WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Aug 13, 2017 – 09:36 AM EDT

PDB ID	:	4BR4
Title	:	Legionella pneumophila NTPDase1 crystal form I, open, apo
Authors	:	Zebisch, M.; Schaefer, P.; Lauble, P.; Straeter, N.
Deposited on	:	unknown
Resolution	:	1.45 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.

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 1.45 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.
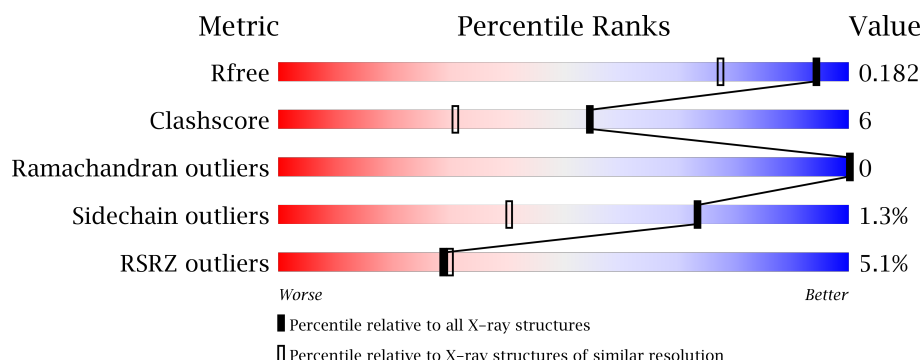
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	1510 (1.48-1.44)
Clashscore	112137	1573 (1.48-1.44)
Ramachandran outliers	110173	1555 (1.48-1.44)
Sidechain outliers	110143	1555 (1.48-1.44)
RSRZ outliers	101464	1516 (1.48-1.44)

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	365	5% / 84% / 13% / ••

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3263 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 ECTONUCLEOSIDE TRIPHOSPHATE DIPHOSPHOHY-DROLASE I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	356	2875	1834	472	554	15	0	10	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	29	MET	-	expression tag	UNP Q5ZUA2
A	30	ASP	-	expression tag	UNP Q5ZUA2
A	31	HIS	-	expression tag	UNP Q5ZUA2
A	32	HIS	-	expression tag	UNP Q5ZUA2
A	33	HIS	-	expression tag	UNP Q5ZUA2
A	34	HIS	-	expression tag	UNP Q5ZUA2
A	35	HIS	-	expression tag	UNP Q5ZUA2
A	36	HIS	-	expression tag	UNP Q5ZUA2
A	137	ASP	GLU	conflict	UNP Q5ZUA2
A	149	VAL	ALA	conflict	UNP Q5ZUA2

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
2	A	2	2	2	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
3	A	3	3	3	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	382	Total 383	O 383	0	1

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 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: ECTONUCLEOSIDE TRIPHOSPHATE DIPHOSPHOHYDROLASE I
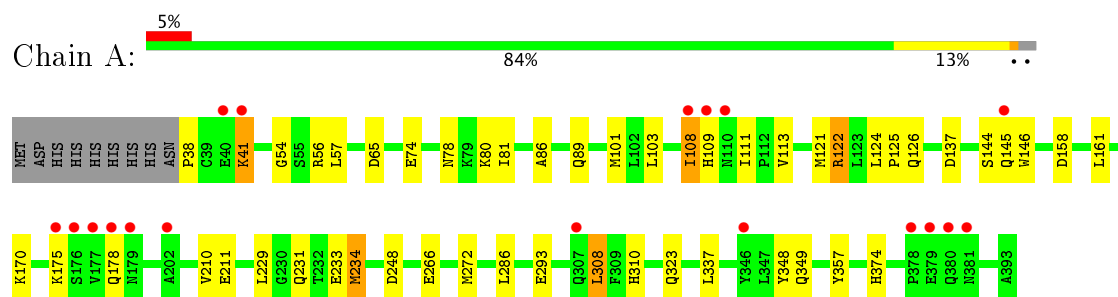
GLOBAL-STATISTICS INFOmissingINFO

4 Model quality (i)

4.1 Standard geometry (i)

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

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	1.08	4/2982 (0.1%)	1.04	8/4065 (0.2%)

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	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	293	GLU	CD-OE2	-6.02	1.19	1.25
1	A	234	MET	CG-SD	5.21	1.94	1.81
1	A	231	GLN	CD-NE2	-5.08	1.20	1.32
1	A	158	ASP	CG-OD2	5.01	1.36	1.25

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	122	ARG	NE-CZ-NH2	-8.12	116.24	120.30
1	A	248	ASP	CB-CG-OD1	7.82	125.33	118.30
1	A	65	ASP	CB-CG-OD1	5.98	123.68	118.30
1	A	211	GLU	OE1-CD-OE2	-5.78	116.37	123.30
1	A	108	ILE	O-C-N	5.45	131.42	122.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	122	ARG	Sidechain

4.2 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	2875	0	2747	34	0
2	A	2	0	0	0	0
3	A	3	0	0	0	0
4	A	383	0	0	13	0
All	All	3263	0	2747	34	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 6.

The worst 5 of 34 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:266:GLU:HG2	4:A:2277:HOH:O	1.74	0.86
1:A:144:SER:O	4:A:2102:HOH:O	1.98	0.80
1:A:56[B]:ARG:HG3	1:A:80:LYS:HD3	1.68	0.76
1:A:233:GLU:OE1	4:A:2234:HOH:O	2.06	0.71
1:A:108:ILE:HD11	1:A:111:ILE:HD13	1.73	0.71

There are no symmetry-related clashes.

4.3 Torsion angles (i)

4.3.1 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	364/365 (100%)	358 (98%)	6 (2%)	0	100　100

There are no Ramachandran outliers to report.

4.3.2 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	321/325 (99%)	317 (99%)	4 (1%)	75　44

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	41	LYS
1	A	78	ASN
1	A	234	MET
1	A	308	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	126	GLN
1	A	318	ASN
1	A	323	GLN
1	A	374	HIS

4.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

4.5 Carbohydrates (i)

There are no carbohydrates in this entry.

4.6 Ligand geometry (i)

Of 5 ligands modelled in this entry, 5 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Fit of model and data (i)

5.1 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	356/365 (97%)	-0.10	18 (5%)	29	30	11, 18, 48, 80	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	177	VAL	7.6
1	A	378	PRO	5.9
1	A	109	HIS	3.4
1	A	176	SER	3.3
1	A	179	ASN	3.2

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

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

5.3 Carbohydrates (i)

There are no carbohydrates in this entry.

5.4 Ligands (i)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
2	CL	A	1394	1/1	0.99	0.04	-4.03	20,20,20,20	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
2	CL	A	1395	1/1	0.99	0.05	-	32,32,32,32	0
3	MG	A	1398	1/1	0.98	0.08	-	32,32,32,32	1
3	MG	A	1397	1/1	0.95	0.13	-	41,41,41,41	1
3	MG	A	1396	1/1	1.00	0.13	-	19,19,19,19	1

5.5 Other polymers (i)

There are no such residues in this entry.