Full wwPDB X-ray Structure Validation Report ⓘ

Jun 17, 2024 – 10:51 PM EDT

PDB ID	:	5WWL
Title	:	Crystal structure of the Schizogenesis pombe kinetochore Mis12C subcomplex
Authors	:	Wang, C.; Zhou, X.; Wu, M.; Zhang, X.; Zang, J.
Deposited on	:	2017-01-02
Resolution	:	2.40 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:

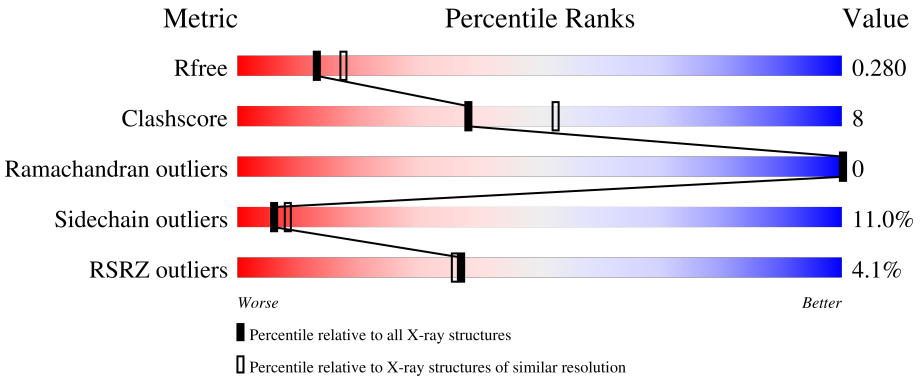
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 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 <=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	M	215	5% 77% 20% ·
2	N	175	3% 69% 17% · 11%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3031 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 Centromere protein mis12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	M	214	Total	C	N	O	S	0	0	0
			1768	1126	302	336	4			

- Molecule 2 is a protein called Kinetochore protein nnf1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	N	155	Total	C	N	O	S	0	0	0
			1263	794	219	242	8			

There are 35 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
N	141	GLU	-	expression tag	UNP Q09858
N	142	GLU	-	expression tag	UNP Q09858
N	143	LYS	-	expression tag	UNP Q09858
N	144	ALA	-	expression tag	UNP Q09858
N	145	THR	-	expression tag	UNP Q09858
N	146	GLU	-	expression tag	UNP Q09858
N	147	ASP	-	expression tag	UNP Q09858
N	148	CYS	-	expression tag	UNP Q09858
N	149	ILE	-	expression tag	UNP Q09858
N	150	SER	-	expression tag	UNP Q09858
N	151	ARG	-	expression tag	UNP Q09858
N	152	MET	-	expression tag	UNP Q09858
N	153	GLN	-	expression tag	UNP Q09858
N	154	SER	-	expression tag	UNP Q09858
N	155	LEU	-	expression tag	UNP Q09858
N	156	ILE	-	expression tag	UNP Q09858
N	157	GLN	-	expression tag	UNP Q09858
N	158	LYS	-	expression tag	UNP Q09858
N	159	LEU	-	expression tag	UNP Q09858
N	160	GLU	-	expression tag	UNP Q09858

Continued on next page...

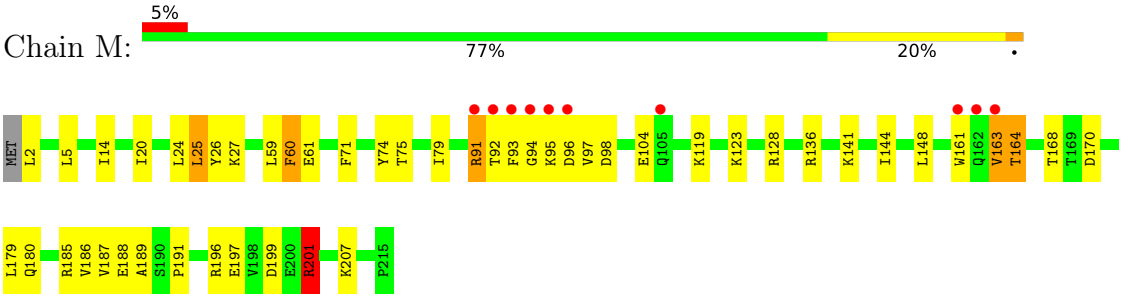
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
N	161	LYS	-	expression tag	UNP Q09858
N	162	THR	-	expression tag	UNP Q09858
N	163	VAL	-	expression tag	UNP Q09858
N	164	TYR	-	expression tag	UNP Q09858
N	165	GLY	-	expression tag	UNP Q09858
N	166	MET	-	expression tag	UNP Q09858
N	167	ASN	-	expression tag	UNP Q09858
N	168	GLU	-	expression tag	UNP Q09858
N	169	LYS	-	expression tag	UNP Q09858
N	170	ASN	-	expression tag	UNP Q09858
N	171	LEU	-	expression tag	UNP Q09858
N	172	ALA	-	expression tag	UNP Q09858
N	173	LYS	-	expression tag	UNP Q09858
N	174	GLU	-	expression tag	UNP Q09858
N	175	ALA	-	expression tag	UNP Q09858

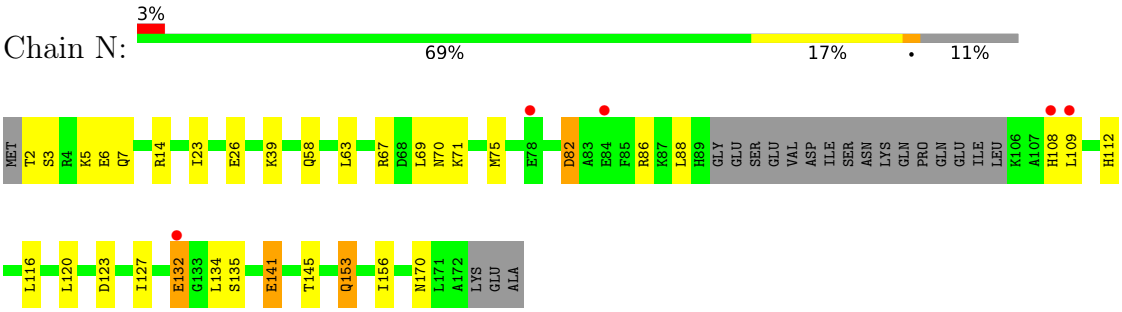
3 Residue-property plots (i)

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: Centromere protein mis12



• Molecule 2: Kinetochore protein nnf1



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	41.50Å 85.61Å 123.79Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.01 – 2.40 70.41 – 2.40	Depositor EDS
% Data completeness (in resolution range)	95.9 (50.01-2.40) 95.9 (70.41-2.40)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.16 (at 2.40Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, R_{free}	0.234 , 0.272 0.245 , 0.280	Depositor DCC
R_{free} test set	848 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å^2)	42.0	Xtriage
Anisotropy	0.214	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 40.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	3031	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.21% 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)

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	M	0.73	0/1801	1.05	14/2430 (0.6%)
2	N	0.73	0/1278	0.86	0/1706
All	All	0.73	0/3079	0.97	14/4136 (0.3%)

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	M	0	1

There are no bond length outliers.

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	M	163	VAL	N-CA-C	11.31	141.53	111.00
1	M	196	ARG	NE-CZ-NH1	8.52	124.56	120.30
1	M	163	VAL	CB-CA-C	-7.23	97.66	111.40
1	M	196	ARG	NE-CZ-NH2	-6.33	117.13	120.30
1	M	201	ARG	NE-CZ-NH1	6.24	123.42	120.30
1	M	161	TRP	N-CA-C	6.08	127.41	111.00
1	M	191	PRO	N-CA-C	5.98	127.64	112.10
1	M	201	ARG	NE-CZ-NH2	-5.95	117.32	120.30
1	M	164	THR	N-CA-C	5.63	126.21	111.00
1	M	25	LEU	CB-CG-CD1	5.47	120.30	111.00
1	M	25	LEU	CA-CB-CG	5.46	127.86	115.30
1	M	199	ASP	CB-CG-OD1	5.40	123.16	118.30
1	M	60	PHE	CB-CG-CD1	5.28	124.50	120.80
1	M	128	ARG	NE-CZ-NH1	5.10	122.85	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	M	189	ALA	Mainchain

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	M	1768	0	1770	35	0
2	N	1263	0	1285	24	0
All	All	3031	0	3055	46	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 (46) 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:M:163:VAL:CG1	1:M:163:VAL:O	1.76	1.24
1:M:91:ARG:NH2	2:N:82:ASP:HB3	1.58	1.19
1:M:91:ARG:H	1:M:91:ARG:HD2	1.23	1.02
1:M:91:ARG:CZ	2:N:82:ASP:HB3	1.91	0.99
2:N:108:HIS:O	2:N:112:HIS:HD2	1.45	0.99
2:N:108:HIS:O	2:N:112:HIS:CD2	2.17	0.97
1:M:163:VAL:O	1:M:163:VAL:HG12	1.13	0.93
1:M:91:ARG:HH21	2:N:82:ASP:HB3	1.51	0.75
1:M:91:ARG:NH2	2:N:82:ASP:CB	2.45	0.73
1:M:91:ARG:H	1:M:91:ARG:CD	1.97	0.72
1:M:26:TYR:OH	1:M:61:GLU:OE2	2.05	0.67
1:M:20:ILE:HD12	2:N:63:LEU:HD11	1.80	0.64
1:M:163:VAL:O	1:M:163:VAL:HG13	1.94	0.63
2:N:23:ILE:O	2:N:23:ILE:HG13	1.99	0.63
1:M:185:ARG:NH1	1:M:197:GLU:OE2	2.33	0.61
1:M:91:ARG:C	1:M:93:PHE:H	2.08	0.57
1:M:144:ILE:HD11	1:M:186:VAL:HG11	1.88	0.56
1:M:123:LYS:HB3	2:N:127:ILE:HD13	1.88	0.55
1:M:74:TYR:CE1	1:M:79:ILE:CD1	2.90	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:M:91:ARG:HA	2:N:86:ARG:NH1	2.24	0.52
2:N:2:THR:HG23	2:N:7:GLN:HG3	1.91	0.52
2:N:141:GLU:O	2:N:145:THR:HG23	2.11	0.51
1:M:94:GLY:HA3	1:M:97:VAL:HG12	1.92	0.50
1:M:91:ARG:NE	2:N:82:ASP:HB3	2.26	0.50
1:M:180:GLN:HG2	2:N:156:ILE:HG21	1.95	0.49
1:M:91:ARG:HD2	1:M:91:ARG:N	2.08	0.47
1:M:123:LYS:CB	2:N:127:ILE:HD13	2.45	0.47
1:M:92:THR:O	1:M:93:PHE:HB2	2.17	0.45
1:M:141:LYS:HA	2:N:145:THR:HG22	1.96	0.45
2:N:3:SER:OG	2:N:6:GLU:HB2	2.16	0.45
2:N:71:LYS:O	2:N:75:MET:HG3	2.15	0.45
2:N:132:GLU:HA	2:N:132:GLU:OE2	2.17	0.45
1:M:20:ILE:CD1	2:N:63:LEU:HD11	2.47	0.44
2:N:153:GLN:HE21	2:N:153:GLN:HA	1.82	0.44
1:M:91:ARG:HH21	2:N:82:ASP:CB	2.22	0.44
1:M:91:ARG:HB2	1:M:94:GLY:H	1.82	0.44
1:M:2:LEU:O	1:M:5:LEU:HB3	2.18	0.43
2:N:5:LYS:HB2	2:N:70:ASN:OD1	2.18	0.43
1:M:91:ARG:CD	1:M:91:ARG:N	2.74	0.43
1:M:71:PHE:O	1:M:75:THR:HG23	2.20	0.42
2:N:123:ASP:O	2:N:127:ILE:HG13	2.20	0.42
1:M:14:ILE:HD11	1:M:75:THR:OG1	2.19	0.42
1:M:197:GLU:O	1:M:201:ARG:HG3	2.20	0.41
1:M:119:LYS:O	1:M:123:LYS:HG3	2.20	0.41
1:M:91:ARG:C	1:M:93:PHE:N	2.73	0.40
1:M:91:ARG:H	1:M:91:ARG:HH11	1.69	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	M	212/215 (99%)	200 (94%)	12 (6%)	0	100	100
2	N	151/175 (86%)	148 (98%)	3 (2%)	0	100	100
All	All	363/390 (93%)	348 (96%)	15 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	M	196/197 (100%)	176 (90%)	20 (10%)	7	10
2	N	141/159 (89%)	124 (88%)	17 (12%)	5	6
All	All	337/356 (95%)	300 (89%)	37 (11%)	6	8

All (37) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	24	LEU
1	M	25	LEU
1	M	27	LYS
1	M	59	LEU
1	M	60	PHE
1	M	91	ARG
1	M	95	LYS
1	M	96	ASP
1	M	98	ASP
1	M	104	GLU
1	M	136	ARG
1	M	148	LEU
1	M	164	THR
1	M	168	THR
1	M	170	ASP
1	M	179	LEU
1	M	187	VAL
1	M	188	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	M	201	ARG
1	M	207	LYS
2	N	14	ARG
2	N	26	GLU
2	N	39	LYS
2	N	58	GLN
2	N	67	ARG
2	N	69	LEU
2	N	82	ASP
2	N	88	LEU
2	N	109	LEU
2	N	116	LEU
2	N	120	LEU
2	N	132	GLU
2	N	134	LEU
2	N	135	SER
2	N	141	GLU
2	N	153	GLN
2	N	170	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
2	N	58	GLN
2	N	89	HIS
2	N	112	HIS
2	N	153	GLN
2	N	167	ASN
2	N	170	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 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

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^{th} 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	M	214/215 (99%)	0.29	10 (4%)	31	30	25, 47, 92, 137	2 (0%)
2	N	155/175 (88%)	0.15	5 (3%)	47	46	29, 50, 76, 98	0
All	All	369/390 (94%)	0.23	15 (4%)	37	36	25, 48, 83, 137	2 (0%)

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	M	161	TRP	11.7
1	M	93	PHE	6.6
1	M	162	GLN	6.3
1	M	163	VAL	5.5
1	M	92	THR	3.9
1	M	94	GLY	3.6
1	M	91	ARG	3.5
1	M	95	LYS	2.9
2	N	109	LEU	2.7
1	M	96	ASP	2.6
2	N	78	GLU	2.6
1	M	105	GLN	2.5
2	N	132	GLU	2.4
2	N	84	GLU	2.1
2	N	108	HIS	2.0

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.