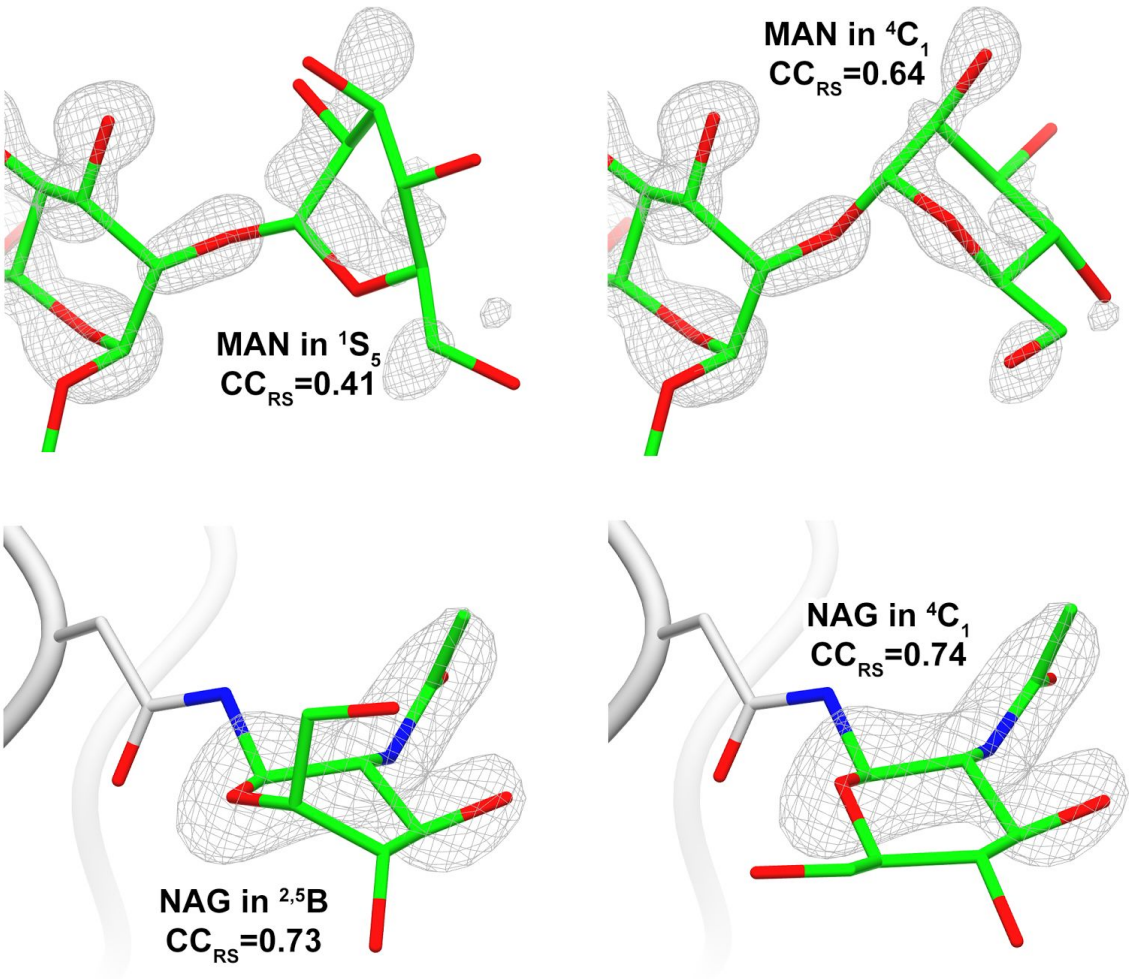

Figure 1. N-glycan D-pyranosides in higher-energy conformations as deposited in the PDB (left) and their corrected counterparts (right). Grey mesh: omit mFo-DFc density contoured at 4σ. Top panel: MAN 616 (chain A) from PDB 3QVP. 3d7d, 1.1 Å-resolution data. This sugar was distorted to a $^{1}S_{5}$ conformation by the refinement software (CNS[10]) after being modeled with a β linkage instead of the expected α one, and therefore required rebuilding before it could be refined to the minimal-energy conformation ($^{4}C_{1}$). Bottom panel: NAG 1757 (chain A) from PDB 3D7D, refined with Refmac5 to the $^{2,5}B$ conformation in partial density. Since this conformation minimizes deviations from ideal bond lengths, angles and harmonic torsions, the model could not be corrected by restraining those parameters. Application of the monoperiodic torsion set in Refmac5 produced by Privateer produced the expected minimal energy conformation ($^{4}C_{1}$).