



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 09:30 pm BST

PDB ID : 5FVC  
Title : Structure of RNA-bound decameric HMPV nucleoprotein  
Authors : Renner, M.; Bertinelli, M.; Leyrat, C.; Paesen, G.C.; Saraiva de Oliveira, L.F.;  
Huiskonen, J.T.; Grimes, J.M.  
Deposited on : 2016-02-05  
Resolution : 4.17 Å(reported)

This is a wwPDB X-ray Structure 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/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

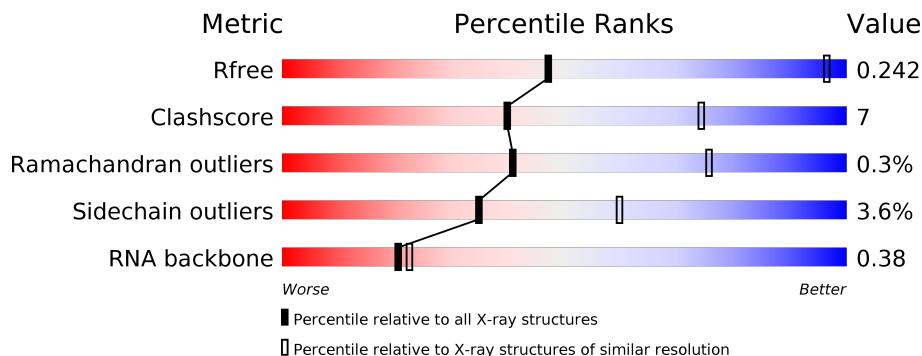
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.17 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)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1034 (4.60-3.76)
Clashscore	141614	1030 (4.54-3.80)
Ramachandran outliers	138981	1006 (4.58-3.78)
Sidechain outliers	138945	1037 (4.60-3.76)
RNA backbone	3102	1049 (5.04-3.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower 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	401	68% (0 types), 17% (1 type), 14% (2+ types)
1	B	401	72% (0 types), 18% (1 type), 8% (2+ types)
1	C	401	68% (0 types), 18% (1 type), 14% (2+ types)
1	D	401	73% (0 types), 18% (1 type), 8% (2+ types)
1	E	401	72% (0 types), 18% (1 type), 9% (2+ types)
1	F	401	72% (0 types), 18% (1 type), 9% (2+ types)

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	401	
1	H	401	
1	I	401	
1	J	401	
2	K	70	

## 2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 29357 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 HMPV NUCLEOPROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	346	Total 2679	C 1693	N 466	O 510	S 10	0	0	0
1	B	367	Total 2848	C 1801	N 492	O 544	S 11	0	0	0
1	C	346	Total 2682	C 1696	N 466	O 509	S 11	0	0	0
1	D	369	Total 2863	C 1809	N 495	O 548	S 11	0	0	0
1	E	364	Total 2815	C 1775	N 489	O 541	S 10	0	0	0
1	F	366	Total 2838	C 1793	N 491	O 544	S 10	0	0	0
1	G	368	Total 2850	C 1797	N 494	O 548	S 11	0	0	0
1	H	351	Total 2719	C 1718	N 473	O 517	S 11	0	0	0
1	I	366	Total 2834	C 1788	N 490	O 545	S 11	0	0	0
1	J	364	Total 2829	C 1790	N 489	O 539	S 11	0	0	0

There are 70 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	395	LYS	-	expression tag	UNP Q91F57
A	396	HIS	-	expression tag	UNP Q91F57
A	397	HIS	-	expression tag	UNP Q91F57
A	398	HIS	-	expression tag	UNP Q91F57
A	399	HIS	-	expression tag	UNP Q91F57
A	400	HIS	-	expression tag	UNP Q91F57
A	401	HIS	-	expression tag	UNP Q91F57
B	395	LYS	-	expression tag	UNP Q91F57
B	396	HIS	-	expression tag	UNP Q91F57

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	397	HIS	-	expression tag	UNP Q91F57
B	398	HIS	-	expression tag	UNP Q91F57
B	399	HIS	-	expression tag	UNP Q91F57
B	400	HIS	-	expression tag	UNP Q91F57
B	401	HIS	-	expression tag	UNP Q91F57
C	395	LYS	-	expression tag	UNP Q91F57
C	396	HIS	-	expression tag	UNP Q91F57
C	397	HIS	-	expression tag	UNP Q91F57
C	398	HIS	-	expression tag	UNP Q91F57
C	399	HIS	-	expression tag	UNP Q91F57
C	400	HIS	-	expression tag	UNP Q91F57
C	401	HIS	-	expression tag	UNP Q91F57
D	395	LYS	-	expression tag	UNP Q91F57
D	396	HIS	-	expression tag	UNP Q91F57
D	397	HIS	-	expression tag	UNP Q91F57
D	398	HIS	-	expression tag	UNP Q91F57
D	399	HIS	-	expression tag	UNP Q91F57
D	400	HIS	-	expression tag	UNP Q91F57
D	401	HIS	-	expression tag	UNP Q91F57
E	395	LYS	-	expression tag	UNP Q91F57
E	396	HIS	-	expression tag	UNP Q91F57
E	397	HIS	-	expression tag	UNP Q91F57
E	398	HIS	-	expression tag	UNP Q91F57
E	399	HIS	-	expression tag	UNP Q91F57
E	400	HIS	-	expression tag	UNP Q91F57
E	401	HIS	-	expression tag	UNP Q91F57
F	395	LYS	-	expression tag	UNP Q91F57
F	396	HIS	-	expression tag	UNP Q91F57
F	397	HIS	-	expression tag	UNP Q91F57
F	398	HIS	-	expression tag	UNP Q91F57
F	399	HIS	-	expression tag	UNP Q91F57
F	400	HIS	-	expression tag	UNP Q91F57
F	401	HIS	-	expression tag	UNP Q91F57
G	395	LYS	-	expression tag	UNP Q91F57
G	396	HIS	-	expression tag	UNP Q91F57
G	397	HIS	-	expression tag	UNP Q91F57
G	398	HIS	-	expression tag	UNP Q91F57
G	399	HIS	-	expression tag	UNP Q91F57
G	400	HIS	-	expression tag	UNP Q91F57
G	401	HIS	-	expression tag	UNP Q91F57
H	395	LYS	-	expression tag	UNP Q91F57
H	396	HIS	-	expression tag	UNP Q91F57

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	397	HIS	-	expression tag	UNP Q91F57
H	398	HIS	-	expression tag	UNP Q91F57
H	399	HIS	-	expression tag	UNP Q91F57
H	400	HIS	-	expression tag	UNP Q91F57
H	401	HIS	-	expression tag	UNP Q91F57
I	395	LYS	-	expression tag	UNP Q91F57
I	396	HIS	-	expression tag	UNP Q91F57
I	397	HIS	-	expression tag	UNP Q91F57
I	398	HIS	-	expression tag	UNP Q91F57
I	399	HIS	-	expression tag	UNP Q91F57
I	400	HIS	-	expression tag	UNP Q91F57
I	401	HIS	-	expression tag	UNP Q91F57
J	395	LYS	-	expression tag	UNP Q91F57
J	396	HIS	-	expression tag	UNP Q91F57
J	397	HIS	-	expression tag	UNP Q91F57
J	398	HIS	-	expression tag	UNP Q91F57
J	399	HIS	-	expression tag	UNP Q91F57
J	400	HIS	-	expression tag	UNP Q91F57
J	401	HIS	-	expression tag	UNP Q91F57

- Molecule 2 is a RNA chain called RNA.

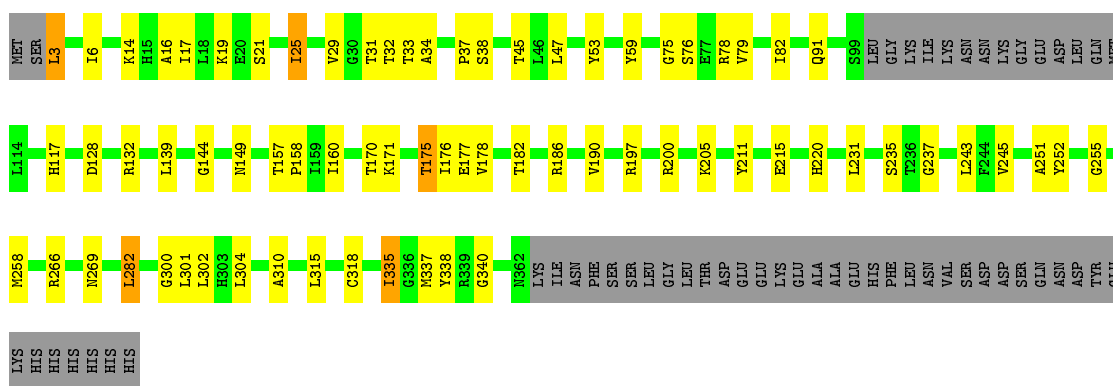
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	K	70	1400	630	210	490	70	0	0	0

### 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.

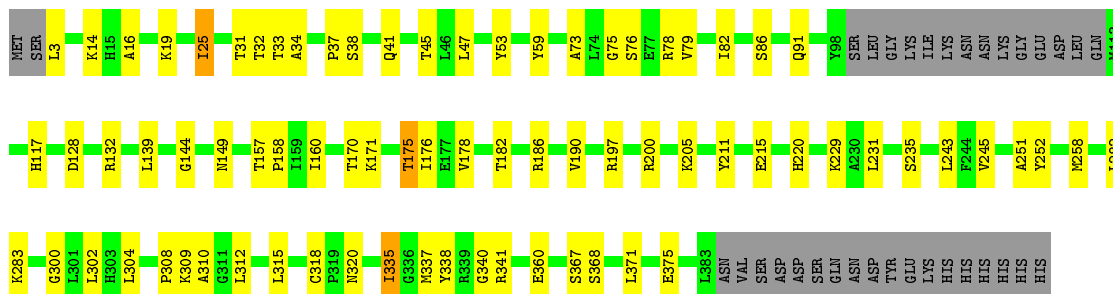
- Molecule 1: HMPV NUCLEOPROTEIN

Chain A: 



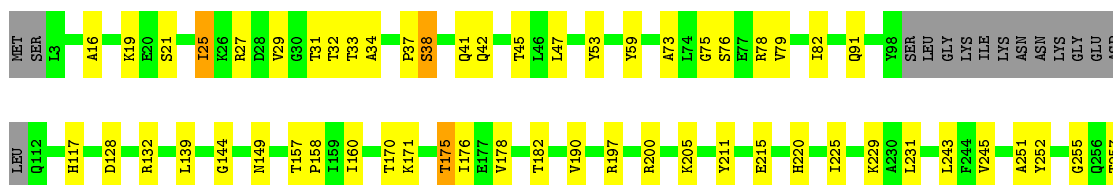
- Molecule 1: HMPV NUCLEOPROTEIN

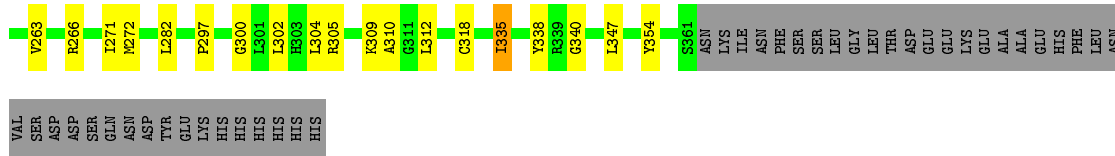
Chain B: 



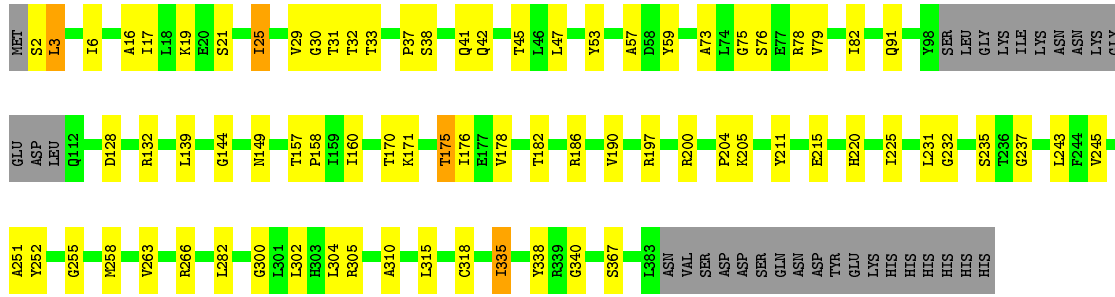
- Molecule 1: HMPV NUCLEOPROTEIN

Chain C: 

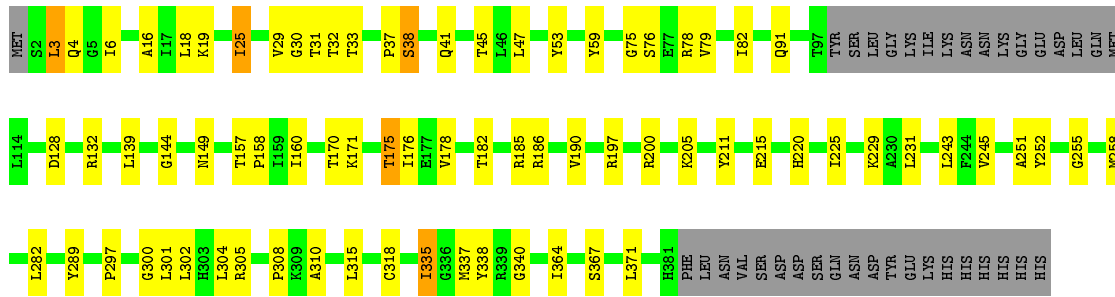




• Molecule 1: HMPV NUCLEOPROTEIN



• Molecule 1: HMPV NUCLEOPROTEIN



• Molecule 1: HMPV NUCLEOPROTEIN

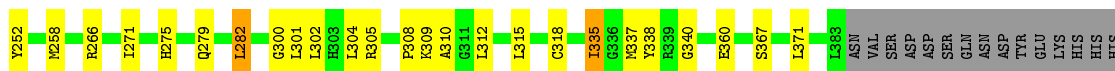


• Molecule 1: HMPV NUCLEOPROTEIN



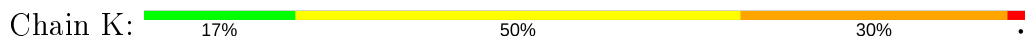






HIS  
HIS  
HIS

• Molecule 2: RNA



C64  
C65  
C66  
C67  
C68  
C69  
C70

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	202.01Å 233.21Å 203.64Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	101.19 – 4.17 101.19 – 4.17	Depositor EDS
% Data completeness (in resolution range)	99.9 (101.19-4.17) 99.9 (101.19-4.17)	Depositor EDS
$R_{merge}$	0.22	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.41 (at 4.15Å)	Xtrriage
Refinement program	BUSTER 2.10.2	Depositor
R, $R_{free}$	0.191 , 0.230 0.212 , 0.242	Depositor DCC
$R_{free}$ test set	1798 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	189.2	Xtrriage
Anisotropy	0.088	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 254.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.26$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	29357	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	215.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.54% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 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	A	0.47	0/2722	0.68	0/3678
1	B	0.45	0/2894	0.68	0/3908
1	C	0.45	0/2725	0.66	0/3681
1	D	0.47	0/2909	0.69	0/3928
1	E	0.46	0/2859	0.68	0/3861
1	F	0.45	0/2884	0.68	0/3895
1	G	0.45	0/2895	0.68	0/3909
1	H	0.46	0/2762	0.69	1/3730 (0.0%)
1	I	0.47	0/2878	0.68	1/3886 (0.0%)
1	J	0.51	1/2874 (0.0%)	0.71	3/3881 (0.1%)
2	K	1.61	21/1539 (1.4%)	1.02	0/2376
All	All	0.58	22/29941 (0.1%)	0.71	5/40733 (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
1	B	0	1

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	J	47	LEU	CG-CD1	9.19	1.85	1.51
2	K	7	C	C1'-N1	8.28	1.61	1.48
2	K	28	C	C1'-N1	7.26	1.59	1.48
2	K	14	C	C1'-N1	7.00	1.59	1.48
2	K	42	C	C1'-N1	6.91	1.59	1.48

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	J	47	LEU	CD1-CG-CD2	-11.03	77.43	110.50
1	J	47	LEU	CB-CG-CD2	-6.10	100.63	111.00
1	H	25	ILE	CB-CA-C	-5.98	99.65	111.60
1	I	306	GLN	CG-CD-OE1	5.38	132.37	121.60
1	J	47	LEU	CB-CG-CD1	-5.10	102.33	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	320	ASN	Sidechain

## 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	2679	0	2731	44	2
1	B	2848	0	2894	45	0
1	C	2682	0	2737	48	0
1	D	2863	0	2907	53	2
1	E	2815	0	2861	56	0
1	F	2838	0	2879	52	0
1	G	2850	0	2892	48	0
1	H	2719	0	2777	54	1
1	I	2834	0	2880	54	0
1	J	2829	0	2874	58	0
2	K	1400	0	770	42	0
All	All	29357	0	29202	427	3

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 427 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:J:47:LEU:CG	1:J:47:LEU:CD1	1.85	1.52

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:J:47:LEU:CD1	1:J:47:LEU:CD2	2.14	1.24
1:A:3:LEU:HD23	1:J:282:LEU:HD11	1.31	1.11
1:E:41:GLN:HB3	1:F:29:VAL:HG11	1.31	1.07
1:I:41:GLN:HB3	1:J:29:VAL:HG11	1.39	1.04

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:337:MET:SD	1:D:57:ALA:O[7_747]	1.93	0.27
1:A:177:GLU:OE2	1:D:204:PRO:CG[7_747]	2.01	0.19
1:H:145:ASN:OD1	1:H:145:ASN:OD1[4_557]	2.02	0.18

## 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 X-ray entries followed by that with respect to entries of similar resolution.

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	342/401 (85%)	321 (94%)	20 (6%)	1 (0%)	41 75
1	B	363/401 (90%)	342 (94%)	20 (6%)	1 (0%)	41 75
1	C	342/401 (85%)	321 (94%)	20 (6%)	1 (0%)	41 75
1	D	365/401 (91%)	342 (94%)	22 (6%)	1 (0%)	41 75
1	E	360/401 (90%)	337 (94%)	22 (6%)	1 (0%)	41 75
1	F	362/401 (90%)	341 (94%)	20 (6%)	1 (0%)	41 75
1	G	364/401 (91%)	343 (94%)	20 (6%)	1 (0%)	41 75
1	H	347/401 (86%)	325 (94%)	21 (6%)	1 (0%)	41 75
1	I	362/401 (90%)	341 (94%)	20 (6%)	1 (0%)	41 75
1	J	358/401 (89%)	336 (94%)	21 (6%)	1 (0%)	41 75
All	All	3565/4010 (89%)	3349 (94%)	206 (6%)	10 (0%)	41 75

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	335	ILE
1	B	335	ILE
1	D	335	ILE
1	E	335	ILE
1	F	335	ILE

### 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 X-ray entries followed by that with respect to entries of similar resolution.

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	293/343 (85%)	282 (96%)	11 (4%)	33	58
1	B	311/343 (91%)	300 (96%)	11 (4%)	36	60
1	C	293/343 (85%)	284 (97%)	9 (3%)	40	62
1	D	313/343 (91%)	301 (96%)	12 (4%)	33	58
1	E	308/343 (90%)	296 (96%)	12 (4%)	32	57
1	F	310/343 (90%)	298 (96%)	12 (4%)	32	57
1	G	312/343 (91%)	301 (96%)	11 (4%)	36	60
1	H	298/343 (87%)	288 (97%)	10 (3%)	37	60
1	I	310/343 (90%)	300 (97%)	10 (3%)	39	62
1	J	309/343 (90%)	298 (96%)	11 (4%)	35	60
All	All	3057/3430 (89%)	2948 (96%)	109 (4%)	35	60

5 of 109 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	215	GLU
1	F	235	SER
1	J	38	SER
1	E	282	LEU
1	F	25	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such

sidechains are listed below:

Mol	Chain	Res	Type
1	D	320	ASN
1	E	320	ASN
1	H	320	ASN
1	C	320	ASN
1	G	320	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	K	69/70 (98%)	23 (33%)	7 (10%)

5 of 23 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	K	4	C
2	K	8	C
2	K	11	C
2	K	12	C
2	K	15	C

5 of 7 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	K	32	C
2	K	67	C
2	K	39	C
2	K	18	C
2	K	46	C

## 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.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.