



## wwPDB/EMDatabank EM Map/Model Validation Summary Report ⓘ

Aug 20, 2017 – 08:27 PM EDT

PDB ID : 5A5T  
EMDB ID: : EMD-3056  
Title : Structure of mammalian eIF3 in the context of the 43S preinitiation complex  
Authors : des-Georges, A.; Dhote, V.; Kuhn, L.; Hellen, C.U.T.; Pestova, T.V.; Frank, J.; Hashem, Y.  
Deposited on : unknown  
Resolution : 6.00 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](mailto: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-20029824

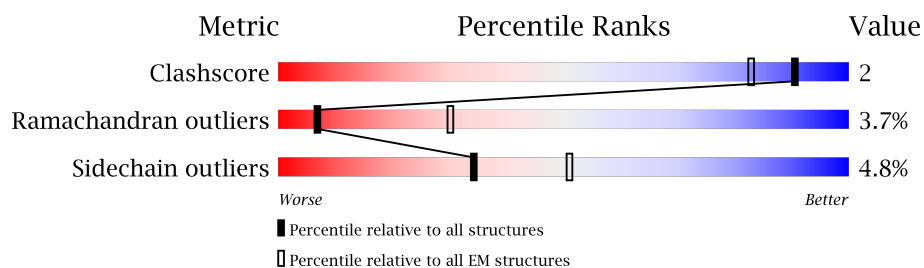
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 6.00 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  $\geq 3$ , 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	1362	38% 5% • 56%
2	C	843	57% 7% • 34%
3	E	445	78% 13% • • 6%
4	F	364	65% 9% • 25%
5	H	352	78% 12% • • 8%
6	K	218	93% 6% •
7	L	564	57% 7% • 34%
8	M	374	84% 12% • •

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 25432 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 EUKARYOTIC TRANSLATION INITIATION FACTOR 3 SUBUNIT A.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	600	Total	C	N	O	S	0	1
			4935	3107	893	914	21		

- Molecule 2 is a protein called EUKARYOTIC TRANSLATION INITIATION FACTOR 3 SUBUNIT C.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	558	Total	C	N	O	S	0	1
			4529	2842	805	849	33		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	577	TYR	ALA	conflict	UNP G1U971

- Molecule 3 is a protein called EUKARYOTIC TRANSLATION INITIATION FACTOR 3 SUBUNIT E.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	E	420	Total	C	N	O	S	0	1
			3466	2220	587	639	20		

- Molecule 4 is a protein called EUKARYOTIC TRANSLATION INITIATION FACTOR 3 SUBUNIT F.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	F	272	Total	C	N	O	S	0	0
			2111	1330	359	410	12		

- Molecule 5 is a protein called EUKARYOTIC TRANSLATION INITIATION FACTOR 3 SUBUNIT H.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	H	324	Total	C	N	O	S	0	0
			2624	1654	452	503	15		

- Molecule 6 is a protein called EUKARYOTIC TRANSLATION INITIATION FACTOR 3 SUBUNIT K.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	K	216	Total	C	N	O	S	0	1
			1738	1109	286	330	13		

- Molecule 7 is a protein called UNCHARACTERIZED PROTEIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	L	373	Total	C	N	O	S	0	1
			3110	2010	520	563	17		

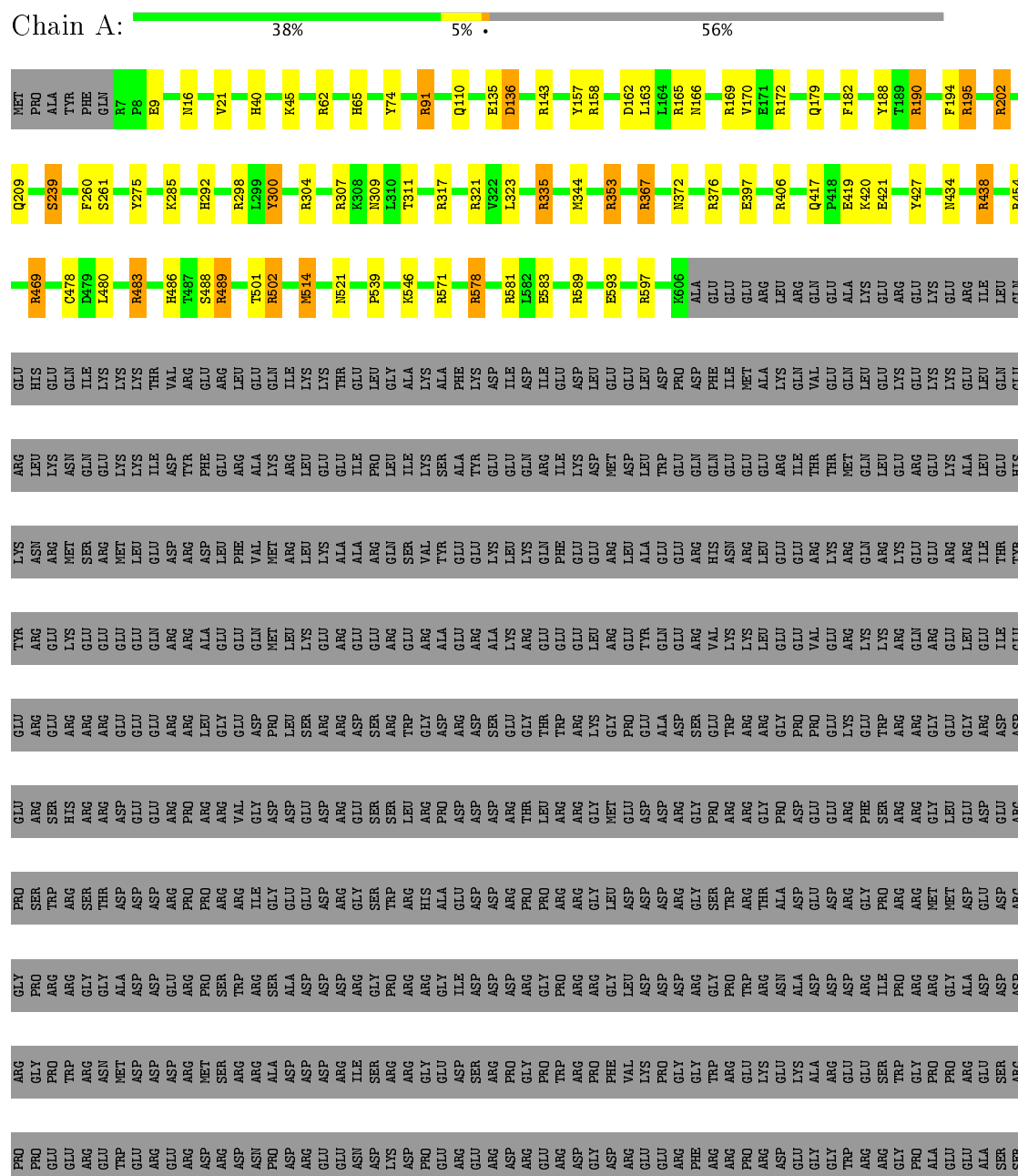
- Molecule 8 is a protein called EUKARYOTIC TRANSLATION INITIATION FACTOR 3 SUBUNIT M.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	M	366	Total	C	N	O	S	0	1
			2919	1850	494	558	17		

### 3 Residue-property plots

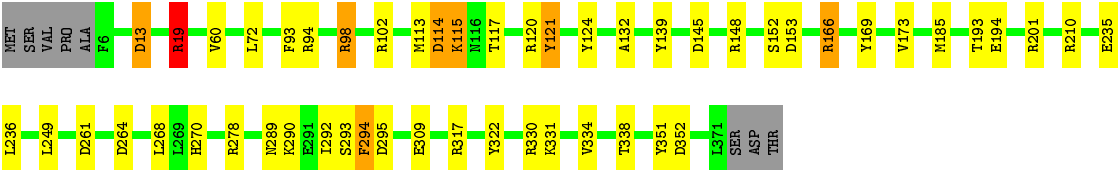
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according 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: EUKARYOTIC TRANSLATION INITIATION FACTOR 3 SUBUNIT A









## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	87192	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	EACH PARTICLE	Depositor
Microscope	FEI/PHILIPS CM300FEG/HE	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	25	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	4500	Depositor
Magnification	30120	Depositor
Image detector	GATAN K2 (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	0.70	0/5021	1.22	40/6781 (0.6%)
2	C	0.70	0/4608	1.19	35/6219 (0.6%)
3	E	1.61	8/3539 (0.2%)	1.29	43/4788 (0.9%)
4	F	0.69	0/2149	1.23	14/2920 (0.5%)
5	H	0.71	0/2675	1.09	7/3609 (0.2%)
6	K	0.69	0/1773	1.09	5/2398 (0.2%)
7	L	0.74	0/3186	1.18	28/4298 (0.7%)
8	M	0.70	0/2964	1.20	19/4000 (0.5%)
All	All	0.88	8/25915 (0.0%)	1.20	191/35013 (0.5%)

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	22
2	C	1	18
3	E	1	11
4	F	0	6
5	H	1	8
6	K	0	4
7	L	2	12
8	M	1	13
All	All	6	94

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	189	LEU	CB-CG	53.28	3.07	1.52
3	E	217	TRP	CD2-CE3	30.41	1.85	1.40
3	E	217	TRP	CD2-CE2	30.28	1.77	1.41
3	E	217	TRP	CE2-CZ2	28.28	1.87	1.39

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	217	TRP	CE3-CZ3	22.94	1.77	1.38

The worst 5 of 191 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	F	128	ARG	NE-CZ-NH1	14.34	127.47	120.30
3	E	251	PRO	CA-C-N	-13.42	87.67	117.20
3	E	251	PRO	C-N-CA	11.49	150.42	121.70
3	E	251	PRO	CA-C-O	11.39	147.53	120.20
3	E	189	LEU	CB-CG-CD1	11.26	130.14	111.00

5 of 6 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	C	513	LYS	CA
3	E	67	PRO	CA
5	H	223	ALA	CA
7	L	418	PRO	CA
7	L	439	PRO	CA

5 of 94 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	165	ARG	Sidechain
1	A	169	ARG	Sidechain
1	A	190	ARG	Sidechain
1	A	195	ARG	Sidechain
1	A	91	ARG	Sidechain

## 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	4935	0	5017	3	0
2	C	4529	0	4533	5	0
3	E	3466	0	3446	74	0
4	F	2111	0	2105	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	H	2624	0	2592	5	0
6	K	1738	0	1706	3	0
7	L	3110	0	3084	5	0
8	M	2919	0	2950	1	0
All	All	25432	0	25433	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 2.

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)
3:E:217:TRP:CE3	3:E:217:TRP:CZ3	1.77	1.71
3:E:217:TRP:CH2	3:E:217:TRP:CZ3	1.75	1.62
3:E:217:TRP:CE2	3:E:217:TRP:CZ2	1.87	1.60
3:E:217:TRP:CH2	3:E:217:TRP:CZ2	1.79	1.59
3:E:217:TRP:CE3	3:E:217:TRP:CD2	1.86	1.59

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	598/1362 (44%)	544 (91%)	36 (6%)	18 (3%)	5	37
2	C	556/843 (66%)	512 (92%)	28 (5%)	16 (3%)	5	38
3	E	418/445 (94%)	362 (87%)	37 (9%)	19 (4%)	3	28
4	F	270/364 (74%)	234 (87%)	20 (7%)	16 (6%)	2	23
5	H	322/352 (92%)	277 (86%)	29 (9%)	16 (5%)	2	27
6	K	214/218 (98%)	200 (94%)	12 (6%)	2 (1%)	20	63

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
7	L	371/564 (66%)	331 (89%)	26 (7%)	14 (4%)	4	32
8	M	364/374 (97%)	320 (88%)	30 (8%)	14 (4%)	4	32
All	All	3113/4522 (69%)	2780 (89%)	218 (7%)	115 (4%)	7	33

5 of 115 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	40	HIS
1	A	136	ASP
1	A	166	ASN
1	A	478	CYS
1	A	539	PRO

### 5.3.2 Protein sidechains ⓘ

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	551/1245 (44%)	528 (96%)	23 (4%)	34	64
2	C	503/750 (67%)	480 (95%)	23 (5%)	31	62
3	E	384/406 (95%)	361 (94%)	23 (6%)	22	55
4	F	239/282 (85%)	231 (97%)	8 (3%)	43	70
5	H	293/311 (94%)	270 (92%)	23 (8%)	15	46
6	K	190/193 (98%)	188 (99%)	2 (1%)	78	89
7	L	342/516 (66%)	325 (95%)	17 (5%)	28	60
8	M	327/335 (98%)	309 (94%)	18 (6%)	25	58
All	All	2829/4038 (70%)	2692 (95%)	137 (5%)	34	61

5 of 137 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	E	270	ARG
4	F	329	MET

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
8	M	152	SER
3	E	331	ASP
4	F	115	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
3	E	216	HIS
3	E	349	GLN
7	L	546	GLN
3	E	85	GLN
5	H	328	GLN

### 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	E	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	E	251:PRO	C	252:HIS	N	1.68