Acta Cryst. (1993). D**49**, 497–504

X-ray Structure of Monoclinic Turkey Egg Lysozyme at 1.3 Å Resolution

BY KAZUAKI HARATA

Biomolecules Department, National Institute of Bioscience and Human Technology, 1-1 *Higashi, Tsukuba, Ibaraki* 305, *Japan*

(*Received* 8 *March* 1993; *accepted* 2 *June* 1993)

Abstract

Monoclinic crystals of turkey egg lysozyme (TEL, E.C. 3.2.1.17) were obtained from 2.2 *M* ammonium sulfate solution at pH 4.2. They belong to space group $P2_1$ with unit-cell dimensions $a = 38.07$, $b = 33.20$, $c = 46.12$ Å and $\beta = 110.1°$, and contain one molecule in the asymmetric unit ($V_m = 1.91$ Å^3 Da^{-1}). The three-dimensional structure of TEL was solved by the method of multiple isomorphous replacement with anomalous scattering. Area detector data to 1.5 Å resolution from native and heavy-atom derivatives were used for the structure determination. The structure was refined by the simulated-annealing method with diffraction data of 10–1.30 Å resolution. The conventional *R* factor was 0.189. The root-mean-square deviations from ideal bond distances and angles were 0.016 Å and 2.9°, respectively. The backbone structure of TEL is very similar to that of hen egg lysozyme (HEL) and the difference in seven amino-acid residues does not affect the basic folding of the polypeptide chain. Except for the region from Gly101 to Gly104, the geometry of the active-site cleft is conserved between TEL and HEL. The Gly101 residue is located at the end of the sugar-binding site and the structural change in this region between TEL and HEL is considered to be responsible for the difference in their enzymatic properties.

Introduction

The three-dimensional structure of lysozymes has been extensively investigated since X-ray analysis first revealed the crystal structure of hen egg lysozyme (HEL) (Blake *et al.*, 1965) and elucidated a possible mechanism of enzyme action (Blake *et al.*, 1967). Several structures of chicken-type lysozymes (Moult *et al.*, 1976; Aschaffenburg *et al.*, 1980; Hogle *et al.*, 1981; Artymiuk & Blake, 1981; Artymiuk, Blake, Rice & Wilson, 1982; Berthou, Lifchitz, Artymiuk & Jolles, 1983; Kundrot & Richards, 1987; Kodandapani, Suresh & Vijayan, 1990), including antigen–antibody complexes (Amit, Mariuzza, Phillips & Poljak, 1986; Sheriff *et al.*, 1987), have been reported. The active site of HEL consists of six subsites designated *A* to *F* and binds not only substrate polysaccharides but also mono- and oligosaccharides of substrate analogues (Blake *et al.*, 1967; Kelly, Sielecki, Sykes, James & Phillips, 1979). Turkey egg lysozyme (TEL) belongs to the same class of hen lysozyme and differs in seven of the 129 amino-acid residues (LaRue & Speck, 1970) in its primary structure (Table 1). The most important difference between these two lysozymes with respect to the enzyme action is in the 101st residue. The active site of HEL has several carboxylate groups which contribute to the saccharide binding and catalytic reaction (Johnson *et al.*, 1988). The Asp101 residue is located at the subsite *A* in HEL and it has been suggested that it forms hydrogen bonds with the substrate saccharide (Blake *et al.*, 1967; Ford, Johnson, Machin, Phillips & Tjian, 1974; Cheetham, Artymiuk & Phillips, 1992). In TEL, the 101st residue is replaced by glycine which has no side-chain group with which to form a hydrogen bond with the substrate. This causes the change in the pH dependence of substrate binding (Banergee & Rupley, 1975) and the free energy for substrate binding of TEL is smaller than that of HEL at acidic pH.

Table 1. *Difference in amino-acid sequence between turkey egg lysozyme (TEL) and hen egg lysozyme (HEL)*

Residue No.	3	15	41	73	99	101	121
TEL	Tyr	Leu	His	Lys	Ala	Gly	His
HEL	Phe	His	Gln	Arg	Val	Asp	Gln

The first crystallographic study of TEL was reported by Bott & Sarma (1976) who crystallized the enzyme at pH 8.0 from 18% NaCl solution and the hexagonal crystals obtained were used for a structural study. Recently, Howell, Almo, Parsons, Hadju & Petsko (1992) have reported the full structure of hexagonal TEL determined at 2.5 Å resolution. Alternatively, we have found that TEL crystallizes in a monoclinic form at acidic pH levels in the presence of ammonium sulfate. The monoclinic crystal diffracts to 1.3 Å resolution and the high-resolution structure analysis has been performed.

Materials and methods

Crystallization

TEL was purchased from Sigma Chemical Company and used without further purification. Crystals were prepared from a 2.2 *M* ammonium sulfate solution at pH 4.2 by the batch method. To 2 ml of 10% aqueous 1-propanol solution, 100 mg of TEL was dissolved, then 5 ml of 3 *M* ammonium sulfate solution was slowly added. The pH of the solution was adjusted to 4.2 with 0.1 *N* sulfuric acid. The solution was passed through a 0.45 μm filter and divided into 1 ml portions in a small culture dish which was allowed to stand at room temperature in a sealed plastic container. Sword-like crystals appeared in a few days and sheaves of large crystals were grown in two weeks. The crystal was monoclinic and the space group was $P2_1$ with $Z = 2$ (V_m = 1.91 Å^3 Da^{-1}). Unit-cell dimensions were $a = 38.07$, $b = 33.20$, $c = 46.12$ Å and $\beta = 110.1°$. Crystals were stable in a 3 *M* ammonium sulfate solution at pH 4.2. Therefore, soaking experiments for the preparation of heavy-atom derivatives were carried out under these conditions.

MIR phase determination

Three heavy-atom reagents were used to prepare heavy-atom derivatives: $HgCl_2$ (10 m*M*, 2 d), K_2PtCl_4 (1 m*M*, 30 h) and PCMBS (5 m*M*, 5 d). A crystal of the PCMBS derivative was immersed in 1 m*M* K_2PtCl_4 solution for a further 30 h and used as the fourth derivative. Intensity data for native crystal and the four derivatives were collected on a Nicolet *P*3/F diffractometer to a maximum resolution of 3 Å with graphite-monochromated Cu $K\alpha$ radiation (40 kV and 25 mA) and ω-scan mode. The scan speed was selected between 2.0 and 16.0° min^{-1} according to the peak intensity. The absorption correction was made by using the program implemented in the data-collection package of the *P*3/F system. The electron-density map was calculated using multiple isomorphous replacement (MIR) phases (figure of merit = 0.78). The map clearly showed the molecular boundary although continuity of electron density between adjacent molecules was found at two regions where molecules are in close contact. The amino-acid sequence was unambiguously traced in the region of residues 1–117. However, the electron density of some portions in the region of 118–129 was poor and the discontinuity of electron density prevented the correct assignment of amino-acid positions.

The crystal of TEL was not significantly deteriorated by soaking in $HgCl_2$ and PCMBS. The phase determination at 1.5 Å resolution was performed by using these two heavy-atom derivatives. The soaking experiments were carried out in the same conditions as described above, but crystals were kept for two months in the heavy-atom solution. Intensity data were collected to 1.3 Å resolution for the native crystal and 1.5 Å resolution for the heavy-atom crystals on an Enraf–Nonius FAST diffractometer equipped with an Elliott GX21 generator (40 kV, 60 mA, and focal spot size 0.3 × 3 mm). Three crystals were used for the collection of native data. A set of unique reflections to 1.30 Å resolution was obtained for each crystal and three data sets were merged to produce a data set of 25 232 reflections that corresponds to 91.3% of unique reflections. For each heavy-atom derivative, data collection to 1.50 Å resolution was carried out by the same procedure. Data sets of 13 029 unique reflections (PCMBS derivative) and 13 456 reflections ($HgCl_2$ derivative) were obtained. Statistics of reflection data for the native crystal are given in Table 2. The heavy-atom parameters used for the 1.5 Å MIR phasing are given in Table 3. The figure of merit was 0.60.

Table 2. *Statistics for diffraction data to* 1.3 Å *resolution*

Crystal	Measured data	Unique data	R_{merge}*
No. 1	49318	21799 (78.8%)	0.102
No. 2	50560	20119 (72.8%)	0.119
No. 3	50199	21680 (78.4%)	0.092
Merge of three data sets		25232 (91.3%)	0.051

* $R_{merge} = \sum[I(hkl)_i - G_j I(hkl)_j]/\sum I(hkl)_i$, where $I(hkl)$ is the average value of observed intensities, $I(hkl)_i$, and G_j is the scale factor for the jth group.

Table 3. *Heavy-atom parameters used in the* 1.5 Å *MIR phasing*

x, y, z = fractional coordinates; Oc. = relative occupancy; B = isotropic temperature factor.

		Oc.	x	y	z	B (Å^2)
PCMBS	Hg1	0.436	0.0654	0.2500	0.2626	14.6
	Hg2	0.199	0.1865	0.3518	0.0255	18.7
	Hg3	0.205	0.1700	0.3680	0.0634	19.6
	Hg4	0.126	0.1656	0.3601	0.4214	35.1
$HgCl_2$	Hg1	0.284	0.4942	−0.0098	0.1027	13.9
	Hg2	0.189	0.3346	0.3513	0.3991	14.0
	Hg3	0.153	0.1764	0.3662	0.4308	14.4

Model building and refinement

The sequence for the 118–129 region of the peptide could be traced from the electron-density map at 1.5 Å resolution. Except for Thr47, all the peptide groups were unambiguously assigned and most of the side-chain groups were located on the map. The structure was preliminary refined at 1.5 Å resolution by the stereochemically constrained least-squares method to an R value of 0.32. The simulated-annealing (SA) method using the program *X-PLOR* (Brünger, Kuriyan & Karplus, 1987) was applied for the refinement of the structure. A summary of the refinement is given in Fig. 1. The refinement consists of 13 stages. At each stage, the energy minimization was performed with 40 cycles for the coordinates and

20 cycles for the temperature factors. In the first and second stages, diffraction data to 1.5 Å were used. After the first simulated annealing followed by energy minimization, water molecules were picked from Fourier and difference Fourier maps. Peaks higher than 0.4 e $Å^{-3}$ in the difference Fourier map and having suitable interatomic contacts were taken into account, but water molecules with temperature factors higher than 60 $Å^2$ were omitted. After the second stage of refinement including 75 fully occupied water molecules, the R value was 0.193 for reflections with $|F_o| > 3\sigma(F)$ in the resolution range 10–1.5 Å. The root-mean-square (r.m.s.) deviations from ideal bond distances and angles at this stage were 0.012 Å and 2.8°, respectively. The atomic coordinates obtained after the second stage have been deposited with the Protein Data Bank.*

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1LZ3, R1LZ3SF). Free copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Supplementary Publication No. SUP 37090). A list of deposited data is given at the end of this issue.

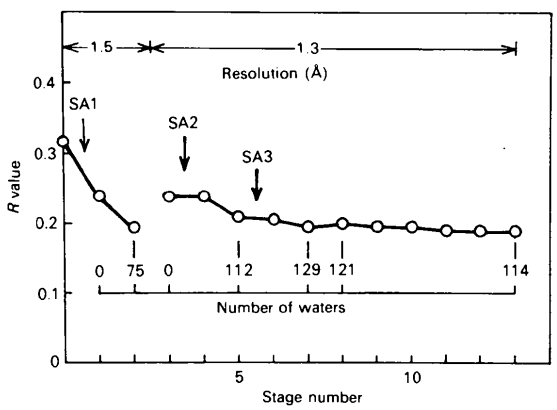

Fig. 1. Plot of the R value against the stage number of the refinement. At each stage, 40 cycles of energy minimization for the coordinates and 20 cycles for the temperature factor were performed. The *X-PLOR* refinement parameters used are as follows: Wa = 58832 (stages 1 and 2), 63909 (stages 3–13); tolerance = 0.005; parameters from the temperature factor bsigma = 1.5 (main chain), 2.0 (side chain); asigma = 2.0 (main chain), 2.5 (side chain); rweight = 0.060 (stage 1), 0.006 (stage 2), 0.071 (stages 3–5), 0.078 (stage 6), 0.020 (stages 7–11), 0.007 (stages 12 and 13); weight for energy terms, weight = 1.0 (stages 1–8), 0.8 (stage 9), 0.7 (stage 10), 0.5 (stages 11–13). SA1: first simulated annealing. The molecular dynamics simulation was performed at 3000 K for 1 ps with 1 fs step $^{-1}$, then the system was cooled to 300 K and the simulation was continued for 0.25 ps with 1 fs step $^{-1}$. SA2: second simulated annealing. The temperature of the system was decreased stepwise in order of 3000 K (1 ps, 1 fs step $^{-1}$), 1000 K (1 ps, 1 fs step $^{-1}$) and 300 K (0.5 ps, 1 fs step $^{-1}$). SA3: third simulated annealing. The harmonic restraint with 335 kJ mol^{-1} $Å^{-2}$ was imposed for water molecules. The simulation was performed at 1000 K for 1 ps (1 fs step $^{-1}$), then at 300 K for 0.5 ps (1 fs $step^{-1}$). For details, see Brünger (1992).

The further refinement was performed using the diffraction data of 10–1.30 Å resolution and 22 160 reflections with $|F_o| > 3\sigma(F)$. At the fourth stage, simulated annealing and energy minimization was performed. At the sixth stage, water molecules were included in the simulated-annealing calculation, where the harmonic restraint (335 kJ mol^{-1} $Å^{-2}$) was imposed on the water molecules. During stages 8–13, the restraints imposed to maintain the correct stereochemistry were gradually loosened. The weight for the energy terms was finally reduced to 0.5 and the restraint weight for the temperature factor was reduced to 0.007 which corresponds to 10% of the default value. Before each of stages 5, 7, 8 and 13, Fourier and difference Fourier maps were calculated and the electron density of the protein and water molecules were examined. At the final stage of the refinement, 114 fully occupied water molecules were included in the structure model. The R value was 0.189. The maximum values for positive and negative residual electron density were 0.76 and −0.48 e $Å^{-3}$, respectively. The r.m.s. residual density of the difference electron-density map was 0.087 e $Å^{-3}$ and no significant peak was found. The r.m.s. deviations of bond distances and angles from their ideal values were 0.016 Å and 2.9°, respectively. The refined atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven. Except for *X-PLOR*, the computer programs used were those developed in the author's laboratory.

Results and discussion

Crystals

The monoclinic crystal of TEL was obtained from a 2.2 *M* ammonium sulfate solution at pH 4.2 in the presence of 10% 1-propanol. 1-Propanol is not essential for the formation of the monoclinic crystal, but in the absence of 1-propanol, the crystals produced were fine needles and were not suitable for X-ray diffraction. Addition of 1-propanol, which decreases the rate of crystal growth, was effective for the formation of large crystals. Hexagonal crystals were also obtained at pH 4.2 in the presence of 3 *M* sodium nitrate. The cell dimensions were $a = b = 71.26$ and $c = 82.98$ Å. These values are similar to those reported by Bott & Sarma (1976) and Howell *et al.* (1992), although the dimension of the c axis is 2 Å shorter than they observed. The hexagonal crystal diffracted to a maximum resolution of 2.0 Å; therefore, the present monoclinic crystal is superior for the structure determination at high resolution.

Refinement of the structure

The structure refinement was carried out initially at 1.5 then at 1.3 Å resolution. The r.m.s. difference

for the superposition of equivalent C^{α} atoms before and after the 1.3 Å refinement was 0.083 Å. This indicates that the extension of the diffraction data from 1.5 to 1.3 Å does not significantly affect the main-chain coordinates. The large atomic movement was observed in several residues having a flexible side-chain group. Fig. 2 shows a plot of R value against $\sin\theta/\lambda$. The Luzzati plot (Luzzati, 1952) indicates that the mean error of positional coordinates is 0.15–0.2 Å. A Ramachandran map (Ramakrishnan & Ramachandran, 1965) is shown in Fig. 3. All φ and ψ torsion angles fall within the normal values.

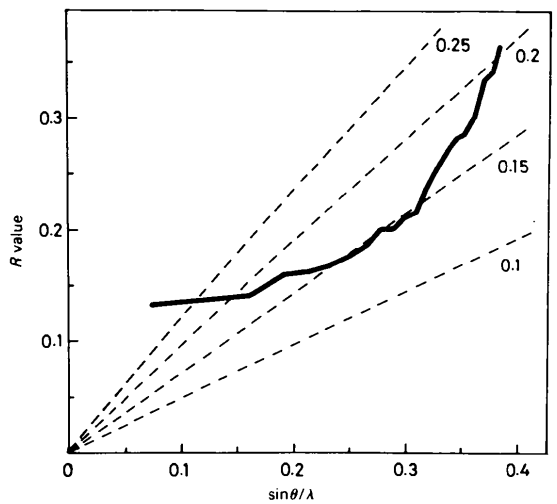

Fig. 2. Plot of the R factor as a function of $\sin\theta/\lambda$. Dashed lines indicate theoretical curves according to Luzzati, which correspond to the upper estimate of coordinate errors of 0.1, 0.15, 0.2 and 0.25 Å.

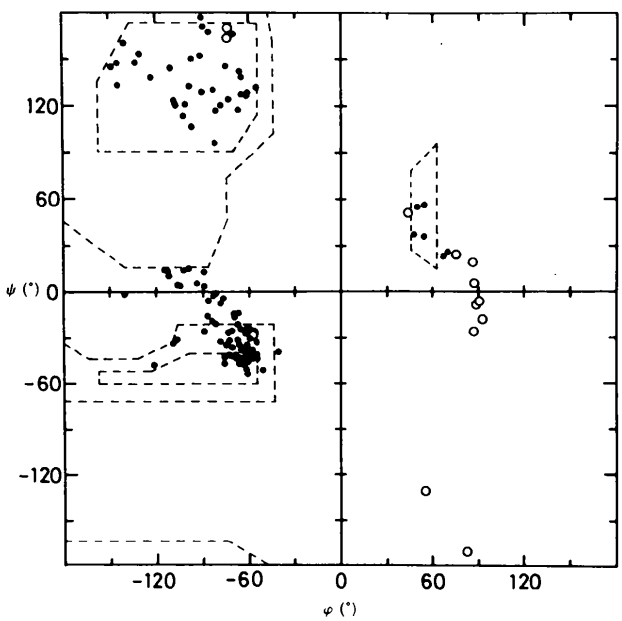

Fig. 3. Ramachandran plot for the main-chain torsion angles (φ,ψ). Glycine residues are shown as open circles.

Molecular structure

The backbone structure of TEL is shown in Fig. 4. Three long α-helices, Gly4–Gly16, Ser24–Ser36 and Ile88–Ser100, which are closely packed, form the core of the molecule. A short α-helix, Trp108–Cys115, and two short 3_{10}-helices, Pro79–Leu83 and Val120–Arg125, are also found. Three strands of Ala42–Asn46, Ser50–Gly54 and Gln57–Ser60 form antiparallel β-sheets. Twelve β-turns are classified into six types (Venkatacharam, 1968), four type I turns, three type II turns, and one for each of type I′, type II′, type III and type III′ (Fig. 5).

The temperature factor for each peptide and side-chain group is plotted in Fig. 5. The average temperature factor is 15.8 for all atoms, 12.8 for peptide groups and 19.0 $Å^2$ for side-chain groups. The relatively high temperature factor of the main chain is observed in the regions of Thr47–Gly49 and Arg125–Leu129. The former corresponds to the turn region connecting two β-strands and the latter is the C-terminal region.

Intramolecular interactions of side-chain groups are given in Table 4. There are 36 main-chain–side-chain contacts and 13 side-chain–side-chain contacts less than 3.2 Å. These pairs are considered to form hydrogen bonds or salt linkages to stabilize the tertiary structure. Many water molecules also contribute to stabilizing the protein structure. Most of the bound water molecules are distributed on the surface of the protein molecule. There are 37 water molecules which form at least two hydrogen bonds with TEL. These water molecules are mostly found in the active-site cleft and loop regions.

Crystal packing

The molecule contacts with six near neighbours four of which are symmetry related by the twofold screw axis (Fig. 6). Intermolecular distances less than 3.2 Å are listed in Table 5. There are three close contacts between molecules related by a twofold screw axis, Gly16···Asn77, Phe34···Asp119 and Gly71···Asn103, in which main-chain peptide groups are within the distance of a hydrogen-bonding contact. The molecule has contacts with two neighbours along the b axis. Seven water molecules are found to form a hydrogen-bond bridge linking two protein molecules. A solvent channel passes through the crystal along the b axis at $x = 0.25$ and $z = 0.75$. The active site of the molecule is open to this solvent region.

Table 4. *Intramolecular contacts* (Å) *less than* 3.2 Å *involving side chains*

	Distance		Distance
NZ(Lys1)···OE1(Glu7)	3.0	OG1(Thr51)···NH2(Arg68)	2.9
OH(Tyr3)···O(Ser86)	2.7	OH(Tyr53)···OD1(Asp66)	2.7
N(Gly4)···OE2(Glu7)	3.0	O(Gly54)···NE2(Gln57)	2.9
NH1(Arg5)···O(Arg125)	3.1	O(Leu56)···NE1(Trp108)	2.8
NH2(Arg5)···O(Trp123)	2.9	OD1(Asn59)···N(Arg61)	3.0
NH2(Arg5)···O(Arg125)	3.0	OD1(Asn59)···N(Trp62)	3.0
NZ(Lys13)···OD2(Asp18)	3.1	O(Ser60)···OG(Ser72)	2.9
O(Leu15)···Nz(Lys96)	2.8	OG(Ser60)···OG1(Thr69)	2.8
O(Leu17)···NE1(Trp28)	3.2	O(Cys64)···OG(Ser72)	2.8
OD1(Asp18)···N(Leu25)	3.1	OD1(Asn65)···N(Asp66)	3.0
NH1(Arg21)···O(Gly101)	3.1	OD1(Asn65)···N(Gly67)	2.9
NH2(Arg21)···O(Ala99)	2.6	OD2(Asp66)···N(Arg68)	3.1
OH(Tyr23)···N(Met105)	3.1	OD2(Asp66)···N(Thr69)	3.1
O(Ser24)···ND2(Asn27)	3.1	OD2(Asp66)···OG1(Thr69)	2.7
OD1(Asn27)···NE1(Trp111)	3.0	ND2(Asn74)···N(Asn77)	3.2
O(Ala32)···OG(Ser36)	2.9	ND2(Asn74)···N(Ile78)	3.0
OG(Ser36)···O(Ile55)	2.8	ND2(Asn74)···O(Ile78)	3.0
OD1(Asn39)···N(His41)	3.0	OG(Ser85)···O(Asp87)	2.9
O(Ala42)···NE2(Gln57)	3.1	OD2(Asp87)···N(Thr89)	3.1
ND2(Asn44)···OE1(Gln57)	3.0	OD2(Asp87)···OG1(Thr89)	2.6
ND2(Asn46)···OG(Ser50)	2.8	O(Lys96)···OG(Ser100)	2.9
ND2(Asn46)···OD2(Asp52)	3.0	O(Asn103)···OD1(Asn106)	3.0
OD1(Asp48)···N(Ser50)	3.2	O(Asn106)···NH1(Arg112)	2.6
OD1(Asp48)···OG(Ser50)	2.7	ND2(Asn106)···NZ(Lys116)	3.0
N(Thr51)···OG(Ser60)	3.0	O(Cys115)···OG1(Thr118)	2.8

Table 5. *Intermolecular distances* (Å) *less than* 3.2 Å

	Distance	Symmetry operator
Protein protein interactions		
NH2(Arg45)···N(Gly117)	3.1	$x, 1+y, z$
NZ(Lys33)···O(Gly117)	2.9	$1-x, \frac{1}{2}+y, 1-z$
O(Phe34)···N(Asp119)	3.0	$1-x, \frac{1}{2}+y, 1-z$
OD1(Asn37)···O(Lys116)	3.1	$1-x, \frac{1}{2}+y, 1-z$
O(Thr43)···NH1(Arg114)	3.2	$1-x, \frac{1}{2}+y, 1-z$
NH2(Arg114)···OD1(Asp119)	2.8	$1-x, \frac{1}{2}+y, 1-z$
NH1(Arg61)···O(Ser100)	3.0	$1-x, \frac{1}{2}+y, 2-z$
O(Gly71)···N(Gly102)	3.2	$1-x, \frac{1}{2}+y, 2-z$
O(Gly71)···N(Asn103)	2.8	$1-x, \frac{1}{2}+y, 2-z$
NH1(Arg21)···NH1(Arg61)	3.2	$1-x, \frac{1}{2}+y, 2-z$
O(Arg14)···OD1(Asn74)	3.2	$2-x, \frac{1}{2}+y, 2-z$
N(Gly16)···O(Asn77)	3.0	$2-x, -\frac{1}{2}+y, 2-z$
Contacts *via* single-bridging water		
O(Val2)···Water···O(Gly126)	2.7, 2.8	$2-x, -\frac{1}{2}+y, 1-z$
OD1(Asn27)···Water···NH1(Arg45)	3.2, 2.8	$x, 1+y, z$
O(Thr43)···Water···NE(Arg114)	3.1, 2.9	$1-x, \frac{1}{2}+y, 1-z$
OD1(Asn44)···Water···O(Arg114)	2.9, 2.9	$1-x, \frac{1}{2}+y, 1-z$
O(Pro70)···Water···N(Gly102)	2.8, 2.8	$1-x, \frac{1}{2}+y, 2-z$
OG(Ser81)···Water···ND2(Asn93)	2.8, 3.1	$2-x, \frac{1}{2}+y, 2-z$
OD1(Asn113)···Water···NH2(Arg125)	2.8, 2.8	$1-x, \frac{1}{2}+y, 1-z$

Amino-acid replacement and conformational change

The primary structure of TEL differs in seven amino-acid residues from that of HEL. These residues, which are well defined on the electron-density map, are located on the surface of the molecule and exposed to solvent. The hydroxyphenyl group of Tyr3 is in the same orientation as the phenyl group Phe3 in HEL and is hydrogen bonded to the carbonyl O atom of Ser86. The replacement of the 15th residue, His, with Leu changes the charge distribution of the molecular surface and also changes the hydration of this region because of the hydrophobic side-chain group in Leu15. The replacement of the 41st residue, Gln, with His also affects the charge distribution of the molecular surface. The 73rd residue is located at a flexible region of the molecule. It is doubtful that the amino-acid replacement from Arg to Lys is responsible for the conformational change of this region since the side-chain group is fully exposed to solvent and has no intramolecular contact with other residues. The Gly101–Gly104 region of TEL shows a β-turn with a 4→1 hydrogen bond. By changing the 101st residue from Asp to Gly, the main chain becomes more flexible, and, as a result, its conformation is changed to another stable form. The 99th residue is located at the end of a helix region and the difference between Ala and Val does not affect the participation of this residue in the formation of the helix structure. The Gln121 residue of HEL is replaced by His in TEL. The imidazolyl group of His121 is exposed to solvent and only affects the charge distribution of the molecular surface.

The least-squares calculation for the superposition of the equivalent C^{α} atoms gives an r.m.s. difference of 0.65 Å for the structure of hexagonal TEL (Parsons & Phillips, 1988), 0.47 Å for the tetragonal HEL (Kundrot & Richards, 1987), 0.84 Å for the triclinic HEL (Ramanadham, Sieker & Jensen, 1990), and 0.66 and 0.72 Å for the monoclinic HEL (Rao, Hogle & Sundaralingam, 1983). The basic folding of the polypeptide chain is same as that of HEL found in these crystals. The structure of the Lys1–Arg45

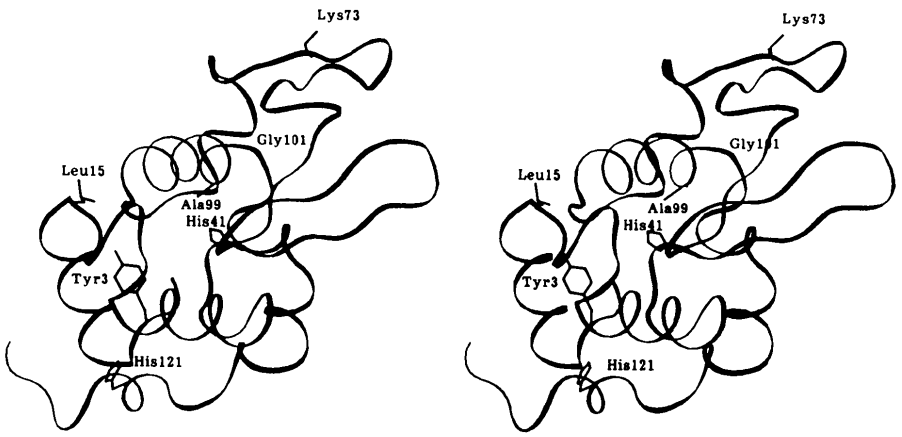

Fig. 4. Stereoview of the backbone structure of TEL. Seven amino-acid residues which differ from those of HEL are highlighted with side-chain structures (thin lines).

Table 6. *Comparison of* (φ, ψ) *values* (°) *for the residues* 100–106

Residue No.		TEL[a] Monoclinic	TEL[b] Hexagonal	HEL[c] Triclinic	HEL[d] Tetragonal	HEL[e] Monoclinic Mol. 1	HEL[e] Monoclinic Mol. 2
Ser100	φ	-78	-79	-79	-88	-42	-77
	ψ	-7	-29	-28	-6	-45	-14
Gly101	φ	-69	-78	-90	-75	-35	-73
(Asp)	ψ	170	166	40	-3	-90	-85
Gly102	φ	-57	-31	75	118	-152	-117
	ψ	-35	-67	19	-31	173	-63
Asn103	φ	-99	-63	-115	-102	35	-112
	ψ	15	-21	146	12	49	61
Gly104	φ	56	69	-92	58	37	33
	ψ	-131	-129	-155	-145	-138	-121
Met105	φ	76	79	-76	-71	-100	-106
	ψ	-5	4	6	-9	50	24
Asn106	φ	-65	-72	-66	-63	-89	-81
	ψ	-14	1	-15	-19	-19	22

References: (*a*) Present work. (*b*) Parsons & Phillips (1988). (*c*) Ramanadham *et al.* (1990). (*d*) Kundrot & Richards (1987). (*e*) Rao *et al.* (1983).

region, which includes two long helices, is almost the same in TEL and HEL. The structure of the Leu75–Ala99 region is also well fitted. On the other hand, there are some regions which are flexible in their backbone structure. The relatively large difference in C^{α} positions is found in regions Thr47–Gly49, Thr69–Ser72 and Ser100–Asn103. These regions are exposed to solvent and comprised of a turn or loop structure. The conformation of the Thr47–Gly49 region, which links two β-strands, closely resembles that of monoclinic and tetragonal HEL, while a large difference of the equivalent C^{α} positions (3.0 Å for Thr47) is observed in comparison with triclinic HEL. The structural difference is also observed in the Thr69–Ser72 region between monoclinic TEL and hexagonal TEL. The Thr69–Ser72 region of HEL has conformational mobility and the main-chain conformation differs among HEL stuctures of different crystal forms.

The Ser100–Gly104 region includes Gly101, which is Asp in HEL. A typical type I β-turn is observed in this region of TEL. The residues of Gly101–Gly104 in hexagonal TEL also form a structure similar to that of a type I β-turn although the values of φ and ψ (Table 6) deviate considerably from their ideal values. In contrast, the corresponding φ and ψ values in HEL differ from those of typical β-turn structures. The variation of φ and ψ angles indicates high conformational flexibility to this region. The possibility of the presence of disorder has been highlighted in the triclinic crystal of HEL (Ramanadham *et al.*, 1990). The high mobility of these regions has also been shown by the molecular-dynamics simulation of HEL and human lysozyme (Post *et al.*, 1986; Ichiye, Olafson, Swaminathan & Karplus, 1986).

The temperature factor is a measure of the relative mobility of a particular atom or group (Artymiuk *et al.*, 1979). Statistical analysis has shown that the positional difference between the equivalent atoms in two structures is correlated with the temperature factor (Bott & France, 1990). In addition to the C-terminal region, three regions, Thr47–Gly49, Thr69–Ser72 and Ser100–Asn103, frequently exhibit

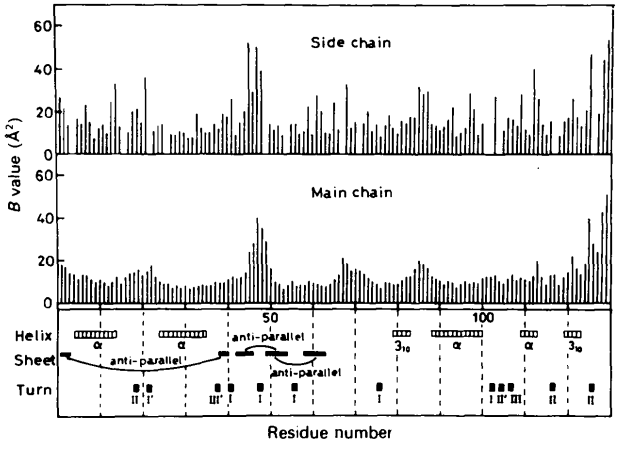

Fig. 5. Temperature factors averaged for the side chain (upper) and main chain (below). Typical secondary structures are shown.

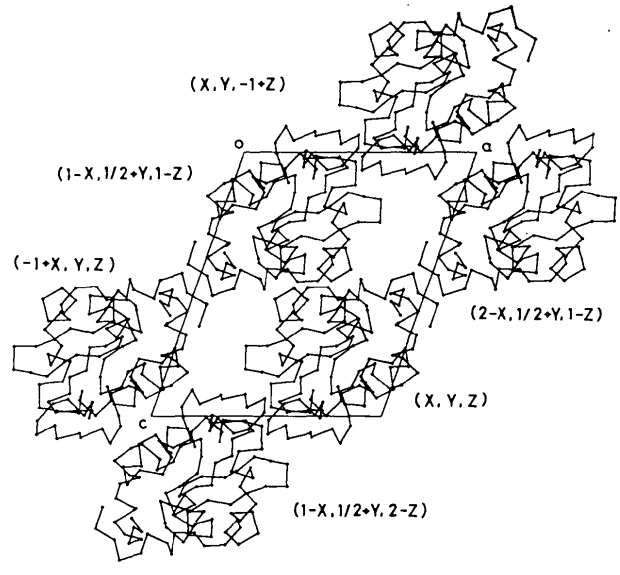

Fig. 6. Crystal packing of TEL viewed along the *b* axis.

high temperature factors in TEL and HEL. These regions, which coincide with most movable regions, may be in more than one conformation in solution and be fixed to a conformation suitable for a certain type of packing in the crystallization.

Active site

The structure of the active-site region is shown in Fig. 7. The cleft of the substrate-binding site includes water molecules which form a hydrogen-bond network with hydrophilic side-chain groups (Fig. 8). Except for the Gly101–Gly104 region, the main-chain structure of this region does not significantly differ from that of HEL. The side-chain groups of two catalytically important residues, Glu35 and Asp52, are in the same orientation as those found in HEL. The carboxyl group of Glu35 is hydrogen bonded to a water molecule, while the carboxyl group of Asp52 forms hydrogen bonds with Asn46. The Trp62 residue has been considered to play an important role in the substrate binding. The chemical modification of Trp62 in HEL to oxindolealanine reduces the enzyme activity and changes the mode of substrate binding (Blake *et al.*, 1981). In the TEL

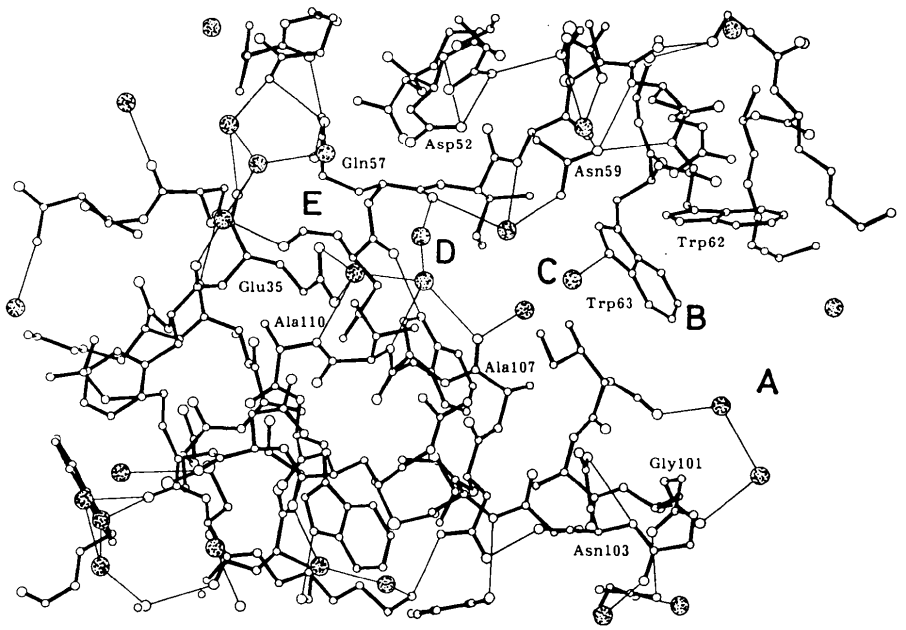

Fig. 7. Structure of the active-site region. Water molecules are shown by shaded circles. Intermolecular distances less than 3.2 Å, involving side-chain atoms and water molecules, are shown by thin lines. Subsites for the binding of sugar residues are denoted *A–E*.

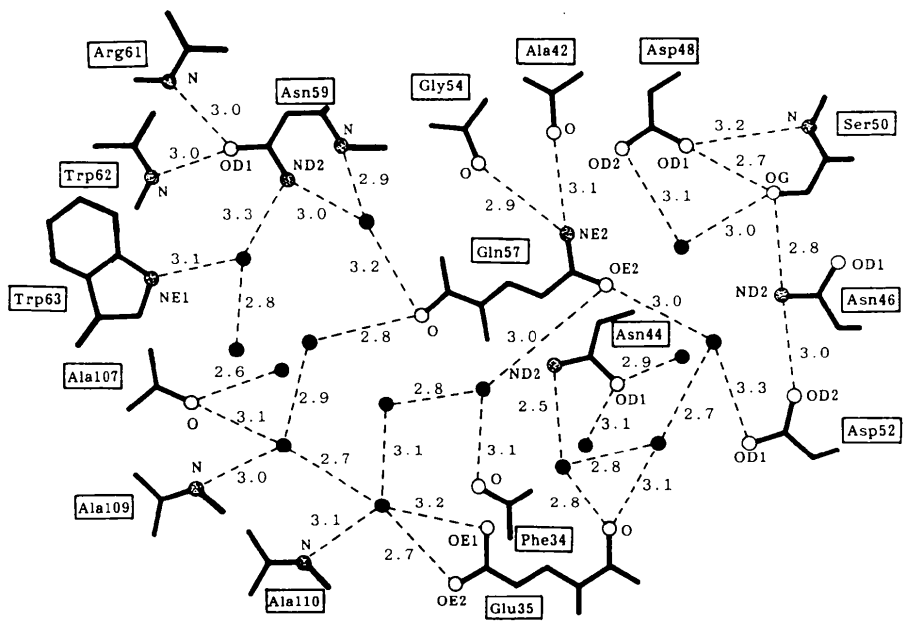

Fig. 8. Hydrogen-bond network in the active-site cleft. Distances less than 3.3 Å are denoted by dashed lines. Water molecules are shown by full circles.

structure, the indole moiety of Trp62 is rotated *ca* 30° from that found in triclinic and monoclinic HEL's, indicating the high flexibility of this side-chain group. The movement of Trp62 has been observed when HEL binds substrate sugars (Blake *et al.*, 1967; Perkins, Johnson, Machin & Philips, 1978). The indole moiety of Trp63, which is almost in the same orientation as that of HEL, forms a hydrogen bond with a water molecule. Therefore, except for highly flexible side-chain groups, the geometry of the active site does not significantly differ from that of HEL. This suggests that the substrate molecule can be bound to subsites *B–D* in the same manner as that found in the oligosaccharide complexes of HEL. The *N*-acetylglucosamine molecule has been found to be bound to the *C* subsite of HEL (Blake *et al.*, 1967). In solution under various pH conditions, the similarity has been highlighted in the binding of *N*-acetylglucosamine to HEL and TEL (Yang, Kuramitsu, Nakae, Ikeda & Hamaguchi, 1976). However, when the substrate or its analogue is bound to subsite *A*, the change in the 101st residue may affect the mode of binding. The tri-*N*-acetylchitotriose molecule occupies subsites *A*, *B* and *C* in the crystalline state of HEL (Blake *et al.*, 1967; Cheetham *et al.*, 1992). The side-chain groups of Asp101 and Asn103 form hydrogen bonds with sugar residues. In the crystalline complex of TEL with tri-*N*-acetylchitotriose (Harata, unpublished work), subsites *B* and *C* accommodate sugar residues in the same manner as that found in the HEL complex, but the sugar residue located at subsite *A* is in a different orientation and has no direct contact with the protein molecule. The binding of trisaccharide changes the β-turn structure of residues Gly101–Gly104 from type I to type III. The side-chain group of Asn103 is linked to a sugar residue by a water-mediated hydrogen-bond bridge. Flexibility in this region may be required to bind the substrate with hydrogen-bond formation and to release the reaction product by breaking hydrogen bonds. The lack of the 101st carboxyl group in TEL also explains the difference in the pH dependence of tri-*N*-acetylchitotriose binding in solution (Banergee & Rupley, 1975). The similarity and local changes in the structure between TEL and HEL suggest that the difference in substrate binding and catalytic reaction is mainly ascribed to the change in the 101st residue.

References

Amit, A. G., Mariuzza, R. A., Phillips, S. E. V. & Poljak, R. J. (1986). *Science*, **233**, 747–753.

Artymiuk, P. J. & Blake, C. C. F. (1981). *J. Mol. Biol.* **152**, 737–762.

Artymiuk, P. J., Blake, C. C. F., Grace, D. E. P., Oatley, S. J., Phillips, D. C. & Sternberg, M. J. E. (1979). *Nature (London)*, **280**, 563–568.

Artymiuk, P. J., Blake, C. C. F., Rice, D. W. & Wilson, K. S. (1982). *Acta Cryst.* B**38**, 778–783.

Aschaffenburg, R., Blake, C. C. F., Dickie, H. M., Gayen, S. K., Keegan, R. & Sen, A. (1980). *Biochim. Biophys. Acta*, **625**, 64–71.

Banergee, S. K. & Rupley, J. A. (1975). *J. Biol. Chem.* **250**, 8267–8274.

Berthou, J., Lifchitz, A., Artymiuk, P. J. & Jolles, P. (1983). *Proc. R. Soc. London Ser. B*, **217**, 471–489.

Blake, C. C. F., Cassells, R., Dobson, C. M., Poulsen, F. M., Williams, R. J. P. & Wilson, K. S. (1981). *J. Mol. Biol.* **147**, 73–95.

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**, 365–377.

Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. (1965). *Nature (London)*, **206**, 757–763.

Bott, R. & France, J. (1990). *Protein Eng.* **3**, 649–657.

Bott, R. & Sarma, R. (1976). *J. Mol. Biol.* **106**, 1037–1046.

Brünger, A. T. (1992). *X-PLOR Manual.* Version 3.0. Yale Univ., New Haven, USA.

Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). *Science*, **235**, 458–460.

Cheetham, J. C., Artymiuk, P. J. & Phillips, D. C. (1992). *J. Mol. Biol.* **224**, 613–628.

Ford, L. O., Johnson, L. N., Machin, P. A., Phillips, D. C. & Tjian, R. (1974). *J. Mol. Biol.* **88**, 349–371.

Hogle, J., Rao, S. T., Mallikarjunan, M., Beddell, C., McMullan, R. K. & Sundaralingam, M. (1981). *Acta Cryst.* B**37**, 591–597.

Howell, P. L., Almo, S. C., Parsons, M. R., Hadju, J. & Petsko, G. A. (1992). *Acta Cryst.* B**48**, 200–207.

Ichiye, T., Olafson, B. D., Swaminathan, S. & Karplus, M. (1986). *Biopolymers*, **25**, 1909–1937.

Johnson, L. N., Cheetham, J., McLaughlin, P. J., Acharya, K. R., Barford, D. & Phillips, D. C. (1988). *Curr. Top. Microbiol. Immunol.* **139**, 81–134.

Kelly, J. A., Sielecki, A. R., Sykes, B. D., James, M. N. G. & Phillips, D. C. (1979). *Nature (London)*, **282**, 875–878.

Kodandapani, R., Suresh, C. G. & Vijayan, M. (1990). *J. Biol. Chem.* **265**, 16126–16131.

Kundrot, C. E. & Richards, F. M. (1987). *J. Mol. Biol.* **193**, 157–170.

LaRue, J. N. & Speck, J. C. Jr (1970). *J. Biol. Chem.* **245**, 1985–1991.

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

Moult, J., Yonath, A., Traub, W., Smilansky, A., Podjarny, A., Rabinovich, D. & Saya, A. (1976). *J. Mol. Biol.* **100**, 179–195.

Parsons, M. R. & Phillips, S. E. V. (1988). Protein Data Bank, entry code 2LZ2.

Perkins, S. J., Johnson, L. N., Machin, P. A. & Phillips, D. C. (1978). *Biochem. J.* **173**, 607–616.

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.

Ramakrishnan, C. & Ramachandran, G. N. (1965). *Biophys. J.* **5**, 909–933.

Ramanadham, M., Sieker, L. C. & Jensen, L. H. (1990). *Acta Cryst.* B**46**, 63–69.

Rao, S. T., Hogle, J. & Sundaralingam, M. (1983). *Acta Cryst.* C**39**, 237–240.

Sheriff, S., Silverton, E. W., Padlan, E. A., Cohen, G. H., Smith-Gill, S. J., Finzel, B. C. & Davies, D. R. (1987). *Proc. Natl Acad Sci. USA*, **84**, 8075–8079.

Venkatacharam, C. M. (1968). *Biopolymers*, **6**, 1425–1436.

Yang, Y., Kuramitsu, S., Nakae, Y., Ikeda, K. & Hamaguchi, K. (1976). *J. Biochem.* **80**, 425–434.