



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Feb 19, 2018 – 08:39 PM EST

PDB ID : 6BM0
EMDB ID: : EMD-7114
Title : Cryo-EM structure of human CPSF-160-WDR33 complex at 3.8 Å resolution
Authors : Sun, Y.; Zhang, Y.; Hamilton, K.; Walz, T.; Tong, L.
Deposited on : 2017-11-12
Resolution : 3.80 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
<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20030736

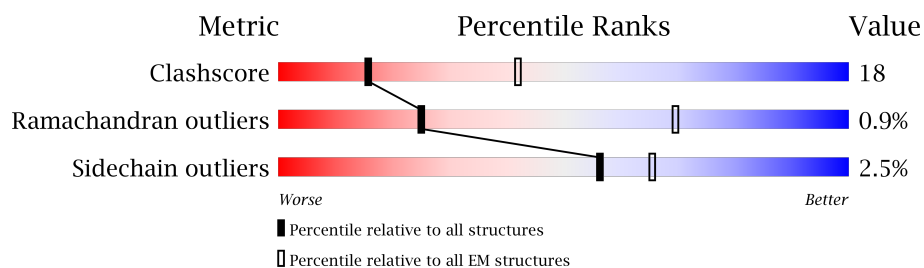
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.80 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	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	1443	
2	B	587	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 12244 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 Cleavage and polyadenylation specificity factor subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1175	Total	C	N	O	S	0	0
			9318	5991	1595	1677	55		

- Molecule 2 is a protein called pre-mRNA 3' end processing protein WDR33.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	363	Total	C	N	O	S	0	0
			2926	1850	532	525	19		

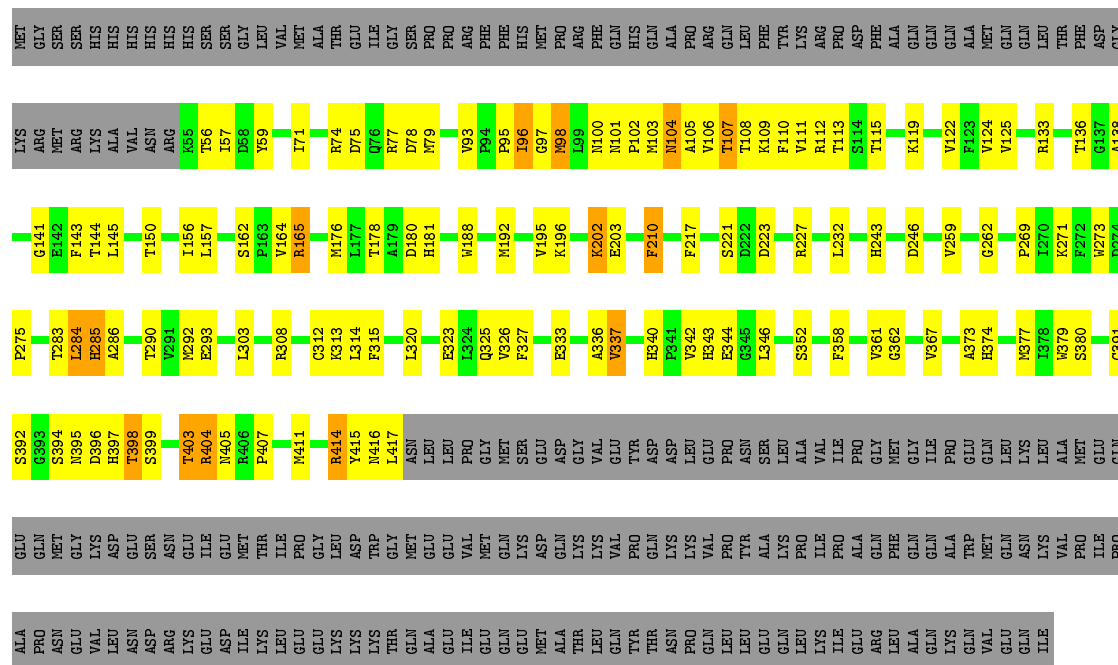
There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-14	MET	-	expression tag	UNP Q9C0J8
B	-13	GLY	-	expression tag	UNP Q9C0J8
B	-12	SER	-	expression tag	UNP Q9C0J8
B	-11	SER	-	expression tag	UNP Q9C0J8
B	-10	HIS	-	expression tag	UNP Q9C0J8
B	-9	HIS	-	expression tag	UNP Q9C0J8
B	-8	HIS	-	expression tag	UNP Q9C0J8
B	-7	HIS	-	expression tag	UNP Q9C0J8
B	-6	HIS	-	expression tag	UNP Q9C0J8
B	-5	HIS	-	expression tag	UNP Q9C0J8
B	-4	SER	-	expression tag	UNP Q9C0J8
B	-3	SER	-	expression tag	UNP Q9C0J8
B	-2	GLY	-	expression tag	UNP Q9C0J8
B	-1	LEU	-	expression tag	UNP Q9C0J8
B	0	VAL	-	expression tag	UNP Q9C0J8



- Molecule 2: pre-mRNA 3' end processing protein WDR33

Chain B:  41% 19% 1% 38%



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	205373	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	47	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	38462	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

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 >2	RMSZ	# Z >2
1	A	0.35	0/9534	0.58	0/12942
2	B	0.37	0/3010	0.61	0/4080
All	All	0.36	0/12544	0.58	0/17022

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	9318	0	9366	285	0
2	B	2926	0	2832	168	0
All	All	12244	0	12198	446	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 18.

The worst 5 of 446 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:A:628:LEU:HD11	1:A:682:LEU:CD2	1.26	1.61
1:A:318:VAL:HG21	1:A:362:PHE:CD2	1.58	1.38

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:626:ILE:HG12	1:A:668:PHE:CE2	1.58	1.36
1:A:617:TYR:HE2	1:A:680:HIS:ND1	1.33	1.26
1:A:816:ARG:HH21	1:A:1410:LEU:CD2	1.52	1.22

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	1149/1443 (80%)	1045 (91%)	95 (8%)	9 (1%)	22 65
2	B	361/587 (62%)	316 (88%)	40 (11%)	5 (1%)	13 54
All	All	1510/2030 (74%)	1361 (90%)	135 (9%)	14 (1%)	25 62

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	378	GLU
1	A	379	PRO
1	A	654	PRO
1	A	812	PRO
1	A	813	VAL

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	1028/1235 (83%)	1005 (98%)	23 (2%)	57	81
2	B	315/514 (61%)	304 (96%)	11 (4%)	41	73
All	All	1343/1749 (77%)	1309 (98%)	34 (2%)	56	79

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	682	LEU
1	A	1220	ASN
2	B	403	THR
1	A	1191	PHE
1	A	318	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1095	GLN
1	A	1188	GLN
2	B	325	GLN
1	A	900	ASN
2	B	285	HIS

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.