wwPDB X-ray Structure Validation Summary Report (i)

Jun 18, 2024 – 09:12 AM EDT

PDB ID	:	4HBL
Title	:	Crystal structure of AbfR of Staphylococcus epidermidis
Authors	:	Liu, X.; Sun, X.; Gan, J.; Lan, L.; Yang, C.-G.
Deposited on	:	2012-09-28
Resolution	:	2.50 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

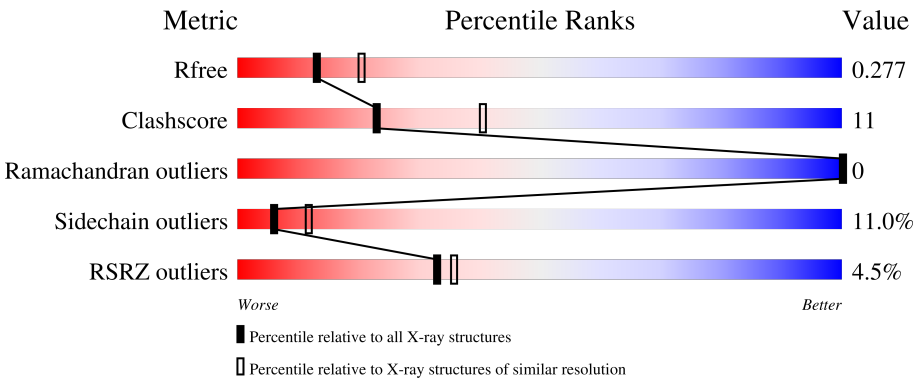
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.50 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	149	3% 70% 21% • 7%
1	B	149	3% 66% 27% • 5%
1	C	149	5% 66% 25% • • 5%
1	D	149	5% 64% 30% 5% •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4732 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 Transcriptional regulator, MarR family.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	138	1141	726	188	221	2	4	0	0	0
1	B	141	1170	742	196	226	2	4	0	0	0
1	C	142	1179	747	197	229	2	4	0	0	0
1	D	147	1224	773	207	238	2	4	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	EXPRESSION TAG	UNP Q5HKZ1
A	-1	ASN	-	EXPRESSION TAG	UNP Q5HKZ1
A	0	ALA	-	EXPRESSION TAG	UNP Q5HKZ1
A	44	MSE	LEU	ENGINEERED MUTATION	UNP Q5HKZ1
A	72	MSE	LEU	ENGINEERED MUTATION	UNP Q5HKZ1
B	-2	SER	-	EXPRESSION TAG	UNP Q5HKZ1
B	-1	ASN	-	EXPRESSION TAG	UNP Q5HKZ1
B	0	ALA	-	EXPRESSION TAG	UNP Q5HKZ1
B	44	MSE	LEU	ENGINEERED MUTATION	UNP Q5HKZ1
B	72	MSE	LEU	ENGINEERED MUTATION	UNP Q5HKZ1
C	-2	SER	-	EXPRESSION TAG	UNP Q5HKZ1
C	-1	ASN	-	EXPRESSION TAG	UNP Q5HKZ1
C	0	ALA	-	EXPRESSION TAG	UNP Q5HKZ1
C	44	MSE	LEU	ENGINEERED MUTATION	UNP Q5HKZ1
C	72	MSE	LEU	ENGINEERED MUTATION	UNP Q5HKZ1
D	-2	SER	-	EXPRESSION TAG	UNP Q5HKZ1
D	-1	ASN	-	EXPRESSION TAG	UNP Q5HKZ1
D	0	ALA	-	EXPRESSION TAG	UNP Q5HKZ1
D	44	MSE	LEU	ENGINEERED MUTATION	UNP Q5HKZ1
D	72	MSE	LEU	ENGINEERED MUTATION	UNP Q5HKZ1

- Molecule 2 is water.

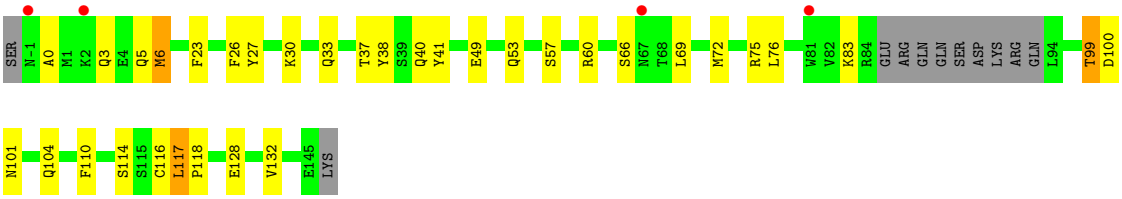
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	3	Total O 3 3	0	0
2	B	7	Total O 7 7	0	0
2	C	5	Total O 5 5	0	0
2	D	3	Total O 3 3	0	0

3 Residue-property plots (i)

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.

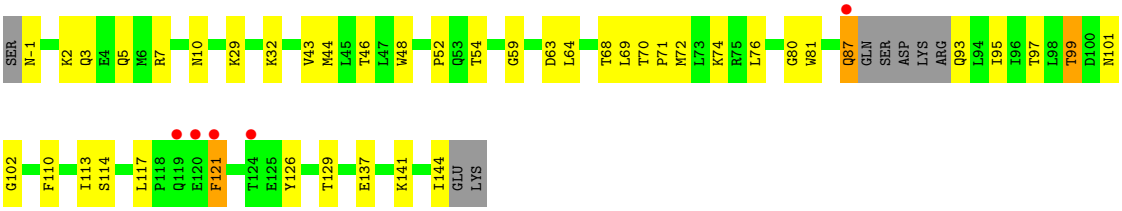
• Molecule 1: Transcriptional regulator, MarR family

Chain A:



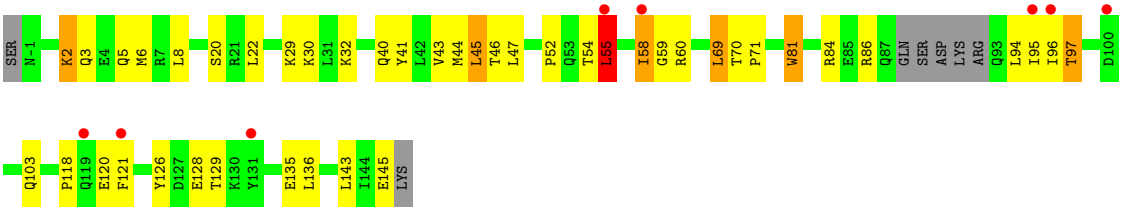
• Molecule 1: Transcriptional regulator, MarR family

Chain B:



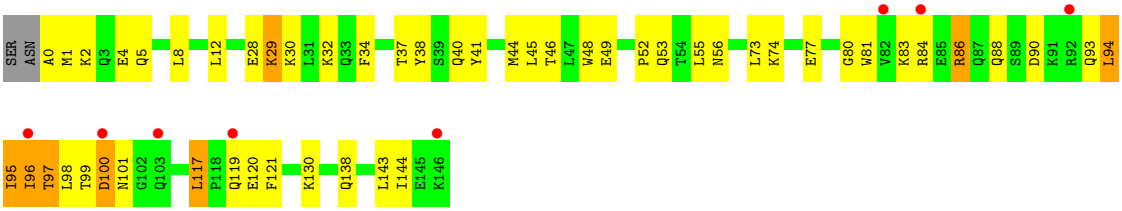
• Molecule 1: Transcriptional regulator, MarR family

Chain C:



• Molecule 1: Transcriptional regulator, MarR family

Chain D:


SER
ASN
A0
M1
K2
Q3
E4
Q5
L8
L12
E28
K29
K30
L31
K32
Q33
F34
T37
Y38
S39
Q40
Y41
M44
L45
T46
L47
W48
E49
P52
Q53
T54
L55
N56
L73
K74
E77
G80
W81
V82
K83
R84
E85
R86
Q87
Q88
S89
D90
K91
R92
Q93
L94
I95
I96
T97
L98
T99
D100
N101
G102
Q103
L117
P118
Q119
E120
F121
K130
Q138
L143
I144
E145
K146

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	127.78Å 50.62Å 107.45Å 90.00° 110.39° 90.00°	Depositor
Resolution (Å)	30.00 – 2.50 44.68 – 2.50	Depositor EDS
% Data completeness (in resolution range)	95.0 (30.00-2.50) 95.1 (44.68-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.82 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.210 , 0.259 0.213 , 0.277	Depositor DCC
R_{free} test set	1090 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	55.0	Xtriage
Anisotropy	0.549	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 56.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4732	wwPDB-VP
Average B, all atoms (Å^2)	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.46% 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 (i)

5.1 Standard geometry (i)

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	0.68	0/1157	0.75	0/1554
1	B	0.76	1/1186 (0.1%)	0.81	0/1592
1	C	0.68	1/1195 (0.1%)	0.75	1/1604 (0.1%)
1	D	0.68	1/1241 (0.1%)	0.75	0/1663
All	All	0.70	3/4779 (0.1%)	0.76	1/6413 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	81	TRP	CD2-CE2	5.55	1.48	1.41
1	B	48	TRP	CD2-CE2	5.54	1.48	1.41
1	D	34	PHE	C-N	5.06	1.42	1.33

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	55	LEU	CA-CB-CG	6.80	130.94	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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	1141	0	1126	29	0
1	B	1170	0	1155	37	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1179	0	1161	29	0
1	D	1224	0	1212	36	0
2	A	3	0	0	0	0
2	B	7	0	0	2	0
2	C	5	0	0	0	0
2	D	3	0	0	0	0
All	All	4732	0	4654	107	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 11.

The worst 5 of 107 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:A:6:MSE:CE	1:B:113:ILE:HB	2.01	0.91
1:A:6:MSE:HE1	1:B:113:ILE:HB	1.53	0.88
1:B:99:THR:HG22	1:B:102:GLY:H	1.37	0.87
1:A:110:PHE:O	1:A:114:SER:HB2	1.77	0.84
1:B:80:GLY:O	1:B:99:THR:HG21	1.77	0.84

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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	134/149 (90%)	132 (98%)	2 (2%)	0	100	100
1	B	137/149 (92%)	135 (98%)	2 (2%)	0	100	100
1	C	138/149 (93%)	134 (97%)	4 (3%)	0	100	100
1	D	145/149 (97%)	139 (96%)	6 (4%)	0	100	100
All	All	554/596 (93%)	540 (98%)	14 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

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	128/135 (95%)	118 (92%)	10 (8%)	12 24
1	B	131/135 (97%)	122 (93%)	9 (7%)	15 30
1	C	132/135 (98%)	112 (85%)	20 (15%)	3 5
1	D	137/135 (102%)	118 (86%)	19 (14%)	3 6
All	All	528/540 (98%)	470 (89%)	58 (11%)	6 12

5 of 58 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	69	LEU
1	D	120	GLU
1	C	135	GLU
1	D	119	GLN
1	D	96	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	25	GLN
1	D	87	GLN
1	D	138	GLN
1	D	88	GLN
1	D	53	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

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	134/149 (89%)	0.34	4 (2%)	50	53	44, 74, 107, 126	0
1	B	137/149 (91%)	0.45	5 (3%)	42	46	40, 58, 93, 129	0
1	C	138/149 (92%)	0.35	8 (5%)	23	24	45, 77, 101, 119	0
1	D	143/149 (95%)	0.43	8 (5%)	24	25	43, 83, 122, 136	0
All	All	552/596 (92%)	0.39	25 (4%)	33	36	40, 73, 109, 136	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	119	GLN	5.2
1	A	-1	ASN	3.8
1	B	121	PHE	3.6
1	C	121	PHE	3.4
1	B	120	GLU	3.3

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.