Full wwPDB X-ray Structure Validation Report ⓘ

Jun 16, 2024 – 10:56 AM EDT

PDB ID	:	4XM4
Title	:	Structure of Chaetomium Mex67:Mtr2 subunits
Authors	:	Aibara, S.; Valkov, M.; Stewart, M.
Deposited on	:	2015-01-14
Resolution	:	2.95 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:

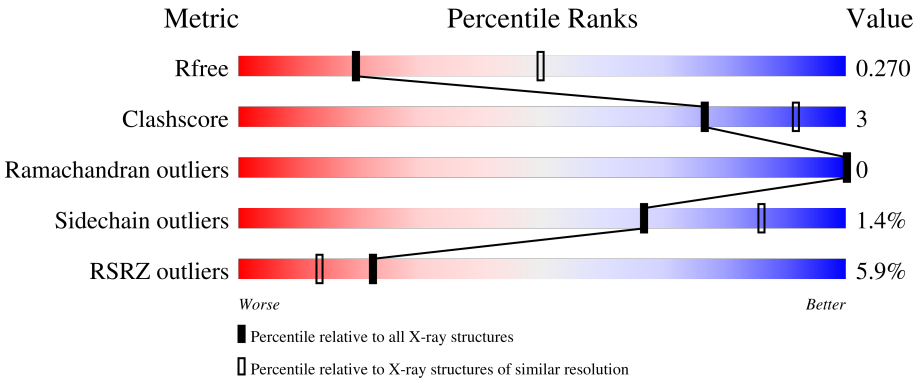
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.95 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	3104 (3.00-2.92)
Clashscore	141614	3462 (3.00-2.92)
Ramachandran outliers	138981	3340 (3.00-2.92)
Sidechain outliers	138945	3343 (3.00-2.92)
RSRZ outliers	127900	2986 (3.00-2.92)

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	A	386	% 43% · 54%
1	B	386	% 44% · 53%
1	C	386	3% 37% · 62%
1	D	386	5% 36% 5% 59%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 10475 atoms, of which 5201 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 Mex67.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	176	Total	C	H	N	O	S	0	0	0
			2745	891	1349	240	258	7			
1	B	183	Total	C	H	N	O	S	0	0	0
			2834	920	1389	249	269	7			
1	C	147	Total	C	H	N	O	S	0	0	0
			2369	746	1190	206	225	2			
1	D	157	Total	C	H	N	O	S	0	0	0
			2527	789	1273	219	244	2			

There are 4 discrepancies between the modelled and reference sequences:

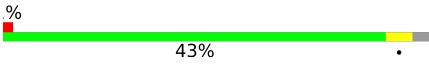
Chain	Residue	Modelled	Actual	Comment	Reference
A	179	SER	-	expression tag	UNP G0SET4
B	179	SER	-	expression tag	UNP G0SET4
C	179	SER	-	expression tag	UNP G0SET4
D	179	SER	-	expression tag	UNP G0SET4

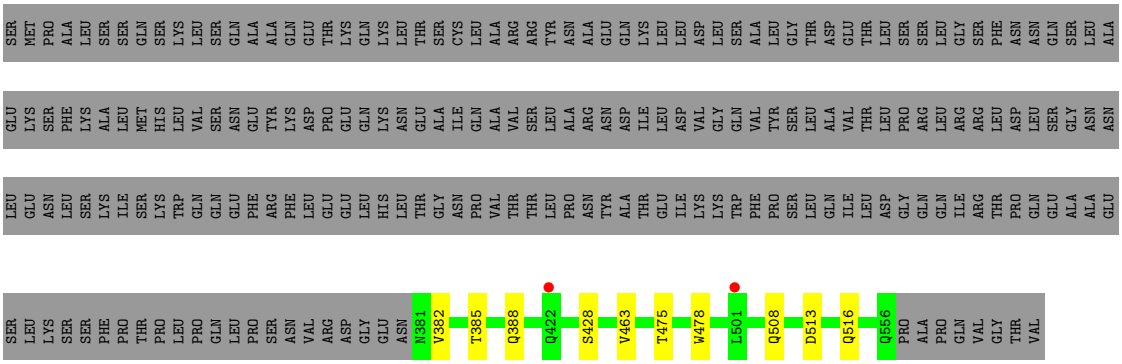
3 Residue-property plots (i)

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: Mex67

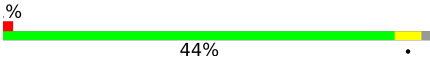
Chain A:

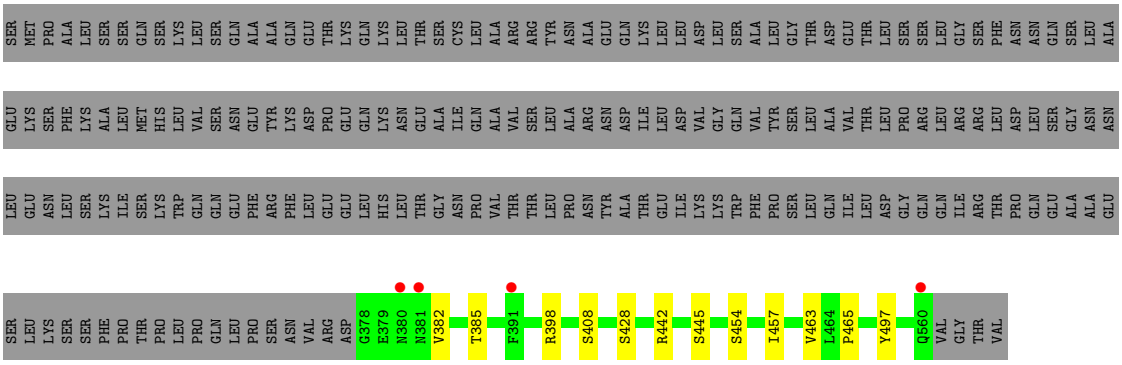




- Molecule 1: Mex67

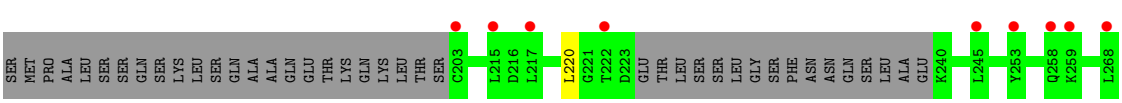
Chain B:

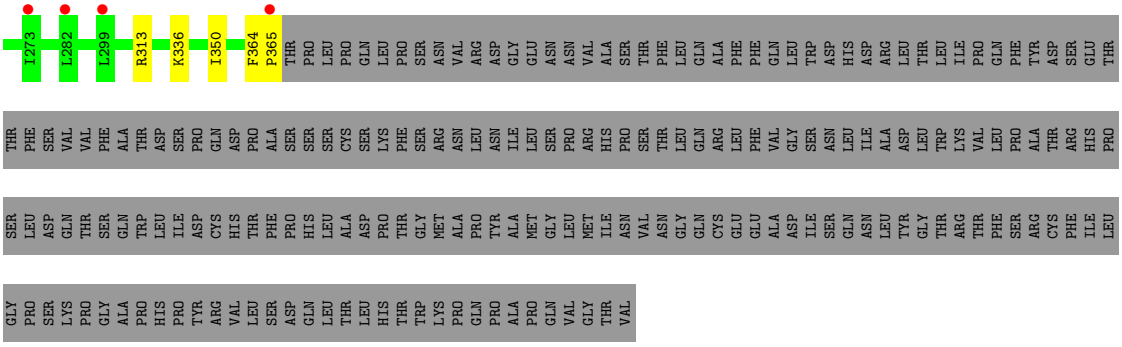




- Molecule 1: Mex67

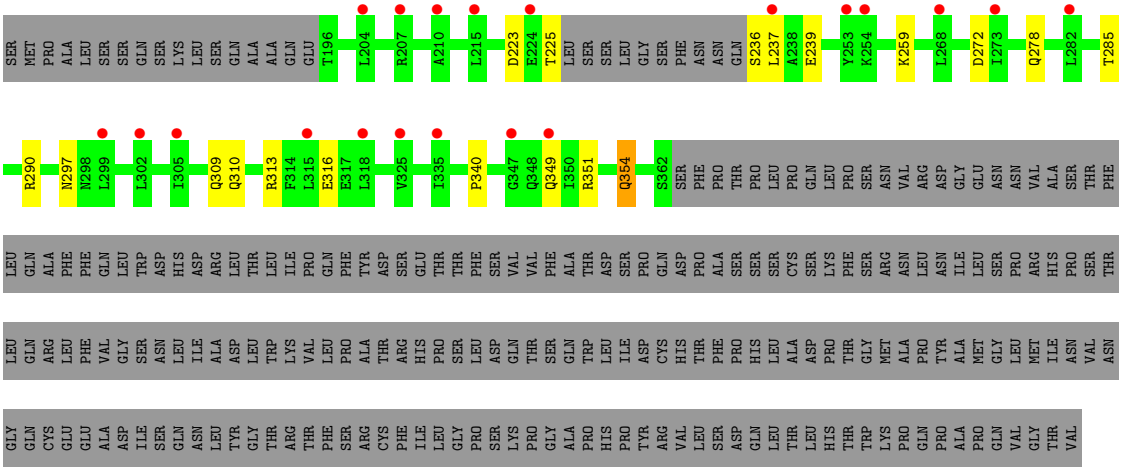
Chain C:





• Molecule 1: Mex67

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	43.75Å 96.09Å 194.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.04 – 2.95 48.04 – 2.95	Depositor EDS
% Data completeness (in resolution range)	98.9 (48.04-2.95) 99.0 (48.04-2.95)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.54 (at 2.96Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.227 , 0.265 0.232 , 0.270	Depositor DCC
R_{free} test set	924 reflections (5.17%)	wwPDB-VP
Wilson B-factor (Å^2)	69.0	Xtriage
Anisotropy	0.423	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 56.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	10475	wwPDB-VP
Average B, all atoms (Å^2)	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.25% 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

5.1 Standard geometry

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	0.24	0/1440	0.42	0/1969
1	B	0.23	0/1491	0.41	0/2040
1	C	0.22	0/1199	0.39	0/1621
1	D	0.22	0/1272	0.39	0/1718
All	All	0.23	0/5402	0.40	0/7348

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

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	1396	1349	1347	8	0
1	B	1445	1389	1389	7	0
1	C	1179	1190	1190	2	0
1	D	1254	1273	1273	17	0
All	All	5274	5201	5199	31	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 3.

All (31) 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:236:SER:O	1:D:237:LEU:HG	1.57	1.04
1:B:382:VAL:O	1:B:385:THR:HG22	1.64	0.95
1:A:382:VAL:O	1:A:385:THR:HG22	1.67	0.95
1:D:237:LEU:HD12	1:D:237:LEU:O	1.70	0.90
1:D:237:LEU:HD13	1:D:278:GLN:OE1	1.78	0.83
1:A:475:THR:O	1:B:442:ARG:NH2	2.26	0.69
1:D:236:SER:O	1:D:237:LEU:CG	2.41	0.65
1:D:237:LEU:O	1:D:237:LEU:CD1	2.43	0.64
1:A:508:GLN:OE1	1:B:445:SER:OG	2.15	0.64
1:D:236:SER:C	1:D:237:LEU:HG	2.18	0.62
1:C:336:LYS:NZ	1:C:350:ILE:O	2.33	0.61
1:B:398:ARG:NH2	1:B:465:PRO:O	2.34	0.60
1:D:239:GLU:HG2	1:D:278:GLN:HA	1.84	0.60
1:A:382:VAL:C	1:A:385:THR:HG22	2.28	0.54
1:D:259:LYS:NZ	1:D:285:THR:O	2.38	0.54
1:A:478:TRP:O	1:B:442:ARG:NH2	2.42	0.52
1:D:340:PRO:O	1:D:351:ARG:NH1	2.46	0.48
1:D:223:ASP:HB3	1:D:225:THR:HG23	1.95	0.47
1:D:290:ARG:NH1	1:D:316:GLU:OE2	2.47	0.47
1:D:354:GLN:OE1	1:D:354:GLN:N	2.48	0.47
1:D:310:GLN:O	1:D:313:ARG:NH1	2.47	0.46
1:D:272:ASP:N	1:D:297:ASN:OD1	2.49	0.45
1:B:428:SER:HB3	1:B:463:VAL:HG12	1.98	0.45
1:D:237:LEU:HD12	1:D:237:LEU:C	2.35	0.45
1:B:454:SER:HA	1:B:457:ILE:CD1	2.47	0.45
1:A:382:VAL:O	1:A:385:THR:CG2	2.52	0.44
1:D:237:LEU:HD22	1:D:272:ASP:O	2.18	0.44
1:C:364:PHE:N	1:C:365:PRO:HD2	2.35	0.42
1:A:428:SER:HB3	1:A:463:VAL:HG12	2.02	0.42
1:D:237:LEU:HD13	1:D:278:GLN:CD	2.40	0.40
1:A:513:ASP:OD2	1:A:516:GLN:NE2	2.53	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	174/386 (45%)	171 (98%)	3 (2%)	0	100	100
1	B	181/386 (47%)	178 (98%)	3 (2%)	0	100	100
1	C	143/386 (37%)	140 (98%)	3 (2%)	0	100	100
1	D	153/386 (40%)	152 (99%)	1 (1%)	0	100	100
All	All	651/1544 (42%)	641 (98%)	10 (2%)	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	A	158/344 (46%)	157 (99%)	1 (1%)	86	94
1	B	163/344 (47%)	161 (99%)	2 (1%)	71	88
1	C	131/344 (38%)	129 (98%)	2 (2%)	65	85
1	D	140/344 (41%)	137 (98%)	3 (2%)	53	80
All	All	592/1376 (43%)	584 (99%)	8 (1%)	67	86

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	388	GLN
1	B	408	SER
1	B	497	TYR
1	C	220	LEU
1	C	313	ARG
1	D	309	GLN
1	D	349	GLN
1	D	354	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are

no such sidechains identified.

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	A	176/386 (45%)	0.29	2 (1%) 80 65	30, 57, 85, 116	0
1	B	183/386 (47%)	0.46	4 (2%) 62 45	36, 62, 95, 130	0
1	C	147/386 (38%)	0.58	13 (8%) 10 5	55, 95, 125, 136	0
1	D	157/386 (40%)	0.70	20 (12%) 3 2	72, 92, 115, 130	0
All	All	663/1544 (42%)	0.50	39 (5%) 22 13	30, 76, 114, 136	0

All (39) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	254	LYS	4.1
1	D	305	ILE	4.1
1	B	560	GLN	3.4
1	C	273	ILE	3.3
1	C	268	LEU	3.3
1	D	204	LEU	3.1
1	B	380	ASN	3.0
1	C	258	GLN	3.0
1	C	259	LYS	2.8
1	D	315	LEU	2.8
1	D	302	LEU	2.7
1	D	335	ILE	2.7
1	C	203	CYS	2.7
1	C	299	LEU	2.7
1	D	273	ILE	2.7
1	D	253	TYR	2.6
1	D	268	LEU	2.5
1	D	318	LEU	2.5
1	B	381	ASN	2.5
1	D	207	ARG	2.4
1	C	365	PRO	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	245	LEU	2.4
1	D	210	ALA	2.4
1	D	237	LEU	2.4
1	C	217	LEU	2.4
1	C	222	THR	2.3
1	D	325	VAL	2.3
1	A	422	GLN	2.3
1	D	299	LEU	2.2
1	D	282	LEU	2.2
1	C	282	LEU	2.2
1	D	347	GLY	2.2
1	D	349	GLN	2.2
1	C	215	LEU	2.1
1	C	253	TYR	2.1
1	A	501	LEU	2.0
1	B	391	PHE	2.0
1	D	224	GLU	2.0
1	D	215	LEU	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.