



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

Feb 20, 2018 – 04:36 PM EST

PDB ID : 6ALG  
EMDB ID: : EMD-8584  
Title : CryoEM structure of HK022 Nun - E.coli RNA polymerase elongation complex  
Authors : Kang, J.Y.; Darst, S.A.  
Deposited on : 2017-08-07  
Resolution : 3.70 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-20030736

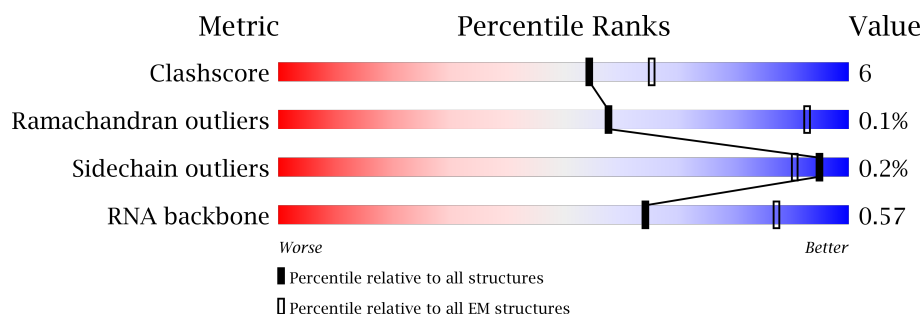
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.70 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	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026
RNA backbone	3398	335

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	29	90% 10%
2	B	29	72% 28%
3	R	20	35% 15% 50%
4	G	239	79% 13% 8%
4	H	239	80% 10% 9%
5	I	1342	81% 17% .
6	J	1407	76% 18% 5%
7	K	91	73% 8% 20%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
8	N	109	 17% . 79%

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 26327 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 DNA chain called DNA (29-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	29	Total	C	N	O	P	0	0
			586	279	108	171	28		

- Molecule 2 is a DNA chain called DNA (29-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	29	Total	C	N	O	P	0	0
			583	279	99	177	28		

- Molecule 3 is a RNA chain called RNA (5'-R(P\*CP\*GP\*GP\*AP\*GP\*AP\*GP\*GP\*UP\*A)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
3	R	10	Total	C	N	O	P	0	0
			221	98	45	68	10		

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	G	219	Total	C	N	O	S	0	0
			1690	1058	298	328	6		
4	H	217	Total	C	N	O	S	0	0
			1673	1046	296	325	6		

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	235	GLU	-	expression tag	UNP P0A7Z4
G	236	VAL	-	expression tag	UNP P0A7Z4
G	237	LEU	-	expression tag	UNP P0A7Z4
G	238	PHE	-	expression tag	UNP P0A7Z4
G	239	GLN	-	expression tag	UNP P0A7Z4
H	235	GLU	-	expression tag	UNP P0A7Z4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	236	VAL	-	expression tag	UNP P0A7Z4
H	237	LEU	-	expression tag	UNP P0A7Z4
H	238	PHE	-	expression tag	UNP P0A7Z4
H	239	GLN	-	expression tag	UNP P0A7Z4

- Molecule 5 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	I	1319	Total	C	N	O	S	0	0
			10402	6528	1814	2017	43		

- Molecule 6 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	J	1335	Total	C	N	O	S	0	0
			10384	6524	1851	1959	50		

- Molecule 7 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	K	73	Total	C	N	O	S	0	0
			582	355	111	115	1		

- Molecule 8 is a protein called Transcription termination factor nun.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	N	23	Total	C	N	O	S	0	0
			203	125	43	34	1		

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
9	J	1	Total	Mg	0
			1	1	

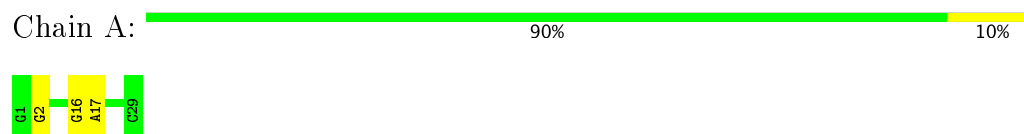
- Molecule 10 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
10	J	2	Total	Zn	0
			2	2	

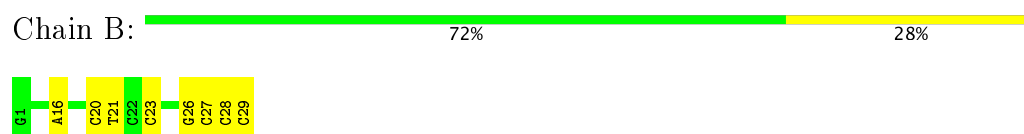
### 3 Residue-property plots [i](#)

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.

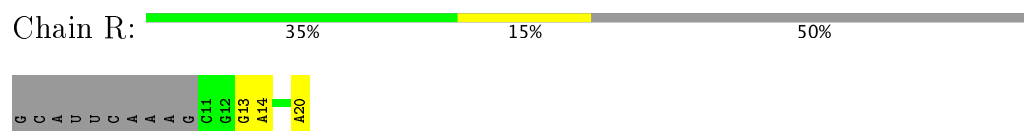
- Molecule 1: DNA (29-MER)



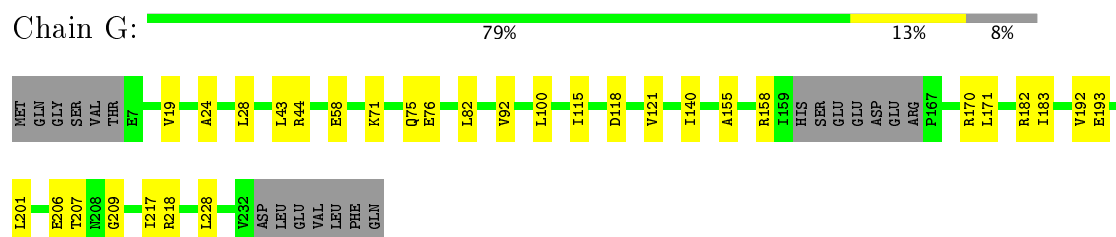
- Molecule 2: DNA (29-MER)



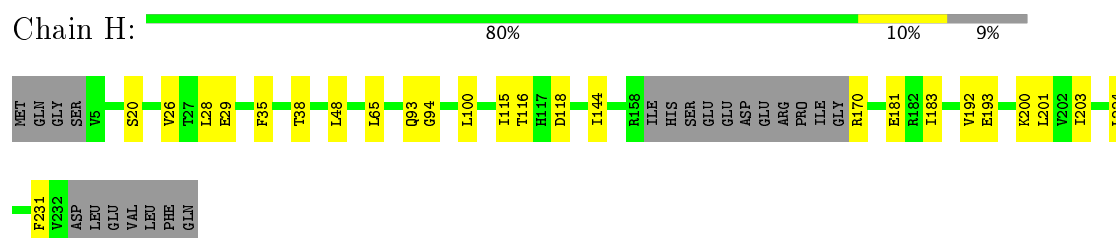
- Molecule 3: RNA (5'-R(P\*CP\*GP\*GP\*AP\*GP\*AP\*GP\*GP\*UP\*A)-3')



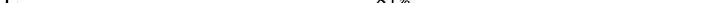
- Molecule 4: DNA-directed RNA polymerase subunit alpha

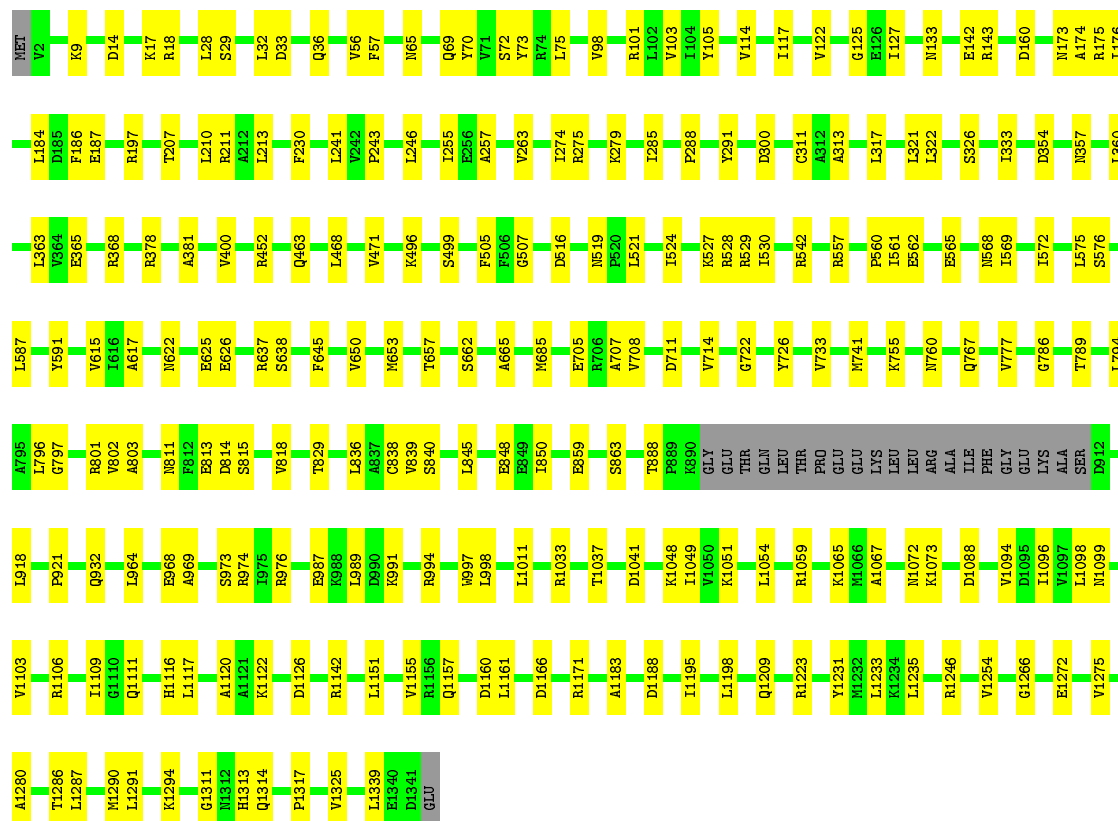


- Molecule 4: DNA-directed RNA polymerase subunit alpha

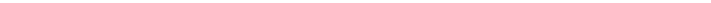


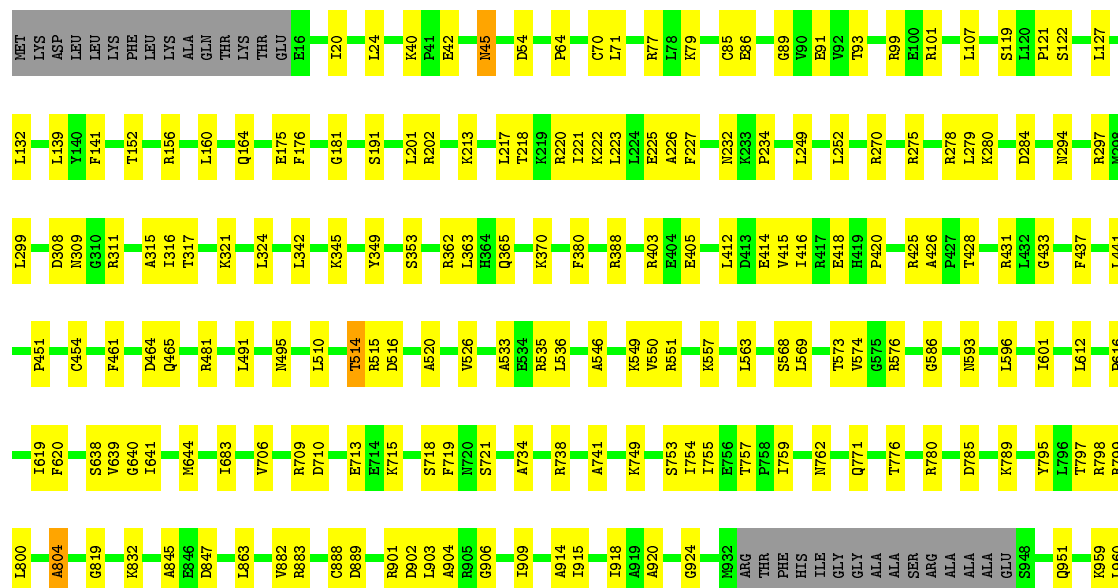
- Molecule 5: DNA-directed RNA polymerase subunit beta

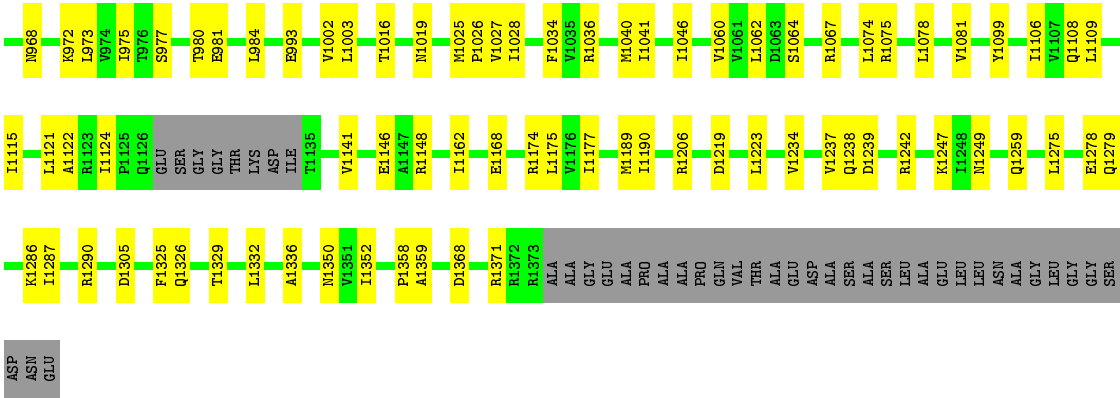
Chain I:  81% 17% .



- Molecule 6: DNA-directed RNA polymerase subunit beta'

Chain J:  76% 18% 5%





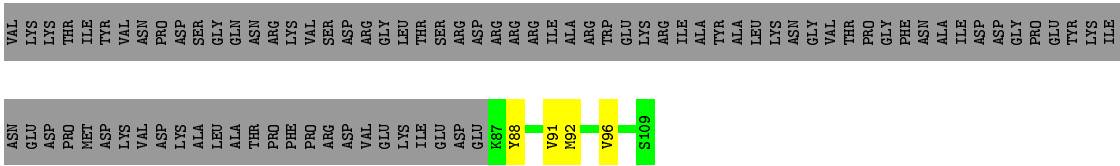
- Molecule 7: DNA-directed RNA polymerase subunit omega

Chain K:



- Molecule 8: Transcription termination factor nun

Chain N:





## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	192000	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE; Used CTFFIND4	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	89	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, MG

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	0.80	0/656	0.97	0/1009
2	B	1.06	0/650	1.02	0/1000
3	R	1.04	0/248	1.05	0/386
4	G	0.41	0/1710	0.59	0/2316
4	H	0.34	0/1692	0.60	0/2293
5	I	0.45	0/10568	0.61	2/14260 (0.0%)
6	J	0.42	0/10541	0.61	3/14232 (0.0%)
7	K	0.37	0/584	0.57	0/786
8	N	0.44	0/209	0.60	0/279
All	All	0.48	0/26858	0.64	5/36561 (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
4	G	0	1
4	H	0	1
6	J	0	2
All	All	0	4

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	J	710	ASP	CB-CG-OD1	7.65	125.19	118.30
6	J	54	ASP	CB-CG-OD1	5.35	123.12	118.30
5	I	845	LEU	CA-CB-CG	5.21	127.28	115.30
5	I	516	ASP	CB-CG-OD2	5.20	122.98	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	J	363	LEU	CA-CB-CG	5.11	127.04	115.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	G	118	ASP	Peptide
4	H	20	SER	Peptide
6	J	416	ILE	Peptide
6	J	804	ALA	Peptide

## 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	586	0	326	3	0
2	B	583	0	329	6	0
3	R	221	0	109	3	0
4	G	1690	0	1731	19	0
4	H	1673	0	1710	13	0
5	I	10402	0	10416	140	0
6	J	10384	0	10604	159	0
7	K	582	0	593	8	0
8	N	203	0	193	4	0
9	J	1	0	0	0	0
10	J	2	0	0	0	0
All	All	26327	0	26011	322	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 6.

The worst 5 of 322 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)
6:J:223:LEU:O	6:J:227:PHE:HB2	1.81	0.79

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:J:550:VAL:O	6:J:569:LEU:HA	1.87	0.74
5:I:1286:THR:O	5:I:1290:MET:HB2	1.92	0.69
2:B:16:DA:H1'	6:J:426:ALA:HB1	1.75	0.67
5:I:964:LEU:O	5:I:968:GLU:HB2	1.94	0.67

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	
4	G	215/239 (90%)	197 (92%)	18 (8%)	0	100	100
4	H	213/239 (89%)	198 (93%)	15 (7%)	0	100	100
5	I	1315/1342 (98%)	1202 (91%)	110 (8%)	3 (0%)	51	85
6	J	1329/1407 (94%)	1235 (93%)	93 (7%)	1 (0%)	55	88
7	K	71/91 (78%)	66 (93%)	5 (7%)	0	100	100
8	N	21/109 (19%)	20 (95%)	1 (5%)	0	100	100
All	All	3164/3427 (92%)	2918 (92%)	242 (8%)	4 (0%)	58	88

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	I	625	GLU
5	I	626	GLU
6	J	175	GLU
5	I	1317	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	
4	G	187/206 (91%)	187 (100%)	0	100	100
4	H	185/206 (90%)	183 (99%)	2 (1%)	78	90
5	I	1137/1157 (98%)	1136 (100%)	1 (0%)	94	98
6	J	1120/1168 (96%)	1117 (100%)	3 (0%)	94	98
7	K	63/75 (84%)	63 (100%)	0	100	100
8	N	22/97 (23%)	22 (100%)	0	100	100
All	All	2714/2909 (93%)	2708 (100%)	6 (0%)	95	98

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	I	1151	LEU
6	J	514	THR
6	J	45	ASN
4	H	170	ARG
6	J	431	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
5	I	1209	GLN
5	I	1313	HIS
6	J	805	GLN
5	I	1157	GLN
6	J	1108	GLN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	R	9/20 (45%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

## 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 ⓘ

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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.