wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 05:11 AM GMT

PDB ID	:	2PR9
Title	:	Mu2 adaptin subunit (AP50) of AP2 adaptor (second domain), complexed with GABAA receptor-gamma2 subunit-derived internalization peptide DEEY-GYECL
Authors	:	Vahedi-Faridi, A.; Haucke, V.; Kittler, J.T.; Kukhtina, V.; Moss, S.J.; Saenger, W.; Chen, G.-J.; Tretter, V.; Smith, K.; Yan, Z.; McAinsh, K.; Arancibia-Carcamo, L.
Deposited on	:	2007-05-04
Resolution	:	2.51 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

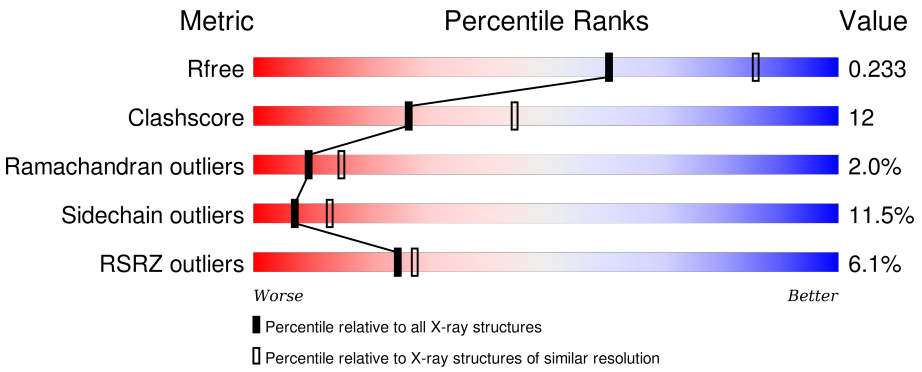
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.51 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}	91344	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

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.

Mol	Chain	Length	Quality of chain
1	A	299	4%, 60%, 19%, •, •, 15%
2	P	10	30%, 10%, 50%, 10%, 30%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2174 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-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	254	2029	1305	353	357	14	0	0	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	137	MET	-	EXPRESSION TAG	UNP P84092
A	138	GLY	-	EXPRESSION TAG	UNP P84092
A	139	SER	-	EXPRESSION TAG	UNP P84092
A	140	SER	-	EXPRESSION TAG	UNP P84092
A	141	HIS	-	EXPRESSION TAG	UNP P84092
A	142	HIS	-	EXPRESSION TAG	UNP P84092
A	143	HIS	-	EXPRESSION TAG	UNP P84092
A	144	HIS	-	EXPRESSION TAG	UNP P84092
A	145	HIS	-	EXPRESSION TAG	UNP P84092
A	146	HIS	-	EXPRESSION TAG	UNP P84092
A	147	SER	-	EXPRESSION TAG	UNP P84092
A	148	SER	-	EXPRESSION TAG	UNP P84092
A	149	GLY	-	EXPRESSION TAG	UNP P84092
A	150	LEU	-	EXPRESSION TAG	UNP P84092
A	151	VAL	-	EXPRESSION TAG	UNP P84092
A	152	PRO	-	EXPRESSION TAG	UNP P84092
A	153	ARG	-	EXPRESSION TAG	UNP P84092
A	154	GLY	-	EXPRESSION TAG	UNP P84092
A	155	SER	-	EXPRESSION TAG	UNP P84092
A	156	HIS	-	EXPRESSION TAG	UNP P84092
A	157	MET	-	EXPRESSION TAG	UNP P84092

- Molecule 2 is a protein called GABA(A) receptor subunit gamma-2 peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	P	10	86	52	10	23	1	0	0	0

- Molecule 3 is water.

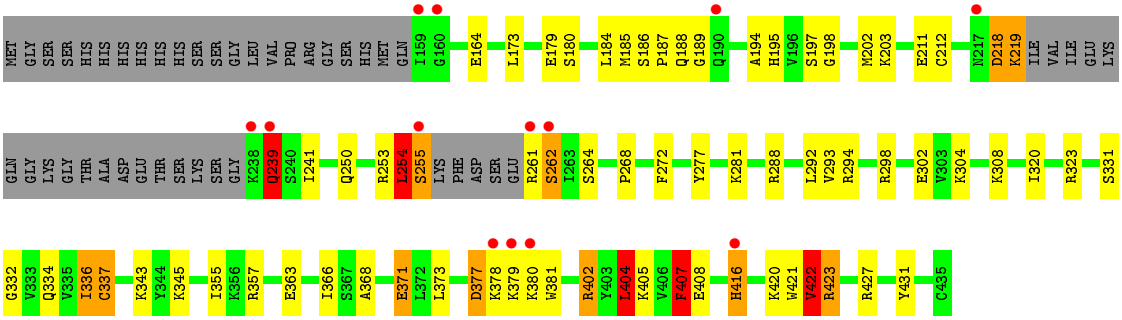
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	54	54	54	0	0
			Total	O		
3	P	5	5	5	0	0

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.

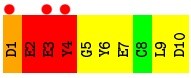
- Molecule 1: AP-2 complex subunit mu-1

Chain A:



- Molecule 2: GABA(A) receptor subunit gamma-2 peptide

Chain P:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 64	Depositor
Cell constants a, b, c, α, β, γ	126.30Å 126.30Å 74.65Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	9.97 – 2.51 9.97 – 2.51	Depositor EDS
% Data completeness (in resolution range)	100.0 (9.97-2.51) 99.7 (9.97-2.51)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.09 (at 2.50Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.203 , 0.240 0.195 , 0.233	Depositor DCC
R_{free} test set	912 reflections (4.16%)	DCC
Wilson B-factor (Å^2)	62.8	Xtriage
Anisotropy	0.030	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 48.9	EDS
Estimated twinning fraction	0.044 for h,-h-k,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 22817 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2174	wwPDB-VP
Average B, all atoms (Å^2)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.56% 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	1.75	27/2070 (1.3%)	1.29	17/2786 (0.6%)
2	P	1.92	1/87 (1.1%)	1.66	1/115 (0.9%)
All	All	1.75	28/2157 (1.3%)	1.30	18/2901 (0.6%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2
2	P	0	2
All	All	0	4

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	255	SER	C-O	31.73	1.83	1.23
1	A	219	LYS	C-O	21.06	1.63	1.23
1	A	261	ARG	NE-CZ	11.54	1.48	1.33
1	A	261	ARG	CZ-NH2	10.54	1.46	1.33
1	A	255	SER	CA-C	9.48	1.77	1.52

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	255	SER	CA-C-O	-12.71	93.40	120.10
1	A	261	ARG	NE-CZ-NH2	-11.91	114.35	120.30
1	A	427	ARG	NE-CZ-NH1	9.32	124.96	120.30
1	A	219	LYS	CA-C-O	-8.54	102.17	120.10
2	P	3	GLU	N-CA-C	8.27	133.32	111.00

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	218	ASP	Peptide
1	A	262	SER	Peptide
2	P	2	GLU	Peptide
2	P	3	GLU	Peptide

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	2029	0	2095	42	0
2	P	86	0	65	14	0
3	A	54	0	0	13	1
3	P	5	0	0	0	0
All	All	2174	0	2160	52	1

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 12.

The worst 5 of 52 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:255:SER:C	1:A:255:SER:CA	1.77	1.49
1:A:255:SER:C	1:A:255:SER:O	1.83	1.14
1:A:185:MET:HE2	3:A:479:HOH:O	1.68	0.91
2:P:5:GLY:H	2:P:6:TYR:HA	1.37	0.90
1:A:402:ARG:HD2	3:A:474:HOH:O	1.78	0.80

All (1) 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:463:HOH:O	3:A:463:HOH:O[4_565]	2.08	0.12

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	248/299 (83%)	237 (96%)	8 (3%)	3 (1%)	16 29
2	P	8/10 (80%)	4 (50%)	2 (25%)	2 (25%)	0 0
All	All	256/309 (83%)	241 (94%)	10 (4%)	5 (2%)	9 15

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	P	3	GLU
1	A	254	LEU
1	A	239	GLN
1	A	378	LYS
2	P	2	GLU

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	225/268 (84%)	202 (90%)	23 (10%)	9 17
2	P	9/9 (100%)	5 (56%)	4 (44%)	0 0
All	All	234/277 (84%)	207 (88%)	27 (12%)	7 13

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	293	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	377	ASP
2	P	3	GLU
1	A	298	ARG
1	A	203	LYS

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

Mol	Chain	Res	Type
1	A	190	GLN
1	A	250	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)

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	254/299 (84%)	-0.17	13 (5%)	32	36	36, 49, 67, 94	0
2	P	10/10 (100%)	1.19	3 (30%)	1	0	46, 56, 70, 72	0
All	All	264/309 (85%)	-0.12	16 (6%)	25	27	36, 50, 67, 94	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	238	LYS	4.9
1	A	239	GLN	4.0
1	A	416	HIS	3.4
1	A	378	LYS	3.3
1	A	255	SER	3.1

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.