wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 24, 2021 – 10:11 pm GMT

PDB ID	:	7P0L
Title	:	Crystal structure of S.pombe Mdb1 BRCT domains in complex with a H2A phosphopeptide
Authors	:	Day, M.; Oliver, A.W.; Pearl, L.H.
Deposited on	:	2021-06-29
Resolution	:	1.97 Å(reported)

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

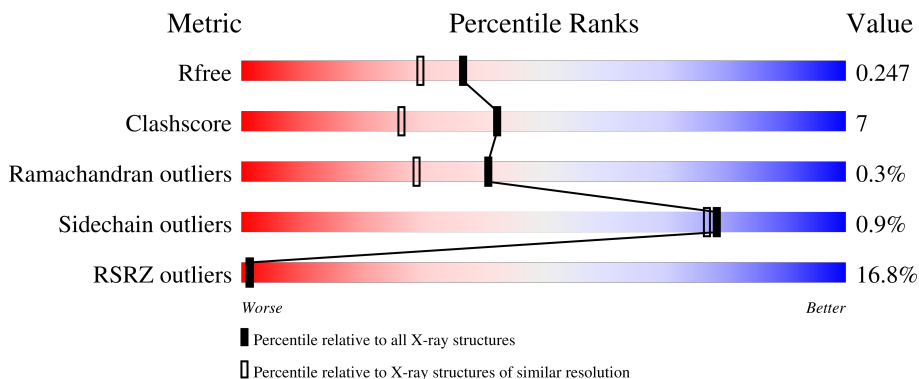|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4 (270009), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 1.97 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		0.247
Clashscore		7
Ramachandran outliers		0.3%
Sidechain outliers		0.9%
RSRZ outliers		16.8%

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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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	199	7% / 89% / 9% •
1	B	199	25% / 82% / 16% •
2	C	11	9% / 45% / 9% / 45%
2	D	11	36% / 27% / 18% / 55%

2   Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 6364 atoms, of which 3080 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 DNA damage response protein Mdb1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	195	Total 3122	C 980	H 1568	N 272	O 297	S 5	0	0	0
1	B	195	Total 2903	C 939	H 1429	N 245	O 285	S 5	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	383	MET	-	initiating methionine	UNP O14079
B	383	MET	-	initiating methionine	UNP O14079

- Molecule 2 is a protein called Histone H2A-beta.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	6	Total 101	C 30	H 48	N 8	O 14	P 1	0	0	0
2	D	5	Total 79	C 24	H 35	N 6	O 13	P 1	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	122	SER	-	insertion	UNP P04910
C	123	GLY	-	insertion	UNP P04910
C	124	ARG	-	insertion	UNP P04910
C	125	THR	-	insertion	UNP P04910
D	122	SER	-	insertion	UNP P04910
D	123	GLY	-	insertion	UNP P04910
D	124	ARG	-	insertion	UNP P04910
D	125	THR	-	insertion	UNP P04910

- Molecule 3 is water.

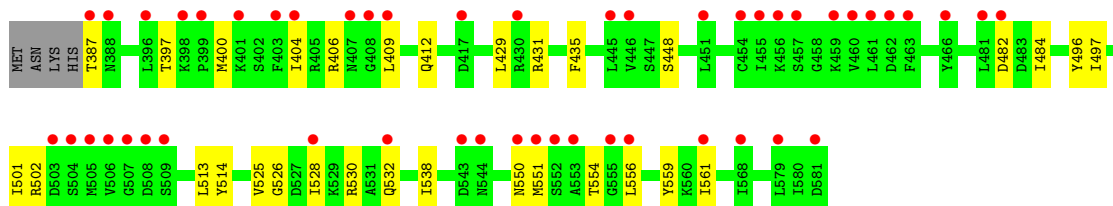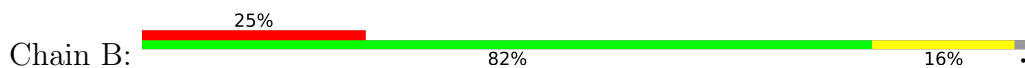| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	92	Total 92	O 92	0	0
3	B	60	Total 60	O 60	0	0
3	C	2	Total 2	O 2	0	0
3	D	5	Total 5	O 5	0	0

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.
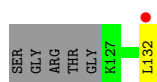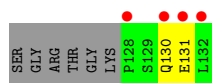
• Molecule 1: DNA damage response protein Mdb1
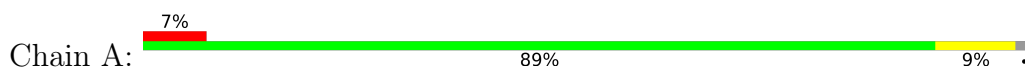
Chain A:

• Molecule 1: DNA damage response protein Mdb1

Chain B:

• Molecule 2: Histone H2A-beta

Chain C:

• Molecule 2: Histone H2A-beta

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	41.38Å 42.05Å 96.29Å 90.00° 91.87° 90.00°	Depositor
Resolution (Å)	48.12 – 1.97 48.12 – 1.97	Depositor EDS
% Data completeness (in resolution range)	98.6 (48.12-1.97) 98.6 (48.12-1.97)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.26 (at 1.97Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.208 , 0.247 0.208 , 0.247	Depositor DCC
R_{free} test set	1158 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	35.8	Xtriage
Anisotropy	0.253	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< |L| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.032 for -k,-h,-l 0.029 for k,h,-l 0.039 for h,-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6364	wwPDB-VP
Average B, all atoms (Å^2)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.24% 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 (i)

5.1 Standard geometry (i)

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	0.32	0/1578	0.54	0/2124
1	B	0.30	0/1498	0.53	0/2029
2	C	0.33	0/42	0.43	0/52
2	D	0.25	0/33	0.47	0/40
All	All	0.31	0/3151	0.53	0/4245

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 (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	1554	1568	1568	16	0
1	B	1474	1429	1429	27	0
2	C	53	48	48	1	0
2	D	44	35	35	2	0
3	A	92	0	0	1	0
3	B	60	0	0	5	0
3	C	2	0	0	0	0
3	D	5	0	0	1	0
All	All	3284	3080	3080	43	0

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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 7.

The worst 5 of 43 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:B:497:ILE:O	1:B:502:ARG:NH1	1.90	1.03
1:B:513:LEU:HD21	1:B:561:ILE:HD11	1.41	0.99
1:A:462:ASP:OD2	1:B:559:TYR:OH	1.98	0.80
1:A:502:ARG:HA	1:A:510:ILE:HD12	1.66	0.76
2:D:130:GLN:OE1	3:D:201:HOH:O	2.05	0.75

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	193/199 (97%)	190 (98%)	3 (2%)	0	100	100
1	B	193/199 (97%)	186 (96%)	6 (3%)	1 (0%)	29	16
2	C	3/11 (27%)	2 (67%)	1 (33%)	0	100	100
2	D	2/11 (18%)	2 (100%)	0	0	100	100
All	All	391/420 (93%)	380 (97%)	10 (3%)	1 (0%)	41	29

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	406	ARG

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	172/178 (97%)	171 (99%)	1 (1%)	86	85
1	B	154/178 (86%)	152 (99%)	2 (1%)	69	64
2	C	5/8 (62%)	5 (100%)	0	100	100
2	D	4/8 (50%)	4 (100%)	0	100	100
All	All	335/372 (90%)	332 (99%)	3 (1%)	78	77

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	543	ASP
1	B	448	SER
1	B	482	ASP

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 ⓘ

2 non-standard protein/DNA/RNA residues are modelled in this entry.

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.

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(Å²)	Q<0.9
1	A	195/199 (97%)	0.84	13 (6%)	17	19	24, 37, 66, 96	0
1	B	195/199 (97%)	1.59	49 (25%)	0	0	26, 53, 87, 98	0
2	C	5/11 (45%)	2.75	1 (20%)	1	0	63, 65, 72, 89	0
2	D	4/11 (36%)	4.31	4 (100%)	0	0	61, 66, 74, 75	0
All	All	399/420 (95%)	1.26	67 (16%)	1	1	24, 45, 85, 98	0

The worst 5 of 67 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	132	LEU	10.2
1	B	460	VAL	10.0
2	D	132	LEU	9.1
1	B	506	VAL	9.0
1	A	506	VAL	7.7

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(Å²)	Q<0.9
2	SEP	C	129	10/11	0.94	0.13	36,50,63,72	0
2	SEP	D	129	10/11	0.94	0.16	41,49,66,66	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.