



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 24, 2024 – 01:09 pm BST

PDB ID : 9H35  
Title : Crystal structure of the YTHDC2 YTH domain  
Authors : Bedi, R.K.; Caffisch, A.  
Deposited on : 2024-10-15  
Resolution : 2.68 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](mailto: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 : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

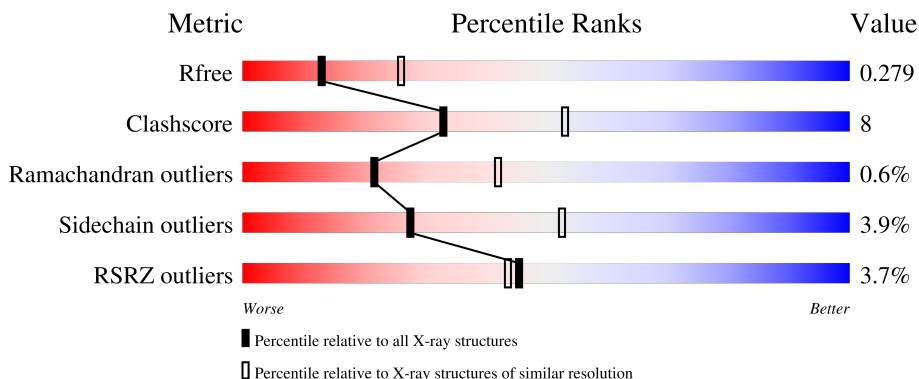
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.68 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))
$R_{free}$	164625	4708 (2.70-2.66)
Clashscore	180529	5138 (2.70-2.66)
Ramachandran outliers	177936	5071 (2.70-2.66)
Sidechain outliers	177891	5071 (2.70-2.66)
RSRZ outliers	164620	4708 (2.70-2.66)

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  $\geq 3$ , 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	141	 82% 15% ..
1	B	141	 2% 74% 21% 6%
1	C	141	 6% 73% 18% • 7%
1	D	141	 6% 80% 14% • •

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4170 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 3'-5' RNA helicase YTHDC2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	138	1094	704	191	196	3	0	0	0
1	B	133	1036	670	179	184	3	0	0	0
1	C	131	993	639	170	181	3	0	0	0
1	D	136	1009	648	172	186	3	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1284	GLY	-	expression tag	UNP Q9H6S0
B	1284	GLY	-	expression tag	UNP Q9H6S0
C	1284	GLY	-	expression tag	UNP Q9H6S0
D	1284	GLY	-	expression tag	UNP Q9H6S0


- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	21	Total	O	0	0
			21	21		
2	B	12	Total	O	0	0
			12	12		
2	C	3	Total	O	0	0
			3	3		
2	D	2	Total	O	0	0
			2	2		

### 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.

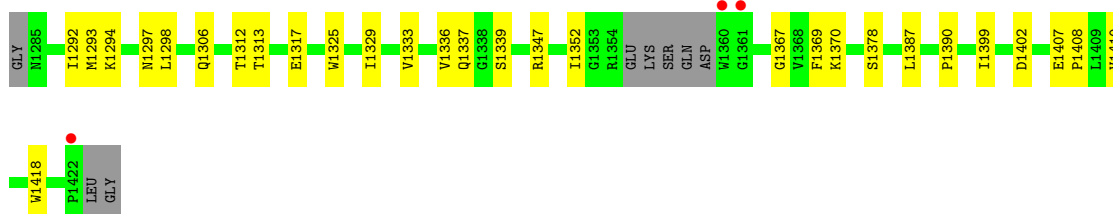
- Molecule 1: 3'-5' RNA helicase YTHDC2

Chain A: 




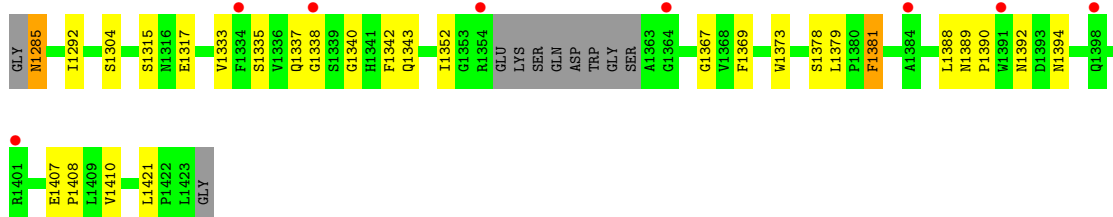
- Molecule 1: 3'-5' RNA helicase YTHDC2

Chain B: 




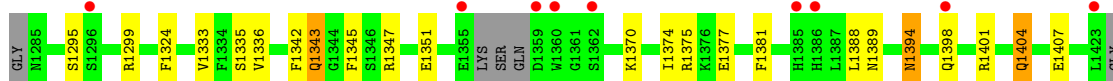
- Molecule 1: 3'-5' RNA helicase YTHDC2

Chain C: 



- Molecule 1: 3'-5' RNA helicase YTHDC2

Chain D: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	115.97Å 115.97Å 205.85Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.79 – 2.68 48.79 – 2.68	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.79-2.68) 99.8 (48.79-2.68)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.21 (at 2.69Å)	Xtrriage
Refinement program	PHENIX 1.21.1_5286	Depositor
R, $R_{free}$	0.227 , 0.284 0.222 , 0.279	Depositor DCC
$R_{free}$ test set	1139 reflections (4.80%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	77.5	Xtrriage
Anisotropy	0.350	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 62.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	4170	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	74.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.65% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	0.48	0/1124	0.65	0/1521
1	B	0.50	0/1062	0.67	0/1437
1	C	0.45	0/1020	0.55	0/1388
1	D	0.44	0/1034	0.57	0/1406
All	All	0.47	0/4240	0.61	0/5752

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	1094	0	1049	10	0
1	B	1036	0	990	21	0
1	C	993	0	897	19	0
1	D	1009	0	902	18	0
2	A	21	0	0	0	0
2	B	12	0	0	1	0
2	C	3	0	0	0	0
2	D	2	0	0	1	0
All	All	4170	0	3838	61	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 8.

All (61) 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:D:1336:VAL:H	1:D:1343:GLN:NE2	1.79	0.80
1:D:1336:VAL:H	1:D:1343:GLN:HE22	1.33	0.74
1:B:1367:GLY:N	2:B:1501:HOH:O	2.12	0.69
1:B:1306:GLN:HE21	1:C:1285:ASN:HD21	1.39	0.69
1:D:1375:ARG:NH1	1:D:1377:GLU:OE1	2.32	0.63
1:A:1356:LYS:HD3	1:B:1297:ASN:HB3	1.82	0.62
1:A:1294:LYS:HB2	1:A:1402:ASP:HB2	1.82	0.61
1:B:1294:LYS:HB2	1:B:1402:ASP:HB2	1.82	0.61
1:A:1317:GLU:HG2	1:A:1352:ILE:HG21	1.83	0.61
1:A:1387:LEU:HD22	1:A:1410:VAL:HG23	1.84	0.59
1:A:1350:SER:HB3	1:A:1370:LYS:HB3	1.85	0.58
1:D:1389:ASN:HD21	1:D:1407:GLU:H	1.51	0.57
1:D:1333:VAL:HG22	1:D:1345:PHE:HB3	1.86	0.57
1:C:1340:GLY:HA2	1:C:1381:PHE:CD2	2.41	0.55
1:A:1360:TRP:HZ3	1:B:1337:GLN:HG3	1.71	0.55
1:D:1336:VAL:HG23	1:D:1343:GLN:HE21	1.72	0.53
1:B:1325:TRP:CZ3	1:D:1370:LYS:HB2	2.44	0.53
1:D:1347:ARG:HD3	2:D:1502:HOH:O	2.08	0.52
1:C:1407:GLU:OE1	1:C:1410:VAL:HG23	2.10	0.51
1:D:1335:SER:HB2	1:D:1342:PHE:HD1	1.75	0.51
1:D:1389:ASN:ND2	1:D:1407:GLU:H	2.10	0.50
1:C:1292:ILE:HD13	1:C:1333:VAL:HB	1.92	0.50
1:B:1370:LYS:HE2	1:D:1324:PHE:O	2.12	0.50
1:D:1401:ARG:O	1:D:1404:GLN:HB2	2.11	0.49
1:A:1360:TRP:CZ3	1:B:1337:GLN:HG3	2.49	0.48
1:C:1317:GLU:HG2	1:C:1352:ILE:HD13	1.96	0.47
1:C:1388:LEU:O	1:C:1390:PRO:HD3	2.14	0.47
1:B:1312:THR:OG1	1:B:1313:THR:N	2.48	0.47
1:B:1329:ILE:HG23	1:B:1347:ARG:HG3	1.97	0.47
1:B:1317:GLU:HG2	1:B:1352:ILE:HG21	1.97	0.46
1:B:1352:ILE:HG12	1:B:1369:PHE:CD1	2.51	0.45
1:A:1407:GLU:HG2	1:A:1408:PRO:HD2	1.97	0.45
1:B:1292:ILE:HD13	1:B:1333:VAL:HB	1.99	0.45
1:B:1387:LEU:HD23	1:B:1387:LEU:HA	1.85	0.45
1:C:1388:LEU:HB3	1:C:1394:ASN:O	2.17	0.44
1:C:1304:SER:HG	1:C:1373:TRP:HZ2	1.63	0.44
1:C:1335:SER:HA	1:C:1343:GLN:OE1	2.17	0.44
1:D:1388:LEU:HB3	1:D:1394:ASN:O	2.18	0.44
1:C:1389:ASN:HB3	1:C:1392:ASN:HB3	1.99	0.44
1:C:1352:ILE:HG12	1:C:1369:PHE:CD1	2.53	0.44

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:1379:LEU:HA	1:C:1421:LEU:HD11	2.00	0.44
1:D:1374:ILE:HG22	1:D:1375:ARG:HG2	2.01	0.43
1:D:1381:PHE:HD2	1:D:1398:GLN:CB	2.32	0.43
1:B:1407:GLU:HG2	1:B:1408:PRO:HD2	2.01	0.43
1:C:1407:GLU:HG2	1:C:1408:PRO:HD2	2.00	0.43
1:D:1295:SER:HB3	1:D:1336:VAL:HG22	2.01	0.43
1:B:1399:ILE:HG22	1:B:1399:ILE:O	2.19	0.42
1:B:1390:PRO:HG2	1:B:1407:GLU:CB	2.49	0.42
1:C:1337:GLN:OE1	1:C:1338:GLY:N	2.51	0.42
1:B:1292:ILE:O	1:B:1293:MET:HE2	2.20	0.42
1:B:1306:GLN:HE21	1:C:1285:ASN:ND2	2.11	0.42
1:A:1292:ILE:HG12	1:A:1406:LEU:HD21	2.02	0.42
1:B:1336:VAL:HB	1:B:1339:SER:HB3	2.02	0.42
1:C:1421:LEU:HD23	1:C:1421:LEU:HA	1.87	0.42
1:D:1295:SER:O	1:D:1336:VAL:HA	2.20	0.41
1:A:1375:ARG:HH21	1:A:1377:GLU:CD	2.24	0.41
1:B:1407:GLU:O	1:B:1410:VAL:HG12	2.20	0.41
1:C:1335:SER:HB2	1:C:1342:PHE:CD1	2.56	0.41
1:D:1343:GLN:N	1:D:1343:GLN:OE1	2.54	0.41
1:C:1352:ILE:HG23	1:C:1367:GLY:HA2	2.03	0.41
1:C:1389:ASN:ND2	1:C:1392:ASN:HB2	2.36	0.40

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	136/141 (96%)	126 (93%)	7 (5%)	3 (2%)	5	13
1	B	129/141 (92%)	122 (95%)	7 (5%)	0	100	100
1	C	127/141 (90%)	124 (98%)	3 (2%)	0	100	100
1	D	132/141 (94%)	124 (94%)	8 (6%)	0	100	100

Continued on next page...



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	524/564 (93%)	496 (95%)	25 (5%)	3 (1%)	22	42

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1394	ASN
1	A	1355	GLU
1	A	1354	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	115/125 (92%)	111 (96%)	4 (4%)	31	57
1	B	106/125 (85%)	103 (97%)	3 (3%)	38	65
1	C	98/125 (78%)	94 (96%)	4 (4%)	26	50
1	D	96/125 (77%)	91 (95%)	5 (5%)	19	41
All	All	415/500 (83%)	399 (96%)	16 (4%)	27	52

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1301	LEU
1	A	1350	SER
1	A	1395	LYS
1	A	1418	TRP
1	B	1298	LEU
1	B	1378	SER
1	B	1418	TRP
1	C	1285	ASN
1	C	1315	SER
1	C	1378	SER
1	C	1381	PHE
1	D	1299	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	1343	GLN
1	D	1351	GLU
1	D	1394	ASN
1	D	1404	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1300	ASN
1	B	1306	GLN
1	D	1343	GLN
1	D	1389	ASN

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	138/141 (97%)	-0.24	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	47, 62, 84, 99	0
1	B	133/141 (94%)	-0.15	3 (2%) <span style="border: 1px solid blue; padding: 2px;">61</span> <span style="border: 1px solid blue; padding: 2px;">60</span>	48, 64, 83, 100	0
1	C	131/141 (92%)	0.27	8 (6%) <span style="border: 1px solid red; padding: 2px;">28</span> <span style="border: 1px solid red; padding: 2px;">26</span>	60, 84, 110, 118	0
1	D	136/141 (96%)	0.37	9 (6%) <span style="border: 1px solid red; padding: 2px;">26</span> <span style="border: 1px solid red; padding: 2px;">24</span>	58, 86, 112, 122	0
All	All	538/564 (95%)	0.06	20 (3%) <span style="border: 1px solid gray; padding: 2px;">45</span> <span style="border: 1px solid gray; padding: 2px;">44</span>	47, 71, 107, 122	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	1359	ASP	3.7
1	B	1360	TRP	3.3
1	B	1361	GLY	3.1
1	D	1423	LEU	3.0
1	B	1422	PRO	2.9
1	C	1391	TRP	2.9
1	D	1398	GLN	2.9
1	D	1296	SER	2.9
1	C	1354	ARG	2.8
1	C	1364	GLY	2.7
1	D	1385	HIS	2.6
1	D	1360	TRP	2.6
1	D	1386	HIS	2.4
1	C	1398	GLN	2.3
1	C	1384	ALA	2.2
1	D	1362	SER	2.2
1	C	1401	ARG	2.1
1	D	1355	GLU	2.1
1	C	1338	GLY	2.1
1	C	1334	PHE	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 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.