



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 7, 2019 – 09:46 AM EDT

PDB ID : 6J0X  
Title : Crystal Structure of Yeast Rtt107 and Mms22  
Authors : Wan, B.; Wu, J.; Lei, M.  
Deposited on : 2018-12-27  
Resolution : 2.31 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:

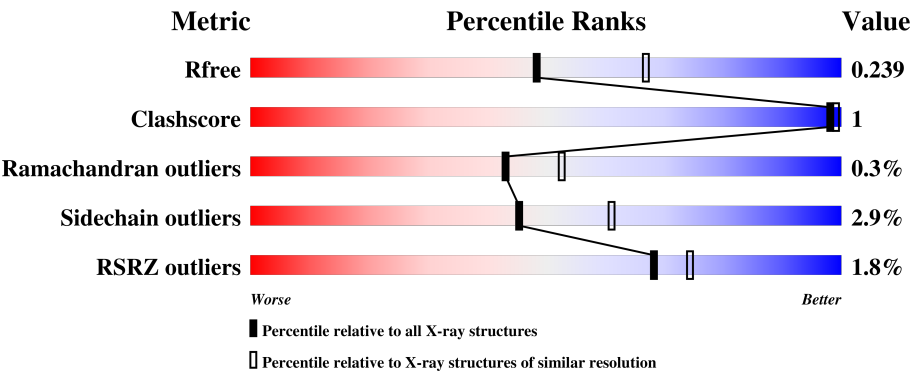
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.4
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.4

# 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 2.31 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 <sub>free</sub>	111664	5225 (2.34-2.30)
Clashscore	122126	5849 (2.34-2.30)
Ramachandran outliers	120053	5790 (2.34-2.30)
Sidechain outliers	120020	5789 (2.34-2.30)
RSRZ outliers	108989	5109 (2.34-2.30)

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. 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	513	<div><div>2%</div><div>86%</div><div>5%</div><div>9%</div></div>
1	B	513	<div><div>2%</div><div>87%</div><div>•</div><div>9%</div></div>
1	C	513	<div><div>•</div><div>84%</div><div>6%</div><div>10%</div></div>
1	D	513	<div><div>•</div><div>87%</div><div>•</div><div>10%</div></div>
2	E	16	<div><div>19%</div><div>100%</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	16	<div><div></div><div>6%</div><div>100%</div></div>
2	G	16	<div><div></div><div>19%</div><div>100%</div></div>
2	H	16	<div><div></div><div>94%</div><div>6%</div></div>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 16129 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 Regulator of Ty1 transposition protein 107.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	465	Total	C	N	O	S	0	0	0
			3814	2485	625	686	18			
1	B	466	Total	C	N	O	S	0	0	0
			3822	2489	626	689	18			
1	C	463	Total	C	N	O	S	0	0	0
			3797	2474	621	684	18			
1	D	462	Total	C	N	O	S	0	0	0
			3799	2477	624	680	18			

- Molecule 2 is a protein called Peptide from E3 ubiquitin-protein ligase substrate receptor MMS22.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	E	16	Total	C	N	O	0	0	0
			143	95	19	29			
2	F	16	Total	C	N	O	0	0	0
			143	95	19	29			
2	G	16	Total	C	N	O	0	0	0
			143	95	19	29			
2	H	16	Total	C	N	O	0	0	0
			143	95	19	29			

- Molecule 3 is water.

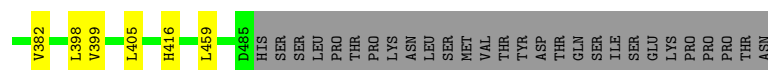
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	127	Total	O	0	0
			127	127		
3	E	5	Total	O	0	0
			5	5		
3	B	82	Total	O	0	0
			82	82		
3	F	2	Total	O	0	0
			2	2		

*Continued on next page...*

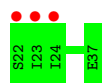
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	25	Total 25	O 25	0	0
3	G	1	Total 1	O 1	0	0
3	D	80	Total 80	O 80	0	0
3	H	3	Total 3	O 3	0	0

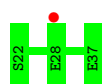




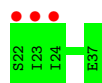
- Molecule 2: Peptide from E3 ubiquitin-protein ligase substrate receptor MMS22



- Molecule 2: Peptide from E3 ubiquitin-protein ligase substrate receptor MMS22



- Molecule 2: Peptide from E3 ubiquitin-protein ligase substrate receptor MMS22



- Molecule 2: Peptide from E3 ubiquitin-protein ligase substrate receptor MMS22



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	72.28Å 99.67Å 167.06Å 90.00° 94.92° 90.00°	Depositor
Resolution (Å)	49.43 – 2.31 85.51 – 2.31	Depositor EDS
% Data completeness (in resolution range)	99.0 (49.43-2.31) 99.1 (85.51-2.31)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.60 (at 2.32Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, $R_{free}$	0.204 , 0.239 0.203 , 0.239	Depositor DCC
$R_{free}$ test set	5086 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	36.0	Xtriage
Anisotropy	0.220	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 33.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	16129	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 36.23 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.1617e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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	0.43	0/3921	0.61	0/5320
1	B	0.44	0/3929	0.58	0/5331
1	C	0.43	0/3903	0.60	0/5295
1	D	0.44	0/3905	0.61	0/5296
2	E	0.47	0/148	0.56	0/202
2	F	0.45	0/148	0.53	0/202
2	G	0.50	0/148	0.54	0/202
2	H	0.47	0/148	0.60	0/202
All	All	0.44	0/16250	0.60	0/22050

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	3814	0	3798	7	0
1	B	3822	0	3802	5	0
1	C	3797	0	3773	7	0
1	D	3799	0	3784	5	0
2	E	143	0	122	0	0
2	F	143	0	122	0	0
2	G	143	0	122	0	0
2	H	143	0	122	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	127	0	0	0	0
3	B	82	0	0	1	0
3	C	25	0	0	0	0
3	D	80	0	0	0	0
3	E	5	0	0	0	0
3	F	2	0	0	0	0
3	G	1	0	0	0	0
3	H	3	0	0	0	0
All	All	16129	0	15645	23	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 1.

All (23) 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:382:VAL:HG21	1:D:399:VAL:HG21	1.86	0.58
1:D:270:HIS:HB3	1:D:317:ILE:HA	1.92	0.51
1:B:29:ARG:NH2	1:B:33:MET:SD	2.85	0.49
1:D:110:ARG:NE	1:D:110:ARG:HA	2.27	0.48
1:C:270:HIS:HB3	1:C:317:ILE:HA	1.94	0.48
1:C:211:TYR:CE1	1:C:215:MET:HG3	2.49	0.48
1:D:382:VAL:HG12	1:D:416:HIS:HB2	1.95	0.48
1:C:401:ILE:HD12	1:C:453:TYR:CZ	2.49	0.47
1:A:146:LEU:HD23	1:A:153:LEU:HD13	1.98	0.45
1:A:141:LEU:C	1:A:141:LEU:HD23	2.38	0.44
1:C:281:THR:HB	1:C:282:PRO:HD3	2.00	0.44
1:A:146:LEU:CD2	1:A:153:LEU:HD13	2.47	0.44
1:C:382:VAL:HG21	1:C:399:VAL:HG21	1.98	0.44
1:B:141:LEU:C	1:B:141:LEU:HD23	2.39	0.43
1:A:51:LYS:NZ	1:B:47:ASN:OD1	2.52	0.43
1:A:177:VAL:HG11	1:A:201:TRP:CH2	2.54	0.42
1:C:146:LEU:HD23	1:C:153:LEU:HD22	2.00	0.42
1:B:85:VAL:HG13	1:B:91:ILE:HB	2.01	0.42
1:D:219:LEU:HD11	1:D:227:CYS:SG	2.60	0.41
1:B:250:LYS:HB2	3:B:617:HOH:O	2.20	0.41
1:C:246:VAL:HG12	1:C:247:ILE:HG23	2.03	0.41
1:A:281:THR:HB	1:A:282:PRO:HD3	2.03	0.40
1:A:139:TYR:OH	1:A:400:GLU:OE2	2.31	0.40

There are no symmetry-related clashes.

## 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	461/513 (90%)	447 (97%)	12 (3%)	2 (0%)	36	44
1	B	462/513 (90%)	448 (97%)	13 (3%)	1 (0%)	49	61
1	C	457/513 (89%)	440 (96%)	16 (4%)	1 (0%)	49	61
1	D	456/513 (89%)	442 (97%)	13 (3%)	1 (0%)	49	61
2	E	14/16 (88%)	13 (93%)	1 (7%)	0	100	100
2	F	14/16 (88%)	14 (100%)	0	0	100	100
2	G	14/16 (88%)	13 (93%)	1 (7%)	0	100	100
2	H	14/16 (88%)	13 (93%)	1 (7%)	0	100	100
All	All	1892/2116 (89%)	1830 (97%)	57 (3%)	5 (0%)	43	53

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	249	ASP
1	A	249	ASP
1	B	249	ASP
1	A	437	PRO
1	C	437	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	433/480 (90%)	419 (97%)	14 (3%)	42	57

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	434/480 (90%)	424 (98%)	10 (2%)	53	70
1	C	431/480 (90%)	415 (96%)	16 (4%)	37	51
1	D	430/480 (90%)	419 (97%)	11 (3%)	49	66
2	E	15/15 (100%)	15 (100%)	0	100	100
2	F	15/15 (100%)	15 (100%)	0	100	100
2	G	15/15 (100%)	15 (100%)	0	100	100
2	H	15/15 (100%)	14 (93%)	1 (7%)	18	23
All	All	1788/1980 (90%)	1736 (97%)	52 (3%)	45	61

All (52) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	27	SER
1	A	50	LEU
1	A	133	SER
1	A	136	LYS
1	A	169	PRO
1	A	217	LYS
1	A	219	LEU
1	A	249	ASP
1	A	306	ASP
1	A	374	LEU
1	A	398	LEU
1	A	405	LEU
1	A	413	LYS
1	A	459	LEU
1	B	50	LEU
1	B	133	SER
1	B	217	LYS
1	B	219	LEU
1	B	265	LEU
1	B	335	SER
1	B	374	LEU
1	B	398	LEU
1	B	405	LEU
1	B	459	LEU
1	C	2	SER
1	C	50	LEU
1	C	133	SER
1	C	148	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	155	ASN
1	C	159	ASN
1	C	217	LYS
1	C	219	LEU
1	C	238	GLN
1	C	265	LEU
1	C	295	ASP
1	C	374	LEU
1	C	398	LEU
1	C	405	LEU
1	C	413	LYS
1	C	459	LEU
1	D	110	ARG
1	D	126	GLN
1	D	133	SER
1	D	148	LEU
1	D	210	LEU
1	D	219	LEU
1	D	265	LEU
1	D	374	LEU
1	D	398	LEU
1	D	405	LEU
1	D	459	LEU
2	H	34	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	67	GLN
1	B	155	ASN
1	B	396	GLN
1	C	256	GLN
1	C	326	ASN
1	D	41	GLN
1	D	67	GLN
1	D	438	GLN

### 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 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	465/513 (90%)	0.04	8 (1%) 70 76	20, 37, 76, 121	0
1	B	466/513 (90%)	0.07	10 (2%) 63 70	22, 45, 83, 119	0
1	C	463/513 (90%)	-0.02	4 (0%) 84 87	21, 41, 78, 114	0
1	D	462/513 (90%)	0.02	5 (1%) 80 84	22, 40, 76, 109	0
2	E	16/16 (100%)	0.88	3 (18%) 1 1	32, 57, 91, 96	0
2	F	16/16 (100%)	0.54	1 (6%) 20 26	34, 57, 80, 82	0
2	G	16/16 (100%)	0.92	3 (18%) 1 1	36, 58, 92, 96	0
2	H	16/16 (100%)	0.55	0 100 100	27, 59, 78, 80	0
All	All	1920/2116 (90%)	0.05	34 (1%) 68 74	20, 41, 80, 121	0

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	251	GLN	5.2
2	E	22	SER	4.9
1	A	252	THR	4.8
2	G	22	SER	4.7
1	C	181	THR	4.5
1	D	181	THR	4.5
2	E	23	ILE	4.2
1	B	251	GLN	3.7
1	D	257	LEU	3.6
1	D	52	ASP	3.6
2	G	24	ILE	3.5
1	B	439	ASN	3.4
1	D	24	ILE	3.1
2	G	23	ILE	3.1
1	A	181	THR	2.9
1	B	252	THR	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	49	ASN	2.8
1	C	485	ASP	2.8
1	C	51	LYS	2.8
1	C	439	ASN	2.7
2	E	24	ILE	2.6
1	B	438	GLN	2.5
1	B	432	LYS	2.5
1	A	156	TYR	2.4
1	B	181	THR	2.3
1	A	439	ASN	2.3
1	B	107	ARG	2.3
1	A	438	GLN	2.3
1	B	5	LEU	2.3
1	B	9	GLN	2.2
2	F	28	GLU	2.2
1	B	108	HIS	2.1
1	A	160	ARG	2.1
1	A	49	ASN	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.