Structures of partridge egg-white lysozyme with and without tri-*N*-acetylchitotriose inhibitor at 1.9 Å resolution

M.A. TURNER¹ AND P.L. HOWELL^{1,2}

¹ Division of Biochemistry Research, Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada

² Department of Biochemistry, Medical Sciences Building, University of Toronto, Toronto, Ontario M5S 1A8, Canada

(RECEIVED October 7, 1994; ACCEPTED December 7, 1994)

Abstract

The three-dimensional structures of native partridge egg-white lysozyme (PEWL) and PEWL complexed with tri-*N*-acetylchitotriose inhibitor have been determined crystallographically and refined at 1.9 Å resolution. Crystals of native and complexed protein are isomorphous and have space group and cell dimensions that are identical to those of hen egg-white lysozyme (HEWL) under similar crystallization conditions. Full occupancy of the trisaccharide in the inhibitor complex has allowed definitive modeling and refinement of all three sugar residues, located at subsites A, B, and C in the PEWL active site. A comparison has been made with HEWL/inhibitor complexes in which coordinates were either not refined (Blake CCF, et al., 1967, *Proc R Soc B* 167:378–388) or were refined at partial occupancy (Cheetham JC, Artymiuk PJ, Phillips DC, 1992, *J Mol Biol* 224:613–628). Although the loop comprising residues 70–75 is located on the surface of the protein and not near the active site, it appears to be affected indirectly by trisaccharide binding such that the loop shifts toward the active site and becomes relatively immobilized. The source of this loop movement appears to be the anchoring of Trp 62, located in the active site cleft, as it forms a hydrogen bond with O6 of the *N*-acetylglucosamine at site C. Good electron density for the trisaccharide in the PEWL complex structure shows that Asp 101 is involved in hydrogen bonding interactions with the terminal sugar residue.

Keywords: partridge egg-white lysozyme; protein/inhibitor complex

The structures of a number of avian lysozymes are being examined to assess conditions that may be suitable for potential time-resolved Laue experiments (Howell et al., 1992; Howell, 1995; M.A. Turner & P.L. Howell, unpubl. results). For example, under virtually identical conditions of buffer and NaCl precipitant, lysozymes from hen and turkey egg whites crystallize in tetragonal (P₄,2₁,2) and hexagonal (P₆,22) space groups with vastly different intermolecular packing. Although the structures of type c enzymes from various species are generally similar at the molecular level (hen egg-white lysozyme [HEWL], Blake et al., 1967; turkey egg-white lysozyme [TEWL], Howell et al., 1992; pigeon egg lysozyme, Yao et al., 1992; Japanese quail egg lysozyme, Chitarra et al., 1993; guinea hen egg lysozyme and pheasant egg lysozyme, Lescar et al., 1994), different amino acid sequences may influence the packing arrangement of those molecules within the unit cell. The partridge egg-white lysozyme (PEWL) sequence differs from that of HEWL by three residues:

HEWL Ala 11, Asp 18, and Glu 121 are Val, Asn, and His, respectively, in PEWL. Residues 18 and 121 are located on the molecular surface and have the potential to confer packing preference.

The enzymatic mechanism by which lysozyme catalyzes the cleavage of the β-1,4-glycosidic linkage of its natural substrate, a polysaccharide of alternating *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (GlcMur) residues, was proposed for HEWL on the basis of an early unrefined 2-Å structure of a tri-*N*-acetylglucosamine complex and subsequent modeling of the hexasaccharide substrate into the active site. This model required the binding of sugar residues at subsites A–F (Blake et al., 1967; Phillips, 1967). On the basis of this model, an enzymatic mechanism was suggested that required puckering of ring D to relieve substrate/protein contacts within the active site. Distortion of the D sugar ring to a higher energy half chair conformation was proposed to be important mechanistically in weakening the scissile C1–O4 bond of the glycosidic linkage between rings D and E. The structure of HEWL complexed with a transition-state analogue, tetrasaccharide-lactone ((GlcNAc)₄-lactone), was solved to test the proposed mechanistic requirement for D-ring puck-

Reprint requests to: P.L. Howell, Division of Biochemistry Research, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada; e-mail: howell@resunix.ri.sickkids.on.ca

ering (Ford et al., 1974). More recently, a refined high-resolution structure of the HEWL complex with the alternating trisaccharide GlcMur-GlcNAc-GlcMur confirmed the distortion of the D sugar ring (Strynadka & James, 1991). In this case, steric requirements of the larger muramic acid residues restrict this inhibitor to binding exclusively at subsites B, C, and D.

Examination of previously reported lysozyme/inhibitor models was intended to serve as a source of reference during interpretation of maps calculated from time-resolved Laue data. However, because only unrefined coordinates were available for (GlcNAc)₄-lactone (Ford et al., 1974) and β -GlcNAc and (GlcNAc)₃ (Imoto et al., 1972) inhibitor models, a high-resolution structure determination of PEWL/(GlcNAc)₃ was undertaken. During the course of this work, results from the refined HEWL/(GlcNAc)₃ complex became available where the trisaccharide was modeled at partial occupancy and the structure refined against data to 1.75 Å resolution (Cheetham et al., 1992). Discrepancies between the unrefined models (Imoto et al., 1972; Ford et al., 1974), this work, and the partial occupancy (GlcNAc)₃ model (Cheetham et al., 1992) were immediately apparent. Overall RMS differences in sugar coordinates ranged from 0.9 to 1.3 Å, with discrepancies at site A ranging from 1.2 to 1.7 Å between unrefined and partially occupied models. These differences in the placement of inhibitor models gives rise to alternative protein/carbohydrate interaction schemes that are outlined in Table 1.

Results and discussion

Quality of the structures

In both structures, main-chain $2|F_o| - |F_c|$ electron density was continuous. No ϕ/ψ angles fall outside the allowed regions on a Ramachandran plot. Both coordinate sets give "better" geometry statistics than those used to form the database on which the PROCHECK suite of programs for geometry analysis is based (Laskowski et al., 1993). Agreement of the model with structural data according the method of Luzzati (1952) indicates mean coordinate error in the range of 0.13–0.23 Å and 0.11–0.19 Å for the native and complex structures, respectively.

Structure of PEWL

The overall structure of PEWL is similar to that of HEWL (PDB 6LYZ; Diamond, 1974), with average RMS main-chain (including carbonyl oxygen atoms) deviations of 0.29 Å. Zones of largest variation occur in the flexible loops, especially residues 70–75 as well as at residue 102. The three amino acid differences between PEWL and HEWL have no effect on the overall structure; the Val 11 residue of PEWL that substitutes for HEWL Ala is accommodated without structural rearrangement and Asp 18 and Glu 121 of HEWL, which are Asn and His, respectively, in PEWL, are located on the surface of the

Table 1. Intermolecular distances (Å) involving sugar residues

Direct H bonds						
Sugar atom	Protein atom	PEWL/(GlcNAc) ₃	HEWL/(GlcNAc) ₃ ^a	HEWL/(GlcNAc) ₃ ^b	HEWL/ (GlcNAc) ₄ -lactone ^c	HEWL/ GlcMur-GlcNAc- GlcMur ^d
Ring C	O7	Asn 59 N	2.81	2.8	2.7	2.9
	O6	Trp 62 NE1	2.76	2.9	2.7	2.9
	O3	Trp 63 NE1	3.12	2.9	3.3	3.0
	N2	Ala 107 O	2.78	2.8	3.1	2.9
Ring B	O6	Asp 101 OD1	2.48	2.4	2.6	2.4
	O3	Asp 119 OD2		3.0	3.0	
Ring A	N2	Asp 101 OD1	2.97			
	O3	Arg 125 NH2 ^c	3.02			
	O6	Asn 103 OD1		2.4		
O3 (C ring)–O5 (B ring)			2.93	3.2	3.0	3.1
O3 (B ring)–O5 (A ring)			2.85	3.5	2.3	2.6
Water-mediated contacts						
Sugar atom	Water	Protein atom	PEWL/(GlcNAc) ₃	HEWL/(GlcNAc) ₃ ^a		
Ring C	O1	Wat 307	Gln 57 O	2.49	2.89	
	O6	Wat 250	Asp 48O D2	3.04	2.85	
			Asn 59O D1		2.99	
	O6	Wat 299	His 121 NE2 ^c	2.91	2.82	
Ring A	O4	Wat 304	Asp 101 O	2.72	2.91	

^a Cheetham et al. (1992).

^b Imoto et al. (1972) (unrefined structure).

^c Ford et al. (1974) (unrefined structure).

^d Strynadka and James (1991). Note that GlcMur-GlcNAc-GlcMur binds at sites B, C, and D due to the larger steric requirements of the muramic acid residues (Blake et al., 1967).

^e In neighboring asymmetric unit.

Fig. 1. Plot of the RMS main- and side-chain distances for PEWL versus PEWL/(GlcNAc)₃. Main chain is defined as C α , C, and N atoms. RMS main- and side-chain differences are represented by bold and thin lines, respectively. Residues 59, 62, 63, 101, 107, and 125, which are involved in direct contact with (GlcNAc)₃, are highlighted with an asterisk.

molecule. None of these residues are involved in intermolecular interactions within the crystal nor are they located in the vicinity of the active site.

Given the high degree of similarity of structure and cell packing between HEWL and PEWL, it is not surprising that the active site of PEWL is also occluded as a result of intermolecular contacts. Thus, this PEWL crystal form is an unlikely candidate for time-resolved Laue studies. However, in light of the discrepancies in earlier inhibitor models, we chose to continue with these refinements in order to supply ourselves with reliable structures to use as a source of reference for future data and map evaluation.

Comparison of native and inhibitor complex structures

Overall, little structural change is apparent upon binding of trisaccharide with average RMS main- and side-chain differences of 0.22 and 0.49 Å, respectively (see Fig. 1). An overall ordering of the lysozyme structure is typical of complexation with sac-

charide inhibitors (Ford et al., 1974; Strynadka & James, 1991; Cheetham et al., 1992). Although $\Delta B_{\text{native-complex}}$ for main-chain PEWL atoms is only -3.5 Å^2 , the quality of electron density was improved in the complex as compared to native maps. The region with the most striking differences in main-chain positions is the loop consisting of residues 70–75. Strynadka and James (1991) noted a decrease in average main-chain B -factor of approximately -10 Å^2 over this loop upon formation in the HEWL/GlcMur-GlcNAc-GlcMur complex. In both native and complexed PEWL, average main-chain B -factors over this loop ($B_{\text{ave}} = 26 \text{ Å}^2$) are similar to those observed in the HEWL/GlcMur-GlcNAc-GlcMur complex ($\sim 29 \text{ Å}^2$) thus a similar decrease in B -factor upon inhibitor binding is not observed here. This is attributed to relatively lower B -factors in the native PEWL structure as compared to the native HEWL structure.

A small main-chain movement ($\Delta C\alpha$; 0.5 Å) and lower average B -factor ($\Delta B_{\text{main-chain}}$; -4.21 Å^2) at Asp 101 is also observed upon inhibitor binding. Asp 101 is located in a flexible (Post et al., 1986) lip region at the "A" end of the active site and makes both direct and water-mediated hydrogen bonding contacts with sugar residues in the complex.

Of interest are the movements of side chains within the active site. In particular, the side chain of Trp 62 is stabilized (ΔB ; -8.4 Å^2) in the saccharide complex relative to the native structure after moving through 19° about the $C\beta$ position to give an average side-chain shift of 0.91 Å relative to the native conformation. Main-chain temperature factors for residues 60–62 decrease from 17.7 to 11.9 Å² for native and inhibitor-bound structures, respectively. Figure 2 illustrates what appears to be the long-range effect of this interaction. Trp 62 is positioned at the apex of a turn where, upon interaction with the inhibitor, it moves ($\Delta C\alpha$; 0.5 Å) toward the active site cavity, "pulling" the strand of β -structure immediately preceding it (Arg 61, $\Delta C\alpha$; 0.4 Å and Ser 60, $\Delta C\alpha$; 0.3 Å) down as well. The shift of residues 60–62 and maintenance of hydrogen bonds through Arg 61 O...Arg 73 N and Trp 60 O...Lys 75 N may cause a sufficient perturbation to result in the peptide flip (relative to native) at Arg 73 also observed in the HEWL/GlcMur-GlcNAc-GlcMur complex system (Strynadka & James, 1991). The flipped Arg 73 O is then in position to hydrogen bond to a water molecule (Wat 220) not present in the native structure. In order to accommodate this

Fig. 2. Stereo view showing the effect of inhibitor binding on the conformation of Trp 62 and neighboring residues and the subsequent effect on loop 70–75. Native PEWL is represented by thin lines and the PEWL/(GlcNAc)₃ complex by bold lines.

water, the Ser 72 side chain moves out of the way, also forming a hydrogen bond with water 220 and resulting in a 1.0-Å shift in the C α coordinates of Ser 72. In the complex, the side chain of Arg 61 is also hydrogen bonded directly to loop 70–75 through Arg 61 NH2...Asp 48 O, Arg 61 NH1...Pro 70 O, and Arg 61 NE...Asp 48 OD2. In the native structure, Arg 61 NH2 is involved in a water-mediated hydrogen bond to Pro 70 O. With the maintenance of the main-chain hydrogen bonds, the repositioning of Trp 62 appears to be responsible for distortion of loop 70–75 and its slight shift toward the center of the molecule.

Large RMS differences in side-chain positions are observed for residues Arg 14, Asn 18, Arg 21, His 121, and Arg 128, which are located on the protein surface. These differences are not significant to the active site and are not involved in packing effects.

The sugar sites

Figure 3, showing the $2|F_o| - |F_c|$ density calculated from the final model, illustrates convincing density for all three sugar residues. Note that site A is also particularly well modeled. This site in inhibitor complex structures of HEWL has often been difficult to model, a result of either low occupancy and/or disordered binding (Ford et al., 1974; Cheetham et al., 1992). However, under the crystallization conditions used for the PEWL/(GlcNAc)₃ complex, full occupancy was expected and observed. The average isotropic temperature factors for each sugar residue are 13.67 Å², 21.91 Å², and 28.47 Å² for rings C, B, and A, respectively.

The trisaccharide assumes an open conformation and exhibits all geometric features typical of small molecule β 1,4-linked oligosaccharide crystal structures (Chu & Jeffrey, 1967; Jeffrey, 1990). Torsion angles about the β -glycosidic linkage (ϕ : O5-C1-O4'-C4'; -73.8° [B-C], -72.3° [A-B] and ψ : C1-O4'-C4'-C5'; -123.2° [B-C], -121.1° [A-B]) fall within typical ranges for undistorted linear polysaccharides (Jeffrey, 1990) and values for helical twist, ψ_H , (42.5° [B-C], 46.2° [A-B])³ are consistent with geometries that allow formation of the interresidue hydrogen bond between O3(*n*) and O5(*n* - 1) characteristic of β 1,4 linkages. The sugar residue at site C is the β -anomer and is essentially identical to that reported for the HEWL/(GlcNAc)₃ complex (Imoto et al., 1972; Strynadka & James, 1991; Cheetham et al., 1992) and also for the corresponding monosaccharides in HEWL complexes of β -GlcNAc and β -Me-GlcNAc (Imoto et al., 1972; Perkins et al., 1978).

Average RMS differences in sugar coordinates between PEWL and HEWL (Cheetham et al., 1992) complexes based on least-squares alignment of the corresponding protein molecules are 0.35, 0.70, and 1.4 Å for rings C, B, and A, respectively. The RMS differences become greater as the geometry of the saccharide in the partially refined HEWL complex structure deviates further from small molecule observations. The ϕ -value for the A-B linkage in the refined HEWL/(GlcNAc)₃ structure (-38.99°) does not fall within the typically observed range and both O3(*n*)...O5(*n* - 1) distances are longer (O3[C]...O5[B];

Fig. 3. (GlcNAc)₃ trisaccharide model and final electron density map calculated with coefficients $2|F_o| - |F_c|e^{i\phi}$. Sugar residues are labeled according to lysozyme subsite (see text). Map is contoured at 1.2 σ , which corresponds to $1.9e^{-3}/\text{\AA}^3$.

3.2 Å and O3[B]...O5[A]; 3.5 Å) than ideal hydrogen bond contacts. The geometry of the trisaccharide in the unrefined HEWL/(GlcNAc)₃ complex was not examined in detail because bond distances and angles deviate significantly from regularized target values.

The protein/inhibitor interactions observed in the PEWL complex illustrated in Figure 4 are also listed in Table 1, along with those from previously reported HEWL/inhibitor structures. The hydrogen bonding scheme for ring C is the same in all complexes. In PEWL/(GlcNAc)₃, only one hydrogen bond is observed between ring B and the protein namely from O6 to Asp 101. The additional hydrogen bond reported by Cheetham et al. (1992), between O3 of GlcNAc B and Asp 119, does not exist in the PEWL/(GlcNAc)₃ complex or any other HEWL structure. The availability of high-quality maps has allowed hydrogen bonds between ring A and PEWL to be described such that N2 makes a hydrogen bond to Asp 101 and O3 to Arg 125 in an adjacent symmetry-related molecule. In view of the ideal geometry of (GlcNAc)₃ in the PEWL complex, we conclude that interaction with the symmetry-related Arg 125 side chain does not distort the trisaccharide conformation; rather this interaction is likely to be flexible enough to accommodate an undistorted trisaccharide geometry. The interaction between O6 and Asn 103 observed by Cheetham et al. (1992) is also not observed in any other structure. In view of the quality of PEWL/inhibitor data and electron density maps of a full-occupancy inhibitor, as well as the indications of high model quality, we suggest that discrepancies between the present inhibitor structure and that of the refined HEWL complex may be a result of difficulties in refinement at partial occupancy, leading to a different structural interpretation.

Trp 62

Relatively low temperature factors, good geometry, and excellent electron density of the refined inhibitor complex model allow an analysis of the interactions made between the protein and

³ Observed ϕ range is -71° to -96° and ψ -range is -127° to -164° (Jeffrey, 1990). Helical twist angle ψ_H is defined by the average of ψ_1 and ψ_2 where $\psi_1 = \text{O5-C1-C4'-C3'}$ and $\psi_2 = \text{C2-C1-C4'-C5'}$ (Mo & Jensen, 1978).

Fig. 4. Stereo view showing the direct and water-mediated hydrogen bonds between protein and inhibitor in the PEWL/(GlcNAc)₃ complex as listed in Table 1. For the sake of clarity, only the portion of the residue (whether main or side chain) involved in bonding is shown. Side chains from symmetry-related molecules are distinguished by a # symbol.

sugar residues. The role of Trp 62 in the interaction with substrate and inhibitors has been the subject of much speculation in light of van der Waals or hydrophobic interactions between the aromatic side chain and the nonpolar face of the B sugar ring (Strynadka & James, 1991; Cheetham et al., 1992). Quijcho (1986) generalizes that such van der Waals contacts are characteristic of many sugar-binding proteins and are a source of substrate selectivity. Indeed, as Vyas (1991) points out, ring-stacking interactions involving aromatic residues are present in all but one protein/carbohydrate complex structure determined so far. In the structure of the PEWL/(GlcNAc)₃ complex, the closest approach made between the tryptophan ring and the apolar face of sugar B is 2.55 Å between the axial hydrogen on C5 (H5) and a point roughly in the center of the benzene ring of the Trp 62 side chain (see Fig. 5A,B). An unsophisticated (in that the delocalization of tryptophan's conjugated π -electrons was not considered and neither were so-called hydrophobic effects such as the change in entropy resulting from displacement of water molecules upon sugar binding) potential energy calculation made with use of XPLOR employing CHARMM 22 parameters suggests that the contribution of the van der Waals energy to the

total energy of the B-ring–Trp 62 interaction is, in fact, favorable by only -4 kcal/mol⁻¹. The interatomic distances between H5 and the Trp 62 ring carbons are, within error, at the van der Waals limit providing a total energy of -0.2 kcal/mol⁻¹. It seems likely that the tryptophan side chain interacts most strongly through its hydrogen bond with O6 of sugar C (O6...NE1; 2.76 Å), which contributes to the total C-ring–Trp 62 electrostatic interaction energy of -14 kcal/mol⁻¹, whereas the overall C-ring–Trp 62 interaction is estimated to be -11 kcal/mol⁻¹. As a comparison, the total energy of interaction between the Trp 62 side chain and ring A, with which no strong interaction is expected, was calculated to be -2 kcal/mol⁻¹. That the Trp 62 side chain interacts with the apolar “face” of the B sugar ring seems to be a consequence of the hydrogen bond formation with ring C. In other words, Trp 62 makes a strong hydrogen bonding contact with ring C and its flatness enables it to effectively avoid steric clashes with ring B.

Movement of the Trp 62 side chain was noted above for the PEWL/inhibitor complex. Similar large movements are reported for the Trp 62 side chain upon monosaccharide binding, where only the C ring O6 hydrogen bond can form and no sugar resi-

Fig. 5. Side view (A) and top view (B) of the interaction between Trp 62 and GlcNAc B. Carbon atoms are light blue, nitrogen atoms are dark blue, oxygen atoms are red, and hydrogen atoms are white.

due exists in the B subsite. In this case, so-called ring-stacking interactions are not a factor, yet the long-range effects of Trp 62 motion are evidenced by the same contraction experienced by loop 70–75 (Perkins et al., 1978) shown in Figure 2 for the PEWL complex. An NMR study of Trp 62 mutants carried out by Maenaka et al. (1994) indicated that 62 Leu, Ile, Val, Ala, and Gly were less active in catalyzing the cleavage of a PNP derivative of (GlcNAc)₃ and this was attributed to the importance of ring-stacking interactions of the tryptophan side chain with the apolar surface of sugar B. Interestingly, all but the Trp 62 Gly mutant showed increased bacteriolytic activity over the wild-type protein through a pH range of 4.5–8. These results are inconsistent with ring-stacking theory, but are not explained. It would be informative, as a means of comparison, to investigate the effect on the reaction of mutants that are able to act as strong hydrogen bond donors to O6 of ring C. Are aromatic residues important in so-called ring-stacking interactions or does their inherent planarity simply confer the best ability for avoiding steric clashes with incoming substrate? A more rigorous calculation of the energy associated with tryptophan/sugar interactions for lysozyme and other sugar binding proteins is in progress.

Other side chains that are directly affected by sugar binding are those of Arg 125, which moves to make a hydrogen bond between its NH₂ and the O3 atom of the A sugar. Also, the side chain of Asp 101 must rotate out of the active site in order to make a hydrogen bond with the O6 atom of sugar residue B. This movement of the Asp 101 side chain may be sufficiently powerful to alter main-chain coordinates in the vicinity and may explain the 0.5-Å RMS shift in main-chain position for Gly 102 between the native and complexed structures. The side chain of Asn 103 assumes a different rotamer conformation although there is no evidence for either direct or water-mediated contact with the inhibitor that may be responsible for this movement.

Displacement of solvent from the active site

Three water molecules present in the native structure of PEWL are displaced from the active site by the C-sugar ring, consistent with observations for HEWL (Strynadka & James, 1991; Cheetham et al., 1991). These water molecules are identified as Wat 220, located at a position corresponding to 0.43 Å from the O7 atom of GlcNAc ring C, Wat 259 at 0.96 Å from C8, and Wat 221, which hydrogen bonds to Trp 63 NE1 but is displaced to accommodate sugar ring C. Thus, two water molecules can be said to be displaced directly by sugar atoms and the third water molecule (Wat 221) is displaced in a manifestation of the hydrophobic effect. The O3 atom of ring C then forms a weak hydrogen bond to Trp 63 NE1. The comparison of solvation at site B of each structure is slightly more complicated. In the native structure, Gln 57O makes a hydrogen bond with Wat 254, whereas in the complex, that hydrogen bond is transferred to Wat 307, which is located 2.5 Å from the native Wat 254 position. Wat 307, in turn, mediates a hydrogen bonded link with O1 of sugar C. No solvent molecules identified at site A or B in the native structure can be described as having been directly displaced by incoming substrate.

Concluding remarks

The quality of the models reported herein for PEWL and the PEWL/(GlcNAc)₃ complex has permitted the resolution of dis-

crepancies observed between previous inhibitor models and has provided a confident evaluation of protein/inhibitor interactions for (GlcNAc)₃ subsites A and B. Our examination of previous results has also highlighted the importance of exercising care when refining portions of a model at partial occupancy regardless of the availability of high-resolution data. The problem of partial occupancy is anticipated to be salient with respect to time-resolved Laue data refinement.

We offer an alternative perception of “ring-stacking” interactions, terminology that suggests a significant attractive force between an aromatic amino acid side chain and the apolar face of a sugar residue. It is possible, in the case of lysozyme/saccharide interactions, that van der Waals’ contacts, hydrophobic effects, or ring-stacking interactions are not the driving force for interaction with ring B, rather the capability of Trp 62 to donate a hydrogen bond to ring C has the added feature that its planar side chain is better able to avoid steric clashes with incoming substrate at subsite B.

Materials and methods

PEWL was isolated and purified from partridge eggs purchased from Lakeside Game Farm, Lakeside, Ontario, according to a combination of two protocols described previously by Harper et al. (1987) and Archer et al. (1990). Briefly, egg whites were diluted fivefold with 0.1 M ammonium acetate and stirred for 30 min. The solution was then filtered to remove the nonsoluble fraction and equilibrated at 4 °C overnight with CM52 in the same buffer. The supernatant was removed and the resin was washed with 0.1 M ammonium acetate until the OD of the wash reached zero. The lysozyme was eluted from the CM52 with 0.5 M ammonium acetate at pH 9.4. Fractions were pooled and concentrated with an Amicon unit and dialyzed overnight in 10 mM sodium acetate at pH 2.5. The protein was further purified by fast protein liquid chromatography on a MONO-S HR 16/10 column using a gradient of 0.68–0.9 M sodium chloride in 10 mM sodium citrate, pH 2.5, over 200 min. Lysozyme-containing fractions were pooled and concentrated, dialyzed overnight against distilled deionized water, lyophilized, and stored at –20 °C until used. Fine grade chitotriose was purchased from Seikagaku Kogyo Co.

Crystallization and data collection

Crystals of PEWL were grown in 10-μL hanging drops from equal volumes of a 40-mg/mL protein solution in distilled water equilibrated with a 0.86 M NaCl (5% w/v) solution in 0.1 M Na acetate buffer at pH 4.6. Crystals of the (GlcNAc)₃ inhibitor complex were grown in 10-μL hanging drops containing equal volumes of a solution composed of 20 mg/mL protein and 15 mg/mL chitotriose (roughly a 1:20 protein:inhibitor ratio) in distilled water equilibrated with 1.1 M NaCl in 0.1 M Na acetate buffer at pH 4.6. Crystals grew within 5 days and data were measured 2 weeks after crystallization plates were set up. Native and complex crystals show distinctly different morphologies giving early indication of inhibitor binding.

Intensity data for both crystals were measured at room temperature using a Siemens/Xentronics multiwire area detector mounted on a Rigaku RU-200 rotating anode generator (37 kV, 70 mA). Intensity data were processed and scaled using the pro-

Table 2. Crystallization and data measurement statistics

	Native PEWL	PEWL/ (GlcNAc) ₃ complex
Size of crystal (mm ³)	0.4 × 0.4 × 0.6	0.4 × 0.4 × 0.8
Cell dimensions (Å)	<i>a</i> = <i>b</i> = 79.12 <i>c</i> = 37.86	<i>a</i> = <i>b</i> = 78.49 <i>c</i> = 38.07
Space group	P4 ₃ 2 ₁ 2	P4 ₃ 2 ₁ 2
Reflections	35,110	71,658
Unique reflections	8,570	9,828
Completeness at resolution 8–1.9 Å (2.0–1.9 Å)	86% (62%)	97% (83%)
<i>R</i> -merge ^a	6.2%	4.8%
<i>R</i> -iso ^b	NA	10.5%

$$^a R\text{-merge} = (\sum I - \langle I \rangle) / \sum I.$$

$$^b R\text{-iso} = (\sum |F_{\text{complex}} - F_{\text{native}}|) / (\sum |F_{\text{native}}|).$$

gram XDS (Kabsch, 1988a, 1988b). Table 2 summarizes crystal characterization and data collection statistics.

Structure solution and refinement

Native PEWL

Crystals of PEWL were determined to be isomorphous with those obtained from HEWL under similar conditions of crystallization. Thus, coordinates for HEWL (Brookhaven Protein Data Bank entry 6LYZ) were used in the isomorphous replacement solution of the PEWL structure. The initial *R*-factor for HEWL coordinates with PEWL data was 42.4% and was brought to 23.0% after one round of simulated annealing refinement using a 2,000 to 300 K slow-cooling protocol and subsequent isotropic *B*-factor refinement with use of XPLOR 3.1 (Brünger et al., 1987). The necessary amino acid substitutions were made at this stage to produce the PEWL sequence and eight further rounds of simulated annealing XPLOR refinement were alternated with rebuilding sessions using O (Jones et al., 1991), during which water molecules were included in the model and refined. A series of simulated annealing omit maps were used to aid in positioning side chains of residues Asn 18, Asn 19, Asn 37, Arg 73, Asp 101, Asn 103, Arg 112, Asn 113, Arg 125, and Arg 128 and the main- and side-chain atoms of residues 70–75. All but the side chain of Asn 77, which is located on a molecular surface lining a solvent channel some 9.5 Å wide, were built into reasonable density; coordinates for the Asn 77 side chain beyond the Cβ atom are not included in the model. Alternate side-chain conformations for Arg 21 are included in the model at 50% occupancy. The final *R*-factor for 7,729 data ($F \geq 2\sigma F$) between 8.0 and 1.9 Å resolution was 18.2%.

PEWL/(GlcNAc)₃ complex

PEWL protein coordinates from the first round of XPLOR refinement of the native enzyme were used as the starting model in refinement against data from the inhibitor complex. After one round of restrained least-squares refinement using PROFFT (Hendrickson & Konnert, 1980), the *R*-factor was reduced from 35.7% to 28.0%. The GlcNAc residue corresponding to sugar

Table 3. Refinement statistics

	Native	Complex
Reflections ($F \geq 2\sigma F$)	7,729	9,072
Resolution (Å)	8.0–1.9	8.0–1.9
<i>R</i> -factor ($F \geq 2\sigma F$) ^a	18.2%	15.7%
<i>R</i> -free ^b	26.7%	22.1%
No. of protein atoms	1,006	1,004
No. of sugar atoms		43
No. of solvent atoms	121	136
Average <i>B</i> -factor (Å ²) protein atoms only	16.5	14.6
RMS bonds (Å)	0.010	0.009
RMS angle (°)	1.85	1.80
RMS <i>B</i> -factor (Å ²) bonded atoms	1.4	1.8

$$^a R = (\sum |F_o| - |F_c|) / \sum |F_o|.$$

^b *R*-free = $(\sum |F_{os}| - |F_{cs}|) / (\sum |F_{os}|)$ where “s” refers to a randomly selected subset of data not included in refinement.

C was modeled into the resulting $|F_o| - |F_c|$ and $2|F_o| - |F_c|$ electron density maps. Addition of water molecules also commenced at this point, taking care not to include solvent sites within the protein active site. Unambiguous density for the B sugar was obvious after the next round of PROFFT refinement and the third sugar moiety (site A) was modeled after one further round. For the purpose of these rounds of refinement, ideal parameters for *N*-acetylglucosamine were incorporated into the PROFFT dictionary and the sugar sites were refined as monosaccharides with the omission of the O4 atom from the B and C rings as required to model the glycosidic linkages. Ten further rounds of simulated annealing refinement were done with XPLOR with parameters for GlcNAc with the glycosidic bond linked explicitly. Simulated annealing omit maps were used to place side chains of residue Arg 125 and the main and side chains of loops 46–48 and 70–75. In general, the structure of the complex is more ordered than that of native protein and all side chains could be positioned with confidence. The final *R*-factor for 9,072 data ($F \geq 2\sigma F$) between 8.0 and 1.9 Å resolution was 15.7%.

Refinement statistics for both structures are summarized in Table 3.

Acknowledgments

We thank K.B. Dole for carrying out the protein isolation and purification. This work was funded by a grant from the Medical Research Council of Canada to P.L.H.

References

- Archer DB, Jeenes DJ, MacKenzie DA, Brightwell G, Lambert N, Lowe G, Radford SE, Dobson CM. 1990. Hen egg white lysozyme expressed in and secreted from *Aspergillus niger* is correctly processed and folded. *Biotechnology* 8:741–745.
- Blake CCF, Johnson LN, Mair GA, North ACT, Phillips DC, Sarma VR. 1967. Crystallographic studies of the activity of hen egg-white lysozyme. *Proc R Soc B* 167:378–388.
- Brünger AT, Kuriyan J, Karplus M. 1987. Crystallographic *R* factor refinement by molecular dynamics. *Science* 235:458–460.
- Cheetham JC, Artymuk PJ, Phillips DC. 1992. Refinement of an enzyme complex with inhibitor bound at partial occupancy: Hen egg-white ly-

- sozyme and tri-*N*-acetylchitotriose at 1.75 Å resolution. *J Mol Biol* 224: 613–628.
- Chitarra V, Alzari PM, Bentley GA, Bhat TN, Eiselé JL, Houdusse A, Lescar J, Souchon H, Poljak RJ. 1993. Three-dimensional structure of a heteroclitic antigen–antibody cross-reaction complex. *Proc Natl Acad Sci USA* 90:7711–7715.
- Chu SSC, Jeffrey GA. 1967. The refinement of the crystal structures of β -D-glucose and cellobiose. *Acta Crystallogr B* 24:830–838.
- Diamond R. 1974. Real-space refinement of the structure of hen egg white lysozyme. *J Mol Biol* 82:371–391.
- Ford LO, Johnson LN, Machin PA, Phillips DC, Tjian R. 1974. Crystal structure of a lysozyme–tetrasaccharide lactone complex. *J Mol Biol* 88: 349–371.
- Harper M, Lema F, Boulet G, Poljak RJ. 1987. Antigen specificity and cross reactivity of monoclonal anti-lysozyme antibodies. *Mol Immunol* 24: 97–108.
- Hendrickson WA, Konnert JH. 1980. Incorporation of stereochemical information into crystallographic refinement. In: Diamond R, Ramaseshan S, Venkatesan K, eds. *Computing in crystallography*. Bangalore, India: Indian Academy of Sciences, International Union of Crystallography. pp 13.01–13.23.
- Howell PL. 1995. Structure determination of hexagonal turkey egg white lysozyme at 1.67 Å resolution. *Acta Crystallogr D* 51. Forthcoming.
- Howell PL, Almo SA, Parsons MR, Hajdu J, Petsko GA. 1992. Structure determination of turkey egg white lysozyme using Laue diffraction data. *Acta Crystallogr B* 48:200–207.
- Imoto T, Johnson LN, North ACT, Phillips DC, Rupley JA. 1972. Vertebrate lysozymes. In: Boyer P, ed. *The enzymes*, 3rd ed, vol VII. New York: Academic Press. pp 665–868.
- Jeffrey GA. 1990. Crystallographic studies of carbohydrates. *Acta Crystallogr B* 46:89–103.
- Jones TA, Zou JY, Cowan SW, Kjeldgaard M. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 47:110–119.
- Kabsch W. 1988a. Automatic indexing of rotation diffraction patterns. *J Appl Crystallogr* 21:67–71.
- Kabsch W. 1988b. Evaluation of single-crystal X-ray diffraction data from a position sensitive detector. *J Appl Crystallogr* 21:916–924.
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26:283–291.
- Lescar J, Souchon H, Alzari PM. 1994. Crystal structures of pheasant and guinea fowl egg-white lysozymes. *Protein Sci* 3:788–798.
- Luzzati V. 1952. Traitement statistique des erreurs dans la détermination des structures cristallines. *Acta Crystallogr* 5:802–807.
- Maenaka K, Kawai G, Watanabe K, Sunada F, Kumagai I. 1994. Functional and structural role of a tryptophan generally observed in protein–carbohydrate interaction: Trp-62 of hen egg white lysozyme. *J Biochem* 269:7070–7075.
- Mo F, Jensen LH. 1978. The crystal structure of a β -(1,4) linked disaccharide, α -*N,N'*-diacetylchitobiose monohydrate. *Acta Crystallogr B* 34: 1562–1569.
- Perkins SJ, Johnson LN, Machin PA, Phillips DC. 1978. Crystal structures of egg-white lysozyme of hen in acetate-free medium and of lysozyme complexes with *N*-acetylglucosamine and β -methyl *N*-acetylglucosaminide. *Biochem J* 173:607–616.
- Phillips DC. 1967. The hen egg-white lysozyme molecule. *Proc Natl Acad Sci USA* 57:484–495.
- Post CB, Brooks BR, Karplus M, Dobson CM, Artymiuk PJ, Cheetham JC, Phillips DC. 1986. Molecular dynamics simulations of native and substrate-bound lysozyme. *J Mol Biol* 190:455–479.
- Quioco FA. 1986. Carbohydrate-binding proteins: Tertiary structures and protein–sugar interactions. *Annu Rev Biochem* 55:287–315.
- Strynadka NCJ, James MNG. 1991. Lysozyme revisited: Crystallographic evidence for distortion of an *N*-acetylmuramic acid residue bound in site D. *J Mol Biol* 220:401–424.
- Vyas NK. 1991. Atomic features of protein–carbohydrate interactions. *Curr Opin Struct Biol* 1:732–740.
- Yao M, Tanaka I, Hikichi K, Nitta K. 1992. Crystallization and preliminary X-ray structure analysis of pigeon egg-white lysozyme. *J Biochem* 111: 1–3.