



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 6, 2026 – 03:11 AM UTC

PDB ID : 3SHT / pdb_00003sht
Title : Crystal structure of human MCPH1 tandem BRCT domains
Authors : Shao, Z.H.; Li, F.D.; Yan, W.
Deposited on : 2011-06-17
Resolution : 1.95 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

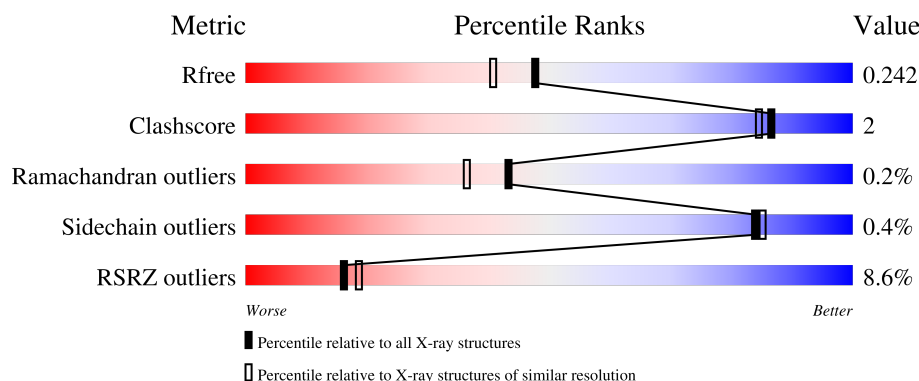
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.95 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}	180053	3494 (1.96-1.96)
Clashscore	190562	3612 (1.96-1.96)
Ramachandran outliers	187476	3587 (1.96-1.96)
Sidechain outliers	187428	3587 (1.96-1.96)
RSRZ outliers	180081	3495 (1.96-1.96)

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 $\leq 5\%$. 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	206	3% 87% 9%
1	B	206	8% 87% 9%
1	C	206	12% 84% 6% 10%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4796 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 Microcephalin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	188	Total	C	N	O	S	0	0	0
			1457	939	248	260	10			
1	B	187	Total	C	N	O	S	0	0	0
			1450	934	247	259	10			
1	C	186	Total	C	N	O	S	0	0	0
			1442	928	246	258	10			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	630	MET	-	expression tag	UNP Q8NEM0
A	631	GLY	-	expression tag	UNP Q8NEM0
A	632	HIS	-	expression tag	UNP Q8NEM0
A	633	HIS	-	expression tag	UNP Q8NEM0
A	634	HIS	-	expression tag	UNP Q8NEM0
A	635	HIS	-	expression tag	UNP Q8NEM0
A	636	HIS	-	expression tag	UNP Q8NEM0
A	637	HIS	-	expression tag	UNP Q8NEM0
A	638	MET	-	expression tag	UNP Q8NEM0
A	761	ALA	VAL	SEE REMARK 999	UNP Q8NEM0
B	630	MET	-	expression tag	UNP Q8NEM0
B	631	GLY	-	expression tag	UNP Q8NEM0
B	632	HIS	-	expression tag	UNP Q8NEM0
B	633	HIS	-	expression tag	UNP Q8NEM0
B	634	HIS	-	expression tag	UNP Q8NEM0
B	635	HIS	-	expression tag	UNP Q8NEM0
B	636	HIS	-	expression tag	UNP Q8NEM0
B	637	HIS	-	expression tag	UNP Q8NEM0
B	638	MET	-	expression tag	UNP Q8NEM0
B	761	ALA	VAL	SEE REMARK 999	UNP Q8NEM0
C	630	MET	-	expression tag	UNP Q8NEM0
C	631	GLY	-	expression tag	UNP Q8NEM0
C	632	HIS	-	expression tag	UNP Q8NEM0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	633	HIS	-	expression tag	UNP Q8NEM0
C	634	HIS	-	expression tag	UNP Q8NEM0
C	635	HIS	-	expression tag	UNP Q8NEM0
C	636	HIS	-	expression tag	UNP Q8NEM0
C	637	HIS	-	expression tag	UNP Q8NEM0
C	638	MET	-	expression tag	UNP Q8NEM0
C	761	ALA	VAL	SEE REMARK 999	UNP Q8NEM0

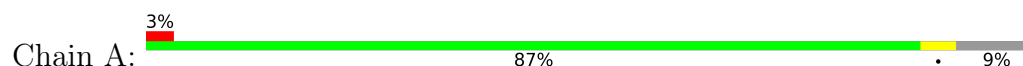
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	156	Total 156	O 156	0	0
2	B	158	Total 158	O 158	0	0
2	C	133	Total 133	O 133	0	0

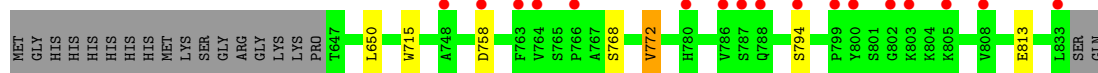
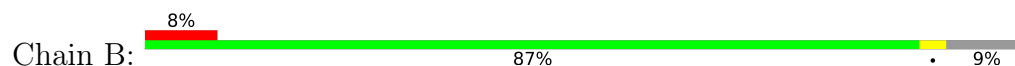
3 Residue-property plots [i](#)

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.

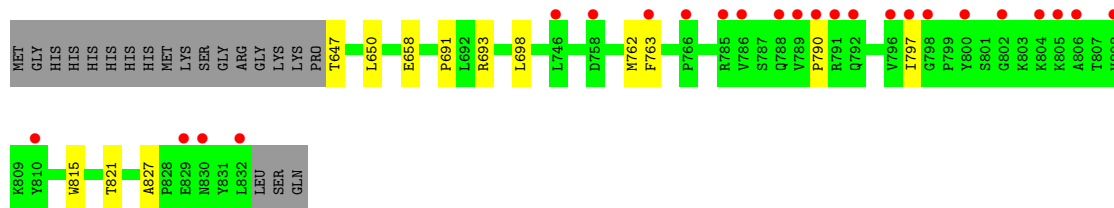
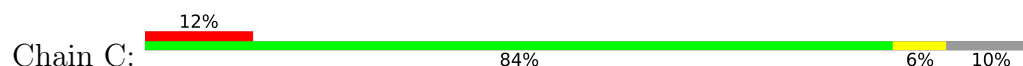
- Molecule 1: Microcephalin



- Molecule 1: Microcephalin



- Molecule 1: Microcephalin



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	46.47Å 113.67Å 74.58Å 90.00° 91.04° 90.00°	Depositor
Resolution (Å)	32.58 – 1.95 32.58 – 1.95	Depositor EDS
% Data completeness (in resolution range)	97.2 (32.58-1.95) 97.2 (32.58-1.95)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.40 (at 1.95Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.202 , 0.242 0.202 , 0.242	Depositor DCC
R_{free} test set	2767 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å ²)	24.6	Xtriage
Anisotropy	0.018	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 41.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.032 for h,-k,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	4796	wwPDB-VP
Average B, all atoms (Å ²)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.32% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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](#)

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	0.84	0/1498	0.89	0/2043
1	B	0.87	0/1490	0.93	2/2032 (0.1%)
1	C	0.77	0/1482	0.91	2/2021 (0.1%)
All	All	0.83	0/4470	0.91	4/6096 (0.1%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	772	VAL	N-CA-C	7.04	117.76	110.36
1	C	827	ALA	CA-C-N	6.13	126.58	119.47
1	C	827	ALA	C-N-CA	6.13	126.58	119.47
1	B	715	TRP	N-CA-C	5.38	117.23	111.36

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	1457	0	1479	5	2
1	B	1450	0	1471	3	1
1	C	1442	0	1460	10	0
2	A	156	0	0	2	1
2	B	158	0	0	1	0
2	C	133	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	4796	0	4410	18	2

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 2.

All (18) 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:738:PRO:HG2	2:A:354:HOH:O	1.96	0.65
1:C:762:MET:HE2	1:C:797:ILE:HG13	1.79	0.64
1:C:693:ARG:HD3	1:C:821:THR:HG21	1.82	0.61
1:A:709:SER:HB3	1:A:730:LEU:HD21	1.85	0.58
1:C:693:ARG:CD	1:C:821:THR:HG21	2.35	0.57
1:A:769:SER:HA	1:A:770:PRO:C	2.35	0.52
1:C:647:THR:HA	2:C:424:HOH:O	2.09	0.51
1:C:763:PHE:CE2	1:C:790:PRO:HG3	2.47	0.49
1:C:693:ARG:HD2	1:C:698:LEU:HD21	1.95	0.48
1:C:762:MET:HE2	1:C:797:ILE:CG1	2.43	0.48
1:C:762:MET:HE3	1:C:815:TRP:HZ3	1.79	0.46
1:B:768:SER:HB2	1:B:813:GLU:HG2	1.97	0.45
1:C:762:MET:CE	1:C:815:TRP:HZ3	2.31	0.44
1:B:758:ASP:HB2	2:B:117:HOH:O	2.17	0.44
1:A:774:LYS:HD3	1:A:774:LYS:HA	1.83	0.43
1:B:650:LEU:HD23	1:B:650:LEU:C	2.43	0.43
1:A:681:GLU:HG3	2:A:284:HOH:O	2.19	0.42
1:C:650:LEU:C	1:C:650:LEU:HD23	2.45	0.41

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:803:LYS:NZ	2:A:354:HOH:O[1_455]	2.18	0.02
1:A:794:SER:OG	1:B:794:SER:OG[1_455]	2.19	0.01

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	186/206 (90%)	184 (99%)	2 (1%)	0	100	100
1	B	185/206 (90%)	183 (99%)	2 (1%)	0	100	100
1	C	184/206 (89%)	181 (98%)	2 (1%)	1 (0%)	24	16
All	All	555/618 (90%)	548 (99%)	6 (1%)	1 (0%)	43	36

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	691	PRO

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	165/180 (92%)	165 (100%)	0	100	100
1	B	164/180 (91%)	163 (99%)	1 (1%)	78	79
1	C	163/180 (91%)	162 (99%)	1 (1%)	78	79
All	All	492/540 (91%)	490 (100%)	2 (0%)	84	85

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

Mol	Chain	Res	Type
1	B	772	VAL
1	C	658	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	665	GLN
1	A	721	HIS
1	C	759	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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

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	188/206 (91%)	0.08	7 (3%) 45 52	12, 23, 35, 42	2 (1%)
1	B	187/206 (90%)	0.43	17 (9%) 15 17	13, 24, 38, 50	2 (1%)
1	C	186/206 (90%)	0.63	24 (12%) 7 9	15, 27, 41, 51	6 (3%)
All	All	561/618 (90%)	0.38	48 (8%) 16 19	12, 25, 38, 51	10 (1%)

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	791	ARG	9.3
1	C	792	GLN	8.6
1	C	785	ARG	6.4
1	C	829	GLU	5.1
1	B	805	LYS	4.8
1	B	833	LEU	4.6
1	B	788	GLN	3.8
1	C	798	GLY	3.7
1	C	790	PRO	3.6
1	C	788	GLN	3.4
1	C	805	LYS	3.2
1	C	789	VAL	3.2
1	A	647	THR	3.2
1	B	794	SER	3.1
1	A	833	LEU	3.0
1	B	787	SER	2.9
1	C	763	PHE	2.9
1	B	799	PRO	2.8
1	C	832	LEU	2.8
1	B	758	ASP	2.8
1	A	732	HIS	2.7
1	C	766	PRO	2.6
1	C	796	VAL	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	805	LYS	2.5
1	A	749	GLY	2.5
1	B	786	VAL	2.5
1	B	808	VAL	2.5
1	C	758	ASP	2.5
1	C	797	ILE	2.4
1	C	804	LYS	2.4
1	A	748	ALA	2.4
1	C	800	TYR	2.3
1	B	763	PHE	2.3
1	B	764	VAL	2.3
1	C	810	TYR	2.3
1	C	806	ALA	2.3
1	C	786	VAL	2.3
1	B	780	HIS	2.3
1	B	766	PRO	2.2
1	B	802	GLY	2.2
1	C	808	VAL	2.2
1	C	802	GLY	2.2
1	A	646	PRO	2.1
1	B	748	ALA	2.1
1	B	800	TYR	2.1
1	C	746	LEU	2.1
1	B	803	LYS	2.0
1	C	830	ASN	2.0

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

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

6.3 Carbohydrates [i](#)

There are no oligosaccharides 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.