



Full wwPDB X-ray Structure Validation Report ⓘ

Jan 4, 2024 – 04:14 am GMT

PDB ID : 5FPI
Title : Mu2 adaptin subunit of the AP2 adaptor (C-terminal domain) complexed with Integrin alpha4 internalisation peptide QYKSILQE
Authors : Owen, D.J.; Evans, P.R.; Ivaska, J.
Deposited on : 2015-11-30
Resolution : 2.77 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

<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
Xtriage (Phenix) : 1.13
EDS : 2.36
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

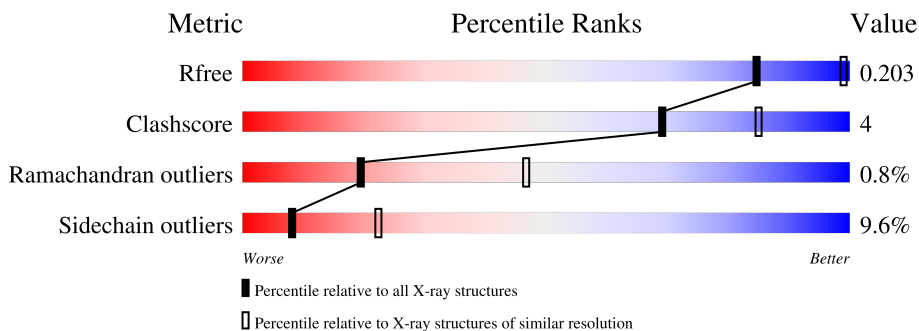
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.77 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}	130704	4107 (2.80-2.76)
Clashscore	141614	4575 (2.80-2.76)
Ramachandran outliers	138981	4487 (2.80-2.76)
Sidechain outliers	138945	4489 (2.80-2.76)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	446	48% (green), 8% (yellow), 8% (orange), 36% (grey)
2	B	8	62% (green), 12% (yellow), 25% (grey)

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2175 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 AP-2 COMPLEX SUBUNIT MU.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	258	2078	1335	366	363	14	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	237	MET	-	insertion	UNP P84092
A	237A	GLU	-	insertion	UNP P84092
A	237B	GLN	-	insertion	UNP P84092
A	237C	LYS	-	insertion	UNP P84092
A	237D	LEU	-	insertion	UNP P84092
A	237E	ILE	-	insertion	UNP P84092
A	237F	SER	-	insertion	UNP P84092
A	237G	GLU	-	insertion	UNP P84092
A	237H	GLU	-	insertion	UNP P84092
A	237I	ASP	-	insertion	UNP P84092
A	237J	LEU	-	insertion	UNP P84092

- Molecule 2 is a protein called INTEGRIN ALPHA-4 SUBUNIT.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	6	52	35	8	9	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	42	Total	O	0	0
			42	42		
3	B	3	Total	O	0	0
			3	3		

4 Data and refinement statistics i

Property	Value	Source
Space group	P 64	Depositor
Cell constants a, b, c, α , β , γ	125.65Å 125.65Å 74.41Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	108.82 – 2.77 62.83 – 2.77	Depositor EDS
% Data completeness (in resolution range)	99.3 (108.82-2.77) 99.2 (62.83-2.77)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.96 (at 2.77Å)	Xtrriage
Refinement program	REFMAC 5.8.0123	Depositor
R, R_{free}	0.174 , 0.224 0.208 , 0.203	Depositor DCC
R_{free} test set	862 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å ²)	86.3	Xtrriage
Anisotropy	0.042	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 48.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.033 for h,-h-k,-l	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	2175	wwPDB-VP
Average B, all atoms (Å ²)	61.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.26% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 [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	0.96	0/2119	1.07	9/2848 (0.3%)
2	B	1.16	0/52	1.17	1/68 (1.5%)
All	All	0.96	0/2171	1.08	10/2916 (0.3%)

There are no bond length outliers.

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	253	ARG	NE-CZ-NH1	6.69	123.65	120.30
1	A	202	MET	CG-SD-CE	5.98	109.77	100.20
1	A	294	ARG	NE-CZ-NH2	-5.65	117.47	120.30
1	A	189	GLY	N-CA-C	5.62	127.16	113.10
2	B	3	LYS	CD-CE-NZ	5.58	124.54	111.70
1	A	170	ARG	NE-CZ-NH1	5.51	123.06	120.30
1	A	434	ARG	NE-CZ-NH2	-5.42	117.59	120.30
1	A	162	ARG	NE-CZ-NH1	5.35	122.97	120.30
1	A	269	ASP	CB-CG-OD2	-5.29	113.54	118.30
1	A	199	ARG	NE-CZ-NH1	5.16	122.88	120.30

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	2078	0	2170	16	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	52	0	59	0	0
3	A	42	0	0	3	0
3	B	3	0	0	0	0
All	All	2175	0	2229	16	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 4.

All (16) 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:256:LYS:O	1:A:260:GLU:N	2.16	0.78
1:A:218:ASP:OD2	1:A:261:ARG:NH1	2.32	0.62
1:A:399:LEU:HD23	1:A:400:LYS:N	2.16	0.60
1:A:323:ARG:NH2	1:A:348:GLU:OE1	2.38	0.53
1:A:380:LYS:HG2	1:A:381:TRP:N	2.25	0.52
1:A:256:LYS:C	1:A:260:GLU:N	2.68	0.47
1:A:163:ARG:HG2	3:A:2001:HOH:O	2.15	0.47
1:A:399:LEU:HD23	1:A:399:LEU:C	2.35	0.47
1:A:294:ARG:HD3	3:A:2015:HOH:O	2.17	0.45
1:A:212:CYS:HA	1:A:405:LYS:O	2.17	0.45
1:A:278:ARG:HG3	1:A:278:ARG:HH11	1.84	0.43
1:A:249:HIS:HD1	1:A:251:CYS:HG	1.68	0.42
1:A:209:MET:N	1:A:408:GLU:OE1	2.47	0.42
1:A:311:PHE:O	1:A:312:LYS:C	2.58	0.42
1:A:265:PHE:CD1	1:A:265:PHE:N	2.88	0.42
1:A:199:ARG:HD3	3:A:2009:HOH:O	2.21	0.41

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	252/446 (56%)	234 (93%)	16 (6%)	2 (1%)	19	47
2	B	4/8 (50%)	4 (100%)	0	0	100	100
All	All	256/454 (56%)	238 (93%)	16 (6%)	2 (1%)	19	47

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	189	GLY
1	A	376	ASN

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	233/398 (58%)	210 (90%)	23 (10%)	8	21
2	B	6/8 (75%)	6 (100%)	0	100	100
All	All	239/406 (59%)	216 (90%)	23 (10%)	8	22

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	163	ARG
1	A	180	SER
1	A	188	GLN
1	A	197	SER
1	A	213	LYS
1	A	221	VAL
1	A	240	SER
1	A	241	ILE
1	A	253	ARG
1	A	256	LYS
1	A	261	ARG
1	A	276	ARG
1	A	292	LEU
1	A	300	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	304	LYS
1	A	309	SER
1	A	323	ARG
1	A	331	SER
1	A	341	LYS
1	A	376	ASN
1	A	377	ASP
1	A	380	LYS
1	A	434	ARG

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	188	GLN
1	A	376	ASN

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

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.