Biochem. J. (1978) **173**, 607–616
Printed in Great Britain

Crystal Structures of Egg-White Lysozyme of Hen in Acetate-Free Medium and of Lysozyme Complexes with *N*-Acetylglucosamine and β-Methyl *N*-Acetylglucosaminide

By STEPHEN J. PERKINS,*‡ LOUISE N. JOHNSON,* PELLA A. MACHIN† and DAVID C. PHILLIPS*
**Laboratory of Molecular Biophysics, Department of Zoology, South Parks Road, Oxford OX1 3PS, U.K., and*
†*Science Research Council Daresbury Laboratory, Daresbury, Warrington, Cheshire WA4 4AD, U.K.*

(*Received* 21 *November* 1977)

The binding of β-methyl *N*-acetylglucosaminide (βMeGlcNAc) to egg-white lysozyme of hen in the tetragonal crystal form was studied by X-ray-diffraction techniques to a resolution of 0.25 nm. The binding of the β-methyl glycoside is almost identical with the binding of β-*N*-acetylglucosamine (βGlcNAc). Real-space refinement of the lysozyme–α/βGlcNAc and lysozyme–βMeGlcNAc complexes allowed preliminary analysis of the conformational changes observed on binding monosaccharide inhibitors, especially in the region involving tryptophan-62 and residues 70–76. Tetragonal lysozyme crystals, grown in the absence of acetate ions, were examined by X-ray diffraction to 0.25 nm resolution. The resulting difference Fourier synthesis shows no firm evidence for bound acetate ions and indicates only minor conformational changes in the side-chain positions of aspartic acid-101 and asparagine-103. The close similarity of the lysozyme structures in the presence and absence of acetate is contrary to expectations from previous n.m.r. studies.

X-ray-crystallographic studies on egg-white lysozyme (EC 3.2.1.17) of hen in the tetragonal crystal form (Blake *et al.*, 1965, 1967*a*), together with studies on the binding of monosaccharide and trisaccharide inhibitors and model-building studies, have led to the identification of six sugar-binding sites (A–F) in the active-site cleft of the enzyme (Blake *et al.*, 1967*b*; reviewed in Imoto *et al.*, 1972). The X-ray studies on the binding of acetamido monosaccharides show that these compounds bind in the region of subsite C and make specific hydrogen bonds between the NH and CO groups of the *N*-acetyl group and the protein main-chain CO and NH groups of residues Ala-107 and Asn-59 respectively (Blake *et al.*, 1967*b*). Lysozyme has been used as a model enzyme to compare the information obtained from n.m.r. studies on the enzyme in solution with that obtained by X-ray studies on the enzymes in the crystal (Perkins, 1977). N.m.r. experiments on the binding of monosaccharides αGlcNAc, βGlcNAc, αMeGlcNAc and βMeGlcNAc to lysozyme suggest that the sugars bind to more than one of the six subsites on the enzyme (Perkins, 1977). The apparent chemical shifts of the inhibitor proton resonances when the inhibitor is fully bound to the enzyme indicate that, although the binding modes of the four sugars are similar, the binding of αGlcNAc is different from the binding of the two methyl glycosides. The n.m.r. data for βGlcNAc are less precise, because these were obtained from studies of a mixture of the α- and β-anomers. Nevertheless the data indicate that the binding of βGlcNAc is more similar to the binding of the glycosides than that of αGlcNAc.

The crystal studies on the binding of the mutarotation mixture of α- and β-GlcNAc (α/βGlcNAc) to lysozyme had previously shown that the sugars bind in two slightly different orientations in subsite C, one of which corresponds to the binding of the β-anomer (termed subsite C-β; Fig. 1*a*) and the other to the binding of the α-anomer (termed subsite C-α; Fig. 1*b*). In subsite C-α the sugar is rotated (with respect to subsite C-β) about the axis of hydrogen bonds formed between the acetamido group of the sugar and main-chain atoms of the protein, so that it is possible to form a hydrogen bond between the α-hydroxy group and the main-chain NH group of residue 109 (Blake *et al.*, 1967*b*; Imoto *et al.*, 1972). Low-resolution experiments with α-methyl *N*-acetyl-6-iodoglucosaminide showed that the α-methyl glycoside bound in subsite C-β and was excluded

Abbreviations used: αMeGlcNAc, α-methyl *N*-acetylglucosaminide; βMeGlcNAc, β-methyl *N*-acetylglucosaminide.

‡ Present address: Institut für Molekularbiologie und Biophysik, Eidgenossiche Technische Hochschule, Hönggerberg, CH-8093 Zürich, Switzerland.

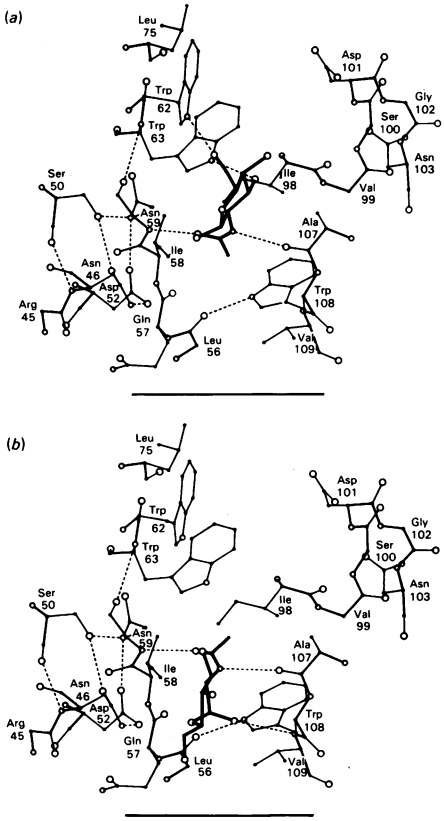

Fig. 1. *Binding of βGlcNAc to subsite C-β (a) and of αGlcNAc to subsite C-α (b) at the active site of hen egg-white lysozyme*

The sugar is shown with heavy lines and the surrounding protein with light lines. Hydrogen bonds are dotted (from Imoto *et al.*, 1972). The view is down [110] with positive *z* pointing towards the top of the page. The bar is equivalent to 1 nm.

from subsite C-α, presumably because of interference between the enzyme and the additional methyl group (Blake *et al.*, 1967*b*). To provide a more direct comparison with the n.m.r. studies, we have carried out a high-resolution study of the binding of βMeGlcNAc to lysozyme in the absence of acetate ions. N.m.r. evidence suggests that the binding modes of the methyl glycosides are similar, but βMeGlcNAc binds more strongly than αMeGlcNAc (Perkins, 1977).

Tetragonal lysozyme crystals are usually grown at pH4.7 in the presence of 40mM-sodium acetate. There have been reports that acetate ions affect the apparent dissociation constants for the binding of GlcNAc (Sykes, 1969) and the binding of Gd(III) (Jones *et al.*, 1974) to lysozyme. Accordingly n.m.r. experiments have been carried out in acetate-free medium. In the present study a comparison is made between native lysozyme crystals in the presence of 40mM-sodium acetate and lysozyme crystals grown in the absence of acetate.

Preliminary work on the refinement of the native structure of egg-white lysozyme of hen in this laboratory has been completed (D. E. P. Grace & D. C. Phillips, unpublished work) and should eventually provide a definitive set of co-ordinates from which the conformational changes of the enzyme on binding of inhibitors may be assessed. We have attempted to obtain a preliminary measure of these changes by using the refinement method of Diamond (1974) with electron-density maps calculated with terms $m(2F_{PI} - F_P) \exp i\alpha_{Iso}$ where F_{PI} and F_P are the structure amplitudes of the protein–inhibitor complex and protein respectively and m and α_{Iso} are the figure of merit and the phase determined by the heavy-atom isomorphous-replacement method. The starting co-ordinates used were the RS5D set of co-ordinates obtained by Diamond (1974) by using the isomorphous-replacement phases. Such refinement of the lysozyme–inhibitor complexes is necessarily preliminary, since it utilizes experimental rather than calculated phases. Nevertheless by using the same constraints as those applied by Diamond (1974) we have attempted to obtain a measure of the shifts in position relative to the RS5D set of atomic positions. These refined co-ordinates for the lysozyme–inhibitor complexes have been used to predict ring-current shifts experienced by the protein and inhibitor molecule. The calculated shifts have been compared with the chemical shifts when the inhibitor is fully bound to the enzyme observed in the n.m.r. experiments, with encouraging results (Perkins, 1977; Perkins *et al.*, 1977).

Experimental

Crystallization

Tetragonal crystals of egg-white lysozyme of hen, space group $P4_32_12$, were grown at pH5.1 from solutions of 40–50mg of lysozyme (previously freed from acetate; Perkins, 1977)/ml, in 4 or 5% NaCl solutions at 20°C. No buffers, acetate or otherwise, were used. The solutions before crystal growth were equivalent to 2.8–3.4mM-lysozyme in NaCl concentrations of 0.68–0.86M. The absence of acetate was confirmed by the use of ^{1}H n.m.r. spectra at 270MHz on the same batches of dialysed lysozyme as used in the crystallization procedure (Perkins, 1977).

Data collection

Four sets of crystallographic data were used, two of which had been obtained with crystals grown from solutions containing 40mM-sodium acetate (A and B), and two sets collected for this study with acetate-free crystals obtained as above (C and D). (A),

Native lysozyme, pH4.7, 40mM-acetate buffer, 5% NaCl; these data were used in the calculation of the 0.2nm-resolution electron-density maps from which the interpretation of the structure was made (Blake *et al.*, 1965, 1967*a*); (B), as in (A), but soaked in 500mM-α/βGlcNAc for 24h, after which data were collected at 0.2nm resolution (Blake *et al.*, 1967*b*); (C), native lysozyme, pH5.1, acetate-free, 5% NaCl; (D), as for (C), but soaked in 500mM-βMeGlcNAc for 24h in a quartz capillary tube: the crystal remained clear with minor cracking. Before the final mounting in (C) and (D), the pH of the crystal mother liquor was checked with both a pH-meter and narrow-range pH paper to ensure that it was as near as possible to pH5.1. The crystals in samples (C) and (D) were chosen to be near to 0.8mm × 0.8mm × 0.8mm in size (0.5mm^3).

The X-ray-diffraction data to 0.25nm resolution were collected from a single crystal for each of samples (C) and (D) on a Hilger and Watts linear diffractometer modified to measure five reflections simultaneously (Phillips, 1964; Arndt *et al.*, 1964), by using a Philips fine-focus X-ray tube with the generator run at 40kV and 16mA. The source-to-crystal distance was 150mm. The crystal was mounted with the crystallographic *a*-axis parallel to the rotation axis. The crystal-to-counter-arm distance was adjusted to 365mm to measure five reflections simultaneously with the separation of the counters 7.5mm. The crystal-to-counter arm was filled with helium. The oscillation range for measurements was 1°30′.

The lattice dimensions for samples (C) and (D) were found to be $a = b = 7.92$nm, $c = 3.80$nm, and $a = b = 7.84$nm, $c = 3.83$nm, respectively. Values of $a = b = 7.88$nm and $c = 3.81$nm were used in the subsequent data processing.

The relative absorption was measured by recording the variation in intensity of the strong axial (800) reflection as a function of ϕ (North *et al.*, 1968). The absorption varied from 1.000 to 1.445 for (C) and from 1.000 to 1.240 for (D).

Radiation damage was monitored, first by measuring eight standard reflections evenly distributed throughout reciprocal space at the start and end of each set of five reciprocal-lattice levels, and secondly by scanning along the b^* and c^* reciprocal-lattice axes before and after the data collection (Snape, 1974). From the first method, the overall damage was 6% for the crystal (C) over about 80h of non-stop data collection, and varied in a range up to 16% for the derivative crystal (D) in about 135h of data collection (to which has been added 45h as a break midway during collection). The dependence of the radiation damage on $\sin\theta$ (where θ is the Bragg angle) was found to vary from 0 to 18% from 0.8nm resolution to 0.25nm resolution on both b^* and c^* axes for (C) and likewise from 0 to 7% for (D).

Data between 0.8 and 0.25nm resolution were collected automatically in the flat-cone setting on the linear diffractometer by measuring five reflections quasi-simultaneously for a quarter sphere of reciprocal space. Each reflection was thus measured up to four times. The low-resolution data were also collected on the linear diffractometer, by using the instrument in the three-counter mode and adjusting the settings by hand when the quasi-simultaneity of the three reflections became poor.

Data processing

The diffractometer data for samples (C) and (D) were corrected for background, Lorentz-polarization, absorption factors (North, 1965; North *et al.*, 1968) and a linear (with time) radiation damage. The scale factors for each of the *h* reciprocal-lattice levels were calculated from the equivalent reflections in different levels of reciprocal space by the method of Hamilton *et al.* (1965). All the data were scaled to the basic native set (A). The fractional change in structure-factor amplitude was 0.100 for sample (C) and 0.161 for sample (D), and the mean isomorphous difference was 40 for sample (C) and 63 for sample (D). Difference Fourier syntheses were computed by using the terms $m(F_{PI} - F_P)\exp i\alpha_{Iso}$ where α_{Iso} is the phase of native protein as determined by multiple isomorphous-replacement method.

Real-space refinement

The atomic positions for the lysozyme–α/βGlcNAc and lysozyme–βMeGlcNAc complexes were refined by the real-space refinement method of Diamond (1974) by using Fourier syntheses based on the expression $m(2F_{PI} - F_P)\exp i\alpha_{Iso}$ for the acentric terms and $mF_{PI}\exp i\alpha_{Iso}$ for the centric reflections (Luzzati, 1953).

The quantity minimized by the refinement program is $(\rho_o - \rho_m)^2$, where ρ_o is the observed electron density and ρ_m is the electron density calculated from the model:

$$\rho_m(r) = k \sum_i Z_i a_i^{-3} \exp[-\pi(r - r_i)^2/a_i^2] + d$$

in which r is a general position vector, r_i is the position vector of the ith atom, Z_i its electron count and a_i its apparent radius. k is an overall scale factor and d is a background electron density. The values of the parameters used in the present study are given in Table 1 and are the same as those used by Diamond (1974). The parameters refined were Z_i and r_i. The starting co-ordinates for the protein were the set RS5D obtained by Diamond (1974) in his real-space refinement of the native lysozyme molecule by using a Fourier synthesis based on isomorphous phases. For acetate-free lysozyme one cycle of refinement was carried out. For lysozyme–βMeGlcNAc one cycle of refinement of both the protein and the sugar co-ordinates was followed by a refinement of the sugar

Table 1. *Values used in the real-space refinement program (Diamond, 1974)*

Zone length	5
Margin width	8
Fixed radius of all atoms	1.5
Relative weights of C:N:O:S	6:7:8:16
Parametric constants for the dihedral angles	
ϕ, ψ (main chain)	3.7
χ_1, χ_4 (side chain)	3.2
τ (N_i, C_i^{α}, C_i)	0.33
ω (main-chain peptide)	0.50
χ_5 (arginine)	0.25
ϕ_1–ϕ_3 (proline)	0.10
τ (cystine)	0.33
Filter ratio for translational, rotational and height refinement parameters	0.01

co-ordinates alone, allowing rotations about the C-5–C-6 and C-2–N-2 bonds to determine the positions of C-6 and the acetamido group. For lysozyme–α/βGlcNAc one cycle of refinement of the protein atoms alone was followed by a cycle for both the protein and sugar atoms.

The sugar co-ordinates for the first cycle of refinement were derived from the co-ordinates of the crystal structure of αGlcNAc (Johnson, 1966) rotated and translated to fit the unregularized co-ordinates of the hand-built model of the lysozyme–α/βGlcNAc complex (Imoto *et al.*, 1972) by a least-squares procedure.

Results

Interpretation of difference Fourier syntheses

The difference Fourier synthesis for α/βGlcNAc has been interpreted in terms of the binding of βGlcNAc to subsite C-β and of αGlcNAc to subsite C-α as described previously (Blake *et al.*, 1967*b*; Imoto *et al.*, 1972). The difference Fourier synthesis for βMeGlcNAc shows that the sugar has bound only to subsite C-β (Fig. 2). Comparison of the corresponding sections of the difference Fourier syntheses for α/βGlcNAc and βMeGlcNAc shows that the electron densities in the region of subsite C-β are almost identical. The small conformational changes in the protein that were previously noted on the binding of α/βGlcNAc are also observed for βMeGlcNAc. In particular the electron-density features that indicate movement of the main-chain loop of residues 70–74 and those that indicate movement of Trp-62 are present in both difference Fourier syntheses. They are illustrated for the lysozyme–α/βGlcNAc complex in Fig. 3.

The difference Fourier synthesis for acetate-free lysozyme shows no definite feature that can be assigned to an acetate ion. As a preliminary check

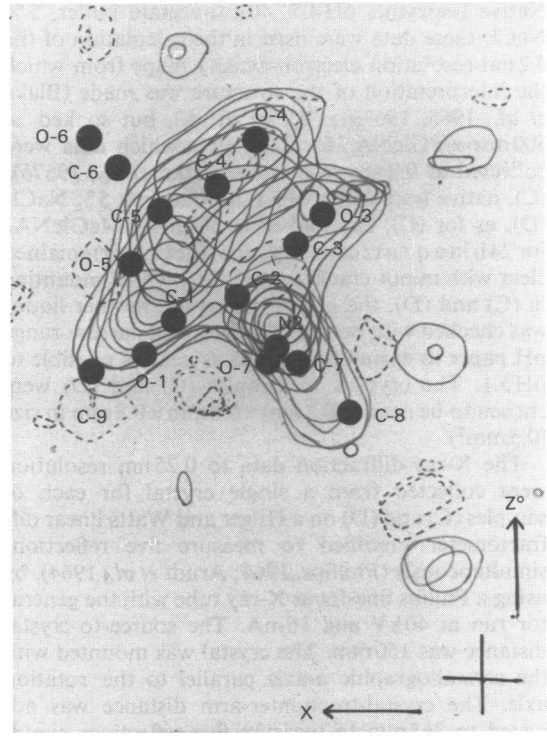

Fig. 2. *Sections $y = 34/120$ to $39/120$ of the difference Fourier synthesis for the lysozyme–βMeGlcNAc complex showing the electron density in subsite C-β and the atomic positions obtained for the sugar by real space refinement* Contour interval is ± 250 electrons/nm^3 with lowest contour at 500 electrons/nm^3. The C-6 and O-6 atoms lie outside the lowest contour, but are still within positive density. The cross in the bottom-right corner defines the position $x = 0$, $z = \frac{1}{2}$.

that this lack of density did not arise from errors in the phases that had been experimentally determined by the heavy-atom isomorphous-replacement method, a difference Fourier synthesis was computed on the basis of a set of phases (termed SF7) obtained during the course of the refinement of the lysozyme structure (D. E. P. Grace & D. C. Phillips, unpublished work). At this stage in the refinement some 40 water molecules had been placed in addition to the protein atoms and the conventional R factor (which represents the fractional agreement between observed and calculated structure factors) was 0.26. The difference Fourier synthesis based on calculated phases also shows no significant features that can be assigned to acetate ions. Both difference Fourier syntheses have small positive and negative features near the positions of side chains of Asp-101 and Asn-103, which may reflect some disturbance at these residues. In addition there is a small negative peak approx. 0.3 nm from

the side chain of Asn-113, which may represent displacement of a bound acetate ion by a water molecule in acetate-free lysozyme. The depth of this negative peak is about half that expected for an acetate ion. There are no features in the acetamido specificity pocket of site C nor are there any near the catalytic residues Asp-52 and Glu-35.

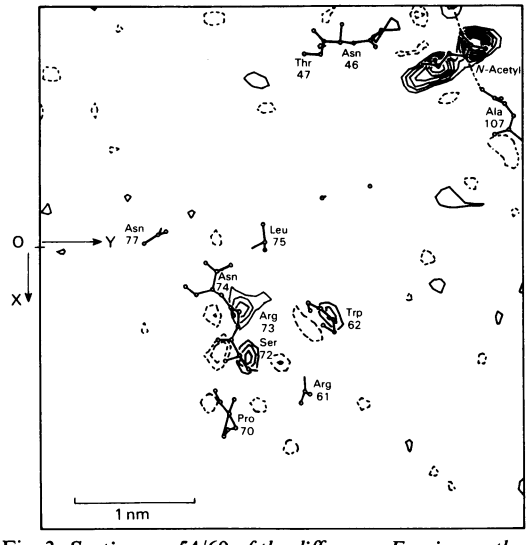

Fig. 3. *Section z = 54/60 of the difference Fourier synthesis for the lysozyme–α/βGlcNAc complex*
The atomic positions shown correspond to those for sections z = 53/60–55/60 obtained after the second cycle of real-space refinement. The movements of the main-chain region residues 72–74 and the indole ring of Trp-62 are apparent. The density for part of αGlcNAc is shown in the top right corner of the Figure and corresponds to the acetamido specificity region in a symmetry-related molecule.

Real-space refinement

The application of real-space refinement methods to the lysozyme–inhibitor complexes is preliminary. We have attempted to find an optimum fit for the protein and inhibitor molecules to the electron-density maps within the constraints of stereochemistry. The results have not been analysed by subsequent double difference Fourier syntheses or by comparison of observed and calculated structure amplitudes. It was judged that this detailed analysis could best be carried out when the refinement of the native protein molecule and its water structure in the tetragonal crystal form was complete. However, it is apparent that the Diamond (1974) method of refinement is most effective in producing a good fit of the shifted co-ordinates to the electron-density map. The movement of Trp-62, together with the difference Fourier synthesis, is illustrated in Fig. 3 for the lysozyme–α/βGlcNAc complex. The indole ring of the native structure lies between regions of positive and negative electron density in the RS5D co-ordinates. It is shifted by the real-space refinement program towards positive density by an amount corresponding to 0.04–0.05 nm. The program, as run without detailed manual intervention, was not so effective when the electron density of side chains was ill-defined. Large shifts occurred for several arginine residues on the surface of the molecule, for example, which did not appear to be explicable in terms of the

Table 2. *Co-ordinates of GlcNAc and βMeGlcNAc after real-space refinement*
Co-ordinates are expressed in a right-handed system of axes in nm. These supersede the data given in Imoto *et al.* (1972), which are for the hand-built model. Positions of O-6, C-7, O-7 and C-8 for α/βGlcNAc are less precisely defined than for βMeGlcNAc (see the text).

	αGlcNAc			βGlcNAc			βMeGlcNAc		
	x	*y*	*z*	*x*	*y*	*z*	*x*	*y*	*z*
C-1	0.553	2.429	2.218	0.616	2.484	2.591	0.612	2.480	2.599
C-2	0.569	2.390	2.366	0.486	2.414	2.629	0.477	2.424	2.644
C-3	0.655	2.492	2.440	0.409	2.496	2.734	0.414	2.515	2.751
C-4	0.793	2.499	2.372	0.498	2.513	2.857	0.511	2.523	2.870
C-5	0.769	2.551	2.246	0.622	2.595	2.813	0.640	2.590	2.819
O-5	0.686	2.449	2.163	0.690	2.512	2.714	0.694	2.500	2.718
O-4	0.871	2.603	2.440	0.426	2.598	2.953	0.453	2.615	2.968
C-6	0.895	2.558	2.144	0.724	2.616	2.923	0.749	2.601	2.924
O-6	0.963	2.436	2.142	0.763	2.496	2.981	0.778	2.477	2.981
N-2	0.435	2.387	2.429	0.401	2.402	2.509	0.386	2.421	2.528
C-7	0.404	2.283	2.514	0.334	2.283	2.485	0.326	2.301	2.494
O-7	0.488	2.194	2.522	0.361	2.192	2.565	0.390	2.196	2.475
C-8	0.273	2.304	2.587	0.236	2.292	2.370	0.220	2.317	2.387
O-3	0.680	2.442	2.574	0.295	2.418	2.777	0.295	2.450	2.801
O-1	0.485	2.550	2.213				0.675	2.388	2.515
C-9							0.800	2.435	2.457

electron-density map. These shifts did not affect residues at the active site.

(*a*) *Sugar binding*. The co-ordinates obtained by real-space refinement for the sugars α/βGlcNAc and βMeGlcNAc are given in Table 2. As indicated above, the positions for βGlcNAc and βMeGlcNAc are similar and the root-mean-square difference of the co-ordinates is 0.029nm (with the exceptions of the N-2, C-7, O-7, C-8 and O-6 atoms). The overlapping electron-density features of αGlcNAc and βGlcNAc did not appear to complicate the course of the real-space refinement when the atoms in each of the subsites were allocated half the usual atomic number (Table 1). In the refinement of the protein and sugar co-ordinates, the two anomers were allowed to move independently of each other and the protein, and there was no tendency for poorly defined atoms to drift into neighbouring areas of higher electron density. The similarity of the co-ordinates for βGlcNAc and βMeGlcNAc indicates that the position of the β-anomer of GlcNAc is not affected by competition of the α-anomer. For βMeGlcNAc rotations about C-2–N-2 and C-5–C-6 bonds were allowed. The angle of rotation about the C-2–N-2 bond differed by +13° from that observed in αGlcNAc (Johnson, 1966). There was no difference in the angle of rotation about the C-5–C-6 bond, presumably because there is no good density for O-6. The rotation of the acetamido groups was not attempted for α/βGlcNAc because these atoms of the anomers occupy overlapping positions.

The contacts within 0.4nm (atom-to-atom distance) between βMeGlcNAc and lysozyme are summarized in Table 3. The same hydrogen bonds are made between the sugar and the enzyme as previously described for subsite C-β (Blake *et al.*, 1967*b*), namely hydrogen bonds between the acetamido N-2 and O-7 atoms and the lysozyme main-chain CO group of Ala-107 and NH of Asn-59 respectively, a hydrogen bond between O-3 of the sugar and the $N^{\varepsilon 1}$ of Trp-63, and a hydrogen bond between O-6 and the $N^{\varepsilon 1}$ of Trp-62. In addition there are some 32 van der Waals contacts within 0.4nm and some 107 contacts between 0.4 and 0.5nm.

(*b*) *Conformational changes in the protein*. The conformational changes in the protein for acetate-free lysozyme, the lysozyme–α/βGlcNAc and the acetate-free-lysozyme–βMeGlcNAc complexes are summarized in Figs. 4, 5 and 6 respectively by means of plots of movement of α-carbon position against residue number. The positions of the α-carbon atoms are likely to be influenced by both side-chain and main-chain movements and hence form an approximate representation of the conformation changes of the protein molecule in the presence of the inhibitor. The main polypeptide backbone is relatively well ordered and hence the use of shifts due to α-carbon atom as an approximate measure of conformational changes is unlikely to be complicated by side-chain disorder. The three plots suggest that there is a general shift in α-carbon atom positions of about 0.02nm, whereas certain residues exhibit very much greater shifts in their α-carbon atoms. As already noted, the difference Fourier syntheses for acetate-free lysozyme gave no indication of conformational changes apart from the region of Asp-101 and Asp-103. It is interesting that a shift is seen in the α-carbon position of Gly-49 where no differences were observed in the difference Fourier synthesis. This is the region where the greatest difficulty was encountered in the

Table 3. *Contacts to lysozyme atoms within 0.4nm of the βMeGlcNAc atoms after real-space refinement of the lysozyme–βMeGlcNAc complex*

There are 107 contacts between 0.4 and 0.5nm from the 16 inhibitor atoms of βMeGlcNAc. Asterisked values in the Table indicate the hydrogen-bonds between βMeGlcNAc and lysozyme. The numbering of the protein atoms follows the convention of the IUPAC–IUB Commission on Biochemical Nomenclature.

Inhibitor atom	Protein atom		Distance (nm)
C-1	Ala-107	O	0.38
C-2	Ala-107	O	0.34
C-3	Ala-107	O	0.34
C-4	Trp-62	$C^{\delta 1}$	0.40
		$N^{\varepsilon 1}$	0.40
C-6	Trp-62	$N^{\varepsilon 1}$	0.38
O-6	Trp-62	$C^{\delta 1}$	0.34
	Trp-62	$N^{\varepsilon 1}$	0.26*
	Trp-62	$C^{\varepsilon 2}$	0.36
N-2	Ala-107	O	0.25*
	Ala-107	C	0.36
C-7	Ala-107	O	0.34
O-7	Ile-58	C^{α}	0.37
	Ile-58	C	0.40
	Asn-59	N	0.33*
	Asn-59	C^{β}	0.36
	Trp-63	$C^{\delta 1}$	0.32
	Trp-63	$N^{\varepsilon 1}$	0.34
C-8	Ala-107	O	0.34
	Trp-108	C^{γ}	0.35
	Trp-108	$C^{\delta 1}$	0.31
	Trp-108	$N^{\varepsilon 1}$	0.30
	Trp-108	$C^{\varepsilon 2}$	0.33
	Trp-108	$C^{\zeta 2}$	0.40
	Trp-108	$C^{\delta 2}$	0.36
O-3	Trp-63	$N^{\varepsilon 1}$	0.33*
	Ala-107	O	0.39
	Ala-107	C^{β}	0.40
O-1	Asp-52	$O^{\delta 1}$	0.39
	Asn-59	C^{β}	0.35
	Asn-59	C^{γ}	0.39
	Asn-59	$N^{\delta 2}$	0.40
C-9	Asn-46	$N^{\delta 2}$	0.40
	Asp-52	$O^{\delta 2}$	0.37
	Asp-52	$O^{\delta 1}$	0.36
	Asn-59	$N^{\delta 2}$	0.40

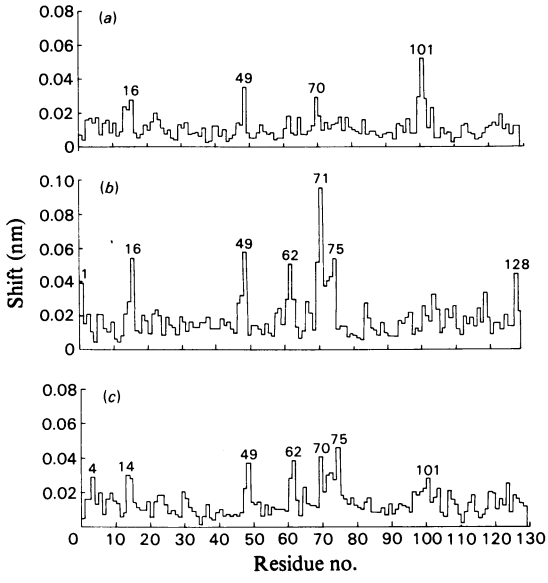

Fig. 4. *Histograms to show the movement in nm of the α-carbon positions in the lysozyme complexes compared with native lysozyme after one cycle of real-space refinement*
(*a*) Acetate-free lysozyme; (*b*) lysozyme–α/βGlcNAc; (*c*) lysozyme–βMeGlcNAc.

original interpretation of the electron-density map (Blake *et al.*, 1967*a*), and the shift in this residue as well as those for Gly-16 and Pro-70, which are not correlated with features in the difference map, may well be spurious. The shifts obtained in the real-space refinement of acetate-free lysozyme may be taken as an estimate of apparent shifts in atoms that reflect the error in measurement of intensities and differences in resolution of the data for acetate-free lysozyme (0.25 nm) compared with the resolution of the data for native enzyme (0.2 nm). For lysozyme–α/βGlcNAc (Fig. 5) and lysozyme–βMeGlcNAc (Fig. 6) significant shifts in the position of α-carbon atoms of Gly-49, Trp-62, Pro-70, Gly-71 and Leu-75 are seen. These shifts are all correlated with features in the difference Fourier synthesis and correspond to movement at the β-bend (Gly-49) connecting the antiparallel chains of the β-sheet that links part of the left-hand side of the cleft to movements of the upper part of the β-pleated sheet (Trp-62), and of the loop of main chain between residues 70–75 above the β-pleated sheet. For the lysozyme–α/βGlcNAc complex there are large movements also for Lys-1, Gly-16 and Arg-128.

The changes in position of several biologically important loci on the protein on binding the GlcNAc inhibitors are summarized in Table 4. Their significance may be assessed by comparison with values obtained for acetate-free lysozyme. The indole rings of Trp-62, Trp-63 and Trp-108 flank the upper half of the active-site cleft in the region of subsites C-β anc B. Of these the benzene ring of the indole of Trp-6d moves the greatest distance on binding the GlcNA2 inhibitors, but the angular movement of the mean-square planes is greatest for Trp-63 (Table 4). The movement of Trp-63 and Trp-108 are at the level of precision of the real-space refinement, but comparison with the difference Fourier synthesis shows that they are significant. The movements of the four disulphide bridges are also given in Table 4. Cys-6–Cys-127 moves by up to 0.04 nm in the lysozyme–βMeGlcNAc complex, but the movements of the other disulphide bridges are small. Table 4 shows that the separation of the two catalytically important carboxy groups of Asp-52 and Glu-35 remains constant at 0.7 nm to within 0.01 nm for the three lysozyme refinements, and that the separation of the important atoms forming the acetamido specificity site of subsite C-β (NH group of Asn-59 to CO group of Ala-107) does not alter significantly.

The results of the real-space refinement are not sufficiently precise at this stage to allow definitive statements on the role of strain on the binding of inhibitors to lysozyme. The dihedral angles of the lysozyme–βMeGlcNAc complex were calculated after real-space refinement. It was found that the dihedral angle (ω) of the peptide bond differed from 180° by more than 15° for 47 residues. Of these, ten residues (Gly-4, Gly-16, Ala-31, Asp-48, Gly-71, Arg-73, Cys-76, Ser-85, Cys-115, Arg-125) had ω values differing from 180° by more than 25°. Although there is evidence that the peptide bond need not be strictly planar (Winkler & Dunitz, 1971), such large deviations from planarity seem improbable. It appears that the constraints used in the refinement (Table 1) were not sufficiently severe for the lysozyme–inhibitor complexes at 0.25 nm resolution, although these constraints were adequate for the refinement of native lysozyme at 0.2 nm resolution (Diamond, 1974). For example, Diamond (1974) noted that relaxing the filter level for eigen shifts with eigen values greater than 10^{-2} times the largest eigen value present (as used to obtain the RS5D co-ordinates and as used in the present study) to 10^{-4} times the largest eigen value present resulted in some 30 main-chain nitrogen atoms moving by more than 0.025 nm, and in each case these movements were associated with ω values substantially different from 180°. Of these Diamond (1974) concluded that only seven residues (12, 33, 36, 41, 53, 76, 123) showed any evidence for a twisted peptide after examination of electron-density maps.

In the present study the co-ordinates obtained from the real-space refinement were processed by using the MODELFIT regularization program (Dodson *et al.*, 1976). The program adjusts the co-ordinates

Table 4. *Summary of movements of protein conformation*

Distances are given in nm. For α/βGlcNAc–lysozyme the first cycle of refinement was for protein atoms alone, and the second cycle included protein and inhibitor atoms. For the other complexes one cycle with both protein and inhibitor atoms was used (see the text).

	RS5D	Lysozyme (acetate-free)	α/βGlcNAc–lysozyme (after two cycles)	βMeGlcNAc–lysozyme
Movement of Trp indole $C^{\zeta 3}$ and $C^{\eta 2}$ atoms from the RS5D co-ordinates				
Trp-62	—	0.02, 0.02	0.04, 0.05	0.04, 0.03
Trp-63	—	0.01, 0.01	0.02, 0.03	0.01, 0.01
Trp-108	—	0.02, 0.01	0.02, 0.02	0.02, 0.03
Angle between mean-square planes of the Trp rings in the RS5D co-ordinates and the derivatives				
Trp-28	—	0°	1°	1°
Trp-62	—	4°	4°	6°
Trp-63	—	3°	7°	7°
Trp-108	—	3°	2°	3°
Trp-111	—	2°	2°	1°
Trp-123	—	2°	2°	1°
Movement of C^{δ} atom of Glu-35	—	0.02	0.02	0.02
Movement of C^{γ} atom of Asp-52	—	0.01	0.03	0.02
Distance between C^{δ} atom (Glu-35) and C^{γ} atom (Asp-52)	0.70	0.70	0.71	0.69
Acetamido specificity site (NH of Asn-59 to CO of Ala-107)	0.77	0.76	0.77	0.75
Movement of disulphide bridges:				
Cys-6–Cys-127	—	0.02, 0.02	0.03, 0.02	0.04, 0.04
Cys-30–Cys-115	—	0.01, 0.01	0.01, 0.01	0.00, 0.01
Cys 64–Cys-80	—	0.02, 0.01	0.01, 0.02	0.03, 0.02
Cys-76–Cys-94	—	0.00, 0.01	0.01, 0.02	0.02, 0.02
Mean	—	0.012	0.015	0.021
Root-mean-square movement of all atoms	—	0.028	0.035 (1st cycle) 0.026 (2nd cycle)	0.036

by a least-squares method so as to obtain a properly weighted fit to both positions from the refinement and the molecular parameters assumed from known stereochemistry. The standard deviations assigned to the x, y, z co-ordinates were 0.02 nm. The root-mean-square shift in co-ordinates was 0.0097 nm for residues 1–65 and 0.0125 nm for residues 64–129. These shifts are small compared with the overall root-mean-square shift for all atoms in the real-space refinement (0.036 nm). The distances given in Table 4 were not altered. MODELFIT necessarily regularized the stereochemistry of the polypeptide chain, and after refinement only three residues had ω values that differed from 180° by more than 15° (Phe-3, $\omega = 164°$; Trp-63, $\omega = 163°$; Pro-70, $\omega = 165°$). Comparison of the MODELFIT co-ordinates with those obtained by real-space refinement and the lysozyme–βMeGlcNAc electron-density map in the region of the 'strained' peptides showed that the fit for the MODELFIT co-ordinates was almost as good as that obtained by real-space refinement, with one striking exception. Evidently the real-space refinement program had not been sufficiently guided and had made shifts in atoms that resulted in a small improvement of the fit to the electron density at the expense of the dihedral angles. The one exception was the peptide bond between Pro-70 and Gly-71. The nitrogen atom of Gly-71 ($\omega = -166°$ in the RS5D co-ordinates) lies outside the positive electron density for the lysozyme–βMeGlcNAc complex, indicating that a shift is required, as noted above. Consideration of the difference electron density and the stereochemistry of the molecule in this region suggests that a *cis*-peptide for Pro-70 is unlikely. The real-space refinement places this atom well in the electron density after a shift of 0.06 nm, but ω now has the unrealistic value of −137°. MODELFIT restores ω to a reasonable value (−177°) by a shift of 0.046 nm, but the N atom is now out of the electron density. Evidently the precise conformation of the protein in this region (residues 70–71) is not yet clear in the lysozyme–monosaccharide complex. The movement of the region residues 70–76 may be described as a sideways movement towards the cleft. This movement is difficult to achieve by the real-space refinement program by using solely rotational angles for a

flexible zone of five residues at any one time without distortion. Further work is required in both the native and complex structures on the conformation of this part of the chain. This region was difficult to fit in the original electron-density map of the native structure (see Levitt, 1974) and has been found to have a somewhat different conformation in the triclinic crystal structure (see below).

(*c*) *Effect of lattice contacts on conformational changes.* A correlation of the protein-conformational changes with the solvent exposure of the residues in the crystal lattice has been made. Shrake & Rupley (1973) have described the relative solvent exposures of the amino acids in lysozyme and the extent to which these values are decreased by lattice contacts in the tetragonal crystal form. The co-ordinates used in their calculations were those obtained by the regularization of the hand-built model of lysozyme by Diamond's model-building program (Diamond, 1966; Imoto *et al.*, 1972). Many of the large movements observed in the lysozyme–inhibitor complexes are correlated with side chains that have a large solvent exposure both in the free form of lysozyme and in the crystal lattice, such as, residues 1, 73 and 74. Changes are also observed in residues that have a very small solvent exposure in the free form of lysozyme, such as residues 71 and 72. Since these residues are partially shielded in the protein, their shifts are likely to be of significance in the inhibitor-binding process. Residues 14 and 62 are examples of side chains that are partially shielded by intermolecular contacts in the lattice and yet are still able to move on formation of the enzyme–inhibitor complex. It appears therefore that lattice contacts do not necessarily impede small shifts of atoms.

Discussion

The crystallographic studies show that there is no significant change in conformation between lysozyme in the presence of 40mM-acetate buffer and lysozyme in acetate-free medium. Small conformational changes for the side chains of Asp-101 and Asn-103 are indicated in the difference Fourier synthesis and in the refinement, and there is some evidence for displacement of an acetate ion from Asn-113. N.m.r. studies have shown that the presence of acetate ions increases the apparent dissociation constants of lysozyme and GlcNAc (Sykes, 1969) and Gd(III) (Jones *et al.*, 1974), which suggests that acetate ions might bind in the region of subsites C and D and possibly in the acetamido specificity pocket. Shindo & Cohen (1976) observed that an n.m.r. signal derived from acetic acid shifted downfield on binding Co^{2+} and also concluded that the acetate anion is bound adjacent to the active-site region. No binding in this region was observed in the present crystal studies. Moreover n.m.r. studies in our laboratory have suggested that GlcNAc binding to lysozyme is little affected by acetate (Perkins, 1977). It is possible that the failure to detect binding in the crystal could be explained by the presence of 0.68–0.86M-NaCl in the crystal, which may prevent any non-specific ionic interaction between the protein and acetate ions. However, such an explanation seems unlikely, since acetate appears to bind firmly to lysozyme and can only be removed after dialysis for 2 days a tpH 3. Any bound acetate ion in the native lysozyme crystals will be replaced by water molecules or possibly chloride ions in the acetate-free crystals. Thus the change in electron density is almost at the limit of detectability in the difference Fourier syntheses used in the present study. The positions of bound acetate ions, if present, should eventually be detected in the process of the refinement of the lysozyme structure.

The crystallographic studies show that βMeGlcNAc binds in subsite C-β in acetate-free lysozyme with co-ordinates that are almost identical with those for βGlcNAc in subsite C-β in lysozyme in the presence of acetate. Evidently binding at subsite C-β is not affected by the presence or absence of acetate, nor by the presence or absence of the β-*O*-methyl group, nor by competition, in the case of βGlcNAc, by αGlcNAc. βMeGlcNAc is excluded from subsite C-α, presumably because binding in this subsite would bring the β-*O*-methyl group too close to the carboxy group of Glu-35.

Comparison of lysozyme structures in the tetragonal and triclinic crystal forms has shown two differences of relevance to this study. One is the failure of GlcNAc to bind in subsite C-α in triclinic lysozyme (Kurachi *et al.*, 1976). This difference may be rationalized by the observation that subsites D, E and F appear to be blocked by a neighbouring molecule in the triclinic crystal form, whereas in the tetragonal-crystal form only sites E and F are blocked. Secondly it appears that the loop of residues 70–74 to the left of the active-site cleft (as viewed conventionally) is in a slightly different conformation in triclinic lysozyme and does not appear to move on binding monosaccharide inhibitors, although movement of Trp-62 is observed (L. H. Jensen, personal communication). It is not known at this stage if this difference arises from different packing arrangements in the crystal or from other effects.

Indeed the explanation for the conformational movements of the region 70–76 in tetragonal lysozyme on binding monosaccharides is not obvious. This region is relatively exposed in the enzyme and does not appear to be affected by lattice contacts, with the exception of the side chain of Arg-73 (Shrake & Rupley, 1973). The loop consisting of residues 70–76 is linked to the active-site-binding residues Trp-62 and Trp-63 through main-chain hydrogen bonds from the NH group of Cys-76 to the CO group of Trp-63 and a weak hydrogen bond from the NH

group of Ile-75 to the CO group of Trp-62, whereas the remainder of the loop is in Van der Waals contact with the main chain of residues 61–63 without making specific contacts. Thus it may be that the movement of Trp-62 together with a change in the dielectric constant of the active-site cleft on binding monosaccharides in subsite C acts as a trigger for a movement of residue 75, which in turn causes a shift of the whole of the region residues 70–76 towards the cleft. The present studies seem to suggest that there is some degree of strain in the region (residues 70–76) on binding βMeGlcNAc, although further work is required to confirm this observation. It is interesting that movements in the region residues 70–76 are not observed in the ternary complex of lysozyme–Gd(III)–GlcNAc in which the Gd(III) cation occupies two mutually exclusive sites, with one position associated with Asp-52 and the other with Glu-35, and GlcNAc binds in subsite C-β (Perkins, 1977). Thus interactions with the Gd(III) in subsite D appear to prevent movements of residues 70–76.

In an independent study of lysozyme crystallized in the presence of 3 M-urea in the tetragonal crystal form (Snape *et al.*, 1974; Snape, 1974), it was found that a urea molecule bound at the acetamido specificity pocket in site C and that possibly as many as four other urea molecules were also associated with the active-site region. The lysozyme–urea complex showed many small conformational changes, almost all of which appear to be correlated with regions previously denoted as strained by Diamond (1974). In particular the movement of the loop consisting of residues 59–80 was similar to that observed for native lysozyme–inhibitor complexes, and no further conformational changes were observed in this region on binding the transition-state analogue, the δ-lactone derivative of tetra-*N*-acetylglucosamine. Snape (1974) has proposed that the conformational changes observed on binding urea to lysozyme may be associated with relaxation of strained features and that the movement of the loop consisting of residues 59–80, in particular, is triggered by the binding of a molecule in the acetamido specificity site. A detailed analysis of the conformation of the lysozyme–urea complex is required to support this attractive hypothesis.

The results discussed in this paper are of direct relevance in the interpretation of n.m.r. studies of the inhibitor–enzyme complexes. The positions of the monosaccharides, GlcNAc and βMeGlcNAc, and the conformational changes that include the aromatic rings in the active site of Trp-62, Trp-63 and Trp-108 to varying extents have been used to predict the changes seen in ^{1}H n.m.r. spectra of the protein–inhibitor complex by means of ring-current calculations and estimates of other shielding effects. In spite of the preliminary nature of the real-space refinements of the lysozyme–saccharide complexes, such calculations are giving useful insights into the behaviour of lysozyme in solution (Perkins *et al.*, 1977; Perkins, 1977).

We thank Dr. R. A. Dwek for useful discussions on the n.m.r. studies. S. J. P. acknowledges financial support from the Medical Research Council (1973–1976). This paper is a contribution from the Oxford Enzyme Group, which is supported by the Science Research Council.

References

Arndt, U. W., North, A. C. T. & Phillips, D. C. (1964) *J. Sci. Instrum.* **41**, 421–425

Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. (1965) *Nature (London)* **206**, 757–761

Blake, C. C. F., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. (1967*a*) *Proc. R. Soc. London Ser. B* **167**, 365–377

Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. (1967*b*) *Proc. R. Soc. London Ser. B* **167**, 378–388

Diamond, R. (1966) *Acta Crystallogr.* **5**, 644–653

Diamond, R. (1974) *J. Mol. Biol.* **82**, 371–391

Dodson, E. J., Isaacs, N. W. & Rollett, J. S. (1976) *Acta Crystallogr.* **A32**, 311–315

Hamilton, W. C., Rollett, J. S. & Sparks, R. A. C. (1965) *Acta Crystallogr.* **18**, 129–130

Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C. & Rupley, J. A. (1972) *Enzymes 3rd Ed.* **7**, 665–868

Johnson, L. N. (1966) *Acta Crystallogr.* **21**, 885–891

Jones, R., Dwek, R. A. & Forsen, S. (1974) *Eur. J. Biochem.* **47**, 271–283

Kurachi, K., Sieker, L. H. & Jensen, L. H. (1976) *J. Mol. Biol.* **101**, 11–24

Levitt, M. (1974) *J. Mol. Biol.* **82**, 393–420

Luzzati, V. (1953) *Acta Crystallogr.* **6**, 142–152

North, A. C. T. (1965) *Acta Crystallogr.* **18**, 212–216

North, A. C. T., Phillips, D. C. & Mathews, F. S. (1968) *Acta Crystallogr. Sect.* A **24**, 351–359

Perkins, S. J. (1977) Ph.D. Thesis, University of Oxford

Perkins, S. J., Johnson, L. N., Phillips, D. C. & Dwek, R. A. (1977) *FEBS Lett.* **82**, 17–22

Phillips, D. C. (1964) *J. Sci. Instrum.* **41**, 123–129

Shindo, H. & Cohen, J. S. (1976) *Proc. Natl. Acad. Sci. U.S.A.* **73**, 1979–1983

Shrake, A. & Rupley, J. A. (1973) *J. Mol. Biol.* **79**, 351–371

Snape, K. W. (1974) Ph.D. Thesis, University of Oxford

Snape, K. W., Tijian, R., Blake, C. C. F. & Koshland, D. E. (1974) *Nature (London)* **250**, 295–298

Sykes, B. D. (1969) *Biochemistry* **8**, 1110–1116.

Winkler, F. K. & Dunitz, J. D. (1971) *J. Mol. Biol.* **59**, 169–182