Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure analysis of endosialidase NF at 0.98 Å resolution

Eike C. Schulz,[a] Piotr Neumann,[a] Rita Gerardy-Schahn,[b] George M. Sheldrick[c] and Ralf Ficner[a]*

[a]Abteilung für Molekulare Strukturbiologie, Institut für Mikrobiologie und Genetik, Georg-August-Universität Göttingen, Justus-von-Liebig-Weg 11, 37077 Göttingen, Germany, [b]Institut für Zelluläre Chemie, Medizinische Hochschule Hannover, Carl-Neuberg-Strasse 1, 30625 Hannover, Germany, and [c]Abteilung für Stukturchemie, Universität Göttingen, Tammanstrasse 4, 37077 Göttingen, Germany

Correspondence e-mail: rficner@gwdg.de

Received 18 September 2009
Accepted 16 November 2009

PDB Reference:
endosialidase NF, 3ju4.

Endosialidase NF (endoNF) is a bacteriophage-derived endosialidase that specifically degrades α-2,8-linked polysialic acid. The structure of a new crystal form of endoNF in complex with sialic acid has been refined at 0.98 Å resolution. The 210 kDa homotrimeric multi-domain enzyme displays outstanding stability and resistance to SDS. Even at atomic resolution, only a minor fraction of side chains possess alternative conformations. However, multiple conformations of an active-site residue imply that it has an important catalytic function in the cleavage mechanism of polysialic acid.

1. Introduction

In order to protect themselves against phage infection, neuroinvasive pathogenic bacteria such as *Neisseria meningitidis* and *Escherichia coli* K1 are covered by a thick layer of polysialic acid (polySia; Scholl *et al.*, 2005). However, bacteriophages belonging to the K1 family specifically recognize and degrade the polySia capsule of their host bacteria during the infection process (Gerardy-Schahn *et al.*, 1995; Jakobsson *et al.*, 2007; Long *et al.*, 1995; Machida *et al.*, 2000; Mühlenhoff *et al.*, 2003; Scholl *et al.*, 2001). The phage tailspikes, which exhibit glycosidase (sialidase) activity and specifically catalyze the hydrolysis of α-2,8-linked polySia (Petter & Vimr, 1993), are responsible for the degradation of the bacterial capsule. The structure of the bacteriophage K1F endosialidase (endoNF) has been solved previously at a resolution of 1.9 Å in its apo state and at 2.55 Å resolution as a complex with oligomeric sialic acid (Stummeyer *et al.*, 2005). EndoNF is a homotrimeric multi-domain protein in which every monomer comprises an N-terminal linker domain, a six-bladed β-propeller, a β-barrel domain and a stalk domain, which intertwines into a triple β-helix in the functional protein. This triple β-helix confers a remarkable stability on the protein, making it resistant to sodium dodecyl sulfate (SDS) even at relatively high temperatures (Mühlenhoff *et al.*, 2003; Stummeyer *et al.*, 2005). The catalytic domains of viral, bacterial and eukaryotic exosialidases are commonly six-bladed β-propellers and share conserved active-site conformations and substrate-binding properties (Mühlenhoff *et al.*, 2003; Stummeyer *et al.*, 2005; Burmeister *et al.*, 1992; Buschiazzo *et al.*, 2000; Crennell *et al.*, 1993; Varghese *et al.*, 1983). In endoNF, substrate-binding sites have been observed in the stalk domain and in the β-barrel domain (for details, see Stummeyer *et al.*, 2005).

Table 1
Data-collection and refinement statistics of endoNF.

Values in parentheses are for the highest resolution shell.

Data collection	
Space group	*R*3
Unit-cell parameters (Å, °)	a = 119.02, b = 119.02, c = 175.67, $\alpha = \beta$ = 90.0, γ = 120.0
Resolution (Å)	19.36–0.98 (1.08–0.98)
Wavelength (Å)	0.93400
$R_{\rm merge}$ (%)	7.3 (41.5)
$I/\sigma(I)$	12.4 (3.4)
Completeness (%)	97.0 (99.6)
Reflections (observed)	1712231 (429843)
Reflections (unique)	516832 (134204)
Multiplicity	3.3 (3.2)
Refinement	
Resolution (Å)	19.36–0.98
No. of reflections	506460
$R_{\rm work}/R_{\rm free}$ (%)	11.6/13.3
$R_{\rm work}/R_{\rm free}$ (>4σ) (%)	9.4/11.1
No. of atoms	
Protein	5502
Ligand/ion	27
Water	1654
B factors (Å^2)	
Protein	9.3
Ligand/ion	14.2
Water	32.0
R.m.s. deviations	
Bond lengths (Å)	0.016
Bond angles (°)	0.031
Ramachandran statistics	
Allowed (%)	100
Favoured (%)	96.3
Outliers	None

Despite its similarity to bacterial sialidases, the substrate binding in the active site of endoNF and its catalytic mechanism have not yet been resolved. To understand the structural basis of reaction mechanisms, proteins are often cocrystallized with substrate analogues that cannot be processed. However, a noncleavable polySia analogue is currently not available. In order to gain insight into the substrate-binding properties and to understand the cleavage mechanism of endoNF, the protein was crystallized in the presence of the reaction product Sia_3. A crystal structure of the 210 kDa protein endoNF in complex with sialic acid could be determined at a resolution of 0.98 Å and is currently one of the largest structures refined to this resolution.

2. Methods

2.1. Protein purification and crystallization

Molecular cloning and protein purification of endoNF were performed as described by Stummeyer *et al.* (2005). Purified endoNF (at 10 mg ml^{-1} in 10 m*M* Tris–HCl pH 7.5, 50 m*M* NaCl) and Sia_3 (150 m*M* α-2,8-linked sialic acid; Nacalai Tesque, USA) were directly mixed in a 1:10(*v*:*v*) ratio and incubated for 15 min at 277 K prior to crystallization. Crystals were grown by microseeding at 293 K in sitting-drop vapour-diffusion plates, combining equal volumes of precipitant solution [16%(*w*/*v*) PEG 8000, 0.1 *M* Tris–HCl pH 7.2, 3% 2-propanol] and the prepared protein–Sia_3 solution. Crystals were cryoprotected by soaking them in precipitant solution containing 12%(*v*/*v*) 2,3-butanediol and were flash-frozen in liquid nitrogen prior to data collection.

2.2. Data collection and processing

X-ray data collection was performed at 100 K; diffraction images were collected on the microfocus beamline 29.1 at the ESRF, Grenoble. Diffraction images were indexed and scaled using *XDS* and *XSCALE*, respectively (Kabsch, 1993). Data integration was performed as ten consecutive *XDS* runs, during which refined unit-cell parameters, crystal orientation, mosaicity and beam divergence were used as starting values for the next run. This refinement resulted in an appreciably better output data set in comparison to the data set obtained after only a single *XDS* run.

2.3. Structure refinement

The structure was solved by molecular replacement (with *MOLREP*) using the previously published structure as a search model (PDB code 1v0e; Lebedev *et al.*, 2008). The coordinates were initially refined with reasonable stereochemistry to a resolution of 1.4 Å using *REFMAC* (Murshudov *et al.*, 1997). Subsequently, the resolution was extended to the full resolution range and refinement took place using *SHELXL* (Sheldrick, 2008; Table 1). Additional and disordered residues were manually built into the structure using *Coot* (Emsley & Cowtan, 2004). Alternating steps of anisotropic refinement and minor structure adjustment were performed until the *R* values converged. Molecular images were generated in *PyMOL* (DeLano, 2002) and the ligand interaction was analyzed with *LIGPLOT* (Wallace *et al.*, 1995).

3. Results

EndoNF crystallizes in multiple polymorphic forms. A new crystal form of endoNF could be obtained that belonged to space group *R*3 with one monomer in the asymmetric unit, while a crystal symmetry operation generated the other monomers of the functional homotrimer (Table 1). All residues of the protein as well as four residues of the thrombin-cleavage site are clearly visible in the electron-density maps. After initial refinement, alternate residue conformations, the sialic acid molecule and H atoms were built into the difference map density (Fig. 1).

3.1. Stability and disorder

In comparison to the reported structure, only minor differences in the backbone (with an r.m.s.d. of 0.3 Å calculated for the C$^{\alpha}$ atoms) can be observed, but further residues at the N-terminus representing the thrombin-cleavage site are visible in the high-resolution structure.

Owing to the compact intertwined all-β fold, endoNF forms a multitude of intermonomer and intramonomer interactions between side-chain and main-chain atoms, which are reflected by relatively low *B* factors. Accordingly, only 44 residues (6.5%), which are mainly located on the surface of endoNF,

display alternate conformations. In addition to the N- and C-termini, increased flexibility of the protein can also be observed in the β-barrel domain (residues 428–440 and 485–500), in a confined area of the β-propeller domain (residues 545–550) and in the stalk domain (residues 828–833) (Fig. 2). Analysis of the crystalline assembly shows that these areas are involved in crystal-packing contacts. The N-terminus of one homotrimer is in contact with the β-barrel domain of the next molecule; this is associated with higher B factors in this domain.

3.2. The active site

In the active site, which is located at the centre of the sialidase β-propeller domain (for details, see Stummeyer *et al.*, 2005), the side chains are generally very rigid and with one exception no alternate conformations can be observed (Fig. 3). However, with respect to the previously described structure there are a few conformational changes in the active-site residues. Different side-chain conformations can be determined for His628 and Arg596. In case of Arg596, this leads to a shift of the guanidinum group by 0.9 Å and thus to a loss of the hydrogen bond to Glu581.

The most significant change can be observed for His350. In addition to the previously observed conformation, two further conformations can be observed in the high-resolution crystal structure. Relative to the reported conformation of His350, the imidazole ring of the alternate states is shifted by 2.5–3 Å towards the centre of the active site (Fig. 3).

3.3. Sialic acid-binding site

The electron density clearly shows the presence of a sialic acid molecule in the stalk domain of endoNF (Fig. 1). In comparison to the low-resolution structure, the rotamers of Gln853 and Asn870 are flipped. Furthermore, additional water-mediated hydrogen bonds can be observed between Arg837, Thr846 and the *N*-acetyl group of sialic acid. It is important to notice that the sialic acid molecule is clearly in the β configuration, indicating that the observed sialic acid moiety is the product of a cleavage reaction.

4. Discussion

The high-resolution crystal structure of endoNF yields an electron-density map of outstanding quality that even enables

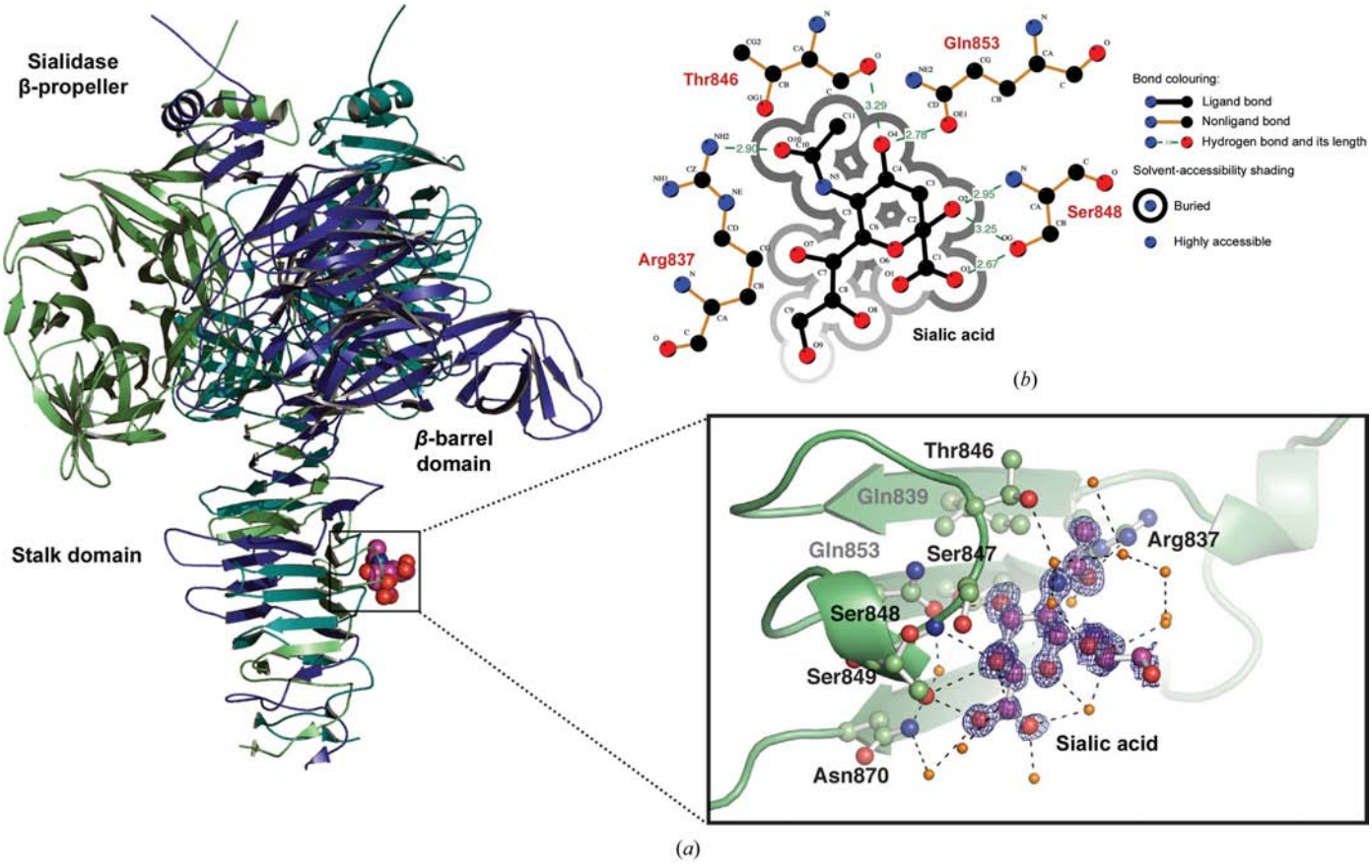

Figure 1
(*a*) Crystal structure of endoNF; the monomers are indicated in different colours. The protein consists of an N-terminal linker domain formed by the only α-helix of the structure, the catalytic sialidase β-propeller domain, into which the β-barrel domain is inserted, and a stalk domain that is built by the intertwining of all three domains into a compact triple β-helix. Sialic acid is shown as magenta spheres at the stalk binding site. The close-up view illustrates the interaction with the stalk domain and displays the electron density contoured at 1σ. Orange spheres correspond to water molecules. (*b*) *LIGPLOT* analysis of sialic acid binding. Green dashed lines indicate all hydrogen bonds. Shades of grey show the solvent-accessibility of the ligand.

the positioning of H atoms in the difference density map (Fig. 3). Despite the high resolution of the data, the different crystal form and altered crystallization conditions, only minor changes relative to the low-resolution structure can be observed in endoNF. The low number of alternate side-chain conformations is consistent with the extraordinary stability of the protein. The close packing in the crystalline assembly results in low protein *B* factors and diffraction to very high resolution. At the biological level, the extreme structural stability can be considered to be an evolutionary adaptation to the extracellular environment. Furthermore, the rigid structure supports the recognition and binding of the extremely flexible polySia ligand.

4.1. PolySia cleavage

Previously, polySia binding has been observed at the β-barrel domain and the stalk domain (Stummeyer *et al.*, 2005). However, the finer details of the ligand geometry were obscured by the available resolution. The high-resolution crystal structure allowed an unambiguous determination of all the interactions between sialic acid and the stalk domain of endoNF (Fig. 1). The sialic acid molecule binds at the same position in both structures, despite their different space groups and crystallization conditions. Most importantly, the observed sialic acid is a single residue in the β configuration. Since the oligomeric reaction product Sia_3 used for cocrystallization was in the α configuration, this clearly indicates that a cleavage reaction has occurred. Although there are three potential binding sites in endoNF (Stummeyer *et al.*, 2005), only two were occupied in the low-resolution structure. Despite the high concentration of Sia_3 in the crystallization conditions, one sialic acid could be identified at the binding site of the stalk domain but none could be identified in the β-barrel domain or the β-propeller domain. It may be that the long incubation time led to a further reaction of catalytically active endoNF that cleaved its major reaction product Sia_3 (Morley *et al.*, 2009; Pelkonen *et al.*, 1989; Schwarzer *et al.*, 2009). However, significant difference density was observed in the active site, suggesting that it is in fact occupied by multiple conformations of sialic acid. This is supported by His350, which is present in three distinct conformations, each of which could resemble a binding mode to a polySia molecule in the active site (Fig. 3). This raises the question of which role His350 plays in the cleavage mechanism of endoNF.

4.2. Mechanistic implications

Mechanistically, glycosidases have been separated into two generally distinct groups: inverting glycosidases and retaining glycosidases (Davies & Henrissat, 1995; Ly & Withers, 1999; Rye & Withers, 2000; Zechel & Withers, 2001). Exo-sialidases, which remove terminal sialic acid moieties from glycosyl conjugates, typically belong to the family of retaining glycosidases, which act using a two-step double-displacement mechanism *via* a covalent glycosyl-enzyme intermediate (Amaya *et al.*, 2004; Watts *et al.*, 2006). In contrast, endoNF has recently been described as an inverting sialidase (Morley *et al.*, 2009). In principle, inverting cleavage reactions allow either an S_N2-type reaction or an S_N1-type reaction. An S_N2-type reaction depends on the initial attack of a nucleophile at the anomeric centre, but it cannot be deduced from the current structural data which residue in the active site could act as such a nucleophile. An S_N1-type reaction, on the other hand, requires a general acid to attack the glycosidic bond. Interestingly, the pH profile of endoNF shows that the enzyme achieves maximum activity at pH 5 (Morley *et al.*, 2009). The high-resolution structure revealed that all conserved active-site residue side chains are present in only one conformation. The only prominent exception is His350, which displays three possible side-chain conformations. Since histidines play important catalytic roles in many enzymes, acting as general

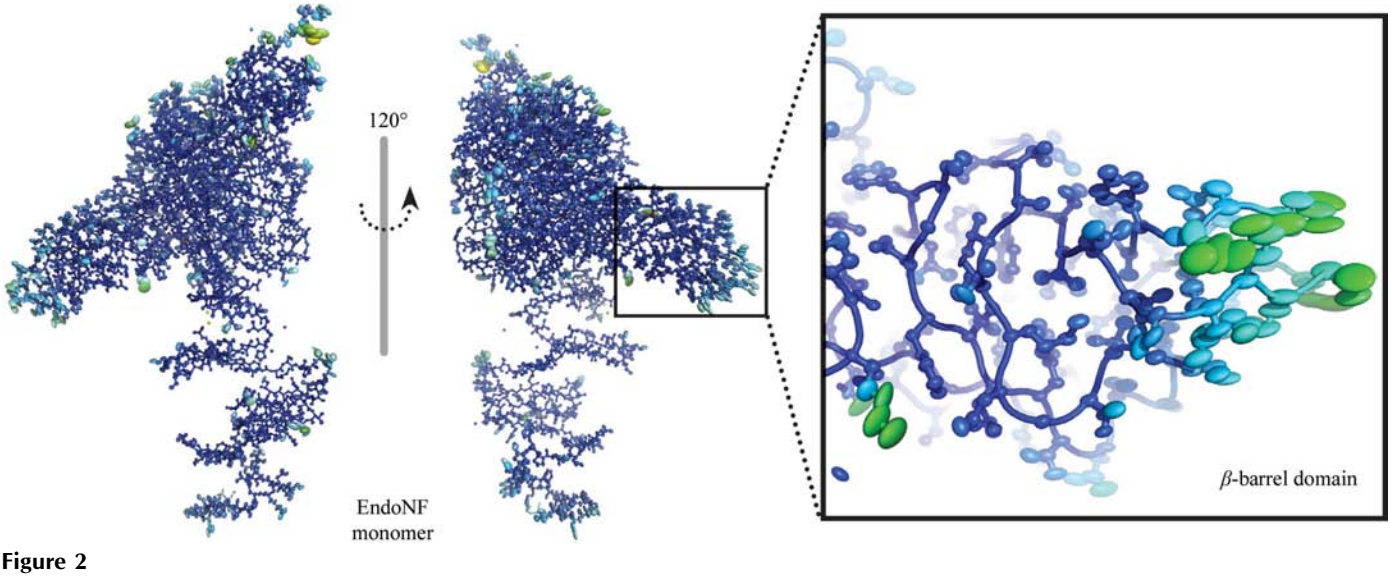

Figure 2
Two endoNF monomers coloured according to *B* factor. The enlarged view shows the anisotropic displacement of surface residues with 50% probability ellipsoids in the β-barrel domain to describe the anharmonic motion.

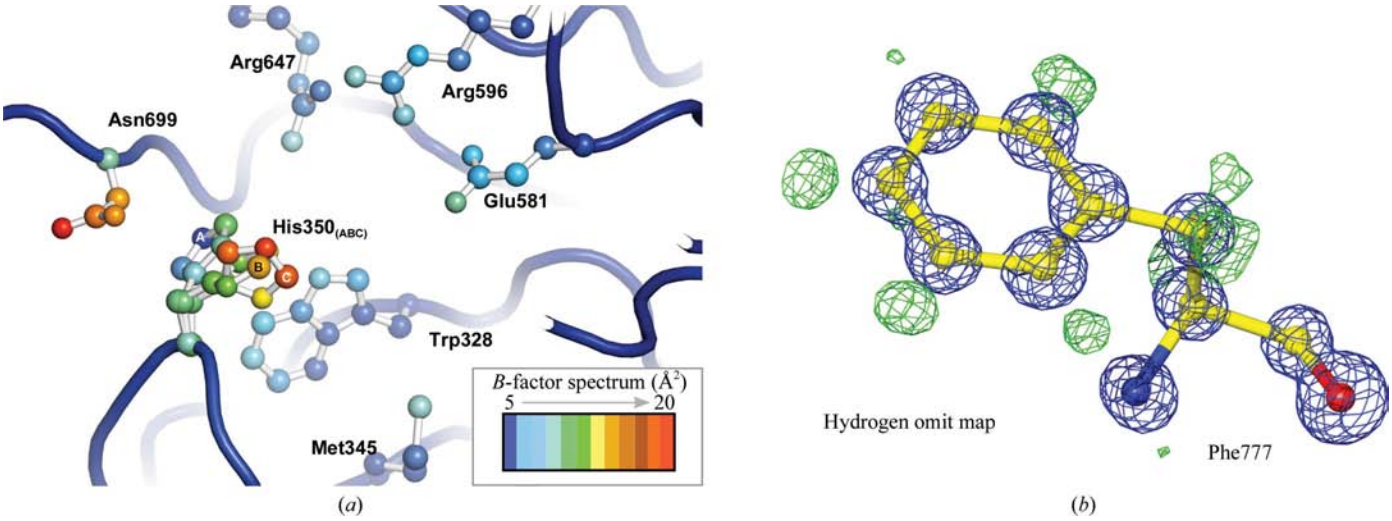

Figure 3
(*a*) Close-up view of the active site. The backbone is shown as tubes and important active-site residues are shown in ball-and-stick representation. The colouring is according to *B* factor, with blue corresponding to the lowest and red to the highest *B* factor. Importantly, His350 is present in three distinct conformations (A, B and C). (*b*) Close-up view of Phe777 in a hydrogen omit map. The difference density for the H atoms is clearly visible as a green mesh. The $2F_o - F_c$ map was contoured at 4σ; the $F_o - F_c$ map was contoured at 3σ.

acid or base, it is not unlikely that His350 also plays such a role in endoNF. Considering the pH profile, it can be assumed that His350 resides in the protonated state, acting as a general acid. However, as long as precise structural data that display the product complex or a noncleavable substrate analogue in the active site are lacking, it will not be possible to unequivocally rule out one or other of the reaction mechanisms. Extensive mutational studies will be necessary in order to assess the role of the respective active-site residues in the cleavage of polySia.

The authors would like to thank the beamline scientists of the ESRF Grenoble for support during data collection. We are grateful to our colleagues Katharina Stummeyer and Achim Dickmanns for many helpful discussions and their continuous support.

References

Amaya, M. F., Watts, A. G., Damager, I., Wehenkel, A., Nguyen, T., Buschiazzo, A., Paris, G., Frasch, A. C., Withers, S. G. & Alzari, P. M. (2004). *Structure*, **12**, 775–784.
Burmeister, W. P., Ruigrok, R. W. & Cusack, S. (1992). *EMBO J.* **11**, 49–56.
Buschiazzo, A., Tavares, G. A., Campetella, O., Spinelli, S., Cremona, M. L., París, G., Amaya, M. F., Frasch, A. C. & Alzari, P. M. (2000). *EMBO J.* **19**, 16–24.
Crennell, S. J., Garman, E. F., Laver, W. G., Vimr, E. R. & Taylor, G. L. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 9852–9856.
Davies, G. & Henrissat, B. (1995). *Structure*, **3**, 853–859.
DeLano, W. L. (2002). *The PyMOL Molecular Viewer.* http://www.pymol.org.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Gerardy-Schahn, R., Bethe, A., Brennecke, T., Mühlenhoff, M., Eckhardt, M., Ziesing, S., Lottspeich, F. & Frosch, M. (1995). *Mol. Microbiol.* **16**, 441–450.
Jakobsson, E., Jokilammi, A., Aalto, J., Ollikka, P., Lehtonen, J. V., Hirvonen, H. & Finne, J. (2007). *Biochem. J.* **405**, 465–472.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Lebedev, A. A., Vagin, A. A. & Murshudov, G. N. (2008). *Acta Cryst.* D**64**, 33–39.
Long, G. S., Bryant, J. M., Taylor, P. W. & Luzio, J. P. (1995). *Biochem. J.* **309**, 543–550.
Ly, H. D. & Withers, S. G. (1999). *Annu. Rev. Biochem.* **68**, 487–522.
Machida, Y., Hattori, K., Miyake, K., Kawase, Y., Kawase, M. & Iijima, S. (2000). *J. Biosci. Bioeng.* **90**, 62–68.
Morley, T., Willis, L., Whitfield, C., Wakarchuk, W. & Withers, S. (2009). *J. Biol. Chem.* **284**, 17404.
Mühlenhoff, M., Stummeyer, K., Grove, M., Sauerborn, M. & Gerardy-Schahn, R. (2003). *J. Biol. Chem.* **278**, 12634–12644.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Pelkonen, S., Pelkonen, J. & Finne, J. (1989). *J. Virol.* **63**, 4409–4416.
Petter, J. G. & Vimr, E. R. (1993). *J. Bacteriol.* **175**, 4354–4363.
Rye, C. S. & Withers, S. G. (2000). *Curr. Opin. Chem. Biol.* **4**, 573–580.
Scholl, D., Adhya, S. & Merril, C. (2005). *Appl. Environ. Microbiol.* **71**, 4872–4874.
Scholl, D., Rogers, S., Adhya, S. & Merril, C. R. (2001). *J. Virol.* **75**, 2509–2515.
Schwarzer, D., Stummeyer, K., Haselhorst, T., Freiberger, F., Rode, B., Grove, M., Scheper, T., Von Itzstein, M., Mühlenhoff, M. & Gerardy-Schahn, R. (2009). *J. Biol. Chem.* **284**, 9465–9474.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Stummeyer, K., Dickmanns, A., Mühlenhoff, M., Gerardy-Schahn, R. & Ficner, R. (2005). *Nature Struct. Mol. Biol.* **12**, 90–96.
Varghese, J. N., Laver, W. G. & Colman, P. M. (1983). *Nature (London)*, **303**, 35–40.
Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). *Protein Eng.* **8**, 127–134.
Watts, A. G., Oppezzo, P., Withers, S. G., Alzari, P. M. & Buschiazzo, A. (2006). *J. Biol. Chem.* **281**, 4149–4155.
Zechel, D. L. & Withers, S. G. (2001). *Curr. Opin. Chem. Biol.* **5**, 643–649.