



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 10, 2018 – 11:13 PM EST

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 Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030736  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030736

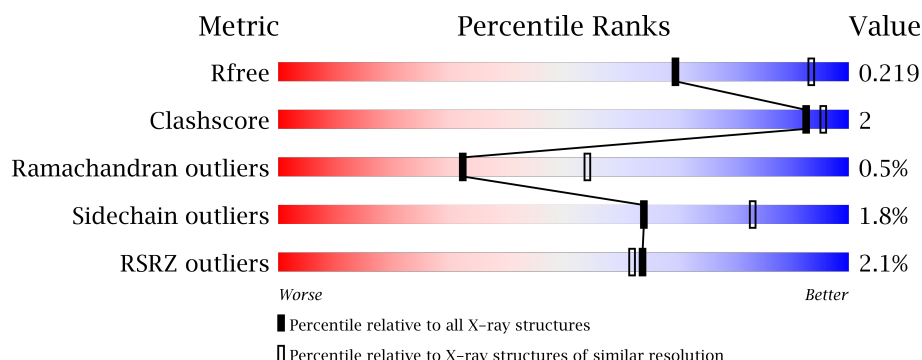
# 1 Overall quality at a glance

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))
$R_{free}$	100719	3689 (2.60-2.52)
Clashscore	112137	4096 (2.60-2.52)
Ramachandran outliers	110173	4037 (2.60-2.52)
Sidechain outliers	110143	4037 (2.60-2.52)
RSRZ outliers	101464	3700 (2.60-2.52)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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  $\leq 5\%$ . 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	<div> <div>3%</div> <div> <div></div> <div>90%</div> <div>8%</div> <div>.</div> </div> </div>
1	B	224	<div> <div>%</div> <div> <div></div> <div>93%</div> <div>.</div> <div>.</div> </div> </div>
1	C	224	<div> <div>5%</div> <div> <div></div> <div>90%</div> <div>5%</div> <div>.</div> <div>.</div> </div> </div>
1	D	224	<div> <div>2%</div> <div> <div></div> <div>93%</div> <div>6%</div> <div>.</div> </div> </div>
1	E	224	<div> <div></div> <div> <div></div> <div>92%</div> <div>5%</div> <div>.</div> <div>.</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	224	<div><div><div>%</div><div><div></div></div><div>91%</div><div>7%</div><div></div></div></div>

## 2 Entry composition

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

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

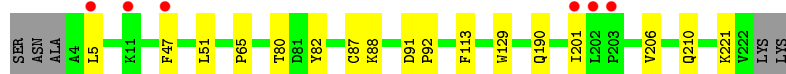
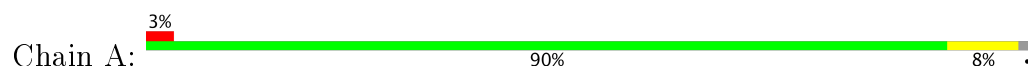
- Molecule 2 is water.

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

### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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



- Molecule 1: DsbA-like protein



- Molecule 1: DsbA-like protein



- Molecule 1: DsbA-like protein

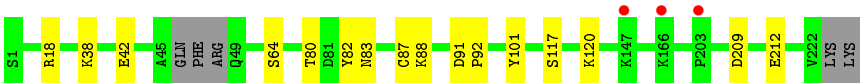


- Molecule 1: DsbA-like protein



- Molecule 1: DsbA-like protein





## 4 Data and refinement statistics

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	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
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.46 (at 2.55Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.171 , 0.222 0.171 , 0.219	Depositor DCC
$R_{free}$ test set	3009 reflections (4.85%)	DCC
Wilson B-factor (Å <sup>2</sup> )	45.0	Xtriage
Anisotropy	0.432	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 47.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	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 <sup>2</sup> )	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.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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	0.42	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(°)	Ideal(°)
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

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.

All (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
1:F:83:ASN:ND2	1:F:117:SER:OG	2.39	0.55
1:E:80:THR:HG1	1:E:87[A]:CYS:HB3	1.73	0.53
1:D:82:TYR:O	1:D:88:LYS:NZ	2.38	0.53
1:B:5:LEU:HD12	1:B:5:LEU:N	2.25	0.52
1:E:48:ARG:NH1	1:E:199:ASP:OD2	2.43	0.52
1:C:80:THR:HG1	1:C:87[A]:CYS:HB3	1.74	0.51
1:A:80:THR:HG1	1:A:87[A]:CYS:HB3	1.77	0.50
1:F:101:TYR:OH	1:F:212:GLU:OE1	2.27	0.49
1:A:206:VAL:HG12	1:A:210:GLN:HB3	1.95	0.48
1:D:103:ASP:N	1:D:103:ASP:OD1	2.40	0.47
1:F:80:THR:HG1	1:F:87[A]:CYS:HB3	1.79	0.47
1:B:80:THR:HG1	1:B:87[A]:CYS:HB3	1.80	0.47
1:D:199:ASP:OD1	1:D:218:GLN:NE2	2.48	0.47
1:F:91:ASP:HB3	1:F:92:PRO:HD3	1.98	0.45
1:D:30:GLU:OE2	1:F:18:ARG:NH2	2.50	0.44
1:A:113:PHE:CD1	1:E:185:ARG:HD3	2.52	0.44
1:E:103:ASP:OD1	1:E:103:ASP:N	2.37	0.44
1:A:65:PRO:HA	1:A:129:TRP:CE2	2.53	0.44
1:C:146:LYS:HD3	1:C:150:LEU:HD23	2.00	0.44
1:C:5:LEU:N	1:C:5:LEU:HD12	2.33	0.44
1:F:82:TYR:O	1:F:88:LYS:HE3	2.18	0.43
1:A:82:TYR:O	1:A:88:LYS:NZ	2.51	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:47:PHE:CE2	1:A:51:LEU:HD11	2.54	0.43
1:E:130:LYS:NZ	1:E:172:ASP:OD2	2.53	0.42
1:C:51:LEU:O	1:C:52:ALA:HB2	2.21	0.41
1:A:91:ASP:HB3	1:A:92:PRO:HD3	2.02	0.41
1:A:5:LEU:HD13	1:B:9:GLN:HG2	2.03	0.41
1:E:197:ILE:HG22	1:E:218:GLN:HG3	2.02	0.41
1:C:91:ASP:HB3	1:C:92:PRO:HD3	2.04	0.40
1:D:206:VAL:HG12	1:D:210:GLN:HB3	2.03	0.40
1:A:206:VAL:CG1	1:A:210:GLN:HB3	2.50	0.40
1:A:113:PHE:CE1	1:E:185:ARG:HD3	2.57	0.40

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%)	32	54
1	B	219/224 (98%)	215 (98%)	3 (1%)	1 (0%)	32	54
1	C	214/224 (96%)	208 (97%)	3 (1%)	3 (1%)	13	24
1	D	221/224 (99%)	219 (99%)	2 (1%)	0	100	100
1	E	221/224 (99%)	215 (97%)	5 (2%)	1 (0%)	32	54
1	F	217/224 (97%)	211 (97%)	6 (3%)	0	100	100
All	All	1311/1344 (98%)	1283 (98%)	22 (2%)	6 (0%)	32	54

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	52	ALA
1	C	49	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	221	LYS
1	C	6	ASN
1	B	112	PRO
1	E	112	PRO

### 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	189/191 (99%)	189 (100%)	0	100	100
1	B	189/191 (99%)	182 (96%)	7 (4%)	39	62
1	C	186/191 (97%)	180 (97%)	6 (3%)	44	68
1	D	190/191 (100%)	188 (99%)	2 (1%)	78	90
1	E	190/191 (100%)	186 (98%)	4 (2%)	59	80
1	F	188/191 (98%)	185 (98%)	3 (2%)	68	85
All	All	1132/1146 (99%)	1110 (98%)	22 (2%)	64	82

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	5	LEU
1	B	53	SER
1	B	84[A]	CYS
1	B	84[B]	CYS
1	B	120	LYS
1	B	172	ASP
1	B	209	ASP
1	C	5	LEU
1	C	51	LEU
1	C	53	SER
1	C	84[A]	CYS
1	C	84[B]	CYS
1	C	116	GLU
1	D	175	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	209	ASP
1	E	39	LYS
1	E	53	SER
1	E	103	ASP
1	E	172	ASP
1	F	64	SER
1	F	120	LYS
1	F	209	ASP

Some 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 carbohydrates 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 ⓘ

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	219/224 (97%)	0.62	6 (2%) 55 52	27, 41, 73, 107	0
1	B	219/224 (97%)	0.62	2 (0%) 84 84	30, 45, 82, 117	0
1	C	216/224 (96%)	0.67	12 (5%) 25 22	32, 50, 82, 106	0
1	D	221/224 (98%)	0.62	5 (2%) 61 58	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 74	29, 49, 79, 97	0
All	All	1315/1344 (97%)	0.61	28 (2%) 64 62	27, 46, 79, 117	0

All (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
1	C	203	PRO	3.0
1	A	201	ILE	2.8
1	A	47	PHE	2.7
1	A	11	LYS	2.7
1	C	44	GLN	2.7
1	A	202	LEU	2.6
1	B	190	GLN	2.5
1	D	209	ASP	2.5
1	C	147	LYS	2.4
1	D	89	ARG	2.4
1	D	47	PHE	2.3
1	C	12	GLU	2.3
1	D	202	LEU	2.3
1	F	147	LYS	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	157	ASP	2.2
1	C	22	VAL	2.2
1	F	166	LYS	2.2
1	C	13	VAL	2.2
1	C	106	VAL	2.2
1	C	153	ALA	2.1
1	A	203	PRO	2.1
1	F	203	PRO	2.0
1	D	214	ILE	2.0

## 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 carbohydrates 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.