Acta Cryst. (1996). D**52**, 105–114

Structure of a Bulgecin-Inhibited g-Type Lysozyme from the Egg White of the Australian Black Swan. A Comparison of the Binding of Bulgecin to Three Muramidases

SOLVEIG KARLSEN,[a] EDWARD HOUGH,[a] ZI H. RAO[b] AND NEIL W. ISAACS[c]*

[a]*Protein Crystallography Group, Institute of Mathematical and Physical Science, University of Tromsø, N-9037 Tromsø, Norway,* [b]*Laboratory of Molecular Biophysics and Oxford Centre for Molecular Sciences, Oxford University, Oxford OX1 3QU, England, and* [c]*Protein Crystallography, Department of Chemistry, University of Glasgow, Glasgow G12 8QQ, Scotland*

(Received 10 February 1995; accepted 22 June 1995)

Abstract

Bulgecin A, a bacterial metabolite, has been shown to bind in the active-site groove of the chicken-type lysozyme from the rainbow trout (RBTL) and in the lysozyme-like C-terminal domain, of a soluble lytic transglycosylase (C-SLT) from *Escherichia coli*. These enzymes are muramidases that cleave the glycosidic bonds in the glycan strands of the murein polymer. Here we report the crystal structure of a complex between the goose-type lysozyme from the egg white of the Australian black swan (SEWL) and bulgecin A at 2.45 Å resolution. As is the case for the C-SLT/bulgecin and RBTL/bulgecin complexes, the ligand binds with the *N*-acetylglucosamine ring in subsite *C* and the proline moiety in site *D* where it interacts with the catalytic glutamic acid. The taurine residue interacts with the β-sheet region. Comparisons of the three bulgecin complexes show that the inhibitor has the same binding mode to the muramidases with similar protein–ligand interactions, particularly for SEWL and RBTL. From our results, it seems likely that bulgecin, in general, inhibits enzymes with lysozyme-like domains and thus might represent a novel class of natural antibiotics that act on murein-degrading rather than murein-synthesizing enzymes.

1. Introduction

Bulgecins are bacterial metabolites produced by *Pseudomonas acidophila* and *mesoacidophila*, which have bulge-inducing and lysis-enhancing activity when used in combination with β-lactam antibiotics (Imada, Kintaka, Nakao & Shinagawa, 1982; Shinagawa, Kasahara, Wada, Harada & Asai, 1984; Shinagawa, Maki, Kintaka, Imada & Asai, 1985). The compounds, seen in Fig. 1, are sulfonated glycopeptides consisting of an *N*-acetylglucosamine (NAG) unit and a *b*-methoxyhydroxyproline ring linked by a β-glycosidic linkage. The differences between bulgecin A, the main component referred to as 'bulgecin', and bulgecins B and C are found in the *R* groups which are numbered 1, 2 and 3, respectively.

The integrity of the bacterial cell wall depends on the balanced action of several peptidoglycan (murein)-synthesizing and degrading enzymes (Rogers, Perkins & Ward, 1980; Höltje & Schwarz, 1985). The bacterial muramidases, like lysozymes and transglycosylases, cleave the β-1,4-glycosidic bond in the glycan strands of the murein polymer. The lysozymes hydrolyse this linkage, but the transglycosylases catalyse an additional intramolecular transfer reaction to produce small 1,6-anhydro-muropeptides (Höltje, Mirelman, Sharon & Schwarz, 1975). As the muramidases are insensitive to natural and synthetic antibiotics which inhibit the enzymes responsible for peptide crosslinks in the peptidoglycan polymer (Waxman & Strominger, 1983), they form potential targets for the development of new class of antibiotics directed towards the glycosidic bonds in murein.

Recently, the crystal structure of a soluble lytic transglycosylase (SLT70) from *E. coli* has been determined and shown to have a lysozyme-like C-terminal domain in addition to two domains which contain 27 helices

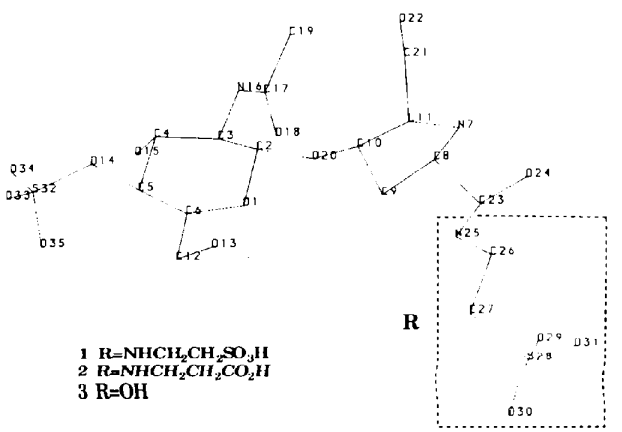

Fig. 1. Atomic numbering of a bulgecin A molecule. The structures of bulgecin A, B and C are numbered 1, 2 and 3, respectively.

† Present address: Protein Crystallography, Department of Chemistry, University of Glasgow, Glasgow G12 8QQ, Scotland.

forming a 'superhelical' ring (Thunnissen *et al.*, 1994). Limited and localized sequence homology has been detected between SLT and lysozymes (Engel, Kazemier & Keck, 1991). The active-site cleft, which was located in this domain by X-ray crystallographic analysis of a complex between SLT and bulgecin, contained similar features for substrate binding and cleavage to those in the lysozymes, although some striking differences were also observed (Thunnissen, Isaacs & Dijkstra, 1996). Structure comparisons of the C-terminal domain structure of SLT (C-SLT) with the structures of chicken-type (c-type), goose-type (g-type) and phage-type lysozymes, revealed that C-SLT is more similar in structure to g-type than to either c- or phage-type lysozyme.

Very recently, bulgecin was also found to bind to the active site of the c-type lysozyme from the rainbow trout (RBTL) (Karlsen & Hough, 1996). In both complexes, the bulgecin molecule is bound with the NAG moiety in subsite C, whereas the δ-methoxy-hydroxyproline ring is found in subsite *D* and interacts with the catalytically active residues. It seems, therefore, reasonable to believe that bulgecin acts as a competitive inhibitor for c-type lysozymes as well as transglycosylases. From kinetic studies it was known that bulgecin A inhibits SLT (Templin, Edwards & Höltje, 1992).

Since the active-site regions of the c- and g-type lysozymes show structural similarities but also differences, and the amino-acid sequences appear to be unrelated (Grütter, Weaver & Matthews, 1983; Weaver *et al.*, 1985), it was interesting to see if bulgecin also might bind to the active site of a g-type lysozyme. One catalytic residue, Glu35, in c-type lysozyme corresponds spatially to Glu73 in g-type lysozyme, but there is no clear counterpart for the active site Asp52. This lack of structural similarity has been suggested to be responsible for subtile differences in the mechanism of actions of the two types of lysozymes. c-Type lysozymes cleave peptide-substituted as well as linear peptidoglycan strands and even hydrolyse chitin, the homo-polymer of NAG. g-Type lysozymes, however, have similar specificity for linear glycan strands, but work somewhat better on peptide-substituted cell walls and cannot cleave chitin (Schindler, Mirelman & Sharon, 1977). In common with the lysozyme from the bacteriophage T4, SLT requires the intact peptidoglycan for catalysis, implying that these enzymes have specific peptide-binding sites that participate in substrate recognition (Thunnissen *et al.*, 1994; Matthews, Remington, Grütter & Anderson, 1981; Anderson, Grütter, Remington, Weaver & Matthews, 1981).

In this paper we present results from an analysis of the crystal structure of a complex between bulgecin and the g-type lysozyme from the egg white of the Australian black swan (SEWL), *Cygnus atratus*. SEWL consists of 185 amino-acid residues and closely resembles the Embden goose egg-white lysozyme (Isaacs, Machin & Masakuni, 1985; Simpson, Begg, Dorow & Morgan, 1980). We compare the binding of bulgecin to this lysozyme with binding in the active-site cleft of the c-type lysozyme from the rainbow trout and the C-domain of the soluble lytic transglycosylase (C-SLT) from *E. coli*.

2. Experimental

2.1. *Soaking, data collection and processing*

The lysozyme was purified from the egg white of Australian black swan as previously described (Masakuni, Simpson & Isaacs, 1979), and plate-like orthorhombic crystals ($P2_12_12$) were grown from phosphate buffer (pH 6.5) and 25%(*w*/*v*) NaCl by the hanging-drop technique (Rao, 1989). The inhibitor, bulgecin, was kindly provided by Takeda Chemical Industries Ltd, Japan.

The SEWL/bulgecin complex was prepared by soaking selected native SEWL crystals in mother liquid containing 6 mg ml^{-1} bulgecin for about 24 h at room temperature. The largest crystal, with dimensions approximately 1.20 × 0.60 × 0.15 mm, was slightly cracked, but

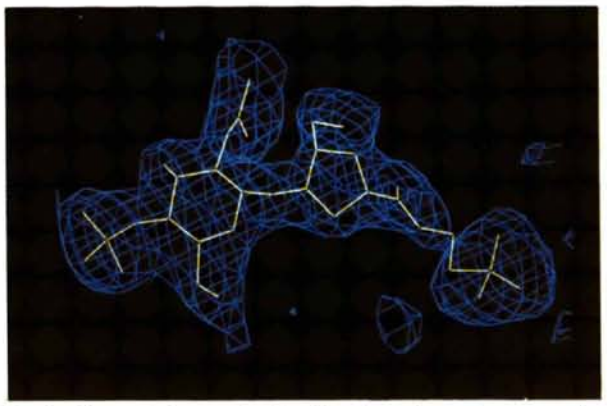

Fig. 2. Observed electron density ($F_o - F_c$ omit map), contoured at 0.12 e Å^{-3} (2σ level) for bulgecin in the active site of SEWL. Bulgecin was omitted from the coordinate file.

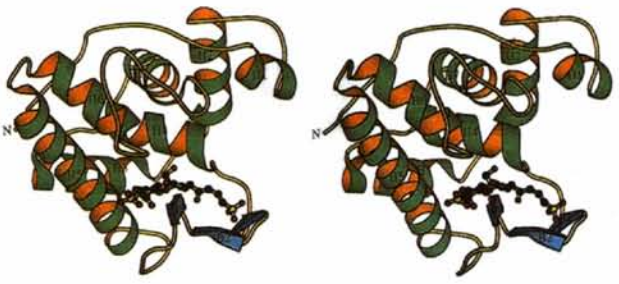

Fig. 3. A ribbon drawing generated with *MOLSCRIPT* (Kraulis, 1991) of SEWL with bulgecin in the active-site groove. The α-helices (H1–H7) are green/orange and the β-strands (B1–B3) are blue. C, O, N and S atoms in bulgecin are shown in grey, red, blue and yellow, respectively.

Table 1. *Summary of results from data collection and merging*

Space group	$P2_12_12$
Cell dimensions (Å)	$a = 91.69$, $b = 65.33$, $c = 38.15$
Resolution range (Å)	32.94–2.45
No. of unique reflections $\|F\| > 3\sigma_F$	6959
R_{merge} (%)*	5.50
Completeness (%)	78.9

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

Table 2. *Summary of final refinement statistics*

Resolution range (Å)	8.00–2.45	
R factor*	17.5	
No. of reflections ($\|F\| > 3\sigma_F$)	6959	
No. of protein atoms	1434	
No. of ligand atoms	35	
No. of solvent atoms	110	
R.m.s. deviations from ideal values		σ
Bond distances (Å)	0.016	0.02
Angle distances (Å)	0.053	0.04
Planar 1–4 distances (Å)	0.073	0.05
Chiral centers (Å^3)	0.045	0.06
Planar groups (Å)	0.014	0.02

* $R = \sum ||F_{obs}| - |F_{calc}|| / \sum |F_{obs}|$.

diffracted well to 2.45 Å. Intensity data were measured from one crystal on a Siemens/Xentronics P100 area detector at room temperature. The data set was collected using a single goniometer setting with a total rotation range of 100° in φ. Reflections were measured with a crystal-to-detector distance of 120 mm and with a width of 0.2° for each rotation image.

Data were processed using the program *XDS* (Kabsch, 1988) with further data reduction carried out using programs in the *CCP*4 program package (Collaborative Computational Project, Number 4, 1994). A summary of results from the data collection and merging is given in Table 1. In total, 6959 unique reflections were measured on 500 frames to 2.45 Å resolution, which gave a completeness of 78.9%. The R_{merge} value is 5.5%.

2.2. *Refinement*

The cell dimensions of the SEWL/bulgecin crystal showed that it was isomorphous with the native SEWL crystals. Therefore, the structure of the complex was determined by difference Fourier methods with phases from the native structure refined to a crystallographic R factor of 16.6% at 1.9 Å resolution (Rao, Ensouf, Isaacs & Stuart, 1995).

Least-squares structure refinement of the complex was carried out with the program *PROLSQ* (Hendrickson, 1985) including restraints for the bulgecin molecule. All water molecules were omitted from the active-site region of the native lysozyme before five cycles of refinement and calculations of an initial Fourier map, $|F_{obs}(\text{complex})| - |F_{calc}(\text{native})|$, were carried out. Coordinates of the bulgecin molecule, obtained from an X-ray structure determination of the desulfated form of bulgecin A (Shinagawa *et al.*, 1984) were used in the initial model of the bound inhibitor and as a template in the restraints dictionary. The bulgecin molecule was fitted into the electron density and atoms in uncertain regions with diffuse density were omitted. Cycles of least-squares refinement were followed by manual rebuilding on an Evans & Sutherland PS300 using the program *FRODO* (Jones, 1985). For the last refinement cycles, all atoms in the bulgecin molecule were refined with full occupancy. In total, 40 cycles of positional and isotropic temperature-factor refinement were carried out, giving a final crystallographic R factor of 17.5% for data between 8 and 2.45 Å resolution and including 110 water molecules.

A summary of the final refinement statistics is given in Table 2. For the protein, root-mean-square deviations from ideal geometry are 0.017, 0.057 and 0.015 Å for bond lengths, angles and planar groups, respectively. These values are comparable with the deviations obtained for the RBTL/bulgecin complex (Karlsen & Hough, 1996).

The coordinates for the structure of the SEWL/bulgecin complex have been deposited with the Protein Data Bank.*

3. Results and discussion

3.1. *Quality of the bulgecin-inhibited structure of SEWL*

A Ramachandran plot of the refined bulgecin-bound SEWL structure generated by the program *PROCHECK* (Laskowski, MacArthur, Moss & Thornton, 1993), shows that 88.5% of the residues, excluding glycines and prolines, are in the most energetically favoured regions. 11.5% (18 residues) of the remainding non-glycine residues, localized in loops, are found in additional allowed areas.

A Luzzati plot (Luzzati, 1952) for the refined SEWL/bulgecin structure indicates that the mean error in atomic position is about 0.17 Å, which is lower than for the rainbow trout lysozyme in complex with bulgecin (0.20 Å) (Karlsen & Hough, 1996).

3.2. *Binding of bulgecin to SEWL*

In addition to well defined electron density for the protein molecule, the final $2|F_o| - |F_c|$ map using phases from the refined structure of the complex between SEWL

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1LSP). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: SE0174).

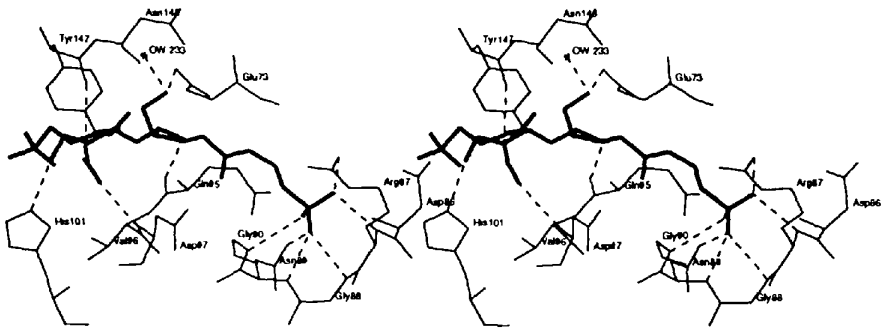

Fig. 4. Hydrogen-bonding interactions between SEWL and bulgecin. The enzyme is shown with thin lines, the ligand with thick lines, hydrogen bonds with broken lines, and the water molecule is depicted by a cross. The figure was generated using *CHAIN* (Sack, 1988).

and bulgecin, clearly shows well defined electron density for a bulgecin molecule in the active-site cleft of the lysozyme. Exceptions are seen for side chains of Lys19 ad Arg87 where the density is diffuse for the outermost atoms. In Fig. 2, the observed electron density for the atoms associated with the bulgecin molecule is contoured at 0.12 e $Å^{-3}$ (2σ level), and average temperature factors for different parts of the molecule are given in Table 3. For clarity, the atomic numbering of a bulgecin A molecule is shown in Fig. 1.

The observed electron density for the bulgecin molecule clearly illustrates that the *N*-acetylglucosamine ring of bulgecin A is bound in a position similar to subsite *C* in the active site of c-type lysozymes, whereas the proline part occupies a region analogous to subsite *D*. The taurine moiety, *R* in Fig. 1, makes hydrogen-bonded interactions with the irregular three-stranded β-sheet region which is known to be conserved in the different types of lysozyme (Grütter *et al.*, 1983; Weaver *et al.*, 1985; Jollès & Jollès, 1984). The complex is illustrated in Fig. 3.

A detailed view of hydrogen-bonding interactions between the swan lysozyme, bound water molecules and the bulgecin ligand in the active-site cleft is seen in Fig. 4, and Table 4 summarizes the protein–ligand interactions in the complex. The C3 acetamido group of the sulfated *N*-acetylglucosamine ring is buried in the protein. N16 and O18 are bound to the main-chain CO group of Tyr147 and the NH group of Asp97, respectively. This tight binding is reflected in the low temperature factor for this group (13.8 $Å^2$). The third interaction between the NAG group of bulgecin and the lysozyme, is seen between O15 and side-chain atom NE2 of His101. The C5 sulfate group does not make any interactions with the protein or bound water molecules, and, therefore, has a relatively high average temperature factor (26.4 $Å^2$).

As for the C3 acetamido group of the NAG moiety, the C11 hydroxymethyl group (C21, O22) of the proline ring makes two hydrogen bonds with SEWL and has a low average temperature factor. O22 interacts with OE1 of the catalytic residue Glu73 (Isaacs *et al.*, 1985) and ND2 of Asn148. A hydrogen bond to water molecule 233

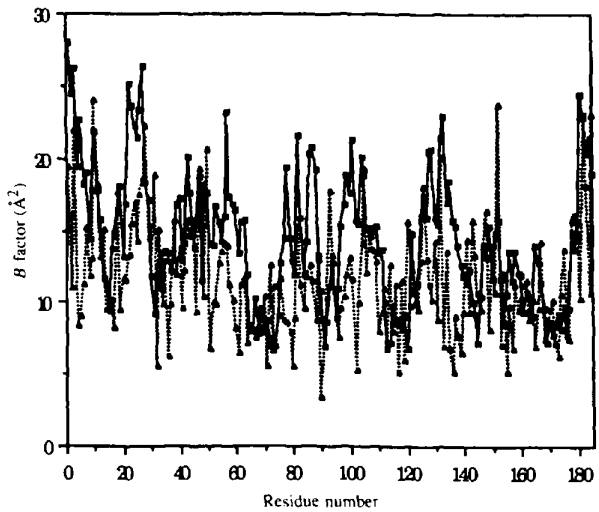

Fig. 5. Variation in isotropic temperature factors [B ($Å^2$)] averaged over main-chain atoms along the polypeptide chain for the native (Rao, Ensouf, Isaacs & Stuart, 1995) (unbroken lines) and the bulgecin-bound structure (broken lines) of SEWL.

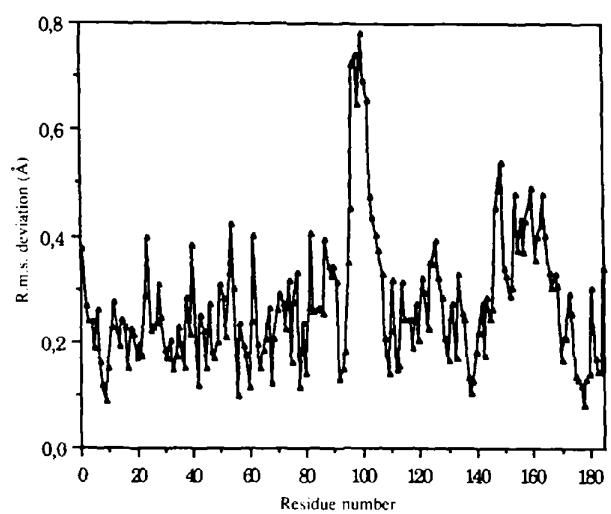

Fig. 6. R.m.s. differences along the main-chain atoms between the native (Rao *et al.*, 1995) and the bulgecin-bound structure of SEWL.

Table 3. *Average temperature factors for bulgecin bound to the active site of SEWL*

Part of the bulgecin molecule	Average temperature factor (Å^2)
N-acetylglucosamine-ring (O1, C2, C3, C4, C5, C6)	13.9
Proline-ring (N7, C8, C9, C10, C11)	10.8
C-3-acetamidoacetyl group (N16, C17, O18, C19)	13.8
Hydroxymethyl group (C12, O13)	23.7
Hydroxymethyl group (C21, O22)	13.6
Sulfate group (S32, O33, O34, O35, O14)	26.4
Peptide group (C23, O24, N25, C26)	34.1
Sulfonyl group (C27, S28, O29, O30, O31)	19.8
All atoms	19.4

Table 4. *Protein–ligand interactions in the SEWL/bulgecin complex*

Bulgecin atom	Protein atom	Distance (Å)
N7	CO Gln95	2.91
O15	NE2 His101	2.80
N16	CO Tyr147	2.83
O18	NH Asp97	2.85
O22	OE1 Glu73	2.75
O22	ND2 Asn148	3.21
O22	OW 233	2.24
O29	NH Arg87	2.83
O30	NH Asn89	2.67
O30	ND2 Asn89	2.89
O31	NH Gly90	3.11

is also seen (see Fig. 4). The sulfonyl group at the end of the taurine moiety is extensively linked to the main-chain NH groups of Arg87, Asn89 and Gly90 and the side-chain ND2 atom of Asn89. No interactions are seen between the peptide group in bulgecin and the lysozyme molecule and this part of the inhibitor has the highest temperature factor (34.1 Å^2).

Table 5. *Endocyclic torsional angles (°) of the NAG and proline ring in unliganded bulgecin and bulgecin bound to RBTL and SEWL*

Torsion angle	Unliganded bulgecin*	RBTL/bulgecin†	SEWL/bulgecin
C6—O1—C2—C3	−61	−56	−25
O1—C2—C3—C4	54	63	43
C2—C3—C4—C5	−51	−50	−54
C3—C4—C5—C6	51	30	46
C4—C5—C6—O1	−57	−23	−32
C5—C6—O1—C2	64	37	21
C10—C11—N7—C8	3	4	−5
C11—N7—C8—C9	24	25	25
N7—C8—C9—C10	−42	−41	−35
C8—C9—C10—C11	44	46	33
C9—C10—C11—N7	−30	−32	−17

* Values of desulfated bulgecin A (Shinagawa *et al.*, 1984). † Values of bulgecin bound to RBTL (Karlsen & Hough, 1996).

Endocyclic torsional angles for the NAG- and δ-methoxy-hydroxyproline ring in unliganded bulgecin and bulgecin in complex with RBTL and SEWL are detailed in Table 5. As for the bulgecin molecule bound to RBTL, the *N*-acetylglucosamine ring of the inhibitor in the active site of SEWL clearly deviates from the energetically favoured chair conformation. The C6—O1—C2—C3, C4—C5—C6—O1 and C5—C6—O1—C2 dihedral angles in particular, are considerably smaller for this sugar ring than for the analogous NAG ring in the unliganded form of bulgecin. The dihedral angles in the proline unit also show deviations from angles for the same five-membered ring in the free bulgecin molecule and bulgecin in complex with RBTL. Comparisons of the bond distances of unliganded bulgecin with the liganded forms of the molecule show that only small changes have occurred upon binding to the active-site cleft.

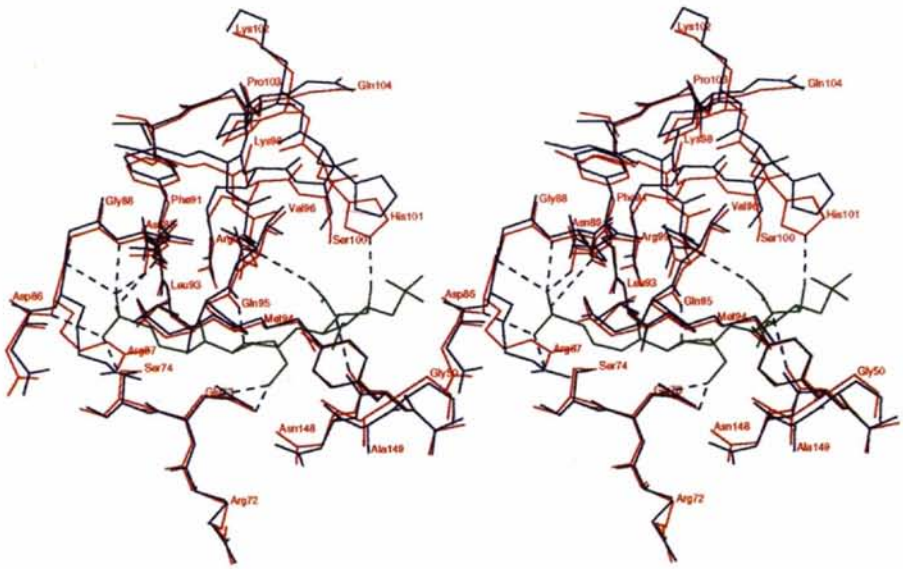

Fig. 7. Conformational changes within the active-site cleft of SEWL on the binding of bulgecin. Residues in the ligand-bound structure (red) are shown superimposed with those in the native protein (blue). The bulgecin molecule is shown in green. The figure was generated using *CHAIN* (Sack, 1988).

3.3. *Changes in temperature factors and conformational changes in SEWL*

Average temperature factors for protein atoms in native SEWL (Rao *et al.*, 1995) and the enzyme in complex with bulgecin are 16.7 and 13.0 Å^2, respectively, and the corresponding values if only main-chain atoms are included are 14.3 and 11.7 Å^2. These are considerably lower than for native and bulgecin-liganded rainbow trout lysozyme (26.3 and 27.2 Å^2 for all protein atoms and 22.0 and 21.9 Å^2 for only main-chain atoms, respectively). This applies also to the bulgecin molecule which has a higher temperature factor when bound to RBTL (30.0 Å^2) than to SEWL (19.4 Å^2). The drop in mean temperature factors for the bulgecin-inhibited structure of SEWL indicates that parts of the lysozyme become more ordered when the ligand binds to the catalytic site.

Fig. 5 shows a temperature-factor plot for the polypeptide backbone in native swan lysozyme (Rao *et al.*, 1995) and in the bulgecin-inhibited structure. R.m.s. deviations along the main-chain for the two lysozyme molecules are illustrated in Fig. 6. The overall r.m.s. differences between the structures are 0.30 and 0.43 Å for main-chain atoms and all protein atoms, respectively. Fig. 6 emphasizes that a specific loop region from residues 97 to 106 between the β-sheet and the fifth helix exhibits the greatest conformational differences. In this area Asp97 and His101 are directly involved in the binding of bulgecin (see Table 4). This is also seen in Fig. 7 which illustrates a superimposition of residues lining the active-site clefts of the native and the bulgecin-bound structure of SEWL. In addition, the temperature-factor plot (see Fig. 5) shows that flexibility is reduced in this area (residues 96 to 102). Reduced flexibility and some conformational changes are also seen for the hairpin loop (residue 85 to 89) between the first and second β-strand where Arg87 and Asn89 are hydrogen bonded to the inhibitor. These results imply that residues which line an active-site cleft and participate in inhibitor or substrate binding, often are found in flexible loop regions and are thereby freer to move towards an incoming ligand than are amino acids in α-helices and β-strands.

The second region in the lysozyme molecule that shows a significant change in position upon ligand binding, is found from residue 147 in the middle of the sixth helix (H6) to residue 165 of the beginning of the seventh α-helix (H7). In this area Tyr147 and Asn148 make hydrogen bonds with the bound bulgecin molecule which presumably move the α-helix closer to the inhibitor. A similar movement is seen for the fourth helix (residues 107–114) in the rainbow trout lysozyme when chitin oligosaccharides (Karlsen & Hough, 1995) and bulgecin (Karlsen & Hough, 1996) bind to the active-site groove. Conformational changes seen for the loop region between H6 and H7, are probably caused by high flexibility.

The *B*-factor plot (Fig. 5) emphasizes that the loops between H1 and H2, H3 and H4, H4 and the β-sheet, and finally between H5 and H6, are less flexible in the liganded SEWL structure.

3.4. *Similarities in the architecture of the active-site cleft of C-SLT, c-type and g-type lysozymes*

The structure determination of the soluble lytic transglycosylase from *E. coli* revealed that the C-terminal domain (C-SLT) clearly resembles the well characterized lysozyme fold (Thunnissen *et al.*, 1994). As for the c-, g- and T4 lysozymes, this ellipsoidal domain is built up almost entirely of α-helices (nine helices) and is divided by a deep cleft into two lobes, an N- and a C terminal. This enzyme also contains the small irregular sheet of three small antiparallel β-strands which is found in all lysozyme types (Grütter *et al.*, 1983; Weaver *et al.*, 1985).

Structural alignments of C-SLT and SEWL, HEWL and T4L (Thunnissen *et al.*, 1996), have shown that the best conserved parts are found in regions that constitute the central core of the enzymes and make up the extended groove that is involved in substrate binding. The overall fold is best conserved in the N-terminal lobe which contributes most of the lining of the active site. This includes the 'catalytic' glutamic acid, the three-stranded β-sheet and the linker helix which connects the two lobes (helix 4, 6 and 3 in SLT, SEWL and HEWL, respectively). In the C-terminal lobe parts of the fifth helix in SLT, the sixth helix in the g-type and the fourth helix in the c-type lysozyme contribute to the active cleft. In addition, regions in the eighth α-helix in the transglycosylase, the final helix (the seventh) in SEWL and the loop following the final helix (the fourth) in HEWL, contribute to the structural core.

Fig. 8 shows a sequence comparison based on the overall structural correspondence between C-SLT, RBTL and SEWL for residues around the active-site cleft that are essential for substrate binding and catalysis. The figure emphasizes that within the structural core there are three conserved regions. These are at the end of the 'catalytic' helix, in the second hairpin loop of the β-sheet and at subsite *C*, where the amino-acid sequence between C-SLT and SEWL is invariant. However, greater variation in primary structure is seen for the rainbow trout lysozyme, indicating that from an evolutionary point of view, the soluble lytic transglycosylase is closer to g-type than to c-type lysozymes. Furthermore, structural alignments (Thunnissen *et al.*, 1996) show that the overall structural correspondence is better between C-SLT and g-type than for c- or phage-type lysozymes.

The first conserved motif contains the 'catalytic' glutamic acid at subsite *D*, followed by an invariant serine residue which is believed to be important for the backbone conformation at the catalytic residue. The conserved serine makes a hydrogen bond to the second

conserved motif in the active-site groove, that forming the second turn of the β-sheet. This interaction, in addition to the two large hydrophobic side chains at positions two and three in the sheet (Leu and Met in SEWL and C-SLT, Ile and Phe in RBTL) which protrude into the interior of the protein, seems to anchor the sheet to the main body of the enzyme (Thunnissen *et al.*, 1996). A hydrogen bond between the invariant residues in the first (Gly) and the fourth (Gln) position, is likely to tighten the β-turn. The last region with almost identical primary structure (invariant for C-SLT and SEWL) is found at subsite *C* which appears to be the most conserved of the different saccharide-binding sites.

Although the 'catalytic' glutamic acid is found in structurally corresponding positions in the three enzymes, the second 'catalytic' residue (Asp52 in RBTL) which is believed to play an important role during lysozyme catalysis (Phillips, 1967; Blake *et al.*, 1967), lacks a clear counterpart in SEWL and C-SLT. The swan lysozyme has two candidates in the β-sheet (Asp86 and Asp97), but neither is in a position comparable to the aspartate in RBTL. The 'catalytic' aspartic acid is obviously absent in the C-domain of SLT, suggesting a different mechanism for this enzyme (Thunnissen *et al.*, 1996).

3.5. *Comparisons of the binding of bulgecin to RBTL, SEWL and SLT*

Atomic coordinates of the complex between the soluble transglycosylase from *E. coli* and bulgecin were kindly provided by Dr Andy-Mark W. H. Thunnissen, Groningen, the Netherlands, enabling us to compare bulgecin binding in three different muramidases. As the structure of this complex only is determined to 2.8 Å resolution, a detailed comparison with the other two bulgecin complexes (RBTL/bulgecin and SEWL/bulgecin solved to 2.0 and 2.45 Å resolution, respectively) is difficult and conclusions must be drawn with caution.

Table 6 gives a survey of interactions between bulgecin and the three enzymes (water molecules are excluded) and the interacting residues are also underlined in Fig. 8. The bulgecin molecule is bound in the same way to the three muramidases. In all three complexes, the inhibitor is bound with the *N*-acetylglucosamine unit in subsite C, whereas the δ-methoxy-hydroxyproline ring is found in the *D* site. The taurine moiety interacts with the β-sheet in the c-and g-type lysozymes, but is not hydrogen bonded to C-SLT. In general, there are clearly fewer interactions between bulgecin and the protein in the C-SLT complex, suggesting that the inhibitor might be more loosely bound to this enzyme than to the lysozymes.

Table 6. *Protein–ligand interactions in the complexes between bulgecin and RBTL, SEWL and SLT*

Bulgecin	RBTL	SEWL	SLT
N7	OD2 Asp52	CO Gln95	—
O15	NE1 Trp63	NE2 His191	OG1 Thr501
N16	CO Ala107	CO Tyr147	—
O18	NH Asn59	NH Asp97	NH Met498
O22	OE2 Glu35	OE1 Glu73	—
O22	NH Val109	ND2 Asn148	—
O24	—	—	OG Ser487
O29	—	NH Arg87	—
O30	NH Thr47	NH Asn89	—
O30	OG1 Thr47	ND2 Asn89	—
O31	ND2 Asn46	NH Gly90	—
O33	—	—	OH Tyr533

The three complexes were superimposed using the least-squares structure comparison routine in the program *O* (Jones, Zou, Cowan & Kjeldgaard, 1991). The three structurally conserved regions in the active-site clefts (see Fig. 8) and the bulgecin molecules were defined as equivalents before the routine automatically optimized the superpositions. Figs. 9(*a*), 8(*b*) and 9(*c*) which give overviews of the superimposed complexes, illustrate that differences in binding mode of bulgecin are mainly found for the taurine moiety which has different orientations in the complexes. More detailed pictures of protein–bulgecin interactions are shown in Figs. 10(*a*), 10(*b*) and 10(*c*). The observation that the

RBTL	Trp	Glu	Ser	Asn	Arg	Asn	Thr	Thr	Asp	Gly	Ile	Phe	Gln	Asn	Tyr	Trp	Val	Asp	Asn	Ala	Trp	Val	Ala	Trp
	34	35	36	44	45	46	47	51	52	54	55	56	57	59	62	63	98	101	103	107	108	109	110	111
	H2	H2	H2	B3	B3	B3	L	B4	B4	B4	L	L	B5	B5	L	L	H3	L	L	L	L	L	H4	H4
SEWL	Arg	Glu	Ser	Asp	Arg	Gly	-	Asn	Gly	Gly	Leu	Met	Gln	Asp	Ser	His	Leu	Asp	Ile	Ala	Tyr	Asn	Ala	Gly
	72	73	74	86	87	88		89	90	92	93	94	95	97	100	101	119	122	124	146	147	148	149	150
	H4	H4	H4	B1	L	L		B2	B2	B2	L	L	B3	B3	L	L	H5	H5	H5	H6	H6	H6	H6	L
SLT	Gln	Glu	Ser	Ser	Pro	Val	-	Gly	Ala	Gly	Leu	Met	Gln	Met	Thr	Ala	Tyr	Tyr	Tyr	Ala	Tyr	Asn	Ala	Gly
	477	478	479	487	488	489		490	491	493	494	495	496	498	501	502	533	536	538	551	552	553	554	555
	H2	H2	H2	B1	L	L		B2	B2	B2	L	L	B3	B3	H3	H3	H4	H4	H4	H5	H5	H5	H5	L

Fig. 8. Sequence comparison of the active sites of RBTL, SEWL and SLT based on the overall structural correspondence. The catalytic residues are marked in bold and the residues that interact directly with bulgecin are underlined. Identical residues are depicted in shaded boxes and regions with α-helix, β-sheet and loop are marked with H, B and L, respectively, and numbered as in Thunnissen *et al.* (1994).

glucosaminyl moiety in bulgecin binds to subsite *C* in all complexes, clearly underlines how important this site is for substrate binding and enzyme specificity. Two conserved interactions are seen between the *N*-acetyl group of the NAG residue and the c- and g-type lysozymes. N16 and O18 are bound to the main-chain CO group of Ala107 in RBTL and Tyr147 in SEWL and the NH group of Asn59 in RBTL and Asp97 in SEWL, respectively. The latter interaction is also found in the C-SLT/bulgecin complex where O18 interacts with the backbone NH group of Met498. Another conserved interaction between the three enzymes and the NAG unit of bulgecin, is seen between O15 of the inhibitor and NE1 of Trp63, NE2 of His101 in SEWL and OG1 of Thr501 in SLT. The only hydrogen bond from the sulfate group of bulgecin and a protein molecule is found in the C-SLT/bulgecin complex where O33 interacts with OH of Tyr533.

In the RBTL/bulgecin complex, N7 in the proline ring is hydrogen bonded to OD2 of the 'catalytic' Asp52. The equivalent interaction in the SEWL/bulgecin complex is to the conserved glutamine residue in the second bend of the β-sheet. This implies that this residue, besides being important in the tightening of the β-turn (Thunnissen *et al.*, 1996), may also be required for substrate binding in some murein hydrolases.

The hydroxyl group on the proline ring of bulgecin is hydrogen bonded in the same way to RBTL as to SEWL. In both complexes the 'catalytic' glutamic acid interacts with O22, and a hydrogen bond to the same inhibitor atom is seen from the main-chain NH group of Val109 in RBTL and the ND2 atom of the structural corresponding Asn148 in SEWL (see Fig. 10*a*). Similar interactions are absent in the C-SLT/bulgecin complex. This complex has, however, a unique hydrogen bond between the O24 atom of peptide group and OG in the side chain of Ser487.

The sulfonyl group in the taurine moiety in bulgecin is extensively linked to the rainbow trout and swan lysozymes, but is not bound to the transglycosylase. In both lysozyme complexes this group makes hydrogen-bonded interactions with the turn between the first and second β-strand, supporting the idea of a 'left-handed' binding mode (Banerjee, Holler, Hess & Rupley, 1975) to the lower part (subsites *E* and *F*) of the active-site cleft.

Fig. 4 shows that the two possible 'catalytic' aspartic acid residues (Asp86 and Asp97) in SEWL are too distant from subsite *D* to have any stabilizing effect on the positively charged oxocarbonium ion that is supposed to develope during lysozyme catalysis (Phillips, 1967; Blake *et al.*, 1967). As mentioned above, the main-chain NH group of Asp97 interacts with the *N*-acetyl group of the NAG ring bound to subsite *C*, but its side chain is extensively hydrogen bonded to Arg99. OD2 interacts with both the backbone NH group of Arg99 and the NE atom of its side chain, and the OD1 atom interacts with NH2 of the same side chain. The distance between OD1 of this aspartic acid and C8 in the proline ring of bulgecin is 4.6 Å, too long for any significant interaction. The second aspartic acid (Asp86) which is solvent exposed and does not make any hydrogen-

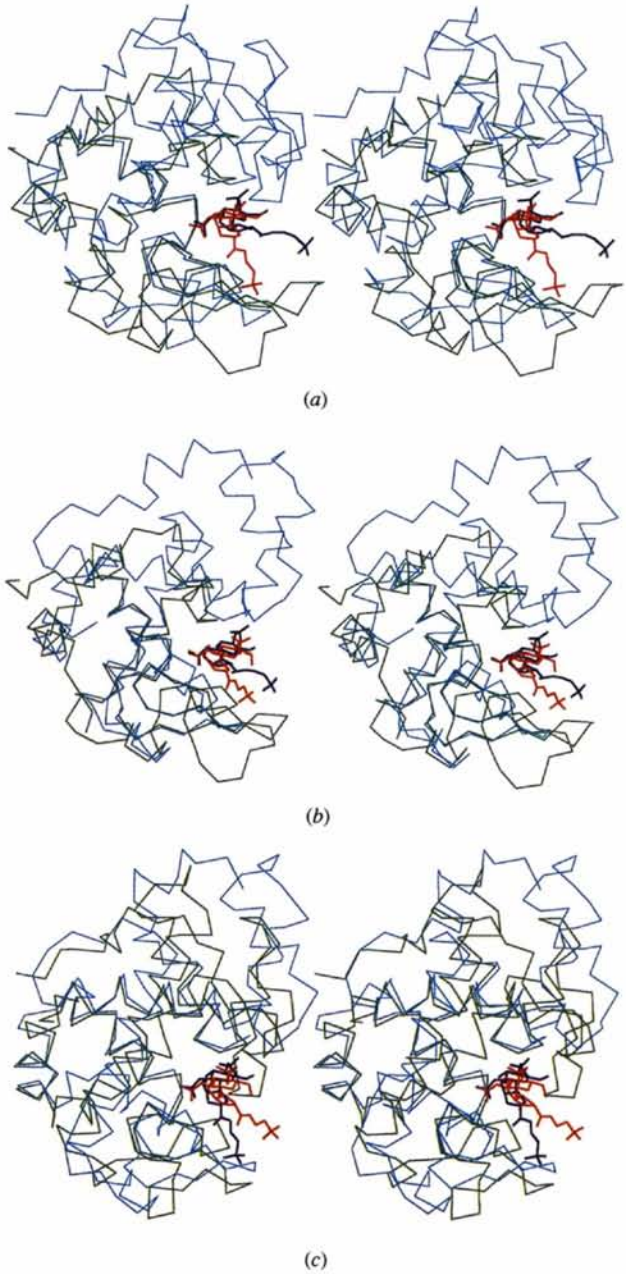

Fig. 9. Superimpositions of the bulgecin molecules in (*a*) RBTL (blue) and SEWL (red), (*b*) RBTL (blue) and SLT (red) and (*c*) SEWL (blue) and SLT (red). The Cα-trace of RBTL is green in (*a*) and (*b*), SLT is cyan in (*b*) and (*c*), and SEWL is cyan and green in (*a*) and (*c*), respectively. The figures were generated using *CHAIN* (Sack, 1988).

bonding interactions within the protein molecule or with the bound bulgecin molecule, is even further from the proline ring (10.8 Å from Cα of Asp86 to C8 of the proline ring). It has been suggested that one of the aspartates in the β-sheet of SEWL might be involved primarily in the activation of a nucleophilic water, rather than in stabilization of the oxocarbonium intermediate (Thunnissen *et al.*, 1996).

Our analysis of the binding of bulgecin to the three muramidases, shows that there are more conserved in-

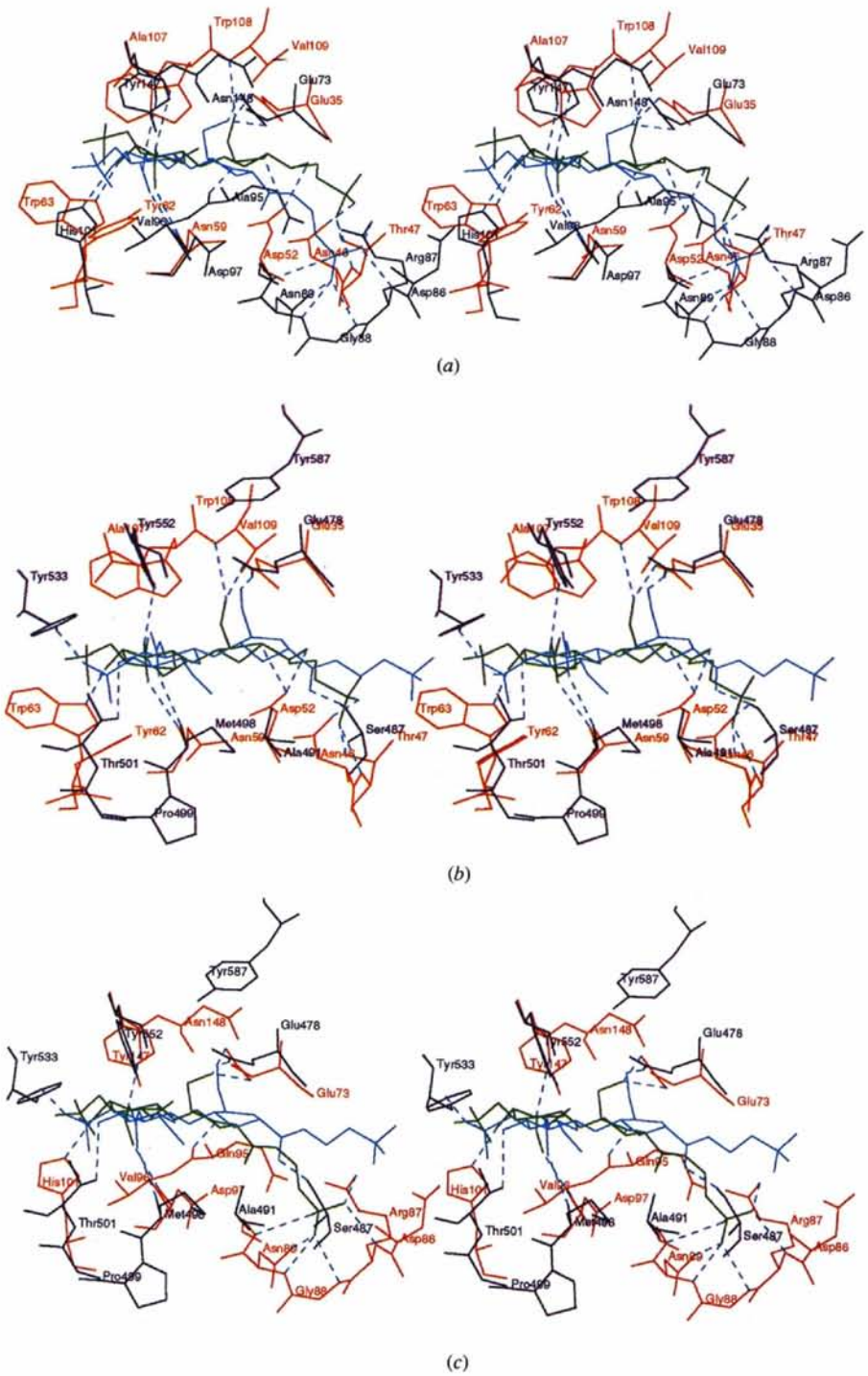

Fig. 10. Superimpositions of the bulgecin molecules in the active sites of (*a*) RBTL (green) and SEWL (cyan), (*b*) RBTL (green) and SLT (cyan) and (*c*) SEWL (cyan) and SLT (green). RBTL is red in (*a*) and (*b*), SLT is blue in (*b*) and (*c*), and SEWL is blue and red in (*a*) and (*c*), respectively. The figures were generated using *CHAIN* (Sack, 1988).

teractions between bulgecin and RBTL and SEWL than between bulgecin and the C-domain of SLT. However, the amino-acid sequences in regions that are important for the architecture of the substrate-binding groove seem to be more homologous for SEWL and C-SLT than for SEWL and RBTL. Fig. 8 shows that many of the interacting residues, particularly for C-SLT, are outside these most conserved motifs, and it seems likely that the bulgecin molecule has somewhat different binding mode in this enzyme than in the others. As mentioned above, the δ-methoxy-hydroxyproline-ring in subsite *D* in the SLT/bulgecin complex does not form the hydrogen-bond with the 'catalytic' glutamic acid which are observed for the other complexes. In addition, the 'catalytic' aspartic acid which is absent in the lytic transglycosylase, indicates that this enzyme has another mechanism for stabilising the positively charged reaction intermediate. In RBTL, Asp52 interacts directly with the proline ring in site *D* where it can easily stabilise an intermediate in order to lengthen its lifetime. The intramolecular attack by the C6 hydroxymethyl group on C1 which has been proposed for C-SLT (Thunnissen *et al.*, 1994), probably occurs very rapidly and is thus less demanding on the lifetime of the oxocarbonium ion. The intermediate might be shielded by the hydrophobic β-sheet from attack by a water molecule.

Structural differences in the active-site cleft of RBTL, SEWL and C-SLT, such as the presence or absence of the second catalytic residue and small differences in the binding mode of bulgecin, can be related to differences in substrate specificity and reaction mechanism. However, as the bulgecin molecule binds with the glucosaminyl group in subsite C and the proline part close to the 'catalytic' glutamic acid in the three enzymes, there is good reason to believe that bulgecin is, in general, an inhibitor for all muramidases.

This work was supported by the Norwegian Research Council (NFR) and the European Molecular and Biology Organisation (EMBO, short-term fellowship for S. Karlsen), and we thank Takeda Industries Ltd, Japan, for providing us with the bulgecin compound and the atomic coordinates. We would also like to thank Dr Andy-Mark W. H. Thunnissen, Groningen, the Netherlands, for providing the coordinates of the C-SLT/bulgecin complex.

References

Anderson, W. F., Grütter, M. G., Remington, S. J., Weaver, L. H. & Matthews, B. W. (1981). *J. Mol. Biol.* **147**, 523–543.

Banerjee, S. K., Holler, E., Hess, G. P. & Rupley, J. A. (1975). *J. Biol. Chem.* **250**, 4355–4367.

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

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Engel, H., Kazemier, B. & Keck, W. (1991). *J. Bacteriol.* **173**, 6773–6782.

Grütter, M. G., Weaver, L. H. & Matthews, B. W. (1983). *Nature (London)*, **303**, 828–831.

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

Höltje, J.-V., Mirelman, D., Sharon, N. & Schwarz, U. (1975). *J. Bacteriol.* **124**, 1067–1076.

Höltje, J.-V. & Schwarz, U. (1985). *Molecular Cytology of Escherichia coli*, pp. 77–119. London: Academic Press.

Imada, A., Kintaka, K., Nakao, M. & Shinagawa, S. (1982). *J. Antibiot.* **35**, 1400–1403.

Isaacs, N. W., Machin, K. J. & Masakuni, M. (1985). *J. Biol. Sci.* **38**, 13–22.

Jollès, P. & Jollès, J. (1984). *J. Mol. Cell. Biochem.* **63**, 165–189.

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

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

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

Karlsen, S. & Hough, E. (1995). *Acta Cryst.* D**51**, 962–978.

Karlsen, S. & Hough, E. (1996). *Acta Cryst.* D**52**. 115–123.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

Laskowski, R. A., MacArthur, M. W., Moss, D. S & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

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

Masakuni, M., Simpson, R. J. & Isaacs, N. W. (1979). *J. Mol. Biol.* **135**, 313–314.

Matthews, B. W., Remington, S. J., Grütter, M. G. & Anderson, W. F. (1981). *J. Mol. Biol.* **147**, 545–558.

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

Rao, Z. H. (1989). PhD thesis, University of Melbourne, Australia.

Rao, Z. H., Ensouf, R., Isaacs, N. & Stuart, D. (1995). *Acta Cryst.* D**51**, 331–336.

Rogers, H. J., Perkins, H. R. & Ward, J. B. (1980). *Microbial Cell Walls and Membranes*, pp. 190–214. London: Chapman and Hall.

Sack, J. S. (1988). *J. Mol. Graphics*, **6**, 244–245.

Schindler, M., Mirelman, D. & Sharon, N. (1977). *Biochim. Biophys. Acta*, **482**, 386–392.

Shinagawa, S., Kasahara, F., Wada, Y., Harada, S. & Asai, M. (1984). *Tetrahedron*, **40**, 3465–3470.

Shinagawa, S., Maki, M., Kintaka, K., Imada, A. & Asai, M. (1985). *J. Antibiot.* **38**, 17–23.

Simpson, R. J., Begg, G. S., Dorow, D. S. & Morgan, F. J. (1980). *Biochemistry*, **19**, 1814–1819.

Templin, M. F., Edwards, D. H. & Höltje, J.-V. (1992). *J. Biol. Chem.* **267**, 20039–20043.

Thunnissen, A.-M. W. H., Dijkstra, A. J., Kalk, K. H., Rozeboom, H. J., Engel, H., Keck, W. & Dijkstra, B. W. (1994). *Nature (London)*, **367**, 750–753.

Thunnissen, A.-M. W. H., Isaacs, N. W. & Dijkstra, B. W. (1996). Submitted.

Waxman, D. J. & Strominger, J. L. (1983). *A. Rev. Biochem.* **52**, 825–869.

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