Worldwide PDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Oct 20, 2020 – 08:03 AM BST

PDB ID	:	6ZOT
Title	:	Crystal structure of YTHDF3 YTH domain in complex with m6A RNA
Authors	:	Bedi, R.K.; Caflisch, A.
Deposited on	:	2020-07-07
Resolution	:	2.70 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.14.6
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.14.6

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 2.70 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.

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution
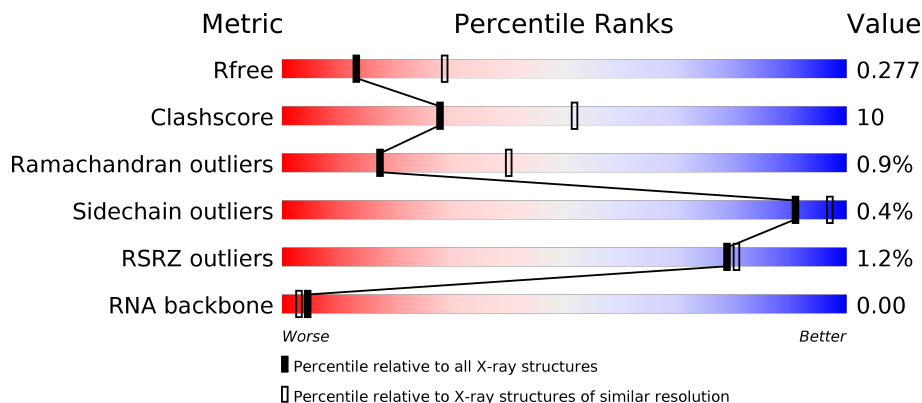
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)
RNA backbone	3102	1159 (3.00-2.40)

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 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	% 64% 11% 25%
1	B	199	2% 66% 23% •• 10%
2	C	5	80% 20%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2825 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 family protein 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	180	Total 1443	C 920	N 252	O 268	S 3	0	0	0
1	A	150	Total 1195	C 759	N 212	O 221	S 3	0	0	0

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	373	MET	-	initiating methionine	UNP Q7Z739
B	374	GLY	-	expression tag	UNP Q7Z739
B	375	HIS	-	expression tag	UNP Q7Z739
B	376	HIS	-	expression tag	UNP Q7Z739
B	377	HIS	-	expression tag	UNP Q7Z739
B	378	HIS	-	expression tag	UNP Q7Z739
B	379	HIS	-	expression tag	UNP Q7Z739
B	380	HIS	-	expression tag	UNP Q7Z739
B	381	SER	-	expression tag	UNP Q7Z739
B	382	SER	-	expression tag	UNP Q7Z739
B	383	GLY	-	expression tag	UNP Q7Z739
B	384	ARG	-	expression tag	UNP Q7Z739
B	385	GLU	-	expression tag	UNP Q7Z739
B	386	ASN	-	expression tag	UNP Q7Z739
B	387	LEU	-	expression tag	UNP Q7Z739
B	388	TYR	-	expression tag	UNP Q7Z739
B	389	PHE	-	expression tag	UNP Q7Z739
B	390	GLN	-	expression tag	UNP Q7Z739
B	391	GLY	-	expression tag	UNP Q7Z739
A	373	MET	-	initiating methionine	UNP Q7Z739
A	374	GLY	-	expression tag	UNP Q7Z739
A	375	HIS	-	expression tag	UNP Q7Z739
A	376	HIS	-	expression tag	UNP Q7Z739
A	377	HIS	-	expression tag	UNP Q7Z739
A	378	HIS	-	expression tag	UNP Q7Z739

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	379	HIS	-	expression tag	UNP Q7Z739
A	380	HIS	-	expression tag	UNP Q7Z739
A	381	SER	-	expression tag	UNP Q7Z739
A	382	SER	-	expression tag	UNP Q7Z739
A	383	GLY	-	expression tag	UNP Q7Z739
A	384	ARG	-	expression tag	UNP Q7Z739
A	385	GLU	-	expression tag	UNP Q7Z739
A	386	ASN	-	expression tag	UNP Q7Z739
A	387	LEU	-	expression tag	UNP Q7Z739
A	388	TYR	-	expression tag	UNP Q7Z739
A	389	PHE	-	expression tag	UNP Q7Z739
A	390	GLN	-	expression tag	UNP Q7Z739
A	391	GLY	-	expression tag	UNP Q7Z739

- Molecule 2 is a RNA chain called RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	4	82	39	15	25	3	0	0	0

- Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	52	Total	O	0	0
			52	52		
3	A	50	Total	O	0	0
			50	50		
3	C	3	Total	O	0	0
			3	3		

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.
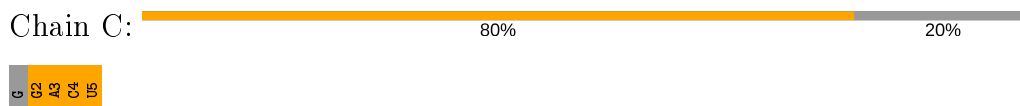
• Molecule 1: YTH domain-containing family protein 3

Chain B:

• Molecule 1: YTH domain-containing family protein 3

Chain A:

• Molecule 2: RNA

Chain C:
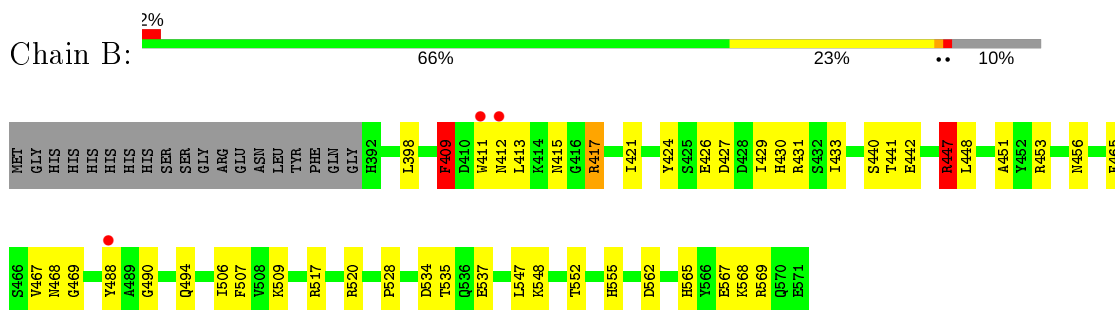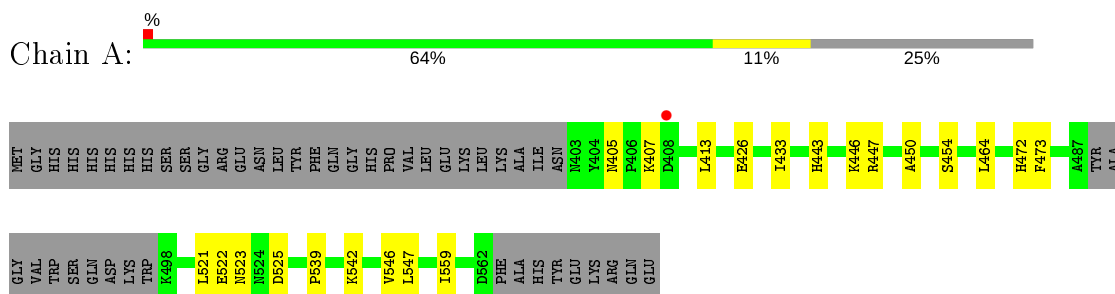
4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 4 21 2	Depositor
Cell constants a, b, c, α, β, γ	100.75Å 100.75Å 72.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.06 – 2.70 45.06 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.7 (45.06-2.70) 99.8 (45.06-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.85 (at 2.69Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.215 , 0.276 0.215 , 0.277	Depositor DCC
R_{free} test set	541 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	36.4	Xtriage
Anisotropy	0.453	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 30.1	EDS
L-test for twinning[2]	$<|L|> = 0.50$, $<L^2> = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	2825	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 3.91% 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: 6MZ

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	0.50	0/1222	0.67	0/1653
1	B	0.59	3/1482 (0.2%)	1.03	10/2012 (0.5%)
2	C	1.67	2/64 (3.1%)	3.20	8/96 (8.3%)
All	All	0.60	5/2768 (0.2%)	1.01	18/3761 (0.5%)

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	B	0	2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	447	ARG	CB-CG	-5.58	1.37	1.52
2	C	5	U	C2-N3	5.52	1.41	1.37
1	B	447	ARG	CG-CD	-5.44	1.38	1.51
2	C	5	U	N3-C4	5.07	1.43	1.38
1	B	442	GLU	CD-OE2	5.04	1.31	1.25

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	447	ARG	NE-CZ-NH2	-18.76	110.92	120.30
2	C	2	G	O4'-C1'-N9	14.81	120.05	108.20
1	B	417	ARG	NE-CZ-NH1	-13.76	113.42	120.30
1	B	417	ARG	NE-CZ-NH2	10.69	125.64	120.30
2	C	4	C	C6-N1-C2	10.06	124.32	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	409	PHE	Sidechain
1	B	447	ARG	Sidechain

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	1195	0	1146	21	0
1	B	1443	0	1347	37	0
2	C	82	0	44	4	0
3	A	50	0	0	5	0
3	B	52	0	0	2	1
3	C	3	0	0	0	0
All	All	2825	0	2537	52	1

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 10.

The worst 5 of 52 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:426:GLU:HG2	1:B:467:VAL:HG21	1.65	0.77
1:A:413:LEU:HD11	1:A:547:LEU:HD12	1.72	0.71
1:B:409:PHE:CE1	1:B:506:ILE:HG21	2.27	0.70
1:B:447:ARG:NH2	1:B:535:THR:O	2.28	0.66
1:A:472:HIS:O	3:A:601:HOH:O	2.13	0.65

All (1) 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)
3:B:650:HOH:O	3:B:650:HOH:O[8_555]	1.99	0.21

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	146/199 (73%)	138 (94%)	8 (6%)	0	100	100
1	B	178/199 (89%)	158 (89%)	17 (10%)	3 (2%)	9	23
All	All	324/398 (81%)	296 (91%)	25 (8%)	3 (1%)	17	40

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	415	ASN
1	B	412	ASN
1	B	456	ASN

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	128/178 (72%)	128 (100%)	0	100	100
1	B	150/178 (84%)	149 (99%)	1 (1%)	84	94
All	All	278/356 (78%)	277 (100%)	1 (0%)	91	97

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

Mol	Chain	Res	Type
1	B	409	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such

sidechains are listed below:

Mol	Chain	Res	Type
1	B	531	ASN

5.3.3   RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	C	2/5 (40%)	1 (50%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	C	4	C

There are no RNA pucker outliers to report.

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
2	6MZ	C	3	2	18,25,26	2.18	3 (16%)	16,36,39	3.35	4 (25%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	6MZ	C	3	2	-	2/5/27/28	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	3	6MZ	C6-N6	8.03	1.48	1.35
2	C	3	6MZ	C5-C4	-2.51	1.34	1.40
2	C	3	6MZ	C2-N3	2.37	1.35	1.32

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	3	6MZ	C1'-N9-C4	-11.42	106.58	126.64
2	C	3	6MZ	N3-C2-N1	-4.80	121.18	128.68
2	C	3	6MZ	C2-N1-C6	3.93	119.96	116.59
2	C	3	6MZ	C9-N6-C6	2.11	124.69	122.87

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	3	6MZ	C3'-C4'-C5'-O5'
2	C	3	6MZ	O4'-C4'-C5'-O5'

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	3	6MZ	3	0

5.5 Carbohydrates (i)

There are no monosaccharides 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 (i)

6.1 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(Å²)	Q<0.9
1	A	150/199 (75%)	-0.43	1 (0%)	87	89	14, 23, 41, 58	0
1	B	180/199 (90%)	-0.32	3 (1%)	70	72	18, 31, 51, 66	0
2	C	3/5 (60%)	1.13	0	100	100	53, 53, 59, 61	0
All	All	333/403 (82%)	-0.36	4 (1%)	79	80	14, 27, 50, 66	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	412	ASN	5.2
1	B	411	TRP	4.1
1	B	488	TYR	2.6
1	A	408	ASP	2.2

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, 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	6MZ	C	3	23/24	0.92	0.17	21,27,51,55	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.