PROTEINS: Structure, Function, and Genetics 22:245–258 (1995)

The Catalytic Domain of a Bacterial Lytic Transglycosylase Defines a Novel Class of Lysozymes

Andy-Mark W.H. Thunnissen,[1] Neil W. Isaacs,[2] and Bauke W. Dijkstra[1]

[1]Laboratory of Biophysical Chemistry and BIOSON Research Institute, Department of Chemistry, University of Groningen, 9747 AG Groningen, the Netherlands; [2]Protein Crystallography, Department of Chemistry, University of Glasgow, Glasgow G12 8QQ, Scotland

ABSTRACT **The 70-kDa soluble lytic transglycosylase (SLT70) from *Escherichia coli* is a bacterial exo-muramidase that cleaves the cell wall peptidoglycan, producing 1,6-anhydro-muropeptides. The X-ray structure of SLT70 showed that one of its domains is structurally related to lysozyme, although there is no obvious similarity in amino acid sequence. To relate discrete structural features to differences in reaction mechanism and substrate/product specificity, we compared the three-dimensional structure of the catalytic domain of SLT70 with the structures of three typical representatives of the lysozyme superfamily: chicken-type hen egg-white lysozyme, goose-type swan egg-white lysozyme, and phage-type lysozyme from bacteriophage T4. We find a particularly close relationship between the catalytic domain of SLT70 and goose-type lysozyme, with not only a significant similarity in overall structure, but even a weak homology in amino acid sequence. This finding supports the notion that the goose-type lysozyme takes up a central position in the lysozyme superfamily and that it is structurally closest to the lysozyme ancestors. The saccharide-binding groove is the most conserved part in the four structures, but only two residues are absolutely preserved: the "catalytic" glutamic acid and a structurally required glycine. The "catalytic" aspartate is absent in SLT70, a difference that can be related to a different mechanism of cleavage of the β-1,4-glycosidic bond. The unique composition of amino acids at the catalytic site, and the observation of a number of differences in the arrangements of secondary structure elements, define the catalytic domain of SLT70 as a novel class of lysozymes. Its fold is expected to be exemplary for other bacterial and bacteriophage muramidases with lytic transglycosylase activity.**

Key words: bacterial muramidase, peptidoglycan, structure comparison, sequence motifs, structure/function relationships, evolutionary relationships, X-ray structure

INTRODUCTION

Protein structures can adopt similar folds even in the absence of sequence similarity. The necessity to preserve a close-packed hydrophobic interior and to conserve the functional properties of the active site imposes more restraints on the protein's tertiary structure than upon its amino acid sequence.[1] The structural relationships among the (α/β)-hydrolase fold enzymes[2] form a recent example of this phenomenon.

Conservation of fold over sequence is also observed in the family of lysozymes. Based on amino acid sequence similarities and specific functional properties, three distinct classes of these enzymes have been discerned: the chicken-type, which includes the lysozymes from avian egg-white and mammalian secretions, the goose-type, found in the eggs of swan, ostrich, and goose, and the phage-type lysozyme, present in the lysate of bacterial viruses (see ref. 3 for a review). While within each of these classes there is significant sequence homology, between the different classes the amino acid sequences seem to be unrelated. Nevertheless, it appeared that the backbone conformations of members of the different classes and their active site architectures were remarkably similar.[4–7] This structural correspondence strongly suggests that the three lysozyme types have diverged from a common precursor.

Recently, we reported the crystal structure of a bacterial exo-muramidase from *Escherichia coli*, known as the 70-kDa soluble lytic transglycosylase (SLT70).[8] This enzyme appears to play an important role in the metabolism of the cell wall peptidoglycan during bacterial growth and division. Like lysozyme, it cleaves the β-1,4-glycosidic bonds between N-acetylmuramic acid (MurNAc) and N-acetylglucosamine (GlcNAc) in the peptidoglycan. However, SLT70 does not simply hydrolyze these bonds, but catalyzes an intra-molecular glycosyl transferase reaction whereby, concomitantly with cleavage, an an-

Received November 29, 1994; revision accepted February 10, 1995.

Address reprint requests to Bauke W. Dijkstra, Laboratory of Biophysical Chemistry, Department of Chemistry, University of Groningen, Nijenborgh 4, 9715 AG Groningen, the Netherlands.

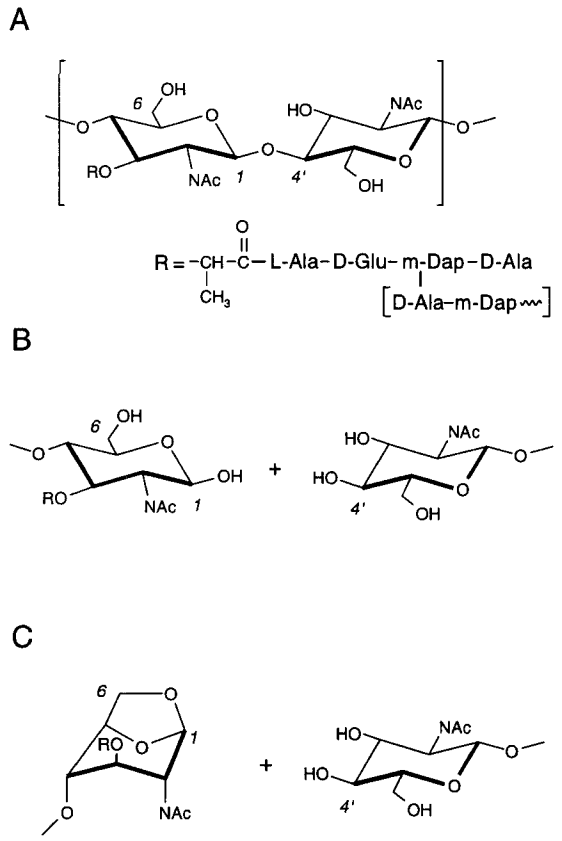


Fig. 1. A comparison of the reaction catalyzed by hen egg-white lysozyme (HEWL) and SLT70 from *E. coli*. **A:** Repeating unit of the cell wall peptidoglycan, the common substrate of the two enzymes. The peptidoglycan polymer is built up of linear glycan strands of alternate GlcNAc and MurNAc residues, cross-linked by short peptides that are attached to the C_3-lactyl group of the MurNAc residues. Lytic cleavage takes place at the β-1,4-glycosidic bond between MurNAc and GlcNAc. The peptide units in *E. coli* peptidoglycan are composed of L-alanine, D-isoglutamic acid, meso-diaminopimelic acid and D-alanine. **B:** The products of the glycosyl hydrolase reaction catalyzed by HEWL, containing reducing muramic acid residues. **C:** The products of the glycosyl transferase reaction catalyzed by SLT70, containing 1,6-anhydro-muramic acid residues.

hydro bond is formed between carbons C_1 and C_6 of the N-acetylmuramic acid[9,10] (Fig. 1). No significant sequence similarity could be detected between SLT70 and various muramidases, including the lysozymes.[11] Therefore it came as a surprise to find in the three-dimensional structure of this monomeric enzyme a C-terminal domain with a lysozyme-like structure in which the active site is located.[8]

Apart from SLT70, two other lytic transglycosylases have been found in *E. coli:* a 35-kDa enzyme that, like SLT70, is located in the periplasmic space[12] and a 38-kDa membrane-bound enzyme.[13] Considering their involvement in the bacterial cell wall metabolism, the transglycosylases are highly interesting targets for the development of new types of antibiotics. Nevertheless, only a few studies have been performed on the mechanism by which these enzymes attack the peptidoglycan, and their precise physiological function is still elusive. The lysozymes, on the other hand, form a well-studied class of enzymes and the structure/function relationships of hen egg-white lysozyme and lysozyme from bacteriophage T4 (in particular) have been studied intensively. Therefore, to substantiate the structure/function and evolutionary relationships of SLT70, we carried out a detailed comparison of its C-terminal domain (C-SLT) with the high-resolution crystal structures of chicken-type hen egg-white lysozyme (HEWL), goose-type lysozyme from the egg-white of the Australian black swan (SEWL), and the phage-type lysozyme from bacteriophage T4 (T4L). We conclude that C-SLT belongs to a novel, distinct class of lysozymes with lytic transglycosylase activity that is characterized by the absence of the catalytic base or nucleophile and by the occurrence of a number of differences in the arrangement of secondary structure elements.

MATERIALS AND METHODS

Atomic Coordinate Data and Structural Analysis

The protein model of SLT70 from *E. coli* used in this comparison was solved by x-ray crystallography to 2.7 Å resolution with an R-factor of 22.8% as described.[8] Structures for HEWL and T4L were taken from the Brookhaven Protein Data Bank[14] as entries 6LYZ and 3LZM. These models are fully refined to a resolution of 2.0 Å[15] and 1.7 Å resolution,[16] respectively. The crystal structure of SEWL has been described to 2.8 Å resolution,[17] but has meanwhile been refined to a resolution of 1.9 Å.[18] Coordinates of HEWL complexed with the trisaccharide MurNAc-GlcNAc-MurNAc[19] and of a T4L mutant containing a covalently bound muropeptide[20] were kindly provided by the authors. Coordinates of T4L with a modeled $(GlcNAc)_3$ were taken from Anderson et al.[21] Although very recently a crystal structure of goose-type lysozyme complexed with $(GlcNAc)_3$ has appeared,[22] at the time of our investigation such a structure was not available, and therefore we built a model of the SEWL-trisaccharide complex. For this purpose, the coordinates of MurNAc-GlcNAc-MurNAc were transformed into the SEWL coordinate system by superposing the backbones of SEWL and HEWL complexed with the trisaccharide. Subsequently the position and orientation of the trisaccharide were corrected to optimize the interactions with protein atoms.

Secondary structure was assigned using the program DSSP.[23] Solvent accessibilities were calculated as the area of the van der Waals surface of the protein atoms that can be contacted with a water probe of 1.4 Å radius, using the algorithm of Lee and Richards.[24] Residues with side chains less than 7%

accessible, compared with values calculated from G-X-G peptides in extended conformation, were taken as buried, based on studies described elsewhere.[25,26]

Structure Alignment Methods

The global structure comparison of C-SLT and the three lysozymes was done with Cα-atoms. Two different procedures were used to define topologically equivalent residues or "common core" regions for each protein pair. The first procedure made use of the method developed by Rossmann and Argos,[27] which allows for the small rigid-body movements of secondary structure elements commonly observed in distantly related structures. In this method equivalences are automatically defined and updated according to both the distances between potentially equivalent Cα atoms and local directions of the main chains in the least-squares fitted structures. The same method has been used in structural comparison studies of the three lysozymes types.[5,16] The other procedure made use of the least-squares structure comparison routine implemented in the program "O."[28] Starting from a small set of user-defined equivalences (see Table II), this routine automatically optimizes the global superposition and the number of structural equivalences of two proteins. It uses an algorithm that involves a search for structure fragments that can be aligned within a given cut-off limit. Atoms (in each protein pair) were considered equivalent if they were separated by a distance of less than 3.8 Å and were located within a consecutive region consisting of at least three residues. The "O"-routine is generally much more stable than the Rossmann-Argos method but finds only the closer equivalences. All superpositions were inspected by computer graphics for their accuracy, and, where needed, manual corrections were performed. To explore the overall relationships between the active sites of the four proteins we superposed the coordinates of the different lysozyme-trisaccharide complexes on the model of C-SLT using the appropriate transformation matrices obtained from the Rossmann-Argos alignment.

RESULTS AND DISCUSSION

Overall Structure of SLT70

The structure of SLT70 is built up almost entirely of α-helices that are assembled into three distinct domains generating a doughnut-like shape with overall dimensions of ~85 × 75 × 55 Å (Fig. 2). The first two domains (residues 1–448) have 26 helices in a superhelical arrangement forming a ring-shaped structure with a large central hole of ~30 Å in diameter. The function of this ring is still unknown, but it is assumed to play a role in murein-binding and/or regulation of the exo-lytic activity of the enzyme. On top of one side of the ring lies the catalytic C-terminal domain (residues 449–618), leaving the hole accessible. This domain has an ellipsoidal shape with approximate dimensions of ~50 × 35 × 30 Å. A deep groove running across the center divides the domain into two lobes. Three helices (Cα1 to Cα3) are found in the "lower" N-terminal lobe, while the "upper" C-terminal lobe contains five helices (Cα5 to Cα9). The two lobes are linked by the long helix Cα4. Besides the nine helices, C-SLT also contains a very irregular sheet of three small anti-parallel β-strands that is located in the "lower" lobe following helix Cα2. The active site was found in the groove of C-SLT with the catalytic residue Glu-478 located at the C-terminal end of helix Cα2.[8]

The Common Fold of C-SLT and the Three Lysozymes

The overall features of the C-domain are clearly reminiscent of the lysozyme fold. To analyze the structural similarity in more detail we pairwise superposed the Cα-backbone of C-SLT to those of SEWL, HEWL, and T4L, according to the Rossman-Argos alignment method[27] (Fig. 3). A sequence alignment based on the overall structural correspondence is shown in Figure 4. The number and root-mean-square (RMS) distances of equivalent Cα-atoms (Table I) correspond to values expected for distantly related proteins with weak or no sequence homology.[29] The majority of secondary structure elements are aligned in the superpositions, but there are also several regions where the structures diverge and the homologous α-helices and β-strands vary considerably in length. Overall, the best agreement is found with the backbone of SEWL, which has 112 out of the 185 Cα atoms in common with C-SLT, with an RMS deviation of 2.1 Å. The backbone of T4L shows the weakest overall similarity to C-SLT, with 90 Cα atoms that superimpose with an RMS distance of 3.1 Å. It should be noted, though, that the RMS differences may in part be due to the relatively low resolution of the SLT70 structure.

The best conserved part of the lysozyme fold is represented by the regions in the C-SLT backbone that are common to all three lysozymes. In the N-terminal lobe these are the "catalytic" helix Cα2, the three-stranded β-sheet as well as the linker helix Cα4; in the C-terminal lobe the best conserved regions include only the C-terminal end of helix Cα5 and part of helix Cα8 (Figs. 4, 5A). Hence, it seems that the N-terminal lobe is much more conserved in overall fold than the C-terminal lobe, which is confirmed by an individual structure alignment of the two lobes (Table II). This suggests that either the N-terminal lobe is a very stable folding unit, or that it is more involved in the basic function of the lysozyme enzymes (or both). The first possibility is not supported by studies on the in vitro folding pathways of HEWL and T4L, which indicate that the folding units in lysozyme do not correspond with the N-ter-

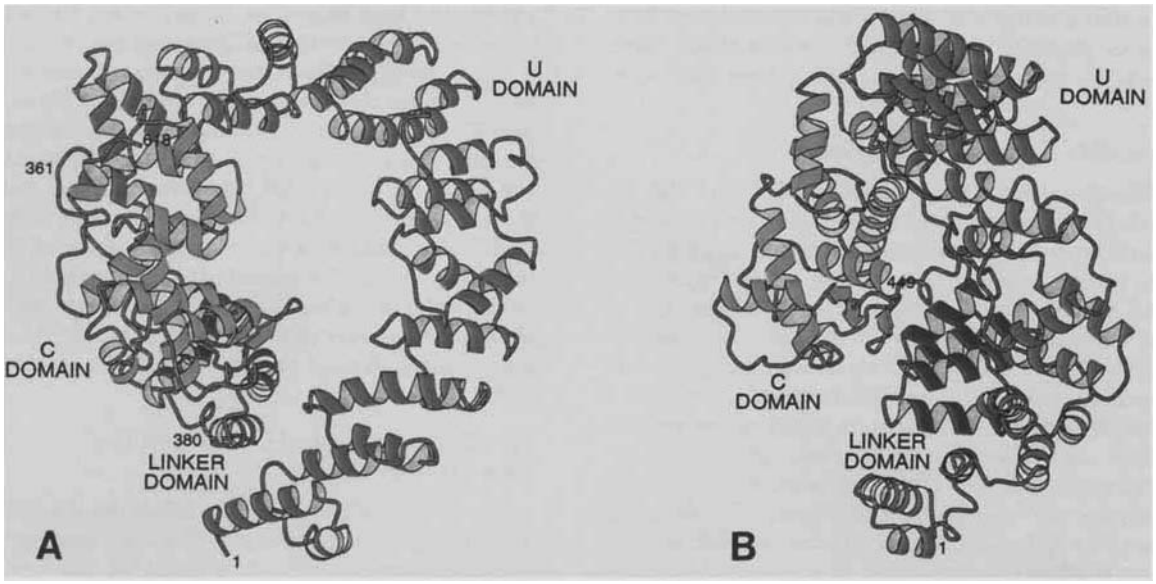


Fig. 2. Ribbon representation of the structure of SLT70. **A:** Overall view of the doughnut looking down the central hole. The doughnut is formed by the U- (residues 1–361) and linker domains (residues 380–448), with the C-terminal domain (residues 449–618) lying on top of one of the sides. **B:** The same doughnut but viewed side-on, rotated approximately 60 degrees around the vertical axis compared with the view in A. The pictures were generated using the Molscript[45] program.

minal and C-terminal lobes.[30,31] The second possibility is more likely. The best conserved regions in C-SLT are all organized around the extended groove between the two lobes that is involved in substrate binding and that contains the catalytic site (Fig. 5A). The major part of this cleft is constituted by the conserved regions located in the N-terminal domain.

Regions Outside the Common Lysozyme Fold

Outside the conserved regions the Cα-backbone of C-SLT can be divided into parts that are equivalent to only one or two of the three lysozymes and parts that are unique to SLT70. For example, in the N-terminal lobe helix Cα1 is equivalent to helix α3 of SEWL and helix α1 of HEWL, whereas it is absent in T4L. Conversely, the C-terminal lobe contains regions that are common to SEWL and T4L but not to HEWL. In fact most of the C-terminal lobe is absent in HEWL, which is clearly the simplest of the four enzymes. In C-SLT and T4L the C-terminal lobes are about equal in size. Nevertheless, the structural correspondence between C-SLT and SEWL is much better. It is dominated by the close similarity of the long helix Cα8 of C-SLT and helix α7 of SEWL, for which essentially no equivalents are present in T4L and HEWL. Based on the Rossmann-Argos alignment, only 15% of the C-SLT backbone (26 residues) has no counterpart in one or more of the lysozymes. These unique residues are mainly confined to two regions, in helix Cα3 that immediately follows the β-sheet and at the C-terminal end of the polypeptide that includes helix Cα9. Both these regions are situated at the periphery of the C-SLT domain and do not significantly contribute to the protein core (Fig. 5A). Helix Cα3 may be involved in substrate-binding regarding its location at the end of the extended groove, which in the lysozymes is known to contain the polysaccharide binding sites. Helix α2 in T4L is also located in between the β-sheet and the linker helix, but it is at a different position than Cα3 in SLT, further away from the saccharide-binding groove.

Conserved Amino Acid Residues

From our comparison it appears that C-SLT is more similar in structure to goose-type than to either chicken- or phage-type. Not only is the overall structural correspondence for the C-SLT/SEWL pair significantly higher than for the other pairs, but it is also possible to detect a small homology in amino acid sequence. Between C-SLT and SEWL there are 27 identical residues compared with 15 for C-SLT versus HEWL and 8 for C-SLT versus T4L (Fig. 4, Table I). Only two amino acids are identical in all four sequences. The first one is, not surprisingly, the "catalytic" glutamic acid. The second identical residue is a glycine located in the second turn of the β-sheet (Gly-493 in C-SLT). It has (φ,ψ)-angles not allowed for in other residues. In all proteins the "catalytic" Glu is followed by a serine, except for T4L, where this residue is a glycine. The serine is hydrogen bonded with its Oγ atom to a main chain C=O in the second turn of the β-sheet. This β-turn (residues 493–496, GLMQ in SLT70) protrudes into the

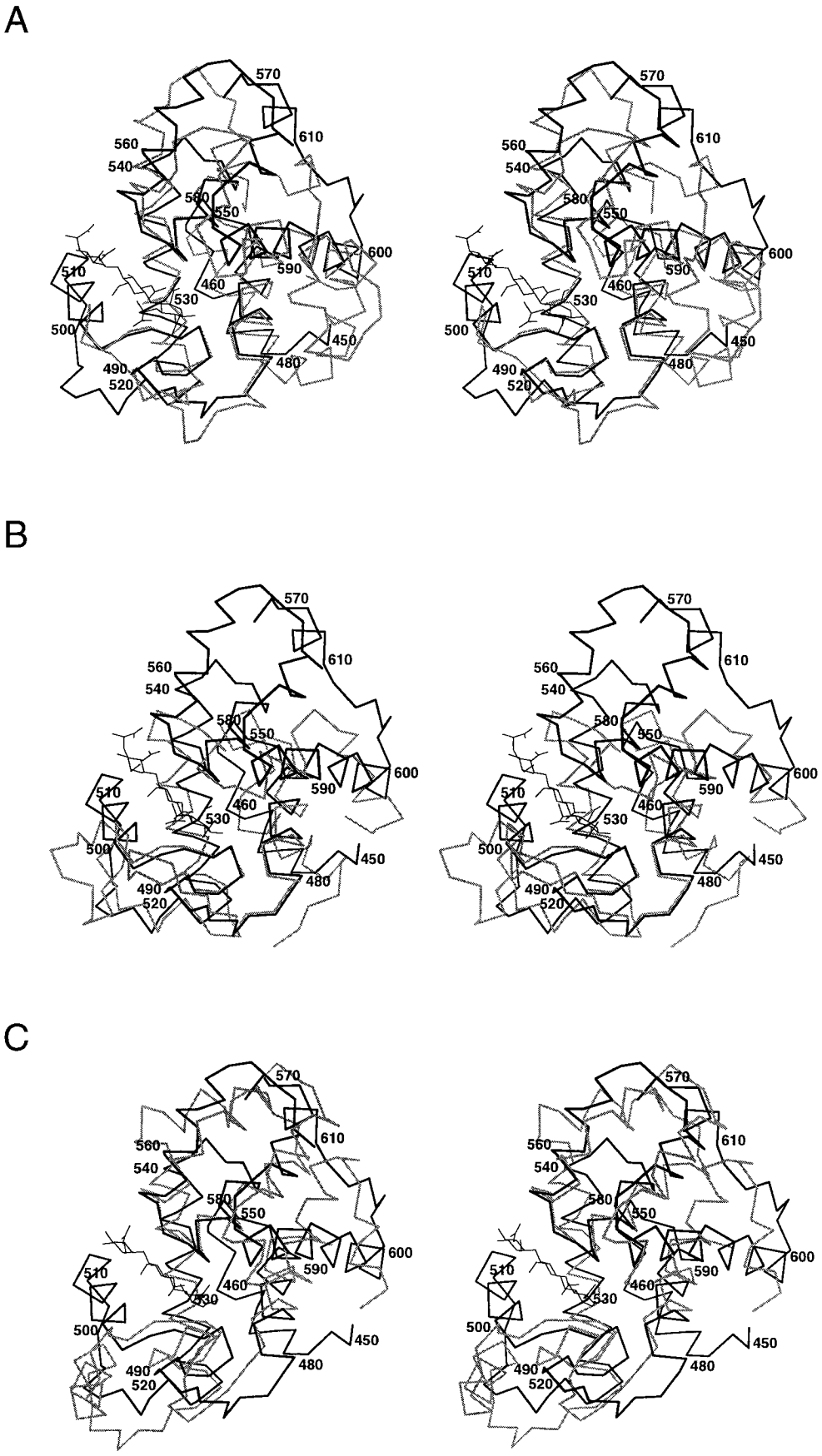


Fig. 3. Stereo views of the pairwise Cα-backbone superpositions of C-SLT to SEWL (**A**), HEWL (**B**), and T4L (**C**). Backbones of the lysozymes are in gray lines, and those of C-SLT are in black. Labels refer to residues in C-SLT. For each protein pair a trisaccharide (thin black lines), similar to that in Figure 6, is shown bound in the active site groove.

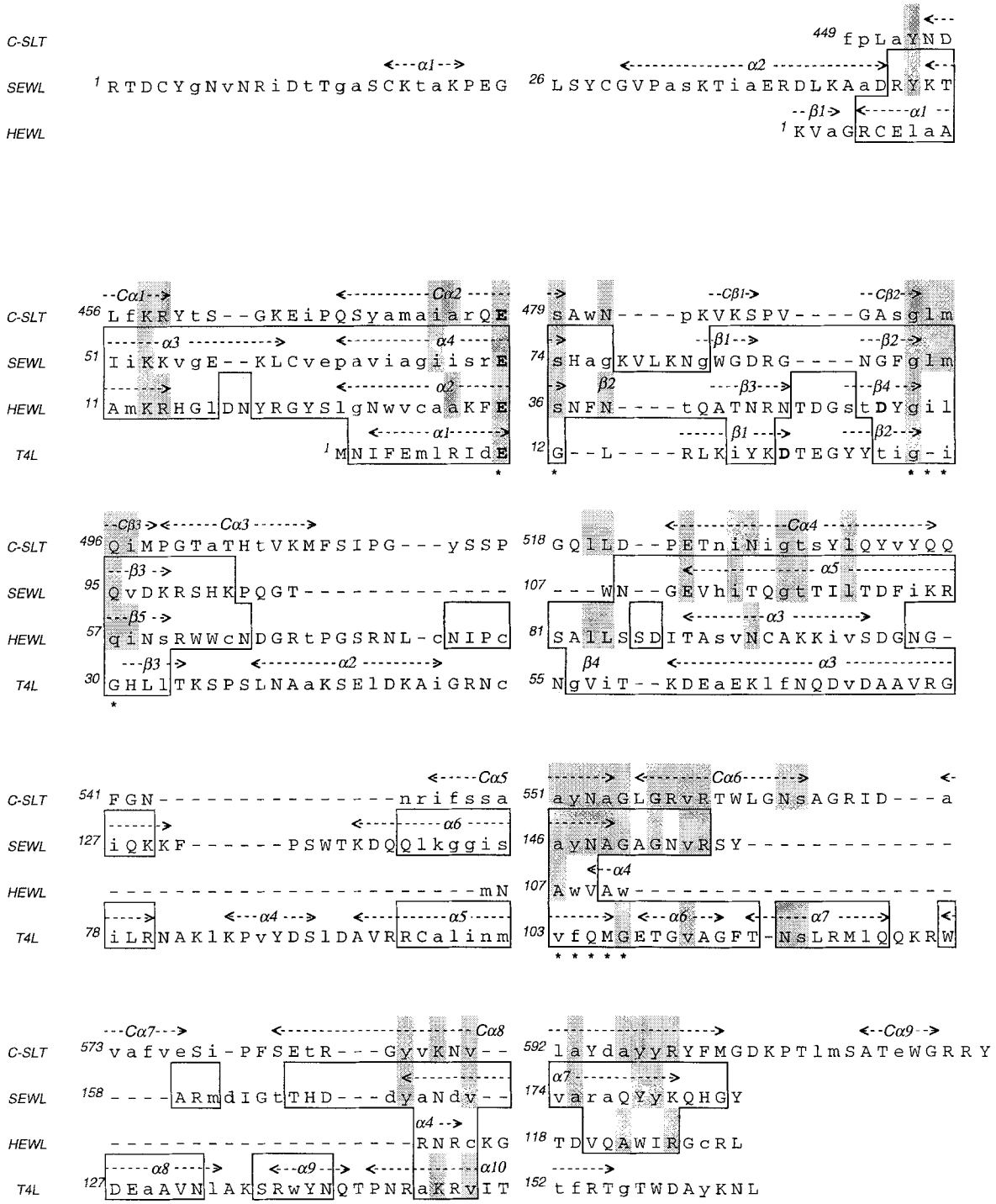


Fig. 4. Sequence comparison of C-SLT, SEWL, HEWL, and T4L based on the overall structural correspondence. The regions in the lysozymes that are topologically equivalent to C-SLT are identified by boxes; identical residues are shaded. Lower case characters point out residues that are buried. Depicted in bold face are the "catalytic" residues in the different enzymes. Asterisks indicate the three conserved motifs that are discussed in the text. The initial 40 residues in SEWL, which include helices α1 and α2, have no counterpart in C-SLT nor in the two other lysozymes. In SLT70, the N-terminus of the C-domain is connected to the linker domain, but no equivalences could be identified between this part and the N-terminal extension in SEWL. The mutual alignments of the three lysozymes are in outline very similar to that reported by Weaver et al.[7] There is a dissimilarity, though, in the alignment of the β-sheet, which in our case contains only two gaps in the first and second β-turn, respectively, whereas Weaver et al. seem to have unnecessarily broken up the β-sheet. Our result is supported by a lysozyme structure comparison study using distance plot analysis.[46]

TABLE I. Statistics of the Overall Structure Alignment of C-SLT and the Three Lysozymes

Protein pair		No. of residues		Topologically equivalent residues*					
1	2	1	2	No.	% of backbone†	RMS‡ (Å)	No. identical	No. buried	% of buried core**
C-SLT	SEWL	170	185	112	60.5	2.1	27	41	77
C-SLT	HEWL	170	129	91	70.5	2.6	15	20	76
C-SLT	T4L	170	164	90	54.9	3.1	8	24	69
SEWL	HEWL	185	129	85	65.9	2.7	13	26	76
SEWL	T4L	185	164	82	50.0	3.4	9	25	60
HEWL	T4L	129	164	61	47.2	3.3	5	16	61

*Topologically equivalent residues (common core) were identified by the Rossmann-Argos alignment method.[27]
†The percentage of the backbone of the smaller protein common to the other protein.
‡RMS, root mean square deviation in position of corresponding Cαs.
**The percentage of buried core is defined for the smaller protein as the number of buried residues in the common core divided by the total number of buried residues.

TABLE II. Statistics of the Individual Structure Alignment of the N- and C-Terminal Lobes in C-SLT and the Three Lysozymes*

Protein pair		N-terminal lobe					C-terminal lobe				
		No. of residues		Common core			No. of residues		Common core		
1	2	1	2	No.	% of backbone†	RMS‡ (Å)	1	2	No.	% of backbone†	RMS‡ (Å)
C-SLT	SEWL	91	130	61	68	1.5	79	55	34	42	1.2
C-SLT	HEWL	91	100	71	79	1.8	79	29	7	9	1.1
C-SLT	T4L	91	80	45	50	1.9	79	84	20	25	1.1

*The alignment was carried out using the least-squares fitting routine of the program "O."[28] The C-terminal end of the linker helix was choosen to divide the polypeptide chains in two parts: the C-terminal lobe starts after residue 539 in C-SLT, 130 in SEWL, 100 in HEWL, and 80 in T4L. The starting set of equivalent residues comprised the residues that were part of the core common to all four proteins identified by the Rossmann-Argos aligment (Fig. 4, residues 469–479, 495–499, 523–530, 551–553, and 588–591, following the C-SLT numbering).
†The percentage of backbone of C-SLT common to the other protein.
‡RMS, root mean square deviation in position of corresponding Cαs.

protein interior with two large hydrophobic side chains (Leu and Met in C-SLT and SEWL, Ile and Leu in HEWL), anchoring the sheet to the main protein body. The buried hydrogen bond of the serine probably serves to fix the backbone conformation at the position of the "catalytic" glutamic acid. In addition to a glycine at position one and a large hydrophobic residue at positions two and three, the second turn in the β-sheet seems also to be constrained at position four. In three enzymes there is a glutamine at that position (Fig. 4). The functional importance of this residue, however, is not completely clear. Its side chain folds back towards the glycine at position one of the turn and is hydrogen bonded to the carbonyl oxygen of that residue. This implies a structural role for the glutamine: involvement in the tightening of the β-turn. Another possibility would be that this residue is required for substrate binding.[15]

Sixteen of the identical residues between C-SLT and SEWL have their side chains buried, forming about 24% of the buried cores in both proteins. Twelve of them constitute a structural unit near the basis of the active site cleft, where they provide for the close-packing of four of the five homologous helices and the second turn of the β-sheet (Fig. 5B). The equivalent residues in the cores of the other lysozymes show a considerable variation in size, although the hydrophobic character of the side chains is mostly conserved. The overall packing of the secondary structure elements is nevertheless very similar in all four enzymes. This agrees with mutational and structural analyses on various proteins, including T4L and HEWL, which show that protein interiors can be quite tolerant to amino acid substitutions (for a review, see ref. 32).

Active Site Relationships

As defined by Blake et al.,[33] the active site groove of lysozyme contains six subsites, designated A through F, each of which can accommodate a single saccharide unit. Crystallographic binding studies with a series of mono- and oligosaccharides have resulted in a detailed description of the active site geometry and protein-saccharide binding interactions in HEWL and T4L, showing close similarities in the arrangement of key residues in substrate binding and catalysis.[4,5,21,34] Oligosaccharides bind with their non-reducing end directed towards subsite A. The scissile bond straddles subsites D and E with the two catalytic carboxyl groups disposed on either side (Glu-35/Asp-52 and Glu-11/Asp-20 in HEWL

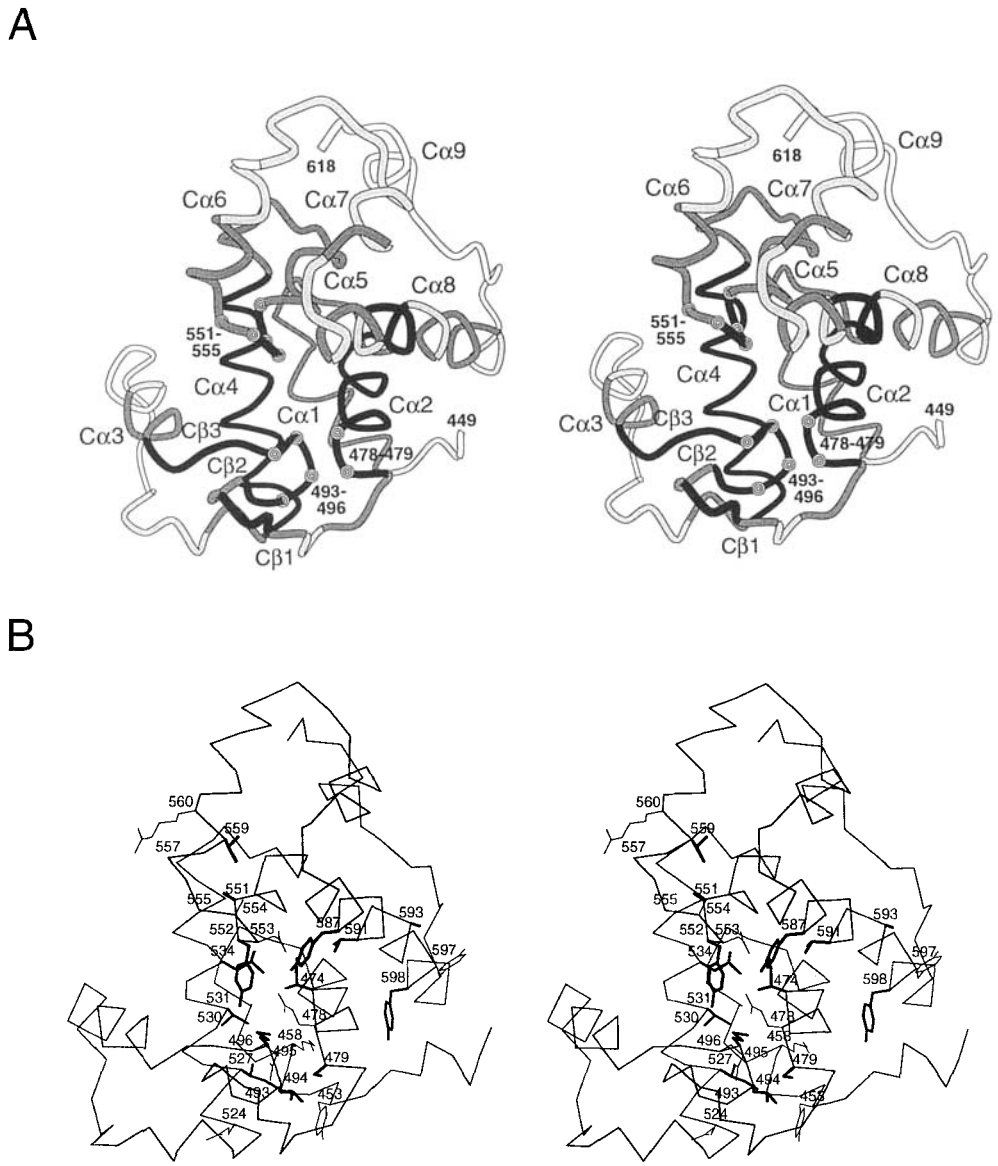


Fig. 5. The conserved features in the C-SLT structure. **A:** Stereo representation of the backbone of C-SLT drawn in different shades according to the degree of conservation. Regions that are equivalent to one or two of the lysozymes are in light and dark gray, respectively. In black are the most conserved regions that are equivalent to all three lysozymes, while the unique parts of the C-SLT fold are in white. The three conserved motifs are labeled and their residues indicated by spheres. **B:** Arrangement of the 27 residues in the C-SLT structure that are identical in SEWL. Side chains that are buried are shown in thick lines and solvent-accessible side chains in thin lines. The pictures were generated using the Molscript[45] program.

and T4L, respectively). The reaction mechanism of cleavage is commonly described in terms of general acid catalysis with the glutamic acid acting as the proton donor.[33] It is supposed that during the reaction a positive charge develops at the D sugar that is stabilized by electrostatic interactions with the negatively charged carboxylate of the aspartate. Although for HEWL this reaction mechanism seems to be essentially correct,[19] at least two different mechanisms have been proposed for T4L with the "catalytic" aspartate acting as a nucleophile to generate a covalent glycosyl-enzyme intermediate[35,36] or as a base activating a water molecule for nucleophilic attack in a single displacement reaction.[20] A further discrepancy was observed for the goose-type lysozyme, which has two aspartates (Asp-86 and Asp-97) in the β-sheet, but neither one is in a position comparable to the "catalytic" Asp-52 of HEWL or Asp-20 of T4L.[7,17,22]

To analyze the active site relationships of SLT70 and the lysozymes, we superimposed HEWL, T4L, and SEWL complexed with (different) non-produc-

tive trisaccharides on the C-domain of SLT70. As shown in Figure 6, the trisaccharides can fit in the active site of SLT70 in a manner very similar to that in the lysozymes, where they occupy the B, C, and D binding sites. Similarities in the active site regions include in particular the "catalytic" glutamic acid at subsite D and the amino acids that bind the N-acetyl group of the GlcNAc residue at subsite C. Going towards the outer ends of the active site groove, the deviations between C-SLT and the lysozymes become larger, as they do if C-SLT is compared with SEWL. In SLT70 and SEWL there is a good hydrogen bond from the OH of a tyrosine (Tyr-587 and Tyr-169 in SLT70 and SEWL, respectively) to the Oε of the glutamic acid. This interaction probably serves to stabilize the position of this key catalytic residue. A similar situation occurs in T4L, with Arg-145 forming an ion pair with Glu-11. The "catalytic" Glu-35 in HEWL is in a mostly hydrophobic environment although its carboxyl group makes a strong hydrogen bond with the backbone NH of Ala-110.[19] Subsite C appears to be the most conserved of the different saccharide binding sites. The N-acetyl group of GlcNAc is hydrogen bonded to a main chain NH and C=O group (residues 498 and 552 in SLT70), and it makes a van der Waals contact with an aromatic side chain (Tyr-552 in SLT70). The conservation of these interactions in the different enzymes underlines the importance of subsite C in determining the cleavage specificity of the enzymes. Since any substituent on O_3 of the sugar prevents binding on steric grounds, only GlcNAc, and not MurNAc, can bind to subsite C, resulting in the cleavage of the glycosidic bond after a MurNAc residue at subsite D.[33]

The most remarkable difference is observed at the β-sheet region near subsite D. The β-sheet of C-SLT does not contain an aspartate, nor any other obvious residue that could take up the catalytic role of the second carboxylate in lysozyme. The carboxylate group of Glu-583 opposite to the β-sheet is only 4–5 Å away from the presumed reaction center, but site-directed mutagenesis showed that the negative charge of this residue is not critical for catalysis.[8] It is tempting to speculate about the relation of this structural difference to the different mechanism in SLT70. In the catalytic mechanism described for HEWL, stabilization of the oxocarbonium intermediate is required to increase its lifetime. This would allow the leaving group to diffuse away and a nucleophilic water molecule to enter the active site to complete the reaction.[19,33] In the reaction catalyzed by SLT70, cleavage of the β-1,4-glycosidic bond is followed by the formation of an anhydro bond between carbons C1 and C6 of the MurNAc residue. Therefore in SLT70 it is more likely that, instead of a nucleophilic water, the intramolecular C_6 hydroxyl group attacks the MurNAc ring at the β-face of the C_1 atom. Such an intramolecular attack would probably require a shorter lifetime of an oxocarbonium intermediate, and thus less stabilization, as it is no longer dependent on the diffusion-controlled displacement of the leaving group with a water molecule. At the same time, during the reaction, SLT70 must shield the substrate at the α-face of the C_1 atom from attack by a water molecule. It could be envisaged that this is accomplished by the interaction of the hydrophobic β-sheet with the sugar ring at subsite D. Formation of 1,6-anhydro bonds by SLT70 has been unambiguously demonstrated.[9] Anhydro bond formation has not been observed for SEWL or HEWL.[37] Nevertheless, the difference in position of the aspartates in the β-sheet between SEWL and HEWL points to a difference in reaction mechanism for these latter two enzymes.[22] An interesting possibility is that hydrolysis of the glycosidic bond by goose-type lysozyme proceeds via a single displacement reaction with inversion of configuration at the anomeric carbon, whereas HEWL works as a retaining enzyme via a double-displacement mechanism. Thus, one of the aspartates in the β-sheet of SEWL might primarily be involved in the activation of a nucleophilic water molecule, instead of stabilization of the oxocarbonium intermediate.

Characteristics Defining the C-SLT Fold

The catalytic domain of SLT70 shares with the lysozymes a structural core of five secondary structure elements that are arranged in a topologically identical order around the substrate binding cleft and that comprise a three-stranded β-sheet and two α-helices in the N-terminal lobe and two small α-helical regions in the C-terminal lobe. Within this structural core three conserved motifs can be distinguished that are fundamental for the architecture of the active site and that contain the residues essential for catalysis and substrate binding. These motifs are: 1) E-S (residues 478–479 in SLT70) at the end of the "catalytic" helix; 2) G-L-M-Q (residues 493–496) in the second turn of the β-sheet; and 3) A-Y-N-A-G (residues 551–555) at subsite C (Figs. 4–6). Between SLT70 and SEWL, these motifs are invariant; in HEWL and T4L, they show a few variations in amino acid residues. The conservation of overall structure and of parts of the amino acid sequences indicates the global structural relationships of these enzymes. Nevertheless, the C-SLT shows a number of significant differences with the goose, chicken, and phage-type lysozymes. Most prominent is the absence of a "catalytic" aspartate in the β-sheet and the enhanced hydrophobicity of the β-sheet in C-SLT. Other differences relate to the arrangement and size of secondary structure elements, in particular helix Cα3 near the substrate binding groove and helix Cα9 at the C-terminus, which are both absent or in a very different orientation in SEWL, HEWL, and T4L (Figs. 3, 5).

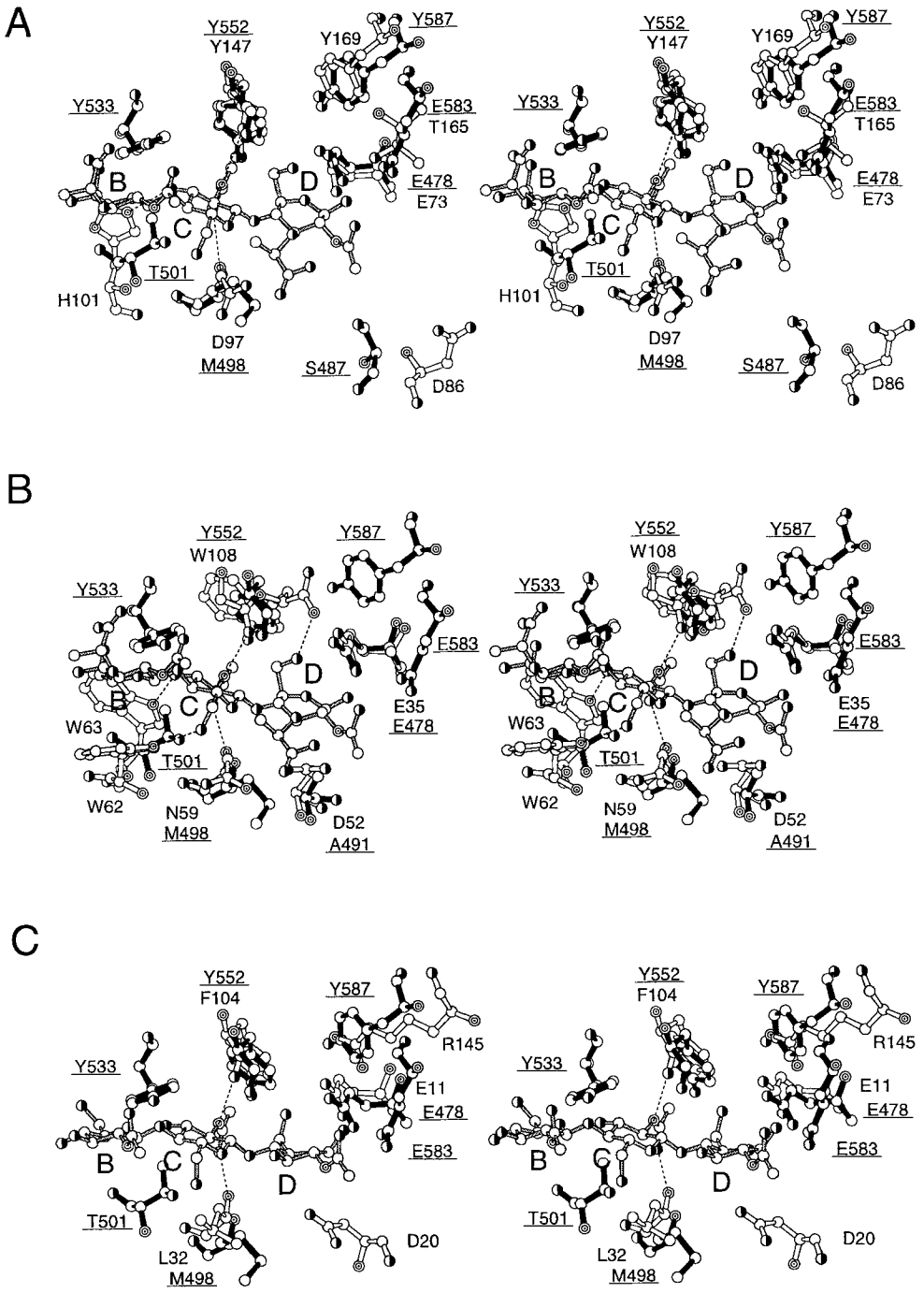


Fig. 6. Pairwise comparison of the active sites of SLT70 and the three lysozymes complexed with a trisaccharide. **A:** Stereo pair comparing the environment of MurNAc-GlcNAc-MurNAc in the B, C, and D subsites of SLT70 and SEWL. The binding mode of the modeled trisaccharide is very similar to that of $(GlcNAc)_3$ bound in goose lysozyme,[22] except for the sugar bound at subsite D, which has a normal chair conformation in the x-ray model. **B:** Similar view comparing the environment of MurNAc-GlcNAc-MurNAc in the SLT70 and HEWL. **C:** Similar view comparing the environment of $(GlcNAc)_3$ in SLT70 and T4L. Amino acids and saccharides are shown in ball-and-stick representations with residues of SLT70 drawn in solid bonds and residues of lysozyme drawn in open bonds. The superposition of the amino acid side chains are all based on the overall alignment of the α-carbon backbones (Fig. 3). The sugar ring of the saccharide bound at subsite D is in a half-chair or sofa conformation. Oxygen atoms are drawn in half-filled circles, nitrogen atoms in concentric circles, and carbon and sulphur atoms in open circles. Residue labels of SLT70 are underlined. Dashed lines indicate (possible) hydrogen bond interactions between lysozyme and the trisaccharide. The pictures were generated using the Molscript[45] program.

Structural Basis for Differences in Substrate Specificity

SLT and the lysozymes show differences not only in reaction mechanism but also in substrate specificity. Whereas all three lysozymes are endo-muramidases, SLT70 has exo-lytic activity, starting cleavage presumably from the non-reducing GlcNAc end of the glycan strands.[38] Furthermore, SLT70 will only cleave intact peptidoglycan from *E. coli* cell walls containing the peptides that cross-link adjacent oligosaccharide strands. In this respect it resembles T4L, which also requires the peptide for catalysis, preferably that from its natural host *E. coli*.[37] In contrast, chicken-type lysozyme can cleave equally well peptide-substituted and unsubstituted cell walls and even works on chitin, a linear polymer of only GlcNAc residues. Goose-type lysozyme has a specificity on linear peptidoglycan similar to that of chicken-type lysozyme, although it works somewhat better on peptide cross-linked cell walls and it can not cleave chitin.

The differences in substrate requirements indicate the presence of a specific peptide binding site that participates in the recognition of the substrate.[5] In T4L this binding site has recently been identified by x-ray crystallography and is located in the C-terminal lobe of the enzyme.[20] The extended peptide that is linked to the O_3 atom of the MurNAc at subsite D is bound in an open groove on the surface of the enzyme formed by helices α6, α7, α8, and α9 in T4L. Figure 7 shows a superposition of the analogous regions in SLT70, SEWL, and HEWL. Only SLT70 has a region that is comparable in size and conformation to the peptide-binding site in T4L. In SEWL there is a 18 residue loop at this position and HEWL completely lacks a counterpart to the peptide-binding region of T4L. The differences in structure correlate nicely with the substrate requirements of the different enzymes. The backbone similarity of SLT70 and T4L at the peptide binding region shows how these two enzymes have adapted to the same natural substrate, while choosing a different set of side chains for making the protein-peptide interactions.

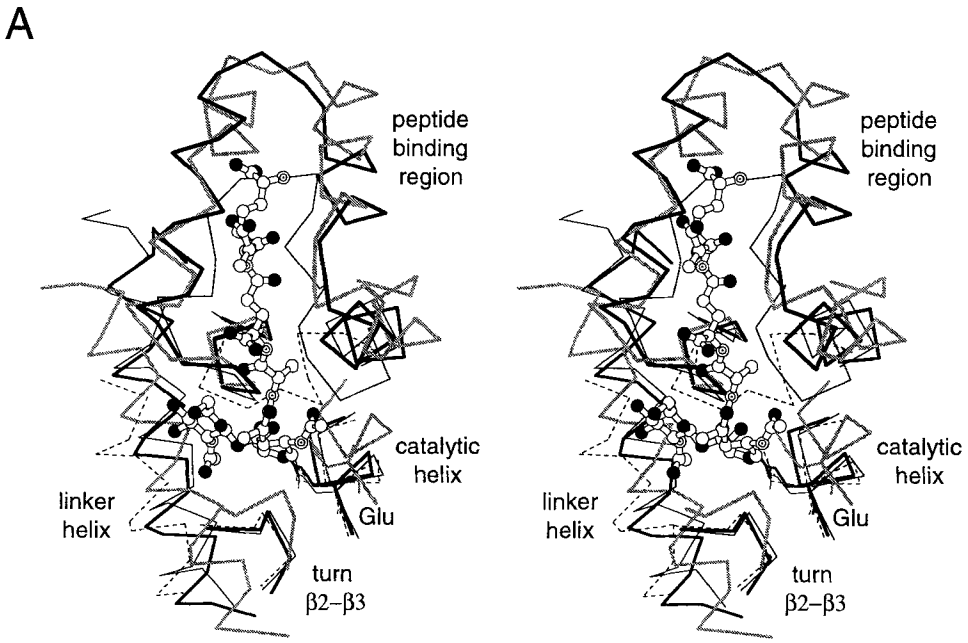
The exo-lytic activity of C-SLT is most probably imposed by the two α-helical domains in SLT70 that form the doughnut rather than by changes in the C-domain itself. A peptidoglycan strand bound in the extended groove of the C-domain is directed towards the doughnut such that it may pass through the central hole. However, due to the network-like structure of the substrate, most sites in the cell wall polymer will be inaccessible for the SLT70 doughnut, except for the loose ends of the glycan strands. The doughnut may also have an important role in the regulation of lytic activity of SLT70. This is of crucial importance for the bacterium, since uncontrolled and random cleavage of the cell wall is potentially suicidal.[39] Whether additional proteins are involved in the regulation of SLT70 activity, perhaps through interactions with the doughnut, is not known. Construction of a recombinant SLT70 that misses the ring structure may help to clarify the functional role of this domain.

Other Proteins Homologous to SLT70

Interestingly, a database search with the three conserved sequence motifs that emerged from this comparison study revealed a number of prokaryotic and phage proteins that were not detected using the complete sequences of SLT70 and the lysozymes (A.J. Dijkstra and W. Keck, unpublished data). For most of these proteins the function is still unknown, although, intriguingly, some of them are encoded by a bacterial plasmid involved in conjugational transfer, a process that almost certainly will require local cleavage of the peptidoglycan. An additional enzyme that carries the conserved motifs is the phage T7 internal protein D, which may be involved in lysis of the host bacterial cell.[40] The similarity between SLT70 and this phage protein has already been noted.[11] The relationships in amino acid sequence suggest that there exists a class of enzymes with a structure similar to the catalytic domain of SLT70. In fact, some of the proteins mentioned above were recently described as putative lytic transglycosylases, based on a conserved sequence pattern that is similar to the three motifs described in this paper.[41] However, the presence of the three conserved sequence motifs does not necessarily distinguish these proteins as lytic transglycosylases, but merely indicates that they all have a lysozyme-like fold. More significant is the finding that, like SLT70, all these proteins, except for one, miss an aspartate or glutamate residue in the region that corresponds to the β-sheet in the lysozymes. This unique feature points to the existence of a distinct class of bacterial and phage lysozymes with lytic transglycosylase activity. A definitive classification, however, awaits a complete analysis of activities, and, in the case of the 35-kDa and 38-kDa lytic transglycosylases from *E. coli,* determination of the complete amino acid sequences. At the moment, the only protein known that possesses an enzymatic activity identical to that of SLT70 and whose primary structure is known, is phage lambda endolysin.[42] The amino acid sequence of this enzyme shows similarities with both chicken- and phage-type lysozymes.[43] However, the similarity with the C-SLT amino acid sequence is low, and apart from a "catalytic" glutamic acid residue, the endolysin does not show the conserved motifs that emerge from the comparison of SLT70 with the lysozymes.

CONCLUSIONS

Our results show that the essential features of the lysozyme fold are preserved in C-SLT: a bilobal



B

peptide binding region

```
                 -----Cα5 ------>   <------ Cα6 ---->   <--- Cα7--->   <---------Cα8 ----
C-SLT  547 fssaayNaGLGRvRTWLGNsAGRID---avafvesi-PFSEtR---GyvKNv 591
                 ----- α6 ------->                                    <--- α7 ---
SEWL   143 ggisayNAGAGNvRSY----------------ARmdIGtTHD---dyaNdv 173
                 <--α4                                                α4->
HEWL   105 --mNAwVAw-------------------------------------------RNRc 115
                 ----- α5 ------->  <--- α6 --->  <------α7 ------>   <-----α8 ----->  <--α9--->  <------ α10 ---
T4L    100 linmvfQMGETGvAGFT-NsLRMlQQKRWDEaAVNlAKSRwYNQTPNRaKRv 149
                                    * *                 *   *  ◇    ◇
```

Fig. 7. Comparison of the peptide binding regions in C-SLT and the three lysozymes. **A:** Stereo pair of the superposition of a T4L-muropeptide complex and the Cα-backbones of C-SLT, SEWL, and HEWL, showing the catalytic helix, the second turn of the β-sheet, the linker helix, and part of the C-terminal lobe that is involved in peptide binding. Backbones are defined as follows: C-SLT, thick, black; T4L, thick, gray; SEWL, thin, black; HEWL, dashed, black. The compound shown in ball-and-stick representation is the muropeptide GlcNAc-MurNAc-*L*-Ala-*D*-Glu-*m*-Dpm-*D*-Ala, which was found to be covalently bound in a mutant T4L structure.[20] Atoms are as follows: oxygen, filled circles; nitrogen, concentric circles; carbon, open circles. The picture was generated using the Molscript[45] program. **B:** A sequence comparison of the peptide binding regions in the four enzymes. Residues in T4L directly involved in peptide binding are labeled with diamonds (if interacting via side chain atoms) or asterisks (if interacting via main chain atoms), based on the crystal structure.[20] Lower case characters define residues that are buried.

structure of mostly α-helices and a small irregular β-sheet of three anti-parallel β-strands, a long substrate-binding groove running across the protein surface containing the catalytic site, and a "catalytic" glutamic acid located at the C-terminal end of an α-helix in the N-terminal lobe. This suggest that SLT70 and the lysozymes have diverged from a common ancestor, since it seems highly unlikely that the similarities in overall fold and arrangement of equivalent catalytic and substrate binding residues have arisen independently. Furthermore, the structural relationships among these enzymes show that evolution was constrained mainly by the necessity to preserve the structural integrity of the extended glycan binding site and the position of the "catalytic" acid. Structural differences, like the location or absence of the second catalytic residue and the geometry of the peptide-binding site, can be related to differences in reaction mechanism and/or substrate specificity and have probably been acquired

by adaptation to specific biological functions. Two aspects of the catalytic functioning of SLT70 may be associated in particular with acquired adaptations of the lysozyme fold: regulation of the lytic activity and formation of the 1,6-anhydro bond. While all lysozymes have anti-bacterial activity, SLT70 is an endogenous bacterial enzyme important for the viability of the cell, implying that its lytic activity needs to be strictly controlled. It is likely that the two domains that form the doughnut have a function in the regulating of lytic activity of the catalytic C-domain. The function of the 1,6-anhydro bond formation by SLT70 is not clear, but it has been suggested that it may be a means of the bacterium to conserve the bond energy in the peptidoglycan for further rearrangement reactions.[9] As the differences in the active site are mainly at the level of amino acid side chains, it can be argued that SLT70 has changed from a hydrolase to a lytic transglycosylase due to evolutionary adaptation.

The structure of SLT70 reveals a new class of prokaryotic and phage lysozymes, which seem evolutionarily most closely related to goose-type lysozymes. A similar evolutionary connection with goose-type lysozymes was reported recently for barley chitinase.[44] These findings further strengthen the belief that the goose-type structure takes up a central position in the lysozyme superfamily and is evolutionarily closest to the lysozyme ancestors.[6] Furthermore, they give credence to the existence of a broad family of muramidases and chitinases with a lysozyme-like structure.

ACKNOWLEDGMENTS

We thank M.N.G. James for providing the HEWL-trisaccharide coordinates, and B.W. Matthews for the T4L-muropeptide coordinates. This research was supported by the Netherlands Foundation for Chemical Research with financial aid from the Netherlands Organization for Scientific Research.

REFERENCES

1. Chothia, C., Lesk, A.M. The evolution of protein structures. In: "Cold Spring Harbor Symposia on Quantitative Biology." Vol. 52. New York: The Cold Spring Harbor Laboratory, 1987:399–405.
2. Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J., Sussman, J.L., Verschueren, K.H.G., Goldman, A. The α/β hydrolase fold. Protein Eng. 5:197–211, 1992.
3. Jollès, P., Jollès, J. What's new in lysozyme research? Mol. Cell. Biochem. 63:165–189, 1984.
4. Matthews, B.W., Grütter, M.G., Anderson, W.F., Remington, S.J. Common precursor of lysozymes of hen egg-white and bacteriophage T4. Nature 290:334–335, 1981.
5. Matthews, B.W., Remington, S.J., Grütter, M.G., Anderson, W.F. Relation between hen egg white lysozyme and bacteriophage T4 lysozyme: Evolutionary implications. J. Mol. Biol. 147:545–558, 1981.
6. Grütter, M.G., Weaver, L.H., Matthews, B.W. Goose lysozyme structure: An evolutionary link between hen and bacteriophage lysozymes? Nature 303:828–831, 1983.
7. Weaver, L.H., Grütter, M.G., Remington, S.J., Gray, T.M., Isaacs, N.W., Matthews, B.W. Comparison of goose-type, chicken-type, and phage-type lysozymes illustrates the changes that occur in both amino acid sequence and three-dimensional structure during evolution. J. Mol. Evol. 21: 97–111, 1985.
8. Thunnissen, A.-M.W.H., Dijkstra, A., Rozeboom, H.J., Kalk, K., Engel, H., Keck, W., Dijkstra, B.W. Doughnut-shaped structure of bacterial muramidase revealed by x-ray crystallography. Nature 367:750–753, 1994.
9. Höltje, J.-V., Mirelman, D., Sharon, N., Schwarz, U. Novel type of murein transglycosylase in *Escherichia coli*. J. Bacteriol. 124:1067–1076, 1975.
10. Keck, W., Wientjes, F.B., Schwarz, U. Comparison of two hydrolytic murein transglycosylases of *Escherichia coli*. Eur. J. Biochem. 148:493–497, 1985.
11. Engel, H., Kazemier, B., Keck, W. Murein-metabolizing enzymes from *Escherichia coli:* Sequence analysis and controlled overexpression of the *slt* gene, which encodes the soluble lytic transglycosylase. J. Bacteriol. 173:6773–6782, 1991.
12. Engel, H., Smink, A.J., Wijngaarden, v.L., Keck, W. Murein-metabolizing enzymes from *Escherichia coli:* Existence of a second lytic transglycosylase. J. Bacteriol. 174: 6394–6403, 1992.
13. Romeis, T., Vollmer, W., Höltje, J.-V. Characterization of three lytic transglycosylases in *Escherichia coli*. FEMS Microbiol. 111:141–146, 1993.
14. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer Jr., E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T., Tasumi, M. The protein data bank: A computer-based archival file for macromolecular structures. J. Mol. Biol. 112:535–542, 1977.
15. Imoto, T., Johnson, L.N., North, A.C.T., Phillips, D.C., Rupley, J.A., Vertebrate lysozymes. In: "The Enzymes." Vol. 7. New York: Academic Press, 1972:666–868.
16. Weaver, L.H., Matthews, B.W. Structure of bacteriophage T4 lysozyme refined at 1.7 Å resolution. J. Mol. Biol. 193: 189–199, 1987.
17. Isaacs, N.W., Machin, K.J., Masakuni, M. Three-dimensional structure of goose-type lysozyme from the egg white of the Australian black swan *Cygnus atratus*. Aust. J. Biol. Sci. 38:13–22, 1985.
18. Rao, Z.H. Refinement of black swan goose-type lysozyme. Ph.D. thesis, University of Melbourne, 1989.
19. Strynadka, N.C.J., James, M.N.G. Lysozyme revisited: Crystallographic evidence for distortion of an N-acetylmuramic acid residue bound in site D. J. Mol. Biol. 220:440–424, 1991.
20. Kuroki, R., Weaver, L.H., Matthews, B.W. A covalent enzyme-substrate intermediate with saccharide distortion in a mutant T4 lysozyme. Science 262:2030–2033, 1993.
21. Anderson, W.F., Grütter, M.G., Remington, S.J., Weaver, L.H., Matthews, B.W. Crystallographic determination of the mode of binding of oligosaccharides to T4 bacteriophage lysozyme: Implications for the mechanism of catalysis. J. Mol. Biol. 147:523–543, 1981.
22. Weaver, L.H., Grütter, M.G., Matthews, B.W. The refined structures of goose lysozyme and its complex with a bound trisaccharide show that the "goose-type" lysozymes lack a catalytic aspartate residue J. Mol. Biol. 245:54–68, 1995.
23. Kabsch, W., Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22:2577–2637, 1983.
24. Lee, B., Richards, F.M. The interpretation of protein structures: Estimation of static accessibility. J. Mol. Biol. 55: 379–400, 1971.
25. Hubbard, T.J.P., Blundell, T.L. Comparison of solvent-inaccessible cores of homologous proteins: Definitions useful for protein modelling. Protein Eng. 1:159–171, 1987.
26. Miller, S., Janin, J., Lesk, A.M., Chothia, C. Interior and surface of monomeric proteins. J. Mol. Biol. 196:641–656, 1987.
27. Rossmann, M.G., Argos, P. Exploring structural homology of proteins. J. Mol. Biol. 105:75–96, 1976.
28. Jones, T.A., Zou, J.-Y., Cowan, S.W., Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of the errors in these models. Acta Crystallogr. A 47:110–119, 1991.
29. Chothia, C., Lesk, A.M. The relation between divergence of sequence and structure in proteins. EMBO J. 5:823–826, 1986.
30. Radford, S.E., Dobson, C.M., Evans, P.A. The folding of

hen lysozyme involves partially structured intermediates and multiple pathways. Nature 358:302–307, 1992.
31. Lu, J., Dahlquist, F.W. Detection and characterization of an early folding intermediate of T4 lysozyme using pulsed hydrogen exchange and two-dimensional NMR. Biochemistry 31:4749–4756, 1992.
32. Matthews, B.W. Structural and genetic analysis of protein folding and stability. Curr. Opin. Struct. Biol. 3:589–593, 1993.
33. Blake, C.C.F., Johnson, L.N., Mair, G.A., North, A.C.T., Phillips, D.C., Sarma, V.R. Crystallographic studies of the activity of hen egg-white lysozyme. Proc. R. Soc. Lond. [Biol.] 167:378–388, 1967.
34. Johnson, L.N., Cheetman, J., McLaughlin, P.J., Acharya, K.R., Barford, D., Phillips, D.C. Protein-oligosaccharide interactions: Lysozyme, phosphorylase, amylases. Curr. Top. Microbiol. Immunol. 139:81–134, 1988.
35. Koshland, D.E. Stereochemistry and the mechanism of enzymatic reactions. Biol. Rev. 28:416–420, 1953.
36. Hardy, L.W., Poteete, A.R. Reexamination of the role of Asp-20 in catalysis by bacteriophage T4 lysozyme. Biochemistry 30:9457–9463, 1991.
37. Schindler, M., Mirelman, D., Sharon, N. Substrate-induced evolution of lysozymes. Biochim. Biophys. Acta 482: 386–392, 1977.
38. Beachey, E.H., Keck, W., Pedro de, M.A., Schwarz, U. Exoenzymatic activity of transglycosylase isolated from *Escherichia coli*. Eur. J. Biochem. 116:355–358, 1981.
39. Shockman, G.D., Barrett, J.F. Autolytic enzymes of *Escherichia coli*. Annu. Rev. Microbiol. 37:501–527, 1983.
40. Dunn, J.J., Studier, F.W. Complete nucleotide sequence of bacteriophage T4 DNA and the locations of T7 genetic elements. J. Mol. Biol. 166:477–535, 1983.
41. Koonin, E.V., Rudd, K.E. A conserved domain in putative bacterial and bacteriophage transglycosylases. Trends Biochem. Sci. 19:106–107, 1994.
42. Taylor, A., Das, B.C., van Heijenoort, J. Bacterial cell-wall peptidoglycan fragments produced by phage λ or Vi II endolysin and containing 1,6-anhydro-N-acetylmuramic acid. Eur. J. Biochem. 53:47–54, 1975.
43. Jespers, L., Sonveaux, E., Fastrez, J. Is the bacteriophage lambda lysozyme an evolutionary link or a hybrid between C and V-type lysozymes? J. Mol. Biol. 228:529–538, 1990.
44. Holm, L., Sander, C. Structural similarity of plant chitinase and lysozymes from animals and phage. FEBS Lett. 340:129–132, 1994.
45. Kraulis, P.J. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. 24:946–950, 1991.
46. Taylor, W.R., Orengo, C.A. Protein structure alignment. J. Mol. Biol. 208:1–22, 1989.