Full wwPDB X-ray Structure Validation 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)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

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

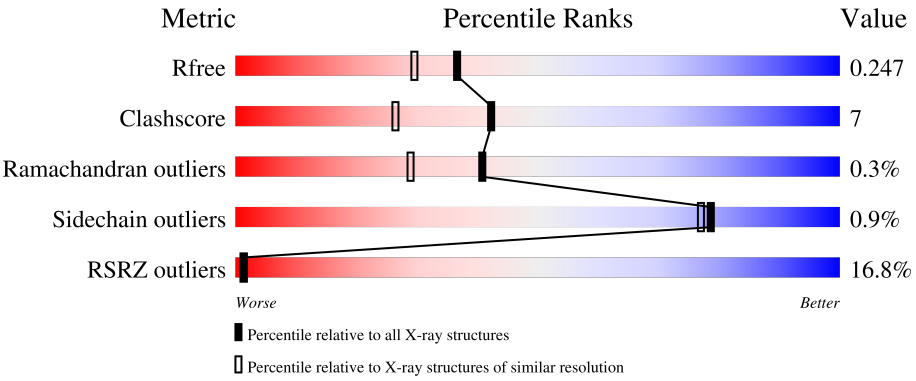
MolProbity	:	4.02b-467
Mogul	:	1.8.4 (270009), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry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

1 Overall quality at a glance (i)

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	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 (i)

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	C	H	N	O	S	0	0	0
			3122	980	1568	272	297	5			
1	B	195	Total	C	H	N	O	S	0	0	0
			2903	939	1429	245	285	5			

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	C	H	N	O	P	0	0	0
			101	30	48	8	14	1			
2	D	5	Total	C	H	N	O	P	0	0	0
			79	24	35	6	13	1			

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

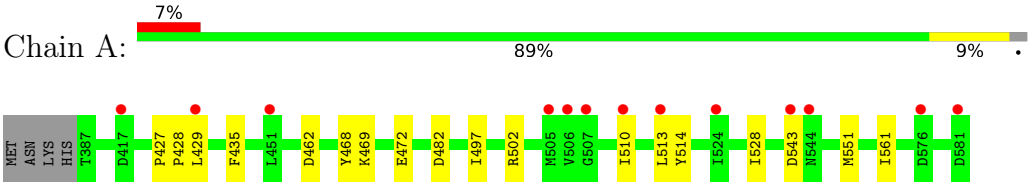
- Molecule 3 is water.

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

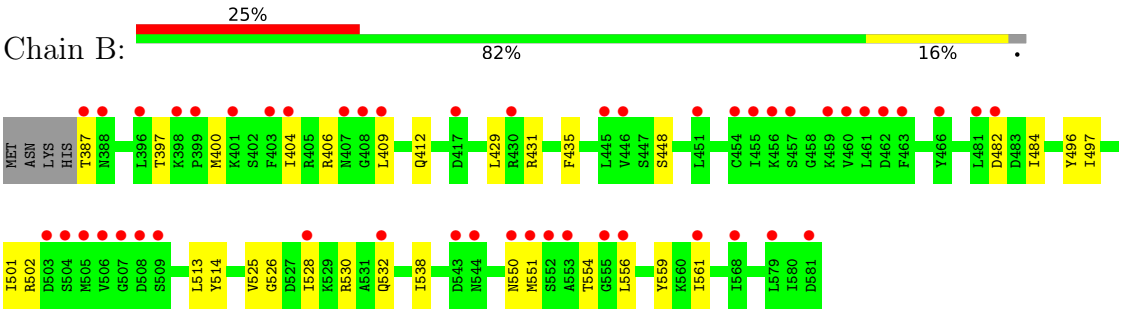
3 Residue-property plots (i)

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.

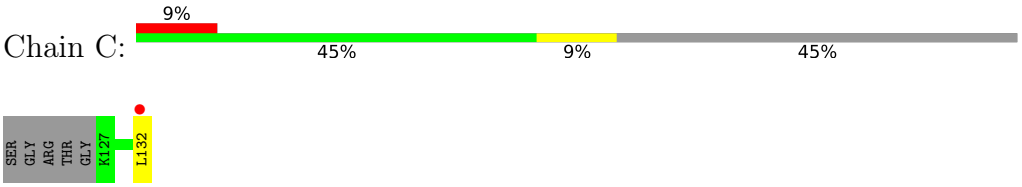
• Molecule 1: DNA damage response protein Mdb1



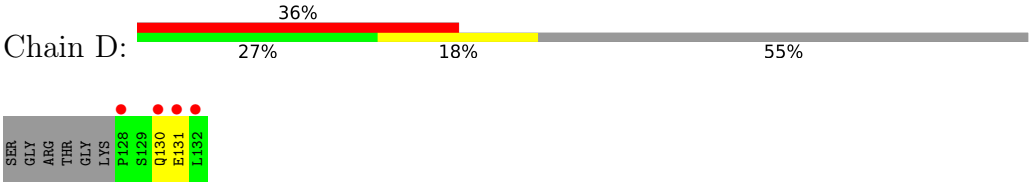
• Molecule 1: DNA damage response protein Mdb1



• Molecule 2: Histone H2A-beta



• Molecule 2: Histone H2A-beta



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

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	0.32	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

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

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.

All (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
1:B:412:GLN:OE1	3:B:601:HOH:O	2.14	0.66
1:B:501:ILE:HD13	1:B:561:ILE:HD13	1.80	0.64
1:B:431:ARG:NH1	3:B:603:HOH:O	2.33	0.62
1:A:513:LEU:CD2	1:A:561:ILE:HD11	2.32	0.59
1:B:554:THR:HG21	3:B:609:HOH:O	2.02	0.58
1:B:397:THR:H	1:B:400:MET:HE2	1.71	0.54
1:B:387:THR:HB	1:B:409:LEU:CD2	2.38	0.54
1:A:510:ILE:HD11	1:A:514:TYR:HE2	1.73	0.53
1:A:497:ILE:HD12	1:A:514:TYR:CE1	2.42	0.53
1:B:550:ASN:O	1:B:554:THR:OG1	2.24	0.53
1:A:469:LYS:HA	1:A:469:LYS:HE2	1.90	0.52
1:B:497:ILE:HD12	1:B:514:TYR:CE1	2.47	0.50
1:A:513:LEU:HD21	1:A:561:ILE:HD11	1.93	0.49
1:A:513:LEU:HD13	2:C:132:LEU:HD21	1.95	0.49
1:A:510:ILE:HD11	1:A:514:TYR:CE2	2.46	0.49
1:B:551:MET:HE3	1:B:556:LEU:HD22	1.94	0.49
1:B:496:TYR:HB2	1:B:525:VAL:CG2	2.43	0.48
1:A:497:ILE:HD13	1:A:513:LEU:HD23	1.96	0.48
1:A:528:ILE:HD11	1:A:551:MET:SD	2.54	0.48
1:B:528:ILE:HD11	1:B:551:MET:HE2	1.96	0.48
1:A:482:ASP:OD2	3:A:601:HOH:O	2.20	0.47
1:B:526:GLY:H	1:B:530:ARG:HH21	1.63	0.46
1:B:397:THR:HG23	1:B:400:MET:CE	2.47	0.44
1:B:397:THR:HG23	1:B:400:MET:HE2	1.99	0.44
1:B:554:THR:HG22	1:B:554:THR:O	2.17	0.44
1:B:429:LEU:HD21	1:B:435:PHE:CZ	2.53	0.44
1:B:532:GLN:NE2	3:B:609:HOH:O	2.51	0.44
1:B:551:MET:CE	1:B:556:LEU:HD22	2.49	0.43
1:B:513:LEU:HD21	1:B:561:ILE:CD1	2.31	0.43
1:B:484:ILE:HD11	3:B:632:HOH:O	2.19	0.43
1:B:538:ILE:HD13	1:B:551:MET:HE3	2.00	0.43
1:A:472:GLU:OE1	1:A:472:GLU:N	2.48	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:429:LEU:HD21	1:A:435:PHE:CZ	2.53	0.42
1:A:427:PRO:HA	1:A:428:PRO:C	2.40	0.42
1:B:431:ARG:H	2:D:131:GLU:CG	2.32	0.42
1:A:468:TYR:O	1:A:469:LYS:HE2	2.20	0.41
1:B:404:ILE:HG23	1:B:409:LEU:HB2	2.02	0.41
1:B:556:LEU:HD12	1:B:556:LEU:N	2.36	0.41

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 (i)

There are no RNA molecules in this entry.

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

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 (i)

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, 95^{th} 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	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

All (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
1	B	555	GLY	6.7
1	B	456	LYS	6.3
1	B	581	ASP	6.0
1	B	553	ALA	5.8
1	B	404	ILE	4.9
1	B	403	PHE	4.7
1	B	482	ASP	4.6
1	B	508	ASP	4.6
1	B	551	MET	4.5
1	A	581	ASP	4.5
1	B	457	SER	4.4
1	B	552	SER	4.4
1	B	396	LEU	4.2
1	B	407	ASN	4.2
1	B	461	LEU	4.2
1	B	417	ASP	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	398	LYS	3.9
1	B	550	ASN	3.8
1	B	459	LYS	3.7
2	D	131	GLU	3.6
1	B	466	TYR	3.5
1	A	451	LEU	3.5
1	B	463	PHE	3.5
1	A	505	MET	3.4
1	B	556	LEU	3.3
1	B	399	PRO	3.0
1	B	503	ASP	2.9
1	B	451	LEU	2.9
1	B	454	CYS	2.8
1	B	528	ILE	2.8
1	A	510	ILE	2.8
1	B	544	ASN	2.8
1	B	509	SER	2.7
1	A	544	ASN	2.7
1	B	387	THR	2.7
1	A	513	LEU	2.7
1	B	507	GLY	2.6
1	A	543	ASP	2.6
1	B	455	ILE	2.5
1	B	409	LEU	2.5
1	B	505	MET	2.5
2	D	128	PRO	2.5
1	B	430	ARG	2.5
1	A	429	LEU	2.4
1	B	568	ILE	2.4
1	A	507	GLY	2.4
1	B	481	LEU	2.3
1	B	561	ILE	2.2
1	B	579	LEU	2.2
1	B	388	ASN	2.1
1	B	446	VAL	2.1
1	B	543	ASP	2.1
2	D	130	GLN	2.1
1	B	408	GLY	2.1
1	B	401	LYS	2.1
1	A	417	ASP	2.1
1	B	462	ASP	2.1
1	B	504	SER	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	576	ASP	2.0
1	B	532	GLN	2.0
1	B	445	LEU	2.0
1	A	524	ILE	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (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. The B-factors column lists the minimum, median, 95^{th} 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	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 (i)

There are no monosaccharides 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.