Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

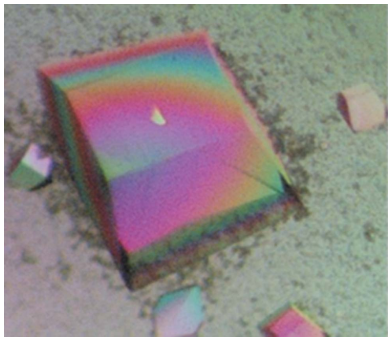
Preliminary joint neutron time-of-flight and X-ray crystallographic study of human ABO(H) blood group A glycosyltransferase

B. Schuman,[a] S. Z. Fisher,[b] A. Kovalevsky,[b] S. N. Borisova,[a] M. M. Palcic,[c] L. Coates,[d] P. Langan[b] and S. V. Evans[a]*

[a]Department of Biochemistry and Microbiology, University of Victoria, PO Box 3800, STN CSC, Victoria, BC V8W 3P6, Canada, [b]Bioscience Division, MS M888, Los Alamos National Laboratory, Los Alamos, NM 87545, USA, [c]Carlsberg Laboratory, Gamle Carlsberg Vej 10, 2500 Valby, Denmark, and [d]Neutron Scattering Science Division, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN 37831, USA

Correspondence e-mail: svevans@uvic.ca

Received 28 October 2010
Accepted 7 December 2010



The biosyntheses of oligosaccharides and glycoconjugates are conducted by glycosyltransferases. These extraordinarily diverse and widespread enzymes catalyze the formation of glycosidic bonds through the transfer of a monosaccharide from a donor molecule to an acceptor molecule, with the stereochemistry about the anomeric carbon being either inverted or retained. Human ABO(H) blood group A α-1,3-*N*-acetylgalactosaminyltransferase (GTA) generates the corresponding antigen by the transfer of *N*-acetylgalactosamine from UDP-GalNAc to the blood group H antigen. To understand better how specific active-site-residue protons and hydrogen-bonding patterns affect substrate recognition and catalysis, neutron diffraction studies were initiated at the Protein Crystallography Station (PCS) at Los Alamos Neutron Science Center (LANSCE). A large single crystal was subjected to H/D exchange prior to data collection and time-of-flight neutron diffraction data were collected to 2.5 Å resolution at the PCS to ∼85% overall completeness, with complementary X-ray diffraction data collected from a crystal from the same drop and extending to 1.85 Å resolution. Here, the first successful neutron data collection from a glycosyltransferase is reported.

1. Introduction

Glycosyltransferase enzymes build complex carbohydrates by the sequential transfer of a monosaccharide unit from an activated donor molecule (typically a nucleotide) to an acceptor molecule in a stereospecific and regiospecific manner (Schuman *et al.*, 2007; Breton *et al.*, 2006; Lairson *et al.*, 2008). The nascent glycosidic linkage can either invert the anomeric stereochemistry of the donor (*e.g.* where an α-linked UDP-Gal forms a β-galactoside) or retain it (*e.g.* where an α-linked UDP-Gal forms a α-galactoside). The mechanism for the latter glycosyltransfer stereospecificity has yet to be unambiguously resolved. There are two leading candidates: an S_N1 mechanism known as S_Ni (Sinnott & Jencks, 1980) and two consecutive S_N2 attacks known as 'double displacement' (Chelsky & Parsons, 1975; Fig. 1). There are lines of evidence against both postulated mechanisms (Lairson *et al.*, 2004; Sinnott, 1990). Exploring the fundamental structure–function relationships of glycosyltransferase enzymes is essential in order to effectively determine the mechanisms of substrate recognition and catalysis. An understanding of these mechanisms can be used to exploit the directed development of biomedically therapeutic glycoconjugates and inhibitor drugs to treat diseases ranging from viral and bacterial infections (Umesiri *et al.*, 2010) to cancer (Werther *et al.*, 1994), genetic disorders (Wennekes *et al.*, 2007) and organ transplantation (Klymiuk *et al.*, 2010) in which specific glycosyltransferases are known to cause or facilitate disease progression.

The glycosyltransferase GTA transfers the *N*-acetylgalactosaminyl monosaccharide residue from UDP to Gal O3 of the H antigen [minimal epitope Fuc-α(1→2)Gal] with retention of the axial stereochemistry to generate the human blood group A antigen and is a model enzyme for studying the retaining glycosyltransfer mechanism (Alfaro *et al.*, 2008; Lee *et al.*, 2005; Letts *et al.*, 2007; Nguyen *et al.*, 2003; Patenaude *et al.*, 2002; Persson *et al.*, 2007; Schuman *et al.*, 2007, 2010; Seto *et al.*, 1999). Evaluation of hydrogen-bonding partners and amino-acid protonation states is critical for

understanding the catalytic mechanisms of GTA and indeed of all retaining glycosyltransferases. GTA active-site residues of particular interest include the ionization state of Glu303, a proposed nucleophile (Patenaude *et al.*, 2002) which must be deprotonated if it is to act as a nucleophile for the double-displacement mechanism, as well as the three residues most strongly associated with changing their hydrogen-bonding partners during the formation of the 'closed' state of the enzyme upon substrate binding: Arg188 loses a bidentate salt bridge to Asp211 upon substrate binding in order to interact with Asp302.

X-ray crystallography is the method of choice for the structure determination of protein structures. X-rays are scattered by the electron clouds associated with atomic nuclei, with an increase in the magnitude of scattering that is proportional to the atomic number. Consequently, while atoms such as C, N and O are readily visible in the electron-density maps of proteins, H atoms, especially the more mobile functionally important ones of particular biological interest, are usually invisible except in cases where a subset of the total number of H atoms can be seen at ultrahigh resolution (Blakeley *et al.*, 2008). In contrast, neutrons diffract strongly from H (scattering length −3.7 fm) compared with other atom types found in proteins (scattering lengths: C, 6.6 fm; O, 5.8 fm; N, 9.4 fm; S, 2.8 fm). This feature makes the visualization and assignment of H-atom locations feasible at medium resolution (Meilleur *et al.*, 2006; Langan *et al.*, 2008; Niimura & Bau, 2008; Blakeley *et al.*, 2008).

In neutron crystallographic studies it is now routine to subject crystals to H/D exchange prior to data collection. D has a larger and positive scattering length (scattering length 6.7 fm) and a smaller incoherent scattering cross-section compared with H. Not only does this H/D exchange dramatically decrease the background arising from incoherent scattering, thus effectively increasing the signal to noise, but it also increases the visibility of H and provides additional information such as solvent accessibility and protein 'breathing' and can identify minimal folding domains in proteins (Bennett *et al.*, 2006). In conventional electron-density maps water molecules usually appear as spherical peaks indicating the position of the O atom. In nuclear density maps, because of the additional density for D atoms it is possible to differentiate between OH^-, H_3O^+ and H_2O (*i.e.* OD^-, D_3O^+ and D_2O) species. In the same way, it is possible to visualize and place H/D atoms on amino-acid residues in order to determine their protonation states and hydrogen-bonding interactions.

X-ray and neutron diffraction data are therefore highly complementary and can be combined during model refinement to obtain more accurate structures that include all of the atoms, including H/D (Adams *et al.*, 2009). The accurate placement and orientation of amino-acid side chains and solvent molecules can result in a detailed understanding of the hydrogen bonding, *via* the deuterium patterns, which can facilitate the elucidation of catalytic mechanisms (Kossiakoff & Spencer, 1981; Niimura *et al.*, 1997; Bennett *et al.*, 2006; Blakeley *et al.*, 2008; Coates *et al.*, 2008; Blum *et al.*, 2009; Adachi *et al.*, 2009; Yagi *et al.*, 2009; Tomanicek *et al.*, 2010; Fisher *et al.*, 2010; Kovalevsky *et al.*, 2010).

Despite a number of kinetic and X-ray crystallographic studies on GTA and other similar enzymes the enzyme mechanism remains unresolved. To understand further the mechanisms of substrate recognition and of catalysis and to probe the role of hydrogen-bonding patterns in the active site of GTA, neutron diffraction studies on H/D-exchanged GTA samples were initiated. Here, we report the first neutron diffraction data-collection strategy and preliminary data of a glycosyltransferase.

2. Crystallization of GTA

Human GTA (gi:85544029) with the N-terminal transmembrane domain truncated at residue Met63 and codon-optimized for BL21 *Escherichia coli* was expressed and purified from cultures using a two-step protocol consisting of ion-exchange chromatography followed by UDP-hexanolamine affinity chromatography as described previously (Marcus *et al.*, 2003). Large crystals were produced in 30–40 µl drops without D_2O in conditions similar to those reported previously (Letts *et al.*, 2007; Alfaro *et al.*, 2008; Schuman *et al.*, 2010): 6–20 mg ml^{-1}

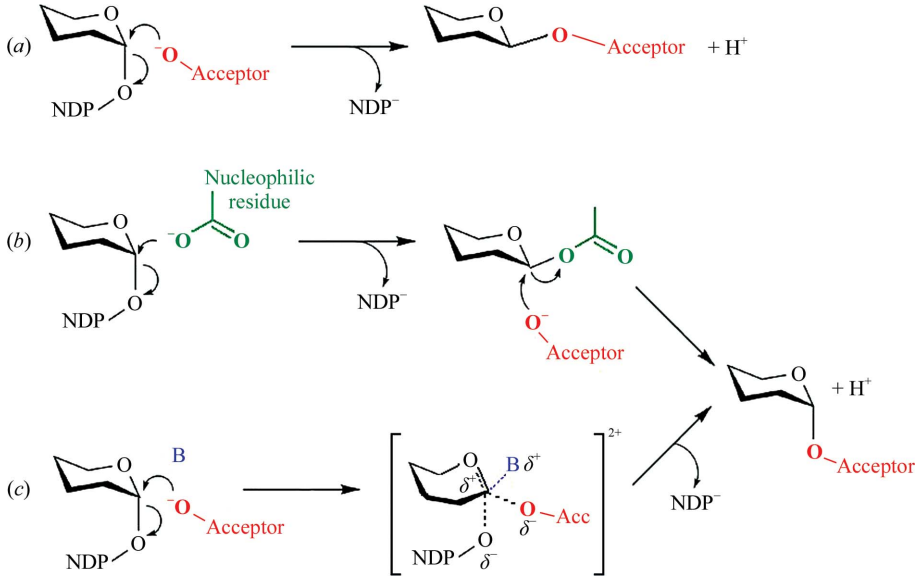


Figure 1
Mechanisms proposed for glycosyltransferases. (*a*) Inverting GTs promote an S_N2 nucleophilic attack of the donor by the acceptor, with resulting inversion of the anomeric bond stereochemistry. (*b*, *c*) Retaining enzymes have two postulated mechanisms. The double-displacement mechanism utilizes two S_N2 inversions: the first is an attack by an enzyme nucleophile on the donor sugar to produce a glycosyl-enzyme intermediate with inverted stereochemistry for the anomeric bond and the second is an attack by the acceptor on the intermediate to form a product with a net retention of stereochemistry. The S_Ni mechanism (*c*) involves a single-step S_N1 transfer that uses the acceptor as a base-stabilized pentavalent 'internal return' intermediate to yield retention of configuration.

protein, 35 m*M* sodium acetate pH 4.6, 45 m*M* *N*-(2-acetamido)-iminodiacetic acid (ADA) pH 7.5, 30 m*M* sodium chloride, 3–4 m*M* magnesium chloride, 3–3.5% PEG 4000 equilibrated against a reservoir consisting of 10% PEG 4000, 50 m*M* ADA, 50 m*M* sodium acetate, 100 m*M* ammonium sulfate, 5 m*M* magnesium chloride. H/D exchange was achieved by vapour diffusion over several months in a quartz capillary by flanking the crystal with D_2O mother liquor.

3. Room-temperature neutron and X-ray diffraction data collection and reduction

A suitable H/D-exchanged crystal of GTA (Fig. 2) was selected for neutron data collection at the Protein Crystallography Station (PCS) at Los Alamos Neutron Science Center (LANSCE; Langan *et al.*, 2004, 2008). After a 24 h test exposure, it was clear that the diffraction quality and signal-to-noise ratio were appropriate for full data collection and structure determination.

Time-of-flight wavelength-resolved Laue diffraction data were collected at room temperature on a Huber κ-circle goniometer at 18 usable settings. Each setting was collected for 18 h and the data extended to better than 2.5 Å resolution. The crystal-to-detector distance was fixed at 730 mm, which corresponded to the cylindrical radius of the curved detector. The detector has a limited vertical area

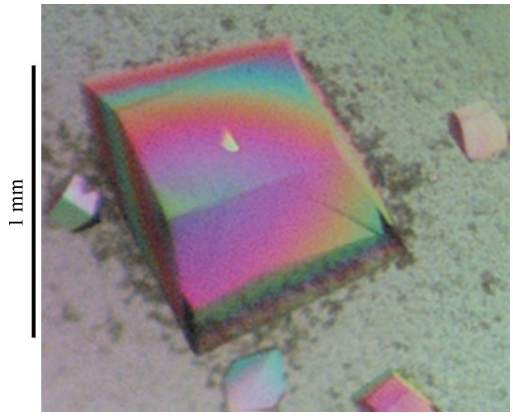


Figure 2
Photograph of a large single GTA crystal (0.95 × 0.6 × 0.7 mm). Neutron diffraction was observed from crystals that were considerably smaller. The scale bar represents 1 mm.

Table 1
Data-collection statistics for neutron and X-ray diffraction of the $C222_1$ GTA crystals at room temperature.

Values in parentheses are for the outer shell.

	X-ray	Neutron
Source	Cu $K\alpha$	LANSCE
Detector	Rigaku R-AXIS IV^{++}	PCS ^{3}He detector
Wavelength/range used (Å)	1.5418	0.7–6.4
Unit-cell parameters (Å)	a = 53.5, b = 151.5, c = 80.6	a = 52.3, b = 151.9, c = 80.5
Unit-cell volume (Å^3)	653283	654195
Resolution (Å)	19.47–1.85 (1.92–1.85)	27.0–2.50 (2.64–2.50)
No. of reflections (measured/unique)	120226/26867	27675 (2219)/9759 (1163)
Multiplicity	4.47 (4.19)	2.8 (1.9)
Mean $I/\sigma(I)$	14.1 (4.1)	3.6 (1.6)
R_{merge} (%)	4.9 (29.5)	25.1 (32.4)
Completeness (%)	94.7 (98.2)	84.1 (70.3)

(120° horizontal span, 16° vertical span) and to cover reciprocal space the crystal was reoriented five times using the κ- and ω-goniometer circles. We performed 30° steps in φ at each orientation for a total of 18 images (Fig. 3; Table 1).

Each image was processed with a version of *d*TREK* (Pflugrath, 1999) that had been modified in-house for use with wavelength-resolved Laue neutron data (Langan & Greene, 2004). Integrated reflections were wavelength-normalized using *LAUENORM* and merged using *SCALA* from the *CCP*4 suite of crystallographic programs (Diederichs & Karplus, 1997; Helliwell *et al.*, 1989; Evans, 2006; Collaborative Computational Project, Number 4, 1994). The polychromatic beam produced by the thermal moderator used at the PCS contains neutrons with wavelengths in the range 0.6–7.0 Å. However, the flux drops off rapidly below 0.7 Å and above 6.4 Å (Langan *et al.*, 2004). Therefore, during data processing only reflections produced by neutrons within the restricted wavelength range 0.7–6.4 Å were used in order to exclude those with poor signal to noise because of low count levels. This led to improved data statistics by using only the most accurately measured thermal neutrons.

A room-temperature X-ray data set from a crystal from the same drop was collected on an R-AXIS IV^{++} area detector at a distance of 72 mm coupled to a MicroMax-002 generator with Osmic Blue optics (Rigaku Americas) and will be used for joint neutron and X-ray structure refinement (Fig. 4; Table 1). The neutron diffraction Laue data set was solved by assuming isomorphism with the structure previously solved using X-ray data. Joint structure refinement of the X-ray and neutron data is in progress using *nCNS* (Adams *et al.*,

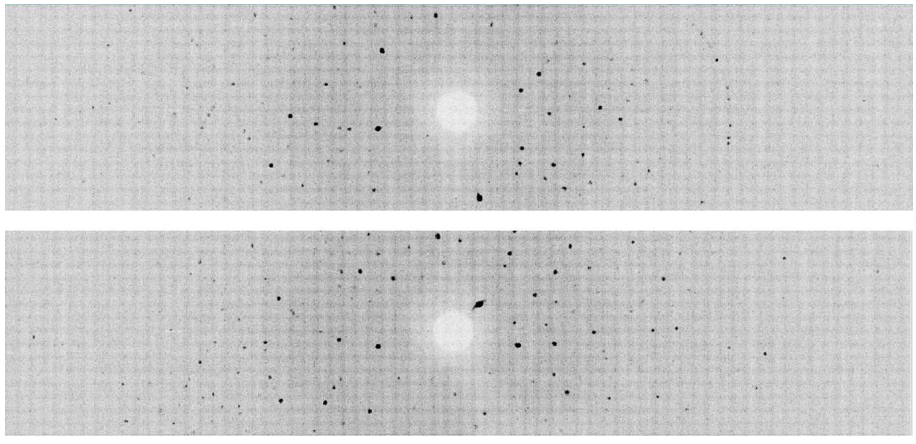


Figure 3
Neutron Laue time-of-flight diffraction images of GTA collected at the PCS at two different φ settings at $\kappa = 0°$. In this representation the time-of-flight data were overlaid to produce a conventional Laue pattern.

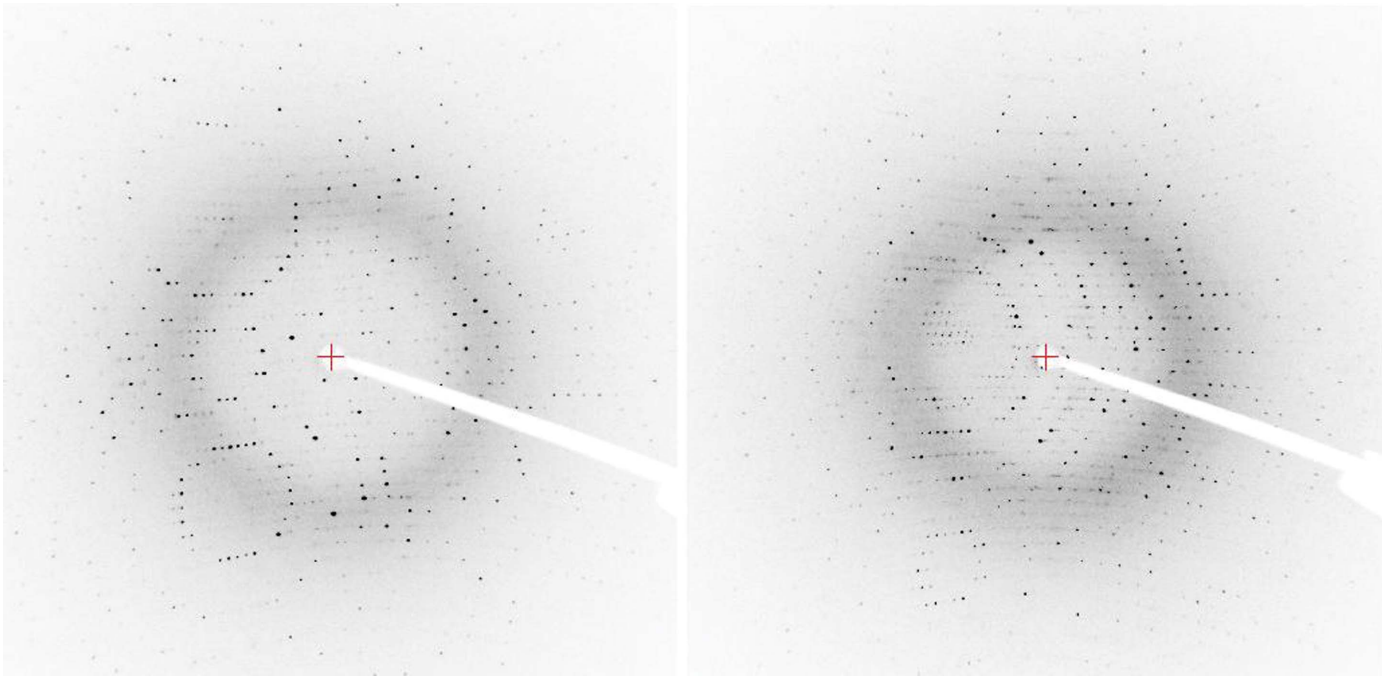


Figure 4
X-ray diffraction images of GTA collected at room temperature with relative angles of 0 and 90° with 0.5° oscillations and 2 min exposure.

2009), a version of *CNS* that has been designed to handle neutron and X-ray data simultaneously, and *phenix.refine* from the *PHENIX* program suite (Adams *et al.*, 2002). Initial neutron and electron-density maps after only one round of rigid-body refinement already reveal some of the details of protonation and hydrogen bonding in the active site of GTA. Analysis of the neutron structure will involve detailed mapping of H atoms, overall H/D exchange and hydrogen-bonding patterns in the GTA active site.

4. Results and discussion

The overall completeness of the neutron data set is ∼85% to 2.5 Å resolution (Table 1), although the crystal produced significant neutron diffraction to about ∼2.2 Å resolution. The neutron diffraction study of GTA reported here is the first of its kind for any glycosyltransferase. Glycosyltransferases are widespread ubiquitous enzymes that are involved in many metabolic processes. A better understanding of their mechanisms of substrate recognition and catalysis will be applicable to a number of biological and biomedical problems. Sample size is a major impediment to more projects being suitable for neutron experiments (Blum *et al.*, 2009). It is therefore significant that it was possible to record neutron data from crystals in the 0.3 mm^3 size range in this study. We also note that the size of the unit cell in the *b*-axis direction is considerably larger than any reported previously in a neutron crystallographic study of an enzyme (although this is mitigated somewhat by the *C*-centring, which results in every second spot being missing). There is only 0.1% difference between the cell volumes of the two data sets and joint refinement is proceeding such that D atoms are observable for residues with exchangeable protons (Fig. 5). We expect the co-refinement of the neutron and X-ray data to reveal many important details about the GTA active-site-residue ionization states and geometry, as well the hydrogen-bonding patterns in the active site that make catalysis possible.

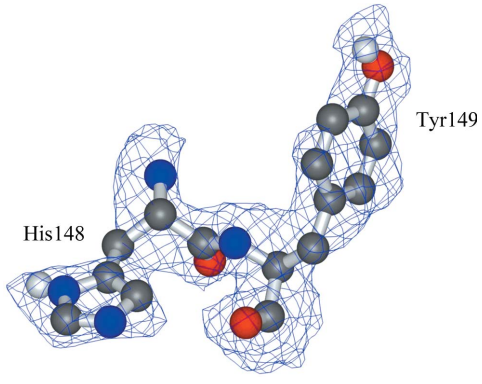


Figure 5
Initial $2mF_o - DF_c$ filled neutron maps of deuterium-exchanged residues after a single round of refinement showing the protonation of His148 and Tyr149.

The PCS is funded by the Department of Energy Office of Biological and Environmental Research (DOE-OBER). Funding from the Canadian Institutes of Health Research MOP-77655 and salary support from the Michael Smith Foundation for Health Research to SVE is acknowledged. Oak Ridge National Laboratory is managed by UT-Battelle LLC under contract No. DE-AC05-000R22725 for the US Department of Energy. PL was partly supported by a grant from the National Institute of General Medical Science of the National Institutes of Health (R01GM071939).

References

Adachi, M. *et al.* (2009). *Proc. Natl Acad. Sci. USA*, **106**, 4641–4646.
Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1948–1954.
Adams, P. D., Mustyakimov, M., Afonine, P. V. & Langan, P. (2009). *Acta Cryst.* D**65**, 567–573.

Alfaro, J. A., Zheng, R. B., Persson, M., Letts, J. A., Polakowski, R., Bai, Y., Borisova, S. N., Seto, N. O., Lowary, T. L., Palcic, M. M. & Evans, S. V. (2008). *J. Biol. Chem.* **283**, 10097–10108.
Bennett, B., Langan, P., Coates, L., Mustyakimov, M., Schoenborn, B., Howell, E. E. & Dealwis, C. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 18493–18498.
Blakeley, M. P., Ruiz, F., Cachau, R., Hazemann, I., Meilleur, F., Mitschler, A., Ginell, S., Afonine, P., Ventura, O. N., Cousido-Siah, A., Haertlein, M., Joachimiak, A., Myles, D. & Podjarny, A. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 1844–1848.
Blum, M. M., Mustyakimov, M., Ruterjans, H., Kehe, K., Schoenborn, B. P., Langan, P. & Chen, J. C. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 713–718.
Breton, C., Snajdrova, L., Jeanneau, C., Koca, J. & Imberty, A. (2006). *Glycobiology*, **16**, 29R–37R.
Chelsky, D. & Parsons, S. M. (1975). *J. Biol. Chem.* **250**, 5669–5673.
Coates, L., Tuan, H. F., Tomanicek, S., Kovalevsky, A., Mustyakimov, M., Erskine, P. & Cooper, J. (2008). *J. Am. Chem. Soc.* **130**, 7235–7237.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Fisher, S. Z., Kovalevsky, A. Y., Domsic, J. F., Mustyakimov, M., McKenna, R., Silverman, D. N. & Langan, P. A. (2010). *Biochemistry*, **49**, 415–421.
Helliwell, J. R., Habash, J., Cruickshank, D. W. J., Harding, M. M., Greenhough, T. J., Campbell, J. W., Clifton, I. J., Elder, M., Machin, P. A., Papiz, M. Z. & Zurek, S. (1989). *J. Appl. Cryst.* **22**, 483–497.
Klymiuk, N., Aigner, B., Brem, G. & Wolf, E. (2010). *Mol. Reprod. Dev.* **77**, 209–221.
Kossiakoff, A. A. & Spencer, S. A. (1981). *Biochemistry*, **20**, 6462–6474.
Kovalevsky, A. Y., Hanson, L., Fisher, S. Z., Mustyakimov, M., Mason, S. A., Forsyth, V. T., Blakeley, M. P., Keen, D. A., Wagner, T., Carrell, H. L., Katz, A. K., Glusker, J. P. & Langan, P. (2010). *Structure*, **18**, 688–699.
Lairson, L. L., Chiu, C. P. C., Ly, H. D., He, S. M., Wakarchuk, W. W., Strynadka, N. C. J. & Withers, S. G. (2004). *J. Biol. Chem.* **279**, 28339–28344.
Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. (2008). *Annu. Rev. Biochem.* **77**, 521–555.
Langan, P., Fisher, Z., Kovalevsky, A., Mustyakimov, M., Sutcliffe Valone, A., Unkefer, C., Waltman, M. J., Coates, L., Adams, P. D., Afonine, P. V., Bennett, B., Dealwis, C. & Schoenborn, B. P. (2008). *J. Synchrotron Rad.* **15**, 215–218.
Langan, P. & Greene, G. (2004). *J. Appl. Cryst.* **37**, 253–257.
Langan, P., Greene, G. & Schoenborn, B. P. (2004). *J. Appl. Cryst.* **37**, 24–31.
Lee, H. J., Barry, C. H., Borisova, S. N., Seto, N. O. L., Zheng, R. B., Blancher, A., Evans, S. V. & Palcic, M. M. (2005). *J. Biol. Chem.* **280**, 525–529.
Letts, J. A., Persson, M., Schuman, B., Borisova, S. N., Palcic, M. M. & Evans, S. V. (2007). *Acta Cryst.* D**63**, 860–865.
Marcus, S. L., Polakowski, R., Seto, N. O. L., Leinala, E., Borisova, S., Blancher, A., Roubinet, F., Evans, S. V. & Palcic, M. M. (2003). *J. Biol. Chem.* **278**, 12403–12405.
Meilleur, F., Myles, D. & Blakeley, M. (2006). *Eur. Biophys. J.* **35**, 611–620.
Nguyen, H. P., Seto, N. O. L., Cai, Y., Leinala, E. K., Borisova, S. N., Palcic, M. M. & Evans, S. V. (2003). *J. Biol. Chem.* **278**, 49191–49195.
Niimura, N. & Bau, R. (2008). *Acta Cryst.* A**64**, 12–22.
Niimura, N., Minezaki, Y., Nonaka, T., Castagna, J. C., Cipriani, F., Høghøj, P., Lehmann, M. S. & Wilkinson, C. (1997). *Nature Struct. Biol.* **4**, 909–914.
Patenaude, S. I., Seto, N. O., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palcic, M. M. & Evans, S. V. (2002). *Nature Struct. Biol.* **9**, 685–690.
Persson, M., Letts, J. A., Hosseini-Maaf, B., Borisova, S. N., Palcic, M. M., Evans, S. V. & Olsson, M. L. (2007). *J. Biol. Chem.* **282**, 9564–9570.
Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.
Schuman, B., Alfaro, J. A. & Evans, S. V. (2007). *Top. Curr. Chem.* **272**, 217–257.
Schuman, B., Persson, M., Landry, R. C., Polakowski, R., Weadge, J. T., Seto, N. O., Borisova, S. N., Palcic, M. M. & Evans, S. V. (2010). *J. Mol. Biol.* **402**, 399–411.
Seto, N. O., Compston, C. A., Evans, S. V., Bundle, D. R., Narang, S. A. & Palcic, M. M. (1999). *Eur. J. Biochem.* **259**, 770–775.
Sinnott, M. L. (1990). *Chem. Rev.* **90**, 1171–1202.
Sinnott, M. L. & Jencks, W. P. (1980). *J. Am. Chem. Soc.* **102**, 2026–2032.
Tomanicek, S. J., Blakeley, M. P., Cooper, J., Chen, Y., Afonine, P. V. & Coates, L. (2010). *J. Mol. Biol.* **396**, 1070–1080.
Umesiri, F. E., Sanki, A. K., Boucau, J., Ronning, D. R. & Sucheck, S. J. (2010). *Med. Res. Rev.* **30**, 290–326.
Wennekes, T., van den Berg, R. J., Donker, W., van der Marel, G. A., Strijland, A., Aerts, J. M. & Overkleeft, H. S. (2007). *J. Org. Chem.* **72**, 1088–1097.
Werther, J. L., Rivera-MacMurray, S., Bruckner, H., Tatematsu, M. & Itzkowitz, S. H. (1994). *Br. J. Cancer*, **69**, 613–616.
Yagi, D., Yamada, T., Kurihara, K., Ohnishi, Y., Yamashita, M., Tamada, T., Tanaka, I., Kuroki, R. & Niimura, N. (2009). *Acta Cryst.* D**65**, 892–899.