



# Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 12:35 pm BST

PDB ID : 5WYL  
Title : Crystal structure of Chaetomium thermophilum Utp10 N-terminal domain in complex with Utp17 C-terminal helices  
Authors : Chen, R.; Zhu, X.; Ye, K.  
Deposited on : 2017-01-13  
Resolution : 2.64 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 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	:	2.11
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)	:	2.11

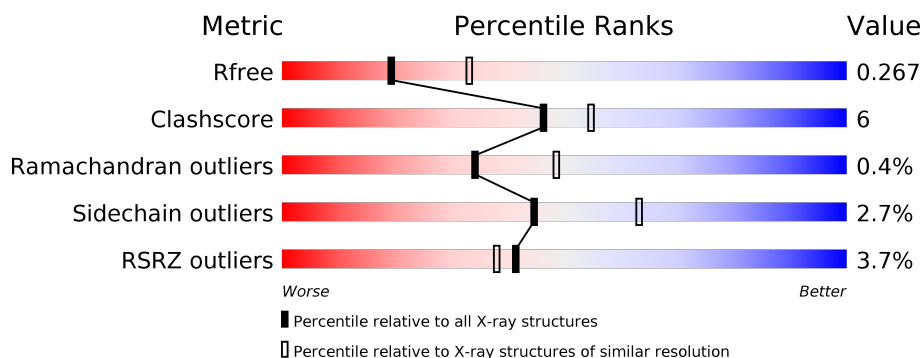
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.64 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	1426 (2.66-2.62)
Clashscore	141614	1472 (2.66-2.62)
Ramachandran outliers	138981	1446 (2.66-2.62)
Sidechain outliers	138945	1446 (2.66-2.62)
RSRZ outliers	127900	1408 (2.66-2.62)

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  $\geq 3$ , 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	471	<div> <div>3%</div> <div> <div></div> <div>73%</div> <div>14%</div> <div>13%</div> </div> </div>
1	C	471	<div> <div>4%</div> <div> <div></div> <div>74%</div> <div>14%</div> <div>11%</div> </div> </div>
2	B	54	<div> <div></div> <div> <div>63%</div> <div>9%</div> <div>28%</div> </div> </div>
2	D	54	<div> <div></div> <div> <div>70%</div> <div>6%</div> <div>24%</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7086 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 Putative uncharacterized protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	411	Total	C	N	O	S	0	0	0
			3184	2041	550	579	14			
1	C	418	Total	C	N	O	S	0	0	0
			3242	2069	572	587	14			

- Molecule 2 is a protein called Putative uncharacterized protein.

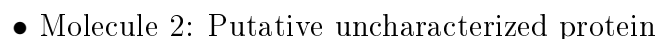
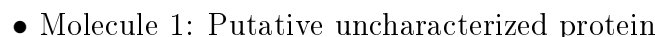
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	39	Total	C	N	O	S	0	0	0
			294	196	45	52	1			
2	D	41	Total	C	N	O	S	0	0	0
			311	205	49	56	1			

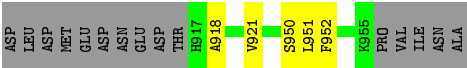
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	30	Total	O	0	0
			30	30		
3	B	1	Total	O	0	0
			1	1		
3	C	18	Total	O	0	0
			18	18		
3	D	6	Total	O	0	0
			6	6		

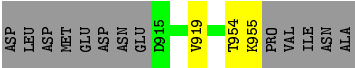


- Molecule 1: Putative uncharacterized protein





- Molecule 2: Putative uncharacterized protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	98.31Å 98.31Å 235.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.01 – 2.64 39.01 – 2.64	Depositor EDS
% Data completeness (in resolution range)	99.9 (39.01-2.64) 99.9 (39.01-2.64)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	13.40 (at 2.65Å)	Xtriage
Refinement program	PHENIX (1.11rc1_2513: ???)	Depositor
R, $R_{free}$	0.212 , 0.267 0.212 , 0.267	Depositor DCC
$R_{free}$ test set	1998 reflections (5.72%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	61.0	Xtriage
Anisotropy	0.066	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 38.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7086	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.96% 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	0.47	0/3240	0.59	0/4394
1	C	0.45	0/3300	0.61	0/4477
2	B	0.41	0/303	0.56	0/416
2	D	0.50	0/320	0.58	0/439
All	All	0.46	0/7163	0.60	0/9726

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	3184	0	3269	38	0
1	C	3242	0	3331	48	0
2	B	294	0	287	4	0
2	D	311	0	303	3	0
3	A	30	0	0	4	0
3	B	1	0	0	0	0
3	C	18	0	0	2	0
3	D	6	0	0	0	0
All	All	7086	0	7190	89	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 6.

All (89) 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:C:83:ARG:NH2	1:C:95:ASP:OD1	2.11	0.84
1:A:83:ARG:NH2	1:A:95:ASP:OD1	2.12	0.81
1:A:336:PRO:O	3:A:501:HOH:O	2.04	0.74
1:C:415:ASN:O	1:C:419:GLN:HG2	1.89	0.72
1:A:319:ARG:O	1:A:353:LYS:NZ	2.24	0.71
1:C:87:THR:HG22	1:C:88:ALA:H	1.55	0.70
1:A:450:GLU:O	1:A:454:LEU:HD22	1.91	0.70
1:A:163:GLU:OE1	3:A:502:HOH:O	2.10	0.69
1:C:413:ASP:HA	1:C:416:ILE:HG22	1.76	0.67
1:C:333:ILE:HD12	1:C:377:VAL:HG11	1.77	0.67
1:C:87:THR:HB	1:C:90:GLU:H	1.62	0.64
1:C:410:VAL:O	1:C:413:ASP:HB2	1.99	0.61
1:C:333:ILE:CD1	1:C:377:VAL:HG11	2.32	0.59
1:C:36:GLU:HG3	1:C:37:PRO:HD2	1.84	0.58
1:C:87:THR:HG22	1:C:88:ALA:N	2.19	0.58
1:C:442:ILE:HA	1:C:445:VAL:HG12	1.85	0.57
1:C:389:GLU:O	1:C:393:LYS:HG3	2.04	0.57
1:C:129:GLU:HG2	1:C:156:PRO:HG3	1.87	0.56
1:C:163:GLU:CD	1:C:163:GLU:H	2.07	0.56
1:C:253:PRO:O	1:C:257:GLU:HG2	2.05	0.56
1:C:190:PRO:O	3:C:501:HOH:O	2.17	0.56
1:A:390:LYS:O	1:A:394:VAL:HG12	2.06	0.55
1:C:372:ILE:HG23	1:C:375:LEU:HD12	1.87	0.55
1:A:249:GLN:NE2	3:A:507:HOH:O	2.40	0.55
1:C:445:VAL:O	1:C:446:ASP:HB2	2.08	0.54
1:A:227:GLY:O	1:A:231:LYS:HG2	2.08	0.54
1:C:132:THR:HG23	1:C:155:LEU:HD22	1.90	0.53
1:A:281:LYS:O	1:A:283:SER:N	2.42	0.53
1:C:157:VAL:HG22	1:C:165:ARG:HH12	1.74	0.53
1:C:279:ALA:HB2	1:C:315:LEU:HD23	1.90	0.52
1:C:425:VAL:O	1:C:429:GLN:HG3	2.09	0.52
1:A:356:ASN:HD22	1:A:391:GLN:HB3	1.75	0.52
1:C:413:ASP:CA	1:C:416:ILE:HG22	2.40	0.52
1:C:116:LYS:NZ	3:C:502:HOH:O	2.32	0.51
1:C:435:GLY:O	1:C:439:ARG:HG2	2.13	0.49
1:A:293:MET:HB3	1:A:328:VAL:HG11	1.95	0.49
1:A:281:LYS:NZ	2:B:921:VAL:HG12	2.27	0.49
1:C:406:ILE:O	1:C:406:ILE:HG13	2.12	0.49
1:C:333:ILE:HG13	1:C:333:ILE:O	2.13	0.48
1:A:383:LEU:HD21	1:A:427:LEU:HD23	1.94	0.48

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:50:SER:O	1:A:54:GLU:HG2	2.13	0.48
1:C:165:ARG:NH2	1:C:168:ASP:OD2	2.46	0.48
1:A:310:VAL:HG12	1:A:345:ILE:HG12	1.95	0.48
1:C:293:MET:HB3	1:C:328:VAL:HG11	1.95	0.48
1:A:284:LEU:HD13	1:A:288:ILE:HG23	1.96	0.48
1:A:442:ILE:HG21	1:A:449:ILE:HD11	1.97	0.46
1:A:324:LEU:HD11	1:A:354:LEU:HD12	1.98	0.45
1:A:205:ARG:NH1	1:A:257:GLU:OE1	2.46	0.45
1:C:270:ALA:O	1:C:273:MET:HB2	2.16	0.45
1:A:277:ILE:HG23	2:B:952:PHE:CE2	2.51	0.45
1:A:350:GLN:NE2	3:A:505:HOH:O	2.50	0.45
1:A:451:GLU:O	1:A:455:LYS:HB3	2.16	0.45
1:C:284:LEU:HD22	1:C:288:ILE:HG21	1.99	0.45
1:C:324:LEU:HG	1:C:357:GLY:HA3	1.99	0.45
1:C:413:ASP:O	1:C:416:ILE:HG22	2.17	0.44
1:A:267:ILE:HA	1:A:267:ILE:HD12	1.91	0.44
1:C:372:ILE:HD11	1:C:416:ILE:HG13	1.99	0.44
1:C:350:GLN:HG3	2:D:954:THR:O	2.18	0.44
1:A:324:LEU:HG	1:A:357:GLY:HA3	1.99	0.44
1:A:356:ASN:ND2	1:A:391:GLN:HB3	2.33	0.44
1:A:290:THR:O	1:A:294:GLU:HG3	2.18	0.44
1:C:362:PHE:HB3	1:C:374:THR:HB	2.00	0.43
1:C:193:LEU:HA	1:C:193:LEU:HD12	1.85	0.43
1:A:86:MET:HB2	1:A:91:ASN:ND2	2.34	0.43
1:C:382:LEU:HD23	1:C:382:LEU:HA	1.86	0.43
1:A:192:LEU:O	1:A:196:ILE:HG13	2.19	0.43
1:C:135:LEU:HD23	1:C:135:LEU:HA	1.79	0.43
2:B:921:VAL:HG11	2:B:952:PHE:CZ	2.53	0.43
1:C:324:LEU:HD11	1:C:354:LEU:HD12	2.00	0.42
1:C:356:ASN:O	1:C:359:VAL:HG12	2.19	0.42
1:A:289:LEU:HA	1:A:289:LEU:HD23	1.86	0.42
1:C:166:PHE:HB3	1:C:185:GLN:CD	2.40	0.42
1:A:381:LEU:O	1:A:384:SER:OG	2.33	0.42
1:C:278:LEU:HB3	1:C:284:LEU:HD11	2.01	0.42
1:A:31:LYS:HB3	1:A:154:ILE:HB	2.02	0.42
1:C:371:ASP:CG	1:C:373:ARG:HG2	2.41	0.42
1:A:309:LEU:O	1:A:313:THR:HG23	2.20	0.41
1:A:59:LEU:HA	1:A:62:LEU:HD12	2.01	0.41
1:A:189:ARG:HA	1:A:189:ARG:HD3	1.80	0.41
1:C:255:LEU:O	1:C:259:MET:HG3	2.20	0.41
1:A:288:ILE:HA	1:A:288:ILE:HD12	1.88	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:919:VAL:CG1	2:D:954:THR:HG22	2.51	0.41
1:A:284:LEU:HB3	1:A:288:ILE:CG2	2.51	0.41
1:C:353:LYS:HE3	1:C:353:LYS:HB2	1.78	0.41
1:C:439:ARG:HG2	1:C:439:ARG:H	1.68	0.41
1:C:346:SER:HA	1:C:349:HIS:O	2.22	0.40
1:A:177:PRO:HA	1:A:178:PRO:HD2	1.98	0.40
1:A:281:LYS:HZ3	2:B:921:VAL:HG12	1.86	0.40
2:D:955:LYS:HB2	2:D:955:LYS:HE3	1.91	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	405/471 (86%)	395 (98%)	9 (2%)	1 (0%)	47	64
1	C	416/471 (88%)	403 (97%)	11 (3%)	2 (0%)	29	43
2	B	37/54 (68%)	36 (97%)	0	1 (3%)	5	6
2	D	39/54 (72%)	38 (97%)	1 (3%)	0	100	100
All	All	897/1050 (85%)	872 (97%)	21 (2%)	4 (0%)	34	48

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	282	GLY
2	B	918	ALA
1	C	87	THR
1	C	408	ASP

### 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	343/397 (86%)	333 (97%)	10 (3%)	42	60
1	C	349/397 (88%)	341 (98%)	8 (2%)	50	68
2	B	30/45 (67%)	28 (93%)	2 (7%)	16	24
2	D	32/45 (71%)	32 (100%)	0	100	100
All	All	754/884 (85%)	734 (97%)	20 (3%)	44	63

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	128	HIS
1	A	203	SER
1	A	230	ASP
1	A	292	PHE
1	A	300	TRP
1	A	347	LYS
1	A	352	ASP
1	A	443	GLN
1	A	451	GLU
1	A	455	LYS
2	B	950	SER
2	B	951	LEU
1	C	163	GLU
1	C	185	GLN
1	C	215	SER
1	C	300	TRP
1	C	415	ASN
1	C	426	ARG
1	C	439	ARG
1	C	446	ASP

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

Mol	Chain	Res	Type
1	A	250	GLN
1	A	356	ASN
1	C	185	GLN
1	C	415	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 carbohydrates 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	411/471 (87%)	0.28	15 (3%) 42 39	40, 59, 93, 108	0
1	C	418/471 (88%)	0.31	19 (4%) 33 30	37, 59, 94, 128	0
2	B	39/54 (72%)	0.18	0 100 100	54, 71, 90, 109	0
2	D	41/54 (75%)	-0.13	0 100 100	41, 51, 73, 97	0
All	All	909/1050 (86%)	0.27	34 (3%) 41 38	37, 59, 94, 128	0

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	414	GLY	8.0
1	C	411	ASP	4.7
1	C	423	ALA	4.5
1	C	415	ASN	4.4
1	C	409	ASN	4.4
1	A	283	SER	4.1
1	C	413	ASP	4.0
1	C	412	ALA	3.9
1	C	410	VAL	3.6
1	A	416	ILE	3.5
1	A	320	SER	3.4
1	A	456	LEU	3.4
1	A	370	GLY	3.4
1	C	416	ILE	3.3
1	C	407	ASP	3.2
1	A	233	ARG	3.1
1	C	428	SER	3.1
1	A	404	HIS	3.1
1	A	321	ALA	3.0
1	A	415	ASN	2.9
1	C	440	THR	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	445	VAL	2.6
1	C	400	LEU	2.6
1	C	425	VAL	2.5
1	C	439	ARG	2.4
1	C	404	HIS	2.4
1	A	402	ALA	2.4
1	A	417	ARG	2.4
1	C	436	ASP	2.3
1	A	406	ILE	2.3
1	A	284	LEU	2.3
1	A	457	GLY	2.2
1	C	418	LYS	2.2
1	A	451	GLU	2.1

## 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 carbohydrates 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.