Acta Crystallographica Section D
**Biological Crystallography**
ISSN 0907-4449

# Sequence-dependent effects of cryoprotectants on the active sites of the human ABO(H) blood group A and B glycosyltransferases

**Asha R. Johal,[a] Brock Schuman,[a] Javier A. Alfaro,[a] Svetlana Borisova,[a] Nina O. L. Seto[b] and Stephen V. Evans[a]***

[a]Department of Biochemistry and Microbiology, University of Victoria, Victoria, BC V8W 3P6, Canada, and [b]Institute for Biological Sciences, National Research Council of Canada, 100 Sussex Drive, Ottawa, ON K1A 0R6, Canada

Correspondence e-mail: svevans@uvic.ca

Received 9 November 2011
Accepted 15 January 2012

**PDB References:** ABO glycosyltransferases, AAAA, MPD cryoprotectant, 3sxg; glycerol cryoprotectant, 3sxe; AABB, MPD, 3sx5; glycerol, 3sx3; ABBA, MPD, 3sx8; glycerol, 3sx7; ABBB, MPD, 3sxb; glycerol, 3sxa; BBBB, MPD, 3sxd; glycerol, 3sxc.

The human ABO(H) A and B blood group glycosyltransferases GTA and GTB differ by only four amino acids, yet this small dissimilarity is responsible for significant differences in biosynthesis, kinetics and structure. Like other glycosyltransferases, these two enzymes have been shown to recognize substrates through dramatic conformational changes in mobile polypeptide loops surrounding the active site. Structures of GTA, GTB and several chimeras determined by single-crystal X-ray diffraction demonstrate a range of susceptibility to the choice of cryoprotectant, in which the mobile polypeptide loops can be induced by glycerol to form the ordered closed conformation associated with substrate recognition and by MPD [hexylene glycol, (±)-2-methyl-2,4-pentanediol] to hinder binding of substrate in the active site owing to chelation of the $Mn^{2+}$ cofactor and thereby adopt the disordered open state. Glycerol is often avoided as a cryoprotectant when determining the structures of carbohydrate-active enzymes as it may act as a competitive inhibitor for monosaccharide ligands. Here, it is shown that the use of glycerol as a cryoprotectant can additionally induce significant changes in secondary structure, a phenomenon that could apply to any class of protein.

## 1. Introduction

The use of low temperatures to improve scattering and to minimize radiation damage has become standard in macromolecular crystallography, but usually requires the introduction into the crystal of small molecules designed to prevent the growth of damaging water ice crystals by promoting the formation of amorphous glasses (Garman & Mitchell, 1996; Mitchell & Garman, 1994). Many of these cryoprotectants have been shown to subtly affect intramolecular and intermolecular interactions. For example, a number of common cryoprotectants have been observed to impart rigidity to protein conformations (Bizzarri & Cannistraro, 1992, 1993) and in a few cases to be essential for protein crystallization (Low *et al.*, 1966; Haas, 1968; Petsko, 1975; Rodgers, 1994, 1997; Sousa, 1995), which has been attributed both to the strengthening of intramolecular protein crystal contacts and the solidification of the solvation lattice itself (Scatchard *et al.*, 1938). However, these molecules have largely been treated as innocuous and there has been little systematic study of their potential effect on structure and structure interpretation.

Two of the most popular cryoprotectants are glycerol and MPD [hexylene glycol, (±)-2-methyl-2,4-pentanediol], which have been used in approximately 23 and 15% of structures

deposited in the Protein Data Bank (PDB) and which appear ordered in 8 and 1% of structures, respectively (Kumar *et al.*, 2010). Significantly, there have been many examples in which glycerol has been reported to mimic carbohydrate binding in lectins, glycoside hydrolases and glycosyltransferases (for typical examples, see Luo *et al.*, 1998; Gourdine *et al.*, 2008; Crennell *et al.*, 1993; Gordon *et al.*, 2006; Alfaro *et al.*, 2008).

Glycosyltransferases are anabolic enzymes that synthesize new glycosidic linkages by sequentially adding monosaccharide units from activated donor molecules such as UDP-Gal in a stereospecific and regiospecific manner to target acceptor molecules. Enzymes of the GT-A fold type typically contain regions of highly labile polypeptide required to envelope and thus recognize specific substrates and cofactors (Bourne & Henrissat, 2001), which makes them useful probes for the study of the influence of cryoprotectants on protein structure.

Two of the enzymes that have been most extensively characterized for their movement during substrate recognition are the retaining human ABO(H) blood group A and B glycosyltransferases GTA and GTB (Hearn *et al.*, 1968; Kobata *et al.*, 1968*a*,*b*; Yamamoto *et al.*, 1990; Palcic, 1994; Palcic *et al.*, 2001; Patenaude *et al.*, 2002; Marcus *et al.*, 2003; Nguyen *et al.*, 2003; Lee *et al.*, 2005; Letts *et al.*, 2006, 2007; Persson *et al.*, 2007; Alfaro *et al.*, 2008; Schuman *et al.*, 2010, 2011), which are GT-A fold-type enzymes that contain an $Mn^{2+}$ centre bound by a D*X*D motif. These two enzymes perform the final glycosylation step in the syntheses of the A and B antigens, respectively, by the $\alpha(1\rightarrow3)$ transfer of the corresponding GalNAc or Gal residue to the H-antigen acceptor [HA; $\alpha$-L-Fuc-(1–2)-$\beta$-D-Gal-O*R*, where *R* is a glycoprotein or glycolipid; Hearn *et al.*, 1968; Kobata *et al.*, 1968*a*,*b*; Watkins, 1980; Yamamoto *et al.*, 1990; Patenaude *et al.*, 2002].

GTA and GTB (so named several years prior to the recognition of the GT-A and GT-B fold families) are the most homologous naturally occurring glycosyltransferases known that transfer distinct naturally occurring donor sugars. They show different susceptibility to molecular motion, but differ by only four 'critical' residues out of 354 amino acids: Arg/Gly176, Gly/Ser235, Leu/Met266 and Gly/Ala268 (Yamamoto *et al.*, 1990; Yamamoto & Hakomori, 1990). The kinetic effect of each of these critical residues on donor and acceptor recognition has been studied through the creation of chimeric GTA/GTB enzymes of all 16 permutations (Yamamoto *et al.*, 1990; Alfaro *et al.*, 2008). GTA and GTB chimeras are designated by a four-letter code according to the identity of each of the four critical residues, where AAAA is wild-type GTA, BBBB is wild-type GTB (Seto *et al.*, 1997), ABBB represents the chimeric enzyme GTB/G176R, ABBA represents GTB/G176R/A268G and AABB represents GTB/G176R/S235G.

Given that the wild-type and chimeric enzymes all utilize the H-antigen acceptor, it was initially thought that the four critical residues would be involved exclusively in donor recognition; however, analysis of enzyme chimeras showed that only the last two of the four residues (Leu/Met266 and Gly/Ala268) have a significant impact on donor recognition (Yamamoto & Hakomori, 1990; Seto *et al.*, 1997, 1999; Kamath *et al.*, 1999; Marcus *et al.*, 2003; Persson *et al.*, 2007). The second critical residue (Gly/Ser235) lies proximal to the H-antigen acceptor and may have an impact on acceptor recognition (Patenaude *et al.*, 2002; Nguyen *et al.*, 2003; Letts *et al.*, 2006). The identity of the first critical residue (Arg/Gly176) is known to affect the rate of enzymatic turnover and it lies at the N-terminus of a labile polypeptide loop consisting of residues 176–188 (Seto *et al.*, 1997; Lee *et al.*, 2005; Alfaro *et al.*, 2008). High-resolution structural studies have revealed that, again like other glycosyltransferases, GTA and GTB also contain a mobile C-terminal tail (residues 346–354), which in conjunction with the internal loop serves to recognize substrates and close about the active site during catalysis (Alfaro *et al.*, 2008).

Interestingly, each of the four residues has been observed to influence the combined lability of the loops in conjunction with the binding of substrate analogues, yet only Arg/Gly176 actually lies within a mobile polypeptide loop. Three distinct conformational states of the mobile loops have been associated with the stepwise binding of substrate (Alfaro *et al.*, 2008). In the absence of substrate, these enzymes generally assume an 'open' conformation in which the loops shift to expose the active site and generally display higher levels of disorder. Upon the addition of donor or simple UDP, some of the chimeras adopt a 'semi-closed' conformation in which the internal mobile loop is more ordered, while the C-terminal tail remains disordered. Addition of an inactive 3-deoxygalactose H-antigen acceptor analogue [DA; $\alpha$-L-Fuc*p*-(1$\rightarrow$2)-$\beta$-D-(3-deoxy)-Gal] induces some of the chimeras to adopt a 'closed' conformation in which both loops become ordered where they can interact with the substrates (Alfaro *et al.*, 2008). The degree to which an enzyme adopts an open or closed conformation has been observed to vary systematically across the chimeras (Alfaro *et al.*, 2008).

One structure observed in a previous study (Alfaro *et al.*, 2008) that broke this trend was the ABBB chimera in complex with UDP and HA, in which the enzyme displayed the closed conformation despite lacking an intact UDP-Gal donor. The presence of a glycerol molecule derived from the cryoprotectant in the donor-binding site led to speculation that its interactions with Arg188 of the labile loop mimicked those made by the galactosyl moiety of the UDP-Gal donor substrate to induce the closed conformation, which prompted this investigation into the influence of cryoprotectants on protein structure.

Here, we use the human ABO(H) blood group glycosyltransferases to probe the effects of the common cryoprotectants glycerol and MPD on the conformations of the mobile polypeptide loops in GTA and GTB as well as the chimeric constructs ABBA, AABB and ABBB.

## 2. Methods and materials

### 2.1. Crystallization and cryogenic conditions

Protein production, purification and crystallization were as described previously (Alfaro *et al.*, 2008). Briefly, 10–

**Table 1**
Data-collection and refinement statistics for the evaluated ABO glycosyltransferases.

All crystals were soaked with UDP and galactose and belonged to space group $C222_1$, with unit-cell parameters in the ranges $a$ = 52.4–52.7, $b$ = 149.1–149.9, $c$ = 78.8–80.0 Å. Values in parentheses are for the highest resolution shell.

| | AAAA (GTA) | | AABB | | ABBA | | ABBB | | BBBB (GTB) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | MPD | Glycerol | MPD | Glycerol | MPD | Glycerol | MPD | Glycerol | MPD | Glycerol |
| Resolution (Å) | 20–1.86 | 20–1.49 | 20–1.43 | 20–1.45 | 20–1.47 | 20–1.42 | 20–1.49 | 20–1.50 | 20–1.55 | 20–1.90 |
| $R_{merge}$† (%) | 5.5 (34.0) | 4.3 (36.4) | 3.1 (30.1) | 4.2 (37.1) | 5.8 (29.1) | 3.0 (29.5) | 3.2 (26.5) | 3.0 (25.2) | 5.5 (42.9) | 5.9 (32.5) |
| Completeness (%) | 93.8 (96.2) | 96.9 (97.8) | 95.6 (89.2) | 96.3 (96.0) | 96.9 (97.8) | 98.2 (95.5) | 95.5 (93.5) | 96.9 (94.7) | 96.9 (94.7) | 95.2 (98.2) |
| Unique reflections | 25287 | 50161 | 55579 | 53446 | 52036 | 58371 | 49296 | 48983 | 44205 | 23770 |
| Refinement | | | | | | | | | | |
| $R_{work}$ (%) | 19.0 | 19.0 | 20.1 | 19.8 | 18.9 | 20.5 | 19.4 | 20.4 | 19.1 | 17.8 |
| $R_{free}$ (%) | 23.9 | 21.9 | 21.8 | 22.7 | 22.2 | 22.8 | 22.1 | 23.3 | 22.7 | 22.8 |
| No. of waters | 282 | 278 | 258 | 260 | 282 | 319 | 267 | 238 | 224 | 185 |
| R.m.s. bonds‡ (Å) | 0.026 | 0.029 | 0.030 | 0.029 | 0.029 | 0.032 | 0.032 | 0.030 | 0.028 | 0.026 |
| R.m.s. angles‡ (°) | 2.036 | 2.256 | 2.393 | 2.380 | 2.384 | 2.491 | 2.562 | 2.365 | 2.267 | 1.987 |
| $Mn^{2+}$ and $PO_4$ occupancy (%) | 100 | 100 | 100 | 100 | 40 | 100 | 50 | 100 | 40 | 40 |
| $B$ factors (Å$^2$) | | | | | | | | | | |
| $Mn^{2+}$ | 33.4 | 18.5 | 19.3 | 14.7 | 29.5 | 25.6 | 18.5 | 19.8 | 25.3 | 27.4 |
| D*X*D§ | 27.0 | 17.3 | 19.0 | 15.7 | 21.5 | 20.9 | 22.9 | 17.4 | 24.9 | 26.8 |
| $PO_4$¶ | 35.3 | 20.6 | 25.8 | 17.6 | 25.7 | 25.8 | 22.7 | 19.4 | 26.8 | 28.2 |
| Glycerol†† | N/A | N/A | N/A | 30.7 | N/A | 36.7 | N/A | 37.9 | N/A | N/A |
| PDB code | 3sxg | 3sxe | 3sx5 | 3sx3 | 3sx8 | 3sx7 | 3sxb | 3sxa | 3sxd | 3sxc |

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. ‡ Root-mean-square deviations. § The D*X*D $B$ factor is an average over the four Asp211 and Asp213 $O^{\delta}$ atoms which coordinate $Mn^{2+}$. ¶ The $PO_4$ $B$ factor is an average over the phosphate $\alpha$ and $\beta$ atoms. †† The glycerol $B$ factor is an average over all atoms.

40 mg ml$^{-1}$ protein was crystallized in 10–15 µl hanging drops containing 1% polyethylene glycol (PEG) 4000, 4.5–5% MPD, 100 m*M* ammonium sulfate, 70 m*M* sodium chloride, 50 m*M* *N*-(2-acetamido)-2-iminodiacetic acid (ADA) pH 7.5, 30 m*M* sodium acetate buffer pH 4.6 and 5 m*M* $MnCl_2$ over a reservoir consisting of 3.7% PEG 4000, 7% MPD, 0.3 *M* ammonium sulfate, 0.25 *M* sodium chloride, 0.2 *M* ADA and 0.1 *M* sodium acetate for a period of 5–10 d at 277 K.

Before ligand soaking, crystals were washed with mother liquor containing 15% cryoprotectant (MPD or glycerol). Ligand soaks contained final concentrations of 40–50 m*M* UDP, 100–150 m*M* galactose and 10 m*M* $MnCl_2$ and all substrates were added incrementally over a period of hours to minimize crystal fracture. Even so, more than half of the crystals shattered during the soaking stage. Before cooling the crystals for data collection, the concentration of the cryoprotectant was increased to 30% glycerol or 20% MPD.

### 2.2. X-ray diffraction data collection, structure determination and refinement

X-ray diffraction data for all crystals were collected at 113 K using a CryoStream 700 crystal cooler. Data were collected on a Rigaku R-AXIS IV$^{++}$ area detector at a distance of 72 mm with exposure times of between 4.0 and 7.0 min for 0.5° oscillations. X-rays were produced by a MicroMax-002 generator (Rigaku/MSC, College Station, Texas, USA) coupled to Osmic 9 'Blue' confocal X-ray mirrors at a power level of 30 W (Osmic, Auburn Hills, Michigan,

| Enzyme | Donor | Overall UDP quality | UDP/$Mn^{2+}$ occupancy (%) | Internal loop 176 181 186 191 196 | C-terminus 346 351 |
|---|---|---|---|---|---|
| Ordered | | | | | |
| AAAA + MPD | Water | Weak $PO_4$ | 100 | EV RAYKR WQDVS MRRME MISdf CERR | VP KNHQA VRNP |
| AAAA + glycerol | Gal | Excellent | 100 | EV RAYKR WQDVS MRRME MISdf CERR | VP KNHQA VRNP |
| Cryoprotectant-dependent | | | | | |
| AABB + MPD | Water | Excellent | 100 | EV RAYKR WQDVS MRRME MISdf CERR | VP KNHQa vrnp |
| AABB + glycerol | Glycerol | Excellent | 100 | EV RAYKR WQDVS MRRME MISdf CERR | VP KNHQA VRNP |
| ABBA + MPD | Water | Fragmented | 40 | EV RAykr wqdvs MRRME MISdf CERR | VP knhqa vrnp |
| ABBA + glycerol | Glycerol | Excellent | 100 | EV RAYKR WQDVS MRRME MISdf CERR | VP KNHQa vrnp |
| ABBB + MPD | Water | Excellent | 50 | EV RAykr wqdvs MRRME MISdf CERR | VP Knhqa vrnp |
| ABBB + glycerol | Glycerol | Excellent | 100 | EV RAYKR WQDVS MRRME MISdf CERR | VP KNHQA VRNP |
| Disordered | | | | | |
| BBBB + MPD | Water | Fragmented | 40 | EV Gaykr wqdvs mRRME MISdf CERR | VP knhqa vrnp |
| BBBB + glycerol | Water | Weak $PO_4$ | 40 | EV Gaykr wqdvs MRRME MISdf CERR | VP KNhqa vrnp |

**Figure 1**
Cryoprotectant-dependent loop ordering in ABO glycosyltransferases. All complexes include UDP and galactose. Black single-letter amino-acid codes correspond to electron density for main-chain and side-chain atoms, green corresponds to electron density for main-chain atoms only and red corresponds to weak electron density for main-chain and side-chain atoms. Residues with single-letter amino-acid codes in lower case were not included in the refined models.

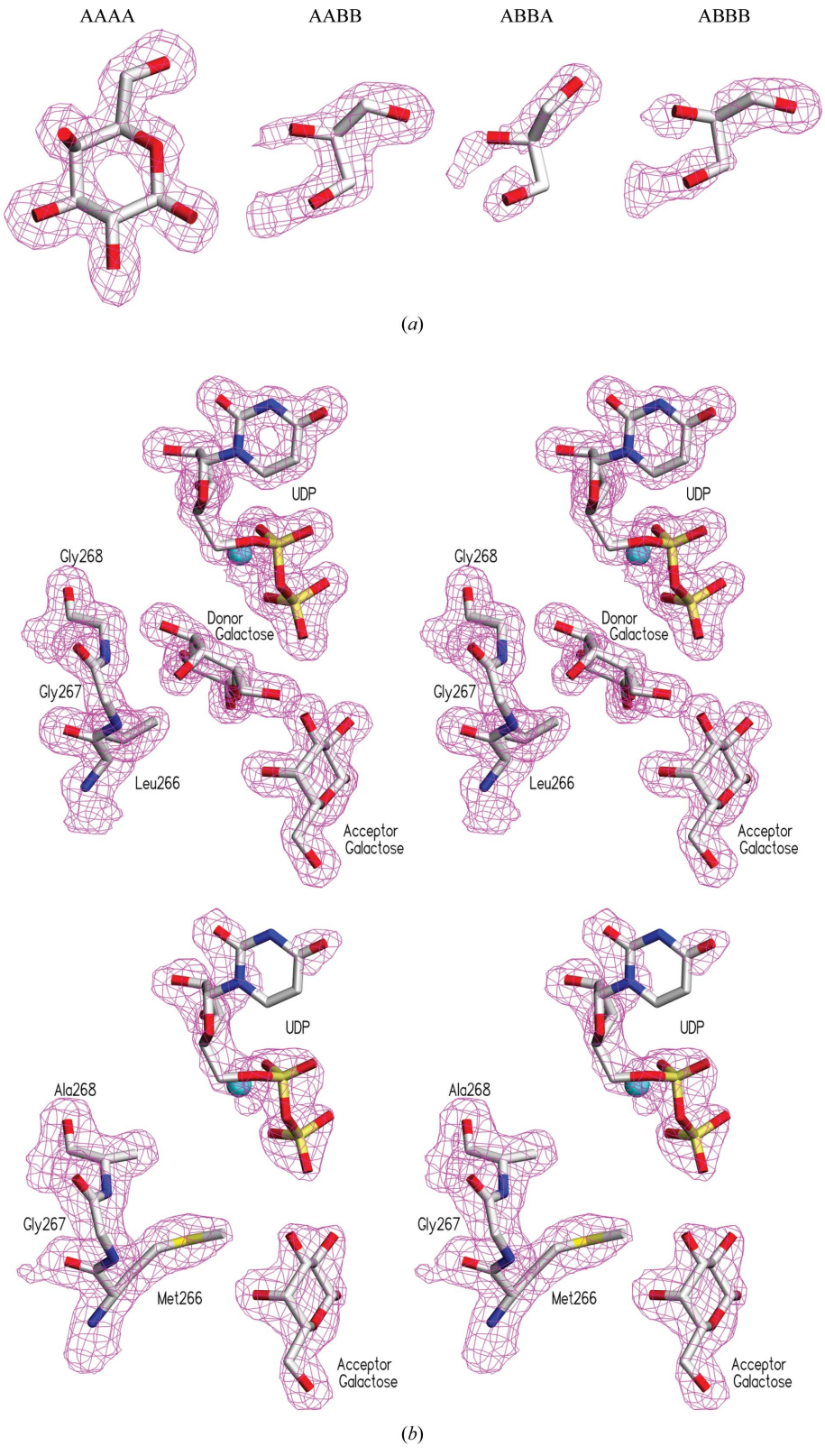


**Figure 2**
$2F_o - F_c$ electron-density maps contoured at $1\sigma$ for structures soaked in glycerol showing (*a*) the donor-site ligands galactose in GTA (AAAA) and glycerol molecules in AABB, ABBA and ABBB and (*b*) stereoviews of the active sites of GTA (top) and GTB (bottom) showing important amino-acid residues, the location of the UDP and acceptor-site galactose molecules and, in GTA, the donor-site galactose.

USA). The data were scaled, averaged and integrated using *d*TREK* (Pflugrath, 1999).

As expected, the structures proved to be nearly isomorphous with the previously determined structures of wild-type GTA and GTB (PDB entries 1lz0 and 1lz7, respectively; Patenaude *et al.*, 2002); they were refined using the *CCP*4 module *REFMAC*5 (Murshudov *et al.*, 2011; Winn *et al.*, 2011) and modelled using *SetoRibbon* (Evans, unpublished work).

In many of the structures it was noted that the isotropic temperature factors of the $Mn^{2+}$ cofactors and UDP substrates were comparable to those of the side-chain O atoms of the D*X*D motifs with which they were associated, especially in those crystals that used glycerol as the cryoprotectant; however, in some of the crystals in which MPD was the cryoprotectant the temperature factors of $Mn^{2+}$ and UDP were initially much higher than those of the side chains of the residues making up their cognate D*X*D motif, indicating that the $Mn^{2+}$ and UDP phosphoryl groups in these structures had partial occupancies. The occupancies of the two groups were adjusted together in increments of 10% until their temperature factors were of the same magnitude as those of the side-chain O atoms of their corresponding D*X*D motifs (Table 1, Fig. 1).

In the three chimeric structures where glycerol was the cryoprotectant, the temperature factors were somewhat higher than those of the protein (Table 1), which was expected given the number of ways that glycerol can orient itself in a glycosyltransferase active site.

## 3. Results

### 3.1. Glycosyltransferase order

Data-collection and refinement statistics for the five enzymes GTA, GTB, ABBB, AABB and ABBA in complex with UDP and Gal are provided in Table 1. The observed electron density surrounding the internal loop (residues 176–188) and the C-terminal loop (residues 346–354) for all complexes is detailed in Fig. 1. With the exception of these two disordered

loops, excellent density for the entire length of the polypeptide is seen in all structures.

There are no significant changes in conformation in the structure of either wild-type GTA (which shows the closed state) or GTB (which shows the open state) with a change in the cryoprotectant; however, order is increased in both the internal and C-terminal loops for all three chimeric enzymes (AABB, ABBA and ABBB) when the cryoprotectant is changed from MPD to glycerol (Fig. 1). This difference is most pronounced in ABBB, in which the loops switch from totally disordered to nearly completely ordered.

### 3.2. Substrate binding

Galactose was used as an acceptor analogue and UDP as a donor analogue in all soaking trials. In all ten structures, a galactose molecule is present in the acceptor-binding site and UDP is bound in the donor-binding site regardless of the identity of the cryoprotectant (although with lower occupancy in structures containing MPD, as above). A second galactose molecule is present in the donor-binding site of GTA when using glycerol as cryoprotectant (Fig. 2). Ordered MPD molecules are never observed; however, a glycerol molecule is present in the donor-binding site of all three chimeric enzymes (Fig. 2). As observed in previous structures, a single $Mn^{2+}$ divalent cation with octahedral geometry is observed coordinated to two UDP phosphate O atoms, one O atom from the carboxylic acid moiety of the D*X*D residue Asp211 and two from that of Asp213. The sixth position of the octahedral coordination is usually occupied by a solvent molecule in those structures containing glycerol and is unoccupied in those structures containing MPD.

## 4. Discussion

### 4.1. Recognition of galactose as an acceptor

Although the H-antigen disaccharide [$\alpha$-L-Fuc-(1–2)-$\beta$-D-Gal-O*R*] has been shown to be the minimum natural acceptor recognized by both GTA and GTB (Watkins, 1991; Lowary & Hindsgaul, 1993), the majority of interactions with the enzyme are made through the galactose residue (Patenaude *et al.*, 2002; Letts *et al.*, 2006; Alfaro *et al.*, 2008; Schwyzer & Hill, 1977) and both enzymes are capable of catalyzing transfer to galactose monosaccharide, albeit at only about 2% and 0.5% of the normal rates (Schwyzer & Hill, 1977; Paulson & Colley, 1989; Watkins, 1980; Letts *et al.*, 2006). The importance to catalysis of the L-fucose residue lies in its specific recognition by amino-acid residues on the C-terminal mobile polypeptide loop (Letts *et al.*, 2006; Alfaro *et al.*, 2008) and thus galactose

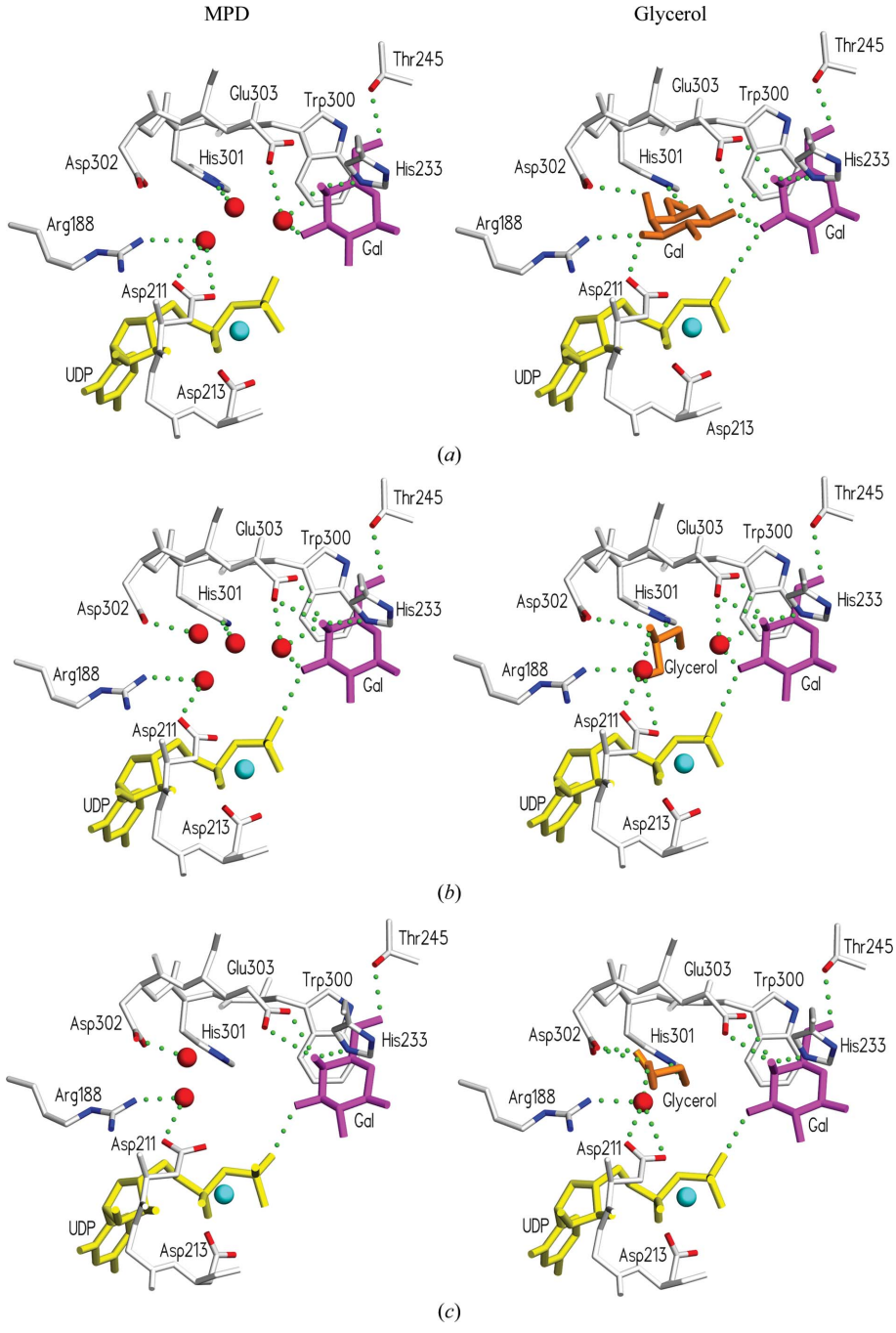


**Figure 3**
The active sites of GTA, GTB and chimeric enzymes in complex with UDP, with structures from crystals soaked in MPD shown on the left and those soaked in glycerol on the right: (*a*) GTA (AAAA), (*b*) AABB, (*c*) ABBA. Donor-site substrates: yellow, UDP; orange, galactose or glycerol; cyan, $Mn^{2+}$. Acceptor-site substrates: magenta, acceptor galactose or 3-deoxy-acceptor analogue.

was used in this study to minimize the influence of the acceptor on loop ordering.

### 4.2. The enzymes display sequence-dependent and cryo-dependent ordering

Despite the fact that three of the four critical residues at which GTA and GTB differ do not form part of the mobile polypeptide loops surrounding the active site, the level of order in these loops is clearly correlated to the identity of these residues. Specifically, these loops show low levels of disorder in GTA, significant levels of disorder in GTB and intermediate levels of disorder in the chimeric enzymes (Alfaro *et al.*, 2008). This differential behaviour has been correlated to crowding in the active site, where GTA has a much larger active site than GTB and so is better able to allow loop ordering around substrate (Alfaro *et al.*, 2008).

In the present study, order in both mobile loops in the wild-type enzymes is largely independent of the cryoprotectant, where GTB shows large areas of disorder and GTA shows very little disorder in both MPD and glycerol (Fig. 1). However, the structures of chimeric enzymes display a clear dependence on the choice of cryoprotectant as well as the identities of the critical amino-acid residues.

While the chimeric enzyme crystals soaked in glycerol show almost GTA-like levels of loop ordering, the corresponding enzyme crystals soaked in MPD show various levels of disorder from GTA-like to GTB-like (Fig. 1). The change is most dramatic in the mobile loops of chimera ABBB, which show almost complete order in glycerol and almost complete disorder in MPD. In contrast, the chimera AABB shows almost no change in structure upon change of cryoprotectant, indicating that the role played by Arg/Gly176 in loop ordering is modulated by the identity of Gly/Ser235.

Interestingly, the location of Gly/Ser235 is at least 10 Å away from residues on the C-terminal loop and more than 14 Å away from residues on the internal mobile loop, so the contribution of residue 235 cannot be simply steric but must involve interactions with substrate or cryoprotectant or both.

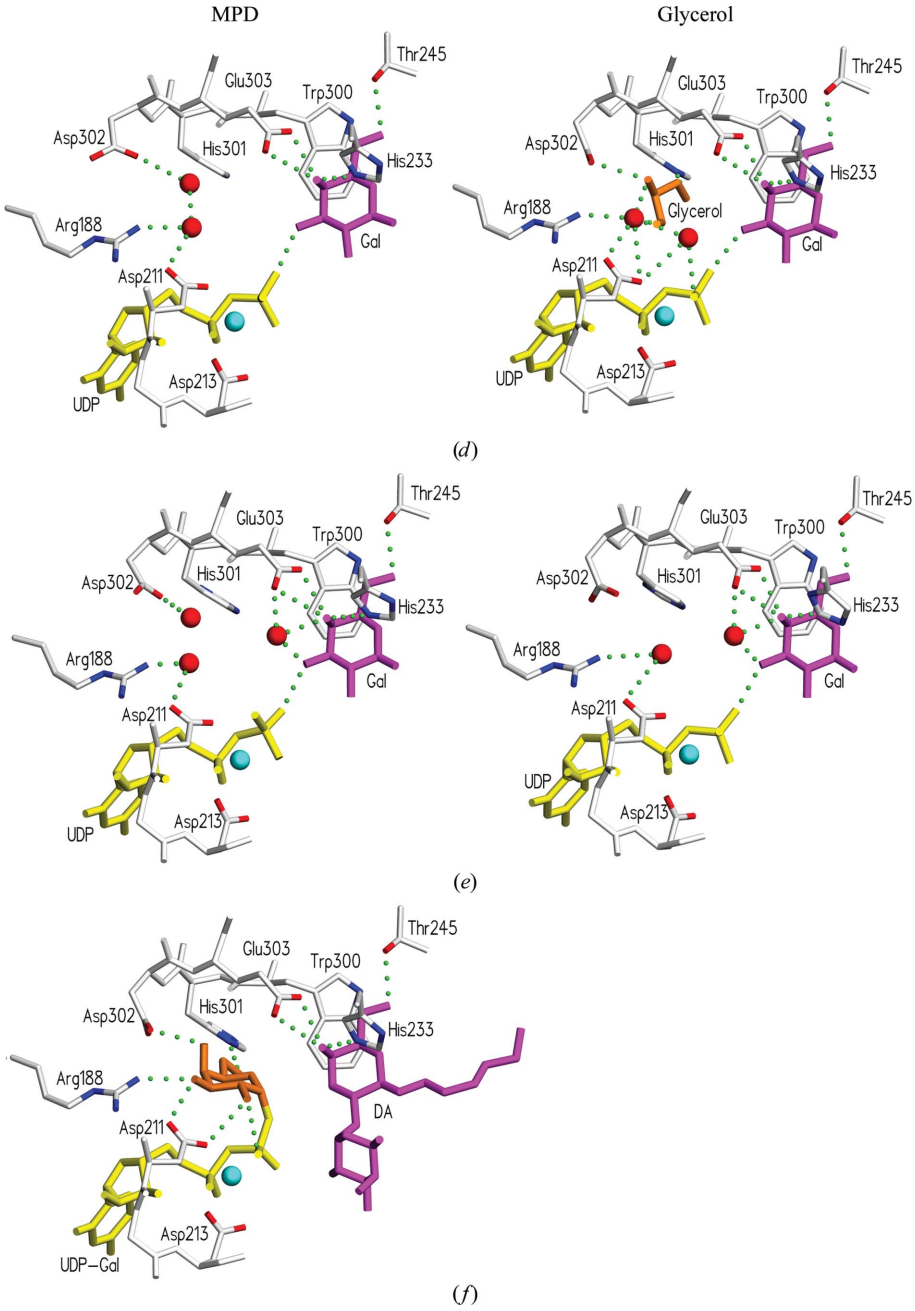


**Figure 3 (continued)**
The active sites of GTA, GTB and chimeric enzymes in complex with UDP, with structures from crystals soaked in MPD shown on the left and those soaked in glycerol on the right: (*d*) ABBB, (*e*) GTB (BBBB). (*f*) For comparison, the structure of AABB in complex with UDP-Gal and a 3-deoxy-acceptor analogue (Alfaro *et al.*, 2008) in the closed conformation determined using MPD as a cryoprotectant is also shown (PDB entry 2rj7). The donor galactosyl moiety forms specific hydrogen bonds to Arg188, Asp211, His301 and Asp302, similar to the interactions made by the glycerol cryoprotectant

### 4.3. Glycerol can mimic the donor sugar to induce the closed state

In general, the closed state can be formed from the open state by the sequential addition of substrate (Alfaro *et al.*, 2008). When UDP or UDP-donor substrate is added to GTA or GTB it makes specific contacts with the internal mobile loop that force the enzymes from the open state to the semi-closed

state. When the acceptor or acceptor analogue is added it makes other specific contacts with the C-terminal mobile polypeptide loop that force the enzyme from the semi-closed state to the closed state. Contact between the internal mobile loop and the C-terminal loop through a $\pi$-orbital stacking interaction stabilizes the closed state.

The induction of the semi-closed or closed states in the chimeric enzymes when cooled in glycerol but not MPD can be attributed to the residence in the donor-binding site of a molecule of glycerol that utilizes many of the same interactions as normally made by the donor sugar. Specifically, the closure of the mobile polypeptide loops to form the catalytically competent conformation has been shown to involve binding of the donor-sugar moiety to Arg188, Asp211, His301 and Asp302 (Fig. 3; Alfaro *et al.*, 2008). The glycerol molecule observed in the active sites of the chimera is able to mimic equivalent interactions to Arg188, Asp211, His301 and Asp302, albeit indirectly through a bridging water molecule in the case of Asp211 and Arg188 (Figs. 3*a*–3*e*, right).

While no cryoprotectant is evident in the corresponding MPD structures, they all have water molecules that form interactions with the same four amino-acid residues; however, they lack the rigid covalent bridge provided by glycerol required for the induction of the closed conformation of the enzyme (Figs. 3*a*–3*e*, left).

### 4.4. MPD can sequester the $Mn^{2+}$ cofactor to inhibit formation of the closed state

The known ability of glycerol to act as an aldose mimic and as a competitive inhibitor in a number of carbohydrate-active enzymes (typical examples can be found in Crennell *et al.*, 1993; Burmeister *et al.*, 1997; Luo *et al.*, 1998; Charnock & Davies, 1999; Vallee *et al.*, 2000; Schmidt *et al.*, 1998; Ng *et al.*, 2002; Gordon *et al.*, 2006; Alfaro *et al.*, 2008; Gourdine *et al.*, 2008) has prompted the introduction of cryoprotectants such as MPD to eliminate competition when substrate-complex structures are sought (Gregoriou *et al.*, 1998; Tsitsanou *et al.*, 1999; Anand *et al.*, 2002; Hammes & Schimmel, 1967). Unfortunately, many molecules similar to MPD are able to form weak coordinate bonds to several metal ions, apparently including the required $Mn^{2+}$ cofactor in GTA and GTB (Carrell *et al.*, 1994; Topcu *et al.*, 2002; Tanase *et al.*, 2000). This effect appears in GTA and GTB as a reduced occupancy of $Mn^{2+}$ (and consequently of UDP) observed in structures utilizing MPD as a cryoprotectant.

In addition to its proposed role as a Lewis acid in catalysis and its interaction with Asp211 and Asp213 of the D*X*D motif (Patenaude *et al.*, 2002; Busch *et al.*, 1998), $Mn^{2+}$ is required to present the UDP-donor phosphates to the enzyme in a conformation suitable to interact with residues in the internal and C-terminal mobile loops and ultimately to form the closed conformation. The use of MPD at concentrations high enough to prove effective for cryoprotection [in this case about 20%(*w*/*v*)] brings it into high excess over the $Mn^{2+}$ concentration, where even a weak ability to chelate the ion will result in lowered ion occupancy in the enzyme, lower UDP coordination and a reduced ability to achieve the closed conformation under conditions otherwise identical to crystals cooled with glycerol (Table 1, Fig. 1).

The lone exceptions are the structures of AABB and AAAA in MPD, which appear to have a fully occupied metal cofactor and may reflect intrinsic stability of these enzymes in the closed state.

### 4.5. GTA displays a galactose molecule in both the donor and acceptor sites

Interestingly, although the conformations of the two wild-type enzymes are not strongly affected by cryoprotectant (GTA always displays the closed conformation and GTB the open conformation), the nature of the cryoprotectant can affect the active site. As noted, the chimeric enzymes soaked in glycerol always contain a glycerol molecule in the donor-binding site, the donor-binding site of GTB is empty and the donor-binding site of GTA contains an ordered galactose molecule (Fig. 3*a*). Interestingly, neither GTA nor GTB protected with MPD show a second galactose in the donor-binding site.

The ability of GTA to stabilize an extra galactose molecule in the donor-binding site can be attributed directly to steric considerations, as the smaller critical residues in GTA provide a more commodious active site. The galactose residue in the donor site displays largely the same interactions as reported for the intact donor UDP-Gal in AABB (Alfaro *et al.*, 2008), involving His301, Asp302, Arg188 and one of two interactions with Asp211. The loss of one Asp211 hydrogen bond skews

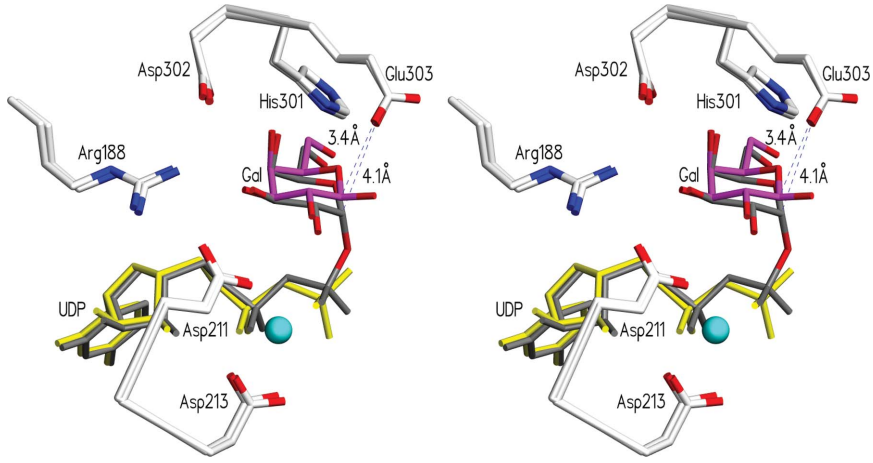


**Figure 4**
Stereoview of the overlap of AABB–UDP-Gal–DA (PDB entry 2rj7; Alfaro *et al.*, 2008) and of GTA–UDP-Gal using glycerol as the cryoprotectant. The modest observed rotation of ~14° about C4 of the galactose molecule results in C1 moving 0.7 Å to bring it within 3.5 Å of Glu303, which is consistent with the postulated role of this amino-acid residue as a nucleophile. Atoms are coloured by element, with carbon white, nitrogen blue, oxygen red and manganese cyan. The donor galactose residue is shown in magenta and UDP-Gal is shown in dark grey.

galactose ∼14° about C4 such that C1 lies ∼0.7 Å closer to the enzyme's postulated nucleophile Glu303 (Fig. 4).

There has been speculation that Glu303 may not be positioned close enough to O3 of the UDP-GalNAc donor to be an appropriate nucleophile (Zhang *et al.*, 2003; Molina *et al.*, 2007); however, this small rotation of about 14° brings the active C1 within 3.5 Å of the Glu303 side chain. Interestingly, Gal O1 in the donor-binding site is stabilized as a $\beta$-anomer by Glu303, whereas the configuration about this O atom would be $\alpha$ in the UDP-GalNAc or UDP-Gal donor.

## 5. Conclusions

GTA and GTB are well known for the mobile loops of polypeptide surrounding their active sites and for their ability to use just a few critical amino acids to distinguish between similar substrates to generate the immunologically distinct A and B blood group antigens. The relative ease with which each enzyme or chimera forms the closed state using donor and acceptor analogues was established by Alfaro *et al.* (2008) to be strongly correlated to the identity of the critical amino acids. However, altering these conditions slightly to remove the L-fucose residue on the acceptor and using simple galactose shows that these same amino-acid differences also result in an exquisite sensitivity of the conformations of these loops to the choice of cryoprotectant.

Interestingly, the conformations of the mobile loops in the wild-type enzymes are relatively unaffected by the choice of cryoprotectant; however, in the chimeric enzymes ABBB, ABBA and AABB the loops are ordered in the presence of glycerol and disordered in the presence of MPD. The effects of the cryoprotectants are opposite. Glycerol can be seen to mimic donor-sugar binding and induce formation of the closed state in these enzymes, while MPD appears to chelate the $Mn^{2+}$ cofactor utilized in donor binding and thus inhibit formation of the closed state.

Finally, although the conformations of GTA and GTB (and presumably other GT-A fold-family glycosyltransferases) appear to display higher than average sensitivity to their environment, it is clear that the assignment of conformational changes in a protein upon introduction of substrate or a single point mutation cannot confidently be made in the absence of an exploration of the effect of cryoprotectants.

SVE is the recipient of a senior scholarship from the Michael Smith Foundation for Health Research. SVE gratefully acknowledges the support of the Canadian Institutes of Health Research (MOP 77655).

## References

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.
Anand, K., Pal, D. & Hilgenfeld, R. (2002). *Acta Cryst.* D**58**, 1722–1728.
Bizzarri, A. R. & Cannistraro, S. (1992). *Biophys. Chem.* **42**, 79–85.
Bizzarri, A. R. & Cannistraro, S. (1993). *Eur. Biophys. J.* **22**, 259–267.
Bourne, Y. & Henrissat, B. (2001). *Curr. Opin. Struct. Biol.* **11**, 593–600.
Burmeister, W. P., Cottaz, S., Driguez, H., Iori, R., Palmieri, S. & Henrissat, B. (1997). *Structure*, **5**, 663–675.
Busch, C., Hofmann, F., Selzer, J., Munro, S., Jeckel, D. & Aktories, K. (1998). *J. Biol. Chem.* **273**, 19566–19572.
Carrell, H. L., Hoier, H. & Glusker, J. P. (1994). *Acta Cryst.* D**50**, 113–123.
Charnock, S. J. & Davies, G. J. (1999). *Biochemistry*, **38**, 6380–6385.
Crennell, S. J., Garman, E. F., Laver, W. G., Vimr, E. R. & Taylor, G. L. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 9852–9856.
Garman, E. F. & Mitchell, E. P. (1996). *J. Appl. Cryst.* **29**, 584–587.
Gordon, R. D., Sivarajah, P., Satkunarajah, M., Ma, D., Tarling, C. A., Vizitiu, D., Withers, S. G. & Rini, J. M. (2006). *J. Mol. Biol.* **360**, 67–79.
Gourdine, J. P., Cioci, G., Miguet, L., Unverzagt, C., Silva, D. V., Varrot, A., Gautier, C., Smith-Ravin, E. J. & Imberty, A. (2008). *J. Biol. Chem.* **283**, 30112–30120.
Gregoriou, M., Noble, M. E., Watson, K. A., Garman, E. F., Krulle, T. M., de la Fuente, C., Fleet, G. W., Oikonomakos, N. G. & Johnson, L. N. (1998). *Protein Sci.* **7**, 915–927.
Haas, D. J. (1968). *Biophys. J.* **8**, 549–555.
Hammes, G. G. & Schimmel, P. R. (1967). *J. Am. Chem. Soc.* **89**, 442–446.
Hearn, V. M., Smith, Z. G. & Watkins, W. M. (1968). *Biochem. J.* **109**, 315–317.
Kamath, V. P., Seto, N. O., Compston, C. A., Hindsgaul, O. & Palcic, M. M. (1999). *Glycoconj. J.* **16**, 599–606.
Kobata, A., Grollman, E. F. & Ginsburg, V. (1968*a*). *Arch. Biochem. Biophys.* **124**, 609–612.
Kobata, A., Grollman, E. F. & Ginsburg, V. (1968*b*). *Biochem. Biophys. Res. Commun.* **32**, 272–277.
Kumar, A., Chiu, H.-J., Axelrod, H. L., Morse, A., Elsliger, M.-A., Wilson, I. A. & Deacon, A. (2010). *Acta Cryst.* F**66**, 1309–1316.
Lee, H. J., Barry, C. H., Borisova, S. N., Seto, N. O., 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.
Letts, J. A., Rose, N. L., Fang, Y. R., Barry, C. H., Borisova, S. N., Seto, N. O., Palcic, M. M. & Evans, S. V. (2006). *J. Biol. Chem.* **281**, 3625–3632.
Low, B. W., Chen, C. C. H., Berger, J. E., Singman, L. & Pletcher, J. F. (1966). *Proc. Natl Acad. Sci. USA*, **56**, 1746–1750.
Lowary, T. L. & Hindsgaul, O. (1993). *Carbohydr. Res.* **249**, 163–195.
Luo, Y., Li, S.-C., Chou, M.-Y., Li, Y.-T. & Luo, M. (1998). *Structure*, **6**, 521–530.
Marcus, S. L., Polakowski, R., Seto, N. O., Leinala, E., Borisova, S., Blancher, A., Roubinet, F., Evans, S. V. & Palcic, M. M. (2003). *J. Biol. Chem.* **278**, 12403–12405.
Mitchell, E. P. & Garman, E. F. (1994). *J. Appl. Cryst.* **27**, 1070–1074.
Molina, P., Knegtel, R. M. & Macher, B. A. (2007). *Biochim. Biophys. Acta*, **1770**, 1266–1273.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Ng, K. K.-S., Kolatkar, A. R., Park-Snyder, S., Feinberg, H., Clark, D. A., Drickamer, K. & Weis, W. I. (2002). *J. Biol. Chem.* **277**, 16088–16095.
Nguyen, H. P., Seto, N. O., Cai, Y., Leinala, E. K., Borisova, S. N., Palcic, M. M. & Evans, S. V. (2003). *J. Biol. Chem.* **278**, 49191–49195.
Palcic, M. M. (1994). *Methods Enzymol.* **230**, 300–316.
Palcic, M. M., Seto, N. O. & Hindsgaul, O. (2001). *Transfus. Med.* **11**, 315–323.
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.

Paulson, J. C. & Colley, K. J. (1989). *J. Biol. Chem.* **264**, 17615–17618.
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.
Petsko, G. A. (1975). *J. Mol. Biol.* **96**, 381–392.
Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.
Rodgers, D. W. (1994). *Structure*, **2**, 1135–1140.
Rodgers, D. W. (1997). *Methods Enzymol.* **276**, 183–203.
Scatchard, G., Hamer, W. J. & Wood, S. E. (1938). *J. Am. Chem. Soc.* **60**, 3061–3070.
Schmidt, A., Schlacher, A., Steiner, W., Schwab, H. & Kratky, C. (1998). *Protein Sci.* **7**, 2081–2088.
Schuman, B., Fisher, S. Z., Kovalevsky, A., Borisova, S. N., Palcic, M. M., Coates, L., Langan, P. & Evans, S. V. (2011). *Acta Cryst.* F**67**, 258–262.
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.
Schwyzer, M. & Hill, R. L. (1977). *J. Biol. Chem.* **252**, 2338–2345.
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.
Seto, N. O., Palcic, M. M., Compston, C. A., Li, H., Bundle, D. R. & Narang, S. A. (1997). *J. Biol. Chem.* **272**, 14133–14138.
Sousa, R. (1995). *Acta Cryst.* D**51**, 271–277.
Tanase, T., Tamakoshi, S., Doi, M., Mikuriya, M., Sakurai, H. & Yano, S. (2000). *Inorg. Chem.* **39**, 692–704.
Topcu, Y., Andac, O., Yilmaz, V. T. & Harrison, W. T. A. (2002). *J. Mol. Struct.* **610**, 99–103.
Tsitsanou, K. E., Oikonomakos, N. G., Zographos, S. E., Skamnaki, V. T., Gregoriou, M., Watson, K. A., Johnson, L. N. & Fleet, G. W. (1999). *Protein Sci.* **8**, 741–749.
Vallee, F., Karaveg, K., Herscovics, A., Moremen, K. W. & Howell, P. L. (2000). *J. Biol. Chem.* **275**, 41287–41298.
Watkins, W. M. (1980). *Adv. Hum. Genet.* **10**, 1–136.
Watkins, W. M. (1991). *Pure Appl. Chem.* **63**, 561–568.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Yamamoto, F., Clausen, H., White, T., Marken, J. & Hakomori, S. (1990). *Nature (London)*, **345**, 229–233.
Yamamoto, F. & Hakomori, S. (1990). *J. Biol. Chem.* **265**, 19257–19262.
Zhang, Y., Swaminathan, G. J., Deshpande, A., Boix, E., Natesh, R., Xie, Z., Acharya, K. R. & Brew, K. (2003). *Biochemistry*, **42**, 13512–13521.