wwPDB X-ray Structure Validation Summary Report (i)

May 12, 2020 – 11:25 pm BST

PDB ID	:	5O2K
Title	:	Native apo-structure of Pseudomonas stutzeri PtxB to 2.1 A resolution
Authors	:	Bisson, C.; Hitchcock, A.
Deposited on	:	2017-05-21
Resolution	:	2.10 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 following versions of software and data (see references (1)) were used in the production of this report:

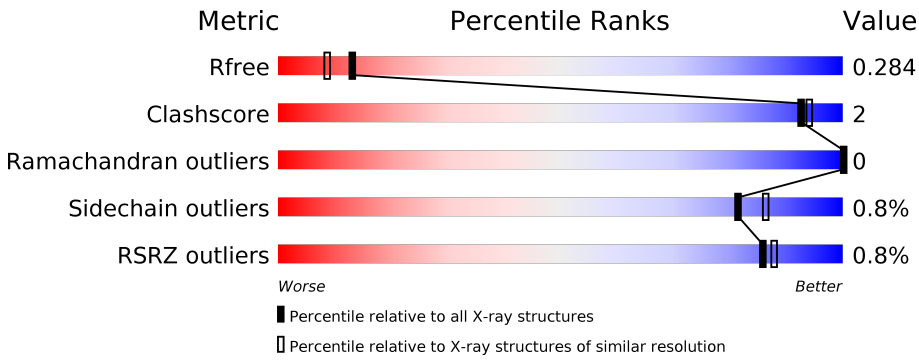
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 >=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	273	95% • •
1	B	273	91% 5% •
1	C	273	88% 7% 5%
1	D	273	91% • 5%
1	E	273	89% 6% 5%
1	F	273	4% 88% 6% 6%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 12355 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 Probable phosphite transport system-binding protein PtxB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	D	259	2003	1271	335	390	7	0	0	0
			Total	C	N	O	S			
1	A	266	2053	1301	342	402	8	0	0	0
			Total	C	N	O	S			
1	B	262	2021	1281	338	395	7	0	0	0
			Total	C	N	O	S			
1	C	259	2003	1271	335	390	7	0	0	0
			Total	C	N	O	S			
1	E	258	1995	1267	334	387	7	0	0	0
			Total	C	N	O	S			
1	F	256	1980	1257	332	384	7	0	0	0

There are 54 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	1	MET	-	initiating methionine	UNP O69052
D	266	LEU	-	expression tag	UNP O69052
D	267	GLU	-	expression tag	UNP O69052
D	268	HIS	-	expression tag	UNP O69052
D	269	HIS	-	expression tag	UNP O69052
D	270	HIS	-	expression tag	UNP O69052
D	271	HIS	-	expression tag	UNP O69052
D	272	HIS	-	expression tag	UNP O69052
D	273	HIS	-	expression tag	UNP O69052
A	1	MET	-	initiating methionine	UNP O69052
A	266	LEU	-	expression tag	UNP O69052
A	267	GLU	-	expression tag	UNP O69052
A	268	HIS	-	expression tag	UNP O69052
A	269	HIS	-	expression tag	UNP O69052
A	270	HIS	-	expression tag	UNP O69052
A	271	HIS	-	expression tag	UNP O69052
A	272	HIS	-	expression tag	UNP O69052

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	273	HIS	-	expression tag	UNP O69052
B	1	MET	-	initiating methionine	UNP O69052
B	266	LEU	-	expression tag	UNP O69052
B	267	GLU	-	expression tag	UNP O69052
B	268	HIS	-	expression tag	UNP O69052
B	269	HIS	-	expression tag	UNP O69052
B	270	HIS	-	expression tag	UNP O69052
B	271	HIS	-	expression tag	UNP O69052
B	272	HIS	-	expression tag	UNP O69052
B	273	HIS	-	expression tag	UNP O69052
C	1	MET	-	initiating methionine	UNP O69052
C	266	LEU	-	expression tag	UNP O69052
C	267	GLU	-	expression tag	UNP O69052
C	268	HIS	-	expression tag	UNP O69052
C	269	HIS	-	expression tag	UNP O69052
C	270	HIS	-	expression tag	UNP O69052
C	271	HIS	-	expression tag	UNP O69052
C	272	HIS	-	expression tag	UNP O69052
C	273	HIS	-	expression tag	UNP O69052
E	1	MET	-	initiating methionine	UNP O69052
E	266	LEU	-	expression tag	UNP O69052
E	267	GLU	-	expression tag	UNP O69052
E	268	HIS	-	expression tag	UNP O69052
E	269	HIS	-	expression tag	UNP O69052
E	270	HIS	-	expression tag	UNP O69052
E	271	HIS	-	expression tag	UNP O69052
E	272	HIS	-	expression tag	UNP O69052
E	273	HIS	-	expression tag	UNP O69052
F	1	MET	-	initiating methionine	UNP O69052
F	266	LEU	-	expression tag	UNP O69052
F	267	GLU	-	expression tag	UNP O69052
F	268	HIS	-	expression tag	UNP O69052
F	269	HIS	-	expression tag	UNP O69052
F	270	HIS	-	expression tag	UNP O69052
F	271	HIS	-	expression tag	UNP O69052
F	272	HIS	-	expression tag	UNP O69052
F	273	HIS	-	expression tag	UNP O69052

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	D	42	42	42	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	45	45	45	0	0
			Total	O		
2	B	44	44	44	0	0
			Total	O		
2	C	61	61	61	0	0
			Total	O		
2	E	59	59	59	0	0
			Total	O		
2	F	49	49	49	0	0

3 Residue-property plots (i)

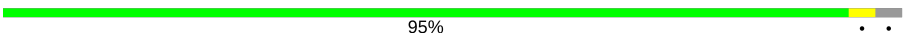
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

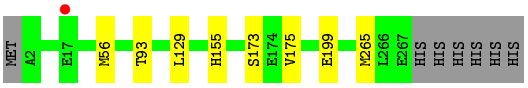
- Molecule 1: Probable phosphite transport system-binding protein PtxB

Chain D:

- Molecule 1: Probable phosphite transport system-binding protein PtxB

Chain A:





- Molecule 1: Probable phosphite transport system-binding protein PtxB

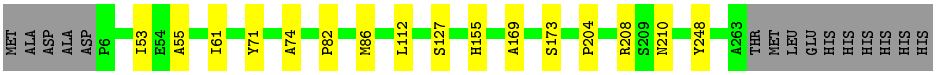
Chain B:

- Molecule 1: Probable phosphite transport system-binding protein PtxB

Chain C:

- Molecule 1: Probable phosphite transport system-binding protein PtxB

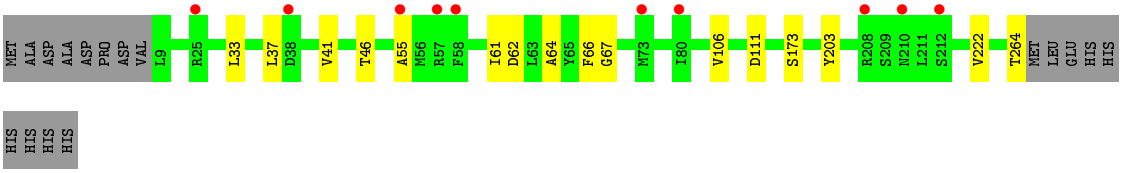
Chain E:



- Molecule 1: Probable phosphite transport system-binding protein PtxB

Chain F:

4%
88%
6%
6%


MET
ALA
ASP
ALA
ASP
PRO
ASP
VAL
L9
R25
L33
L37
D38
V41
T46
A55
M56
R57
F58
I61
D62
L63
A64
Y65
F66
G67
M73
I80
V106
D111
S173
Y203
R208
S209
N210
L211
S212
V222
T264
MET
LEU
GLU
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	87.95Å 136.64Å 90.84Å 90.00° 115.02° 90.00°	Depositor
Resolution (Å)	79.69 – 2.10 79.69 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.6 (79.69-2.10) 99.6 (79.69-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.61 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.236 , 0.282 0.241 , 0.284	Depositor DCC
R_{free} test set	5605 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	25.3	Xtriage
Anisotropy	0.481	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 36.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.028 for l,-k,h	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	12355	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.77% 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	0.60	0/2093	0.76	0/2839
1	B	0.59	0/2061	0.74	0/2796
1	C	0.67	0/2043	0.79	2/2771 (0.1%)
1	D	0.64	0/2043	0.78	0/2771
1	E	0.62	0/2035	0.76	0/2759
1	F	0.60	0/2019	0.77	0/2737
All	All	0.62	0/12294	0.77	2/16673 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	252	ARG	NE-CZ-NH1	5.60	123.10	120.30
1	C	183	ARG	NE-CZ-NH1	5.43	123.02	120.30

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	2053	0	2028	5	0
1	B	2021	0	1995	7	0
1	C	2003	0	1981	8	0
1	D	2003	0	1981	4	0
1	E	1995	0	1978	7	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	1980	0	1964	6	0
2	A	45	0	0	1	0
2	B	44	0	0	0	0
2	C	61	0	0	0	0
2	D	42	0	0	0	0
2	E	59	0	0	0	0
2	F	49	0	0	0	0
All	All	12355	0	11927	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:E:55:ALA:HB3	1:E:61:ILE:HD12	1.83	0.61
1:C:93:THR:HG23	1:C:199:GLU:OE1	2.06	0.55
1:A:93:THR:HG23	1:A:199:GLU:OE1	2.08	0.54
1:E:53:ILE:HG12	1:E:74:ALA:HB2	1.90	0.53
1:B:127:SER:HB3	1:B:155:HIS:CD2	2.44	0.52

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	264/273 (97%)	251 (95%)	13 (5%)	0	100	100
1	B	260/273 (95%)	248 (95%)	12 (5%)	0	100	100
1	C	257/273 (94%)	244 (95%)	13 (5%)	0	100	100
1	D	257/273 (94%)	247 (96%)	10 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	256/273 (94%)	245 (96%)	11 (4%)	0	100	100
1	F	254/273 (93%)	242 (95%)	12 (5%)	0	100	100
All	All	1548/1638 (94%)	1477 (95%)	71 (5%)	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	220/227 (97%)	219 (100%)	1 (0%)	88	92
1	B	216/227 (95%)	216 (100%)	0	100	100
1	C	215/227 (95%)	214 (100%)	1 (0%)	88	92
1	D	215/227 (95%)	213 (99%)	2 (1%)	78	84
1	E	214/227 (94%)	212 (99%)	2 (1%)	78	84
1	F	212/227 (93%)	208 (98%)	4 (2%)	57	63
All	All	1292/1362 (95%)	1282 (99%)	10 (1%)	81	86

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	86	MET
1	E	173	SER
1	F	62	ASP
1	C	7	ASP
1	F	46	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	166	ASN

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 (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	266/273 (97%)	-0.03	1 (0%) 92 93	18, 34, 55, 71	0
1	B	262/273 (95%)	0.05	0 100 100	18, 37, 60, 73	0
1	C	259/273 (94%)	-0.01	1 (0%) 92 93	16, 29, 61, 74	0
1	D	259/273 (94%)	-0.16	0 100 100	18, 31, 49, 61	0
1	E	258/273 (94%)	-0.11	0 100 100	18, 29, 55, 66	0
1	F	256/273 (93%)	0.30	10 (3%) 39 45	17, 35, 79, 105	0
All	All	1560/1638 (95%)	0.00	12 (0%) 86 88	16, 33, 63, 105	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	58	PHE	6.7
1	F	25	ARG	5.0
1	F	210	ASN	3.4
1	F	208	ARG	3.1
1	F	80	ILE	3.1

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.