WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 26, 2023 – 06:11 AM EDT

PDB ID	:	6BNT
Title	:	Crystal structure of AP2 mu1 adaptin C-terminal domain with IRS-1 peptide
Authors	:	Kikuchi, S.; Choi, E.; Yu, H.
Deposited on	:	2017-11-17
Resolution	:	3.20 Å(reported)

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

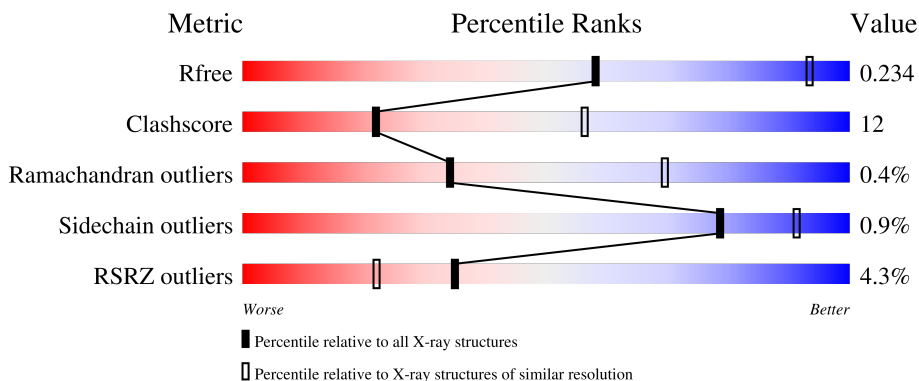|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 3.20 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		0.234
Clashscore		12
Ramachandran outliers		0.4%
Sidechain outliers		0.9%
RSRZ outliers		4.3%

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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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	315	3% — 58% — 20% • 21%
2	B	15	33% — 7% 7% — 53%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 2084 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	250	1993	1284	341	354	14	0	0	0

There are 39 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	121	MET	-	expression tag	UNP Q96CW1
A	122	GLY	-	expression tag	UNP Q96CW1
A	123	SER	-	expression tag	UNP Q96CW1
A	124	SER	-	expression tag	UNP Q96CW1
A	125	HIS	-	expression tag	UNP Q96CW1
A	126	HIS	-	expression tag	UNP Q96CW1
A	127	HIS	-	expression tag	UNP Q96CW1
A	128	HIS	-	expression tag	UNP Q96CW1
A	129	HIS	-	expression tag	UNP Q96CW1
A	130	HIS	-	expression tag	UNP Q96CW1
A	131	SER	-	expression tag	UNP Q96CW1
A	132	SER	-	expression tag	UNP Q96CW1
A	133	GLY	-	expression tag	UNP Q96CW1
A	134	LEU	-	expression tag	UNP Q96CW1
A	135	VAL	-	expression tag	UNP Q96CW1
A	136	PRO	-	expression tag	UNP Q96CW1
A	137	ARG	-	expression tag	UNP Q96CW1
A	138	GLY	-	expression tag	UNP Q96CW1
A	139	SER	-	expression tag	UNP Q96CW1
A	140	HIS	-	expression tag	UNP Q96CW1
A	141	MET	-	expression tag	UNP Q96CW1
A	142	ALA	-	expression tag	UNP Q96CW1
A	143	SER	-	expression tag	UNP Q96CW1
A	144	MET	-	expression tag	UNP Q96CW1
A	145	THR	-	expression tag	UNP Q96CW1
A	146	GLY	-	expression tag	UNP Q96CW1
A	147	GLY	-	expression tag	UNP Q96CW1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	148	GLN	-	expression tag	UNP Q96CW1
A	149	GLN	-	expression tag	UNP Q96CW1
A	150	MET	-	expression tag	UNP Q96CW1
A	151	GLY	-	expression tag	UNP Q96CW1
A	152	ARG	-	expression tag	UNP Q96CW1
A	153	GLY	-	expression tag	UNP Q96CW1
A	154	SER	-	expression tag	UNP Q96CW1
A	155	GLU	-	expression tag	UNP Q96CW1
A	156	PHE	-	expression tag	UNP Q96CW1
A	157	GLY	-	expression tag	UNP Q96CW1
A	158	ARG	-	expression tag	UNP Q96CW1
A	159	PRO	-	expression tag	UNP Q96CW1

- Molecule 2 is a protein called Insulin receptor substrate 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
2	B	7	57	33	7	14	1	2	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	606	CYS	-	expression tag	UNP P35568

- Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	34	34	34	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: AP-2 complex subunit mu

Chain A:
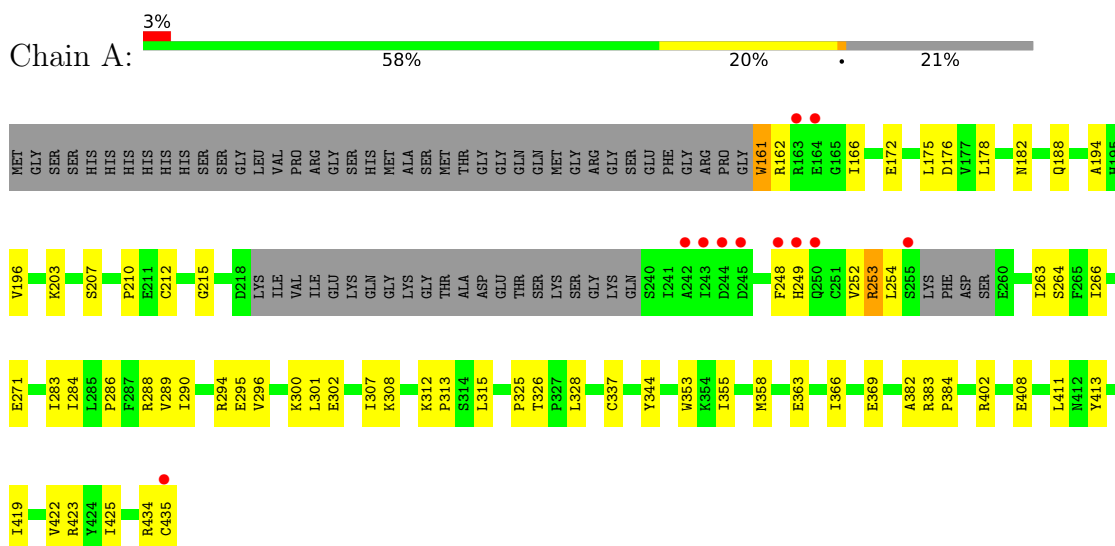
• Molecule 2: Insulin receptor substrate 1

Chain B:
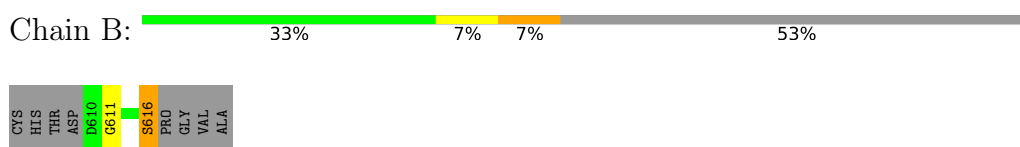
4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	P 64	Depositor
Cell constants a, b, c, α, β, γ	125.33Å 125.33Å 74.82Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	27.64 – 3.20 27.64 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.8 (27.64-3.20) 99.8 (27.64-3.20)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.39 (at 3.17Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.206 , 0.234 0.206 , 0.234	Depositor DCC
R_{free} test set	1113 reflections (9.97%)	wwPDB-VP
Wilson B-factor (Å^2)	148.4	Xtriage
Anisotropy	0.003	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 99.9	EDS
L-test for twinning[2]	$<|L|> = 0.49, <L^2> = 0.32$	Xtriage
Estimated twinning fraction	0.043 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	2084	wwPDB-VP
Average B, all atoms (Å^2)	145.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.64% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5   Model quality ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: SEP

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	0.55	0/2034	0.69	0/2740
2	B	0.77	0/48	1.25	0/63
All	All	0.56	0/2082	0.71	0/2803

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	1993	0	2046	48	0
2	B	57	0	43	1	0
3	A	34	0	0	8	0
All	All	2084	0	2089	49	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 12.

The worst 5 of 49 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:161:TRP:HD1	1:A:162:ARG:H	1.10	0.98
1:A:308:LYS:NZ	1:A:363:GLU:OE2	2.14	0.80
1:A:358:MET:SD	3:A:515:HOH:O	2.42	0.77
1:A:369:GLU:OE1	3:A:501:HOH:O	2.07	0.73
1:A:296:VAL:HB	1:A:300:LYS:HB2	1.78	0.66

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	244/315 (78%)	220 (90%)	24 (10%)	0	100	100
2	B	5/15 (33%)	4 (80%)	0	1 (20%)	0	0
All	All	249/330 (76%)	224 (90%)	24 (10%)	1 (0%)	34	69

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	611	GLY

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	221/278 (80%)	219 (99%)	2 (1%)	78	91

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	5/11 (46%)	5 (100%)	0	100	100
All	All	226/289 (78%)	224 (99%)	2 (1%)	78	91

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	161	TRP
1	A	253	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
2	SEP	B	616	2	8,9,10	1.97	3 (37%)	8,12,14	1.30	1 (12%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SEP	B	616	2	-	2/5/8/10	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	616	SEP	P-O1P	4.15	1.63	1.50
2	B	616	SEP	P-O2P	2.40	1.64	1.54
2	B	616	SEP	P-O3P	2.28	1.63	1.54

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(º)	Ideal(º)
2	B	616	SEP	OG-CB-CA	2.41	110.49	108.14

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	616	SEP	CA-CB-OG-P
2	B	616	SEP	N-CA-CB-OG

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	616	SEP	1	0

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	250/315 (79%)	0.02	11 (4%)	34	21	110, 138, 193, 217	0
2	B	6/15 (40%)	-0.59	0	100	100	132, 136, 138, 142	0
All	All	256/330 (77%)	0.01	11 (4%)	35	22	110, 138, 193, 217	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	244	ASP	5.5
1	A	164	GLU	3.4
1	A	248	PHE	3.4
1	A	242	ALA	3.2
1	A	255	SER	3.1

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	SEP	B	616	10/11	0.92	0.10	162,165,179,192	0

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.