



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Dec 3, 2019 – 08:47 PM EST

PDB ID : 6FE8
EMDB ID: : EMD-4241
Title : Cryo-EM structure of the core Centromere Binding Factor 3 complex
Authors : Zhang, W.J.; Lukoynova, N.; Miah, S.; Vaughan, C.K.
Deposited on : 2017-12-30
Resolution : 3.70 Å(reported)
Based on PDB ID : 2VEQ, 1NEX

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 ⓘ 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

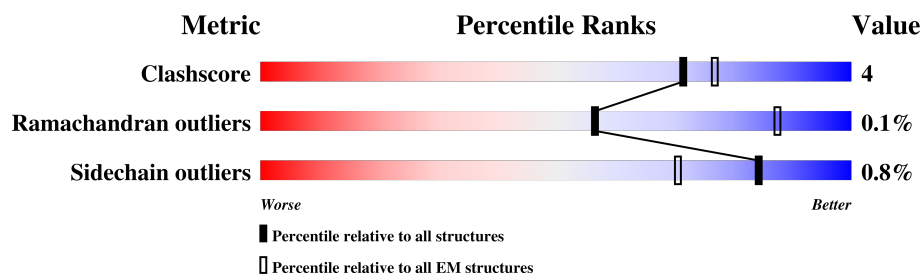
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.70 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	584	
1	B	584	
2	C	197	
3	D	519	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 12479 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 Centromere DNA-binding protein complex CBF3 subunit B.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	524	Total	C	N	O	S	0	0
			4320	2812	697	790	21		
1	B	512	Total	C	N	O	S	0	0
			4249	2766	685	777	21		

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	25	MET	-	initiating methionine	UNP P40969
A	26	GLY	-	expression tag	UNP P40969
A	27	GLY	-	expression tag	UNP P40969
A	28	SER	-	expression tag	UNP P40969
A	29	SER	-	expression tag	UNP P40969
A	30	HIS	-	expression tag	UNP P40969
A	31	HIS	-	expression tag	UNP P40969
A	32	HIS	-	expression tag	UNP P40969
A	33	HIS	-	expression tag	UNP P40969
A	34	HIS	-	expression tag	UNP P40969
A	35	HIS	-	expression tag	UNP P40969
A	36	SER	-	expression tag	UNP P40969
A	37	SER	-	expression tag	UNP P40969
A	38	GLY	-	expression tag	UNP P40969
A	39	LEU	-	expression tag	UNP P40969
A	40	VAL	-	expression tag	UNP P40969
A	41	PRO	-	expression tag	UNP P40969
A	42	ARG	-	expression tag	UNP P40969
A	43	GLY	-	expression tag	UNP P40969
A	44	SER	-	expression tag	UNP P40969
A	45	HIS	-	expression tag	UNP P40969
A	46	MET	-	expression tag	UNP P40969
B	25	MET	-	initiating methionine	UNP P40969
B	26	GLY	-	expression tag	UNP P40969
B	27	GLY	-	expression tag	UNP P40969
B	28	SER	-	expression tag	UNP P40969

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	29	SER	-	expression tag	UNP P40969
B	30	HIS	-	expression tag	UNP P40969
B	31	HIS	-	expression tag	UNP P40969
B	32	HIS	-	expression tag	UNP P40969
B	33	HIS	-	expression tag	UNP P40969
B	34	HIS	-	expression tag	UNP P40969
B	35	HIS	-	expression tag	UNP P40969
B	36	SER	-	expression tag	UNP P40969
B	37	SER	-	expression tag	UNP P40969
B	38	GLY	-	expression tag	UNP P40969
B	39	LEU	-	expression tag	UNP P40969
B	40	VAL	-	expression tag	UNP P40969
B	41	PRO	-	expression tag	UNP P40969
B	42	ARG	-	expression tag	UNP P40969
B	43	GLY	-	expression tag	UNP P40969
B	44	SER	-	expression tag	UNP P40969
B	45	HIS	-	expression tag	UNP P40969
B	46	MET	-	expression tag	UNP P40969

- Molecule 2 is a protein called Suppressor of kinetochore protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	150	Total	C	N	O	S	0	0
			1137	717	199	217	4		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	MET	-	initiating methionine	UNP P52286
C	1	GLY	-	expression tag	UNP P52286
C	195	GLY	-	expression tag	UNP P52286
C	196	SER	-	expression tag	UNP P52286

- Molecule 3 is a protein called Centromere DNA-binding protein complex CBF3 subunit C.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	361	Total	C	N	O	S	0	0
			2773	1825	474	465	9		

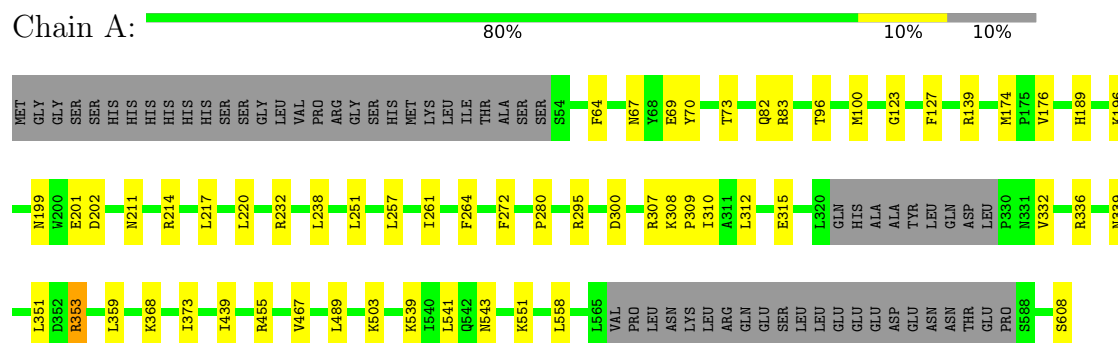
There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	MET	-	initiating methionine	UNP P35203
D	1	GLY	-	expression tag	UNP P35203
D	479	GLY	-	expression tag	UNP P35203
D	480	SER	-	expression tag	UNP P35203
D	481	ARG	-	expression tag	UNP P35203
D	482	SER	-	expression tag	UNP P35203
D	483	GLY	-	expression tag	UNP P35203
D	484	SER	-	expression tag	UNP P35203
D	485	GLU	-	expression tag	UNP P35203
D	486	ASN	-	expression tag	UNP P35203
D	487	LEU	-	expression tag	UNP P35203
D	488	TYR	-	expression tag	UNP P35203
D	489	PHE	-	expression tag	UNP P35203
D	490	GLN	-	expression tag	UNP P35203
D	491	GLY	-	expression tag	UNP P35203
D	492	SER	-	expression tag	UNP P35203
D	493	LYS	-	expression tag	UNP P35203
D	494	ARG	-	expression tag	UNP P35203
D	495	ARG	-	expression tag	UNP P35203
D	496	TRP	-	expression tag	UNP P35203
D	497	LYS	-	expression tag	UNP P35203
D	498	LYS	-	expression tag	UNP P35203
D	499	ASN	-	expression tag	UNP P35203
D	500	PHE	-	expression tag	UNP P35203
D	501	ILE	-	expression tag	UNP P35203
D	502	ALA	-	expression tag	UNP P35203
D	503	VAL	-	expression tag	UNP P35203
D	504	SER	-	expression tag	UNP P35203
D	505	ALA	-	expression tag	UNP P35203
D	506	ALA	-	expression tag	UNP P35203
D	507	ASN	-	expression tag	UNP P35203
D	508	ARG	-	expression tag	UNP P35203
D	509	PHE	-	expression tag	UNP P35203
D	510	LYS	-	expression tag	UNP P35203
D	511	LYS	-	expression tag	UNP P35203
D	512	ILE	-	expression tag	UNP P35203
D	513	SER	-	expression tag	UNP P35203
D	514	SER	-	expression tag	UNP P35203
D	515	SER	-	expression tag	UNP P35203
D	516	GLY	-	expression tag	UNP P35203
D	517	ALA	-	expression tag	UNP P35203
D	518	LEU	-	expression tag	UNP P35203

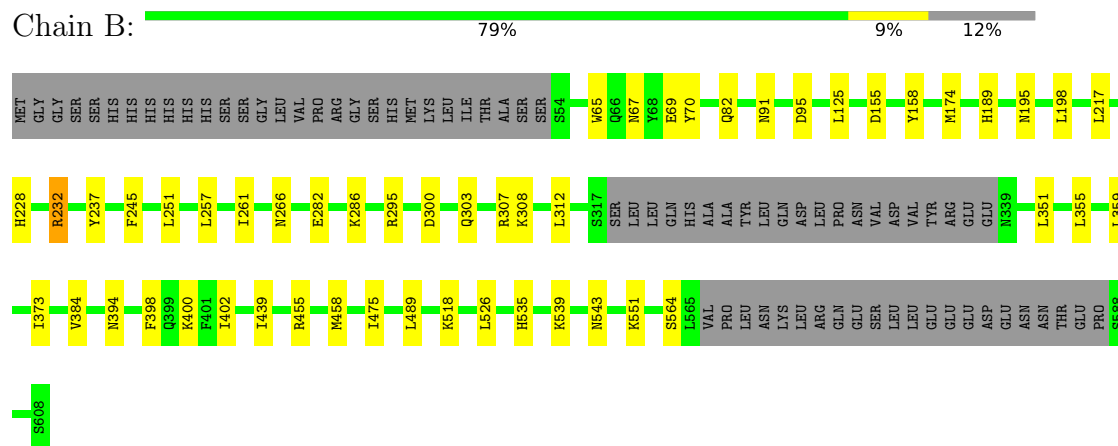
3 Residue-property plots

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.

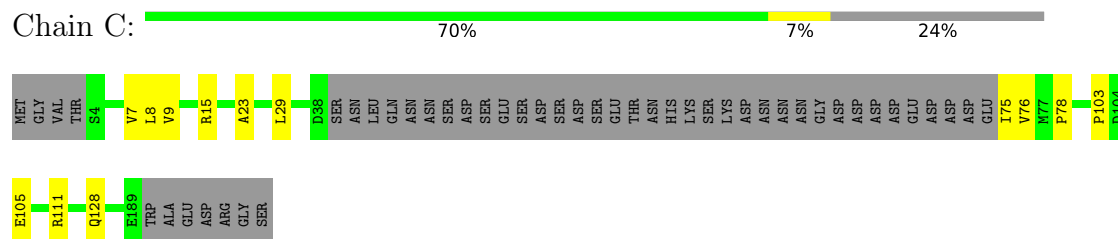
- Molecule 1: Centromere DNA-binding protein complex CBF3 subunit B



- Molecule 1: Centromere DNA-binding protein complex CBF3 subunit B

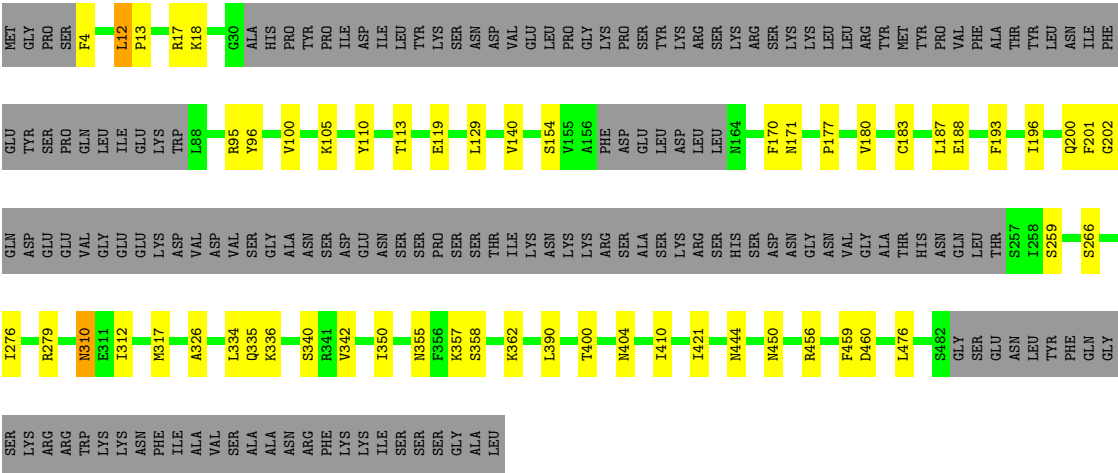


- Molecule 2: Suppressor of kinetochore protein 1



- Molecule 3: Centromere DNA-binding protein complex CBF3 subunit C

Chain D: 59%10%30%



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, POINT	Depositor
Number of particles used	69392, 187606	Depositor
Resolution determination method	FSC 0.143 CUT-OFF, FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION, PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	4.6, 4.06	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	47170	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k), 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	0.31	0/4425	0.51	1/5995 (0.0%)
1	B	0.28	0/4353	0.48	0/5894
2	C	0.26	0/1156	0.49	0/1571
3	D	0.28	0/2832	0.61	2/3851 (0.1%)
All	All	0.29	0/12766	0.52	3/17311 (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
3	D	0	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	12	LEU	CA-CB-CG	6.39	129.99	115.30
1	A	558	LEU	CA-CB-CG	5.88	128.82	115.30
3	D	390	LEU	CA-CB-CG	5.24	127.35	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	D	450	ASN	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	4320	0	4272	33	0
1	B	4249	0	4208	29	0
2	C	1137	0	1071	8	0
3	D	2773	0	2669	31	0
All	All	12479	0	12220	94	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 4.

The worst 5 of 94 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:B:125:LEU:O	1:B:295:ARG:NH1	2.34	0.61
3:D:188:GLU:HA	3:D:193:PHE:HB2	1.82	0.60
3:D:171:ASN:HA	3:D:200:GLN:HB2	1.84	0.59
1:A:123:GLY:O	1:A:127:PHE:HB2	2.03	0.58
1:A:308:LYS:NZ	1:A:310:ILE:O	2.37	0.58

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	518/584 (89%)	500 (96%)	18 (4%)	0	100	100
1	B	506/584 (87%)	493 (97%)	13 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	C	146/197 (74%)	138 (94%)	7 (5%)	1 (1%)	24	66
3	D	353/519 (68%)	307 (87%)	46 (13%)	0	100	100
All	All	1523/1884 (81%)	1438 (94%)	84 (6%)	1 (0%)	56	87

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	103	PRO

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	481/543 (89%)	478 (99%)	3 (1%)	87	94
1	B	475/543 (88%)	472 (99%)	3 (1%)	87	94
2	C	116/180 (64%)	116 (100%)	0	100	100
3	D	271/481 (56%)	266 (98%)	5 (2%)	62	84
All	All	1343/1747 (77%)	1332 (99%)	11 (1%)	84	92

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	232	ARG
1	B	394	ASN
3	D	357	LYS
1	B	67	ASN
3	D	355	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	394	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	399	GLN
3	D	335	GLN
1	B	82	GLN
3	D	355	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 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.