



## wwPDB EM Validation Summary Report ⓘ

May 4, 2024 – 02:04 pm BST

PDB ID : 5GAP  
EMDB ID : EMD-8014  
Title : Body region of the U4/U6.U5 tri-snRNP  
Authors : Nguyen, T.H.D.; Galej, W.P.; Oubridge, C.; Bai, X.C.; Newman, A.; Scheres, S.; Nagai, K.  
Deposited on : 2015-12-15  
Resolution : 3.60 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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.

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:

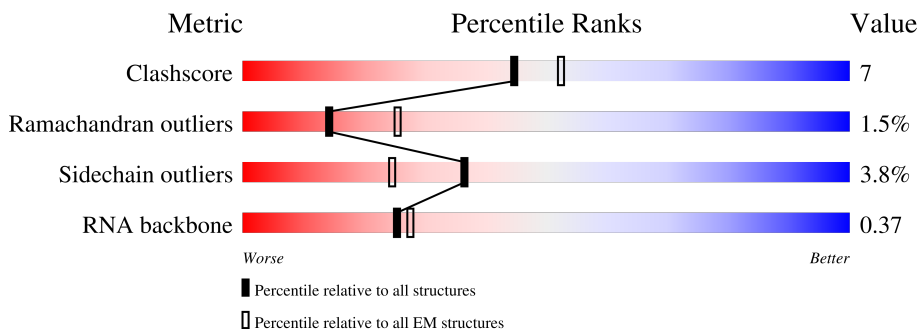
EMDB validation analysis : 0.0.1.dev92  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : **FAILED**  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36.2

# 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 3.60 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.







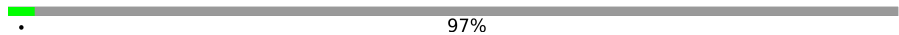
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	V	67	42% (green), 37% (yellow), 19% (orange), 2% (red), 0% (grey)
2	W	112	23% (green), 17% (yellow), 10% (orange), 50% (grey)
3	U	214	6% (green), 2% (yellow), 91% (grey)
4	x	82	100% (green)
5	A	2413	47% (green), 9% (yellow), 44% (grey)
6	H	465	59% (green), 18% (yellow), 23% (grey)
7	J	899	64% (green), 16% (yellow), 1% (orange), 19% (grey)

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
8	D	143	 76% 21% ..
9	F	494	 65% 18% • 16%
10	G	469	 59% 9% 32%
11	K	126	 70% 25% ..
12	B	2163	 • 97%

## 2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 31576 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 RNA chain called U4 snRNA, 5' region, nucleotides 1-67.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	V	67	1426	637	247	475	67	0	0

- Molecule 2 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	W	56	1190	533	210	391	56	0	0

- Molecule 3 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	U	20	414	186	64	144	20	0	0

- Molecule 4 is a protein called unknown protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
4	x	82	410	246	82	82	0	0

- Molecule 5 is a protein called Pre-mRNA-splicing factor 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	A	1349	11066	7094	1901	2031	40	0	0

- Molecule 6 is a protein called U4/U6 small nuclear ribonucleoprotein PRP4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	H	357	2789	1743	501	532	13	0	0

- Molecule 7 is a protein called Pre-mRNA-splicing factor 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	J	729	5822	3726	992	1079	25	0	0

- Molecule 8 is a protein called Spliceosomal protein DIB1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	D	140	1151	728	200	212	11	0	0

- Molecule 9 is a protein called Pre-mRNA-processing factor 31.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
9	F	415	3218	2052	575	580	11	0	0

- Molecule 10 is a protein called U4/U6 small nuclear ribonucleoprotein PRP3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	G	318	2632	1659	469	488	16	0	0

- Molecule 11 is a protein called 13 kDa ribonucleoprotein-associated protein.

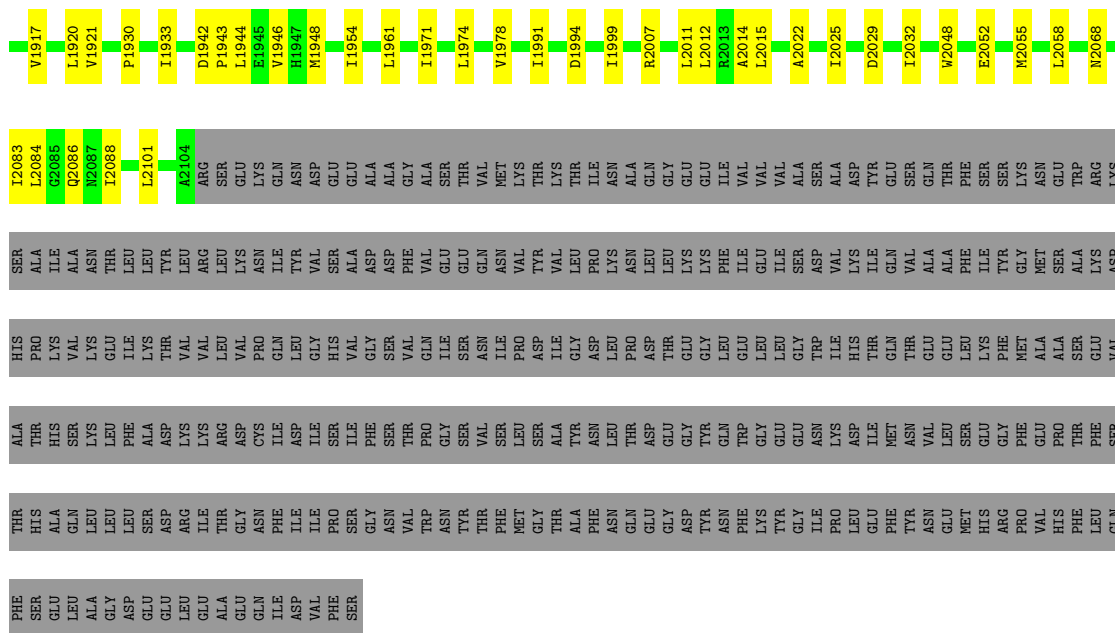
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
11	K	124	936	597	161	174	4	0	0

- Molecule 12 is a protein called Pre-mRNA-splicing helicase BRR2.

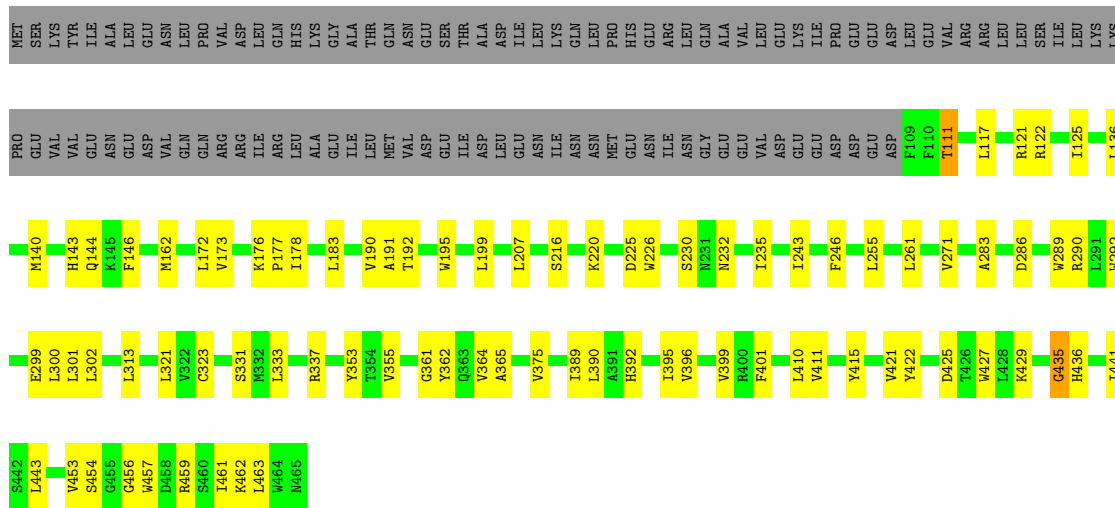
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
12	B	71	522	326	89	106	1	0	0



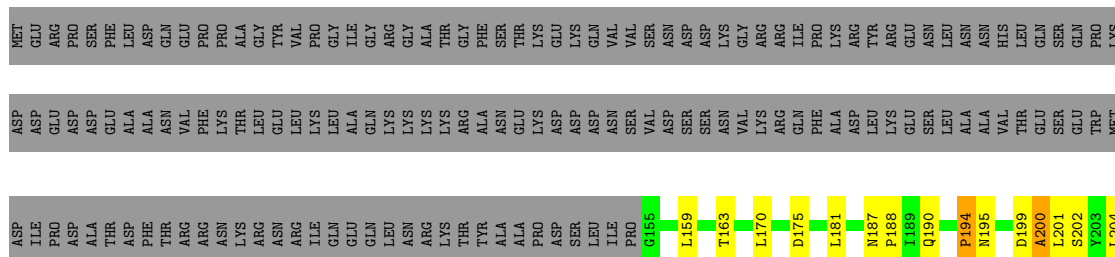




• Molecule 6: U4/U6 small nuclear ribonucleoprotein PRP4

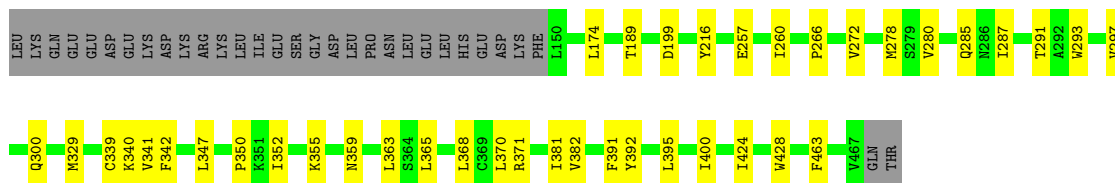


• Molecule 7: Pre-mRNA-splicing factor 6

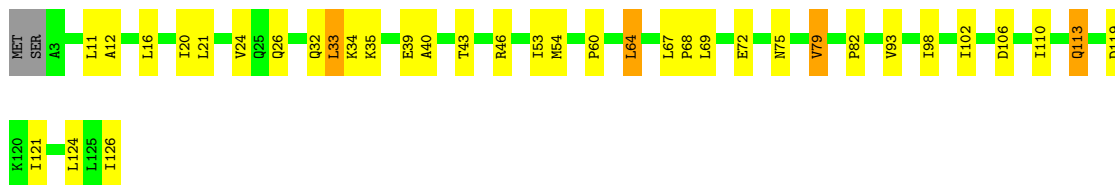




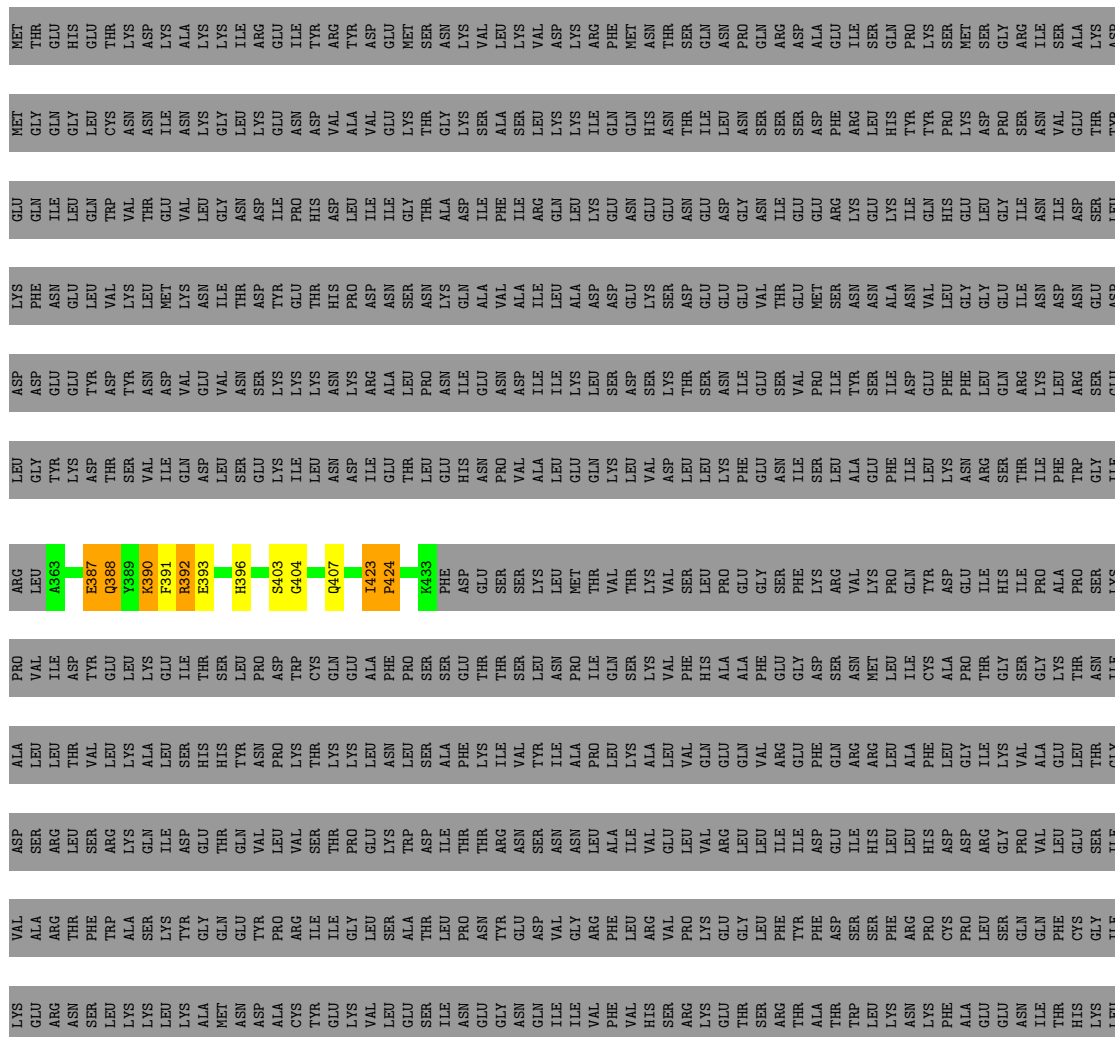




● Molecule 11: 13 kDa ribonucleoprotein-associated protein



● Molecule 12: Pre-mRNA-splicing helicase BRR2





LEU ASN  
THR ALA  
THR THR  
THR ASP  
SER ALA  
MET  
ASP  
ALA  
GLN MET  
LEU ILE  
GLN  
GLY  
VAL ASN  
VAL  
TYR TRP  
ASP  
VAL  
ASP  
GLU  
PRO  
LEU  
LEU  
THR ARG  
GLN  
ILE  
PRO  
ASN  
HIS  
PHE  
ASN  
ASN  
ASN  
SER  
SER  
ASP  
ASN  
LEU  
LYS  
ILE  
ILE  
SER  
GLY  
GLU  
LYS  
CYS  
LYS  
GLU  
ILE  
ILE  
THR  
ASN  
VAL  
GLU  
THR  
VAL  
TYR  
ASP  
ASP  
ILE  
MET  
ALA  
GLU  
ASN  
GLU  
ASP  
ASP  
GLU  
ARG  
GLU  
VAL  
GLU  
ARG  
ASP  
GLU  
GLU  
LYS

LEU  
THR  
THR  
THR  
SER  
GLN  
LEU  
ALA  
GLN  
VAL  
ALA  
ALA  
PHE  
VAL  
GLU  
VAL  
ASN  
SER  
TYR  
LYS  
PRO  
ASN  
VAL  
GLU  
TYR  
LEU  
ALA  
LEU  
ILE  
THR  
TYR  
SER  
SER  
LEU  
PRO  
ASN  
ASN  
PHE  
SER  
ASP  
SER  
SER  
LEU  
ILE  
ILE  
SER  
GLN  
TYR  
GLY  
VAL  
LYS  
LYS  
GLN  
LYS  
PHE  
GLU  
ASP  
ILE  
THR  
THR  
ASN  
ILE  
VAL  
GLU  
THR  
THR  
THR  
THR  
TYR  
ARG  
TYR  
ASP  
VAL  
ILE  
MET  
GLU  
PRO  
GLU  
ASN  
TRP  
ASN  
LEU  
LEU  
GLN  
VAL  
VAL  
GLU  
THR  
SER  
SER  
TYR  
LEU

TYR  
PRO  
PHE  
ASP  
LYS  
LEU  
GLU  
SER  
TRP  
TRP  
LEU  
VAL  
LEU  
LEU  
GLY  
GLU  
VAL  
VAL  
LYS  
LYS  
GLU  
LEU  
TYR  
ALA  
ALA  
ILE  
LYS  
LYS  
VAL  
THR  
THR  
LEU  
ASN  
ASN  
PHE  
SER  
LYS  
GLU  
GLU  
THR  
GLN  
GLN  
SER  
GLN  
TYR  
TYR  
GLU  
LEU  
GLU  
GLU  
PHE  
ASP  
THR  
THR  
PRO  
THR  
SER  
SER  
GLY  
LYS  
HIS  
ASN  
LEU  
THR  
THR  
ILE  
TRP  
CYS  
VAL  
VAL  
CYS  
ASP  
SER  
TYR  
LEU

ASP  
ALA  
ASP  
LYS  
LEU  
SER  
PHE  
GLU  
ILE  
ASN  
VAL  
LYS

## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	140155	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	38	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	35714	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  >5	RMSZ	# Z  >5
1	V	0.36	0/1593	0.79	1/2480 (0.0%)
2	W	0.38	0/1328	0.84	2/2061 (0.1%)
3	U	0.32	0/459	0.76	0/710
5	A	0.42	1/11327 (0.0%)	0.73	2/15348 (0.0%)
6	H	0.38	0/2845	0.71	0/3843
7	J	0.44	0/5934	0.80	2/8039 (0.0%)
8	D	0.42	0/1172	0.75	1/1578 (0.1%)
9	F	0.43	0/3273	0.80	2/4413 (0.0%)
10	G	0.41	0/2687	0.69	0/3611
11	K	0.45	0/949	0.81	0/1292
12	B	0.42	0/529	0.69	0/716
All	All	0.42	1/32096 (0.0%)	0.76	10/44091 (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
5	A	0	3
9	F	0	1
12	B	0	1
All	All	0	5

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	825	SER	CB-OG	5.40	1.49	1.42

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	J	613	LEU	CA-CB-CG	7.12	131.69	115.30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	979	SER	C-N-CD	-6.95	105.32	120.60
1	V	17	A	C2'-C3'-O3'	6.90	124.74	113.70
2	W	55	G	C2'-C3'-O3'	5.73	122.87	113.70
9	F	74	LEU	CA-CB-CG	5.58	128.14	115.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	A	1278	VAL	Peptide
5	A	1994	ASP	Peptide
5	A	979	SER	Peptide
12	B	423	ILE	Peptide
9	F	154	SER	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	V	1426	0	716	22	0
2	W	1190	0	603	14	0
3	U	414	0	213	4	0
4	x	410	0	89	0	0
5	A	11066	0	11078	138	0
6	H	2789	0	2725	53	0
7	J	5822	0	5792	103	0
8	D	1151	0	1138	17	0
9	F	3218	0	3297	53	0
10	G	2632	0	2599	27	0
11	K	936	0	987	24	0
12	B	522	0	506	5	0
All	All	31576	0	29743	413	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 7.

The worst 5 of 413 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)
5:A:1578:ALA:HB1	5:A:1602:PRO:HB3	1.40	1.02
5:A:1275:MET:HE2	5:A:1281:ASN:ND2	1.92	0.84
10:G:272:VAL:HG12	10:G:280:VAL:HG21	1.59	0.81
5:A:1067:ASN:HB2	5:A:1083:THR:HG21	1.64	0.80
7:J:199:ASP:O	7:J:201:LEU:N	2.14	0.79

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	
5	A	1343/2413 (56%)	1226 (91%)	104 (8%)	13 (1%)	15	55
6	H	355/465 (76%)	301 (85%)	48 (14%)	6 (2%)	9	45
7	J	719/899 (80%)	650 (90%)	56 (8%)	13 (2%)	8	43
8	D	138/143 (96%)	126 (91%)	11 (8%)	1 (1%)	22	61
9	F	413/494 (84%)	364 (88%)	38 (9%)	11 (3%)	5	35
10	G	316/469 (67%)	283 (90%)	32 (10%)	1 (0%)	41	75
11	K	122/126 (97%)	111 (91%)	9 (7%)	2 (2%)	9	46
12	B	69/2163 (3%)	57 (83%)	8 (12%)	4 (6%)	1	18
All	All	3475/7172 (48%)	3118 (90%)	306 (9%)	51 (2%)	14	47

5 of 51 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	A	1093	LYS
5	A	1278	VAL
5	A	2088	ILE
6	H	362	TYR
7	J	479	GLU



### 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
5	A	1216/2182 (56%)	1194 (98%)	22 (2%)	59 81
6	H	305/410 (74%)	290 (95%)	15 (5%)	25 59
7	J	627/813 (77%)	600 (96%)	27 (4%)	29 63
8	D	129/132 (98%)	121 (94%)	8 (6%)	18 53
9	F	346/445 (78%)	324 (94%)	22 (6%)	17 52
10	G	289/436 (66%)	283 (98%)	6 (2%)	53 78
11	K	102/104 (98%)	90 (88%)	12 (12%)	5 28
12	B	56/1955 (3%)	52 (93%)	4 (7%)	14 48
All	All	3070/6477 (47%)	2954 (96%)	116 (4%)	36 66

5 of 116 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	J	657	ASN
11	K	113	GLN
8	D	139	HIS
11	K	93	VAL
10	G	395	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
7	J	397	GLN
7	J	578	GLN
10	G	300	GLN
7	J	537	GLN
7	J	673	GLN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	V	66/67 (98%)	29 (43%)	6 (9%)
2	W	54/112 (48%)	20 (37%)	9 (16%)
3	U	19/214 (8%)	8 (42%)	1 (5%)
All	All	139/393 (35%)	57 (41%)	16 (11%)

5 of 57 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	V	2	U
1	V	11	A
1	V	18	A
1	V	19	U
1	V	20	A

5 of 16 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	W	84	C
2	W	83	A
2	W	45	A
2	W	75	A
2	W	32	U

#### 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 monosaccharides 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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
4	x	1
5	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	x	62:UNK	C	101:UNK	N	54.04
1	A	1860:VAL	C	1861:THR	N	4.50

## 6 Map visualisation

This section contains visualisations of the EMDB entry EMD-8014. These allow visual inspection of the internal detail of the map and identification of artifacts.

No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

### 6.1 Orthogonal projections

This section was not generated.

### 6.2 Central slices

This section was not generated.

### 6.3 Largest variance slices

This section was not generated.

### 6.4 Orthogonal standard-deviation projections (False-color)

This section was not generated.

### 6.5 Orthogonal surface views

This section was not generated.

### 6.6 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

## 7 Map analysis

This section contains the results of statistical analysis of the map.

### 7.1 Map-value distribution

This section was not generated.

### 7.2 Volume estimate versus contour level

This section was not generated.

### 7.3 Rotationally averaged power spectrum

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

## 8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

## 9 Map-model fit

This section was not generated.