wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 06:34 PM GMT

PDB ID	:	1BHY
Title	:	LOW TEMPERATURE MIDDLE RESOLUTION STRUCTURE OF P64K FROM MASC DATA
Authors	:	Ramin, M.; Shepard, W.; Fourme, R.; Kahn, R.
Deposited on	:	1998-06-10
Resolution	:	4.18 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 (i)) were used in the production of this report:

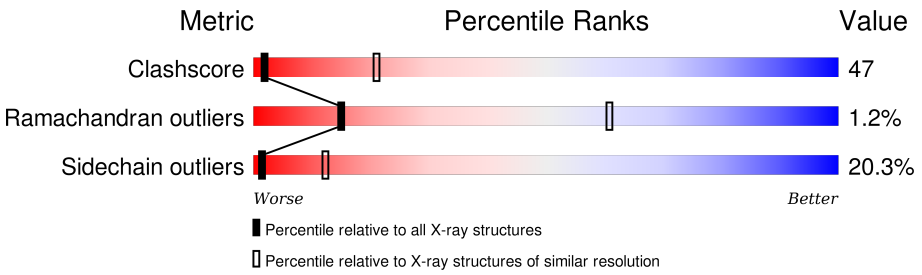
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
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 4.18 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))
Clashscore	102246	1130 (4.76-3.60)
Ramachandran outliers	100387	1076 (4.76-3.60)
Sidechain outliers	100360	1061 (4.76-3.60)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	482	38% 44% 14% •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	FAD	A	600	X	-	-	-

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3616 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.

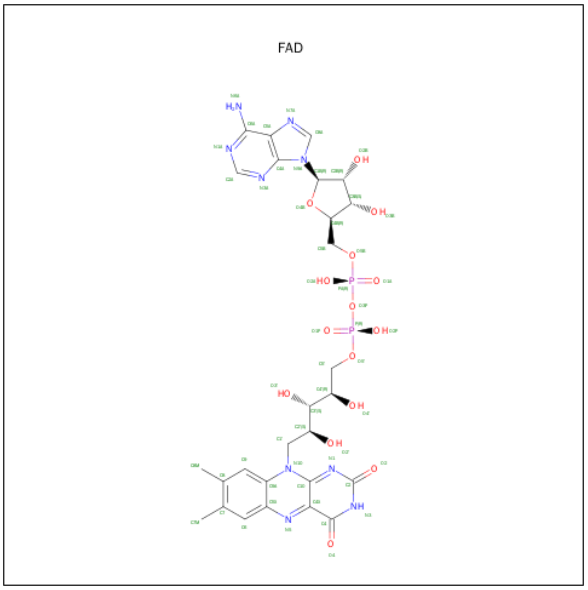
- Molecule 1 is a protein called P64K.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	482	3563	2249	616	679	19	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	498	ALA	GLY	CONFLICT	UNP Q51225
A	553	TYR	CYS	CONFLICT	UNP Q51225

- Molecule 2 is FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: FAD) (formula: $C_{27}H_{33}N_9O_{15}P_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	A	1	53	27	9	15	2	0	0

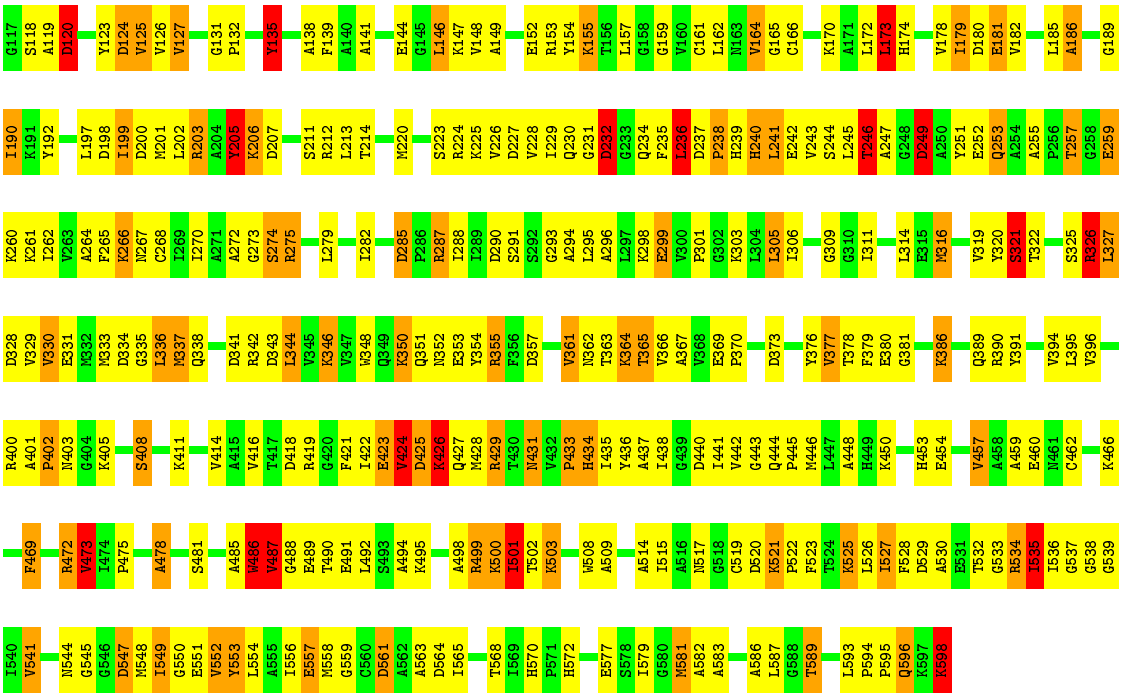
3 Residue-property plots (i)

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.

Note EDS was not executed.

- Molecule 1: P64K

Chain A: 38% 44% 14% •



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	140.62Å 140.62Å 77.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 4.18	Depositor
% Data completeness (in resolution range)	(Not available) (8.00-4.18)	Depositor
R_{merge}	0.06	Depositor
R_{sym}	0.06	Depositor
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.338 , 0.323	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3616	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

5 Model quality (i)

5.1 Standard geometry (i)

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

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	1.35	5/3629 (0.1%)	2.06	102/4917 (2.1%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	598	LYS	C-OXT	7.30	1.37	1.23
1	A	321	SER	CB-OG	6.91	1.51	1.42
1	A	274	SER	CB-OG	6.47	1.50	1.42
1	A	402	PRO	C-N	6.02	1.48	1.34
1	A	598	LYS	C-O	5.45	1.33	1.23

The worst 5 of 102 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	553	TYR	CB-CG-CD1	-19.69	109.19	121.00
1	A	355	ARG	NE-CZ-NH2	-19.16	110.72	120.30
1	A	275	ARG	NE-CZ-NH2	-15.56	112.52	120.30
1	A	534	ARG	NE-CZ-NH2	-14.54	113.03	120.30
1	A	553	TYR	CB-CG-CD2	13.92	129.35	121.00

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	3563	0	3580	340	0
2	A	53	0	31	9	0
All	All	3616	0	3611	340	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 47.

The worst 5 of 340 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:475:PRO:CA	1:A:553:TYR:OH	1.83	1.25
1:A:426:LYS:H	1:A:426:LYS:HE2	1.02	1.13
1:A:275:ARG:NH1	1:A:403:ASN:HD22	1.46	1.13
1:A:386:LYS:H	1:A:386:LYS:HD2	1.10	1.10
1:A:521:LYS:N	1:A:522:PRO:HD3	1.59	1.10

There are no symmetry-related clashes.

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	480/482 (100%)	445 (93%)	29 (6%)	6 (1%)	15 61

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	118	SER
1	A	185	LEU
1	A	238	PRO
1	A	426	LYS
1	A	285	ASP

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	364/364 (100%)	290 (80%)	74 (20%)	1 12

5 of 74 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	330	VAL
1	A	361	VAL
1	A	535	ILE
1	A	334	ASP
1	A	342	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	188	ASN
1	A	230	GLN
1	A	403	ASN
1	A	431	ASN
1	A	517	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 carbohydrates in this entry.

5.6 Ligand geometry (i)

1 ligand 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	FAD	A	600	1	48,58,58	2.44	20 (41%)	54,89,89	3.06	21 (38%)

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	FAD	A	600	1	3/3/9/9	0/30/50/50	0/6/6/6

The worst 5 of 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	600	FAD	C1'-N10	-9.77	1.38	1.48
2	A	600	FAD	O4B-C4B	-3.65	1.36	1.45
2	A	600	FAD	PA-O2A	-3.28	1.41	1.54
2	A	600	FAD	C5'-C4'	-2.99	1.47	1.51
2	A	600	FAD	C9A-C5X	-2.76	1.37	1.42

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	600	FAD	C4-C4X-C10	-8.45	114.53	119.94
2	A	600	FAD	O5'-P-O1P	-5.40	88.65	109.62
2	A	600	FAD	C2B-C1B-N9A	-5.26	106.26	114.29
2	A	600	FAD	C1B-N9A-C4A	-4.97	119.44	126.94
2	A	600	FAD	O2P-P-O1P	-4.34	88.99	112.53

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	600	FAD	C4'
2	A	600	FAD	C2'
2	A	600	FAD	C3'

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	600	FAD	9	0

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)

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates (i)

EDS was not executed - this section will therefore be empty.

6.4 Ligands (i)

EDS was not executed - this section will therefore be empty.

6.5 Other polymers (i)

EDS was not executed - this section will therefore be empty.