Amino-acid sequence and three-dimensional structure of the *Staphylococcus aureus* metalloproteinase at 1.72 Å resolution

A Banbula[1], J Potempa[1], J Travis[2], C Fernandez-Catalán[3], K Mann[3], R Huber[3], W Bode[3] and FJ Medrano[3†*]

Background: Aureolysin is an extracellular zinc-dependent metalloproteinase from the pathogenic bacterium *Staphylococcus aureus*. This enzyme exhibits *in vitro* activity against several molecules of biological significance for the host, indicating that it is involved in the pathology of staphylococcal diseases.

Results: Here we report the amino-acid sequence and inhibitor-free X-ray crystal structure of aureolysin, a member of the thermolysin family of zinc-dependent metalloproteinases. This enzyme, which binds one zinc and three calcium ions, comprises a single chain of 301 amino acids that consists of a β-strand-rich upper domain and an α-helix-rich lower domain.

Conclusions: The overall structure of aureolysin is very similar to that of the other three members of this family whose structures are known – thermolysin (TLN) from *Bacillus thermoproteolyticus*, neutral protease (NP) from *Bacillus cereus* and elastase (PAE) from *Pseudomonas aeruginosa*. But an important difference has been encountered: in contrast to what has been observed in the other three members of this family (TLN, NP and PAE), inhibitor-free aureolysin displays a 'closed' active site cleft conformation. This new structure therefore raises questions about the universality of the hinge-bending motion model for the neutral metalloproteinases.

Addresses: [1]Jagiellonian University, Institute of Molecular Biology, Al. Mickiewicza 3, 31-120 Krakow, Poland, [2]Department of Biochemistry, The University of Georgia, Athens, GA 30602, USA and [3]Max-Plank-Institut für Biochemie, Abteilung Strukturforschung, Am Klopferspitz 18a, D-82152 Martinsried, Germany.

[†]Present address: C.I.B. (C.S.I.C.), C/ Velázquez, 144, 28006 - Madrid, Spain.

*Corresponding author.
E-mail: cibj154@fresno.csic.es

Key words: aureolysin, metalloproteinases, *Staphylococcus aureus*, thermolysin

Received: **26 May 1998**
Revisions requested: **18 June 1998**
Revisions received: **17 July 1998**
Accepted: **24 July 1998**

Structure 15 September 1998, **6**:1185–1193
http://biomednet.com/elecref/0969212600601185

Introduction

The thermolysin family of zinc-dependent metalloproteinases comprises secreted bacterial endopeptidases from both gram-positive and gram-negative bacteria such as *Bacillus*, *Legionella*, *Listeria*, *Pseudomonas*, and *Vibrio* [1]. Three-dimensional structures have been determined for three metalloproteinases of this family, namely, thermolysin (TLN) from *Bacillus thermoproteolyticus* [2], neutral proteinase (NP) from *Bacillus cereus* [3] and elastase (PAE) from *Pseudomonas aeruginosa* [4]. The architecture of these enzymes is very similar — in each of the structures an N-terminal subdomain, mainly consisting of a β-pleated sheet, can be distinguished from a predominantly α-helical C-terminal subdomain. The two domains are connected by a central α helix that provides several residues important for catalysis [5]. In addition to the zinc ion, these proteinases also bind calcium ions that play an important role in maintaining the active conformation of the enzymes.

Thermolysin-family enzymes produced by pathogenic bacterial strains have the potential to contribute to the pathogenicity of the bacteria. Some of these proteinases efficiently activate matrix metalloproteinases [6] and they might therefore play an important role in the proteolysis of extracellular matrix and in tissue destruction at the site of infection. PAE is capable of degrading several biologically significant molecules including complement components, immunoglobulin (Ig) A, IgG and human γ interferon, as well as some connective tissue proteins. The *Legionella pneumophila* metalloproteinase displays cytotoxic and tissue destructive activities, inhibitory effects on phagocytes and proteolytic activity against a broad spectrum of physiologically important substrates [7]. *Vibrio cholerae* hemagglutinin protease may contribute to the pathogenicity of the etiologic agent of cholera by direct cytotoxic activity [8] and/or activation of a cytolytic toxin [9]. In addition, the enzyme has the ability to digest several putative receptors for *V. cholerae* adhesins, thereby facilitating the vibrios detachment from intestinal epithelium and, in this way, possibly contributing to dissemination of the bacterium [10].

Several strains of *Staphylococcus aureus* produce aureolysin (also known as *S. aureus* metalloproteinase), an extracellular zinc-dependent neutral metalloproteinase [11,12]. The exact role of aureolysin in staphylococcal infections requires more detailed investigation. *In vitro*, cleavage of plasma proteinase inhibitors α_1-antichymotrypsin and α_1-proteinase inhibitor by aureolysin [13,14], and direct proteolytic activation of prothrombin in human plasma [15,16] has been observed, indicating an involvement of aureolysin in the pathology of staphylococcal diseases. Aureolysin may also modulate immunological reactions

by affecting stimulation of T and B lymphocytes by polyclonal activators and by inhibiting immunoglobulin production by lymphocytes [17].

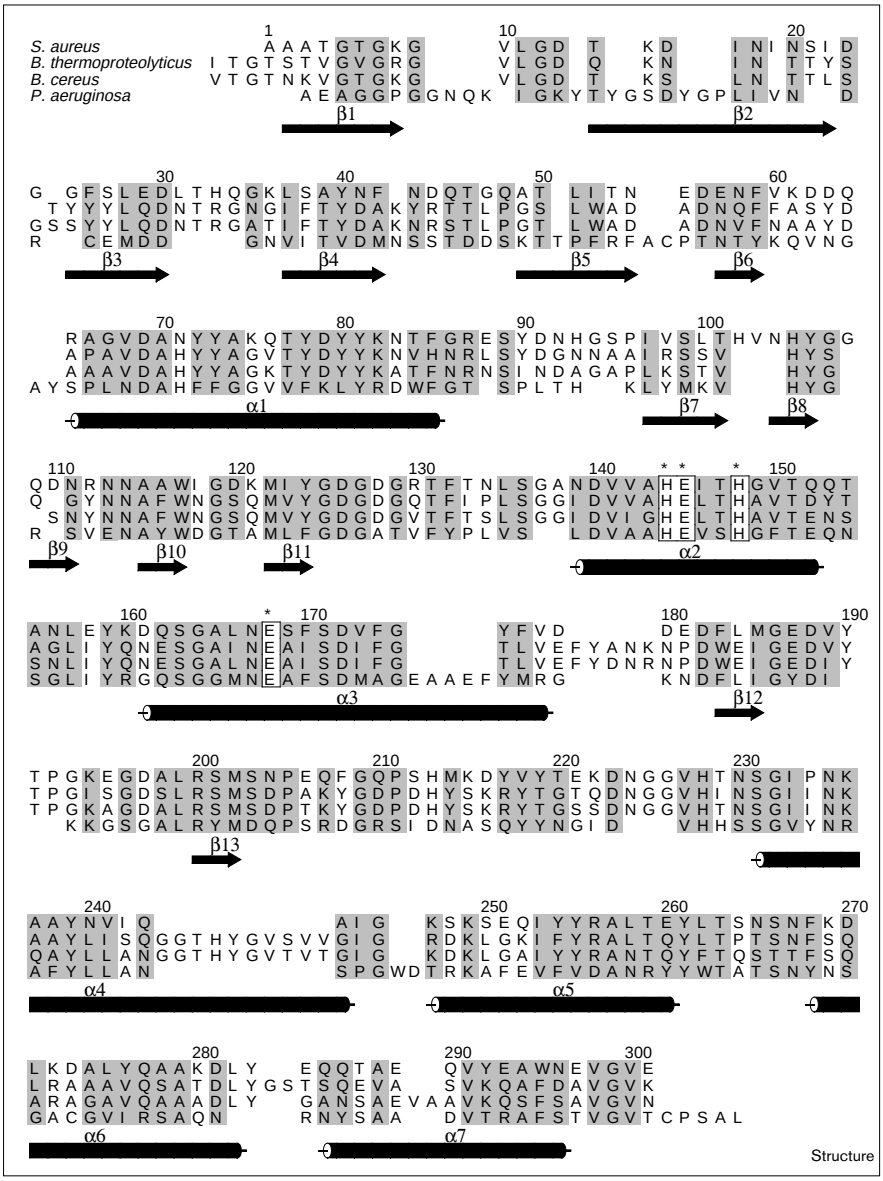
More systematic studies of the aureolysin function in *S. aureus* pathogenicity have been restricted by limited knowledge of the structure of the protein. Biochemical properties of aureolysin suggested that this enzyme might be related to thermolysin — aureolysin's substrate specificity is similar to that of thermolysin [18,19] and, like thermolysin, the enzyme requires the presence of both zinc and calcium ions to maintain catalytic activity [20,21]. Indeed, the amino-acid sequence and tertiary structure of aureolysin described in this paper allow us to classify the enzyme as the fourth structurally characterized member of the thermolysin family.

Results and discussion

Protein sequence

The amino-acid sequence of aureolysin is shown aligned with those of the other three members of the thermolysin family whose three-dimensional structures are known in Figure 1. The sequence of aureolysin was obtained by both amino-acid sequence analysis and X-ray crystallographic model building. The high similarity of the first 40 N-terminal residues of aureolysin with the metalloproteinase from *Staphylococcus epidermidis* (77.5%) allowed us to predict the location of the various peptides being

Figure 1



Sequence alignment of aureolysin from *S. aureus*, thermolysin *from B. thermoproteolyticus*, neutral protease from *B. cereus* and elastase from *P. aeruginosa*. Conserved residues are highlighted in grey and catalytic residues are boxed. Cylinders represent α helices and arrows represent β strands, as present in aureolysin. This figure was prepared with the program ALSCRIPT [33].

sequenced. Afterwards, the sequence was confirmed by modeling against the electron-density map.

Aureolysin is a single-chain enzyme and consists of 301 amino acids. It exhibits a high sequence similarity to the proteinases from the thermolysin family (Figure 1); sequence alignment performed with the program BESTFIT revealed a sequence identity of 48.6%, 51.1% and 36.7%, and a sequence similarity of 54.4%, 57.1% and 44.6% with TLN, PAE and NP, respectively. The high degree of similarity reveals that aureolysin belongs to the thermolysin family of metalloproteinases.

Two small peptide segments comprising amino acids 146–153 and 270–272 were not sequenced; the former segment contains one of the histidines that is located in the active site and involved in the binding of the catalytic zinc ion. The good quality of the electron-density map in these regions together with the sequence equivalence allowed the unequivocal crystallographic determination of the sequence (Figure 1). The peptide segment comprising amino acids 281–293 was obtained from only one tryptic fragment. When this sequence was compared with the electron density of this region, we found that amino acids 286 and 287 did not fit satisfactorily. Position 286, sequenced as a threonine, is located on the surface of the protein. Tentatively, because of the size of the density and the sequence homology with TLN and NP, a glutamine was placed at this position (Figure 1). Position 287, sequenced as alanine, exhibits a density with the shape of a threonine or valine residue and because of the presence of the Nε from Lys249 and a solvent molecule located within bond distance of the lateral chain of this residue, a threonine was placed at this position.

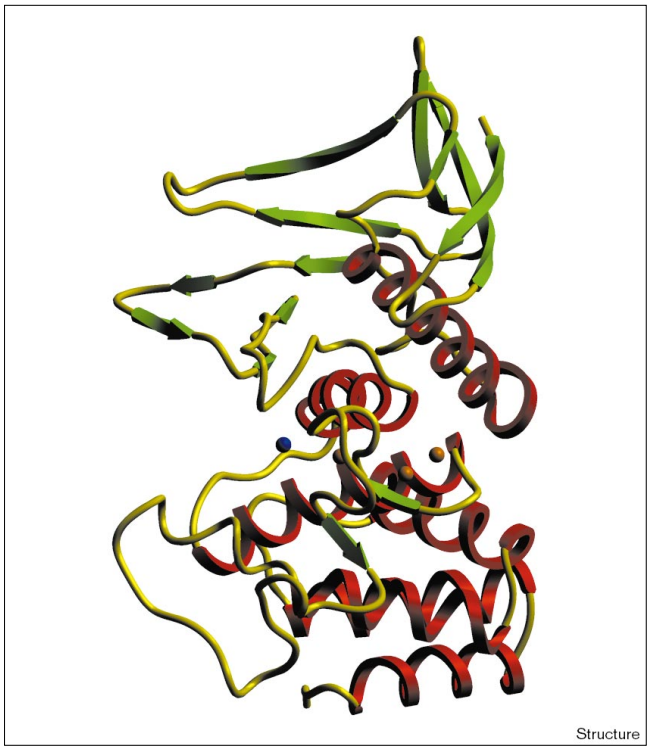
Overall structure

Structurally, aureolysin is clearly a member of the bacterial neutral-metalloproteinase family — the overall folding of the molecule is shown in Figure 2. The overall tertiary structure of aureolysin is similar to the known structures of TLN [2], NP [3] and PAE [4]. (The secondary structure elements of aureolysin are shown in Figure 1.)

The N-terminal subdomain (residues 1–156) consists of a distorted four-stranded β sheet followed by a five-stranded β sheet, a short two-stranded β hairpin, and two α helices. The C-terminal subdomain (residues 157–301) is composed of five α helices and a long loop (residues 179–230) that contains a short antiparallel β sheet. This loop contains all the ligands for the calcium ions. Helices α2 and α3 anchor the active-site zinc ion.

There is a ten-residue deletion in aureolysin, as well as in PAE, with respect to TLN and NP. In the last two proteins, these residues form a loop containing a short two-stranded β sheet (see Figure 1).

Figure 2



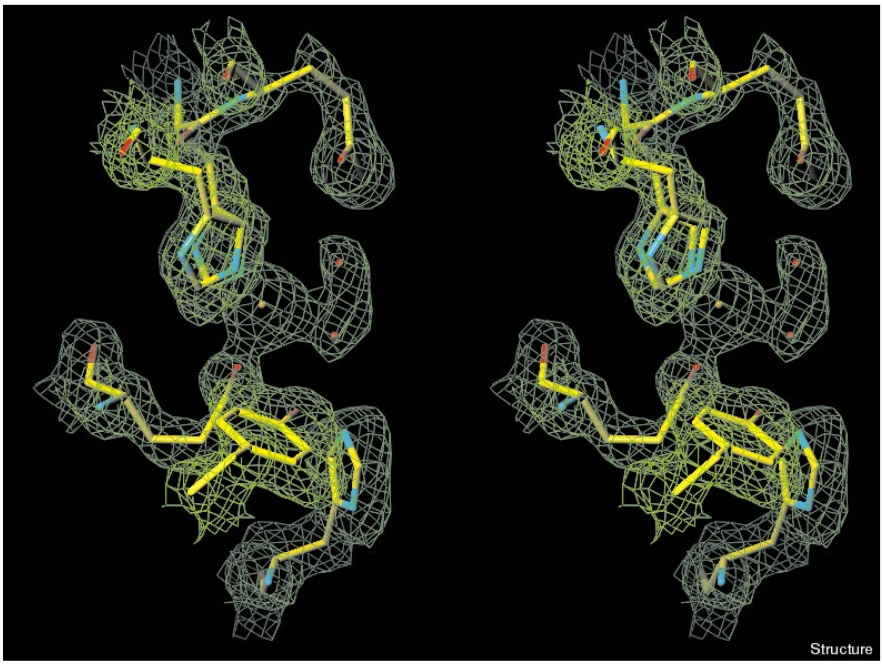
Ribbon representation of the overall structure of aureolysin based on the α-carbon positions of the final model. The α helices are shown in red, the β strands in green, and the loops are in yellow. The calcium ions are represented by brown spheres and the zinc ion is represented by a blue sphere. This figure was prepared with the program SETOR [34].

Active-site and specificity pocket

The active-site zinc ion is coordinated by the imidazole Nε2 atoms of His144 and His148 of helix α2, one carboxylate oxygen of Glu168, provided by helix α3 of the C-terminal subdomain, and two solvent molecules (Figure 3 and Table 1). The catalytic Glu145 binds through both carboxylate oxygen atoms to one of the solvent molecules that is liganded to the catalytic zinc ion. These two active-site solvent molecules are very close to one another (2.2 Å) and have B-factor values (53.0 $Å^2$ and 33.6 $Å^2$) higher than the average B values for the protein atoms in the vicinity of the active site (18 $Å^2$). They may represent a single disordered solvent molecule. The enzymatic mechanism is assumed to be the same as that reported for TLN [5,22].

All the residues involved in binding of the substrate and in the catalytic mechanism are strictly conserved among the four proteins whose three-dimensional structures are known (aureolysin, TLN, NP and PAE; Figure 1). The main specificity determinant seems to be the S1′ subsite, which is a hydrophobic pocket lined by residues Phe132, Leu135, Val141, Met185, Gly186 and Leu199. The only

Figure 3



Stereo view of the enzyme active site including the final $2F_o–F_c$ electron-density map contoured at 1σ. The brown sphere represents the zinc ion and the red spheres represent solvent molecules. This figure was prepared with the program TURBO–FRODO [30].

Table 1

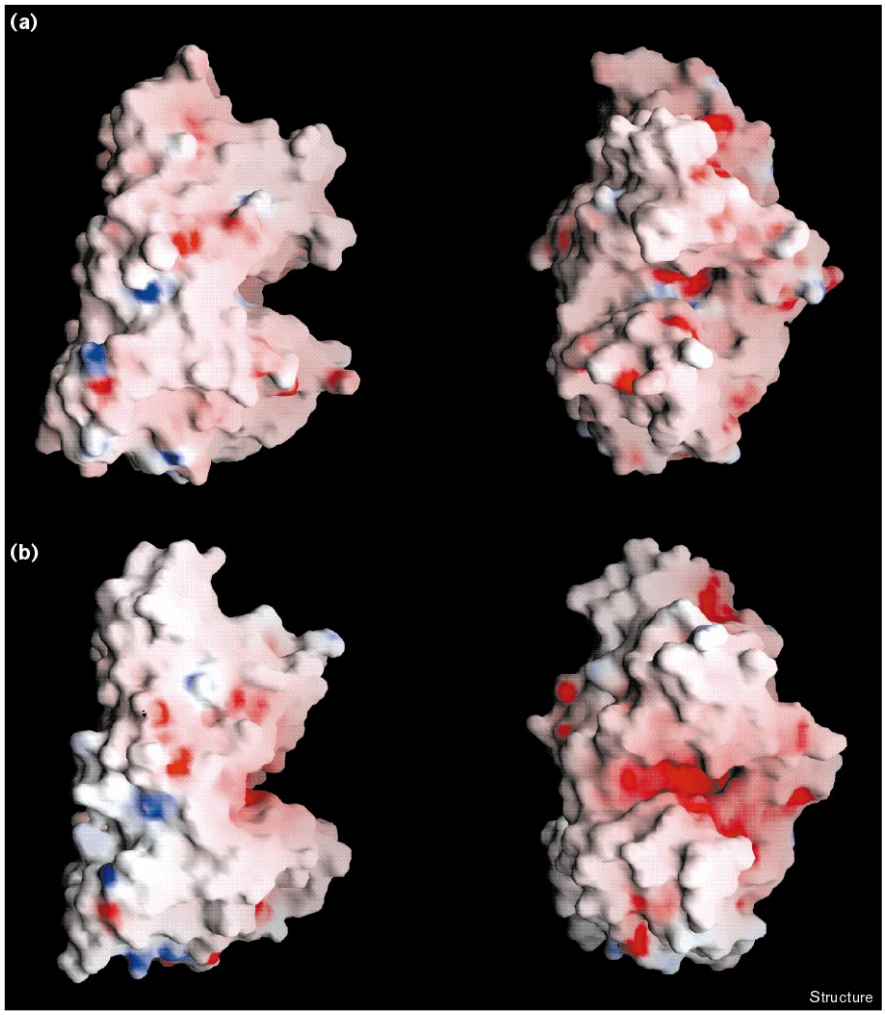
Ion–ligand distances in aureolysin and TLN.

Aureolysin				TLN*			
Ion	Residue	Atom (Å)	Distance (Å)	Ion	Residue	Atom (Å)	Distance (Å)
Zn	His144	Nε2	2.2	Zn	His142	Nε2	2.1
	His148	Nε2	2.2		His146	Nε2	2.1
	Glu168	Oε2	2.0		Glu166	Oε2	2.1
	Sol350†	O	2.2		Sol392†	O	1.9
	Sol406†	O	2.5				
Ca1	Asp140	Oδ1	2.4	Ca1	Asp138	Oδ1	2.5
	Asp179	Oδ1	2.3		Glu177	Oε1	2.5
	Asp182	Oδ2	2.4		Asp185	Oδ2	2.5
	Leu184	O	2.4		Glu187	O	2.3
	Glu187	Oε1	2.6		Glu190	Oε1	2.4
	Glu187	Oε2	2.6		Glu190	Oε2	2.5
	Sol402†	O	2.4		Sol346†	O	2.6
Ca2	Asp179	Oδ1	2.6	Ca2	Glu177	Oε2	2.4
	Asp179	Oδ2	2.6		Asp183	O	2.4
	Glu181	Oε1	2.2		Asp185	Oδ1	2.4
	Asp182	Oδ1	2.4		Glu190	Oε2	2.4
	Glu187	Oε1	2.3		Sol353†	O	2.5
	Sol359†	O	2.2		Sol475†	O	2.2
	Sol360†	O	2.4				
Ca3	Tyr190	O	2.3	Ca4	Tyr193	O	2.4
	Thr191	O	2.6		Thr194	O	2.5
	Thr191	Oγ1	2.4		Thr194	Oγ1	2.5
	Lys194	O	2.3		Ile197	O	2.3
	Asp197	Oδ1	2.5		Asp200	Oδ1	2.2
	Sol374†	O	2.2		Sol354†	O	2.4
	Sol375†	O	2.1		Sol480†	O	2.4

*The values for TLN were taken from [32]. †Sol = solvent molecule.

Figure 4

Solid-surface representation of the electrostatic potential of **(a)** aureolysin and **(b)** thermolysin. The colour coding corresponds to the surface potential, from negative (red) to positive (blue). This figure was prepared with the program GRASP [35].



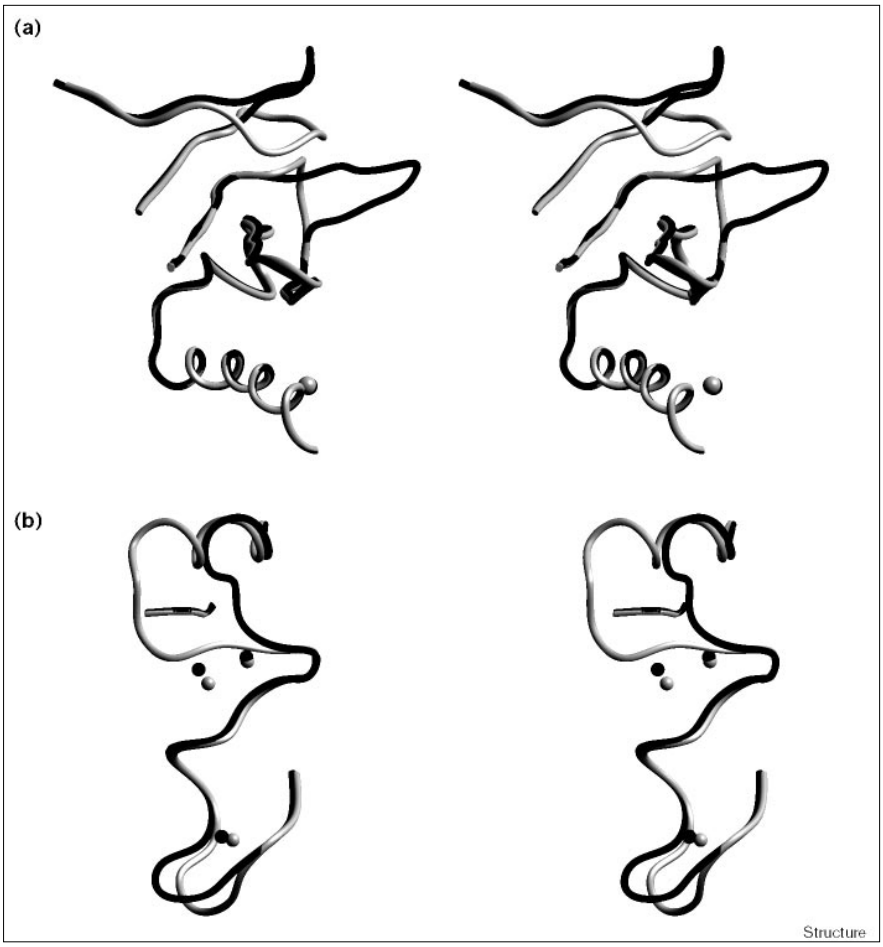
difference from the other three related proteinases is that Met185 replaces an isoleucine residue. In agreement with the specificity of TLN, this pocket is well equipped to accommodate big hydrophobic P1′ residues. One significant change close to the active site occurs at residues Ala116 and Ile118, occupied, respectively, by an aromatic residue (phenylalanine or tyrosine), and by a polar residue (asparagine or aspartate) in TLN, NP and PAE (Figure 1).

The entrance to the active site is a deep and narrow cleft between the N-terminal and the C-terminal subdomains (Figure 4). The residues provided to this cleft by the C-terminal subdomain are structurally equivalent in all four proteins. In contrast, the residues provided by the N-terminal subdomain, are different in aureolysin due to an insertion of five amino acids in the region 100–110 (Figure 5a). This insertion produces a narrower cleft in aureolysin than in the other three proteins. Moreover, this insertion induces a shift in the position of the loop, made up by residues 43–48, with respect to its position in the other proteins (Figure 5a).

Calcium-binding sites

We have observed three bound calcium ions in aureolysin; two calcium ions, Ca1 and Ca2, form a double calcium-binding site (Figure 6a). The residues liganding these calcium ions are all located in the loop 176–187, apart from Asp140, which is located at the N terminus of helix α2. Both calcium ions, which are 3.8 Å apart, are surrounded by seven ligands in an irregular pentagonal bipyramidal coordination (Table 1). A third calcium ion, Ca3, is bound near the double calcium-binding site (Figure 6a). This ion is bound to seven groups provided by residues located in the loop 190–197 and two water molecules. One of the ligands, Thr191, coordinates with both its mainchain carbonyl oxygen and its sidechain hydroxyl group. Ion–ligand distances are given in Table 1. The calcium-binding sites were confirmed by substitution

Figure 5



Superposition of **(a)** the residues from the N-terminal subdomain that contribute to the 'upper' lip of the active-site cleft and **(b)** the calcium-binding sites of aureolysin on thermolysin. The spheres represent the calcium ions. Aureolysin is shown in black and thermolysin in grey. This figure was prepared with the program SETOR [34].

with the more electron-dense lanthanum ions — calcium was replaced with lanthanum in the three binding sites.

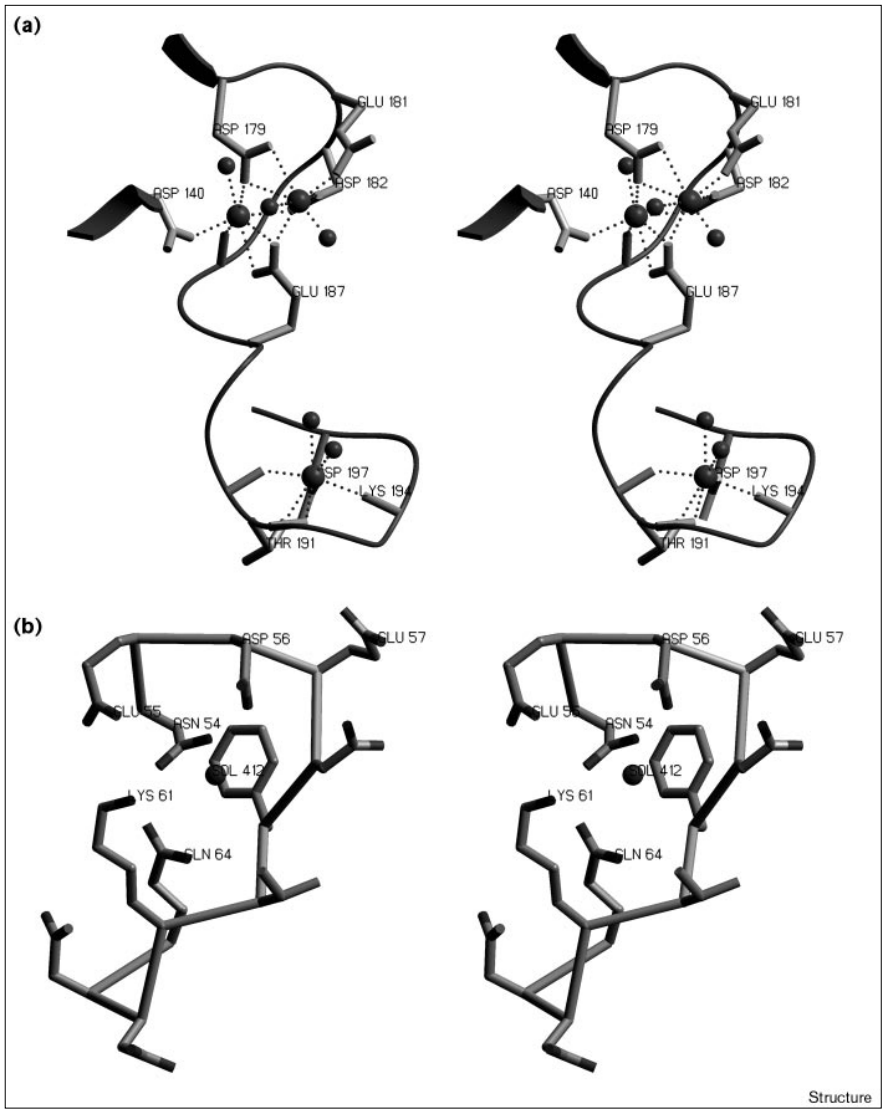
TLN and NP are known to bind four calcium ions, whereas PAE contains only one. The calcium ions present in aureolysin are located close to the active site in positions similar to the double-calcium site, Ca1 and Ca2, and Ca4 of TLN and NP (Figure 5b). The single calcium ion bound to PAE is located near the 'inner' site of the double-calcium site. The interactions in aureolysin are similar to the equivalent calcium sites in the homologous proteins. In the double-calcium site, the main significant difference, which is the result of a five-residue deletion, appears to be the carbonyl group of Asn184 liganding Ca2; in aureolysin, the equivalent interaction is made through one of the carboxyl groups of Glu181 (Figures 6a and 5b). The conformation of the loop that binds Ca3 is very similar to the ones in TLN and NP. Only residues Gly193 and Lys194 are slightly shifted (Figure 5b).

TLN and NP present a fourth calcium-binding site in the N-terminal subdomain but we have not observed any calcium ion bound to the equivalent position in aureolysin. Asn54, Glu55, Asp56, Lys61, Gln64 together with Sol412 exhibit a network of interactions that stabilizes this region (Figure 6b). These substitutions affect mainly acidic residues interacting with the calcium ion present in TLN and NP at this position.

Domain flexibility

Superimposing the entire structure of aureolysin onto TLN, NP and PAE yields root mean square deviations (rmsds) of 0.67 Å, 0.80 Å, and 1.25 Å for 288, 274, and 236 common residues, respectively (Table 2). Independent superimposition of the N-terminal and C-terminal subdomains gives significantly lower rmsds for only the C-terminal subdomain of NP and PAE (Table 2). This indicates an almost equivalent subdomain arrangement in aureolysin and TLN, a small difference with respect to NP and a significant relative-position difference towards PAE. As a result of presence of an extra loop formed by five additional residues in the 'upper' part of the cleft (Figure 5a), the active-site cleft between the two domains is significantly more 'closed' in aureolysin than in the other three proteins

Figure 6



Calcium-binding sites and region of the missing calcium ion of aureolysin. **(a)** Stereoview of the calcium-binding sites, showing the calcium ligands. The larger spheres represent calcium ions and the smaller spheres solvent molecules. **(b)** Stereoview of residues 54–63 and solvent molecule 412. TLN and NP bind a calcium ion in this region. This figure was prepared with SETOR [34].

(Figure 4). If this loop is disregarded, the relative orientation of both subdomains in aureolysin is very similar to that in TLN, giving rise to a more closed conformation of the active-site cleft than in NP and PAE. These results are very similar to those obtained from the comparison of TLN with NP and PAE [23]. Prompted by the fact that TLN presents a bound dipeptide in its active site [23], we carefully studied this region of aureolysin. The electron-density map in the active-site cleft did not indicate the presence of any density in the S1′ subsite that might account for a contiguous peptide such as the one in NP, apart from a couple of solvent molecules [23]. In order to confirm this result, an electron-density map combining phases obtained from the active-site zinc, a mercury derivative, the lanthanum derivative and the model obtained by molecular replacement was calculated. This electron-density map showed density only for the solvent molecules in the S1′ subsite. This evidence confirms that this structure is of the free enzyme. In NP, residues Gly136 and Gly137 have been proposed to be a

Table 2

Root mean square deviations.

	TLN	NP	PAE
Entire structure	0.67 Å (288)	0.80 Å (274)	1.25 Å (236)
N-terminal domain	0.65 Å (141)	0.73 Å (143)	1.27 Å (114)
C-terminal domain	0.59 Å (141)	0.59 Å (142)	0.86 Å (127)

Root mean square deviations between the entire structure and the N-terminal and C-terminal subdomains of aureolysin and TLN, NP and PAE. Numbers in parentheses indicate the number of α-carbon atoms included in the comparison.

flexible hinge. Most of the residues of the aromatic cluster proposed to be involved in the hinge-bending mechanism are conserved in aureolysin [24]; only Gly137, which is substituted with an alanine residue, is not conserved. Here we present the structure of inhibitor-free aureolysin. This structure is very similar to the inhibitor-bound TLN structure. At the present moment it is difficult to say if the closed conformation exhibited by aureolysin represents its active form, or if it is due to the crystal packing. We are currently looking for inhibitors of aureolysin in order to obtain the inhibitor-bound structure of the enzyme.

Biological implications

Staphylococcus aureus is a frequent cause of human disease. This bacterium produces several extracellular proteinases, including a zinc-dependent metalloproteinase, aureolysin. Aureolysin belongs to the family of metalloproteinases that encompasses several enzymes considered to be important virulence factors, including _Pseudomonas aeruginosa_ elastase, _Legionella pneumophila_ and _Listeria monocytogenes_ metalloproteinases, _Vibrio cholerae_ hemagglutinin protease, _Staphylococcus epidermidis_ elastase and the lambda toxin of _Clostridium perfringens_. Relatively little is known about possible pathological functions of aureolysin; however, the fact that this enzyme is homologous to the above proteinases implies that it may play a role both in tissue destruction and in the evasion of host defenses.

The high-resolution structure of aureolysin is very similar to those of TLN, NP and PAE. Despite a similar substrate specificity for bulky hydrophobic groups at P1′, aureolysin, in contrast to its tertiary homologs, has no elastinolytic activity. This difference can be easily explained by the presence of a loop peptide on top of the active-site cleft.

Our knowledge of the structure of aureolysin will prove valuable in further studies on the molecular features of this enzyme that are responsible for the differences in thermal stability, substrate specificity and inhibitor susceptibility, relative to other members of the thermolysin family of metalloproteinases. These studies may help to understand any pathological significance of aureolysin.

Materials and methods

Protein purification

Aureolysin was purified from *Staphylococcus aureus* strain V8–BC 10. The bacteria were grown in a casein hydrolysate yeast extract medium (CCYI medium) [9]. Purification from the culture media was carried out as described previously [19]. Briefly, the culture media was saturated with ammonium sulfate (60%), the precipitated protein was redissolved in 10 mM Tris–HCl (pH 7.6) and 5 mM $CaCl_2$, and then precipitated again with cold acetone (60%); the pellet was redissolved and extensively dialyzed against the same buffer. The protein was then applied to a DE-52 column and eluted with a gradient of 0–0.6 M KCl. Active fractions were pooled and rechromatographed under the same conditions on a TSK-DEAE column. Purity was assessed by SDS–PAGE [25] and gel filtration.

Amino-acid sequence determination

For cleavage with trypsin, the purified protein was dialyzed against 0.2 M ammonium hydrogen carbonate. Trypsin was added to the protein solution in a 1:50 (w/w) enzyme to substrate ratio, and cleavage was allowed to proceed for 16 h at 30°C. The reaction was terminated by addition of trifluoroacetic acid to pH 1. The incubation mixture was centrifuged and the soluble peptides were separated by reversed-phase HPLC (3 mm × 250 mm column, Vydac C18, 5 μm) using a gradient of 0%–42% (v/v) acetonitrile in 0.1% trifluoroacetic acid in 160 min at a flow rate of 0.25 ml/min. For cyanogen bromide (CNBr) cleavage, the protein was dialyzed against 70% formic acid and cleavage was started by addition of a small crystal of CNBr. After 24 h incubation in the dark, the mixture was dried under a stream of N_2. The fragments were dissolved in 0.1 M sodium acetate (pH 5.0).

The N terminus of the protein and selected peptides were sequenced by Edman degradation using Applied Biosystems protein sequencers 470A and 473A according to the manufacturer's instructions.

Crystallization

Crystallization was carried out at 20°C using the sitting-drop vapour-diffusion method. All reservoir solutions contained 0.02% (w/v) sodium azide. The initial protein concentration was 10 mg/ml in 10 mM Tris–HCl (pH 7.6) and 5 mM $CaCl_2$. Drops were prepared by mixing 2 μl of the protein solution with 2 μl of the reservoir solution. The mixture was allowed to equilibrate against 0.4 ml reservoir solution consisting of 0.1 M sodium acetate (pH 5.2) and 40% polyethylene glycol (MW 4000). After about three days, crystals of maximal dimensions 0.5 mm × 0.3 mm × 0.3 mm were obtained, which diffracted beyond 1.72 Å resolution. The crystals belong to the orthorhombic space group $P2_12_12_1$ (the correct space group was determined after the translational searches; see below) with cell constants a = 33.7 Å, b = 79.5 Å, and c = 100.1 Å.

Data collection and processing

Crystals were harvested using the reservoir solution and mounted in thin-walled glass capillary tubes. X-ray diffraction data were collected on a 300 mm MAR-Research image plate detector attached to a Rigaku RU200 rotating-anode X-ray generator providing graphite-monochromatized CuKα radiation. Intensities were integrated with the program DENZO, then scaled and merged with the program SCALEPACK [26]. Intensities were converted to structure factors with programs from the CCP4 suite [27]. The data collection statistics are summarized in Table 3.

Refinement

The structure was solved by molecular replacement using the coordinates of thermolysin from *Bacillus thermoproteolyticus* [2]. The rotational and translational searches were performed with the program AMoRe [28] using data from 10–4 Å resolution, confirming $P2_12_12_1$ as the correct space group. The highest solution corresponded to a correlation coefficient of 37.3 and a crystallographic R_{factor} of 49.1%. The corresponding figures for the next best solution were 19.2 and 53.6%.

After a positional refinement using the program X-PLOR [29], the calculated ($3F_{obs}-2F_{calc}$) electron density permitted modeling of the sidechains of aureolysin using the program TURBO-FRODO [30]. Subsequent cycles of rebuilding and refinement finally revealed a model with a crystallographic R_{factor} of 17.6% for data between 8.0 and 1.72 Å (for all data between 15.0 and 1.72 Å, R_{factor} = 18.2% and R_{free} = 24.7%). Three calcium cations and one zinc cation were tentatively assigned on the basis of the high density and reasonable temperature factors after refinement. 160 solvent molecules were further introduced at stereochemically reasonable positions.

The polypeptide chain is well defined apart from residues 23, 24 and the C-terminal Glu301. In addition, the sidechains of a few surface residues (Gln45, Gln109, Arg130, Glu221, Lys222, and Lys269) are not well defined by proper electron density. A Ramachandran plot of the main-chain torsion angles, performed with the program PROCHECK [31],

Table 3

Data collection and refinement statistics.

Diffraction data	
Space group	$P2_12_12_1$
Cell constants a, b, c (Å)	33.7, 79.5, 100.1
Limiting resolution (Å)	1.72
No of measured reflections	120 813
No of unique reflections	27 738
I / σ(I)	5.6
R_{merge}* (%)	9.1
Completeness (%)	
(15.0–1.72 Å)	93.8
(1.78–1.72 Å)	86.4
Refinement and final model	
Resolution range	8.0–1.72
Reflections used	27 431
Crystallographic R_{factor}† (free R_{factor}‡) (%)	17.6 (23.9)
Rmsd	
Bonds (Å)	0.010
Angles (°)	1.380
No of non-hydrogen protein atoms	2354
Ions	
Calcium	3
Zinc	1
Solvent molecules	160

*$R_{merge} = \Sigma\Sigma|I(h)_i - \langle I(h)\rangle| / \Sigma\langle I(h)\rangle$; $I(h)_i$ is the observed intensity of the ith measurement of reflection h, and $\langle I(h)\rangle$ the mean intensity of reflection h; calculated after loading, and scaling.
†$R_{factor} = (\Sigma||F_o| - |F_c|| / \Sigma|F_o|) \times 100$. ‡$R_{free}$ was calculated randomly omitting 10% of the observed reflections from refinement and R_{factor} calculation.

shows all residues but Ser89 and Asp161 situated in allowed regions. These two residues are located in generously allowed positions.

Accession numbers

The amino-acid sequence of aureolysin has been deposited with the Swiss-Prot Data Bank with accession number P81177. The atomic coordinates have been deposited with the Brookhaven Protein Data Bank with accession code 1bqb.

Acknowledgements

We thank Antonio Romero and FX Gomis-Rüth for their invaluable help. The financial support by the Biotech program of the EC (ERBB10-4CT96-0464) is kindly acknowledged.

References

1. Hase, C.C. & Finkelstein, R.A. (1993). Bacterial extracellular zinc-containing metalloproinases. *Microbiol. Rev.* **57**, 823-837.
2. Matthews, B.W., Colman, P.M., Jansonius, J.N., Titani, K., Walsh, K.A. & Neurath, H. (1972). Structure of thermolysin. *Nature New Biol.* **238**, 41-43.
3. Pauptit, R.A., Karlsson, R., Picot, D., Jenkins, J.A., Niklaus-Reimer, A.-S. & Jansonius, J.N. (1988). Crystal structure of neutral protease from *Bacillus cereus* refined at 3.0 Å resolution and comparison with the homologous but more thermostable enzyme thermolysin. *J. Mol. Biol.* **199**, 525-537.
4. Thayer, M.M., Flaherty, K.M. & McKay, D.B. (1991). Three-dimensional structure of the elastase of *Pseudomonas aeruginosa* at 1.5 Å resolution. *J. Biol. Chem.* **266**, 2864-2871.
5. Hangauer, D.G., Monzingo, A.F. & Matthews, B.W. (1984). An interactive computer graphics study of thermolysin-catalyzed peptide cleavage and inhibition by N-carboxymethyl dipeptides. *Biochemistry* **23**, 5730-5741.
6. Okamoto, T. *et al*., & Maeda, H. (1997). Activation of human matrix metalloproteinases by various bacterial proteinases. *J. Biol. Chem.* **272**, 6059-6066.
7. Dowling, J.A., Saha, A.K. & Glew, R.A. (1992). Virulence factors of the family *Legionellaceae. Microbiol. Rev.* **56**, 32-60.
8. Wu, Z., Milton, D., Nybom, P., Sjo, A., Magnusson, K.E. (1996). *Vibrio cholerae* hemagglutinin/protease (HA/protease) causes morphological changes in cultured epithelial cells and perturbs their paracellular barrier function. *Microb. Pathog.* **21**, 111-123.
9. Nagamune, K., Yamamoto, K., Naka, A., Matsuyama, J., Miwatani, T. & Honda, T. (1996) *In vitro* proteolytic processing and activation of the recombinant precursor of E1 tor cytolysin/hemolysin (pro-HlyA) of *Vibrio cholerae* by soluble hemagglutinin/protease of *V. cholerae*, trypsin, and other proteases. *Infect. Immun.* **64**, 4655-4658.
10. Finkelstein, R.A., Boesman-Finkelstein, M., Chang, Y. & Hase, C.C. (1992). *Vibrio cholerae* hemagglutinin/protease, colonial variation, virulence, and detachment. *Infect. Immun.* **60**, 472-478.
11. Arvidson, S.O., Holme, T. & Wadstrom, T. (1970). Formation of bacteriolytic enzymes in batch and continuous culture of *Staphylococcus aureus*. *J. Bacteriol.* **104**, 227-233.
12. Björklind, A. & Arvidson, S.O. (1977). Occurrence of an extracellular serine proteinase among *Staphylococcus aureus* strains. *Acta Pathol. Microbiol. Immunol. Scand* [B], **85**, 277-280.
13. Potempa, J., Watorek, W. & Travis, J. (1986). The inactivation of human α-1-proteinase inhibitor by proteinases from *Staphylococcus aureus*. *J. Biol. Chem.* **261**, 14330-14334.
14. Potempa, J., Fedak, D., Dubin, A., Mast, A. & Travis, J. (1991). Proteolytic inactivation of a-1-anti-chymotrypsin. Sites of cleavage and generation of chemotactic activity. *J. Biol. Chem.* **266**, 21482-21487.
15. Wegrzynowicz, Z., Hechko, P.B., Jeljaszewicz, J., Neugebauer, M. & Pulverer, G. (1979). Pseudocoagulase activity of staphylococci. *J. Clin. Microbiol.* **9**, 15-19.
16. Wegrzynowicz, Z., Hechko, P.B., Drapeau, G., Jeljaszewicz, J. & Pulverer, G. (1980). Prothrombin activation by metalloproteinase from *Staphylococcus aureus*. *J. Clin. Microbiol.* **12**, 138-139.
17. Prokesova, L., Porwit-Bobr, Z., Baran, K., Potempa, J., Pospisil, M. & John, C. (1991). Effect of metalloproteinase from *Staphylococcus aureus* on *in vitro* stimulation of human lymphocytes. *Immunol. Lett.* **27**, 225-230.
18. Björklind, A. & Jörnvall, H. (1974). Substrate specificity of three different extracellular proteolytic enzymes from *Staphylococcus aureus*. *Biochim. Biophys. Acta* **370**, 524-529.
19. Saheb, S.A. (1976). Purification et caractérisation d'une protéase extracellulaire de *Staphylococcus aureus* inhibée par l'E.D.T.A. *Biochemie* **58**, 793-804.
20. Arvidson, S.O. (1973). Studies on extracellular proteolytic enzymes from *Staphylococcus aureus*. II. Isolation and characterization of an EDTA-sensitive protease. *Biochim. Biophys Acta* **302**, 149-157.
21. Drapeau, G.R. (1978). Role of a metalloprotease in activation of the precursor of staphylococcal protease. *J. Bacteriol.* **136**, 607-613.
22. Monzingo, A.F. & Matthews, B.W. (1984). Binding of N-carboxymethyl dipeptide inhibitors to thermolysin determined by X-ray crystallography: A novel class of transition-state analogues for zinc peptidases. *Biochemistry* **23**, 5724-5729.
23. Holland, D.R., *et al*., & Matthews, B.W. (1992). Structural comparison suggests that thermolysin and related neutral proteases undergo hinge-bending motion during catalysis. *Biochemistry* **31**, 11310-11316.
24. Stark, W., Paupitt, R.A., Wilson, K.S. & Jansonius, J.N. (1992). The structure of neutral protease from *Bacillus cereus* at 0.2-nm resolution. *Eur. J. Biochem.* **207**, 781-791.
25. Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.
26. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.
27. Collaborative Computational Project, No. 4. (1994). The CCP4 Suite: Programs for protein crystallography. *Acta Cryst. D* **50**, 760-763.
28. Navaza, J. (1994) AmoRe: an automated package for molecular replacement *Acta Cryst. A* **50**, 157-163.
29. Brünger, A.T. (1992). X-PLOR Version 3.1, A system for X-ray Crystallography and NMR. Yale University Press, New Haven, CT.
30. Roussel, A. & Cambilleau, C. (1992). TURBO-FRODO, Biographics, LCCMB, Marseille, France.
31. Laskowski, R.A., McArthur, M.W., Moss, D.S. & Thornton, J. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283-291.
32. Holmes, M.A. & Matthews, B.W. (1982). Structure of thermolysin refined at 1.6 Å resolution. *J. Mol. Biol.* **160**, 623-639.
33. Barton, G.J. (1993). ALSCRIPT: a tool to format multiple sequence alignments *Protein Eng.* **6**, 37-40.
34. Evans, S.V. (1993). SETOR: hardware lighted three-dimensional solid model representations of macromolecules. *J. Mol. Graph.* **11**, 134-138.
35. Nicholls, A., Bharadwaj, R. & Honig, B. (1993). GRASP – graphical representation and analysis of surface properties. *Biophys. J.* **64**, A166.