



## wwPDB EM Validation Summary Report ⓘ

Sep 1, 2024 – 12:12 AM JST

PDB ID : 8YB6  
EMDB ID : EMD-39110  
Title : Type I-EHNNH Cascade complex  
Authors : Li, Z.  
Deposited on : 2024-02-11  
Resolution : 3.06 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:

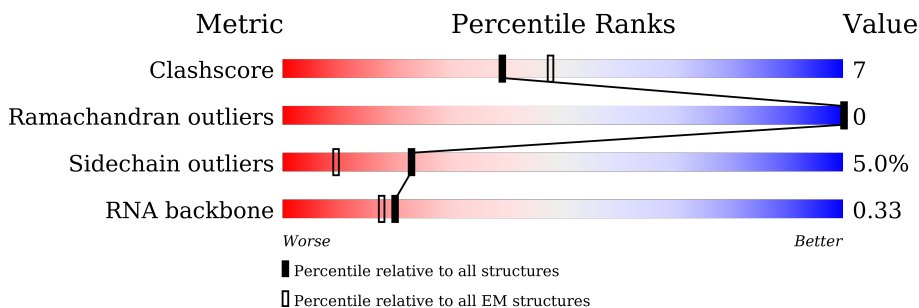
EMDB validation analysis : **FAILED**  
MolProbity : 4.02b-467  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
MapQ : **FAILED**  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.38.2

# 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 3.06 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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415
RNA backbone	6643	2191

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	A	388	73% 15% 10%
2	B	272	77% 19% ..
3	C	61	41% 39% 20%
4	D	378	74% 23% ..
4	E	378	75% 22% ..
4	F	378	81% 17% .
4	G	378	82% 16% .
4	H	378	81% 18% .

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	I	378	
5	J	535	
6	K	174	

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 27758 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 CRISPR system Cascade subunit CasD.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	350	2768	1746	514	490	18	0	0

- Molecule 2 is a protein called CRISPR-associated endoribonuclease Cse3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	267	2166	1392	387	383	4	0	0

- Molecule 3 is a RNA chain called 61-nt crRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	C	61	1303	584	239	420	60	0	0

- Molecule 4 is a protein called CRISPR system Cascade subunit CasC.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	372	2897	1830	506	549	12	0	0
4	E	370	2882	1821	503	546	12	0	0
4	F	374	2911	1839	509	551	12	0	0
4	G	377	2935	1856	512	555	12	0	0
4	H	377	2935	1856	512	555	12	0	0
4	I	268	2072	1319	358	387	8	0	0

- Molecule 5 is a protein called CRISPR-associated protein Cse1 (CRISPR\_cse1).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	J	443	3531	2266	605	641	19	0	0

- Molecule 6 is a protein called CRISPR-associated protein Cse2 (CRISPR\_cse2).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	K	165	1356	878	239	233	6	0	0

- Molecule 7 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

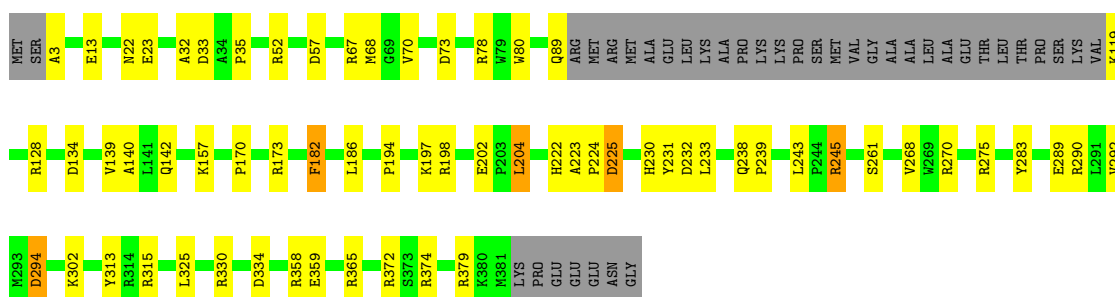
Mol	Chain	Residues	Atoms		AltConf
			Total	Zn	
7	A	1	1	1	0
7	J	1	1	1	0

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


These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

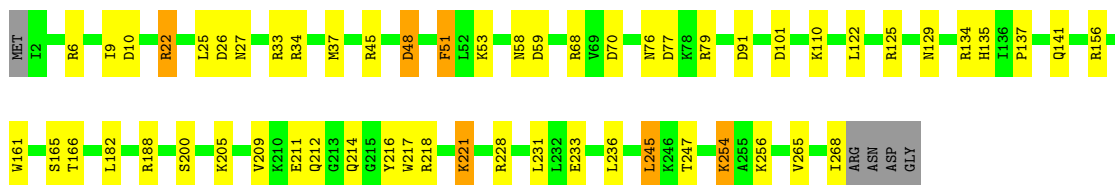
- Molecule 1: CRISPR system Cascade subunit CasD

Chain A: 



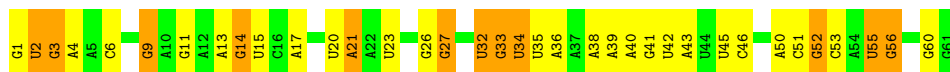
- Molecule 2: CRISPR-associated endoribonuclease Cse3

Chain B: 



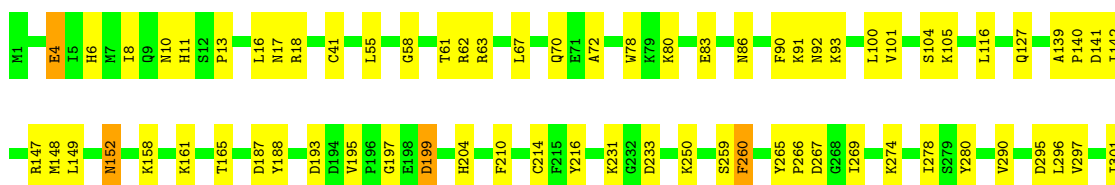
- Molecule 3: 61-nt crRNA

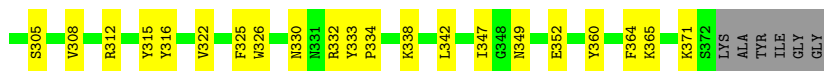
Chain C: 



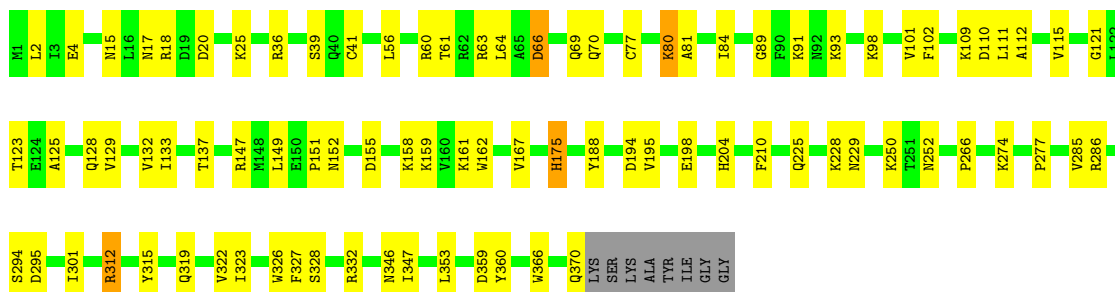
- Molecule 4: CRISPR system Cascade subunit CasC

Chain D: 

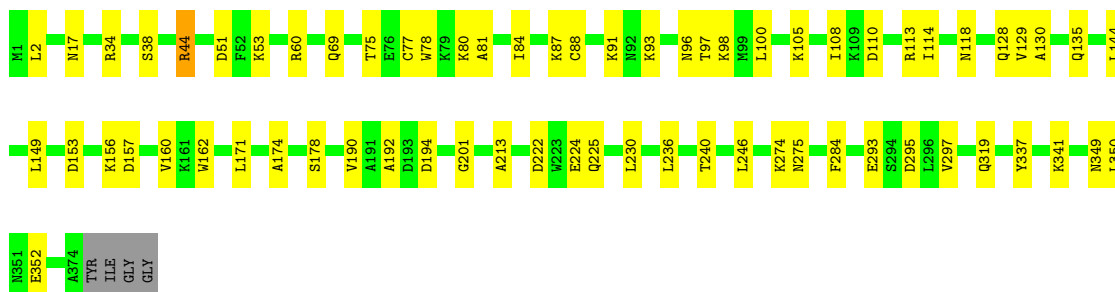
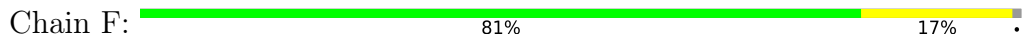




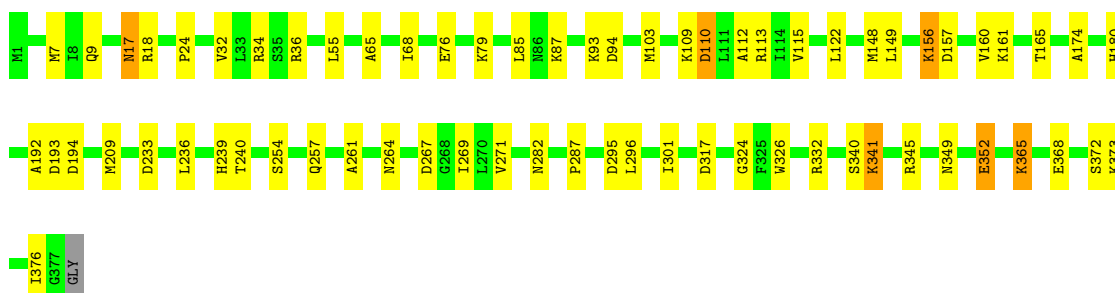
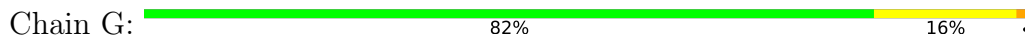
• Molecule 4: CRISPR system Cascade subunit CasC



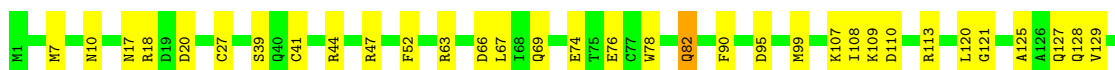
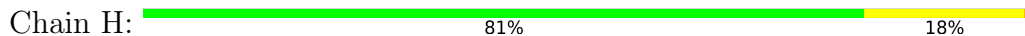
• Molecule 4: CRISPR system Cascade subunit CasC

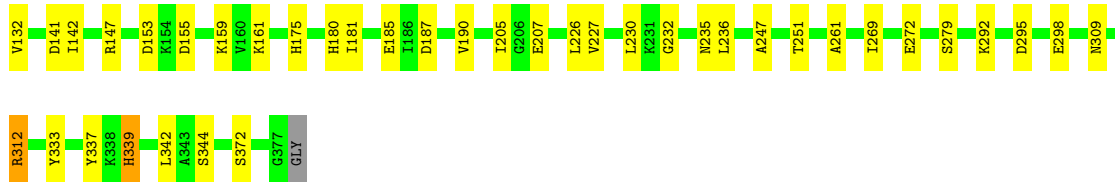


• Molecule 4: CRISPR system Cascade subunit CasC

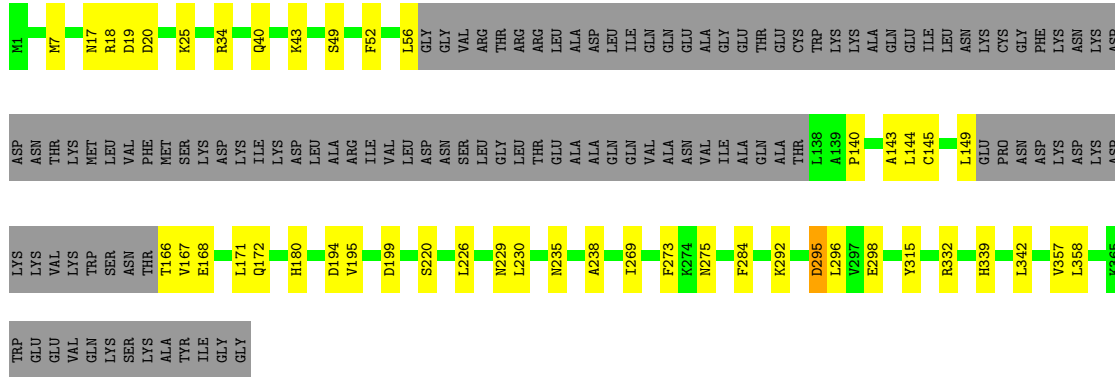


• Molecule 4: CRISPR system Cascade subunit CasC

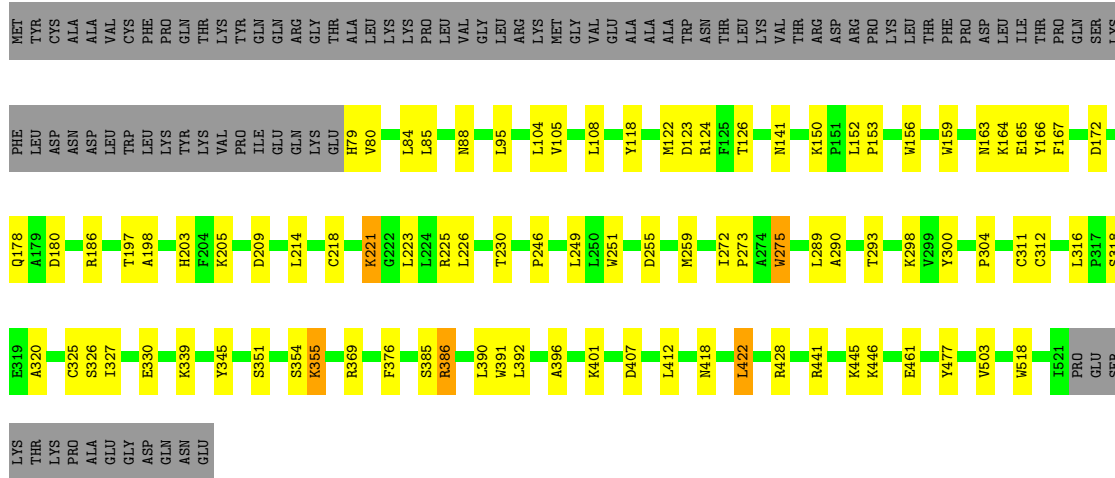




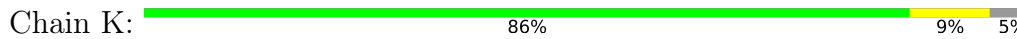
• Molecule 4: CRISPR system Cascade subunit CasC



• Molecule 5: CRISPR-associated protein Cse1 (CRISPR\_cse1)



• Molecule 6: CRISPR-associated protein Cse2 (CRISPR\_cse2)





## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	291872	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	54	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	2200	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k 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

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	A	0.26	0/2837	0.55	0/3848
2	B	0.25	0/2224	0.52	0/3012
3	C	0.30	0/1459	0.88	0/2273
4	D	0.26	0/2954	0.51	0/3995
4	E	0.28	0/2939	0.51	0/3976
4	F	0.28	0/2968	0.49	0/4013
4	G	0.27	0/2993	0.50	0/4047
4	H	0.28	0/2993	0.48	0/4047
4	I	0.27	0/2118	0.49	0/2870
5	J	0.26	0/3632	0.51	0/4942
6	K	0.26	0/1386	0.53	0/1866
All	All	0.27	0/28503	0.54	0/38889

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	2768	0	2765	43	0
2	B	2166	0	2177	41	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	1303	0	659	35	0
4	D	2897	0	2878	48	0
4	E	2882	0	2860	51	0
4	F	2911	0	2896	48	0
4	G	2935	0	2919	37	0
4	H	2935	0	2919	49	0
4	I	2072	0	2036	27	0
5	J	3531	0	3460	54	0
6	K	1356	0	1398	10	0
7	A	1	0	0	0	0
7	J	1	0	0	0	0
All	All	27758	0	26967	385	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 385 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)
4:H:247:ALA:O	4:H:251:THR:HG22	1.33	1.23
2:B:182:LEU:HD23	2:B:247:THR:HG23	1.40	1.02
4:F:114:ILE:HG21	4:F:129:VAL:HG23	1.52	0.91
4:F:105:LYS:O	4:F:108:ILE:HB	1.77	0.83
4:H:247:ALA:O	4:H:251:THR:CG2	2.24	0.82

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	346/388 (89%)	314 (91%)	32 (9%)	0	<b>100</b> <b>100</b>

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	265/272 (97%)	250 (94%)	15 (6%)	0	100	100
4	D	370/378 (98%)	334 (90%)	36 (10%)	0	100	100
4	E	368/378 (97%)	331 (90%)	37 (10%)	0	100	100
4	F	372/378 (98%)	341 (92%)	31 (8%)	0	100	100
4	G	375/378 (99%)	340 (91%)	35 (9%)	0	100	100
4	H	375/378 (99%)	352 (94%)	23 (6%)	0	100	100
4	I	262/378 (69%)	236 (90%)	26 (10%)	0	100	100
5	J	441/535 (82%)	397 (90%)	44 (10%)	0	100	100
6	K	163/174 (94%)	153 (94%)	10 (6%)	0	100	100
All	All	3337/3637 (92%)	3048 (91%)	289 (9%)	0	100	100

There are no Ramachandran outliers to report.

### 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	290/322 (90%)	277 (96%)	13 (4%)	23	50
2	B	234/238 (98%)	218 (93%)	16 (7%)	13	36
4	D	310/313 (99%)	288 (93%)	22 (7%)	12	35
4	E	308/313 (98%)	289 (94%)	19 (6%)	15	39
4	F	311/313 (99%)	301 (97%)	10 (3%)	34	60
4	G	313/313 (100%)	297 (95%)	16 (5%)	20	46
4	H	313/313 (100%)	300 (96%)	13 (4%)	25	52
4	I	220/313 (70%)	208 (94%)	12 (6%)	18	43
5	J	378/459 (82%)	360 (95%)	18 (5%)	21	48
6	K	144/153 (94%)	141 (98%)	3 (2%)	48	69
All	All	2821/3050 (92%)	2679 (95%)	142 (5%)	23	47

5 of 142 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	I	235	ASN
4	I	315	TYR
5	J	300	TYR
4	D	364	PHE
4	D	333	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
4	E	86	ASN
4	F	275	ASN
4	F	319	GLN
4	G	225	GLN
4	H	131	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	C	60/61 (98%)	22 (36%)	1 (1%)

5 of 22 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	C	2	U
3	C	3	G
3	C	9	G
3	C	13	A
3	C	14	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
3	C	2	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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 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 [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.