Cell, Volume *169*

Supplemental Information

Structural Basis of Egg Coat-Sperm Recognition at Fertilization

Isha Raj, Hamed Sadat Al Hosseini, Elisa Dioguardi, Kaoru Nishimura, Ling Han, Alessandra Villa, Daniele de Sanctis, and Luca Jovine

Table S1. Data Collection and Refinement Statistics for VERL Repeats 1-3, Related to Figures 1, 5 and 7

	mMBP–VR1+ [PDB: 5II4]	mMBP–VR1 [PDB: 5II5]	mMBP–VR3 [PDB: 5IIC]	VR2+ [PDB: 5MR2]
Data Collection				
Wavelength (Å)	0.9762	0.9762	0.9762	0.8729
Resolution range (Å)	46.4 - 2.00 (2.07 - 2.00)	37.8 - 1.80 (1.86 - 1.80)	47.0 - 2.90 (3.00 - 2.90)	47.0 - 2.50 (2.59 - 2.50)
Space group	*P* 65 2 2 (179)	*P* 65 2 2 (179)	*P* 1 21 1 (4)	*C* 1 2 1 (5)
Cell dimensions				
a, *b*, *c* (Å)	107.12, 107.12, 195.85	107.09, 107.09, 196.35	58.33, 81.64, 111.63	69.92, 79.19, 81.24
α, *β*, *γ* (°)	90, 90, 120	90, 90, 120	90, 99.60, 90	90, 102.62, 90
Total reflections	217941 (20613)	366414 (35608)	91100 (8909)	67133 (6520)
Unique reflections	44601 (4369)	61782 (6065)	22738 (2230)	15039 (1495)
Multiplicity	4.9 (4.7)	5.9 (5.9)	4.0 (4.0)	4.5 (4.4
Completeness (%)	97.7 (97.8)	99.1 (99.1)	98.2 (97.1)	99.6 (97.1)
Mean I/sigma(I)	8.8 (1.0)	17.5 (1.1)	7.3 (1.2)	7.8 (0.6)
Wilson B-factor (Å^2)	36.16	33.37	68.84	61.40
R_{meas} (%)	11.5 (162.8)	5.4 (168.3)	20.0 (149.8)	15.7 (271.4)
R_{pim} (%)	5.1 (73.4)	2.1 (66.8)	9.9 (73.4)	7.2 (127.1)
$CC_{1/2}$	1.00 (0.57)	1.00 (0.62)	0.99 (0.50)	1.00 (0.22)
CC*	1.00 (0.85)	1.00 (0.88)	1.00 (0.82)	1.00 (0.60)
Refinement				
Reflections	44558 (4369)	61681 (6056)	22701 (2223)	14993 (1455)
Free reflections	2287 (224)	3155 (308)	1189 (117)	1503 (148)
R_{work} (%)	22.13 (39.50)	20.13 (37.21)	25.36 (41.19)	23.42 (39.52)
R_{free} (%)	24.76 (40.39)	23.45 (40.68)	30.98 (42.06)	28.04 (41.51)
CC_{work}	0.95 (0.66)	0.96 (0.73)	0.93 (0.61)	0.94 (0.44)
CC_{free}	0.89 (0.73)	0.93 (0.64)	0.85 (0.44)	0.91 (0.34)
Number of non-hydrogen atoms	3880	3985	7360	2572
macromolecules	3686	3704	7258	2508
ligands	73	63	102	42
solvent	121	218	0	22
Protein residues	477	476	939	318
RMS deviations				
bonds (Å)	0.006	0.012	0.004	0.009
angles (°)	0.92	1.34	0.96	1.38
Ramachandran plot				
favored (%)	99.16	99.16	98.05	98.33
allowed (%)	0.84	0.84	1.95	1.67
outliers (%)	0	0	0	0
Rotamer outliers (%)	0.78	0.52	0.39	1.04
Clashscore	1.07	1.74	2.4	7.55
Average B-factor (Å^2)	51.99	57.85	77.75	76.14
macromolecules	52.36	58.33	77.82	75.51
ligands	52.39	72.01	72.74	124.72
solvent	40.59	45.67	-	55.96

Parameters for the outermost shell are shown in parentheses. I/sigma(I): signal-to-noise ratio; $R_{meas} = \Sigma_{hkl} \sqrt{(n/n-1)}\ \Sigma_i |I_i(hkl) - I(hkl)| / \Sigma_{hkl} \Sigma_i\ I_i(hkl)$; $R_{pim} = \Sigma_{hkl} \sqrt{(1/n-1)}\ \Sigma_i |I_i(hkl) - I(hkl)| / \Sigma_{hkl} \Sigma_i\ I_i(hkl)$, where $I_i(hkl)$ is the intensity for an observation of a reflection and I(hkl) is the average intensity of all symmetry-related observations of a reflection; $CC_{1/2}$: percentage of correlation between intensities from random half-datasets; $CC^* = \sqrt{(2CC_{1/2}/1+CC_{1/2})}$; $R_{work} = \Sigma_{hkl} ||F_{obs}| - k|F_{calc}|| / \Sigma_{hkl} |F_{obs}|$; R_{free}: same as R_{work} calculated from free reflections excluded from refinement; CC_{work}: correlation of the experimental intensities with the intensities calculated from the refined molecular model; CC_{free}: correlation of the experimental intensities of free reflections excluded from the refinement with the intensities calculated from the refined molecular model.

Table S2. Data Collection and Refinement Statistics for ZP2 ZP-N1, Related to Figure 1

	ZP2 ZP-N1 [PDB: 5II6]
Data Collection	
Wavelength (Å)	0.8856
Resolution range (Å)	31.4 - 0.95 (0.98 - 0.95)
Space group	*P* 21 21 21 (19)
Cell dimensions	
a, *b*, *c* (Å)	39.54, 50.01, 51.75
α, *β*, *γ* (°)	90, 90, 90
Total reflections	128064 (11086)
Unique reflections	64506 (5816)
Multiplicity	2.0 (1.9)
Completeness (%)	98.7 (90.2)
Mean I/sigma(I)	13.0 (2.7)
Wilson B-factor (Å^2)	11.02
R_{meas} (%)	3.2 (34.2)
R_{pim} (%)	2.2 (24.2)
$CC_{1/2}$	1.00 (0.77)
CC*	1.00 (0.93)
Refinement	
Reflections	64499 (5816)
Free reflections	3230 (304)
R_{work} (%)	12.68 (21.49)
R_{free} (%)	14.44 (22.69)
CC_{work}	0.97 (0.89)
CC_{free}	0.96 (0.85)
Number of non-hydrogen atoms	1048
macromolecules	901
ligands	0
solvent	147
Protein residues	100
RMS deviations	
bonds (Å)	0.016
angles (°)	1.67
Ramachandran plot	
favored (%)	96.88
allowed (%)	3.12
outliers (%)	0
Rotamer outliers (%)	0.99
Clashscore	4.53
Average B-factor (Å^2)	16.24
macromolecules	14.2
ligands	-
solvent	28.73

Parameters for the outermost shell are shown in parentheses. I/sigma(I): signal-to-noise ratio; $R_{meas} = \Sigma_{hkl} \sqrt{(n/n-1)}\ \Sigma_i|Ii(hkl) - I(hkl)|/\Sigma_{hkl}\Sigma_i\ I_i(hkl)$; $R_{pim} = \Sigma_{hkl} \sqrt{(1/n-1)}\ \Sigma_i|Ii(hkl) - I(hkl)|/\Sigma_{hkl}\Sigma_i\ I_i(hkl)$, where $I_i(hkl)$ is the intensity for an observation of a reflection and I(hkl) is the average intensity of all symmetry-related observations of a reflection; $CC_{1/2}$: percentage of correlation between intensities from random half-datasets; $CC^* = \sqrt{(2CC_{1/2}/1+CC_{1/2})}$; $R_{work} = \Sigma_{hkl}||F_{obs}| - k|F_{calc}||/\Sigma_{hkl}|F_{obs}|$; R_{free}: same as R_{work} calculated from free reflections excluded from refinement; CC_{work}: correlation of the experimental intensities with the intensities calculated from the refined molecular model; CC_{free}: correlation of the experimental intensities of free reflections excluded from the refinement with the intensities calculated from the refined molecular model.

Table S3. Data Collection and Refinement Statistics for Lysin^R, Related to Figure 7

	Orthorhombic lysin^R (S–SAD) [PDB: 5II7]	Orthorhombic lysin^R (high resolution) [PDB: 5II8]	Monoclinic lysin^R [PDB: 5II9]
Data Collection			
Wavelength (Å)	1.542	0.886	1.029
Resolution range (Å)	43.2 - 1.66 (1.72 - 1.66)	31.6 - 0.99 (1.02 - 0.99)	27.3 - 2.11 (2.18 - 2.11)
Space group	*P* 21 21 21 (19)	*P* 21 21 21 (19)	*P* 1 21 1 (4)
Cell dimensions			
a, *b*, *c* (Å)	45.20, 51.19, 80.42	45.21, 51.05, 80.53	39.19, 69.34, 182.75
α, *β*, *γ* (°)	90, 90, 90	90, 90, 90	90, 90.01, 90
Total reflections	135781 (432)	480552 (43518)	125117 (10244)
Unique reflections	19219 (359)	98373 (9051)	52122 (5036)
Multiplicity	7.1 (1.2)	4.9 (4.8)	2.4 (2.0)
Completeness (%)	84.8 (16.2)	94.3 (87.7)	92.0 (88.7)
Mean I/sigma(I)	49.5 (6.5)	11.9 (1.2)	10.4 (2.3)
Wilson B-factor (Å^2)	13.2	11.79	27.47
R_{meas} (%)	3.9 (19.3)	5.9 (199.0)	7.7 (48.6)
R_{pim} (%)	1.3 (13.0)	2.6 (89.3)	4.6 (30.0)
$CC_{1/2}$	1.00 (0.95)	1.00 (0.38)	1.00 (0.76)
CC*	1.00 (0.99)	1.00 (0.74)	1.00 (0.93)
Refinement			
Reflections	19195 (359)	98352 (9051)	52119 (5036)
Free reflections	1929 (43)	9849 (898)	2660 (266)
R_{work} (%)	15.62 (24.67)	16.42 (30.65)	19.79 (25.63)
R_{free} (%)	18.80 (37.33)	19.09 (33.25)	22.97 (29.38)
CC_{work}	0.95 (0.86)	0.96 (0.68)	0.95 (0.80)
CC_{free}	0.94 (0.88)	0.93 (0.62)	0.92 (0.75)
Number of non-hydrogen atoms	1507	1448	7181
macromolecules	1168	1146	6666
ligands	57	77	0
solvent	282	225	515
Protein residues	132	131	784
RMS deviations			
bonds (Å)	0.006	0.015	0.005
angles (°)	0.96	1.59	0.99
Ramachandran plot			
favored (%)	100	100	100
allowed (%)	0	0	0
outliers (%)	0	0	0
Rotamer outliers (%)	0.82	0	0.29
Clashscore	2.46	4.13	1.41
Average B-factor (Å^2)	18.93	18.85	34.58
macromolecules	16.26	14.39	34.55
ligands	37.06	46.01	-
solvent	26.33	32.27	34.99

Parameters for the outermost shell are shown in parentheses. I/sigma(I): signal-to-noise ratio; $R_{meas} = \Sigma_{hkl} \sqrt{(n/n-1)}\ \Sigma_i|I_i(hkl) - I(hkl)|/\Sigma_{hkl}\Sigma_i\ I_i(hkl)$; $R_{pim} = \Sigma_{hkl} \sqrt{(1/n-1)}\ \Sigma_i|I_i(hkl) - I(hkl)|/\Sigma_{hkl}\Sigma_i\ I_i(hkl)$, where $I_i(hkl)$ is the intensity for an observation of a reflection and I(hkl) is the average intensity of all symmetry-related observations of a reflection; $CC_{1/2}$: percentage of correlation between intensities from random half-datasets; $CC^* = \sqrt{(2CC_{1/2}/1+CC_{1/2})}$; $R_{work} = \Sigma_{hkl}||F_{obs}| - k|F_{calc}||/\Sigma_{hkl}|F_{obs}|$; R_{free}: same as R_{work} calculated from free reflections excluded from refinement; CC_{work}: correlation of the experimental intensities with the intensities calculated from the refined molecular model; CC_{free}: correlation of the experimental intensities of free reflections excluded from the refinement with the intensities calculated from the refined molecular model.

Table S4. Data Collection and Refinement Statistics for VERL/Lysin Complexes, Related to Figures 3, 4, 5, 6 and 7

	Triclinic VR3/lysin complex [PDB: 5IIA]	Trigonal VR3/lysin complex [PDB: 5IIB]	VR2+/lysin complex [PDB: 5MR3]
Data Collection			
Wavelength (Å)	0.972	0.976	1.072
Resolution range (Å)	48.7 - 1.70 (1.77 - 1.70)	32.2 - 1.64 (1.70 - 1.64)	43.9 - 1.80 (1.86 - 1.80)
Space group	*P* 1 (1)	*P* 31 2 1 (152)	*P* 1 21 1 (4)
Cell dimensions			
a, *b*, *c* (Å)	60.49, 60.33, 89.17	64.50, 64.50, 129.93	75.87, 87.86, 92.55
α, *β*, *γ* (°)	105.04, 89.04, 113.28	90, 90, 120	90, 100.92, 90
Total reflections	412456 (52385)	210912 (21243)	372220 (37258)
Unique reflections	119274 (14764)	39053 (3892)	109992 (10977)
Multiplicity	3.5 (3.5)	5.4 (5.5)	3.4 (3.4)
Completeness (%)	97.3 (96.3)	99.5 (98.8)	99.5 (99.6)
Mean I/sigma(I)	9.6 (0.8)	12.4 (0.8)	11.5 (1.0)
Wilson B-factor (Å^2)	35.28	34.67	34.22
R_{meas} (%)	5.8 (199.3)	5.4 (231.0)	6.4 (177.3)
R_{pim} (%)	3.1 (104.1)	2.3 (97.6)	3.4 (94.5)
$CC_{1/2}$	1.00 (0.27)	1.00 (0.51)	1.00 (0.31)
CC*	1.00 (0.65)	1.00 (0.82)	1.00 (0.68)
Refinement			
Reflections	119231 (14753)	38972 (3847)	109971 (10966)
Free reflections	3799 (470)	1850 (180)	5503 (548)
R_{work} (%)	17.79 (39.01)	19.38 (40.60)	20.26 (39.31)
R_{free} (%)	20.24 (39.48)	20.82 (44.81)	22.86 (42.76)
CC_{work}	0.97 (0.58)	0.95 (0.64)	0.96 (0.52)
CC_{free}	0.96 (0.59)	0.95 (0.53)	0.95 (0.40)
Number of non-hydrogen atoms	8257	2093	8543
macromolecules	7568	1901	7833
ligands	112	66	150
solvent	577	126	560
Protein residues	922	232	959
RMS deviations			
bonds (Å)	0.008	0.007	0.005
angles (°)	1.19	1.21	1.12
Ramachandran plot			
favored (%)	99.78	99.56	99.36
allowed (%)	0.22	0.44	0.64
outliers (%)	0	0	0
Rotamer outliers (%)	0.85	0	0.12
Clashscore	1.82	0.51	3.37
Average B-factor (Å^2)	47.69	48.17	51.13
macromolecules	47.37	47.23	50.9
ligands	71.69	78.47	95.78
solvent	47.17	46.44	42.32

Parameters for the outermost shell are shown in parentheses. I/sigma(I): signal-to-noise ratio; $R_{meas} = \Sigma_{hkl} \sqrt{(n/n-1)} \Sigma_i |I_i(hkl) - I(hkl)| / \Sigma_{hkl} \Sigma_i I_i(hkl)$; $R_{pim} = \Sigma_{hkl} \sqrt{(1/n-1)} \Sigma_i |I_i(hkl) - I(hkl)| / \Sigma_{hkl} \Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the intensity for an observation of a reflection and I(hkl) is the average intensity of all symmetry-related observations of a reflection; $CC_{1/2}$: percentage of correlation between intensities from random half-datasets; $CC^* = \sqrt{(2CC_{1/2}/1+CC_{1/2})}$; $R_{work} = \Sigma_{hkl} ||F_{obs}| - k|F_{calc}|| / \Sigma_{hkl} |F_{obs}|$; R_{free}: same as R_{work} calculated from free reflections excluded from refinement; CC_{work}: correlation of the experimental intensities with the intensities calculated from the refined molecular model; CC_{free}: correlation of the experimental intensities of free reflections excluded from the refinement with the intensities calculated from the refined molecular model.

Table S5. Protein-Protein Interactions in the Trigonal Structure of the VR3/Lysin Complex [PDB: 5IIB], Related to Figures 3 and 4

Lysin		VR3		
Hydrogen bonds				
R 74	(α2)	F 385		SM (2 H-bonds)
R 74	(α2)	I 387	(βE)	SM
T 78	(α2)	T 382		MM
T 78	(α2)	N 388	(βE)	SS
H 79	(α2)	M 389	(βE)	SM
N 82	(α2)	P 380	(βD)	SM
W 86	(α2)	D 401		SS
K 90	(α2)	D 401		SS (salt bridge)
R 113	(α3)	D 401		SM
Y 135	(α5)	I 387	(βE)	SM
Ion pair				
K 150		E 384		
Hydrophobic interactions				
F 70	(α2)	I 387	(βE)	
V 71	(α2)	M 389	(βE)	
A 81	(α2)	P 380	(βD)	
Y 83	(α2)	P 403		
L 85	(α2)	P 380	(βD)	
W 86	(α2)	P 403		
I 114		L 391	(βE)	
Y 118	(α4)	P 352		
F 119	(α4)	M 389	(βE)	aromatic-sulfur
F 119	(α4)	L 391	(βE)	
F 122	(α4)	Y 344	(βA)	aromatic-aromatic
F 122	(α4)	I 356	(βB)	
M 128		Y 344	(βA)	aromatic-sulfur
P 130		I 387	(βE)	
Y 135	(α5)	F 385		aromatic-aromatic
Cation-π interaction				
F 122	(α4)	K 354	(βB)	
Non-bonded contacts				
Y 75	(α2)	M 389	(βE)	
Q 77	(α2)	T 382		
Q 77	(α2)	E 384		
T 78	(α2)	I 381	(βD)	
H 79	(α2)	M 390	(βE)	
N 82	(α2)	V 379	(βD)	
N 82	(α2)	I 381	(βD)	
N 82	(α2)	G 404		
W 86	(α2)	S 402		
Y 135	(α5)	G 386	(βE)	

SM, side chain-main chain; MM, main chain-main-chain; SS, side chain-side chain. All interactions are within 4 Å distance.

Table S6. Protein-Protein Interactions in the Structure of the VR2+/Lysin Complex [PDB: 5MR3], Related to Figure 6

Lysin		VR2+		
Hydrogen bonds				
R 74	(α2)	H 222		SM (2 H-bonds)
R 74	(α2)	F 224	(βE)	SM (2 H-bonds)
T 78	(α2)	T 219		MM (2 H-bonds)
T 78	(α2)	N 225	(βE)	SS (3 H-bonds)
H 79	(α2)	M 226	(βE)	SM
R 113	(α3)	D 238		SM (2 H-bonds)
Y 135	(α5)	F 224	(βE)	SM
Ion pair				
K 150		E 221		
Hydrophobic interactions				
V 71	(α2)	M 226	(βE)	
Y 75	(α2)	M 226	(βE)	
L 85	(α2)	P 217	(βD)	
W 86	(α2)	P 240		
Y 118	(α4)	P 189		
F 119	(α4)	M 226	(βE)	aromatic-sulfur
F 119	(α4)	L 228	(βE)	
F 122	(α4)	F 181	(βA)	aromatic-aromatic
F 122	(α4)	I 193	(βB)	
M 128		F 181	(βA)	aromatic-sulfur
P 130		F 224	(βE)	
Cation-π interaction				
F 122	(α4)	K 191	(βB)	
Non-bonded contacts				
R 74	(α2)	N 225	(βE)	
R 74	(α2)	E 221		
H 79	(α2)	I 218	(βD)	
N 82	(α2)	P 217	(βD)	
N 82	(α2)	G 241		
N 82	(α2)	V 216	(βD)	
Y 83	(α2)	P 240		
Y 135	(α5)	G 223	(βE)	

SM, side chain-main chain; MM, main chain-main-chain; SS, side chain-side chain. All interactions are within 4 Å distance.