wwPDB/EMDataBank EM Map/Model Validation Summary Report (i)

Sep 9, 2019 – 09:33 AM EDT

PDB ID	:	6OF4
EMDB ID:	:	EMD-20042
Title	:	Precursor ribosomal RNA processing complex, apo-state.
Authors	:	Pillon, M.C.; Hsu, A.L.; Krahn, J.M.; Williams, J.G.; Goslen, K.H.; Sobhany, M.; Borgnia, M.J.; Stanley, R.E.
Deposited on	:	2019-03-28
Resolution	:	3.20 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

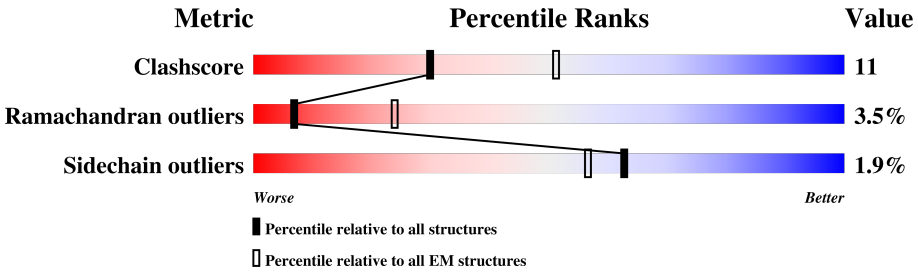
MolProbity	:	4.02b-467
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.4

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.20 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)	EM structures (#Entries)
Clashscore	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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%

Mol	Chain	Length	Quality of chain
1	A	391	42% 5% • 53%
1	D	391	41% 5% • 53%
2	B	640	60% 20% • 18%
2	E	640	59% 21% • 18%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 10998 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Ribonuclease.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	183	Total	C	N	O	S	0	0
			1451	928	258	259	6		
1	D	183	Total	C	N	O	S	0	0
			1451	928	258	259	6		

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-27	MET	-	initiating methionine	UNP G0SGE9
A	-26	SER	-	expression tag	UNP G0SGE9
A	-25	TYR	-	expression tag	UNP G0SGE9
A	-24	TYR	-	expression tag	UNP G0SGE9
A	-23	HIS	-	expression tag	UNP G0SGE9
A	-22	HIS	-	expression tag	UNP G0SGE9
A	-21	HIS	-	expression tag	UNP G0SGE9
A	-20	HIS	-	expression tag	UNP G0SGE9
A	-19	HIS	-	expression tag	UNP G0SGE9
A	-18	HIS	-	expression tag	UNP G0SGE9
A	-17	ASP	-	expression tag	UNP G0SGE9
A	-16	TYR	-	expression tag	UNP G0SGE9
A	-15	ASP	-	expression tag	UNP G0SGE9
A	-14	ILE	-	expression tag	UNP G0SGE9
A	-13	PRO	-	expression tag	UNP G0SGE9
A	-12	THR	-	expression tag	UNP G0SGE9
A	-11	THR	-	expression tag	UNP G0SGE9
A	-10	GLU	-	expression tag	UNP G0SGE9
A	-9	ASN	-	expression tag	UNP G0SGE9
A	-8	LEU	-	expression tag	UNP G0SGE9
A	-7	TYR	-	expression tag	UNP G0SGE9
A	-6	PHE	-	expression tag	UNP G0SGE9
A	-5	GLN	-	expression tag	UNP G0SGE9
A	-4	GLY	-	expression tag	UNP G0SGE9
A	-3	ALA	-	expression tag	UNP G0SGE9
A	-2	MET	-	expression tag	UNP G0SGE9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP G0SGE9
A	0	SER	-	expression tag	UNP G0SGE9
D	-27	MET	-	initiating methionine	UNP G0SGE9
D	-26	SER	-	expression tag	UNP G0SGE9
D	-25	TYR	-	expression tag	UNP G0SGE9
D	-24	TYR	-	expression tag	UNP G0SGE9
D	-23	HIS	-	expression tag	UNP G0SGE9
D	-22	HIS	-	expression tag	UNP G0SGE9
D	-21	HIS	-	expression tag	UNP G0SGE9
D	-20	HIS	-	expression tag	UNP G0SGE9
D	-19	HIS	-	expression tag	UNP G0SGE9
D	-18	HIS	-	expression tag	UNP G0SGE9
D	-17	ASP	-	expression tag	UNP G0SGE9
D	-16	TYR	-	expression tag	UNP G0SGE9
D	-15	ASP	-	expression tag	UNP G0SGE9
D	-14	ILE	-	expression tag	UNP G0SGE9
D	-13	PRO	-	expression tag	UNP G0SGE9
D	-12	THR	-	expression tag	UNP G0SGE9
D	-11	THR	-	expression tag	UNP G0SGE9
D	-10	GLU	-	expression tag	UNP G0SGE9
D	-9	ASN	-	expression tag	UNP G0SGE9
D	-8	LEU	-	expression tag	UNP G0SGE9
D	-7	TYR	-	expression tag	UNP G0SGE9
D	-6	PHE	-	expression tag	UNP G0SGE9
D	-5	GLN	-	expression tag	UNP G0SGE9
D	-4	GLY	-	expression tag	UNP G0SGE9
D	-3	ALA	-	expression tag	UNP G0SGE9
D	-2	MET	-	expression tag	UNP G0SGE9
D	-1	GLY	-	expression tag	UNP G0SGE9
D	0	SER	-	expression tag	UNP G0SGE9

- Molecule 2 is a protein called CLP1_P domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	528	Total	C	N	O	S	0	0
			4048	2586	702	739	21		
2	E	528	Total	C	N	O	S	0	0
			4048	2586	702	739	21		

There are 2 discrepancies between the modelled and reference sequences:

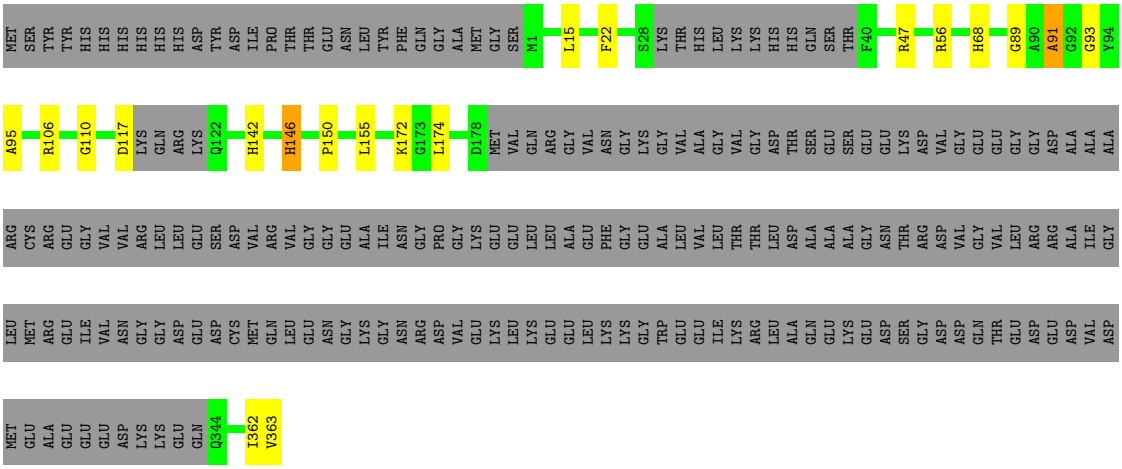
Chain	Residue	Modelled	Actual	Comment	Reference
B	109	MET	-	initiating methionine	UNP G0S263
E	109	MET	-	initiating methionine	UNP G0S263

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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. 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. 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.

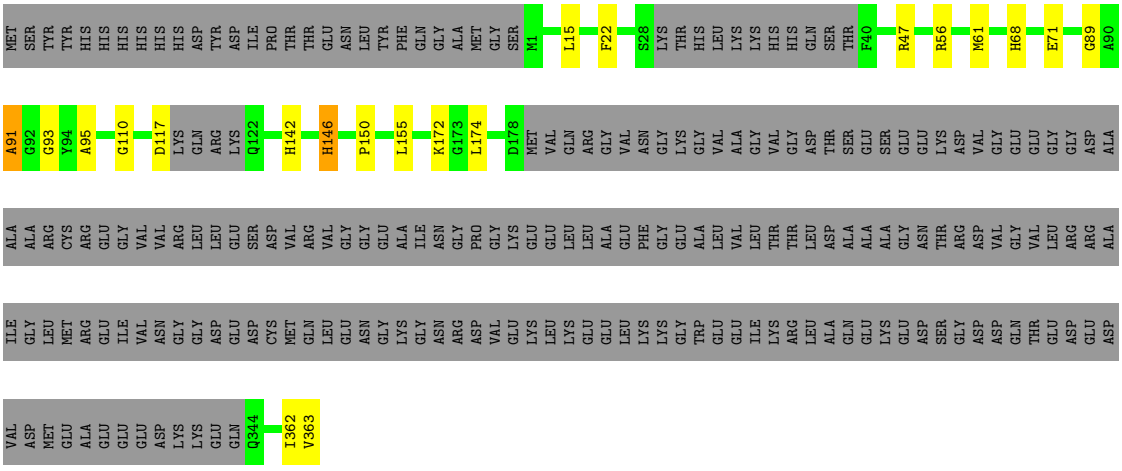
• Molecule 1: Ribonuclease

Chain A:



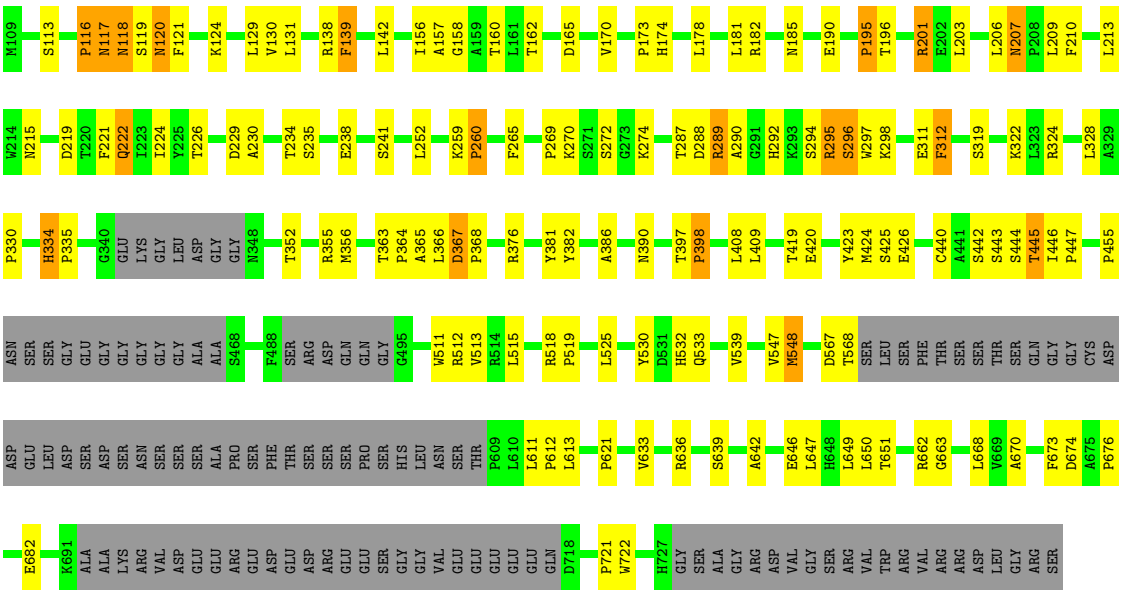
• Molecule 1: Ribonuclease

Chain D:



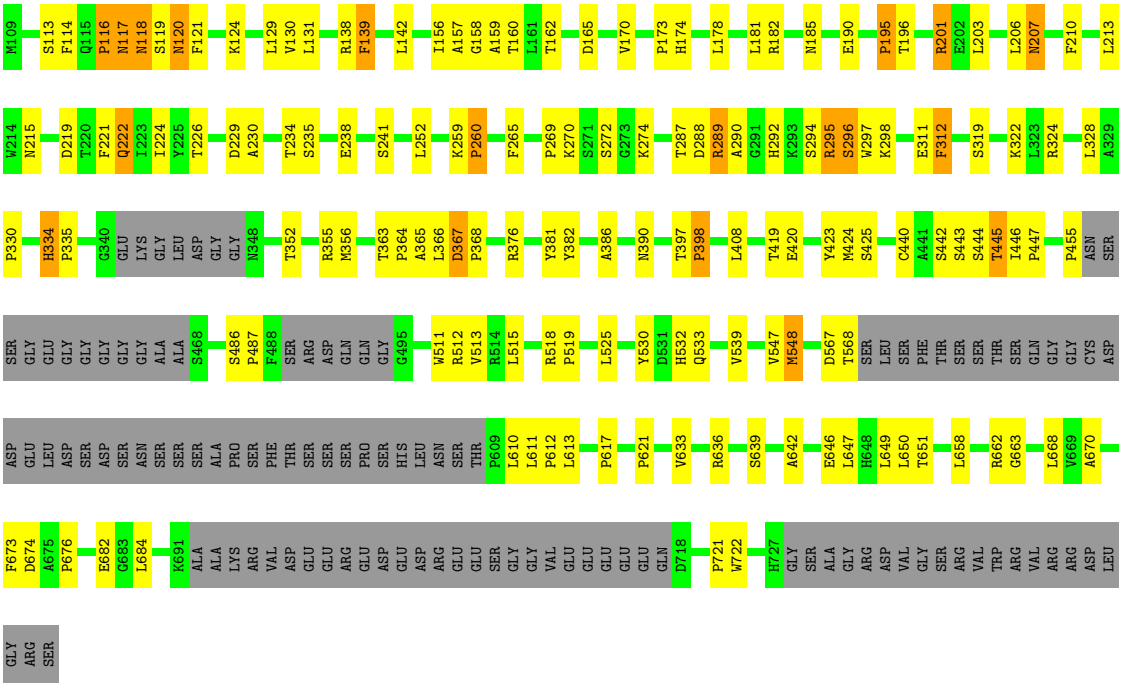
• Molecule 2: CLP1_P domain-containing protein

Chain B:



• Molecule 2: CLP1_P domain-containing protein

Chain E:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	102753	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	40	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	FEI FALCON III (4k x 4k)	Depositor

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	$\#\|Z\| > 2$	RMSZ	$\#\|Z\| > 2$
1	A	0.60	0/1490	0.57	1/2021 (0.0%)
1	D	0.60	0/1490	0.57	1/2021 (0.0%)
2	B	0.50	0/4158	0.63	1/5683 (0.0%)
2	E	0.50	0/4158	0.63	1/5683 (0.0%)
All	All	0.53	0/11296	0.62	4/15408 (0.0%)

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	A	0	1
1	D	0	1
2	B	0	2
2	E	0	2
All	All	0	6

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	146	HIS	C-N-CA	-6.07	106.53	121.70
1	D	146	HIS	C-N-CA	-6.07	106.54	121.70
2	E	455	PRO	N-CA-CB	6.01	110.52	103.30
2	B	455	PRO	N-CA-CB	5.98	110.48	103.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	91	ALA	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
2	B	295	ARG	Peptide
2	B	642	ALA	Peptide
1	D	91	ALA	Peptide
2	E	295	ARG	Peptide

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	1451	0	1402	15	0
1	D	1451	0	1402	16	0
2	B	4048	0	4025	103	0
2	E	4048	0	4025	108	0
All	All	10998	0	10854	238	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 11.

The worst 5 of 238 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)
2:B:419:THR:HG22	2:B:420:GLU:HG3	1.54	0.89
2:E:419:THR:HG22	2:E:420:GLU:HG3	1.54	0.86
1:A:363:VAL:HG11	2:B:639:SER:HB3	1.58	0.84
2:B:130:VAL:HG22	2:B:190:GLU:HG2	1.63	0.80
2:E:130:VAL:HG22	2:E:190:GLU:HG2	1.63	0.80

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 PDB entries followed by that with respect to all EM entries.

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	175/391 (45%)	149 (85%)	26 (15%)	0	100	100
1	D	175/391 (45%)	149 (85%)	26 (15%)	0	100	100
2	B	516/640 (81%)	382 (74%)	110 (21%)	24 (5%)	2	19
2	E	516/640 (81%)	382 (74%)	109 (21%)	25 (5%)	2	19
All	All	1382/2062 (67%)	1062 (77%)	271 (20%)	49 (4%)	7	27

5 of 49 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	118	ASN
2	B	157	ALA
2	B	398	PRO
2	B	662	ARG
2	E	118	ASN

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 PDB entries followed by that with respect to all EM entries.

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	144/316 (46%)	142 (99%)	2 (1%)	69	88
1	D	144/316 (46%)	142 (99%)	2 (1%)	69	88
2	B	446/542 (82%)	437 (98%)	9 (2%)	58	83
2	E	446/542 (82%)	437 (98%)	9 (2%)	58	83
All	All	1180/1716 (69%)	1158 (98%)	22 (2%)	63	84

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	376	ARG
1	D	117	ASP
2	E	376	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	548	MET
1	D	22	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	327	ASN
1	D	146	HIS
2	E	207	ASN
2	B	222	GLN
2	E	144	ASN

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 carbohydrates 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.