W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 27, 2014 – 09:23 PM GMT

PDB ID	:	3HS9
Title	:	Intersectin 1-peptide-AP2 beta ear complex
Authors	:	Vahedi-Faridi, A.; Pechstein, A.; Schaefer, J.G.; Saenger, W.; Haucke, V.
Deposited on	:	2009-06-10
Resolution	:	2.15 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 2.15 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		0.280
Clashscore		17
Ramachandran outliers		0.8%
Sidechain outliers		8.7%
RSRZ outliers		11.5%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
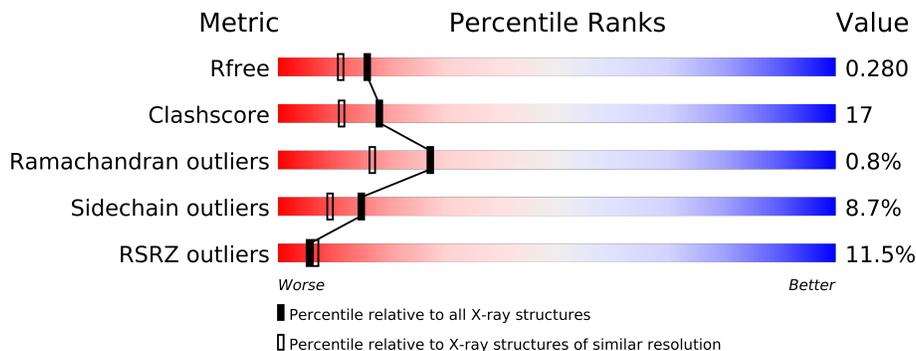
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	1094 (2.18-2.14)
Clashscore	79885	1299 (2.18-2.14)
Ramachandran outliers	78287	1272 (2.18-2.14)
Sidechain outliers	78261	1272 (2.18-2.14)
RSRZ outliers	66119	1094 (2.18-2.14)

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	260	
2	P	12	

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 2085 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 AP-2 complex subunit beta-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	233	1858	1194	312	341	11	0	0	0

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	678	MET	-	EXPRESSION TAG	UNP P62944
A	679	GLY	-	EXPRESSION TAG	UNP P62944
A	680	SER	-	EXPRESSION TAG	UNP P62944
A	681	SER	-	EXPRESSION TAG	UNP P62944
A	682	HIS	-	EXPRESSION TAG	UNP P62944
A	683	HIS	-	EXPRESSION TAG	UNP P62944
A	684	HIS	-	EXPRESSION TAG	UNP P62944
A	685	HIS	-	EXPRESSION TAG	UNP P62944
A	686	HIS	-	EXPRESSION TAG	UNP P62944
A	687	HIS	-	EXPRESSION TAG	UNP P62944
A	688	SER	-	EXPRESSION TAG	UNP P62944
A	689	SER	-	EXPRESSION TAG	UNP P62944
A	690	GLY	-	EXPRESSION TAG	UNP P62944
A	691	LEU	-	EXPRESSION TAG	UNP P62944
A	692	VAL	-	EXPRESSION TAG	UNP P62944
A	693	PRO	-	EXPRESSION TAG	UNP P62944
A	694	ARG	-	EXPRESSION TAG	UNP P62944
A	695	GLY	-	EXPRESSION TAG	UNP P62944
A	696	SER	-	EXPRESSION TAG	UNP P62944
A	697	HIS	-	EXPRESSION TAG	UNP P62944
A	698	MET	-	EXPRESSION TAG	UNP P62944
A	699	ALA	-	EXPRESSION TAG	UNP P62944
A	700	SER	-	EXPRESSION TAG	UNP P62944

- Molecule 2 is a protein called peptide from Intersectin-1, residues 841-851.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	P	11	Total 94	C 61	N 15	O 18	0	0	0

• Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	133	Total 133	O 133	0	0

3 Residue-property plots ⓘ

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.

• Molecule 1: AP-2 complex subunit beta-1

Chain A:
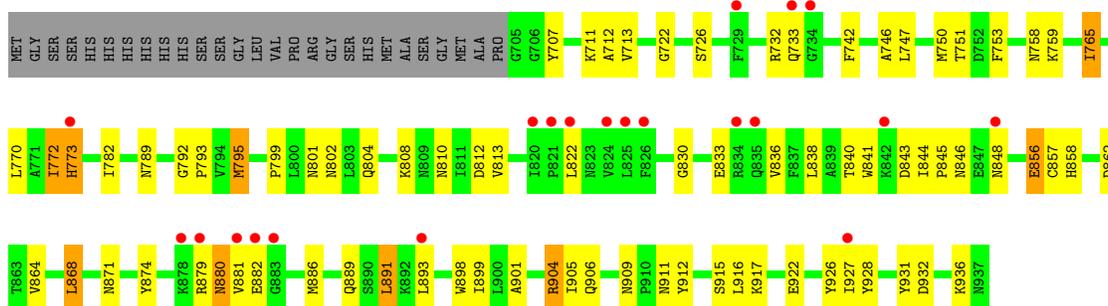
• Molecule 2: peptide from Intersectin-1, residues 841-851

Chain P:
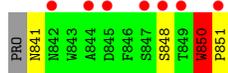
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	122.09Å 47.10Å 58.95Å 90.00° 97.66° 90.00°	Depositor
Resolution (Å)	50.00 – 2.15 27.79 – 1.95	Depositor EDS
% Data completeness (in resolution range)	95.7 (50.00-2.15) 95.8 (27.79-1.95)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	0.04	Depositor
$< I/\sigma(I) >$ [1]	2.22 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.3.0040	Depositor
R, R_{free}	0.249 , 0.283 0.254 , 0.280	Depositor DCC
R_{free} test set	743 reflections (4.42%)	DCC
Wilson B-factor (Å^2)	37.9	Xtriage
Anisotropy	0.852	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 66.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Outliers	0 of 23437 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	2085	wwPDB-VP
Average B, all atoms (Å^2)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.59% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities 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	0.60	0/1901	0.73	0/2583
2	P	0.92	1/99 (1.0%)	0.82	0/137
All	All	0.62	1/2000 (0.1%)	0.73	0/2720

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	P	850	TRP	CB-CG	-5.24	1.40	1.50

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	1858	0	1868	54	0
2	P	94	0	73	12	0
3	A	133	0	0	4	0
All	All	2085	0	1941	66	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 17.

The worst 5 of 66 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:707:TYR:H	1:A:871:ASN:HD21	1.04	0.95
2:P:850:TRP:H	2:P:851:PRO:CD	1.80	0.92
1:A:751:THR:H	1:A:810:ASN:ND2	1.69	0.90
2:P:850:TRP:N	2:P:850:TRP:CD1	2.46	0.80
2:P:850:TRP:N	2:P:851:PRO:CD	2.47	0.74

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	231/260 (89%)	216 (94%)	14 (6%)	1 (0%)	43	39
2	P	9/12 (75%)	5 (56%)	3 (33%)	1 (11%)	1	0
All	All	240/272 (88%)	221 (92%)	17 (7%)	2 (1%)	27	17

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	P	850	TRP
1	A	856	GLU

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	209/230 (91%)	192 (92%)	17 (8%)	17	11
2	P	10/11 (91%)	8 (80%)	2 (20%)	2	1
All	All	219/241 (91%)	200 (91%)	19 (9%)	15	9

wwPDB X-ray Structure Validation Summary Report

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	843	ASP
1	A	868	LEU
1	A	922	GLU
1	A	838	LEU
1	A	928	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	858	HIS
1	A	937	ASN
1	A	880	ASN
1	A	773	HIS
1	A	871	ASN

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	233/260 (89%)	0.61	21 (9%)	10	11	35, 50, 77, 85	0
2	P	11/12 (91%)	2.92	7 (63%)	0	0	20, 70, 72, 73	0
All	All	244/272 (89%)	0.71	28 (11%)	5	7	20, 51, 76, 85	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	P	851	PRO	8.2
2	P	848	SER	4.5
2	P	849	THR	4.2
1	A	733	GLN	3.8
1	A	826	PHE	3.2

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.