WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Oct 11, 2021 – 01:33 AM EDT

PDB ID	:	2RAL
Title	:	Crystal Structure Analysis of double cysteine mutant of S.epidermidis adhesin SdrG: Evidence for the Dock,Lock and Latch ligand binding mechanism
Authors	:	Ponnuraj, K.; Sthanam, N.; Bowden, M.G.; Hook, M.
Deposited on	:	2007-09-17
Resolution	:	2.80 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

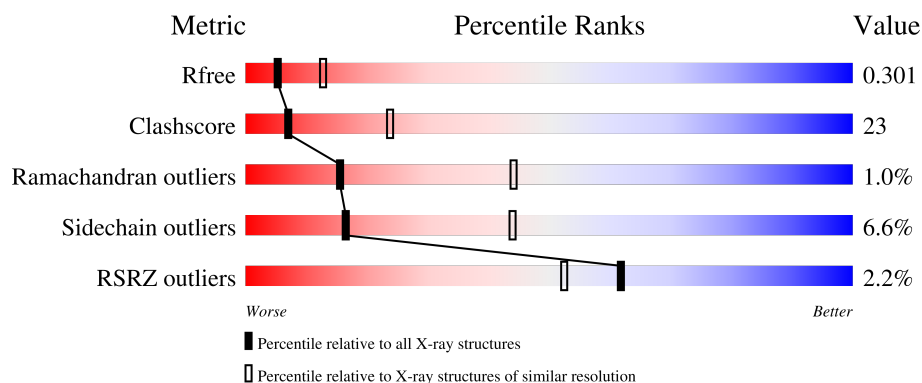|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Rfree		0.301
Clashscore		23
Ramachandran outliers		1.0%
Sidechain outliers		6.6%
RSRZ outliers		2.2%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	340	% 60% 32% • 6%
1	B	340	3% 54% 34% 5% 7%

WORLDWIDE
wwPDB
PROTEIN DATA BANK

2 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 4966 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Serine-aspartate repeat-containing protein G.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	320	Total 2491	C 1537	N 405	O 544	S 5	0	0	0
1	B	315	Total 2475	C 1530	N 404	O 536	S 5	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	256	MET	-	expression tag	UNP Q9KI13
A	257	GLY	-	expression tag	UNP Q9KI13
A	258	ARG	-	expression tag	UNP Q9KI13
A	259	SER	-	expression tag	UNP Q9KI13
A	260	HIS	-	expression tag	UNP Q9KI13
A	261	HIS	-	expression tag	UNP Q9KI13
A	262	HIS	-	expression tag	UNP Q9KI13
A	263	HIS	-	expression tag	UNP Q9KI13
A	264	HIS	-	expression tag	UNP Q9KI13
A	265	HIS	-	expression tag	UNP Q9KI13
A	266	GLY	-	expression tag	UNP Q9KI13
A	267	SER	-	expression tag	UNP Q9KI13
A	268	LEU	-	expression tag	UNP Q9KI13
A	269	VAL	-	expression tag	UNP Q9KI13
A	270	PRO	-	expression tag	UNP Q9KI13
A	271	ARG	-	expression tag	UNP Q9KI13
A	272	GLY	-	expression tag	UNP Q9KI13
A	273	SER	-	expression tag	UNP Q9KI13
A	381	CYS	GLU	engineered mutation	UNP Q9KI13
A	595	CYS	PRO	engineered mutation	UNP Q9KI13
B	256	MET	-	expression tag	UNP Q9KI13
B	257	GLY	-	expression tag	UNP Q9KI13
B	258	ARG	-	expression tag	UNP Q9KI13
B	259	SER	-	expression tag	UNP Q9KI13
B	260	HIS	-	expression tag	UNP Q9KI13

Continued on next page...

Continued from previous page...

| Chain	Residue	Modelled	Actual	Comment	Reference
B	261	HIS	-	expression tag	UNP Q9KI13
B	262	HIS	-	expression tag	UNP Q9KI13
B	263	HIS	-	expression tag	UNP Q9KI13
B	264	HIS	-	expression tag	UNP Q9KI13
B	265	HIS	-	expression tag	UNP Q9KI13
B	266	GLY	-	expression tag	UNP Q9KI13
B	267	SER	-	expression tag	UNP Q9KI13
B	268	LEU	-	expression tag	UNP Q9KI13
B	269	VAL	-	expression tag	UNP Q9KI13
B	270	PRO	-	expression tag	UNP Q9KI13
B	271	ARG	-	expression tag	UNP Q9KI13
B	272	GLY	-	expression tag	UNP Q9KI13
B	273	SER	-	expression tag	UNP Q9KI13
B	381	CYS	GLU	engineered mutation	UNP Q9KI13
B	595	CYS	PRO	engineered mutation	UNP Q9KI13

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1: Serine-aspartate repeat-containing protein G
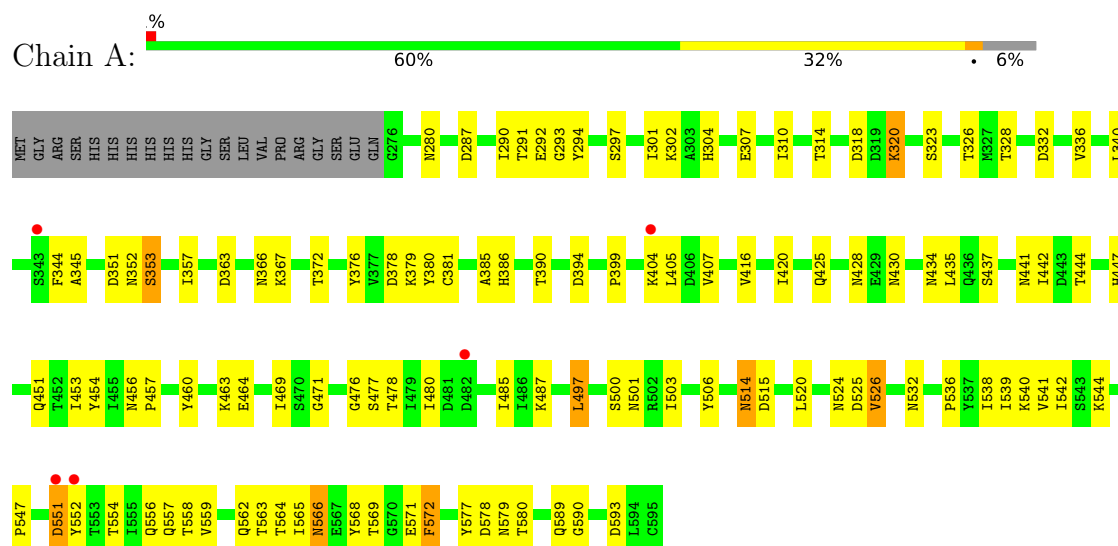
• Molecule 1: Serine-aspartate repeat-containing protein G
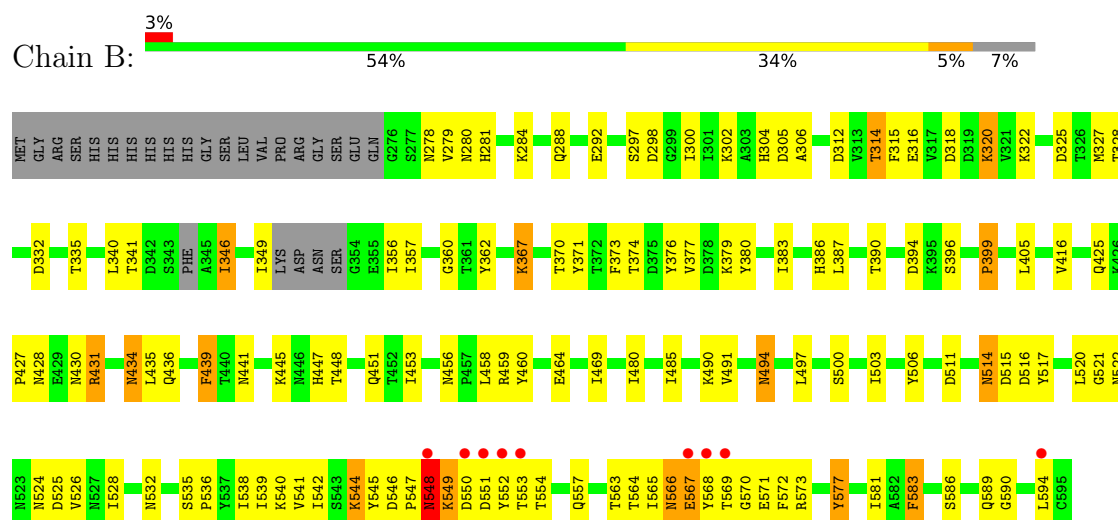
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	61.51Å 94.17Å 129.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.99 – 2.80 32.38 – 2.79	Depositor EDS
% Data completeness (in resolution range)	97.3 (19.99-2.80) 96.7 (32.38-2.79)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.57 (at 2.81Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.242 , 0.303 0.241 , 0.301	Depositor DCC
R_{free} test set	761 reflections (3.97%)	wwPDB-VP
Wilson B-factor (Å^2)	32.7	Xtriage
Anisotropy	0.220	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 29.5	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	4966	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.17% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5   Model quality ⓘ

5.1   Standard geometry ⓘ

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$ >5	RMSZ	#$|Z|$ >5
1	A	0.44	0/2529	0.70	0/3444
1	B	0.44	0/2511	0.68	1/3414 (0.0%)
All	All	0.44	0/5040	0.69	1/6858 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	405	LEU	CA-CB-CG	6.03	129.16	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2   Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2491	0	2342	101	0
1	B	2475	0	2353	122	0
All	All	4966	0	4695	221	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 23.

The worst 5 of 221 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:304:HIS:HD2	1:B:425:GLN:H	1.11	0.94
1:A:304:HIS:HD2	1:A:425:GLN:H	0.94	0.94
1:A:304:HIS:CD2	1:A:425:GLN:H	1.85	0.92
1:A:480:ILE:H	1:A:524:ASN:HD21	1.15	0.91
1:B:456:ASN:HD21	1:B:460:TYR:H	1.20	0.89

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	318/340 (94%)	290 (91%)	26 (8%)	2 (1%)	25	56
1	B	309/340 (91%)	276 (89%)	29 (9%)	4 (1%)	12	36
All	All	627/680 (92%)	566 (90%)	55 (9%)	6 (1%)	15	44

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	352	ASN
1	B	549	LYS
1	B	548	ASN
1	B	399	PRO
1	A	515	ASP

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	286/310 (92%)	270 (94%)	16 (6%)	21	51
1	B	286/310 (92%)	264 (92%)	22 (8%)	13	35
All	All	572/620 (92%)	534 (93%)	38 (7%)	16	44

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	494	ASN
1	B	577	TYR
1	B	514	ASN
1	B	553	THR
1	B	589	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 43 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	402	ASN
1	B	494	ASN
1	B	430	ASN
1	B	456	ASN
1	B	514	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	320/340 (94%)	-0.34	5 (1%)	72	66	14, 23, 33, 43	0
1	B	315/340 (92%)	-0.12	9 (2%)	51	41	18, 27, 37, 43	0
All	All	635/680 (93%)	-0.23	14 (2%)	62	52	14, 25, 36, 43	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	553	THR	5.7
1	B	552	TYR	5.4
1	B	550	ASP	3.5
1	B	551	ASP	3.4
1	B	568	TYR	3.2

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK