



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 5, 2026 – 10:28 PM UTC

PDB ID : 5H8A / pdb_00005h8a
Title : Mmi1 YTH domain
Authors : Chatterjee, D.; Goldgur, Y.; Shuman, S.
Deposited on : 2015-12-23
Resolution : 1.75 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

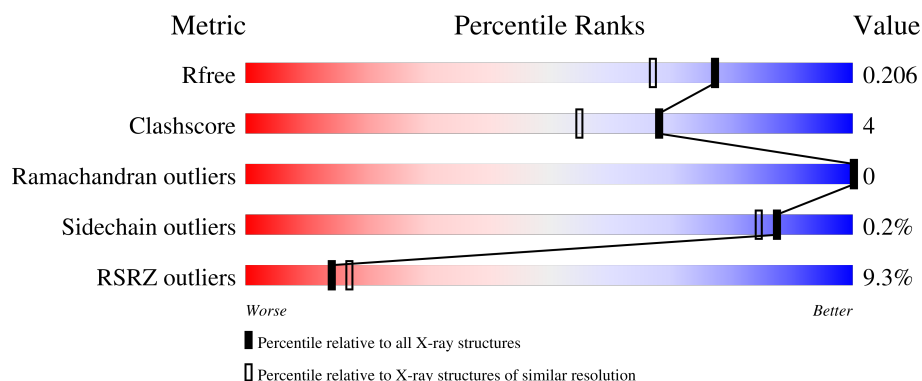
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.75 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	3183 (1.76-1.76)
Clashscore	190562	3299 (1.76-1.76)
Ramachandran outliers	187476	3274 (1.76-1.76)
Sidechain outliers	187428	3274 (1.76-1.76)
RSRZ outliers	180081	3183 (1.76-1.76)

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	178	6% 92% • •
1	B	178	11% 85% 11% •
1	C	178	10% 89% 8% • •
1	D	178	8% 89% 7% •

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6319 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 YTH domain-containing protein mm1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	170	Total	C	N	O	S	0	0	0
			1355	849	241	255	10			
1	B	172	Total	C	N	O	S	0	0	0
			1370	860	243	257	10			
1	C	174	Total	C	N	O	S	0	0	0
			1385	868	248	259	10			
1	D	171	Total	C	N	O	S	0	0	0
			1363	855	242	256	10			

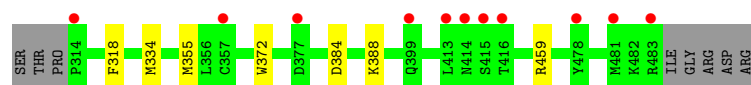
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	247	Total	O	0	0
			247	247		
2	B	231	Total	O	0	0
			231	231		
2	C	193	Total	O	0	0
			193	193		
2	D	175	Total	O	0	0
			175	175		

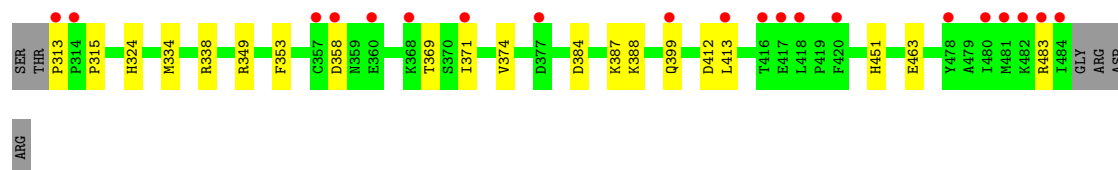
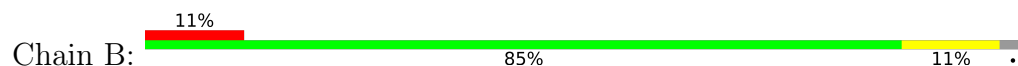
3 Residue-property plots [i](#)

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: YTH domain-containing protein mm1l



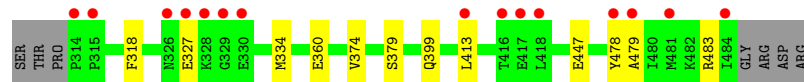
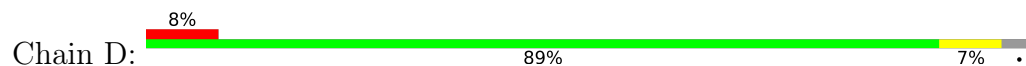
- Molecule 1: YTH domain-containing protein mm1l



- Molecule 1: YTH domain-containing protein mm1l



- Molecule 1: YTH domain-containing protein mm1l



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	49.81Å 79.56Å 195.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.80 – 1.75 48.80 – 1.75	Depositor EDS
% Data completeness (in resolution range)	97.8 (48.80-1.75) 97.3 (48.80-1.75)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.17 (at 1.75Å)	Xtriage
Refinement program	PHENIX (1.10_2155: ???)	Depositor
R, R_{free}	0.187 , 0.205 0.189 , 0.206	Depositor DCC
R_{free} test set	2000 reflections (2.53%)	wwPDB-VP
Wilson B-factor (Å ²)	21.1	Xtriage
Anisotropy	0.541	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 37.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	6319	wwPDB-VP
Average B, all atoms (Å ²)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.69% 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	0.21	0/1383	0.45	0/1861
1	B	0.24	0/1399	0.49	0/1884
1	C	0.20	0/1414	0.48	0/1903
1	D	0.24	0/1391	0.50	2/1872 (0.1%)
All	All	0.22	0/5587	0.48	2/7520 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	327	GLU	CA-C-N	-5.07	109.67	121.52
1	D	327	GLU	C-N-CA	-5.07	109.67	121.52

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	1355	0	1331	7	0
1	B	1370	0	1349	16	0
1	C	1385	0	1365	12	0
1	D	1363	0	1342	8	0
2	A	247	0	0	2	4
2	B	231	0	0	5	3
2	C	193	0	0	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	175	0	0	1	4
All	All	6319	0	5387	39	7

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 4.

All (39) 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:483:ARG:NH1	2:B:501:HOH:O	2.08	0.87
1:A:355:MET:HE2	1:A:372:TRP:HE1	1.42	0.84
1:C:363:ILE:HD13	1:C:404:LEU:HD13	1.59	0.82
1:A:355:MET:SD	2:A:716:HOH:O	2.41	0.77
1:B:324:HIS:CE1	1:B:334:MET:HE3	2.20	0.77
1:C:417:GLU:OE1	2:C:501:HOH:O	2.08	0.72
1:A:384:ASP:OD1	1:A:388:LYS:NZ	2.25	0.70
1:B:374:VAL:HB	1:B:413:LEU:HD11	1.78	0.66
1:C:377:ASP:OD1	2:C:502:HOH:O	2.12	0.66
1:C:338:ARG:NH2	1:D:399:GLN:O	2.30	0.65
1:A:318:PHE:HA	1:A:334:MET:HE1	1.79	0.64
1:B:387:LYS:NZ	1:B:412:ASP:OD2	2.31	0.63
1:C:374:VAL:HB	1:C:413:LEU:HD11	1.86	0.57
1:B:399:GLN:H	1:B:399:GLN:CD	2.13	0.56
1:C:313:PRO:HD3	1:C:478:TYR:CE2	2.42	0.55
1:A:355:MET:HE2	1:A:372:TRP:NE1	2.19	0.54
1:B:358:ASP:HA	1:B:399:GLN:NE2	2.24	0.53
1:B:334:MET:HE1	2:B:524:HOH:O	2.10	0.52
1:B:349:ARG:HD2	1:C:451:HIS:CD2	2.45	0.52
1:C:320:ARG:NH1	2:C:504:HOH:O	2.23	0.51
1:B:369:THR:O	1:B:371:ILE:HD12	2.11	0.51
1:B:451:HIS:HE1	2:B:679:HOH:O	1.93	0.50
1:A:355:MET:CE	1:A:372:TRP:HE1	2.19	0.48
1:D:374:VAL:HB	1:D:413:LEU:HD21	1.96	0.48
1:B:387:LYS:HE2	2:B:670:HOH:O	2.13	0.47
1:C:349:ARG:NH1	1:D:447:GLU:OE1	2.42	0.47
1:D:318:PHE:HA	1:D:334:MET:HE1	1.96	0.46
1:D:379:SER:HB2	1:D:413:LEU:HD22	1.98	0.45
1:B:338:ARG:NH2	1:C:400:THR:HA	2.32	0.45
1:B:384:ASP:OD1	1:B:388:LYS:HE2	2.16	0.45
1:A:459:ARG:NH2	2:A:505:HOH:O	2.50	0.44
1:B:349:ARG:HD3	2:B:617:HOH:O	2.18	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:483:ARG:HD2	1:D:483:ARG:HA	1.93	0.42
1:D:360:GLU:OE1	2:D:501:HOH:O	2.21	0.42
1:B:313:PRO:HA	1:B:315:PRO:HD3	2.03	0.41
1:B:353:PHE:CE2	1:B:463:GLU:HB2	2.56	0.41
1:C:486:ARG:NH1	2:C:513:HOH:O	2.50	0.41
1:D:478:TYR:CD1	1:D:479:ALA:N	2.89	0.41
1:C:399:GLN:HG2	2:C:579:HOH:O	2.21	0.41

All (7) 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)
2:A:680:HOH:O	2:D:646:HOH:O[3_756]	1.89	0.31
2:A:647:HOH:O	2:B:579:HOH:O[4_466]	2.01	0.19
2:A:596:HOH:O	2:B:621:HOH:O[4_566]	2.09	0.11
2:D:649:HOH:O	2:D:672:HOH:O[3_746]	2.09	0.11
2:A:662:HOH:O	2:D:645:HOH:O[3_756]	2.12	0.08
2:B:647:HOH:O	2:B:702:HOH:O[1_655]	2.17	0.03
2:D:630:HOH:O	2:D:639:HOH:O[3_756]	2.18	0.02

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	168/178 (94%)	166 (99%)	2 (1%)	0	100	100
1	B	170/178 (96%)	167 (98%)	3 (2%)	0	100	100
1	C	172/178 (97%)	168 (98%)	4 (2%)	0	100	100
1	D	169/178 (95%)	163 (96%)	6 (4%)	0	100	100
All	All	679/712 (95%)	664 (98%)	15 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	148/155 (96%)	148 (100%)	0	100	100
1	B	150/155 (97%)	150 (100%)	0	100	100
1	C	151/155 (97%)	150 (99%)	1 (1%)	76	67
1	D	149/155 (96%)	149 (100%)	0	100	100
All	All	598/620 (96%)	597 (100%)	1 (0%)	87	84

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	363	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	344	ASN
1	A	454	HIS
1	B	324	HIS
1	B	398	GLN
1	B	451	HIS
1	B	454	HIS
1	C	398	GLN
1	C	402	ASN
1	C	439	ASN
1	C	477	ASN
1	D	365	HIS

5.3.3 RNA ⓘ

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	170/178 (95%)	0.07	11 (6%) 25 29	15, 23, 45, 64	0
1	B	172/178 (96%)	0.43	20 (11%) 9 10	15, 25, 49, 69	0
1	C	174/178 (97%)	0.50	18 (10%) 12 13	15, 27, 57, 72	0
1	D	171/178 (96%)	0.62	15 (8%) 15 19	17, 27, 55, 70	0
All	All	687/712 (96%)	0.41	64 (9%) 14 17	15, 26, 54, 72	0

All (64) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	484	ILE	5.7
1	D	484	ILE	5.6
1	D	478	TYR	5.1
1	D	314	PRO	5.1
1	B	313	PRO	4.9
1	B	314	PRO	4.6
1	C	363	ILE	4.5
1	A	314	PRO	4.5
1	D	481	MET	4.2
1	C	357	CYS	3.9
1	C	486	ARG	3.9
1	D	329	GLY	3.8
1	D	315	PRO	3.7
1	A	483	ARG	3.6
1	C	413	LEU	3.5
1	C	420	PHE	3.5
1	A	399	GLN	3.4
1	C	478	TYR	3.3
1	B	399	GLN	3.3
1	A	481	MET	3.2
1	D	326	ASN	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	414	ASN	3.2
1	C	360	GLU	3.1
1	B	481	MET	3.1
1	C	484	ILE	3.1
1	A	478	TYR	3.0
1	A	357	CYS	3.0
1	A	413	LEU	3.0
1	B	483	ARG	2.9
1	D	418	LEU	2.9
1	C	418	LEU	2.9
1	D	417	GLU	2.8
1	B	357	CYS	2.8
1	D	413	LEU	2.8
1	B	417	GLU	2.7
1	B	482	LYS	2.7
1	D	328	LYS	2.7
1	B	413	LEU	2.7
1	D	416	THR	2.7
1	B	478	TYR	2.6
1	C	481	MET	2.6
1	B	416	THR	2.6
1	B	420	PHE	2.5
1	B	418	LEU	2.5
1	B	377	ASP	2.5
1	C	485	GLY	2.5
1	C	421	TRP	2.5
1	D	330	GLU	2.5
1	C	424	SER	2.5
1	A	416	THR	2.4
1	C	358	ASP	2.4
1	C	361	THR	2.3
1	D	479	ALA	2.2
1	A	377	ASP	2.2
1	B	480	ILE	2.2
1	B	368	LYS	2.2
1	C	416	THR	2.2
1	A	415	SER	2.2
1	D	327	GLU	2.1
1	B	371	ILE	2.1
1	C	417	GLU	2.1
1	B	358	ASP	2.1
1	A	414	ASN	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	360	GLU	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.