Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

1.35 and 2.07 Å resolution structures of the red abalone sperm lysin monomer and dimer reveal features involved in receptor binding

Nicole Kresge,[a] Victor D. Vacquier[b] and C. David Stout[a]*

[a]Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037-1093, USA, and [b]Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92093-0202, USA

Correspondence e-mail: dave@scripps.edu

Received 8 September 1999
Accepted 10 November 1999

PDB References: sperm lysin monomer, 2lis; sperm lysin dimer, 2lyn.

Abalone sperm use lysin to make a hole in the egg's protective vitelline envelope (VE). When released from sperm, lysin first binds to the VE receptor for lysin (VERL) then dissolves the VE by a non-enzymatic mechanism. The structures of the monomeric and dimeric forms of *Haliotis rufescens* (red abalone) lysin, previously solved at 1.90 and 2.75 Å, respectively, have now been refined to 1.35 and 2.07 Å, respectively. The monomeric form of lysin was refined using previously obtained crystallization conditions, while the dimer was solved in a new crystal form with four molecules (two dimers) per asymmetric unit. These high-resolution structures reveal alternate residue conformations, enabling a thorough analysis of the conserved residues contributing to the amphipathic nature of lysin. The availability of five independent high-resolution copies of lysin permits comparisons leading to insights on the local flexibility of lysin and alternative conformations of the hypervariable N-terminus, thought to be involved in species-specific receptor recognition. The new analysis led to the discovery of the basic nature of a cleft formed upon dimerization and a patch of basic residues in the dimer interface. Identification of these features was not possible at lower resolution. In light of this new information, a model explaining the binding of sperm lysin to egg VERL and the subsequent dissolution of the egg VE is proposed.

1. Introduction

Abalone are archeogastropod mollusks of the genus *Haliotis*, seven species of which inhabit the coast of California. Many of these species have overlapping breeding seasons and habitats, yet hybrids are rare (Lindberg, 1992; Owen *et al.*, 1971). The species-specificity of abalone gamete interaction arises in part owing to a 16 kDa sperm protein named lysin. When the sperm contacts the egg it undergoes the acrosome reaction, which releases lysin onto the egg's vitelline envelope (VE). A major fibrous component of the VE is a 1000 kDa gamete-recognition glycoprotein called the VE receptor for lysin (VERL). VERL is rich in acidic amino acids and contains approximately 26 tandem repeats of a species-specific lysin-binding motif of 153 amino acids. Approximately two molecules of lysin bind each VERL repeat ($EC_{50} = 10^{-10}\,M$), resulting in an average of about 50 lysins bound per VERL molecule (Swanson & Vacquier, 1997). Lysin's binding to VERL causes the VE filaments to lose cohesion and splay apart, creating a 3 μm diameter hole through which the sperm swims to fuse with the egg (Lewis *et al.*, 1982; Vacquier *et al.*, 1999; Swanson *et al.*, 1999).

Among the seven California abalone species, there is a high degree of divergence in the sequence of lysin. Although 50%

of the amino-acid positions are conserved in all seven species, replacements in the other 50% are non-conservative in regard to charge and class. N-terminal amino acids 2–12 are hypervariable and differ dramatically in charge distribution between species (Lee & Vacquier, 1992; Lyon & Vacquier, 1999).

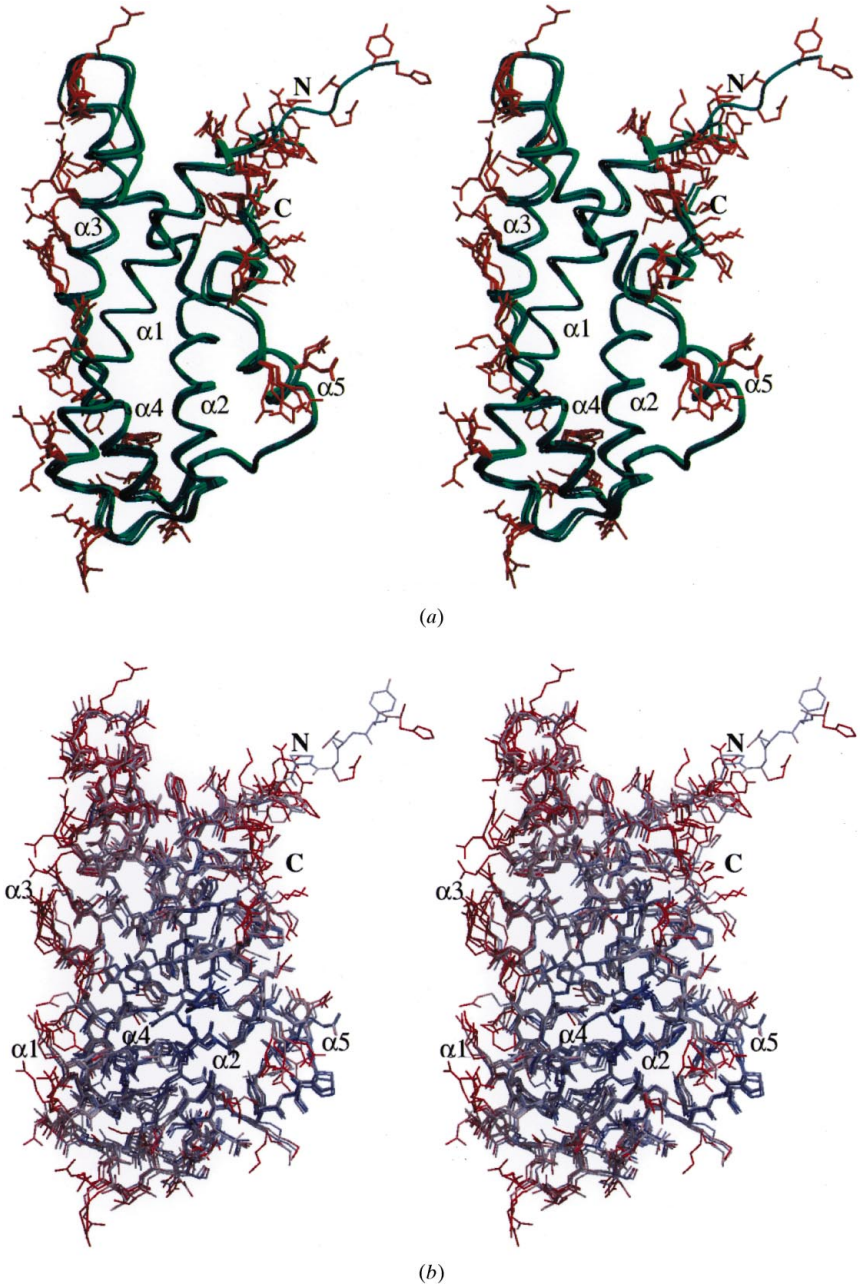

Figure 1
Stereoviews of superpositions of lysin dimer and monomer subunits displaying variability in side chains and in *B* factors. (*a*) Superposition of the five monomers showing side chains with an r.m.s. difference greater than 2σ following least-squares fit of the subunits. Main chains are shown in ribbon form (green) and side chains are shown in stick form (orange). Residues with higher r.m.s. differences are concentrated in the region containing the N-terminus, the loop between helices $\alpha2$ and $\alpha3$, the N-terminal region of helix $\alpha3$, the C-terminal region of helix $\alpha5$ and the C-terminus. (*b*) Overlay of the five independent copies of the monomer colored according to normalized *B* factors. Color is ramped from red to blue *via* white, with high *B* factors colored red and low *B* factors colored blue. Areas with higher *B* factors are concentrated on the solvent-exposed side chains of the molecules, especially in the hypervariable termini and in the adjacent turn between helices $\alpha2$ and $\alpha3$. These figures were created with *MOLSCRIPT* (Kraulis, 1991) and *BOBSCRIPT* (Esnouf, 1997) and were rendered with *RASTER3D* (Merrit & Murphy, 1994).

Lysin exists as a dimer in the acrosomal vesicle and at physiological concentrations in seawater (pH 7.8). Fluorescence resonance energy-transfer experiments show that at pH 7.8 the dimer has an approximate K_D of 1 μM and that the half-time of monomer exchange between dimers is approximately 8 min (Shaw *et al.*, 1995). When released from the acrosomal vesicle, lysin's local concentration on the VE surface is greater than 60 μM, indicating that it is a dimer when it contacts the egg. The addition of isolated VE to a solution of dimers causes immediate dissociation of the dimer, suggesting the lysin monomer is the active species in binding VERL. However, it is not known whether dissociation of the dimer is required for interaction with VERL, or whether it is a result of interaction with VERL.

The crystal structures of the monomeric and dimeric forms of red abalone lysin were previously solved to 1.90 Å (Shaw *et al.*, 1993) and 2.75 Å (Shaw *et al.*, 1995). The structure of the monomer contains an up–down five-helix bundle with a right-handed twist (Fig. 1*a*). The N-terminal hypervariable region extends away from the helical core but is in association with the C-terminus (Fig. 1*a*). Two tracks of basic residues run the length of one side of the molecule, while the opposite surface possesses a solvent-exposed patch of hydrophobic residues. The lysin dimer assumes an extended S-shape when viewed along the direction of the twofold axis relating the dimer subunits (*e.g.* Fig. 4). The dimer is formed by the association of the hydrophobic patches of monomers, which places the basic tracks on opposite sides of the dimer. There is a large cleft on the opposite sides of the twofold axis (Fig. 4).

Using synchrotron radiation and cryo-cooling, data were collected to 1.35 Å resolution on the monomeric form and to 2.07 Å resolution on a new crystal form of the red abalone lysin dimer. These high-resolution structures reveal features of the lysin molecule and its surface that were not discernable in the lower resolution structures. The refined structures were also used for analysis of charged and hydrophobic residue distribution, solvent-accessible surface, exposed hydrophobic surface area and electrostatic potential. The analysis

Table 1
Data collection and refinement statistics.

Values for the highest resolution shell are given in parentheses.

	Monomer	Dimer
Data statistics		
Space group	$P2_12_12_1$	$P4_12_12$
Unit-cell dimensions (Å)		
a	51.41	72.52
b	45.31	72.52
c	81.14	266.08
Resolution range (Å)	18.63–1.35	40.50–2.07
Highest resolution shell (Å)	1.38–1.35	2.18–2.07
Molecules per asymmetric unit	1	4
Observed reflections	259591	948262
Unique reflections	41455	44137
$I/\sigma(I)$	7.2 (8.6)	4.8 (2.8)
Completeness (%)	97.4 (88.7)	99.1 (94.6)
Multiplicity	4.6 (2.4)	6.9 (4.5)
R_{sym}† (%)	4.5 (8.2)	8.7 (25.6)
Refinement statistics		
R factor‡ (%)	13.6	21.3
R_{free}§ (%)	17.8	28.5
Number of non-H atoms	1398	4618
Number of water molecules	284	316
Average isotropic *B* factors (Å^2)¶		
Protein, overall	14.7	Subunit *A*, 31.5 Subunit *B*, 56.5 Subunit *C*, 30.5; Subunit *D*, 29.6
Main chain	11.7 (0.81)	34.5 (3.78)
Side chain	20.4 (5.71)	40.2 (6.20)
Water	32.5	39.7
Ramachandran statistics††		
Most favored regions (%)	98.3	94.9
Additional allowed regions (%)	1.7	4.9
Generously allowed regions (%)		0.2
R.m.s. deviations from ideal geometry‡‡		
Bond distance (Å)	0.013	0.005
Bond angle (°)	2.47	1.55

† $R_{sym} = 100 \sum_h \sum_j |I_{hj} - I_h| / \sum_h \sum_j I_{hj}$, where I_h is the weighted mean intensity of the symmetry-related reflections I_{hj}. ‡ *R* factor $= \sum ||F_{obs}| - |F_{calc}|| / \sum |F_{obs}|$, where $|F_{obs}|$ and $|F_{calc}|$ are the observed and calculated structure-factor amplitudes, respectively. § R_{free} is the *R* factor for a 5% subset of reflections not used in refinement. ¶ R.m.s. deviations in *B* values are indicated in parentheses. †† As calculated in the program *PROCHECK* (Laskowski *et al.*, 1993). ‡‡ As calculated in the program *WHATIF* (Hooft *et al.*, 1996).

reveals (i) alternate conformations for the hypervariable N-terminus relevant to lysin's role in the species-specific recognition of VERL (Lyon & Vacquier, 1999), (ii) large positively charged clefts formed by dimerization that contain species-variable residues, (iii) a patch of basic residues in the dimer interface that account for the low K_D of the dimer and (iv) additional residues involved in the hydrophobic patch. Comparisons of the new structures further reveal the flexible nature of lysin's surface residues and loops.

Based on these results and biochemical data for VERL (Swanson & Vacquier, 1997, 1998), a model is proposed to explain lysin's recognition of VERL and its dissolution of the egg VE. The model involves initial species-specific recognition between the lysin dimer and VERL, followed by monomerization and high-affinity binding between VERL and lysin monomers, resulting in a loss of cohesion between VE filaments.

2. Materials and methods

2.1. Protein preparation and crystallization

Lysin from red abalone was isolated and purified by CM cellulose chromatography (Vacquier & Lee, 1993). The protein was dialyzed against 50 m*M* sodium acetate pH 5.5, 0.1%(*w*/*v*) NaN_3 and was concentrated to 14 mg ml^{-1}. Lysin monomer and dimer crystals were grown at 295 K. Orthorhombic crystals of the monomer were grown as previously described (Table 1; Shaw *et al.*, 1993). The dimeric form of lysin was crystallized using a reservoir solution of 0.75 *M* ammonium sulfate, 20 m*M* NH_4SCN, 52.5 m*M* sodium citrate/boric acid/citric acid buffer pH 4.5 (from a stock solution of 1.5 *M* sodium citrate, 10 m*M* boric acid, 1.5 *M* citric acid), 0.1 m*M* EDTA and 0.02% 1-*S*-octyl-β-D-thioglucopyranoside (OTGP). Drops contained 2.5 µl protein solution and 2.5 µl precipitant, and wells contained 1 ml of reservoir solution. The characteristics of these tetragonal dimer crystals are shown in Table 1.

2.2. Data collection

All data collection was performed at the Stanford Synchrotron Radiation Laboratory (SSRL) beamline 9-1, using a MAR Research image-plate scanner and monochromatic radiation of wavelength 0.98 Å. Low-temperature data were collected using 20% glycerol (for the orthorhombic form) or 35% ethylene glycol (tetragonal form) as cryoprotectants. The data set for the monomeric form of lysin was collected to 1.35 Å resolution in two passes using the orthorhombic crystal form. A 2.07 Å resolution data set was collected in one pass for the lysin dimer using the tetragonal crystal form. Both data sets were indexed and integrated with *MOSFLM* and scaled with *SCALA* (Collaborative Computational Project, Number 4, 1994). Data-collection statistics are given in Table 1.

2.3. Structure solution

The structure of the dimer was solved by molecular replacement with the program *AMoRe* (Navaza & Saludjian, 1997), using the 2.7 Å structure of the red lysin dimer at room temperature (Shaw *et al.*, 1995) as a search model. A cross-rotation function with data in the resolution range 12.0–3.5 Å produced four distinct peaks arising from two solutions and two symmetry-related solutions. These solutions had correlation coefficients of 16.3 and 15.9, consistent with two dimers in the asymmetric unit. A translation function in space group $P4_12_12$ using data in the resolution range 12.0–3.5 Å gave clear solutions for both orientations. Rigid-body refinement resulted in a correlation coefficient of 0.55 and an *R* factor of 0.42 for both individual solutions.

2.4. Model building and refinement

The 1.9 Å structure of the previously solved red lysin monomer at room temperature (Shaw *et al.*, 1993) was used as a starting point for the refinement of the 1.35 Å monomer data set. All refinement calculations were performed with the

*SHELX*97 package (Sheldrick & Schneider, 1997) and all model building was performed with *Xfit* (McRee, 1992). Distance, planarity, chiral volume and anti-bumping restraints were applied from the onset of refinement. The first round of refinement only included data to 2.3 Å. The next round included data to 1.7 Å and the third round of refinement was carried out against all data. Water molecules were added with *Xfit* after the fourth round of refinement. Manual fitting and adjustment of water molecules by inspection of *Xfit* $2|F_o| - |F_c|$ and σ_A-weighted $|F_o| - |F_c|$ maps followed all subsequent rounds of refinement. Anisotropic *B* factors were introduced during the sixth round of refinement and were restrained to have similar values for spatially adjacent atoms. H atoms were added during the seventh round of refinement. A total of 18 rounds of refinement were performed, with the final round including the working (95%) as well as the $R_{\rm free}$ data (5%; Table 1).

Model building and refinement for the dimer were performed as above with *Xfit* and *SHELX*97. The first round of refinement included data to 2.5 Å and the second round was performed against all data. Water molecules were added using *Xfit* after the fourth round of refinement. A total of 12 rounds of refinement were performed, with the final round including all data. Neither H atoms nor anisotropic *B* factors were added to the model.

2.5. Structural superposition and analysis of the monomer and dimer

Structural superpositions and calculations of r.m.s. deviations were carried out using the least-squares-fit option in *SHELXPRO* (Sheldrick & Schneider, 1997). Solvent-accessible areas for individual residues and whole molecules were calculated with *AREAIMOL* and *RESAREA* (Collaborative Computational Project, Number 4, 1994). Analysis of surface features as well as calculations of surface area and distances were performed with *GRASP* (Nicholls *et al.*, 1991). Electrostatic surfaces were calculated with *DELPHI* (Honig & Nicholls, 1995). Contacts between subunits, within subunits and between subunits and solvent molecules were examined with *CONTACT* (Collaborative Computational Project, Number 4, 1994).

The r.m.s. deviations between the previous low-resolution lysin structures and the new high-resolution structures are 1.25 Å for the monomers and 1.55 Å for the dimers.

3. Results and discussion

3.1. Final models

Refinement statistics for the red abalone lysin monomer and dimer are shown in Table 1. The final *R* factor for the monomer is 13.6% and the free *R* factor is 17.8%. The model consists of residues 4–134 of the primary structure, two of which have been modeled in alternate conformations (Arg55 and Lys72). Analysis with *WHATIF* (Hooft *et al.*, 1996) showed that nine hydrogen-bond donors are not involved in hydrogen-bonding interactions; of these, two are unsatisfied buried main-chain hydrogen bonds.

Table 2
R.m.s. deviations (Å) for superposition of lysin molecules.

All values refer to residues 10–134 and are computed with *SHELXPRO* (Sheldrick & Schneider, 1997).

	Subunit *A*	Subunit *B*	Subunit *C*	Subunit *D*
C atoms				
Monomer	0.65	0.52	0.62	0.81
Subunit *A*	—	0.74	0.81	0.56
Subunit *B*	—	—	0.51	0.95
Subunit *C*	—	—	—	0.96
Side chains (all atoms)				
Monomer	1.64	1.57	1.38	1.66
Subunit *A*	—	1.72	1.52	1.15
Subunit *B*	—	—	1.42	1.90
Subunit *C*	—	—	—	1.68

Lysin crystals grown in the $P4_12_12$ space group contain two dimers per asymmetric unit, or four crystallographically independent monomeric subunits. The dimers are designated as containing subunits *AB*, with residues 5–134 (*A*) and 11–134 (*B*), and subunits *CD* with residues 8–134 (*C*) and 7–134 (*D*). Electron density for the loop connecting helices $\alpha2$ and $\alpha3$ (residues 70–80) in subunit *B* is weak. The absent density is a direct result of the extent of bonding between dimers in the lattice, which varies with each monomer subunit. Consequently, subunit *B* has the least number of stabilizing bonds and weaker density, while subunit *D* has the most interactions and the strongest density. This is reflected in the overall *B* factors for the four monomer subunits (Table 1). The final *R* factor and free *R* factor are 21.3 and 28.5%, respectively (Table 1). There are 40 hydrogen-bond donors not involved in hydrogen-bonding interactions, 19 of which are unsatisfied buried main-chain hydrogen bonds.

3.2. Lysin exhibits flexible domains

The four dimer subunits and the monomer (hereafter referred to as 'the five monomers') have very similar overall folds (Figs. 1*a* and 1*b*). When compared, the two least similar molecules have an r.m.s. deviation of 0.96 Å for the C^α atoms and 1.90 Å for the side-chain atoms (Table 2). Superpositions reveal more main-chain variability in the turns and termini, and little variability in the helical regions (Fig. 1*a*). Side-chain positions are more unrestrained, especially external side chains in helices and turns (Fig. 1*a* and Table 2). Residues with higher r.m.s. differences are concentrated in the upper half of the molecule, adjacent to the hypervariable N-terminus, and also along the surface carrying the basic tracks (Fig. 1*a*). Comparison of the *B* factors for residues in the five monomers shows low *B* factors for internal residues and main-chain atoms of surface residues, but higher *B* factors for solvent-exposed residues (Fig. 1*b*). A tendency towards higher *B* factors in the region at the 'top' of the molecule can also be seen. The N- and C-termini are regions of the highest variability in side-chain and main-chain position and exhibit high *B* factors. Electron density for the N-terminus is present to

varying degrees in each of the five monomers, indicating a high extent of structural variability and disorder in these residues. Because the N-terminus adopts a different conformation in all five monomers, it may adopt yet another conformation when bound to VERL. Another variable region is comprised of residues 103–108 located on helix α4 and the turn between helices α4 and α5. These residues have also been shown to be important for species-specific VE dissolution (Lyon & Vacquier, 1999). Therefore, while there is an overall correlation between *B* factors and conformation variability as expected, it is also clear that some parts of the molecule are relatively more flexible than others. These inherently more flexible regions of the lysin molecule could aid in the initial species-specific contacts with VERL by mediating an induced-fit interaction with the receptor.

Side-chain flexibility is exemplified in a split side chain on helix α2. Arg55 is a semi-buried residue which adopts two conformations (I and II) with equal occupancy in the monomer and in dimer subunits *A* and C (Fig. 2). In contrast, in subunits *B* and *D* Arg55 assumes only conformation I. Both conformations of Arg55 are stabilized by three hydrogen bonds to neighboring residues or to waters (Fig. 2). In subunits *B* and *D*, there are additional contacts between Arg55 and neighboring waters. These additional interactions in conformation I may compensate for the loss of interactions in conformation II. Arg55 is a surface residue believed to be involved in bonding interactions with acidic residues on VERL. The ability of this conserved residue to adopt alternate conformations may be relevant to an induced-fit interaction of lysin with VERL.

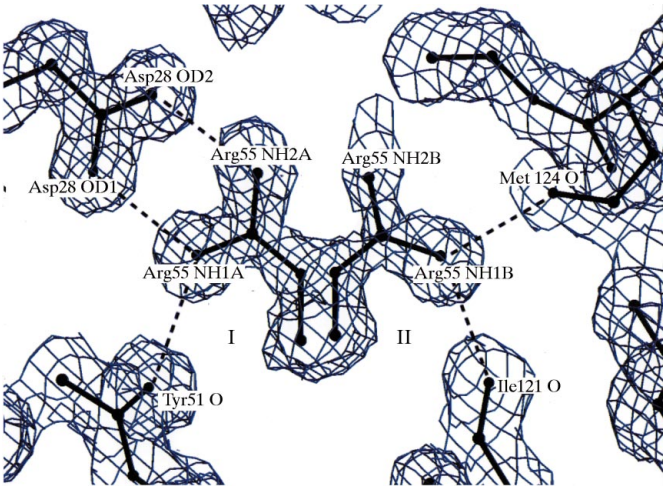

Figure 2
View of Arg55 in the lysin monomer with a $2|F_o| - |F_c|$ electron-density map contoured at 1σ. In the monomer (shown) and subunits *A* and *C*, Arg55 adopts two conformations (I and II). In conformation I, NH1 bonds to Asp28 OD1 (2.9 Å) and Tyr51 O (3.1 Å), while Arg55 NH2 bonds to Asp28 OD2 (2.8 Å). In conformation II, Arg55 NH1 bonds to Met124 O (3.0 Å) and Ile121 O (2.7 Å), while Arg55 NH2 binds a neighboring water molecule (not shown, 3.3 Å). In subunits *B* and *D*, Arg55 adopts only conformation I. This figure was created with *MOLSCRIPT* (Kraulis, 1991) and *BOBSCRIPT* (Esnouf, 1997) and was rendered with *RASTER3D* (Merrit & Murphy, 1994).

3.3. Lysin is an amphipathic molecule

Red abalone lysin has 12 arginines and 13 lysines, making 19% of its residues basic. Of these residues, nine arginines and five lysines are conserved in the lysins of all seven species of California abalone (Lee & Vacquier, 1992). An additional three of these residues are lysine to arginine substitutions, preserving the basic nature of the side chains. Mapping the locations of the basic residues onto the surface of the lysin monomer reveals two distinct tracks of basic residues going down one face of the molecule (Shaw *et al.*, 1993). Comparisons of the locations and sizes of the basic tracks indicate similarities in all five monomers. The average lengths of the tracks are 44.1 (0.4) Å for track 1 and 42.7 (2.2) Å for track 2. Because lysine and arginine have long flexible side chains, their positions in the five monomers are variable in some cases (Fig. 1*a*). This results in different surface patterns for the tracks on different subunits, but an overall similar area covered by the tracks. Track 1 contains an average of 1214.0 (56.9) Å^2 of solvent-accessible area, while track 2 contains 1471.2 (25.1) Å^2.

On the face opposite the basic tracks is a hydrophobic patch of solvent-exposed aromatic and aliphatic residues not normally found on the surface of proteins (Fig. 3). Previously, it was thought that there were only 11 residues involved in this patch (Shaw *et al.*, 1993). However, in light of the new high-resolution monomeric and dimeric structures and surface-area

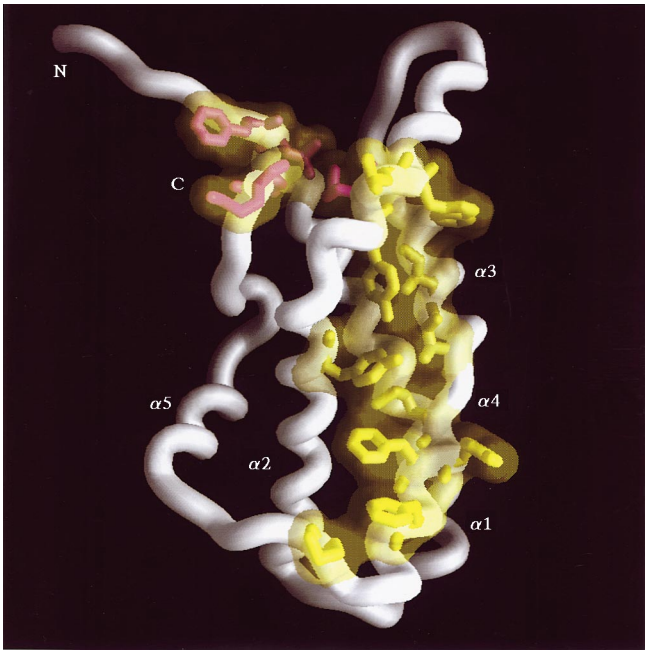

Figure 3
Solvent-accessible surface (yellow) and side chains (yellow and pink) of the aromatic and aliphatic residues forming the hydrophobic patch. The view is rotated by approximately 180° from that in Figs. 1(*a*) and 1(*b*). Side chains are shown (from top to bottom) for new patch residues (pink) Phe10, Leu11, Met66, Tyr133, Met134 and for original patch residues (yellow) Leu67, Trp68, Tyr65, Ile92, Ile96, Tyr57, Met98, Tyr100, Phe101, Phe104, Met110. The main chain is depicted in white. This figure was created with *GRASP* (Nicholls *et al.*, 1991).

calculations, five additional residues have been added to the patch (Fig. 3). The patch includes four tyrosine, two leucine, one tryptophan, two isoleucine, four methionine and three phenylalanine residues. In all five copies, the surface area and length of the patches are similar and the positions of these exposed side chains remain constant. The average solvent-accessible area of the patch is 1360.0 (77.8) Å^2 and its average length is 33.9 (0.9) Å. In the *AB* and *CD* dimers, the patch is largely occluded from solvent by dimerization (Shaw *et al.*, 1993). Of the 16 patch residues, ten are strictly conserved among the seven species and two of the non-conserved residues have undergone substitutions that preserve their hydrophobic nature (Lee & Vacquier, 1992).

The basic tracks and hydrophobic patch are on opposite faces of the lysin monomer. Consequently, lysin is an amphipathic molecule (Shaw *et al.*, 1993). Because both the hydrophobic patch and the basic tracks are conserved in all seven species of California abalone lysin, the amphipathic character of lysin is also preserved. This species conservation suggests that lysin's basic and hydrophobic residues are involved in a general binding interaction with VERL and dissolution of the VE. After initial species-specific binding, the hydrophobic patch and basic track residues could interact with VERL in a conserved manner. The basic residues could stereospecifically sever the inter-VE hydrogen bonds, while the hydrophobic patch residues could bind to the VE glycoproteins, competing for their hydrophobic sites. Together, these interactions would result in the unraveling of the VE fibers.

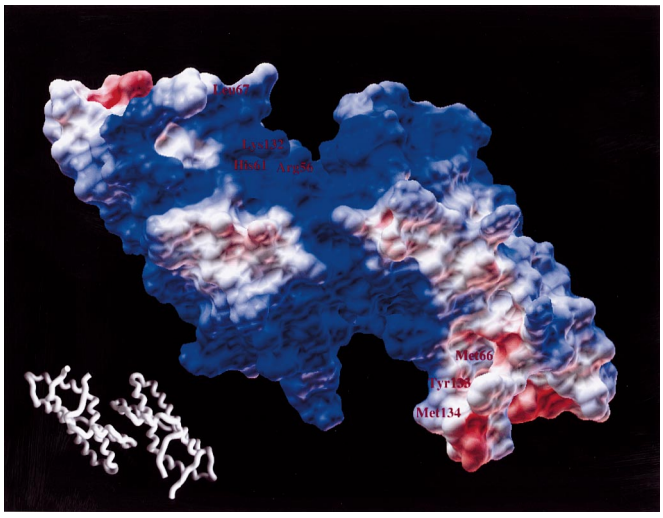

Figure 4
Electrostatic surface of the cleft formed by lysin dimerization. The inset small ribbon diagram shows the orientation of the dimer. The basic tracks are located on the sides of the dimer. The cleft residues are Arg56, Gln59, Thr60, His61, Ala63, Asn64, Met66, Leu67, Lys132, Tyr133 and Met134 from one monomer and Thr41, Leu42, Ser43, Gln46, Lys106, Asn109, Ile111 and Pro112 from the other monomer. The surface is colored according to electrostatic potential. The deepest shades of blue and red correspond to potentials of $+3k_BT/e$ and $-k_BT/e$, respectively, where k_B is the Boltzmann constant, T is the temperature and e is the electronic charge. Neutral points are colored white. This figure was created with *DELPHI* (Honig & Nicholls, 1995) and *AVS* (AVS Inc., Waltham, MA, USA).

Table 3
Comparison of features of the two dimers.

	Dimer *AB*	Dimer *CD*
Dimer interface interactions		
Direct hydrogen bonds	3	5
Water-mediated hydrogen bonds	7	3
Water molecules in interface	15	12
Cleft solvent-accessible area (Å^2)		
Cleft 1	1205	1220
Cleft 2	1135	1238

3.4. A tenuous dimer

To form the dimer, lysin monomers associate *via* a surface interface of basic residues and intercalated hydrophobic residues (Shaw *et al.*, 1995). Dimers *AB* and *CD* have different numbers of intermolecular hydrogen bonds between monomers, suggesting some flexibility in this interface (Table 3). The water molecules in the dimer interface also are distributed unevenly and display a higher concentration nearer the charged residues. Five of the residues involved in dimer interactions differ significantly in position with respect to the monomeric subunits; three of the five are involved in the central stack of residues (Met110, Phe101, Tyr117), indicating a possible dimer-induced conformational change in these side chains. Overall, the main-chain and side-chain positions of the two dimers are comparable: their C$^\alpha$ r.m.s. deviation is 1.35 Å and the r.m.s. deviation for all atoms is 1.99 Å.

From analysis of the van der Waals surface of the interface residues, two features become apparent. The first is the presence of a pocket in the subunit interface. This pocket contains residues Phe52 and Pro112 of both subunits and results in a hollow region in the center of the interface that is

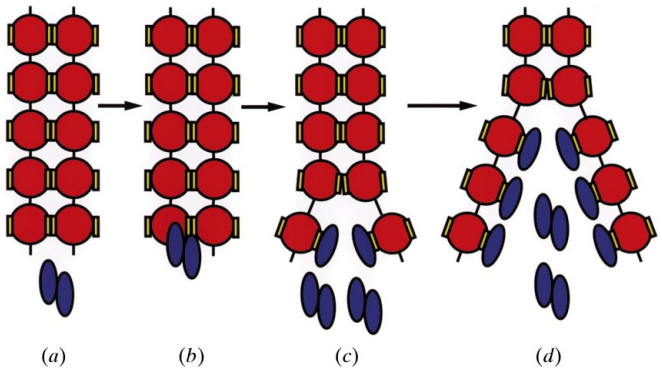

Figure 5
Model for the interaction of lysin with VERL. The lysin dimer is depicted as two blue ellipses. Each VERL repeat is shown as a red sphere connected to other repeats by primary structure (black sticks). Repeats interact with other repeats on neighboring VERL molecules through hydrophobic interactions and hydrogen bonds (yellow rectangles). (*a*) A lysin dimer contacts one VERL repeat. (*b*) The dimer binds to a VERL repeat *via* its cleft residues. (*c*) Binding to VERL causes lysin to monomerize and results in further interactions involving lysin's basic and hydrophobic residues. (*d*) The tight binding between lysin monomers and VERL repeats exposes other repeats to interactions with lysin dimers, allowing steps (*a*) through (*d*) to recur, resulting in an 'unzipping' of the VE fibers.

occupied by six water molecules in both dimers *AB* and *CD*. The second feature is the presence of four positively charged residues in the interface. The charges on Arg56, Lys108, Lys113 and, perhaps, His61 result in a basic dimer interface. Owing to the absence of negatively charged residues, the interaction between subunits must therefore be less favorable, in spite of the large number of hydrophobic contacts. These observations explain the low K_D for the dimer and the short half-time for monomer interchange. An easily dissociated dimer is also consistent with the overall average *B* factors (Table 1). Lysin is known to monomerize upon binding to VERL (Shaw *et al.*, 1995). Presumably, hydrophobic interactions with VERL in combination with the aforementioned interactions between lysin's basic tracks and acidic VERL domains promote dissociation of the dimer and interaction of the lysin monomer with VERL.

3.5. A positively charged cleft could act as an initial binding site

Dimerization causes the formation of two clefts on opposite sides of the local twofold axis (Fig. 4). Having similar size, shape and residue composition (Table 3), the clefts are approximately 20 Å wide and 13 Å deep, and the average solvent-accessible surface of each cleft is roughly 1200 Å^2, or 8% of the total dimer solvent-accessible surface (Fig. 4). Each cleft consists of 11 residues from one subunit and eight residues from the other. Residues from helix α2 and the N-terminus are contributed by one subunit, while residues in the turns between helices α1 and α2 and helices α4 and α5 are provided by the other subunit. One side of each cleft is bordered by the hypervariable N- and C-termini. These regions are among the most flexible in terms of *B*-factor and r.m.s. differences (Figs. 1*a* and 1*b*). Of the 19 cleft residues in the seven species of California abalone, 60% are absolutely conserved, while 40% have undergone substitution (Lee & Vacquier, 1992). 25% of the solvent-accessible surface of the cleft consists of the basic residues: Arg56, His61 and Lys132 from one subunit and Lys106 from the other subunit (Fig. 4). Two of these basic residues (Arg56 and His61) are also involved in the dimer interface, resulting in a basic area that extends from the cleft into the interface. Arg56 and His61 are conserved in all seven species of California abalone (Lee & Vacquier, 1992). The cleft also contains four residues from the hydrophobic patch (Met66, Leu67, Tyr133 and Met134), each donated by the same subunit, that together contribute 31% of the total solvent-accessible surface of the cleft.

The conserved and flexible natures of the residues in the cleft, its overall positive charge and the proximity of the hypervariable N- and C-termini suggest involvement of the cleft in an initial species-specific binding of the lysin dimer to VERL. Because residues in the cleft are contributed from different parts of the dimers, an asymmetric interaction can occur between one VERL repeat domain and dimeric lysin, consistent with the observed stoichiometry of two lysins bound per VERL repeat (Swanson & Vacquier, 1997). This binding interaction could involve the variable residues in the cleft and on the adjacent N- and C-termini, and could be stabilized by electrostatic and van der Waals contacts between the negatively charged VERL and the positively charged cleft.

3.6. A model for the interaction of lysin with VERL

The information obtained from the high-resolution monomeric and dimeric structures of red abalone lysin suggests the following model for the interaction of lysin with VERL (Fig. 5). The lysin dimer is released from the acrosome and contacts a VERL repeat in the VE (Fig. 5*a*). The flexible hypervariable N-terminus mediates an initial species-specific binding interaction between the dimer and the VERL repeat (Fig. 5*b*). The repeat may also bind to the lysin C-terminal residues in a species-specific manner. These initial interactions and the proximity of the VERL repeat to the lysin dimer cleft allow VERL's acidic residues to interact with one of the basic clefts of the dimer. This interaction creates a more negatively charged environment near the dimer interface and promotes interaction of the basic residues within the dimer interface with negatively charged VERL residues. This interaction further destabilizes the dimer interface and leads to dissociation of the dimer and formation of new contacts between lysin monomers and VERL (Fig. 5*c*). These contacts would be non-species-specific and would involve the residues in the conserved hydrophobic patch and basic tracks. Consequently, the second step of lysin binding, as a monomer, is very tight and essentially irreversible (Lewis *et al.*, 1982; Swanson & Vacquier, 1997). This tight binding disrupts the interactions between VERL and other components of the VE, splaying apart VE fibers and exposing more VERL repeats to interactions with free lysins (Fig. 5*d*). The end result is the formation of a hole in the VE. This mechanism accounts for the cooperative nature of lysin–VERL contacts as observed in VE dissolution assays (Hill coefficient approximately 3; Swanson & Vacquier, 1997). Further, it is consistent with available kinetic (Shaw *et al.*, 1995) and biochemical (Hong & Vacquier, 1986; Lewis *et al.*, 1982; Swanson & Vacquier, 1997) data for the non-enzymatic VE dissolution reaction. Finally, it agrees with the biochemical properties of the component molecules, lysin and VERL, and the structural features of the lysin monomer and dimer.

This work is based upon research conducted at the Stanford Synchrotron Radiation Laboratory (SSRL), which is funded by the Department of Energy, Office of Basic Energy Sciences. The Biotechnology Program is supported by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program and the Department of Energy, Office of Biological and Environmental Research. We would like to thank Peter Kuhn, Mike Soltis and Aina Cohen for their help at the SSRL beamline 9–1, Pamela Williams for generous assistance with data collection and G. Sridhar Prasad for discussions and advice. This work was supported by NSF grant MCB-9816426 to CDS and NIH grant HD12986 to VDV.

References

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 132–134.
Hong, K. & Vacquier, V. D. (1986). *Biochemistry*, **25**, 543–549.
Honig, B. & Nicholls, A. (1995). *Science*, **268**, 1144–1149.
Hooft, R. W. W., Vriend, G., Sander, C. & Abola, E. E. (1996). *Nature (London)*, **381**, 272.
Kraulis, P. (1991). *J. Appl. Cryst.* **24**, 946–950.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **24**, 946–950.
Lee, Y.-H. & Vacquier, V. D. (1992). *Biol. Bull.* **182**, 97–104.
Lewis, C. A., Talbot, C. F. & Vacquier, V. D. (1982). *Dev. Biol.* **92**, 227–239.
Lindberg, D. R. (1992). *Abalone of the World: Biology, Fisheries and Culture*, edited by S. A. Shepherd, M. J. Tegner & S. A. Guzman del Proo, pp. 3–18. Oxford: Blackwell Scientific.
Lyon, J. D. & Vacquier, V. D. (1999). *Dev. Biol.* **214**, 151–159.
McRee, D. E. (1992). *J. Mol. Graph.* **10**, 44–47.
Merrit, E. A. & Murphy, M. E. P. (1994). *Acta Cryst.* D**50**, 869–873.
Navaza, J. & Saludjian, P. (1997). *Methods Enzymol.* **276**, 581–594.
Nicholls, A., Sharp, K. & Honig, B. (1991). *Proteins*, **11**, 281–296.
Owen, B., McLean, J. H. & Meyer, R. J. (1971). *Bull. Los Angeles Mus. Nat. Hist. Sci.*, pp. 1–37.
Shaw, A., Fortes, P. A. G., Stout, C. D. & Vacquier, V. D. (1995). *J. Cell Biol.* **130**, 1117–1125.
Shaw, A., Lee, Y.-H., Stout, C. D. & Vacquier, V. D. (1994). *Semin. Dev. Biol.* **5**, 209–215.
Shaw, A., McRee, D. E., Vacquier, V. D. & Stout, C. D. (1993). *Science*, **262**, 1864–1867.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **276**, 319–343.
Swanson, W. J., Metz, E. C., Stout, C. D. & Vacquier, V. D. (1999). *The Male Gamete*, edited by C. Gagnon, pp. 140–147. Vienna: Cache River Press.
Swanson, W. J. & Vacquier, V. D. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 6724–6729.
Swanson, W. J. & Vacquier, V. D. (1998). *Science*, **281**, 710–712.
Vacquier, V. D. & Lee, Y.-H. (1993). *Zygote*, **1**, 1–16.
Vacquier, V. D., Swanson, W. J., Metz, E. C. & Stout, C. D. (1999). *Adv. Dev. Biochem.* **5**, 49–81.