wwPDB X-ray Structure Validation Summary Report (i)

Sep 20, 2023 – 03:51 PM EDT

PDB ID	:	5IDR
Title	:	Crystal structure of Proteus Mirabilis ScsC in a transitional conformation
Authors	:	Furlong, E.J.; Kurth, F.; Choudhury, H.G.; Martin, J.L.
Deposited on	:	2016-02-24
Resolution	:	2.56 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 (1)) were used in the production of this report:

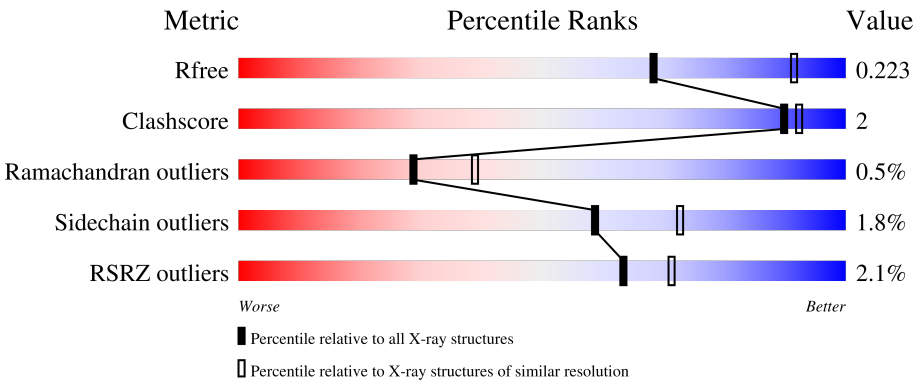
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.56 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	1279 (2.58-2.54)
Clashscore	141614	1327 (2.58-2.54)
Ramachandran outliers	138981	1312 (2.58-2.54)
Sidechain outliers	138945	1312 (2.58-2.54)
RSRZ outliers	127900	1269 (2.58-2.54)

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	224	3% 90% 8% •
1	B	224	% 93% • •
1	C	224	5% 90% 5% • •
1	D	224	2% 93% 6% •
1	E	224	92% 5% ••

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	224	% 91% 7%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 20856 atoms, of which 10512 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 DsbA-like protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	219	3475	1080	1762	290	335	8	0	2	0
1	B	219	3458	1080	1745	290	335	8	0	2	0
1	C	216	3397	1060	1715	283	331	8	0	2	0
1	D	221	3497	1087	1771	293	338	8	0	2	0
1	E	221	3505	1089	1778	293	337	8	0	2	0
1	F	219	3442	1070	1741	287	336	8	0	2	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	SER	-	expression tag	UNP C2LPE2
A	2	ASN	-	expression tag	UNP C2LPE2
B	1	SER	-	expression tag	UNP C2LPE2
B	2	ASN	-	expression tag	UNP C2LPE2
C	1	SER	-	expression tag	UNP C2LPE2
C	2	ASN	-	expression tag	UNP C2LPE2
D	1	SER	-	expression tag	UNP C2LPE2
D	2	ASN	-	expression tag	UNP C2LPE2
E	1	SER	-	expression tag	UNP C2LPE2
E	2	ASN	-	expression tag	UNP C2LPE2
F	1	SER	-	expression tag	UNP C2LPE2
F	2	ASN	-	expression tag	UNP C2LPE2

- Molecule 2 is water.

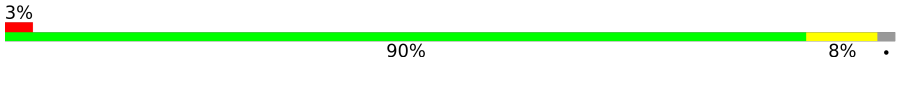
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	17	17	17	0	0
			Total	O		
2	B	7	7	7	0	0
			Total	O		
2	C	19	19	19	0	0
			Total	O		
2	D	9	9	9	0	0
			Total	O		
2	E	17	17	17	0	0
			Total	O		
2	F	13	13	13	0	0

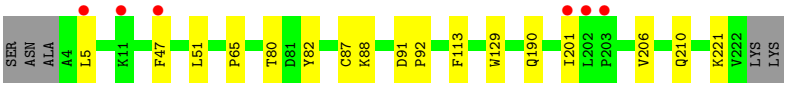
3 Residue-property plots (i)

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: DsbA-like protein

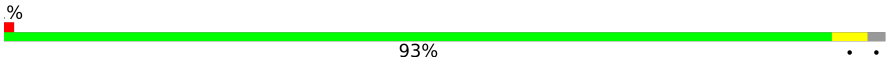
Chain A:

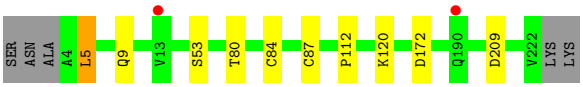




- Molecule 1: DsbA-like protein

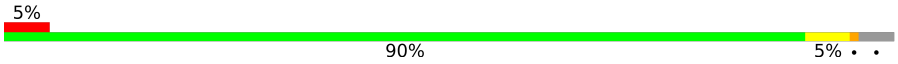
Chain B:

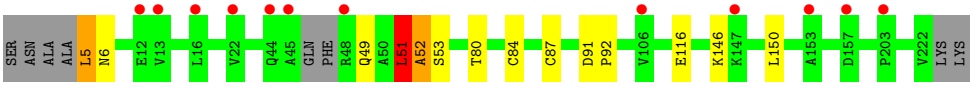




- Molecule 1: DsbA-like protein

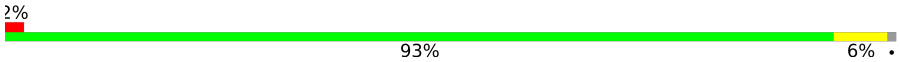
Chain C:





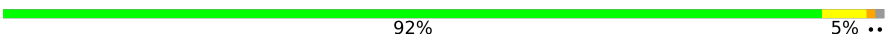
- Molecule 1: DsbA-like protein

Chain D:



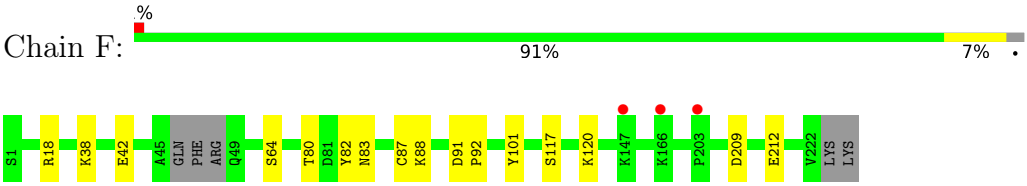
- Molecule 1: DsbA-like protein

Chain E:



- Molecule 1: DsbA-like protein

Chain F:


%
91%
7%
S1
R18
K38
E42
A45
GLN
PHE
ARG
Q49
S64
T80
D81
Y82
N83
C87
K88
D91
P92
Y101
S117
K120
K147
K166
P203
D209
E212
V222
LYS
LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, α, β, γ	193.05Å 193.05Å 105.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.82 – 2.56 136.51 – 2.56	Depositor EDS
% Data completeness (in resolution range)	99.3 (42.82-2.56) 99.3 (136.51-2.56)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.46 (at 2.55Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.171 , 0.222 0.178 , 0.223	Depositor DCC
R_{free} test set	3015 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å^2)	45.0	Xtriage
Anisotropy	0.432	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 47.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.025 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	20856	wwPDB-VP
Average B, all atoms (Å^2)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.91% 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	0.42	0/1736	0.53	0/2349
1	B	0.39	0/1736	0.53	0/2349
1	C	0.36	0/1703	0.54	1/2304 (0.0%)
1	D	0.38	0/1749	0.51	0/2367
1	E	0.39	0/1750	0.55	0/2367
1	F	0.37	0/1722	0.53	0/2330
All	All	0.38	0/10396	0.53	1/14066 (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	C	51	LEU	CA-CB-CG	6.67	130.63	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	1713	1762	1757	12	0
1	B	1713	1745	1757	3	0
1	C	1682	1715	1723	5	0
1	D	1726	1771	1768	7	0
1	E	1727	1778	1775	9	0
1	F	1701	1741	1745	7	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	17	0	0	0	0
2	B	7	0	0	0	0
2	C	19	0	0	0	0
2	D	9	0	0	0	0
2	E	17	0	0	0	0
2	F	13	0	0	0	0
All	All	10344	10512	10525	37	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 2.

The worst 5 of 37 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:190:GLN:O	1:E:185:ARG:NH1	2.13	0.80
1:F:38:LYS:NZ	1:F:42:GLU:OE2	2.24	0.70
1:A:47:PHE:CZ	1:A:201:ILE:HB	2.31	0.65
1:D:47:PHE:CZ	1:D:201:ILE:HB	2.33	0.63
1:D:18:ARG:NH2	1:E:30:GLU:OE2	2.34	0.60

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	219/224 (98%)	215 (98%)	3 (1%)	1 (0%)	29 39
1	B	219/224 (98%)	215 (98%)	3 (1%)	1 (0%)	29 39
1	C	214/224 (96%)	208 (97%)	3 (1%)	3 (1%)	11 14
1	D	221/224 (99%)	219 (99%)	2 (1%)	0	100 100
1	E	221/224 (99%)	215 (97%)	5 (2%)	1 (0%)	29 39

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	217/224 (97%)	211 (97%)	6 (3%)	0	100	100
All	All	1311/1344 (98%)	1283 (98%)	22 (2%)	6 (0%)	29	39

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	52	ALA
1	C	49	GLN
1	A	221	LYS
1	C	6	ASN
1	B	112	PRO

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	189/191 (99%)	189 (100%)	0	100	100
1	B	189/191 (99%)	182 (96%)	7 (4%)	34	45
1	C	186/191 (97%)	180 (97%)	6 (3%)	39	51
1	D	190/191 (100%)	188 (99%)	2 (1%)	73	83
1	E	190/191 (100%)	186 (98%)	4 (2%)	53	67
1	F	188/191 (98%)	185 (98%)	3 (2%)	62	76
All	All	1132/1146 (99%)	1110 (98%)	22 (2%)	59	71

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	209	ASP
1	E	103	ASP
1	E	53	SER
1	E	172	ASP
1	B	209	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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	219/224 (97%)	0.62	6 (2%) 54 63	27, 41, 73, 107	0
1	B	219/224 (97%)	0.62	2 (0%) 84 89	30, 45, 82, 117	0
1	C	216/224 (96%)	0.67	12 (5%) 24 31	32, 50, 82, 106	0
1	D	221/224 (98%)	0.62	5 (2%) 60 68	32, 46, 72, 98	0
1	E	221/224 (98%)	0.57	0 100 100	28, 46, 74, 100	0
1	F	219/224 (97%)	0.55	3 (1%) 75 81	29, 49, 79, 97	0
All	All	1315/1344 (97%)	0.61	28 (2%) 63 71	27, 46, 79, 117	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	45	ALA	6.7
1	C	48	ARG	3.8
1	C	16	LEU	3.6
1	A	5	LEU	3.4
1	B	13	VAL	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.