Crystal Structures of Three Complexes between Chito-oligosaccharides and Lysozyme from the Rainbow Trout. How Distorted is the NAG Sugar in Site *D*?

By Solveig Karlsen and Edward Hough*

Department of Chemistry, Institute of Mathematical and Physical Science, University of Tromsø, N-9037 Tromsø, Norway

Abstract

Like all c-type lysozymes, those from rainbow trout act as 1,4-β-acetyl-muramidases to destroy bacteria by cleaving the polysaccharide chains of alternating N-acetylglucosamine (NAG) and N-acetylmuramic acid (NAM) units in the cell walls. Lysozymes also hydrolyse chitin, the analogous N-acetylglucosamine polymer. The rainbow trout enzymes have been shown to be particularly effective in bacterial defence. We have determined the crystal structures of three complexes between rainbow trout lysozyme (RBTL) and the chito-oligosaccharides $(NAG)_2$, $(NAG)_3$ and $(NAG)_4$ to resolutions of 1.8, 2.0 and 1.6 Å, respectively. Crystals of these complexes were obtained by co-crystallization, and intensity data were collected on a FAST area detector system. Refinement and model building gave final R values of 16.6, 15.9 and 16.5% for the di-, tri- and tetrasaccharide complexes, respectively. The results show that the chito-oligosaccharides bind to sites A, B and C as previously observed for complexes between the hen egg-white lysozyme (HEWL) and a variety of saccharides. The NAG ring in site D is not bound so deeply and is only slightly distorted towards a half-chair conformation as observed for the equivalent NAM residue in HEWL. From our results, there is reason to question the position and the degree of strain of the D saccharide and the mode of binding and importance of two saccharides in sites E and F for correct orientation of sugar D and effective hydrolysis of a productive substrate–lysozyme complex. Simple model building study from our structures implies a 'left-sided' binding mode of $(NAG)_6$ in the lower part of the active site of RBTL.

Introduction

Lysozymes are defined as 1,4-β-N-acetyl-muramidases which cleave β-1,4-linkages of polysaccharide chains in the peptidoglycan component of the cell walls of Gram-positive bacteria (Salton & Ghuysen, 1959, 1960). These long copolymer chains consist of alternating N-acetylmuramic acid (NAM) and N-acetylglucosamine

(NAG) units linked by glycosidic bonds between the C1 atom of the NAM and O4 atom of the NAG unit. Lysozymes also catalyse the hydrolysis of chitin, the homopolymer of N-acetylglucosamine.

Lysozymes play an important role in the non-specific immune defence of higher organisms and are therefore target enzymes for gene transfer in order to improve the immune defence of key species in the fish farming industry. Lysozymes from rainbow trout (*Oncorhynchus mykiss*) are particularly interesting for transgenic experiments as they show higher activity than conventional hen egg-white lysozyme against *Micrococcus luteus* and a variety of pathogenic bacteria which cause problems in the aquaculture of salmon (Grinde, 1989a,b; Grinde, Lie, Poppe & Salte, 1988; Grinde, Jollès & Jollès, 1988). The tertiary structure of rainbow trout lysozyme (RBTL) has been determined at 1.8 Å resolution (Karlsen *et al.*, 1995). In common with most c-type lysozymes, this enzyme consists of a single polypeptide chain of 129 amino-acid residues (Dautigny *et al.*, 1991) with a conserved overall fold.

The currently accepted reaction mechanism of lysozyme is based on the crystallographically observed binding of several mono-, di- and trisaccharide inhibitors in the upper part of the active-site cleft (subsites *A*, *B*,
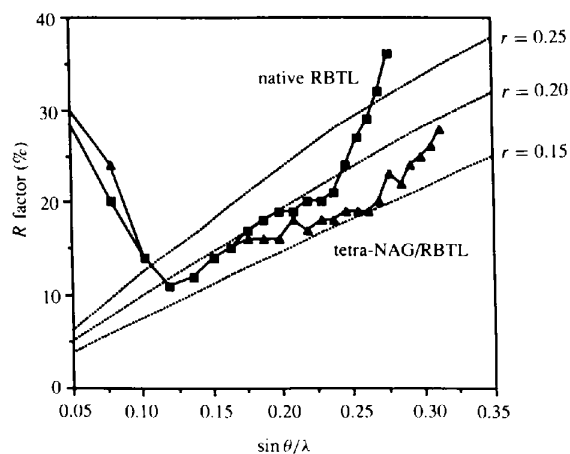

Fig. 1. Luzzati plots (Luzzati, 1952) of the native (squares) and $(NAG)_4$ bound (triangles) structure of RBTL. Atom-coordinate errors, r, are marked to the right.

---

* Corresponding author.

and *C*) of hen egg-white lysozyme and model building of three additional saccharide units into three further subsites, termed *D*, *E* and *F* (Phillips, 1966, 1967; Blake *et al.*, 1967; Blake, 1967). These results suggested that the glycosidic bond between *D* and *E* is hydrolysed through a general acid catalysis where Glu35 donates a proton to the glycosidic O atom and Asp52 stabilizes the positive oxo-carbonium ion which develops during the catalysis. The most controversial step in this catalytic pathway is the ground-state distortion of the sugar ring in the forth subsite (site *D*) from a normal chair toward a half-chair or sofa conformation (Phillips, 1967; Blake *et al.*, 1967). It was proposed that this distortion of the substrate in subsite *D* plays an important role in lysozyme catalysis (reviewed in Imoto, Johnson, North, Phillips & Rupley, 1972), because it reduces the energy difference between the bound substrate and the transition state, thus facilitating cleavage of the substrate.

Alternative mechanisms have been proposed by Koshland (1953) who suggested that a covalent enzyme–substrate intermediate is formed during hydrolysis, and Lowe (1967) who suggested that the C2 *N*-acetylamido group of the site *D* NAG ring participates in the cleavage of the glycosidic bond. The former hypothesis was abandoned because the model studies of Phillips (1966) indicated that the formation of a covalent bond was prohibited by the large gap between the carboxylate group of Asp52 and C1 of the site

D sugar (~3 Å). Later results, however, have shown that several (e→e)β-glycosidases probably catalyse the hydolysis of substrates through a covalently attached intermediate (Sinnott, 1990). The Lowe hypothesis was abandoned because it seemed impossible that the C2 acetamido group could approach C1 without disrupting the structure of the enzyme–substrate complex.

The crystal structures of complexes with different oligosaccharides bound in the active site of hen egg-white lysozyme (Ford, Johnson, Machin, Phillips & Tjian, 1974; Strynadka & James, 1991) lend some support to the idea that strain is an important factor in the catalytic mechanism of lysozymes. Earlier measurements on binding energies of a variety of sugars to all subsites of HEWL (Chipman, Grisaro & Sharon, 1967; in Imoto *et al.*, 1972) had shown that sugar residue *D* makes a negative contribution of about 3–6 kcal mol^{-1} to the binding affinity. In a series of experiments in solution, however, association constants for the binding of various oligosaccharides to lysozymes, indicated that the importance of strain in subsite *D* has probably been overestimated, and that strain can only account for a small part of the observed catalytic effect (Schindler, Assaf, Sharon & Chipman, 1977). Energy-minimization calculations on complexes between HEWL and the homopolymer (NAG)$_6$ and a hexasaccharide of alternating NAG and NAM units, have suggested that there are two distinct binding modes, a left- and right-sided, in the lower part of the active-site cleft (Pincus & Scheraga, 1979, 1981). In the 'left-sided' binding mode, which has the most stable conformation, the saccharide in site *D* lies near the surface of the cleft and has a normal unstrained conformation, whereas the fifth and sixth saccharide rings make contacts with the left side of subsites *E* and *F*. In the 'right-sided' binding mode, the site *D* saccharide is buried more deeply in the cleft
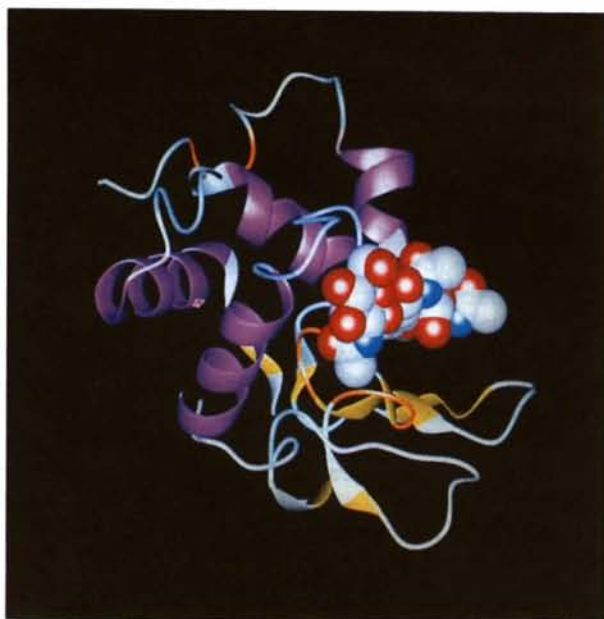

Fig. 2. Ribbon drawing (*QUANTA*) of the rainbow trout lysozyme with a (NAG)$_4$ molecule in the *A* to *D* sites, marked with van der Waals spheres. The α-helices are coloured violet, the β-strands are yellow, and the 3-turn and 4-turn are coloured orange and light blue, respectively. O and N atoms in the tetrasaccharide are marked in red and blue, respectively.

Fig. 3. $F_o - F_c$ omit map contoured at 0.12 e Å^{-3} (2σ) for the (NAG)$_4$ molecule bound in the active-site cleft of RBTL. The oligosaccharide was omitted from the coordinate file.

between the two catalytic acid residues, and a strained sugar conformation in this site seems to improve the interactions between saccharides E and F and protein residues on the right side of the cleft.

From extensive kinetic and thermodynamic studies, it has been proposed that a high polymer substrate binds initially to the left side to form an inactive enzyme–substrate complex, which then rearranges to a catalytically active complex with a distorted sugar ring in site D on the right side of the cleft (Banerjee, Holler, Hess & Rupley, 1975). This was supported by crystallographic studies on the unrefined complex between HEWL and NAM–NAG–NAM (MGM) (Kelly, Sielecki, Sykes & James, 1979) and the binding of the tetrasaccharide (NAG–NAM)$_2$ to tortoise egg-white lysozyme (in Johnson et al., 1988). Since the saccharide units in site D in these complexes were pointing outward from the active-site cleft and exhibit the normal chair conformation, it was suggested that they represent initial Michaelis complexes, rather than the transition-state conformation seen in the HEWL-δ-lactone complex (Ford et al., 1974). Productive binding and conversion to a reactive 'right-sided' complex is proposed to occur only after the binding of two additional hexose units in subsites E and F (reviewed in Johnson et al., 1988). The only crystallographic evidence of saccharide binding to site E and a new site F has been observed for the turkey egg-white lysozyme (Sarma & Bott, 1977), but, unfortunately, this is a low-resolution structure which has poorly defined electron density for the saccharide rings.

In order to further elucidate the mode of binding of chitin oligosaccharides and the conformation of the glucopyranose ring in site D, di-, tri- and tetrasaccharide oligomers of N-acetylglucosamine were co-crystallized with lysozyme from the rainbow trout and the structures of these complexes were determined. We describe the binding of these oligosaccharides to the A, B, C and D sites of the active-site cleft of the lysozyme.

Materials and methods

Co-crystallization of RBTL with (NAG)$_2$, (NAG)$_3$ and (NAG)$_4$

The rainbow trout lysozymes, type I and II, were purified as previously described (Grinde, Jollès et al., 1988). Hanging-drop co-crystallization experiments with lysozyme and three different oligosaccharides, (NAG)$_2$, (NAG)$_3$ and (NAG)$_4$, were carried out. The oligosaccharides were kindly provided by Marit W. Anthonsen, Department of Biotechnology, NTH, Trondheim. Crystals appeared after one to two weeks for all RBTL–saccharide complexes using the same crystallization conditions as for the native enzyme in addition to 10 mg ml^{-1} saccharide, i.e. 18 mg ml^{-1} enzyme, 27.5–32.5% saturation of (NH$_4$)$_2$SO$_4$ and pH

9.5 to 10.5 at 277 K (Karlsen et al., 1995). Crystals of the complexes were a bit bigger than the native crystals, particularly for the RBTL–(NAG)$_4$ complex (about 0.8 mm in the longest dimension).

Data collection and processing

Intensity data from RBTL crystals with (NAG)$_2$, (NAG)$_3$ and (NAG)$_4$ were measured at room temperature with Cu $K\alpha$ radiation on the FAST area detector system. The detector is mounted on a FR571 SERIES 5 rotating-anode generator operating at 40 kV and 70 mA. Reflections were measured with a crystal-to-detector distance of 60 mm and with a width of 0.1° for each individual rotation image. For each scan the crystal was rotated 65° in ω. Data from native lysozyme crystals were collected, as previously described, on an R-AXIS II system with unit-cell dimensions $a = b = 76.68$ and $c = 54.46$ Å (Karlsen et al., 1995). The cell parameters for the RBTL–oligosaccharide complexes are seen in Table 1. The small changes in cell dimensions indicate that the crystals of native lysozyme and the three complexes are isomorphous.

The data for the RBTL–saccharide complexes were processed with MADNES (Plugrath, 1989) and profile fitting was carried out in XDS (Kabsch, 1988). Merging to unique data sets was carried out with programs from the CCP4 software package (Collaborative Computational Project, Number 4, 1994). For the complexes between (NAG)$_2$, (NAG)$_3$ and (NAG)$_4$ and RBTL, a total of 15 737, 12 175 and 22 072 unique reflections were measured to 1.8, 2.0 and 1.6 Å resolution, respectively, with completeness of 91.3, 96.9 and 89.9% and R_{merge} values of 4.5, 3.3 and 3.3, respectively. Results from the data collection and merging are summarized in Table 1.

Refinement

The rainbow trout lysozyme complexes with bound (NAG)$_2$, (NAG)$_3$ and (NAG)$_4$ were solved by difference Fourier methods using phases obtained from the structure of native RBTL refined to 1.8 Å resolution (Karlsen et al., 1995).

All solvent molecules and residues that were expected to interact with the substrates from the active-site region were omitted, before five cycles of least-squares refinement with the program PROLSQ (Hendrickson, 1985) were carried out. Difference Fourier maps, $|F_{obs}(\text{complex})| - |F_{calc}(\text{native})|$, were calculated and models of the oligosaccharides were fitted into the electron density. The molecular template for generating initial coordinates of the substrates and setting the restraints dictionary were constructed from the crystal structure of α-N,N'-diacetylchitobiose (Mo & Jensen, 1978) where the glucopyranose rings are in the full $^4C_1$ conformation. The models were improved during the refinement by inspection and minor manual adjustments of the positions of the saccharide rings on an Evans

Table 1. *Summary of data collection and processing*

Area detector	Native RBTL* R-AXIS II image plate	RBTL–(NAG)₂ FAST	RBTL–(NAG)₃ FAST	RBTL–(NAG)₄ FAST
Resolution (Å)	42.11–1.80	14.63–1.80	14.67–2.00	14.65–1.60
Completeness (%)	91.2	91.3	96.9	89.9
No. of independent reflections ($F > 3\sigma_F$)	16080	15737	12175	22072
Cell parameters (Å)	$a = b = 76.68$ $c = 54.46$	$a = b = 76.46$ $c = 54.22$	$a = b = 76.42$ $c = 54.29$	$a = b = 76.58$ $c = 54.19$
R_{merge} (%)†	4.1	4.5	3.3	3.3

* From Karlsen *et al.* (1994).

† $R_{merge} = \sum |I_i - \langle I \rangle| / \sum I_i$, where I_i is the intensity of an individual reflection and $\langle I \rangle$ is the mean intensity of that reflection.

Table 2. *Summary of the final refinement statistics*

Structure	Native RBTL*	RBTL–(NAG)₂	RBTL–(NAG)₃	RBTL–(NAG)₄	Standard deviation (σ)
Resolution range (Å)	8.0–1.8	8.0–1.8	8.0–2.0	8.0–1.6	
R-factor	0.174	0.166	0.159	0.165	
No. of reflections > 3σ	16080	15737	12175	22072	
No. of protein atoms	999	999	999	999	
No. of sugar atoms	—	29	43	57	
No. of solvent atoms	127	112	117	122	
R.m.s. deviations from ideal values:					
Distance restraints (Å)					
Bond distance (1-2)	0.013	0.013	0.013	0.012	0.02
Angle distance (1-3)	0.044	0.045	0.047	0.040	0.04
Planar distance (1-4)	0.047	0.074	0.075	0.073	0.05
Plane restraints (Å)	0.013	0.013	0.013	0.013	0.02
Chiral centers (Å³)	0.038	0.039	0.037	0.036	0.06

* Values from Karlsen *et al.* (1995).

& Sutherland PS300 using the program *FRODO* (Jones, 1985). In order to ensure that the final refined conformation observed for the *D* ring was not biased, the model of this saccharide unit was changed back to the full ⁴C₁ conformation and ten cycles of refinement were run. The difference maps clearly indicated that the conformation, of the sugar ring in site *D* obtained in the first refinement runs, was correct.

In order to give saccharide rings greater conformational freedom, restraints on the torsional angles were loosened up, and all sugar rings in (NAG)₄ were changed back to the full chair conformation. Ten cycles of refinement were carried out, with and without restraints. Finally, the two refinement runs were repeated, but the *D* ring was refined using a sofa conformation (from Anderson, Grütter, Remington, Weaver & Matthews, 1981) in the dictionary. The four resulting models of the tetrasaccharide and the $|2F_o| - |F_c|$ maps clearly showed that the saccharide in site *D* had the chair conformation.

The occupancy factors for all atoms of the bound chito-oligosaccharides were kept constant. 112, 117 and

122 water molecules were included in the final structures of (NAG)₂, (NAG)₃ and (NAG)₄ bound to the rainbow trout lysozyme, respectively. Table 2 shows a summary of the final refinement statistics for both the native rainbow trout lysozyme and the complexed structures. The r.m.s. deviations for the refined complexes are of the same size as for the native structure. One exception concern the planar 1–4 distances which are 0.074 for the RBTL–(NAG)₂, 0.075 for the RBTL–(NAG)₃ and 0.073 for the RBTL–(NAG)₄ complex, but only 0.047 for the native lysozyme structure.

The coordinates for the structures of the RBTL–oligosaccharides complexes have been deposited with the Protein Data Bank, Brookhaven National Laboratory.*

Results and discussion

Quality of the refined structures

Ramachandran plots (Ramachandran, Ramakrishnan & Sasisekharan, 1963) of the refined RBTL–saccharide complexes generated by the program *PROCHECK* (Laskowski, MacArthur, Moss & Thornton, 1993), show that the majority of the main-chain dihedral angles [92.9% for the RBTL–(NAG)₃ and 92.0% for the (NAG)₂ and (NAG)₄ bound complexes] fall within the most favoured regions. The remaining 8.0, 7.1 and 8.0% of the non-glycine residues in the RBTL complexes with bound di-, tri- and tetrasaccharide, are in additionally allowed regions. The residues in these regions, Ser36, Tyr38, Gln57, Tyr62, Arg68, Asn74, Asn113, Cys115 and Asn117, all lie in loop regions and with more conformational freedom than the amino acids in β-sheet and α-helix regions.

From Luzzati plots (Luzzati, 1952), the mean positional errors are estimated to be about 0.18 Å for all the complexes. In Fig. 1 the Luzatti plots of the native and (NAG)₄ bound structure of RBTL are compared. The figure shows that the mean error in atomic position is less for the complexed structure than for the native form, and that the curve for lysozyme with bound (NAG)₄ does not rise so steeply at high resolution as for the native enzyme.

Interpretation of the electron-density maps

In general, the most poorly ordered residues in the native rainbow trout lysozyme structure, involving the side chains of lysine, arginine, glutamine and aspartic

Table 3. *Temperature factors for $(NAG)_4$ bound to sites A, B, C and D in RBTL*

	A	B	C	D
Sugar moiety	$\langle B\rangle$* (Å^2)	$\langle B\rangle$* (Å^2)	$\langle B\rangle$* (Å^2)	$\langle B\rangle$* (Å^2)
Ring atoms (C1, C2, C3, C4, C5, O5)	30.1	19.8	18.8	33.6
Exocyclic O atoms				
O1	—	—	—	53.9
O3	67.1	29.1	15.1	32.0
O4	46.6	24.1	17.7	24.7
Hydroxymethyl atoms				
C6	46.3	19.4	30.7	23.6
O6	66.8	18.8	41.7	231.5
Acetamido atoms				
(N2, C7, O7, C8)	45.3	29.3	13.5	50.9

* Average temperature factor values ($\langle B\rangle$) are given when more than one atom is included in the group.

acid residues that lie on the molecular surface, are also difficult to locate accurately in the complexed structures. Electron-density maps ($|2F_o| - |F_c|$) contoured at $0.3\,\text{e}\,\text{Å}^{-3}$ (1σ level) show weak density at the ends of the side chains for Arg61, Lys73, Gln118 and Asp119 with B factors of over $50\,\text{Å}^2$ for the most flexible atoms. In common with the native structure, Arg121 has the most diffuse side chain in the maps of the complexes. The electron-density maps of the sugar complexes show clear electron density for all residues in the active-site cleft that interact with the substrates.

Difference ($|2F_o| - |F_c|$) maps based on the refined structures of the RBTL–saccharide complexes, clearly show well defined electron density for an $(NAG)_2$ molecule in the B and C sites, for an $(NAG)_3$ molecule bound to sites B, C and D and an $(NAG)_4$ molecule bound in the A to D sites of the active-site cleft. The three chitin oligosaccharides bind to the rainbow trout lysozyme in a similar mode, so that discussion of protein–carbohydrate interactions and the conformation of the bound saccharide complex will only be discussed in detail for RBTL–$(NAG)_4$. Fig. 2 illustrates the rainbow trout lysozyme in a ribbon drawing with a $(NAG)_4$ molecule shown as van der Waals spheres, bound to the active site.

Observed electron density ($|F_o| - |F_c|$) contoured at $0.12\,\text{e}\,\text{Å}^3$ (2σ level) for each of the atoms of $(NAG)_4$ and corresponding average temperature factors for the enzyme-bound sugar are given in Fig. 3 and Table 3, respectively. The atomic numbering of an N-acetylglucosamine ring is given in Fig. 4. The clearly defined electron density associated with the atoms and the low mean B factor of ring C ($19.0\,\text{Å}^2$) shows that this is the most tightly bound sugar unit of the saccharide. The buried 2-acetamido group that plays an important role in substrate specificity (Imoto *et al.*, 1972) has a mean B factor ($13.5\,\text{Å}^2$) which is lower than that for well defined groups in the protein.

The well defined electron density and low average B factor ($23.4\,\text{Å}^2$) for the atoms in ring B shows that also this sugar unit is highly ordered in the complex.

On the other hand, B factors for the solvent-exposed 2-acetamido group in NAG ring B indicate that this is more mobile than that in ring C. Poorly defined electron density and high B factors are found for O3 and N2 in the acetamido group of the NAG A unit. This is not unexpected, since this NAG unit lies on the surface of the active-site cleft and exhibits the largest average temperature factor of the sugar units ($50.4\,\text{Å}^2$). The acetamido atoms of NAG ring D show the largest average temperature factor for these groups in the tetrasaccharide, probably because this side chain is more exposed to the solvent. The exocyclic O1 atom of the terminal sugar also exhibits a high B factor and diffuse observed electron density, indicating that this atom could be both in an axial (α-anomeric form) and equatorial (β-anomeric form) position. However, the observed density seems to favour an axial O1 position, thus the α-anomer seems to be preferred in the crystals. The NAG unit in site D is not as extensively buried as the sugar moieties in sites B and C, and has a mean temperature factor ($34.3\,\text{Å}^2$), higher than for NAG in site A.

Electron-difference maps for the complexes between rainbow trout lysozyme and $(NAG)_2$ and $(NAG)_3$ clearly show that the two saccharides bind to the active-site cleft in the same way and with the same well defined electron density in the B and C and B, C and D sites, respectively, as $(NAG)_4$. The average temperature factors seen in Table 4 and the observed electron density for the atoms associated with the di- and trisaccharide indicate that, as in the $(NAG)_4$ bound complex, ring C is the most tightly bound and ordered of the sugar rings. The average B factors for the B and C NAG unit in the RBTL–$(NAG)_2$ complex are higher than for the equivalent rings in the RBTL–$(NAG)_4$ complex, suggesting that NAG units which do not have adjacent sugars on both sides, are more loosely bound than NAG units in a polymere substrate. Also the sugar unit bound to site B in the RBTL–$(NAG)_3$ complex clearly exhibits a higher B factor than the same sugar ring in the tetrasaccharide, whereas the sugar in site D is more strongly bound than in $(NAG)_4$. Electron density for the exocyclic O1 atoms of this pyranose ring is most clearly defined in the axial position in the difference map.

Geometry of the oligosaccharides bound to RBTL

All three chitin oligosaccharides bind to the active-site cleft of RBTL in an extended conformation. The geometry of the β-1,4-glycosidic linkages between the
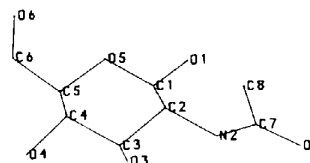

Fig. 4. Atomic numbering of a N-acetylglucosamine sugar ring.

Table 4. *Average temperature factors for the N-acetylglucosamine units of the three chito-oligosaccharides bound to the active-site cleft of RBTL*

Site	RBTL–(NAG)$_2$	RBTL–(NAG)$_3$	RBTL–(NAG)$_4$
A	—	—	50.4
B	33.5	31.5	23.4
C	25.6	23.4	19.0
D	—	38.2	34.3

sugar units in the three complexes is detailed in Table 5. Values from the NAM–NAG–NAM and (NAG)$_3$ bound complexes with HEWL are included for comparison. The glycosidic linkages described by the torsional angles φ and ψ, between the sugar units bound to sites B, C and D in the three RBTL–saccharide complexes, are almost identical. The φ values for the B–C glycosidic linkages for (NAG)$_2$ and (NAG)$_3$ and A–B and B–C linkages for (NAG)$_4$ are all within the range observed (-71 to $-96°$) in crystal structures of other oligosaccharides (Mo & Jensen, 1978; Jeffrey, 1990). However, the φ-value for the C–D linkage in both sugar complexes with RBTL and the HEWL–MGM complex of HEWL are outside this range. The φ value for the A–B linkage for the (NAG)$_3$ molecule in HEWL also deviates from the normal values. Deviations in the glycosidic bridges in lysozymes complexes are, thus, greater than normally observed for unbound saccharides with β-1,4-linkage. The ψ values for the four sugar–lysozyme complexes are all inside the range that is observed (-107 to $-161°$) for previously solved oligosaccharide structures (Mo & Jensen, 1978; Jeffrey, 1990).

The wide range of intraring twist described by the ψ_H values in Table 5 is evidence of a considerably degree of flexibility in the β-1,4-glycosidic bridge. The helical twist parameter varies from a left-handed twist of -25 and $-16°$ for the C–D glycosyl linkages in the (NAG)$_3$ and (NAG)$_4$ bound complexes respectively, to a right-handed twist of $70°$ for the A–B linkage in the HEWL–(NAG)$_3$ complex. Previously solved crystal structures of oligosaccharides have ψ_H values varying from $-12.5°$ for methyl β-cellobioside (Ham & Williams, 1970) to $80.5°$ for a xylobiose unit in an aldotriuronic acid (Moran & Richards, 1973). The left-handed intraring twist of $-25°$ between the C and D sugar units in the RBTL–(NAG)$_3$ complex is more extreme than this. It is not unlikely that the deviation from the normally observed geometry of the glycosidic linkage between the sugar units in subsites C and D, and the large left-handed helical twist for the chitin oligosaccharides bound to the same sites in RBTL, are necessary for optimal binding of the NAG unit in subsite D for effective catalysis. The table shows that the interring twist for the sugar units bound to subsite C and D in the HEWL–MGM complex is $11°$, which lies within the range observed for this parameter in crystal structures of uncomplexed oligosaccharides. Intramolecular hydrogen

Table 5. *Geometry of glycosidic linkages*

Values of φ, ψ are the torsional angles about $C_{(1)}$—$O_{(4')}$ and $O_{(4')}$—$C_{(4')}$ defined by $O_{(5)}$—$C_{(1)}$—$O_{(4')}$—$C_{(4')}$ and $C_{(1)}$—$O_{(4')}$—$C_{(4')}$—$C_{(5')}$, respectively. The value of ψ_H is the helical twist parameter defined as the average of the pseudorotation angles $\psi_1 = O_{(5)}$—$C_{(1)}$—$C_{(4')}$—$C_{(3')}$ and $\psi_2 = C_{(2)}$—$C_{(1)}$—$C_{(4')}$—$C_{(5')}$ (Mo & Jensen, 1978).

Lysozyme	Sugar	φ (°)	ψ (°)	ψ_H (°)	$O_{(5)}$—$O_{(3')}$ (Å)
RBTL	(NAG)$_2$				
	B–C	−96	−112	36	2.83
RBTL	(NAG)$_3$				
	B–C	−88	−116	35	2.86
	C–D	−121	−148	−25	2.70
RBTL	(NAG)$_4$				
	A–B	−90	−111	35	2.89
	B–C	−85	−125	34	2.81
	D–C	−113	−150	−16	—
HEWL	MGM*				
	B–C	−89	−120	30	3.00
	C–D	−108	−110	11	—
HEWL	GGG*				
	A–B	−40	−126	70	3.5
	B–C	−90	−117	28	3.2

* Values of the bound NAM–NAG–NAM (MGM) and (NAG)$_3$ (GGG) in the active-site cleft of HEWL (Strynadka & James, 1991; Cheetham *et al.*, 1992).

bonds between the ring O atom, O5, and the exocyclic O atom, O3', in the adjacent saccharide ring which are characteristic for β-1,4-linkages, are present in the three chitin oligosaccharides bound to RBTL.

Difference electron density for the chitin oligosaccharide molecules bound to RBTL shows that all pyranose rings have a chair conformation. The endocyclic torsion angles, however, for the atoms associated with the NAG sugar rings (see Table 6) show more variation and some of the bond distances fall outside the ranges for equivalent parameters observed in crystal structures of several disaccharides with β-1,4-linkage (Mo & Jensen, 1978; Jeffrey, 1990; Hirotsu & Shimada, 1974). The pyranose rings bound to site D in the (NAG)$_3$ and (NAG)$_4$ molecule in particular, show the largest deviations from ideality, but they still have a chair-conformation (see Table 7). The C(3)—C(4)—C(5)—O(5) and C(4)—C(5)—O(5)—C(1) torsional angles are considerably smaller for this sugar ring making it 'flatter' than the other saccharides. A superposition of a crystal structure of NAG (Mo & Jensen, 1978) with our refined site D NAG saccharide (see Fig. 5), illustrates that the deviation from full $^4C_1$ conformation is mainly caused by the hydrogen-bonded interaction between the endocyclic O5 sugar atom and OD2 of Asp52. Interactions from the hydroxymethyl group of the saccharide and the protein propably also contribute to the stabilization of the sugar conformation.

Fig. 6 shows the observed density and the refined conformation of the D sugar ring of (NAG)$_4$ where a model of an NAG unit with sofa conformation (from Anderson *et al.*, 1981) is superimposed for comparison. The figure clearly emphasises that the refined NAG ring fits the electron density very well, and the hydroxymethyl

Table 6. *Torsional angles for NAG sugar rings in the active-site cleft of RBTL*

Torsional angle (°)	RBTL–(NAG)$_4$				RBTL–(NAG)$_3$			RBTL–(NAG)$_2$	
	A	B	C	D	B	C	D	B	C
O5—C1—C2—C3	65	59	63	49	45	57	55	52	52
C1—C2—C3—C4	−58	−55	−51	−51	−46	−45	−54	−46	−46
C2—C3—C4—C5	55	53	50	46	49	44	38	48	48
C3—C4—C5—O5	−61	−55	−58	−37	−53	−53	−23	−56	−58
C4—C5—O5—C1	74	63	72	36	59	70	29	67	72
C5—O5—C1—C2	−76	−64	−74	−42	−53	−72	−46	−64	−68

Table 7. *A comparison of the endocyclic torsional angles of our refined chito-oligosaccharides bound to RBTL and NAG sugars with chair, half-chair and sofa conformations*

Torsion angles (°)	A, B, C*	D*	Chair conformation †	Half-chair conformation‡	Sofa conformation‡
O5—C1—C2—C3	56	52	58	25	30
C1—C2—C3—C4	−50	−52	−54	−48	−60
C2—C3—C4—C5	50	42	52	62	62
C3—C4—C5—O5	−56	−30	−54	−51	−31
C4—C5—O5—C1	68	33	61	28	0
C5—O5—C1—C2	−67	44	−63	−15	0

* Average values for the endocylic torsional angles of the NAG units in (NAG)$_2$, (NAG)$_3$ and (NAG)$_4$ bound to RBTL.
† Values from Mo & Jensen (1978).
‡ Values from Ford *et al.* (1974).

group on C5, well defined in the electron-density map, is clearly positioned equatorially, as expected for a chair conformation. The sofa conformation does not fit the observed density as well as the refined saccharide ring.

Protein–oligosaccharide interactions

Fig. 7 shows hydrogen-bonding interactions between the protein, bound water molecules and the tetrasaccharide ligand in sites *A* to *D*. The protein–oligosaccharide interactions are detailed in Table 8 which includes details of the HEWL–MGM complex for comparison. Carbohydrate–protein interactions and the binding of the ligand in our three complexes mutually confirm the positions for chito-oligosaccharides in sites *A* to *D* in the active-site cleft of the rainbow trout lysozyme.

This is emphasized in Figs. 8(*a*) and 8(*b*) where the (NAG)$_2$ and (NAG)$_3$ molecules are superimposed on the tetrasaccharide with r.m.s. deviations of only 0.30 and 0.21 Å for subsites *B* and *C* and 0.18, 0.20 and 0.29 Å for subsites *B*, *C* and *D*, respectively.

The mode of binding of our chitin oligosaccharides to sites *A* to *C* in the active-site cleft of RBTL is approximately similar to that reported from previously solved complexes of various oligosaccharides bound to HEWL (Ford *et al.*, 1974; Strynakda & James, 1991; Cheetham, Artymiuk & Phillips, 1992), but the saccharides in site *D* show greater variation in binding pattern and position. The NAG ring of the tetrasaccharide bound to site *A* in the rainbow trout lysozyme lies almost on the surface of the active-site cleft exposed to the solvent, and is the least tightly bound of the sugars. The only
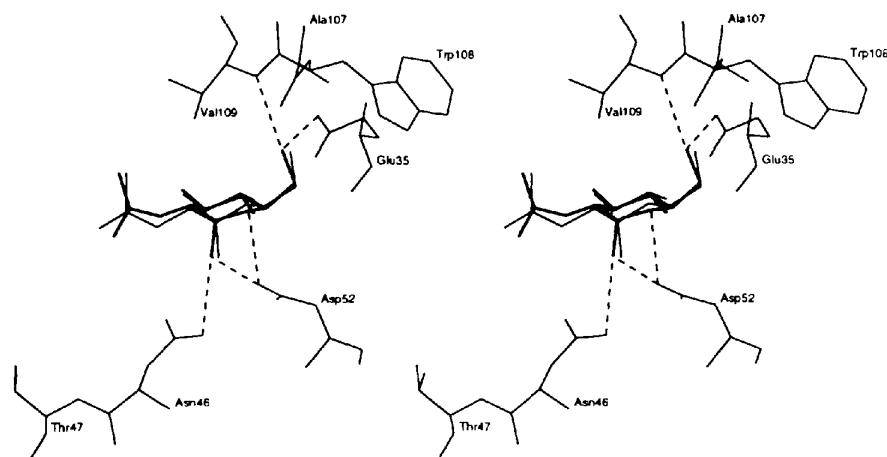

Fig. 5. A superimposition of a small molecular structure of NAG (thin lines) (Mo & Jensen, 1978) with our refined site *D* NAG saccharide (thick lines). The hydrogen bonds are depicted with dashed lines.

Table 8. *Comparison of protein–oligosaccharide interactions in the RBTL–(NAG)$_4$ and HEWL–MGM complexes* (*Strynadka & James*, 1991)

Complex Ligand site	RBTL–(NAG)$_4$			HEWL–MGM		
	Sugar atom	Protein atom	Distance (Å)	Sugar atom	Protein atom	Distance (Å)
A	N2	OD2 Asp101	3.0			
B	O6	OD2 Asp101	2.5	O6	OD1 Asp101	2.4
	O7	ND2 Asn103	3.2	O10	ND2 Asn103	2.8
	O7	OW 152	3.2	O12	ND2 Asn103	3.2
	O7	OW 246	2.6	O6	OW 236	2.6
	O6	OW 153	2.5	O7	OW 270	3.2
C	O3	NE1 Trp63	3.2	O3	NE1 Trp63	3.0
	N2	CO Ala107	3.0	N2	CO Ala107	2.9
	O7	NH Asn59	3.0	O7	NH Asn59	2.9
	O4	OW 152	3.1	O6	NE1 Trp62	2.9
				O6	OW 205	3.1
D	O6	NH Val109	3.0	O6	NH Val109	2.8
	O6	OE2 Glu35	2.9	O1	OE1 Glu35	2.9
	O5	OD2 Asp52	3.2	N2	OD2 Asp52	3.2
	O1	OD2 Asp52	2.2	N2	OD1 Asn46	3.4
	O1	OD1 Asn46	3.1	O1	OW 181	2.7
	N2	OW 150	3.1	O10	OW 200	2.4
				O10	OW 205	2.9

interaction between the protein and this pyranose ring is a hydrogen bond between OD2 of Asp101 and N atom of the 2-acetamido group (see Fig. 7), which extends into a mainly hydrophilic region made up of residues Lys73 and Asp101. The N-acetyl side chains of the sugar moieties in sites A and B are both accessible to solvent.

In site B, the hydroxyl O atom at position 6 in the pyranose ring interacts with the OD2 atom of Asp101 (see Fig. 7). The second hydrogen bond to the protein in this site is found between O7 in the solvent-exposed acetamido group and ND2 of Asn103. The other major binding interaction for site B is the extensive non-polar contact between the apolar face of the NAG ring and the side-chain ring of Tyr62. The planes of these two rings, which are almost parallel, are separated by approximately 4.5 Å. This large hydrophobic stacking interaction and the extra hydrogen bond between the sugar and the protein in site B may explain why the binding energy for this site (–2.7 kcal mol^{-1}) (reviewed in Imoto *et al.*, 1972) seems to be somewhat greater than that for site A (–2.3 kcal mol^{-1}). Three hydrogen bonds between the O6 and O7 atoms of the NAG molecule and structural water molecules are also present. In the RBTL–(NAG)$_3$ complex the hydroxyl O atoms, O4 and O5, of the end sugar ring bound to site B also make hydrogen bonds to water molecules.

A water-mediated hydrogen-bonded network is seen between the hydroxymethyl group of the sugar unit B in (NAG)$_4$ and the main-chain carboxyl group of Val98 and Gly104, where water 153 makes a bridge between the O6 atom of the pyranose ring and water 155. This latter water is hydrogen bonded to residues 98 and 104. The same hydrogen-bonded contacts are found in the RBTL–(NAG)$_3$ complex, indicating that ordered water molecules contribute to binding affinity and specificity by mediating hydrogen bonds between carbohydrate and protein (Vyas, 1991).

Previously solved structures of complexes between various saccharides and hen egg-white lysozyme (Strynadka & James, 1991; Cheetham *et al.*, 1992) have shown that site C, which for steric reasons only can accommodate an NAG residue, is the binding site which is most specific in nature. Thus, there are more hydrogen-bonded sugar–protein interactions, particularly from the buried 2-acetamido group, in this site than in the other sites. Furthermore, the importance of the NAG C residue for substrate specificity is seen from estimates of free energy of binding for this site (–4.6 kcal mol^{-1}) compared with the others (reviewed in Imoto *et al.*, 1972).

The carbohydrate–protein interactions we observe (Fig. 7) for the NAG residue bound in subsite C of RBTL are similar to those found in MGM and (NAG)$_3$ with HEWL (Strynadka & James, 1991; Cheetham *et al.*, 1992) and three of the four hydrogen bonds between the pyranose ring and the protein in these complexes are also found in the RBTL–(NAG)$_4$ complex (see Table 8). N2 and O7 of the buried 2-acetamido group interact with the main-chain CO group of Ala107 and the main-chain NH group of Asn59, respectively. Hydrophobic contacts are made between C2 in this group and the indole ring of Trp108 and the side chains of Ile58 and Val98. NE1 in the plane of the tryptophan ring of residue 63 lies perpendicular to the polar face of the NAG ring and makes a hydrogen bond with the exocyclic O3 of the sugar.

In complexes between oligosaccharides and HEWL, NE of Trp62 also forms a hydrogen bond to O6 of the pyranose ring (Strynadka & James, 1991; Cheetham *et al.*, 1992). However, this interaction is not found in RBTL–(NAG)$_4$, where residue 62 is a tyrosine. The distance between the OH group in the side chain of this residue and the exocyclic O6 atom of the C saccharide ring is 3.25, but the geometry is not suitable for the formation of a hydrogen bond.

A comparison of binding free energies obtained from association constants for a series of oligosaccharides bound to hen egg-white and human lysozyme suggest that subsite C of the human enzyme is 'looser' than the equivalent site in the hen egg-white enzyme due to the replacement of Trp62 in HEWL by Tyr63 (63 due to an insertion) in the human enzyme (Schindler et al., 1977). Furthermore, the binding constants of chitotriose to the two mutants of HEWL, Trp62Tyr and Trp62Phe, were found to be several times lower than that of the wild-type enzyme (Kumagai, Sunada, Takeda & Miura, 1992).

Recent studies on hen egg-white lysozyme where Trp62 was replaced with Tyr, His and Phe by site-directed mutagenesis showed that the substrate-binding modes of these mutants were markedly altered in comparison with that of the wild-type lysozyme (Kumagai, Maenaka, Sunada, Takeda & Miura, 1993). With the chito-oligosaccharide, PNP-(NAG)$_5$, as a substrate, substitution of Trp62 resulted in an increase in productive binding to subsites B–F and in total hydrolytic activity. A feasible explanation for this might be that replacement of Trp62 with Tyr, Phe and His could lead to reduced non-polar interaction with the sugar ring in subsite B and, as seen for the complexes between RBTL and the chito-oligosaccharides, loss of hydrogen bonding to a sugar in subsite C which in turn could change the substrate specificity. These observations confirm the crystallographic result from the complex between HEWL and (NAG)$_3$ where the saccharide is bound in subsites A to C (Cheetham et al., 1992) and not in sites B to D as found for our RBTL–(NAG)$_3$ complex and in the complex of (NAG)$_3$ bound to human lysozyme (Matsushima, Inaka & Morikawa, 1990).

Although it deviates somewhat from an ideal geometry, the sugar ring in site D has the chair conformation (see Fig. 5) and is not distorted toward the sofa conformation found for this residue in the HEWL–MGM complex. The NAG residue in this site packs between residues Trp108 and Val109 from the 'top' and residues in the β-sheet (Asn46, Asp52 and Asn59) from 'below'. There are no apparent collisions between the protein and the chair-formed saccharide ring which could force it into a sofa conformation. However, Fig. 9(a) which illustrates a superimposition of the MGM molecule bound to subsites B to D in HEWL (Strynadka & James, 1991) and the (NAG)$_4$ ligand in the equivalent sites in RBTL, clearly shows that the saccharide units in site D are bound in a different way. A closer view of this binding site (see Fig. 9b) shows that the half-chair NAM residue in MGM is more deeply buried between the catalytic residues than the D NAG-ring of tetra-NAG. The equatorial O1 atom of the NAM residue interacts with OE1 of Glu35, whereas O6 of the NAG unit is hydrogen-bonded to OE2 of the glutamic acid. In both complexes, the hydroxymethyl group is hydrogen-bonded to the main-chain N atom of Val109. The importance of this

hydrogen bond was underscored in studies where a mutant form of HEWL in which Val109 was substituted by proline, showed markedly reduced hydrolytic activity toward oligosaccharide substrates (Inaka, Matsushima & Morikawa, 1990). The presence of a proline in position 109 may preclude hydrogen bonding to the O6 atom of the saccharide in site D and thereby destabilize productive binding of NAG or NAM in this site.

The different positions of the D saccharide units also lead to different interactions with the catalytic aspartic acid (see Fig. 9b). In the HEWL–MGM complex N2 of the acetamido group is hydrogen bonded to OD2 of Asp52, whereas the equivalent interactions are from O1 and O5 in RBTL–(NAG)$_4$ which can easily stabilize a putative carbonium ion in the transition state. In the latter complex, the side chain of Asn46 is hydrogen bonded to the axial O1 atom, but the corresponding interaction is to N2 of the NAM unit in the HEWL complex. The only interaction between the 2-acetamido group and RBTL occurs between the N atom and water 150.

Figs. 9(a) and 9(b) show that most of the amino acids in the active sites of HEWL and RBTL that interact with the bound oligosaccharide substrates, are structurally conserved. Furthermore, the only difference between a NAG and a NAM residue is found at O3 where the latter carries a lactyl group. Therefore, it seems reasonable that the difference in binding mode to subsite D in the two lysozymes has something to do with this extra lactyl group which may force the NAM ring to a 'deeper' more buried position. Results from kinetic analyses of the cleavage rates of a variety of oligosaccharides by HEWL and human lysozyme indicated that no strain is involved in the binding of N-acetylglucosamine and N-acetylxylosamine in site D, and that the lactyl group of NAM rather than the hydroxymethyl group may be responsible for proposed strain at this subsite (Schindler et al., 1977).

Strynadka & James (1991) suggested that distortion of the NAM saccharide in site D of HEWL was caused by close packing of the O3 lactyl group and residues in the tightly hydrogen-bonded β-sheet region (Asp52, Asn46, Asn59, Ser50 and Asp48), by close contact between the hydroxymethyl group of ring D and the acetamido group of ring C and by the strong interactions between the NH group of Val109 and O6 in the quasiaxially oriented hydroxymethyl group. However, in the HEWL–MGM complex the O10 atom of the lactyl group is hydrogen bonded via two water molecules (see Fig. 9b) and not directly to the hydrogen-bonded 'platform'. It is, therefore, difficult to see how this group is so tightly bound to the protein that the NAM ring is forced into a deeper position in the active site and into a distorted high-energy conformation. It is possible to extend the O3 group in our observed position of the NAG unit in site D of RBTL by addition of a lactyl group without collisions with the β-sheet 'platform'.

In contrast to the weaker binding of many oligosaccharides to human lysozyme, the human enzyme binds (NAG–NAM)$_2$ more tightly than HEWL due to less strain in subsite D (Schindler *et al.*, 1977). The difference in the destabilizing effect of NAM in subsite D for the two enzymes may be related as much to the nature of subsite C as it is to the nature of subsite D. It is suggested that 'looser' contacts in subsite C in the human lysozyme which is very similar to the equivalent site in RBTL, compared to HEWL would explain why the NAM unit binds to site D with less strain in the former enzyme. Fig. 9(a) shows that the NAG units bind to HEWL and RBTL in the same way to site C. Thus, it is difficult to see how site C can affect the binding mode and conformation of a NAM saccharide in subsite D.

Even though it is hard to see why the NAM residue is bound more deeply into the active-site cleft of HEWL than is the NAG unit in RBTL, it is possible that the former represents the active transition-state complex, whereas the latter provides a view of an unproductive complex. In addition, the endocyclic O1 atom of the D saccharide in the RBTL complexes is in an axial and not in a equatorial position as expected in a productive complex. However, a superimposition of δ-lactone in sites A to D of HEWL (Ford *et al.*, 1974) and (NAG)$_4$ in RBTL (see Fig. 10c), shows that the sofa-formed D ring of δ-lactone which resembles the supposed transition-state intermediate, has the same orientation as the chair-formed NAG unit in our RBTL complex. There is reason, therefore, to question the well established thesis about the importance of strain in the lysozyme catalysis and the position of the D saccharide unit in a productive complex.

Binding to sites E and F

From the observed position of the β-anomer of the NAG residue bound to subsite D, we extended the tetrasaccharide by two additional sugar rings with chair conformation to sites E and F in the lower part of the cleft (see Fig. 10). The two extra saccharide units seem to make the best interactions and have most reasonable glycosidic link geometry when they lie on the left side of the cleft. Asp52, Asn44 and Asn46 make hydrogen bonds with saccharide E and and Asn44 interacts with sugar ring F. This result agrees well with earlier conformational energy calculations which indicated that oligosaccharides with an undistorted sugar ring positioned near the surface of subsite D, will exhibit a 'left-sided' binding mode to sites E and F (Pincus & Scheraga, 1979).

In an additional experiment, crystals of rainbow trout lysozyme complexed with (NAG)$_4$ were soaked with 5 mM (NAG)$_2$ for 24 h in the hope that the dimer would bind to the unoccupied E and F subsites. Unfortunately, the difference map only showed electron density of a tetramer in subsites A to D, but no sugar binding in the other two sites. Measurements of the free energies of association to the individual sites from A to F imply that the E site which has a binding energy of about -4 kcal mol^{-1}, should be the most preferred binding site after site C (in Imoto *et al.*, 1972). It is, therefore, surprising that the (NAG)$_2$ molecule does not bind to the E and F sites in RBTL–(NAG)$_4$ as these are not blocked by neighbouring molecules and are accessible in the crystal lattice. Co-crystallization with (NAG)$_6$ yielded an RBTL–(NAG)$_4$ complex – perhaps not surprising on the time scale of the crystallization process (about two weeks).

In the absence of firm crystallographic evidence, it is natural to query the hypothesis that the complexes between RBTL and (NAG)$_3$ and (NAG)$_4$ in common with the complexes between (NAG-NAM)$_2$ and the tortoise enzyme (in Johnson *et al.*, 1988) and the unrefined HEWL–MGM complex (Kelly *et al.*, 1979), represent unproductive initial Michaelis complexes in lysozyme catalysis. It is possible that productive binding and conversion to a reactive complex with a strained conformation of the sugar buried in site D, can only be achieved by interactions of two more saccharide units to the right side of subsites E and F.

Temperature factors

In common with the native structure, the regions in RBTL–(NAG)$_4$ that exhibit the greatest flexibility are found in the external loop regions from residue 45 to 50 and residue 65 to 75 which constitute the edges of the active-site cleft and, thus, presumably allow the binding sites to accommodate the incoming substrate.

Fig. 11 which shows a ribbon drawing of the RBTL–(NAG)$_4$ complex coloured according to the difference in temperature factor for Cα atoms between the native and the complexed forms for the lysozyme, indicates that only small regions in the N- and C-terminus show higher mobility (blue in Fig. 12) in the liganded structure. Areas exhibiting the greatest decreases in thermal motion (light red and yellow in Fig. 12) when the tetrasaccharide binds, are those from residue 60 to 73 in the long coiled loop region and residue 100 to 110 lining the active-site cleft. In these regions residues 62, 63, 101, 103, 107 and 109 make hydrogen bonds with the substrate (see Table 8) and it seems, therefore, likely that the motion is reduced by interaction with the substrate. Also a short loop from residue 83 to 88 lying between a 3$_{10}$ helix and the α-helix connecting the two domains of the lysozyme molecule, seems to stiffen upon sugar binding.

A significant reduction in temperature factors for the side chain of Trp62 was observed upon binding of (NAG)$_3$ to the hen egg-white lysozyme (Cheetham *et al.*, 1992). This is not seen in the structure of our

RBTL–saccharide complexes where Tyr62 only exhibits slightly decreased mobility when the saccharides bind. It seems likely that the stacking interactions between the Tyr62 in RBTL and the NAG pyranose ring bound in subsite B are weaker than those between Trp62 in HEWL and the equivalent saccharide ring.

Our crystallographic results agree with earlier dynamic trajectory simulations of free and substrate bound lysozyme from hen egg white which conclude that the fluctuations of residues 101 to 105 lining the active-site cleft are very large in the free enzyme but substantially reduced by the presence of $(NAG)_6$ (Post *et al.*, 1986). Residues lining the active-site cleft and in close contact with the saccharides also show decreased motion in the complexes of HEWL with bound ligands compared with the native enzyme (Strynadka & James, 1991; Cheetham *et al.*, 1992).

Conformational changes in RBTL

The atomic coordinates of the native RBTL were compared with those of the lysozyme molecule in the complexes with $(NAG)_3$ and $(NAG)_4$ using the program *SQUID* (Oldfield, 1992). The root-mean-square deviations between the native structure and both complexes for main-chain atoms and all protein atoms are 0.13 and 0.40 Å, respectively, showing that conformational changes upon saccharide binding are rather small. Since the mean positional error is estimated to be about 0.18 Å from Luzzati plots, the r.m.s. deviation for main-chain atoms may not be significant. Fig. 12, illustrating r.m.s. differences for main-chain atoms of native and $(NAG)_4$ bound RBTL, indicates that the greatest conformational differences are found in two specific regions. These are from Asp66 to Lys73 and Leu100 to Asp119. The latter is directly involved in carbohydrate binding. Residues 100–110 in the lysozyme structure become, as mentioned earlier, more ordered upon sugar binding, which indicates that changes in the positions of these amino acids are real.

Even though it does not directly make contacts with the carbohydrate substrate, the first of the two regions mentioned above (Asp66–Lys73) which is packed against the Asn59–Cys64 carbohydrate-binding loop, shifts in position as well as becoming more ordered. The carbohydrate-binding loop from residue Asn59 to Cys64 exhibits a slight shift in position with the side chain of Asn59, Tyr62 and Trp63 moving significantly toward the sugar units in site B and C on binding the saccharide. Conformational shifts in these regions of the lysozyme molecule are also found in the complexes between hen egg-white lysozyme and different oligosaccharides (Ford *et al.*, 1974; Strynadka & James, 1991; Cheetham *et al.*, 1992). In these complexes the largest conformational changes are seen for Trp62 and Trp63 which have a mean shift for the ring atoms of about 1.0 and 0.7 Å. respectively. This movement narrows the active-site cleft and enhances the protein–inhibitor interactions. Comparisons of the unliganded structure of RBTL with the liganded forms indicate that Tyr62 only moves about 0.5 Å towards the sugar ring in site B. As a consequence the active-site cleft does not close to as great an extent upon carbohydrate binding for RBTL as for HEWL. One explanation can be that the interaction between Tyr62 and the apolar face of the pyranose ring bound to site B is weaker than the interaction between Trp62 and the equivalent sugar in HEWL. Another point is that the hydrogen bond which is seen between the tryptophan and the sugar unit in the same site of the liganded form of HEWL, is no longer possible between the tyrosine and the pyranose ring bound to site C. As mentioned earlier, subsite C of human lysozyme which also contains a tyrosine in position 62, is 'looser' than the equivalent site in the hen egg-white enzyme (Schindler *et al.*, 1977).

Our crystallographic results correlate well with theoretical predictions employing normal mode analysis (Brooks & Karplus, 1985; Gibrat & Gö, 1990). These studies have shown that the motion of the two domains in lysozyme can be described as a hinge-bending motion, *i.e.* the two domains move as rigid bodies to open or close the active-site cleft slightly in order to bury the sugar molecules and enhance the interactions between the protein and saccharide.

The conformations of the side chains of the catalytically active residues (Glu35 and Asp52) are, to within the limits of accuracy of the structure determinations, identical in native RBTL and in the complexed forms of the enzyme.

Changes in the water structure in the active site

Comparison of the water molecules in the active-site cleft of the refined structure of RBTL and the position of the $(NAG)_4$ molecule bound to subsites A to D, shows that the tetrasaccharide displaces 13 water molecules upon binding. Sugars are highly polar organic molecules that are solvated in aqueous solution. Thus, in the formation of complexes, they exchange their solvation shell to form hydrogen-bonded interactions with polar groups in the binding site of the protein (Quiocho, 1986). Concomitantly, water molecules are displaced from the active-site cleft. The NAG moiety bound at site A displaces three water molecules, one positioned 0.43 Å from C3, one positioned 0.48 Å from O7 and one positioned 0.76 Å from C8. Three water molecules are also displaced from site B, one of which was 0.53 Å from the C6 position, one 0.54 Å from O5 and the last approximately 0.54 Å from O5 and 0.83 Å from C4. NAG unit C displaces four water molecules from the active site, positioned 0.67, 0.37, 0.55 and 0.57 Å from C1, O3, O4 and O7, respectively. Some of these displaced solvent molecules are hydrogen bonded to the protein in the apoenzyme, but these interactions were relinquished on the binding of the inhibitors in

favour of stronger interactions between the protein and the sugar molecules.

In the native structure the catalytically active amino acids, Glu35 and Asp52, are hydrogen bonded to each other through an extensive water-mediated network of hydrogen bonds (Karlsen *et al.*, 1995). This is illustrated in Fig. 13(*a*) where OD2 of Asp52 makes a hydrogen bond to water 202 which is hydrogen bonded to water 201. The latter water molecule is also hydrogen bonded to OE2 in the side chain of Glu35. This hydrogen-bonded network is lost in the liganded form of the enzyme (see Fig. 13*b*). Water molecules 201 and 202 are displaced by the NAG unit bound to site *D*, water 201 positioned 1.15 Å from the axial O1 and water 202 positioned exactly at O5. A third water molecule 0.19 Å from O6, is also removed from the active-site cleft. Only

one water molecule (number 142 in native and 143 in the complexed structure) is conserved in the native and both (NAG)$_3$ and (NAG)$_4$ bound structures, and is hydrogen bonded to OD2 of Asp52 and OE1 of Gln57.

Formation of a covalent enzyme–glycosyl intermediate

Nucleophilic attack (SN2) by a carboxylate in glycosyl bond cleavage where a covalent bond is formed between this group and the C1 sugar atom in a glycosyl-enzyme intermediate, is well established for (e−>e)β-glycosidases (Sinnott, 1990). Support for this mechanism has recently been obtained from studies of mutant forms of the T4 lysozyme (Hardy & Poteete, 1991) and HEWL (Lumb *et al.*, 1992) which suggest that the 'catalytic' aspartates have a significant nucleophilic rather merely
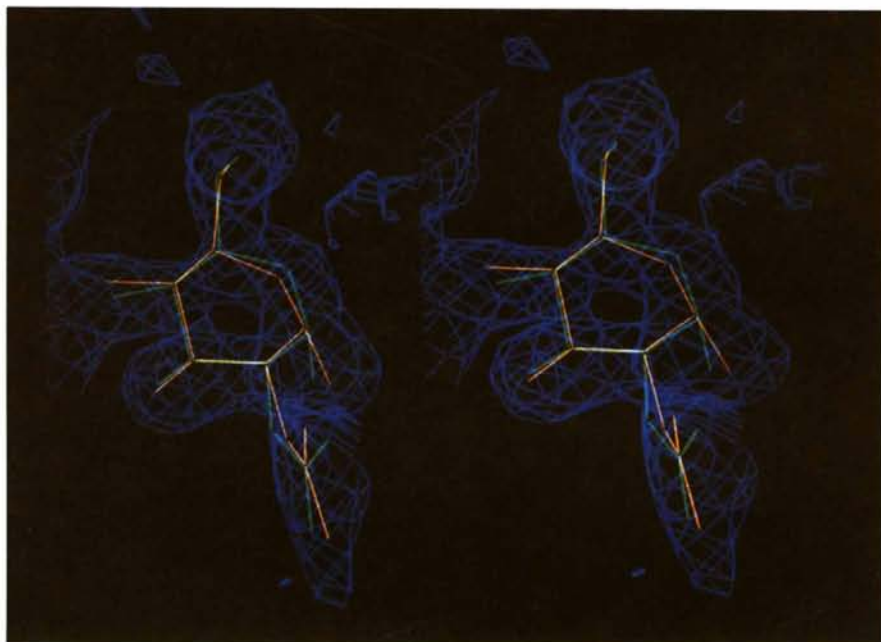

Fig. 6. $F_o - F_c$ omit map contoured at 0.12 e Å^{-3} for the refined NAG ring in site *D* (green). A model of a pyranose ring (orange) with sofa conformation is included for comparison.
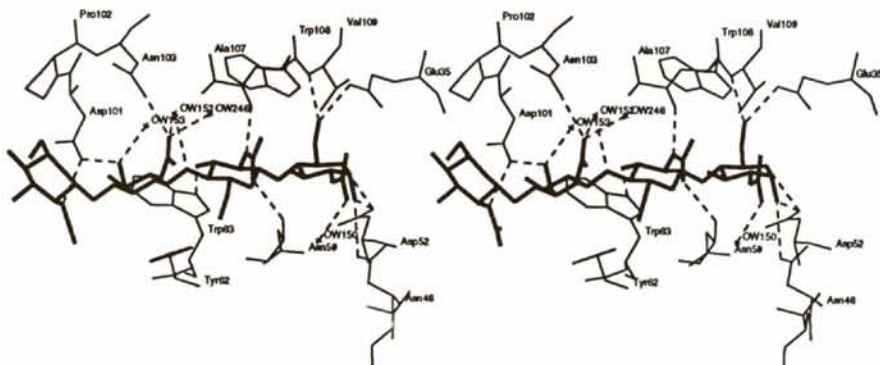

Fig. 7. Hydrogen-bonding interactions between protein atoms and sugar residues within sites *A* to *D* in RBTL. Lysozyme structure is shown with thin lines, sugar residues with thick lines and hydrogen bonds with broken lines. Water molecules are depicted with crosses.

an electrostatic function. The only crystallographic evidence for a covalent enzyme–substrate intermediate has very recently been observed for a complex between a NAG–NAM saccharide linked to a peptide fragment from the *Escherichia coli* cell wall and the T4 lysozyme with the mutation Thr26→Glu (Kuroki, Weaver & Matthews, 1993).

In our RBTL–(NAG)$_4$ complex, the distance between C1 of the saccharide unit in subsite *D* and OD2 in the side chain of Asp52 is 3.2 Å. Adjustments of the the χ_1 and χ_2 angles of the side chain can reduce the distance to 2.8 Å. Formation of a covalent bond between these two atoms during lysozyme catalysis thus seems unlikely. It is possible, however, that binding of two extra saccharides in subsites *E* and *F* might force the site *D* saccharide closer to the 'catalytic' aspartate. Molecular dynamic studies have shown that the two domains of lysozyme move as rigid bodies to slightly open and close the active-site cleft (Brooks & Karplus, 1985; Gibrat & Gö, 1990; Kidera, Inaka & Matsushima, 1992). Because of this flexibility, the whole β-sheet region, of which Asp52 is a part, might approach the bound sugar to enable the carboxylate group of Asp52 to act as a nucleophile. As this β-sheet is conserved in all types of lysozymes known so far (in Weaver *et al.*, 1985), it seems reasonable that it has an important function in the enzyme catalysis.

Concluding remarks

From detailed analysis of the refined structures of complexes between RBTL and chito-oligosaccharides, we can conclude that the saccharides are bound within the active-site cleft of the enzyme by an extensive network of hydrogen bonds between the sugar molecules and protein atoms within the recognition sites. Hydrophobic stacking interactions between aromatic side chains and the non-polar faces of the pyranose rings seem to be important for substrate specificity and stabilization of the complexes, and is observed to be a general feature for protein–saccharide binding (Quiocho, 1986).

The chito-oligosaccharides bind to sites *A*, *B* and *C* in the active-site cleft of RBTL in the same way and with normal chair conformations as observed for complexes between HEWL and a variety of saccharides (Ford *et al.*, 1974; Strynadka & James, 1991; Cheetham *et al.*, 1992). The NAG ring positioned in site *D* is only slightly perturbed from the 'ideal' chair conformation and is not buried so deeply in the cleft as the equivalent NAM unit in HEWL–MGM complex where the ring has a strained half-chair conformation.

Our crystallographic results and comparisons with previously solved structures of HEWL–oligosaccharide complexes, raise doubt about the well established mechanism of lysozyme action. Firstly, it not clear whether a
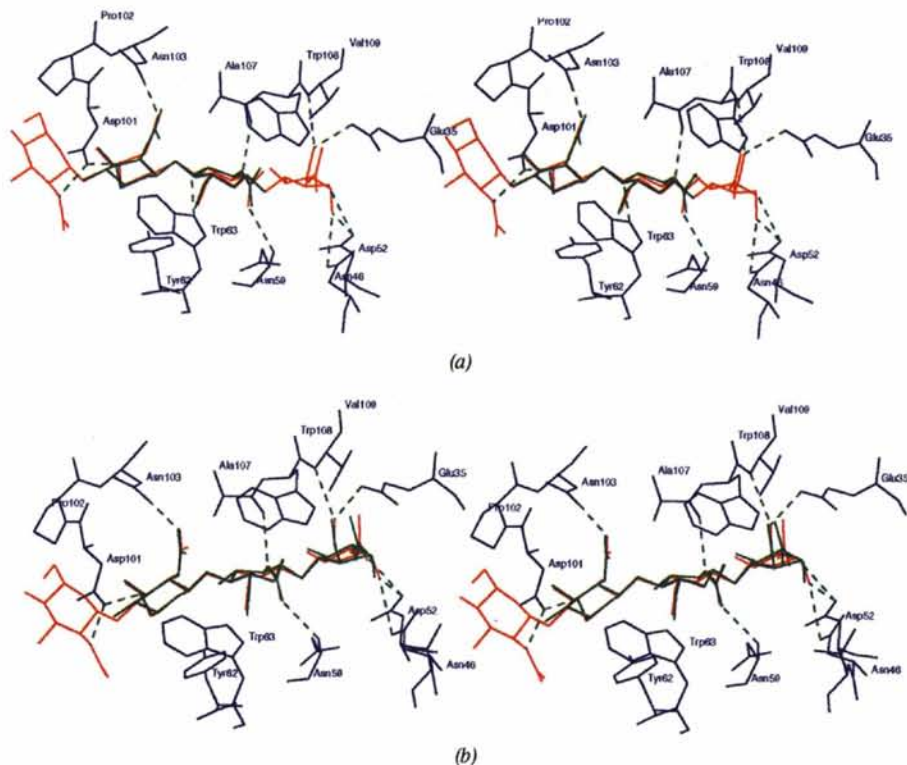

Fig. 8. Superimpostions of (*a*) the (NAG)$_2$ (green) and (*b*) the (NAG)$_3$ (green) molecule on (NAG)$_4$ (red) in sites *B* and *C* and *B*, *C* and *D* of RBTL, respectively.

distorted sugar conformation in site D is necessary for effective catalysis. Secondly, the exact position of the saccharide in site D in an active transition-state complex seems to be unclear. The site D NAM ring of MGM is bound closer to the catalytic Glu35 in HEWL than is the equivalent NAG ring in RBTL, which may imply that this is a productive complex. The site D NAG ring in RBTL, on the other hand, is bound in the same way as the N-acetyl lactone ring in the proposed transition-state analogue, δ-lactone.

Thirdly, the mode of binding to subsites E and F and importance of these sugar residues for effective catalysis is somewhat uncertain. Our, admittedly simple, model building shows that we can readily extend the tetrasaccharide by two extra sugar units from the observed position of the NAG saccharide in site D of RBTL, only when these occupy the left side of subsites E and F. The two extra sugar rings can be placed in the right-hand side of the lower part of the active-site cleft, thereby forcing the sugar residue in site D into a 'deeper' and strained
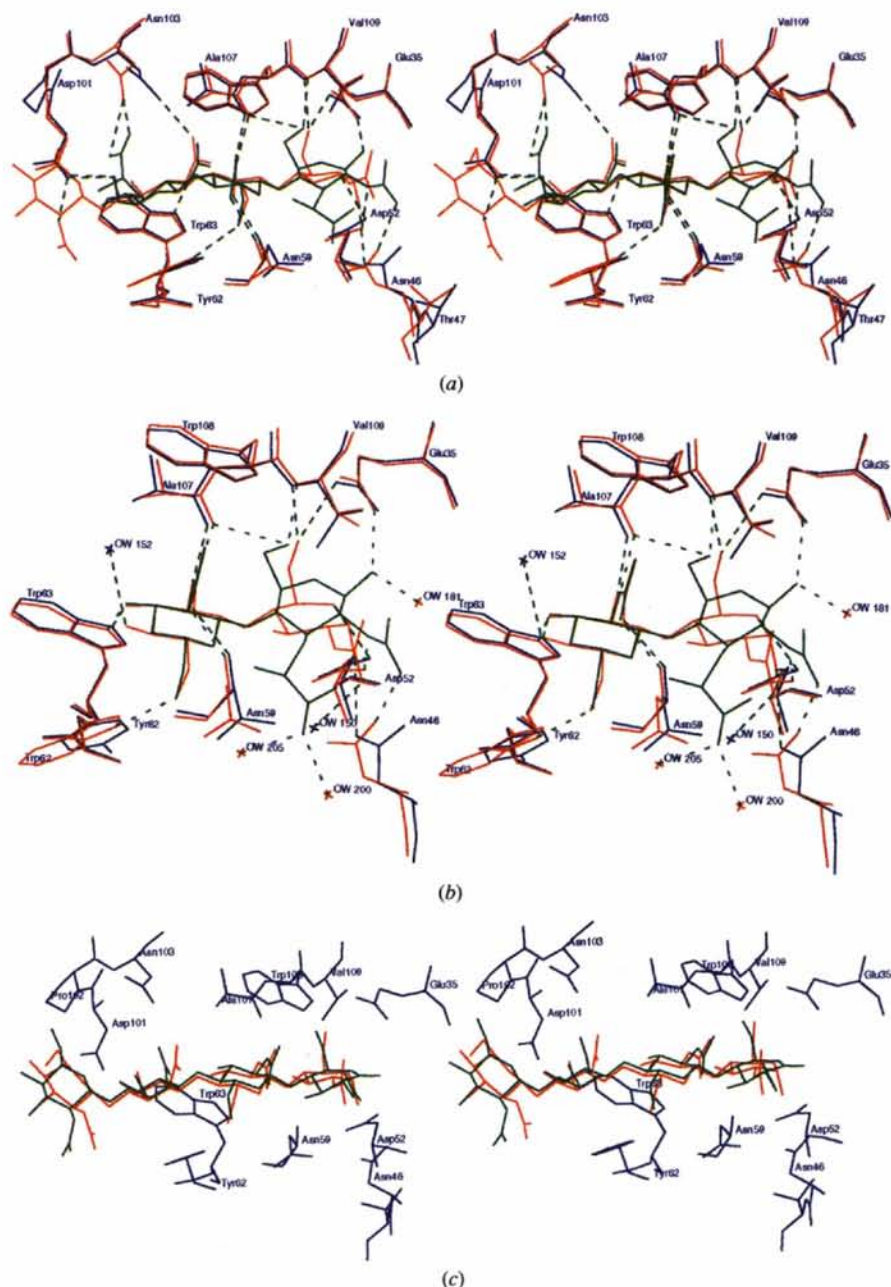

(a)

(b)

(c)

Fig. 9. (a) A superimposition of MGM (NAM–NAG–NAM) (green) bound to HEWL (red) on (NAG)$_4$ (orange) in RBTL (blue). (b) A closer view of the saccharides in the C and D site of HEWL and RBTL (same colours as in (a). (c) A superimposition of δ-lactone (green) on (NAG)$_4$ (red) in sites A to D in RBTL.
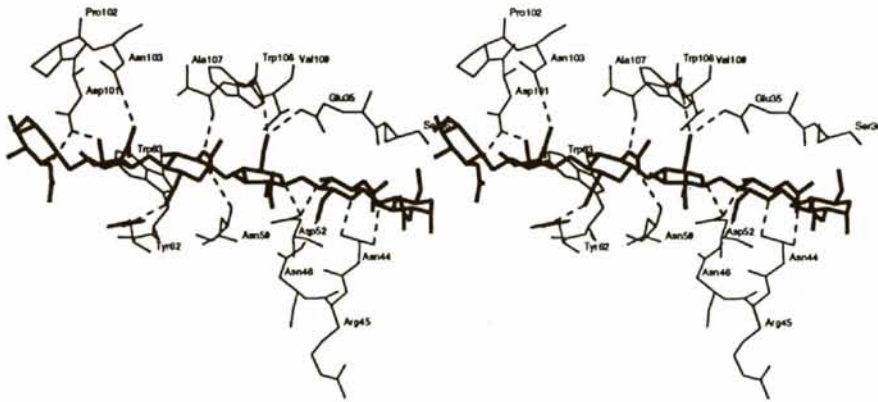
Fig. 10. Proposed binding of a hexa-saccharide (thick lines) of chitin in the active-site cleft of RBTL. The saccharides in sites A to D have the crystallographically determined positions.
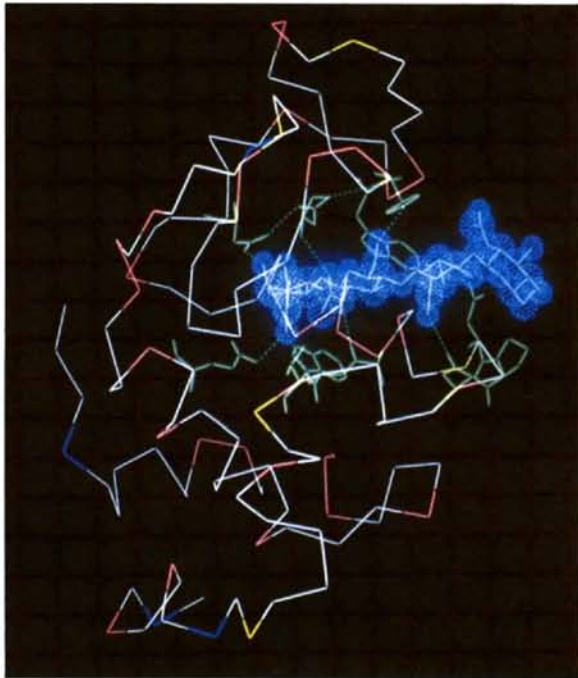

Fig. 11. Cα-drawing of the RBTL–(NAG)₄ complex coloured according to differences in temperature factors between the native and the liganded form of the enzyme. The molecule is coloured as follows: $\Delta B < -3 \, \text{Å}^2$ (blue), $-3 \, \text{Å}^2 < \Delta B < 3 \, \text{Å}^2$ (light blue), $3 \, \text{Å}^2 < \Delta B < 12 \, \text{Å}^2$ (light red) and $\Delta B > 12 \, \text{Å}^2$ (yellow). Residues that interact directly with the ligand are marked in light green and the saccharide is depicted with blue van der Waals spheres.

Fig. 12. R.m.s. differences along the main-chain atoms between the native and the structure of RBTL with bound (NAG)₄.

binding mode to the lower part of the active-site groove, is hydrolysed through a covalently attached intermediate as proposed by Koshland (1953).

conformation but we feel that this is not supported by the majority of the available experimental evidence.

Our crystallographic studies on the RBTL–oligo-saccharide complexes do not, as seems to be the case with most experimental studies on lysozyme catalysis, support Phillips (Phillips, 1966) proposed mechanism of lysozyme action. We feel that it is equally likely that the substrate which binds with all six saccharide units in an unstrained chair conformation and shows a left-sided
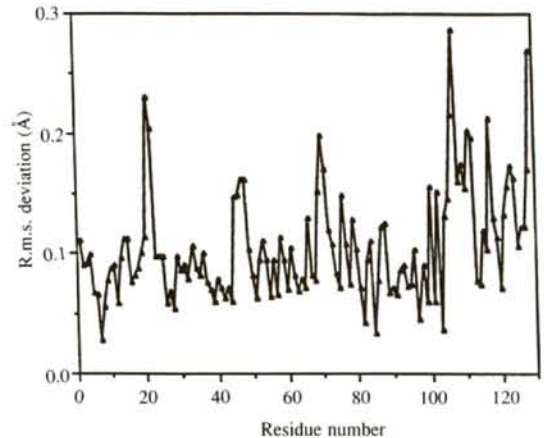
References

ANDERSON, W. F., GRÜTTER, M. G., REMINGTON, S. J., WEAVER, L. H. & MATTHEWS, B. W. (1981). *J. Mol. Biol.* **147**, 523–543.

BANERJEE, S. K., HOLLER, E., HESS, G. P. & RUPLEY, J. A. (1975). *J. Biol. Chem.* **250**, 4355–4367.

BLAKE, C. C. F., JOHNSON, L. N., MAIR, G. A., NORTH, A. C. T., PHILLIPS, D. C. & SARMA, V. R. (1967). *Proc. R. Soc. London Ser. B,* **167**, 378–388.
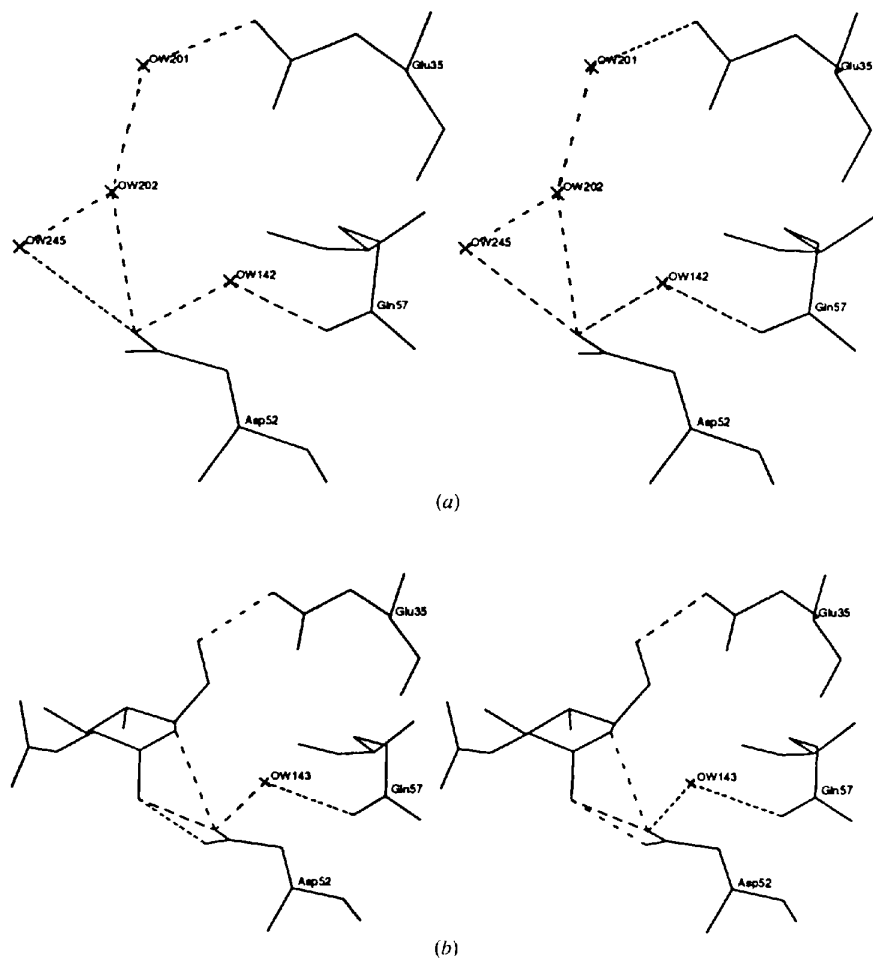
Fig. 13. Ordered water molecules within the D binding site in the active-site cleft of unliganded (a) and (NAG)$_4$-bound (b) RBTL. Water molecules are marked with crosses.

BLAKE, C. C. F. (1967). *Proc. R. Soc. London Ser. B*, **167**, 435–438.

BROOKS, B. & KARPLUS, M. (1985). *Proc. Natl Acad. Sci. USA*, **82**, 4995–4999.

COLLABORATIVE COMPUTATIONAL PROJECT, NUMBER 4 (1994). *Acta Cryst.* D50, 760–763.

CHEETHAM, J. C., ARTYMIUK, P. J. & PHILLIPS, D. C. (1992). *J. Mol. Biol.* **224**, 613–628.

CHIPMAN, D. M., GRISARO, V. & SHARON, N. (1967). *J. Biol. Chem.* **242**, 4388–4394.

DAUTIGNY, A., PRAGER, E. M., PHAM-DINH, D., JOLLÈS, J., PAKDEL, F., GRINDE, B. & JOLLÈS, P. J. (1991). *Mol. Evol.* **32**, 187–198.

FORD, L. O., JOHNSON, L. N., MACHIN, P. A., PHILLIPS, D. C. & TJIAN, R. (1974). *J. Mol. Biol.* **88**, 349–371.

GIBRAT, J.-F. & GO, N. (1990). *Proteins Struct. Funct. Genet.* **8**, 258–279.

GRINDE, B. (1989a). *J. Fish Dis.* **12**, 95–104.

GRINDE, B. (1989b). *FEMS Microbiol. Lett.* **60**, 179–182.

GRINDE, B., LIE, Ø., POPPE, T. & SALTE, R. (1988). *Aquaculture*, **68**, 299–304.

GRINDE, B., JOLLÈS, J. & JOLLÈS, P. (1988). *Eur. J. Biochem.* **173**, 269–273.

HAM, J. T. & WILLIAMS, D. G. (1970). *Acta Cryst.* B26, 1373–1383.

HARDY, L.W. & POTEETE, A.R. (1991). *Biochemistry*, **30**, 9457–9463.

HENDRICKSON, W. A. (1985). *Methods Enzymol. B*, **115**, 252–270.

HIROTSU, K. & SHIMADA, A. (1974). *Bull. Chem. Soc. Jpn*, **47**, 1872–1879.

IMOTO, T., JOHNSON, L. N., NORTH, A. C. T., PHILLIPS, D. C. & RUPLEY, J. A. (1972). *Vertebrate Lysozymes. The Enzymes*, Vol. VII, pp. 665–868. New York & London: Academic Press.

INAKA, K., MATSUSHIMA, M. & MORIKAWA, K. (1990). *Acta Cryst.* A46, C82.

JEFFREY, G. A. (1990). *Acta Cryst.* B46, 89–103.

JOHNSON, L. N., CHEETHAM. J., MCLAUGHIN, P. J., ACHARYA, K. R., BARFORD, D. & PHILLIPS, D. C. (1988). *Curr. Top. Microbiol. Immunol.* **139**, 81–134.

JONES, T. A. (1985). *Methods Enzymol.* **115**, 157–170.

KABSCH, W. (1988). *J. Appl. Cryst.* **21**, 67–71.

KARLSEN, S., ELIASSEN, B. E., HANSEN, L. K., LARSEN, R. L., RIISE, B. W., SMALÅS, A. O., HOUGH, E. & GRINDE, B. (1995). *Acta Cryst.* D51, 354–367.

KELLY, J. A., SIELECKI, A. R., SYKES, B. D. & JAMES, M. N. G. (1979). *Nature (London)*, **282**, 875–878.

KIDERA, A., INAKA, K. & MATSUSHIMA, M. (1992). *J. Mol. Biol.* **225**, 477–486.

KOSHLAND, D. E. (1953). *Biol. Rev.* **28**, 416–420.

KUMAGAI, I., MAENAKA, K., SUNADA, F., TAKEDA, S. & MIURA, K.-I. (1993). *Eur. J. Biochem.* **212**, 151–156.

KUMAGAI, I., SUNADA, F., TAKEDA, S. & MIURA, K.-I. (1992). *J. Biol. Chem.* **267**, 4608–4612.

KUROKI, R., WEAVER, L. H. & MATTHEWS, B. W. (1993). *Science*, **262**, 2030–2033.

LASKOWSKI, R. A., MACARTHUR, M. W., MOSS, D. S. & THORNTON, R. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

LOWE, G. (1967). *Proc. R. Soc. London Ser. B*, **167**, 431–434.

LUMB, K. J., APLIN, R. T., RADFORD, S. E., ARCHER, D. B., JEENES, D. J., LAMBERT, N., MACKENZIE, D. A., DOBSON, C. M. & LOWE, G. (1992). *FEBS Lett.* **296**, 153–157.

Luzzati, P. V. (1952). *Acta. Cryst.* **5**, 802–810.

Matsushima, K., Inaka, K. & Morikawa, K. (1990). *Acta Cryst.* **A46**, C-82.

Mo, F. & Jensen, L. H. (1978). *Acta Cryst.* **B34**, 1562–1569.

Moran, R. A. & Richards, G. F. (1973). *Acta Cryst.* **B29**, 2770–2783.

Oldfield, T. J. (1992). *J. Mol. Graphics*, **10**, 247–252.

Phillips, D. C. (1966). *Sci. Am.* **215**, 78–90.

Phillips, D. C. (1967). *Proc. Natl Acad. Sci. USA*, **57**, 484–495.

Pincus, M. R. & Scheraga, H. A. (1979). *Macromolecules*, **12**, 633–644.

Pincus, M. R. & Scheraga, H. A. (1981). *Biochemistry*, **20**, 3960–3965.

Plugrath, J. W. (1989). *MADNES, Munich Area Detector NE System Users Guide*. Cold Spring Harbor Laboratory Press.

Post, C. B., Brooks, B. R., Karplus, M., Dobson, C. M., Artymiuk, P. J., Cheetham, J. C. & Phillips, D. C. (1986). *J. Mol. Biol.* **190**, 455–479.

Quiocho, F. A. (1986). *Ann. Rev. Biochem.* **55**, 287–315.

Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963). *J. Mol. Biol.* **7**, 95–99.

Salton, M. R. J. & Ghuysen, M. J. (1959). *Biochim. Biophys. Acta*, **36**, 552–554.

Salton, M. R. J. & Ghuysen, M. J. (1960). *Biochim. Biophys. Acta*, **45**, 355–363.

Sarma, R. & Bott, R. (1977). *J. Mol. Biol.* **113**, 555–565.

Schindler, M., Assaf, Y., Sharon, N. & Chipman, D. M. (1977). *Biochemistry*, **16**, 423–431.

Sinnott, M. L. (1990). *Chem. Rev.* **90**, 1171–1202.

Strynadka, N. C. J. & James, M. N. G. (1991). *J. Mol. Biol.* **220**, 401–424.

Vyas, N. K. (1991). *Curr. Op. Struct. Biol.* **1**, 732–740.

Weaver, L. H., Grütter, M. G., Remington, S. J., Gray, T. M., Isaacs, N. W. & Matthews, B. W. (1985). *J. Mol. Evol.* **21**, 97–111.