



wwPDB X-ray Structure Validation Summary Report i

Feb 27, 2014 – 06:48 AM GMT

PDB ID : 3PSP
Title : Crystal structure of PUL and PFU domain
Authors : Liu, Y.; Sun, J.
Deposited on : 2010-12-02
Resolution : 2.42 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

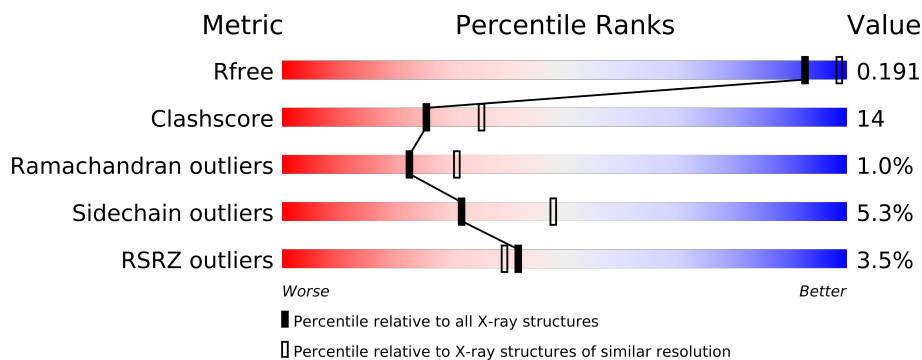
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance


The reported resolution of this entry is 2.42 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}	66092	2496 (2.44-2.40)
Clashscore	79885	3124 (2.44-2.40)
Ramachandran outliers	78287	3067 (2.44-2.40)
Sidechain outliers	78261	3068 (2.44-2.40)
RSRZ outliers	66119	2499 (2.44-2.40)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	425	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2710 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Protein DOA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	317	Total	C	N	O	S	0	0	0
			2549	1638	408	497	6			

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	291	MET	-	EXPRESSION TAG	UNP P36037
A	292	GLY	-	EXPRESSION TAG	UNP P36037
A	293	SER	-	EXPRESSION TAG	UNP P36037
A	294	SER	-	EXPRESSION TAG	UNP P36037
A	295	HIS	-	EXPRESSION TAG	UNP P36037
A	296	HIS	-	EXPRESSION TAG	UNP P36037
A	297	HIS	-	EXPRESSION TAG	UNP P36037
A	298	HIS	-	EXPRESSION TAG	UNP P36037
A	299	HIS	-	EXPRESSION TAG	UNP P36037
A	300	HIS	-	EXPRESSION TAG	UNP P36037
A	301	SER	-	EXPRESSION TAG	UNP P36037
A	302	SER	-	EXPRESSION TAG	UNP P36037
A	303	GLY	-	EXPRESSION TAG	UNP P36037
A	304	LEU	-	EXPRESSION TAG	UNP P36037
A	305	VAL	-	EXPRESSION TAG	UNP P36037
A	306	PRO	-	EXPRESSION TAG	UNP P36037
A	307	ARG	-	EXPRESSION TAG	UNP P36037
A	308	GLY	-	EXPRESSION TAG	UNP P36037
A	309	SER	-	EXPRESSION TAG	UNP P36037
A	310	HIS	-	EXPRESSION TAG	UNP P36037
A	311	MET	-	EXPRESSION TAG	UNP P36037
A	312	ALA	-	EXPRESSION TAG	UNP P36037
A	313	SER	-	EXPRESSION TAG	UNP P36037
A	314	MET	-	EXPRESSION TAG	UNP P36037
A	315	THR	-	EXPRESSION TAG	UNP P36037
A	316	GLY	-	EXPRESSION TAG	UNP P36037
A	317	GLY	-	EXPRESSION TAG	UNP P36037

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	318	GLN	-	EXPRESSION TAG	UNP P36037
A	319	GLN	-	EXPRESSION TAG	UNP P36037
A	320	MET	-	EXPRESSION TAG	UNP P36037
A	321	GLY	-	EXPRESSION TAG	UNP P36037
A	322	ARG	-	EXPRESSION TAG	UNP P36037
A	323	GLY	-	EXPRESSION TAG	UNP P36037
A	324	SER	-	EXPRESSION TAG	UNP P36037

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	161	Total 161	O 161	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α , β , γ	102.50Å 102.50Å 73.31Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	25.62 – 2.42 44.38 – 2.42	Depositor EDS
% Data completeness (in resolution range)	95.3 (25.62-2.42) 99.4 (44.38-2.42)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.03 (at 2.42Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, R_{free}	0.185 , 0.234 0.187 , 0.191	Depositor DCC
R_{free} test set	842 reflections (5.07%)	DCC
Wilson B-factor (Å ²)	37.6	Xtriage
Anisotropy	0.062	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 45.8	EDS
Estimated twinning fraction	0.064 for h,-h-k,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 16675 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	2710	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.58% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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	0.39	0/2600	0.55	0/3533

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2549	0	2523	69	0
2	A	161	0	0	2	0
All	All	2710	0	2523	69	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 14.

The worst 5 of 69 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:493:ASN:O	1:A:495:LYS:N	1.99	0.94
1:A:604:ASP:OD1	1:A:653:LYS:HE3	1.80	0.81
1:A:512:ILE:HG23	1:A:543:ILE:HG22	1.63	0.81
1:A:469:VAL:H	1:A:676:ASN:HD21	1.28	0.79
1:A:661:GLN:HE22	1:A:696:ASN:HD22	1.34	0.72

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	313/425 (74%)	293 (94%)	17 (5%)	3 (1%)	22	31

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	464	MET
1	A	494	GLU
1	A	409	ASN

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	283/372 (76%)	268 (95%)	15 (5%)	32	48

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	531	GLU
1	A	557	ILE
1	A	653	LYS
1	A	513	ASP
1	A	616	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	524	ASN
1	A	594	GLN
1	A	661	GLN
1	A	510	HIS
1	A	616	GLN

5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates 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	317/425 (74%)	-0.16	11 (3%) 42 39	22, 37, 65, 87	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	376	LYS	4.9
1	A	500	GLU	3.2
1	A	498	ASP	3.0
1	A	606	GLU	2.8
1	A	378	ILE	2.7

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 carbohydrates 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.