



wwPDB X-ray Structure Validation Summary Report ⓘ

May 23, 2024 – 02:09 PM EDT

PDB ID : 2W4S
Title : novel RNA-binding domain in *Cryptosporidium parvum* at 2.5 angstrom resolution
Authors : Varrot, A.; Mackereth, C.; Mourao, A.; Sattler, M.; Cusack, S.
Deposited on : 2008-12-02
Resolution : 2.45 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtrriage (Phenix) : 1.13
EDS : **FAILED**
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36.2

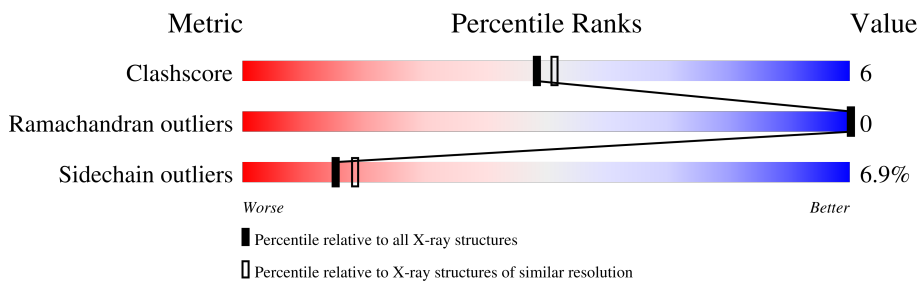
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.45 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))
Clashscore	141614	1613 (2.48-2.44)
Ramachandran outliers	138981	1598 (2.48-2.44)
Sidechain outliers	138945	1598 (2.48-2.44)

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 $\leq 5\%$.

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	113	62% 14% 24%
1	B	113	58% 16% • 25%
1	C	113	61% 13% • 24%
1	D	113	64% 11% • 24%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2820 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 Ankyrin-repeat protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	86	684	432	118	130	4	0	0	0
1	B	85	676	428	117	127	4	0	0	0
1	C	86	684	432	118	130	4	0	0	0
1	D	86	684	432	118	130	4	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	384	GLY	-	expression tag	UNP A0A7G2HK26
A	385	ALA	-	expression tag	UNP A0A7G2HK26
A	386	MET	-	expression tag	UNP A0A7G2HK26
B	384	GLY	-	expression tag	UNP A0A7G2HK26
B	385	ALA	-	expression tag	UNP A0A7G2HK26
B	386	MET	-	expression tag	UNP A0A7G2HK26
C	384	GLY	-	expression tag	UNP A0A7G2HK26
C	385	ALA	-	expression tag	UNP A0A7G2HK26
C	386	MET	-	expression tag	UNP A0A7G2HK26
D	384	GLY	-	expression tag	UNP A0A7G2HK26
D	385	ALA	-	expression tag	UNP A0A7G2HK26
D	386	MET	-	expression tag	UNP A0A7G2HK26

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	31	Total	O	0	0
			31	31		
2	B	26	Total	O	0	0
			26	26		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	23	Total 23	O 23	0	0
2	D	12	Total 12	O 12	0	0

4 Data and refinement statistics

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	42.18Å 52.86Å 99.31Å 90.00° 94.05° 90.00°	Depositor
Resolution (Å)	20.00 – 2.45	Depositor
% Data completeness (in resolution range)	94.5 (20.00-2.45)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.65 (at 2.46Å)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.193 , 0.254	Depositor
Wilson B-factor (Å ²)	35.4	Xtrriage
Anisotropy	0.676	Xtrriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	2820	wwPDB-VP
Average B, all atoms (Å ²)	38.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.99% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² 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	0.59	0/689	0.70	0/925
1	B	0.62	0/681	0.69	0/914
1	C	0.65	0/689	0.77	0/925
1	D	0.55	0/689	0.66	0/925
All	All	0.60	0/2748	0.71	0/3689

There are no bond length outliers.

There are no bond angle outliers.

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	684	0	707	7	0
1	B	676	0	703	8	0
1	C	684	0	707	13	0
1	D	684	0	707	11	0
2	A	31	0	0	0	0
2	B	26	0	0	0	0
2	C	23	0	0	0	0
2	D	12	0	0	1	0
All	All	2820	0	2824	33	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 6.

The worst 5 of 33 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:C:392:ILE:HD13	1:D:419:GLN:HG2	1.74	0.68
1:C:388:THR:HG23	1:D:419:GLN:NE2	2.12	0.64
1:C:392:ILE:CD1	1:D:419:GLN:HG2	2.30	0.62
1:D:397:ILE:HG22	1:D:401:LEU:HD22	1.83	0.59
1:C:470:TRP:CZ2	1:C:474:ARG:HD2	2.40	0.56

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	82/113 (73%)	82 (100%)	0	0	100	100
1	B	81/113 (72%)	81 (100%)	0	0	100	100
1	C	82/113 (73%)	81 (99%)	1 (1%)	0	100	100
1	D	82/113 (73%)	82 (100%)	0	0	100	100
All	All	327/452 (72%)	326 (100%)	1 (0%)	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	76/100 (76%)	71 (93%)	5 (7%)	16	20
1	B	75/100 (75%)	68 (91%)	7 (9%)	9	9
1	C	76/100 (76%)	73 (96%)	3 (4%)	32	42
1	D	76/100 (76%)	70 (92%)	6 (8%)	12	14
All	All	303/400 (76%)	282 (93%)	21 (7%)	15	18

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	455	LEU
1	D	419	GLN
1	D	474	ARG
1	D	420	ARG
1	D	402	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	458	GLN
1	D	433	ASN
1	D	402	GLN
1	B	433	ASN
1	D	419	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

6.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

6.4 Ligands

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers

EDS failed to run properly - this section is therefore empty.