Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 06:09 AM GMT

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 Full wwPDB X-ray Structure Validation 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/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:

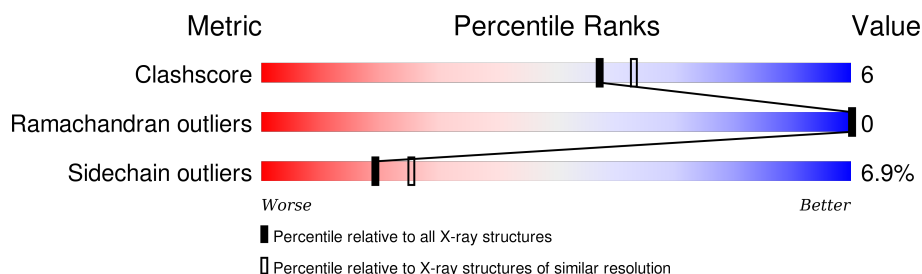|		
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	FAILED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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	Percentile Ranks	Value
Clashscore		6
Ramachandran outliers		0
Sidechain outliers		6.9%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	102246	1030 (2.48-2.44)
Ramachandran outliers	100387	1024 (2.48-2.44)
Sidechain outliers	100360	1024 (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 on the lower bar indicate the fraction of residues that contain outliers for >=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 <=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.

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 (i)

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
1	A	86	Total	C	N	O	S	0	0	0
			684	432	118	130	4			
1	B	85	Total	C	N	O	S	0	0	0
			676	428	117	127	4			
1	C	86	Total	C	N	O	S	0	0	0
			684	432	118	130	4			
1	D	86	Total	C	N	O	S	0	0	0
			684	432	118	130	4			

- 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		
2	C	23	Total	O	0	0
			23	23		
2	D	12	Total	O	0	0
			12	12		

# 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 errors 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.
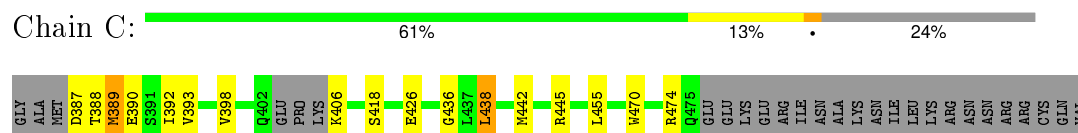
Note EDS failed to run properly.
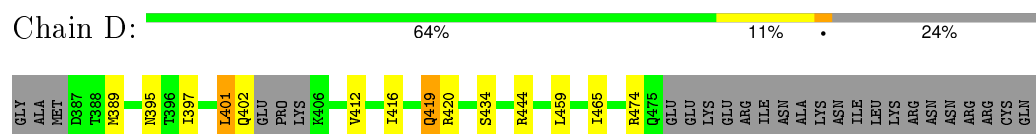
- Molecule 1: ANKYRIN-REPEAT PROTEIN
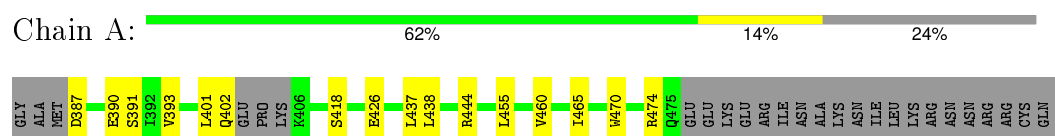
- Molecule 1: ANKYRIN-REPEAT PROTEIN
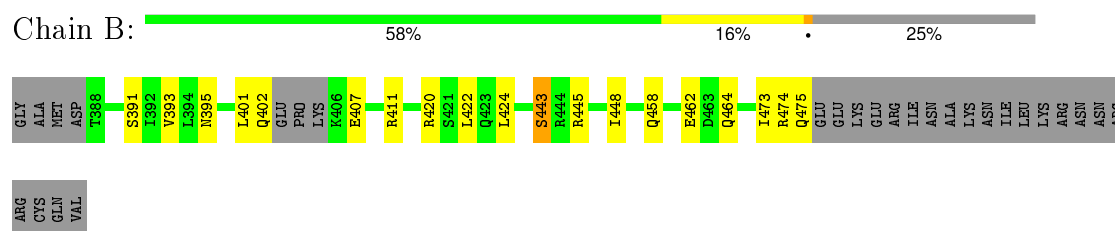
- Molecule 1: ANKYRIN-REPEAT PROTEIN

- Molecule 1: ANKYRIN-REPEAT PROTEIN

4 Data and refinement statistics (i)

EDS failed to run properly - this section will therefore be 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
$< I/\sigma(I) >$ [1]	2.65 (at 2.46Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.193 , 0.254	Depositor
Wilson B-factor (Å^2)	35.4	Xtriage
Anisotropy	0.676	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Outliers	0 of 15195 reflections	Xtriage
Total number of atoms	2820	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage'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 $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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	0.59	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 (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	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.

All (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
1:A:470:TRP:CH2	1:A:474:ARG:HD2	2.42	0.55
1:C:389:MET:O	1:C:393:VAL:HG23	2.08	0.53
1:C:426:GLU:OE2	1:D:395:ASN:HB3	2.10	0.52
1:C:388:THR:HG23	1:D:419:GLN:HE22	1.73	0.52
1:B:411:ARG:NH1	1:B:473:ILE:O	2.42	0.51
1:A:401:LEU:O	1:A:402:GLN:HB2	2.12	0.50
1:C:388:THR:O	1:C:392:ILE:HD12	2.11	0.49
1:D:459:LEU:HB2	1:D:465:ILE:HD12	1.94	0.49
1:B:393:VAL:HG13	1:B:422:LEU:HD13	1.94	0.48
1:A:470:TRP:CZ2	1:A:474:ARG:HD2	2.49	0.48
1:C:393:VAL:HG21	1:C:418:SER:HB2	1.95	0.48
1:A:393:VAL:HG21	1:A:418:SER:HB2	1.94	0.48
1:B:458:GLN:O	1:B:462:GLU:HB2	2.13	0.48
1:D:412:VAL:O	1:D:416:ILE:HG12	2.13	0.48
1:C:387:ASP:HA	1:C:390:GLU:HB2	1.97	0.47
1:B:401:LEU:HD23	1:B:448:ILE:HG12	1.97	0.46
1:C:436:GLY:HA3	1:C:445:ARG:HD3	1.97	0.46
1:A:460:VAL:HG22	1:A:465:ILE:HG13	1.97	0.45
1:A:387:ASP:OD1	1:A:390:GLU:OE2	2.34	0.45
1:C:398:VAL:HG21	1:C:406:LYS:HB3	1.99	0.45
1:B:401:LEU:CD2	1:B:448:ILE:HG12	2.49	0.43
1:C:438:LEU:HD23	1:C:442:MET:HG2	1.99	0.43
1:B:443:SER:OG	1:B:445:ARG:NH1	2.51	0.42
1:B:420:ARG:HH22	1:B:464:GLN:HB3	1.84	0.42
1:D:389:MET:HE3	1:D:389:MET:HA	2.02	0.42
1:D:444:ARG:NH1	2:D:2005:HOH:O	2.44	0.41
1:A:426:GLU:OE2	1:B:395:ASN:HB3	2.21	0.40
1:D:397:ILE:CG2	1:D:401:LEU:HD22	2.49	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	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%)	21	28
1	B	75/100 (75%)	68 (91%)	7 (9%)	11	13
1	C	76/100 (76%)	73 (96%)	3 (4%)	39	54
1	D	76/100 (76%)	70 (92%)	6 (8%)	15	20
All	All	303/400 (76%)	282 (93%)	21 (7%)	19	25

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	391	SER
1	A	437	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	438	LEU
1	A	444	ARG
1	A	455	LEU
1	B	391	SER
1	B	402	GLN
1	B	407	GLU
1	B	424	LEU
1	B	443	SER
1	B	474	ARG
1	B	475	GLN
1	C	389	MET
1	C	438	LEU
1	C	455	LEU
1	D	401	LEU
1	D	402	GLN
1	D	419	GLN
1	D	420	ARG
1	D	434	SER
1	D	474	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

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

EDS failed to run properly - this section will therefore be empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS failed to run properly - this section will therefore be empty.

6.3 Carbohydrates (i)

EDS failed to run properly - this section will therefore be empty.

6.4 Ligands (i)

EDS failed to run properly - this section will therefore be empty.

6.5 Other polymers (i)

EDS failed to run properly - this section will therefore be empty.