WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

May 21, 2024 – 04:08 PM JST

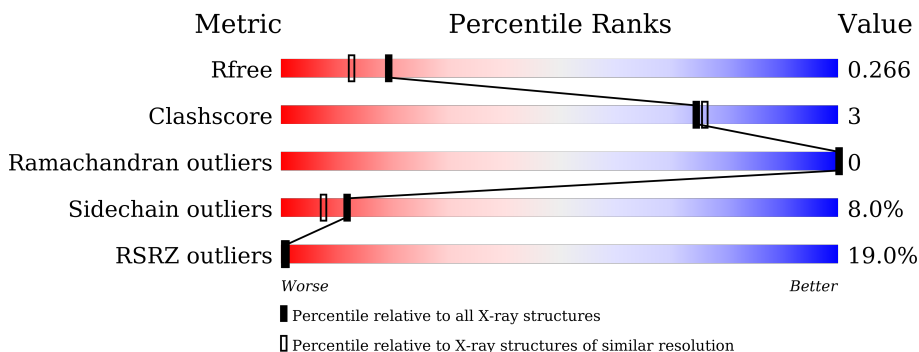|		
PDB ID	:	8YM2
Title	:	Crystal structure of AIDA-1 PTB domain in complex with SynGAP NPxF motif
Authors	:	Wang, X.; Wang, Y.; Cai, Q.; Zhang, M.
Deposited on	:	2024-03-08
Resolution	:	2.00 Å(reported)

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 2.00 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
Rfree		0.266
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		8.0%
RSRZ outliers		19.0%

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))
R_{free}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 <=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	162	19% / 79% / 12% · 8%
2	B	11	73% / 9% / 18%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 1295 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 and sterile alpha motif domain-containing protein 1B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	149	Total 1190	C 759	N 199	O 224	S 8	0	1	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1028	GLY	-	expression tag	UNP Q8BZM2
A	1029	PRO	-	expression tag	UNP Q8BZM2
A	1030	GLY	-	expression tag	UNP Q8BZM2
A	1031	SER	-	expression tag	UNP Q8BZM2

- Molecule 2 is a protein called Ras/Rap GTPase-activating protein SynGAP.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	9	Total 75	C 52	N 11	O 12	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	28	Total 28	O 28	0	0
3	B	2	Total 2	O 2	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Ankyrin repeat and sterile alpha motif domain-containing protein 1B

Chain A:
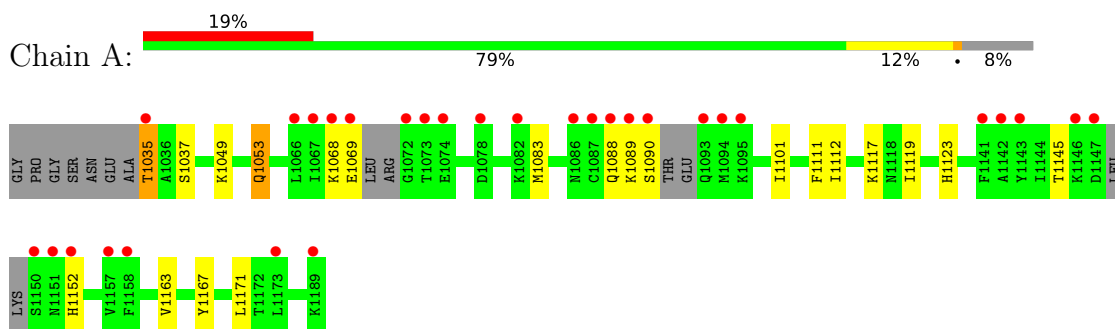
• Molecule 2: Ras/Rap GTPase-activating protein SynGAP

Chain B:
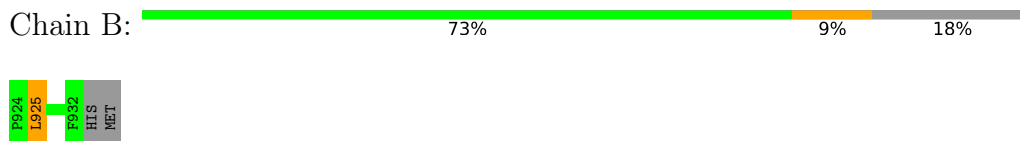
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	I 2 3	Depositor
Cell constants a, b, c, α, β, γ	112.92Å 112.92Å 112.92Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.26 – 2.00 25.25 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.8 (25.26-2.00) 97.9 (25.25-2.00)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.48 (at 2.01Å)	Xtriage
Refinement program	REFMAC 5.8.0352	Depositor
R, R_{free}	0.213 , 0.264 0.222 , 0.266	Depositor DCC
R_{free} test set	750 reflections (4.70%)	wwPDB-VP
Wilson B-factor (Å^2)	41.8	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 57.5	EDS
L-test for twinning[2]	$< \lvert L \rvert > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.034 for -l,-k,-h	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	1295	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.11% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< \lvert L \rvert >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5   Model quality ⓘ

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	0.40	0/1216	0.70	0/1640
2	B	0.38	0/78	0.96	0/105
All	All	0.40	0/1294	0.71	0/1745

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	1190	0	1169	8	0
2	B	75	0	73	1	0
3	A	28	0	0	1	2
3	B	2	0	0	0	0
All	All	1295	0	1242	8	2

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 3.

All (8) 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:1049:LYS:HA	1:A:1053:GLN:HG3	1.90	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1049:LYS:HE2	3:A:1219:HOH:O	2.10	0.52
1:A:1089:LYS:O	1:A:1090:SER:C	2.47	0.52
1:A:1035:THR:C	1:A:1037:SER:H	2.19	0.45
1:A:1123:HIS:CG	1:A:1145:THR:HG21	2.54	0.43
1:A:1112:ILE:HG22	1:A:1119:ILE:HD13	2.01	0.42
1:A:1101:ILE:O	1:A:1111:PHE:HA	2.20	0.41
1:A:1167:TYR:CD1	2:B:925:LEU:HD22	2.57	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:1211:HOH:O	3:A:1216:HOH:O[20_545]	2.05	0.15
3:A:1220:HOH:O	3:A:1221:HOH:O[4_556]	2.18	0.02

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	142/162 (88%)	137 (96%)	5 (4%)	0	100	100
2	B	7/11 (64%)	7 (100%)	0	0	100	100
All	All	149/173 (86%)	144 (97%)	5 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	129/138 (94%)	119 (92%)	10 (8%)	12	8
2	B	9/11 (82%)	8 (89%)	1 (11%)	6	3
All	All	138/149 (93%)	127 (92%)	11 (8%)	12	7

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1035	THR
1	A	1053	GLN
1	A	1068	LYS
1	A	1069	GLU
1	A	1083	MET
1	A	1088	GLN
1	A	1117	LYS
1	A	1152	HIS
1	A	1163	VAL
1	A	1171	LEU
2	B	925	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1088	GLN
1	A	1118	ASN

5.3.3 RNA ⓘ

There are no RNA molecules 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 monosaccharides 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	149/162 (91%)	1.03	30 (20%)	1	0	31, 50, 107, 126	0
2	B	9/11 (81%)	0.03	0	100	100	43, 50, 58, 64	0
All	All	158/173 (91%)	0.97	30 (18%)	1	1	31, 51, 107, 126	0

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1090	SER	7.2
1	A	1035	THR	5.5
1	A	1151	ASN	5.2
1	A	1147	ASP	5.0
1	A	1094	MET	5.0
1	A	1089	LYS	4.7
1	A	1074	GLU	4.6
1	A	1068	LYS	4.4
1	A	1069	GLU	4.3
1	A	1088	GLN	4.3
1	A	1152	HIS	4.2
1	A	1093	GLN	4.1
1	A	1073	THR	3.9
1	A	1150	SER	3.7
1	A	1158	PHE	3.3
1	A	1173	LEU	3.3
1	A	1141	PHE	3.2
1	A	1072	GLY	3.2
1	A	1086	ASN	3.1
1	A	1087	CYS	2.8
1	A	1143	TYR	2.6
1	A	1078[A]	ASP	2.6
1	A	1189	LYS	2.6
1	A	1066	LEU	2.6

Continued on next page...

Mol	Chain	Res	Type	RSRZ
1	A	1157	VAL	2.5
1	A	1095	LYS	2.4
1	A	1082	LYS	2.3
1	A	1142	ALA	2.3
1	A	1146	LYS	2.1
1	A	1067	ILE	2.0

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 monosaccharides 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.