doi:10.1006/jmbi.2000.3533 available online at http://www.idealibrary.com on IDEAL® *J. Mol. Biol.* (2000) **296**, 1225–1234

JMB

The High Resolution Crystal Structure of Green Abalone Sperm Lysin: Implications for Species-specific Binding of the Egg Receptor

Nicole Kresge[1], Victor D. Vacquier[2] and C. David Stout[1]*

[1]*Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037-1093, USA*

[2]*Center for Marine Biotechnology and Biomedicine Scripps Institution of Oceanography, University of California San Diego, La Jolla CA, 92093-0202, USA*

**Corresponding author*

Abalone sperm lysin is a 16 kDa acrosomal protein used by sperm to create a hole in the egg vitelline envelope. Lysins from seven California abalone exhibit species-specificity in binding to their egg receptor, and range in sequence identity from 63 % to 90 %. The crystal structure of the sperm lysin dimer from *Haliotis fulgens* (green abalone) has been determined to 1.71 Å by multiple isomorphous replacement. Comparisons with the structure of the lysin dimer from *Haliotis rufescens* (red abalone) reveal a similar overall fold and conservation of features contributing to lysin's amphipathic character. The two structures do, however, exhibit differences in surface residues and electrostatics. A large clustering of non-conserved surface residues around the waist and clefts of the dimer, and differences in charged residues around these regions, indicate areas of the molecule which may be involved in species-specific egg recognition.

Keywords: abalone sperm lysin; fertilization; gamete recognition proteins; species-specificity; X-ray crystallography

Introduction

Recognition and binding interactions between proteins on the surface of the sperm and the egg are essential for successful fertilization. Events mediated by gamete recognition proteins include: adhesion of the sperm to the egg envelope, induction of the acrosome reaction of sperm, penetration of the egg envelope by the sperm, attachment of the sperm to the egg plasma membrane, and fusion of the sperm and egg cell membranes (Vacquier, 1998). Many of these interactions exhibit species-specificity, meaning fertilization between conspecific gametes occurs more readily than between heterospecific gametes. In both mammals and echinoderms, the greatest barrier to cross-species fertilization is failure of the sperm to bind to the egg envelope (O'Rand, 1988; Roldan & Yanagimachi, 1989; Summers & Hylander, 1975, 1976; Yanagimachi, 1994). Experiments involving marine invertebrates have led to the hypothesis that this species-specificity is due to the rapid evolution of gamete recognition proteins (Biermann, 1998; Hellberg & Vacquier, 1999; Lee *et al*., 1995; Metz *et al*., 1998; Palumbi & Metz, 1991; Vacquier *et al*., 1995).

Abalone sperm lysin is a 16 kDa protein that mediates the attachment of the sperm to the egg vitelline envelope (VE). Lysin is contained in the sperm acrosomal vesicle and is released during the acrosome reaction. Upon its release, lysin binds to the vitelline envelope receptor for lysin (VERL) and, through a non-enzymatic mechanism, disrupts contacts between the VE fibers, creating a hole through which the sperm swims to fuse with the egg plasma membrane (Lewis *et al*., 1982; Swanson & Vacquier, 1997). VERL is a 1000 kDa gamete recognition glycoprotein which is rich in hydrophobic and acidic residues. VERL contains approximately 26 tandem repeats of a 153 amino acid lysin-binding domain. Roughly two molecules of lysin bind each VERL repeat, resulting in an average of about 50 lysins bound per VERL molecule (Swanson & Vacquier, 1997).

The seven species of abalone which live off the coast of California have overlapping breeding seasons and habitats, yet hybrids are rare (Lindberg, 1992; Owen *et al*., 1971). Lysins from these different abalone exhibit species-specificity in binding to and dissolving VEs, and range in sequence identity

Abbreviations used: BSA, bovine serum albumin; VE, vitelline envelope; VERL, vitelline envelope receptor for lysin.

E-mail address of the corresponding author: dave@scripps.edu

0022-2836/00/051225–10 $35.00/0

from 65% to 90% (Lee & Vacquier, 1992; Shaw *et al*., 1994; Vacquier & Lee, 1993). The most variable domain in lysin is the N-terminal residues 1-12 (Figure 1). This domain has been shown to be important in species-specific VE recognition and dissolution (Lyon & Vacquier, 1999). Analysis of amino acid substitutions shows most replacements between any two lysins are non-conservative (Lee & Vacquier, 1992). A high ratio of non-synonymous to synonymous nucleotide substitutions indicates there is adaptive value in diversifying the protein's sequence, and that the lysin sequences are rapidly diverging due to positive Darwinian selection (Metz *et al*., 1998). The selective force for this diversification of lysin may be concerted evolution in the multi-domain VERL (Swanson & Vacquier, 1998).

The crystal structure of the monomeric red abalone lysin reveals an amphipathic, α-helical protein with three distinct features: (1) two parallel tracks of positively charged residues run the length of one side of the molecule, 56% of the amino acids in this basic track are conserved among the seven California species; (2) an hydrophobic patch exists on the molecular face opposite the basic tracks, 63% of these residues are conserved among the seven California species; and (3) the species-specific N-terminal domain extends away from the core of the molecule (Kresge *et al*., 2000; Shaw *et al*., 1993). In solution, lysin exists as a dimer held together by the surface cluster of hydrophobic residues, placing the basic tracks on opposite sides of the dimer (Shaw *et al*., 1995). This interaction optimizes the packing of lysin in the acrosome, allows delivery into an aqueous environment, and prevents non-specific association of the hydrophobic patch with the VE. Once the dimer contacts the VE, it binds to VERL and monomerizes (Kresge *et al*., 2000; Shaw *et al*., l995). The exposed hydrophobic patch and basic tracks are expected to bind irreversibly to the VE glycoproteins and acidic residues, causing the VE fibers to unravel and remain separated by hydrophilic regions of lysin (Kresge *et al*., 2000; Swanson & Vacquier, 1997).

Figure 1. Sequence alignment for red and green abalone lysins. Basic track residues for each species are highlighted in blue, and hydrophobic patch residues are highlighted in yellow. Green bars indicate residues that differ between red and green lysins while residues outlined in green are contained in the waist or cleft regions.

To demonstrate the species-specificity of lysin-VE interactions, we quantitatively assayed the ability of red and green lysins to solubilize VEs from red and green abalone eggs. In order to investigate the relationship between lysin structure and species-specificity, we determined the three-dimensional structure of the lysin dimer from green abalone to 1.71 Å. We then compared this structure to the 2.07 Å crystal structure of the red abalone lysin dimer (Kresge *et al*., 2000). Green abalone lysin is the most divergent from red abalone lysin, with only 65% sequence identity. Comparison of the two structures establishes for the first time the molecular details of species-specific interactions between gamete recognition proteins.

Results

Species-specific VE dissolution

Although both red and green lysins are able to dissolve red abalone VE, the red lysin is much more effective (Figure 2(a)). Some 50% dissolution of the red abalone VE occurred after exposure to red lysin for eight minutes, while it took 18.2 minutes of exposure to an equivalent amount of green lysin to achieve 50% dissolution. With green abalone VE, eight times as much lysin was needed to achieve dissolution. Green lysin dissolved 50% of green abalone VE by 12.8 minutes, while red lysin was inactive at equivalent concentrations (Figure 2(b)).

Final model

Crystallography statistics for the green abalone lysin dimer structure are summarized in Tables 1 and 2. The structure is comprised of two monomeric subunits: subunit A contains residues 11-132, while subunit B contains residues 11-134. There was no density present for residues 1-10 in subunits A and B and residues 133-134 in subunit A. Coordinates and structure factors for the dimer refined at 1.71 Å resolution have been deposited into the PDB with accession codes 3lyn and r3lynsf.

Overall fold description and comparison to red abalone sperm lysin structure

Despite the difference in sequence between the red and green abalone lysin dimers, their overall size and solvent accessible surfaces are very similar, with the green abalone lysin being only slightly smaller than the red (Table 3). This size discrepancy is due in part to the fact that the N-terminal

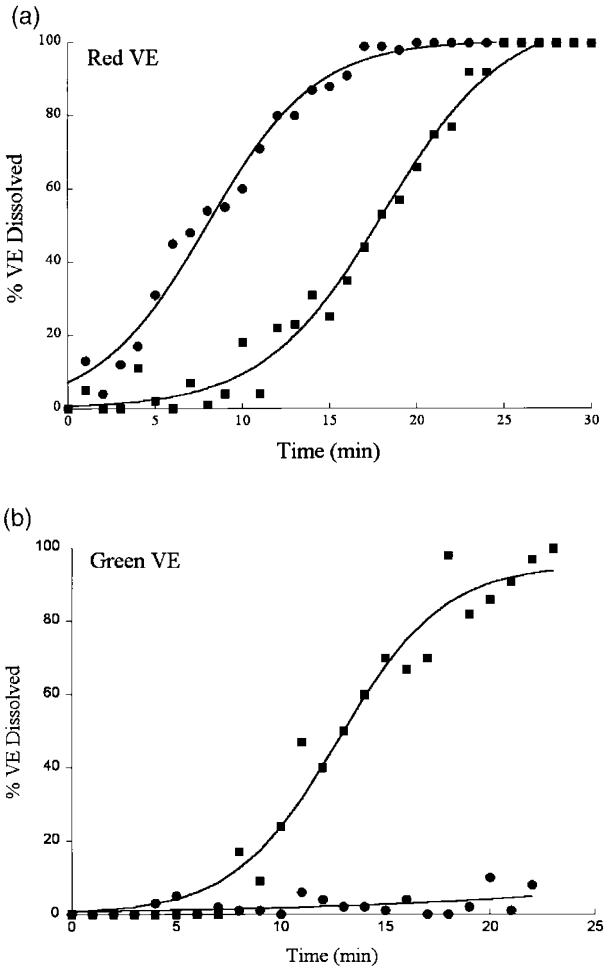

Figure 2. Time course of red and green abalone VE dissolution by green and red abalone lysins. (a) Red abalone VE with 2.15 μg of red (●) and green (■) abalone lysin added at two minute intervals. (b) Green abalone VE with 17.5 μg of red (●) and green (■) abalone lysin added at two minute intervals.

residues 1-10 are missing in the green abalone lysin structure but are present to some extent in the red abalone lysin structure. The overall folds of the lysin monomers from both species are very similar. Both contain five α-helices which form an up-down helical bundle with a right-handed twist (Figure 3). In both species the dimer assumes an extended S-shape with two clefts on either side of the 2-fold axis which relates the monomeric subunits (see Figure 5). Dimerization causes approximately 12% of each green abalone lysin monomer to be buried, while only 9% of each red abalone lysin monomer is buried (Table 3).

Superpositions of red and green abalone lysin monomer subunits have an average rms deviation of 1.14 Å for C^{α} atoms and 1.74 Å for all atoms (Table 4A). The rms deviation for the C^{α} atom superposition of the two dimers is larger (2.03 Å on average) (Table 4B), due to a slight shift in monomer association. The termini and turns show the most variability in main-chain and side-chain position when the molecules are compared (Figure 3). Although all positionally variable residues occur on the surface, there is no concentration of these residues in any one particular area.

The hydrophobic patch

Both the red and green abalone lysin monomers contain a patch of solvent-exposed hydrophobic residues that extends the length of the molecule (Figure 4(a)). This patch is much larger on red abalone lysin than on green abalone lysin (Table 3). The patches are, however, similar in shape and location in both species. Nine of these hydrophobic residues are conserved between the two species, while three others have undergone conservative substitutions (Figure 1).

The basic tracks

Red abalone contains 24 arginine and lysine residues comprising 18% of its total residues, while green abalone lysin has 26 arginine and lysine residues that make up 19% of its total residues (Figure 1). On the face opposite the hydrophobic patch on both the red and green abalone lysins are two parallel tracks of basic residues (Figure 4(b)). Track 1 occurs in approximately the same position in both lysins and is similar in size in both species (Table 3). However, track 2 differs in shape when the monomers are compared (Table 3). Although the total number of residues and surface areas are similar (Table 3), track 2 is more scattered in green abalone lysin, and there are many small differences that cause the overall morphology of track 2 to differ between the two species.

The dimer interface

The monomeric subunits of the red and green abalone dimers associate *via* five intercalating hydrophobic patch residues (Tyr57, Tyr/Phe100, Phe101, Phe104, and Met110) and a small number of charged residues. This association results in the burying of 38% of the hydrophobic patch residues in the green dimer and 30% in the red dimer. Other residues involved in stacking interactions and hydrogen bonds surround these central stacked residues. Both species' dimers contain a total of 14 stacking interactions, and similar numbers of water molecules in their interfaces, but the red abalone lysin dimer has fewer hydrogen bonds between its subunits (Table 3). Overall, there are more contacts in the green abalone lysin dimer interface, consistent with the larger solvent accessible surface that is occluded (Table 3).

The dimer clefts

Dimerization causes the formation of clefts on opposite sides of the 2-fold axis relating the subunits in both the red and green abalone lysin

Table 1. Data collection and phasing statistics (calculated to 3.6 Å resolution)

Data set	Native 1	Native 2	$KAuI_4$	$U(NO_3)_2$
Cell dimensions				
a (Å)	119.17	121.37	121.34	121.17
b (Å)	119.17	121.37	121.34	121.17
c (Å)	51.29	52.32	52.27	52.46
Resolution (Å)	1.71	3.53	3.63	3.65
Multiplicity[a]	4.4(3.2)	9.1(1.7)	13.0(11.5)	11.8(5.1)
$I/\sigma(I)$	9.7(1.9)	7.1(2.10)	5.6(2.2)	4.5(2.3)
Completeness (%)	97.8(92.1)	94.5(68.5)	97.1(97.0)	94.0(82.5)
R_{symm} (%)[b]	5.1(38.7)	9.2(16.0)	14.2(12.5)	11.6(17.5)
Temperature (K)	100	291	291	291
A. *Derivatives*				
Concentration (mM)	-	-	1	50
Soak (days)	-	-	7	4
Number of sites	-	-	2	3
R_{iso} (%)[c]	-	-	19.8	22.5
R_{cullis} (%)[d]	-	-	67.1	60.7
Phasing power[e]	-	-	1.80	2.22
Mean figure of merit	-	-	0.50	0.52

[a] Values for the highest resolution shell are given in parentheses.
[b] $R_{symm} = 100 \times \Sigma_h \Sigma_j |I_{hj} - I_h| / \Sigma_h \Sigma_j I_{hj}$, where I_h is the weighted mean intensity of the symmetry related reflections I_{hj}.
[c] $R_{iso} = 100 \times \Sigma_{hkl} |F_P - F_{PH}| / \Sigma_{hkl} F_P$ where F_P is the native amplitude and F_{PH} is the derivative amplitude.
[d] $R_{cullis} = 100 \times \Sigma_{hkl} \mid |F_P \pm F_{PH}| - F_{H,calc}| / \Sigma_{hkl} |F_P \pm F_{PH}|$, where F_P is the native amplitude, F_{PH} is the derivative amplitude for centric reflections and $F_{H,calc}$ is the calculated heavy atom structure factor.
[e] Phasing power $= F_H/E$ where E is the estimated lack of closure error.

dimers (Figure 5). A fairly narrow waist region (30 Å wide) connects the clefts (Figure 5). The clefts in both species consist of 11 residues from one subunit and eight residues from the other subunit, are of similar size, and have similar solvent accessible surfaces (Table 3). Of the cleft residues, 37 % differ between the two species (Figures 1 and 5(a)). The cleft residues are also more variable in position, resulting in different-shaped clefts in the two species (Figures 3 and 5).

Table 2. Refinement statistics

R-factor (%)[a]	21.5
R_{free} (%)[b]	25.9
Number of reflections	44,062
Number of non-hydrogen atoms	2271
Number of water molecules	218
Average *B*-factor (Å^2)[c]	
Protein, overall	28.4
Main-chain	24.4(3.0)
Side-chain	32.0(5.8)
Water	41.2
rms deviations from ideal geometry[d]	
Bond distance (Å)	0.009
Bond angle (deg.)	1.89
Ramachandran statistics[e]	
Most favored regions (%)	96.3
Additionally allowed regions (%)	3.7
Disallowed regions (%)	0.0

[a] $R\text{-factor} = \Sigma||F_{obs}| - |F_{calc}|| / \Sigma F_{obs}$, where $|F_{obs}|$ and $|F_{calc}|$ are the observed and calculated structure factor amplitudes, respectively.
[b] R_{free} is the *R*-factor for a 5 % randomly selected subset of reflections that were not used in refinement.
[c] rms deviations in *B*-values are indicated in parentheses.
[d] As calculated in the program WHATIF (Hooft *et al.*, 1996).
[e] As calculated in the program PROCHECK (Laskowski *et al.*, 1993).

Spatial distribution of non-conserved residues

A total of 48 out of 134 residues are not conserved between red and green abalone lysins (Figure 1). Some 96 % of these residues are located on the surface of the lysin dimer (Figures 3 and 5(b)). Analysis of the locations of non-conserved residues on the green lysin monomer reveals an even distribution about the molecule (Figure 3). However, when the locations of the same residues are viewed on the green abalone lysin dimer, a clustering around the waist and cleft of the dimer becomes visible (Figure 5). Some 38 % of the residues not conserved between red and green abalone lysins are located in the waist and cleft regions, an area which represents only 8-9 % of the total surface area (Table 3). Similarly, residues 103-108, shown to be important in species-specific dissolution of VEs (Lyon & Vacquier, 1999), are located around the waist of the dimer. Thus, although the non-conserved residues are scattered throughout the primary and tertiary structures (Figures 1 and 3), they are clustered in the quaternary structure. A trend towards a loss of acidic residues in going from red lysin to green lysin, and also a gain of polar aromatic residues in green lysin is also apparent (Figure 1). The effects of this loss of negatively charged residues in green abalone lysin are seen in the differences between the electrostatic surfaces of red and green abalone lysins (Figure 5(b)) *versus* Figure 6(a) and (b)).

Electrostatics of lysin

The electrostatic surface of green abalone lysin reveals that it is a very basic molecule with a few scattered patches of negative and neutral residues

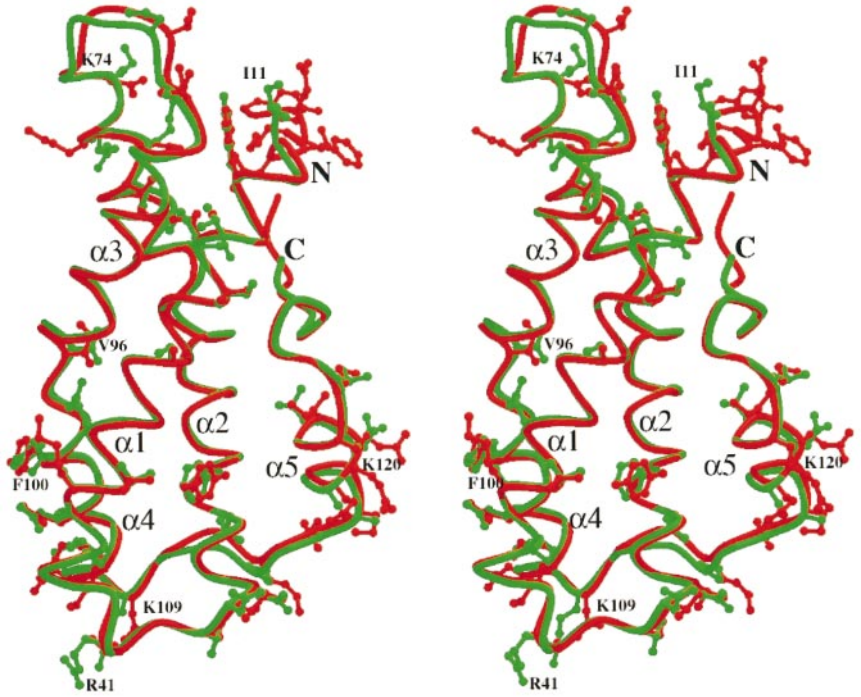

Figure 3. Stereo diagram of the green and red abalone lysin monomer α-carbon chain superposition. Green lysin dimer subunit A is green and red lysin dimer subunit A is red. Side-chains for the 48 residues not conserved between the two species are shown in red on red abalone lysin and in green on green lysin. Non-conserved residues contained in the basic tracks or hydrophobic patch are labeled. These Figures were created with MOLSCRIPT (Kraulis, 1991) and BOBSCRIPT (Esnouf, 1997) and rendered with RASTER3D (Merrit & Murphy, 1994).

(Figure 6(a)). Red lysin, however, is a less basic molecule with fairly significant patches of negative potential and several areas of neutral surface (Figure 6(b)). When viewed along the 2-fold axis relating the dimer subunits, two small patches of negative potential are seen adjacent to the waist of the green abalone lysin dimer (Figure 6(a), left-hand side); these regions are larger in the red abalone lysin dimer (Figure 6(b), left-hand side). When the molecule is viewed from the opposite side, the exposed face of the red lysin dimer reveals patches of neutral residues scattered across the molecular

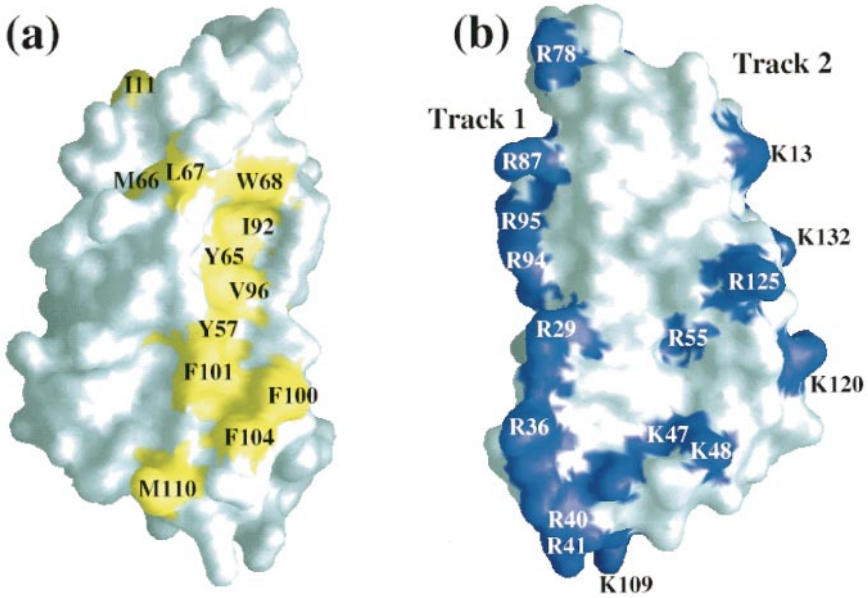

Figure 4. Molecular surface of the green abalone lysin dimer subunit A showing the hydrophobic patch and basic tracks. (a) Surface containing the hydrophobic patch. View is rotated 180° from that shown in Figure 3. In yellow from top to bottom are Ile11, Met66, Leu67, Trp68, Ile92, Tyr65, Val96, Tyr57, Phe101, Phe100, Phe104, and Met110. (b) 180° rotation from (a) displaying the surface containing the basic tracks. The ten residue track 1 (left) contains residues (blue) Arg78, Lys74, Arg71, Lys70, Lys72, Arg87, Arg95, Arg94, Arg29, and Arg36. Track 2 (right) is 13 residues long and consists of Lys13, Lys132, Lys20, Arg125, Arg55, Arg56, Lys120, Lys48, Lys47, Arg40, Arg41, Arg108, and Lys109 (in blue). This Figure was created with GRASP (Nicholls *et al.*, 1991).

Table 3. Comparisons of features of red and green abalone sperm lysins

	Green lysin	Red lysin[a]
Overall dimensions (Å)		
Monomer	42 × 33 × 49	45 × 35 × 55
Dimer	35 × 43 × 77	46 × 50 × 76
Solvent accessible surface (Å^2)		
Monomer	7585	7860
Dimer	13,342	14,380
Dimer interface[b]		
Total protein surface buried upon dimerization (Å^2)	1828	1340
Water molecules	14	15/12
Water-mediated hydrogen bonds	9	7/3
Stacking interactions	14	14/14
Hydrogen bonds	12	3/5
Cleft solvent accessible surface[a] (Å^2)	1245(9 %)	1200(8 %)
Hydrophobic patch		
Residues involved	12	16
Solvent accessible surface (Å^2)	848(11 %)	1373(17 %)
Basic track 1		
Residues involved	10	9
Solvent accessible surface (Å^2)	1391(18 %)	1191(15 %)
Basic track 2		
Residues involved	13	14
Solvent accessible surface (Å^2)	1593(21 %)	1475(%)

[a] Calculations are an average of two red abalone lysin dimers AB and CD (Kresge *et al.*, 2000).
[b] Numbers for red dimer are shown as red dimer AB/red dimer CD.
[c] Numbers in parentheses are percentages of total solvent accessible surface.

surface (Figure 6(b), right-hand side); in contrast, the corresponding face of the green abalone lysin dimer contains significantly smaller patches of neutral residues (Figure 6(a), right-hand side). In both species there is a significant concentration of basic residues within the clefts. In the green abalone lysin, 58 % of the solvent accessible surface of the cleft consists of basic residues Arg56A, His61A, Lys132A, His134A, Arg41B, and Lys109B. In the red abalone lysin dimer there are four basic residues, Arg56A, His61A, Lys132A and Lys106B, in the cleft.

Discussion

The inability of red lysin to dissolve green abalone VE, and the inefficiency of green lysin in dissolving red abalone VE, demonstrates the species-specific interaction of the two lysins with the VE. The fact that red abalone VEs are more easily dissolved than green abalone VEs is consistent with observations from hybridization experiments using freshly spawned abalone sperm and eggs. Leighton & Lewis (1982) found that homospecific fertilization rates were lower for green abalone than for red abalone, and that the most successful heterologous crosses occurred between red abalone eggs and pink, green, or white abalone sperm. Thus, different VEs display different degrees of solubility, with red abalone VEs being extremely soluble and green abalone VEs insoluble to any lysins but their own.

Comparisons of the three-dimensional structures of red and green abalone lysins reveal which features of the molecule are structurally conserved and which are not. Since the species-specificity of lysin is a result of differences between lysin molecules, traits such as electrostatics and residue distribution must contribute to lysin's species-specific binding to VERL. Shared features such as the hydrophobic patch and the basic tracks, on the other hand, are involved in the universal mechanism used by lysins to dissolve VEs.

When the lysin dimer is released from the acrosomal vesicle, its initial interactions with VERL are species-specific. From the comparisons of red and green lysin it can be inferred that much of this initial interaction occurs with the clefts and around the waist of the lysin dimer where there is a high concentration of species-specific residues. This species-specific interaction may be mediated by

Table 4. Comparisons of red and green abalone lysin superpositions

A. *rms deviations (Å) for superpositions of green and red abalone lysin subunits*

	Green B	Red A[a]	Red B	Red C	Red D
C^α atoms					
Green A	1.00	0.91	1.00	1.08	1.06
Green B	-	1.14	1.32	1.31	1.42
All atoms					
Green A	1.69	1.57	1.91	1.76	1.76
Green B	-	1.62	1.98	1.68	1.73

B. *rms deviations (Å) for superpositions of green and red abalone lysin dimers*[b]

	Red AB[c]	Red CD
C^α atoms		
Green AB	1.87	2.19
All atoms		
Green AB	2.26	2.44

[a] Letters refer to the subunits of the red abalone lysin dimer as defined by Kresge *et al.* (2000).
[b] Comparisons were made using residues 11-132 of chain A and 11-134 of chain B for both the red abalone and green abalone dimers.
[c] Letters refer to the subunits of the red abalone lysin dimer as defined by Kresge *et al.* (2000).

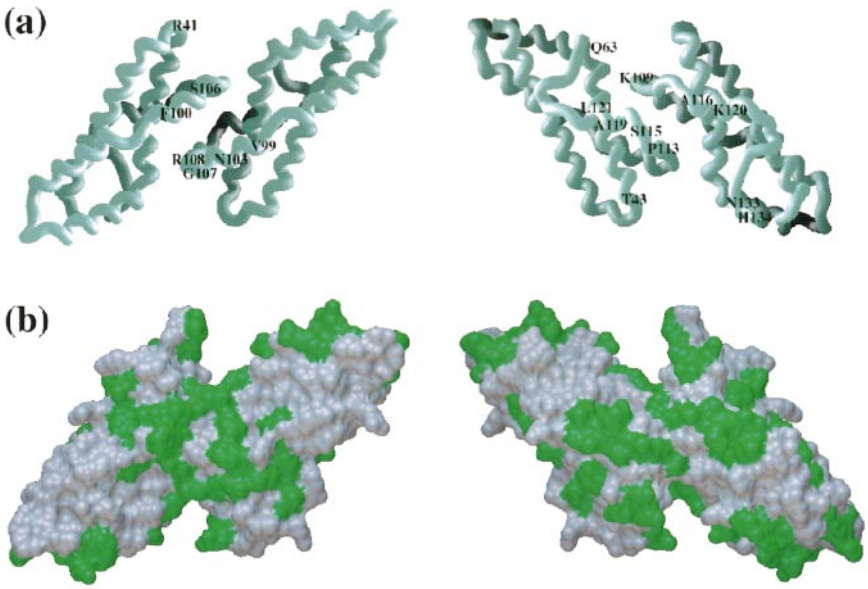

Figure 5. Sequence conservation between red and green abalone lysins. (a) The α-carbon trace showing the orientations of green abalone lysin dimer in (b). The view on the left is rotated about the vertical axis 180° from the view on right. Non-conserved cleft and waist residues are labeled. (b) Surface of the green abalone lysin dimer showing residues (green) which differ between red and green abalone lysin. This Figure was created with AVS (AVS Inc. Waltham, MA).

attractive forces between the basic residues that also occupy these regions and acidic residues in VERL. VERL from red abalone contains 15 glutamic and aspartic acid residues, while VERL from green abalone contains 17 glutamic and aspartic acid residues, making green VERL more acidic than red VERL (Swanson & Vacquier, 1998). Thus, the more acidic green VERL will interact less favorably with the red lysin dimer, due to its negatively charged waist area inhibiting species-specific contacts (Figure 6(b), left-hand side). On the other hand, the more basic green lysin dimer can interact more favorably with the more acidic green VERL, allowing close interaction and species-specific pairings of residues to occur. Consequently, both attractive and repulsive interactions between lysin and VERL facilitate or prevent species-specific binding of lysin to VERL (Figure 6). At the same time, it is important to note that the species-specific residues overlap in the cleft and waist regions of the dimer where the electrostatic differences occur. Therefore, the electrostatic forces promoting species-specific interaction also promote direct contact of these variable residues with VERL. The high concentration of variable residues around the clefts suggests that VERL may bind directly inside a cleft, and that the different cleft shapes may play a role in mediating species-specific contacts between lysin and VERL.

After initial species-specific binding, lysin monomerizes and VE dissolution is initiated (Kresge *et al*., 2000; Shaw *et al*., 1995). This second step of dissolution is one that is general to all lysins and is presumed to involve the conserved hydrophobic patch and basic tracks. Monomerization of the lysin dimer exposes the hydrophobic patch and allows it to interact with the VE glycoproteins. Meanwhile, the basic residues, in part responsible for the species-specific recognition of VERL, remain bound to the acidic residues on VERL. These cooperative interactions cause the repeating subunits of VERL to separate and the fibers comprising the VE to splay apart, forming a hole through which the sperm swims to fertilize the egg. This general dissolution mechanism also explains how exposure to a large amount of green abalone lysin can result in the dissolution of red abalone VEs. If enough green abalone lysin is added, some will monomerize and bind to red abalone VERL, initiating VE dissolution.

Methods

VE dissolution assay

Red and green abalone VE were isolated as described by Lewis *et al*. (1982) and suspended in seawater containing 1 mg/ml bovine serum albumin (BSA), 10 mM N_3Na, 10 mM Tris (pH 7.8). Red and green abalone lysins were purified by CM-cellulose chromatography (Vacquier & Lee, 1993) and dialyzed against seawater containing 10 mM N_3Na, 10 mM Tris (pH 7.8). For dissolution of red abalone VE, 800 μl of VE suspension was placed in a 1 ml cuvette and 2.5 μl (2.15 μg protein) of either lysin was added at two-minute intervals. For green abalone VE, 20 μl (17.5 μg protein) of either lysin was added every two minutes and the decrease in light scattering at 640 nm was followed (Lyon & Vacquier, 1999).

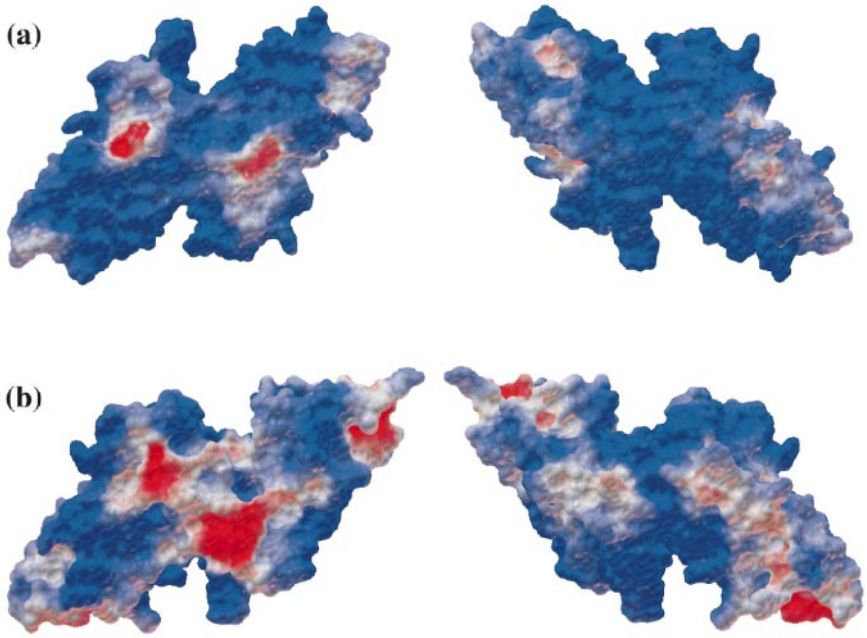

Figure 6. Electrostatic surfaces of the green and red abalone lysin dimers. The surface is colored according to electrostatic potential. The deepest shades of blue and red correspond to potentials of $\geqslant +3\ k_BT/e$ and $\geqslant -3\ k_BT/e$, respectively, where k_B is the Boltzmann constant, T is the temperature, and e is the electronic charge. Neutral points are white. The orientations are the same as described in the legend to Figure 5. (a) Electrostatic surface of the green abalone lysin dimer. (b) Electrostatic surface of the red abalone lysin dimer. This Figure was created with DELPHI (Honig & Nicholls, 1995) and AVS (AVS Inc. Waltham, MA).

Protein preparation and crystallization

Lysin from green abalone was purified by CM cellulose chromatography (Vacquier & Lee, 1993). The protein was dialyzed against 50 mM sodium acetate (pH 4.9) and concentrated to 10 mg/ml. Green abalone lysin crystals were grown at 22 °C using the sitting drop vapor diffusion technique and a reservoir solution of 1.26 M ammonium sulfate, 100 mM sodium acetate (pH 4.0), 200 mM sodium chloride, and 0.01% (w/v) sodium azide. Drops contained 4.0 µl of protein solution and 3.0 µl of precipitant, and wells contained 1 ml of reservoir solution. The crystals produced contained two molecules per asymmetric unit and belonged to space group $P6_1$.

Data collection

Native data were collected at the Stanford Synchrotron Radiation Laboratory (SSRL) beamline 7-1, using a MAR image plate and monochromatic radiation of wavelength 1.08 Å. Low-temperature data were collected by immersing the crystals in mother liquor with 25% (v/v) glycerol prior to freezing at 100 K. A 1.9 Å resolution data set was collected from the crystals. The data set was indexed and integrated with MOSFLM (Collaborative Computational Project, 1994) and scaled with the SCALA (Collaborative Computational Project, 1994) software package. Data collection statistics are shown in Table 1.

Heavy-atom derivatives of the green abalone lysin dimer were prepared by soaking the crystals in mother liquor containing either 1 mM $KAuI_4$ or 50 mM $U(NO_3)_2$ (Table 1). Room temperature native and derivative data sets were collected using CuK^{α} radiation from a Rigaku RU200 rotating anode X-ray generator operated at 40 kV, 80 mA, and equipped with a graphite monochromator. The data were recorded with a Siemens X-1000 (Xentronics) multiwire area detector and processed with the Xengen suite of programs (Howard *et al*., 1985). Data collection statistics for the room temperature native and derivative data are shown in Table 1.

Structure solution

The room temperature native and derivative data sets were used in determining the gold and uranium heavy-atom sites with the program SOLVE (Terwilliger & Berendzen, 1996; Terwilliger & Eisenberg, 1983, 1987; Terwilliger *et al*., 1987). The gold derivative contained two sites and the uranium derivative contained three sites (Table 1). Site refinement and phase calculation were done with SHARP (De La Fortelle & Bricogne, 1997). The resulting electron density map was subjected to solvent flattening with the program SOLOMON (Collaborative Computational Project, 1994). Phasing statistics are shown in Table 1.

Model building and refinement

Map interpretation and model building were done with the program Xfit (McRee, 1999). The model was refined against the high-resolution low temperature SSRL data with X-PLOR (Brünger, 1992). To monitor the progress of refinement, 5% of the reflections were set aside for a free *R*-factor calculation. Four rounds of model building and refinement were done with Xfit and X-PLOR. The first round of refinement included data to 2.3 Å. The next two rounds included data to 2.0 Å, and the final round of refinement was carried out against all the data. The SHELX-97 package (Sheldrick & Schneider,

1997) was then used for subsequent refinement. As with the X-PLOR refinement, all rounds of SHELX refinement were followed by manual fitting of $2|F_o| - |F_c|$ and σ_A-weighted $|F_o| - |F_c|$ maps using Xfit. Distance, planarity, chiral volume and anti-bumping restraints were applied from the onset of SHELX refinement. After the first round of SHELX refinement, water molecules were added using Xfit. A total of nine rounds of refinement was performed, with the final round including the working (95%) as well as the R_{free} data (5%, Table 2).

Structural superposition and analysis

The least-squares fit option in SHELXPRO (Sheldrick & Schneider, 1997) was used for structural superpositions and rms deviation calculations. Solvent accessibility was calculated with AREAIMOL and RESAREA (Collaborative Computational Project, 1994), and surface feature analysis and calculations were done with GRASP (Nicholls *et al*., 1991). Contacts were examined with CONTACT (Collaborative Computational Project, 1994), and electrostatic calculations were done with DELPHI (Honig & Nicholls, 1995).

Acknowledgments

We thank Peter Kuhn, Mike Soltis, and Aina Cohen for help at the SSRL beamline 7-1, and Pamela Williams and Sridhar Prasad for discussions, advice, and assistance with data collection. 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. This work was supported by NSF grant MCB-9816426 to C. D. S. and NIH grant HD12986 to V.D.V.

References

Biermann, C. H. (1998). The molecular evolution of sperm binding in six species of sea urchins (echinoida: strongylocentrotidae). *Mol. Biol. Evol.* **15**, 1761-1771.

Brünger, A. T. (1992). *X-PLOR Manual Version 3.1: a System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.

Collaborative Computational Project Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D,* **50**, 760-763.

De La Fortelle, E. & Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multi-wavelength anomalous diffraction methods. *Methods Enzymol.* **276**, 472-493.

Esnouf, R. M. (1997). An extensively modified version of MOLSCRIPT that includes greatly enhanced coloring capabilities. *J. Mol. Graph.* **15**, 132-134.

Hellberg, M. E. & Vacquier, V. D. (1999). Rapid evolution of fertilization selectivity and lysin cDNA sequences in teguline gastropods. *Mol. Biol. Evol.* **16**, 839-848.

Honig, B. & Nicholls, A. (1995). Classic electrostatics in biology and chemistry. *Science,* **268**, 1144-1149.

Hooft, R. W. W., Vriend, G., Sander, C. & Abola, E. E. (1996). Errors in protein structures. *Nature,* **381**, 272.

Howard, A. J., Nielsen, C. & Xuong, N. H. (1985). Software for a diffractometer with multiwire area detector. *Methods Enzymol.* **114A**, 452-472.

Kraulis, P. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.

Kresge, N., Vacquier, V. D. & Stout, C. D. (2000). 1.35 and 2.07 Å resolution structures of the red abalone sperm lysin monomer and dimer reveal features involved in receptor binding. *Acta Crystallog. sect. D,* **56**, 34-41.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **24**, 946-950.

Lee, Y.-H. & Vacquier, V. D. (1992). The divergence of species-specific abalone sperm lysins is promoted by positive Darwinian selection. *Biol. Bull,* **182**, 97-104.

Lee, Y.-H., Ota, T. & Vacquier, Y. D. (1995). Positive selection is a general phenomenon in the evolution of abalone sperm lysin. *Mol. Biol. Evol.* **12**, 231-238.

Leighton, D. L. & Lewis, C. A. (1982). Experimental hybridization in abalones. *Int. J. Invert. Reprod.* **5**, 273-282.

Lewis, C. A., Talbot, C. F. & Vacquier, V. D. (1982). A protein from abalone sperm dissolves the egg vitelline layer by a nonenzymatic mechanism. *Dev. Biol.* **92**, 227-239.

Lindberg, D. R. (1992). Evolution, distribution and systematics of Haliotidae. In *Abalone of the World: Biology, Fisheries and Culture* (Shepherd, S. A., Tegner, M. J. & Guzman del Proo, S. A., eds), pp. 3-18, Blackwell Scientific, Oxford.

Lyon, J. D. & Vacquier, V. D. (1999). Interspecies chimeric sperm lysins identify regions mediating species-specific recognition of the abalone egg vitelline envelope. *Dev. Biol.* **214**, 151-159.

McRee, D. E. (1999). XtalView/Xfit - a versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156-165.

Merrit, E. A. & Murphy, M. E. P. (1994). Raster3D version 2.0 - a program for photorealistic molecular graphics. *Acta Crystallog. sect. D,* **50**, 869-873.

Metz, E. C., Robles-Sikisaki, R. & Vacquier, V. D. (1998). Nonsynonymous substitution in abalone sperm fertilization genes exceeds substitution in introns and mitochondrial DNA. *Proc. Natl Acad. Sci. USA,* **95**, 10676-10681.

Nicholls, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281-296.

O'Rand, M. G. (1988). Sperm-egg recognition and barriers to interspecies fertilization. *Gamete Res.* **19**, 315-328.

Owen, B., McLean, J. H. & Meyer, R. J. (1971). Hybridization in the eastern Pacific abalone (*Haliotis*). *Bull. Los Angeles Mus. Nat. Hist. Sci.,* 1-37.

Palumbi, S. R. & Metz, E. C. (1991). Strong reproductive isolation between closely related tropical sea urchins (genus *Echinometra*). *Mol. Biol. Evol.* **8**, 227-239.

Roldan, E. R. S. & Yanagimachi, R. (1989). Cross-fertilization between Syrian and Chinese hamsters. *J. Exp. Zool.* **250**, 321-328.

Shaw, A., McRee, D. E., Vacquier, V. D. & Stout, C. D. (1993). The crystal structure of lysin, a fertilization protein. *Science,* **262**, 1864-1867.

Shaw, A., Lee, Y.-H., Stout, C. D. & Vacquier, V. D. (1994). The species-specificity and structure of abalone sperm lysin. *Dev. Biol.* **5**, 209-215.

Shaw, A., Fortes, P. A. G., Stout, C. D. & Vacquier, V. D. (1995). Crystal structure and subunit dynamics of the abalone sperm lysin dimer: egg envelopes dissociate dimers, the monomer is the active species. *J. Cell Biol.* **130**, 1117-1125.

Sheldrick, G. M. & Schneider, T. R. (1997). SHELXL: high-resolution refinement. *Methods Enzymol.* **276**, 319-343.

Summers, R. G. & Hylander, B. L. (1975). Species-specificity of acrosome reaction and primary gamete binding in echinoids. *Exp. Cell. Res.* **96**, 63-68.

Summers, R. G. & Hylander, B. L. (1976). Primary gamete binding: the species-exclusive event of echinoid fertilization. *Exp. Cell. Res.* **100**, 190-194.

Swanson, W. J. & Vacquier, V. D. (1997). The abalone egg vitelline envelope receptor for sperm lysin is a giant multivalent molecule. *Proc. Natl Acad. Sci. USA,* **94**, 6724-6729.

Swanson, W. J. & Vacquier, V. D. (1998). Concerted evolution in an egg receptor for a rapidly evolving abalone sperm protein. *Science,* **281**, 710-712.

Terwilliger, T. C. & Berendzen, J. (1996). Correlated phasing of multiple isomorphous replacement data. *Acta Crystallog. sect. D,* **52**, 749-757.

Terwilliger, T. C. & Eisenberg, D. (1983). Unbiased three-dimensional refinement of heavy-atom parameters by correlation of origin-removed Patterson functions. *Acta Crystallog. sect. A,* **39**, 813-817.

Terwilliger, T. C. & Eisenberg, D. (1987). Isomorphous replacement: effects of errors on the phase probability distribution. *Acta Crystallog. sect. A,* **43**, 6-13.

Terwilliger, T. C., Kim, S.-H. & Eisenberg, D. (1987). Generalized method of determining heavy-atom positions using the difference Patterson function. *Acta Crystallog. sect. A,* **43**, 1-5.

Vacquier, V. D. (1998). Evolution of gamete recognition proteins. *Science,* **281**, 1995-1998.

Vacquier, V. D. & Lee, Y.-H. (1993). Abalone sperm lysin: unusual mode of evolution of a gamete recognition protein. *Zygote,* **1**, 1-16.

Vacquier, V. D., Swanson, W. J. & Hellberg, M. E. (1995). What have we learned about sea urchin sperm bindin? *Dev. Growth Differ.* **37**, 1-10.

Yanagimachi, R. (1994). Mammalian fertilization. In *The Physiology of Reproduction* (Knobil, B. & Neul, J. D., eds), pp. 189-317, Raven Press, New York.

Edited by R. Huber

(Received 10 November 1999; received in revised form 13 January 2000; accepted 13 January 2000)