



wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 12:17 pm BST

PDB ID : 2WJ8
Title : Respiratory Syncytial Virus RiboNucleoProtein
Authors : Tawar, R.G.; Duquerroy, S.; Vonrhein, C.; Varela, P.F.; Damier-Piolle, L.; Castagne, N.; MacLellan, K.; Bedouelle, H.; Bricogne, G.; Bhella, D.; Eleouet, J.; Rey, F.A.
Deposited on : 2009-05-25
Resolution : 3.29 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

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
Mogul : 1.8.5 (274361), CSD as541be (2020)
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

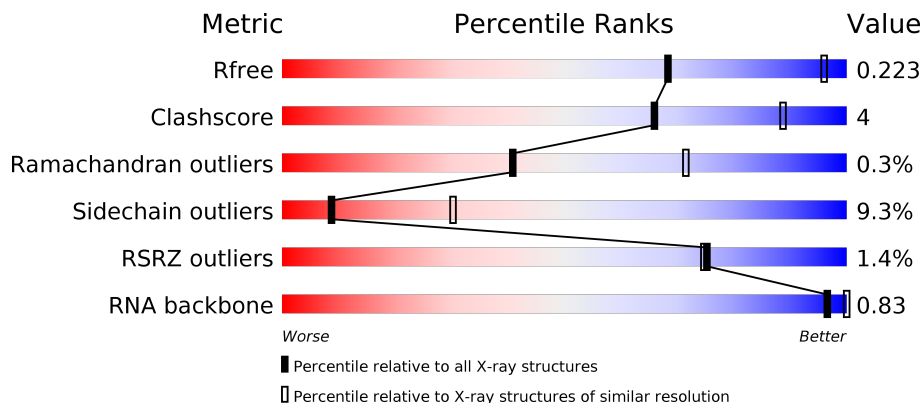
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.29 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)
RNA backbone	3102	1117 (3.70-2.90)

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 ≥ 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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	391	
1	B	391	
1	C	391	
1	D	391	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	E	391	
1	F	391	
1	G	391	
1	H	391	
1	I	391	
1	J	391	
1	K	391	
1	L	391	
1	M	391	
1	N	391	
1	O	391	
1	P	391	
1	Q	391	
1	R	391	
1	S	391	
1	T	391	
2	a	7	
2	b	7	
2	c	7	
2	d	7	
2	e	7	
2	f	7	
2	g	7	
2	h	7	
2	i	7	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	j	7	 29% 71%
2	k	7	 29% 71%
2	l	7	 29% 71%
2	m	7	 29% 71%
2	n	7	 29% 71%
2	o	7	 29% 71%
2	p	7	 29% 71%
2	q	7	 29% 71%
2	r	7	 29% 71%
2	s	7	 29% 71%
2	t	7	 29% 71%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 61928 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 NUCLEOPROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	374	2912	1843	505	548	16	0	0	0
1	B	375	2921	1848	506	551	16	0	0	0
1	C	377	2934	1857	508	553	16	0	0	0
1	D	374	2912	1843	505	548	16	0	0	0
1	E	374	2912	1843	505	548	16	0	0	0
1	F	377	2934	1857	508	553	16	0	0	0
1	G	377	2934	1857	508	553	16	0	0	0
1	H	375	2921	1848	506	551	16	0	0	0
1	I	374	2912	1843	505	548	16	0	0	0
1	J	370	2881	1824	501	540	16	0	0	0
1	K	374	2912	1843	505	548	16	0	0	0
1	L	377	2934	1857	508	553	16	0	0	0
1	M	374	2912	1843	505	548	16	0	0	0
1	N	378	2943	1863	510	554	16	0	0	0
1	O	373	2904	1837	504	547	16	0	0	0
1	P	374	2912	1843	505	548	16	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Q	374	Total	C	N	O	S	0	0	0
			2912	1843	505	548	16			
1	R	376	Total	C	N	O	S	0	0	0
			2926	1851	507	552	16			
1	S	378	Total	C	N	O	S	0	0	0
			2943	1863	510	554	16			
1	T	374	Total	C	N	O	S	0	0	0
			2912	1843	505	548	16			

- Molecule 2 is a RNA chain called RNA (5'-R(*CP*CP*CP*CP*CP*C)-3').

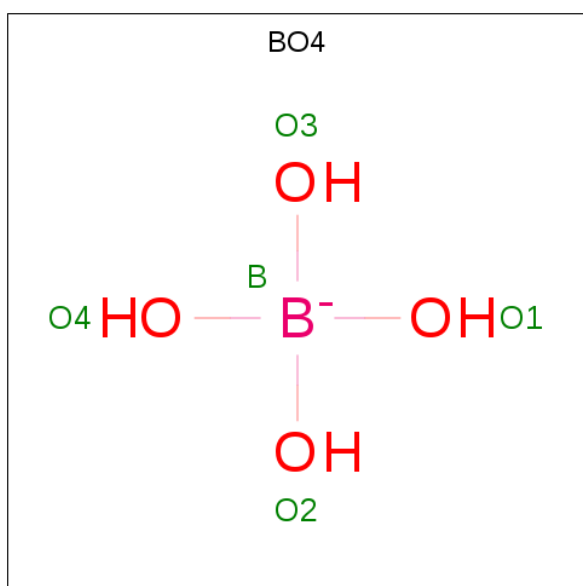
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	a	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			
2	b	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			
2	c	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			
2	d	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			
2	e	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			
2	f	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			
2	g	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			
2	h	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			
2	i	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			
2	j	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			
2	k	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			
2	l	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			
2	m	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			
2	n	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			
2	o	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	p	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			
2	q	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			
2	r	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			
2	s	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			
2	t	7	Total	C	N	O	P	0	0	0
			140	63	21	49	7			

- Molecule 3 is BORATE ION (three-letter code: BO4) (formula: BH_4O_4).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	B O	0	0
			5	1 4		
3	B	1	Total	B O	0	0
			5	1 4		
3	C	1	Total	B O	0	0
			5	1 4		
3	D	1	Total	B O	0	0
			5	1 4		
3	H	1	Total	B O	0	0
			5	1 4		
3	I	1	Total	B O	0	0
			5	1 4		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	K	1	Total B O 5 1 4	0	0
3	L	1	Total B O 5 1 4	0	0
3	M	1	Total B O 5 1 4	0	0
3	N	1	Total B O 5 1 4	0	0
3	O	1	Total B O 5 1 4	0	0
3	T	1	Total B O 5 1 4	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	29	Total O 29 29	0	0
4	B	30	Total O 30 30	0	0
4	C	38	Total O 38 38	0	0
4	D	45	Total O 45 45	0	0
4	E	33	Total O 33 33	0	0
4	F	39	Total O 39 39	0	0
4	G	25	Total O 25 25	0	0
4	H	57	Total O 57 57	0	0
4	I	48	Total O 48 48	0	0
4	J	20	Total O 20 20	0	0
4	K	41	Total O 41 41	0	0
4	L	23	Total O 23 23	0	0
4	M	22	Total O 22 22	0	0

Continued on next page...

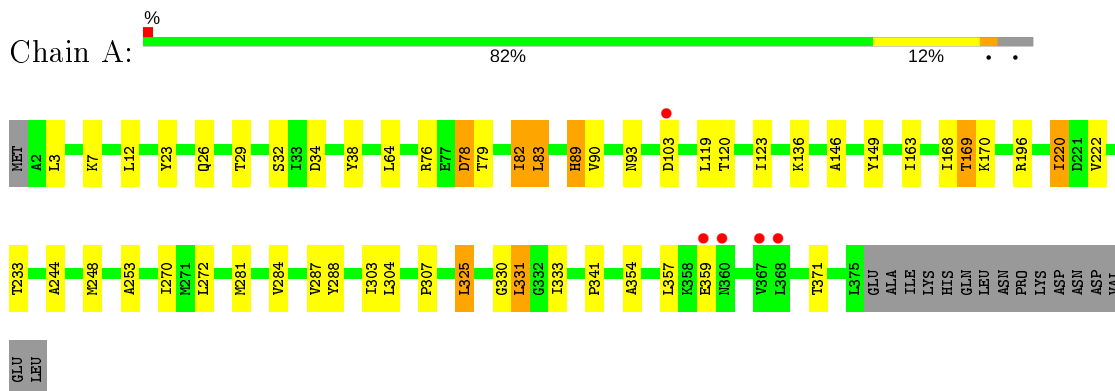
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	N	40	Total 40	O 40	0	0
4	O	34	Total 34	O 34	0	0
4	P	20	Total 20	O 20	0	0
4	Q	22	Total 22	O 22	0	0
4	R	40	Total 40	O 40	0	0
4	S	37	Total 37	O 37	0	0
4	T	42	Total 42	O 42	0	0

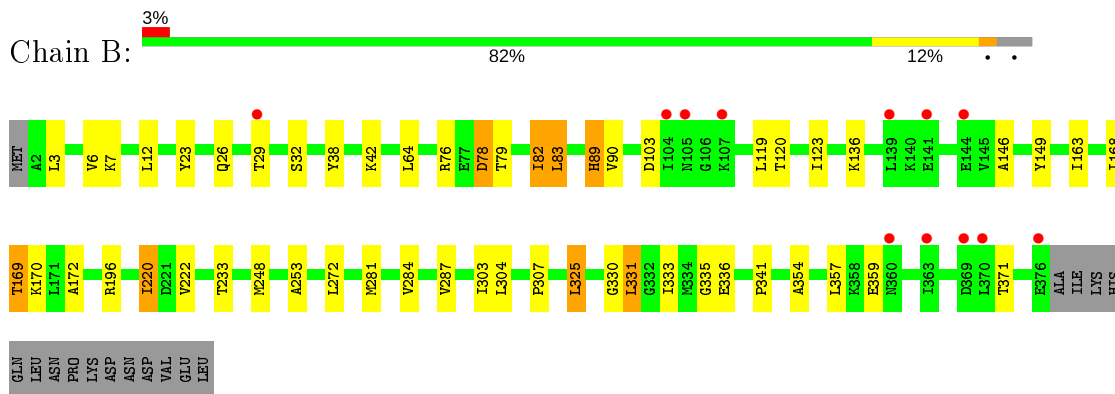
3 Residue-property plots [i](#)

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). 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.

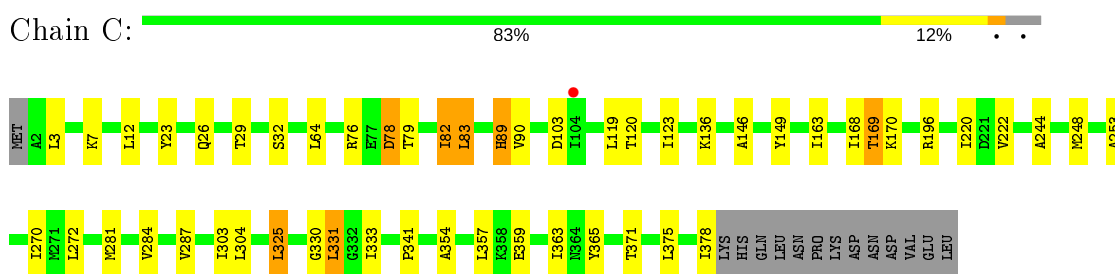
- Molecule 1: NUCLEOPROTEIN



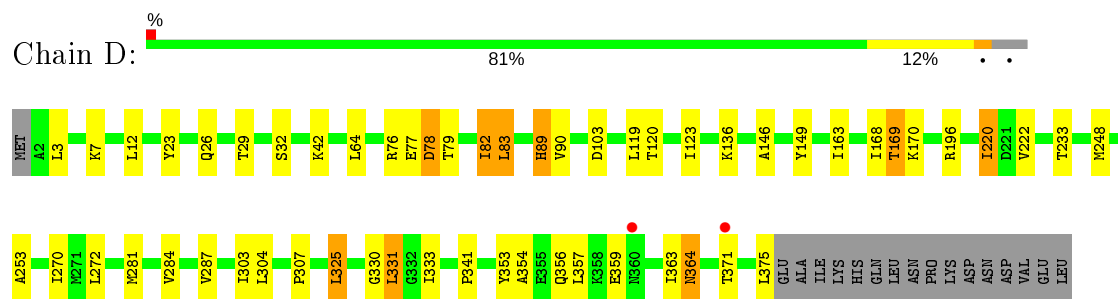
- Molecule 1: NUCLEOPROTEIN



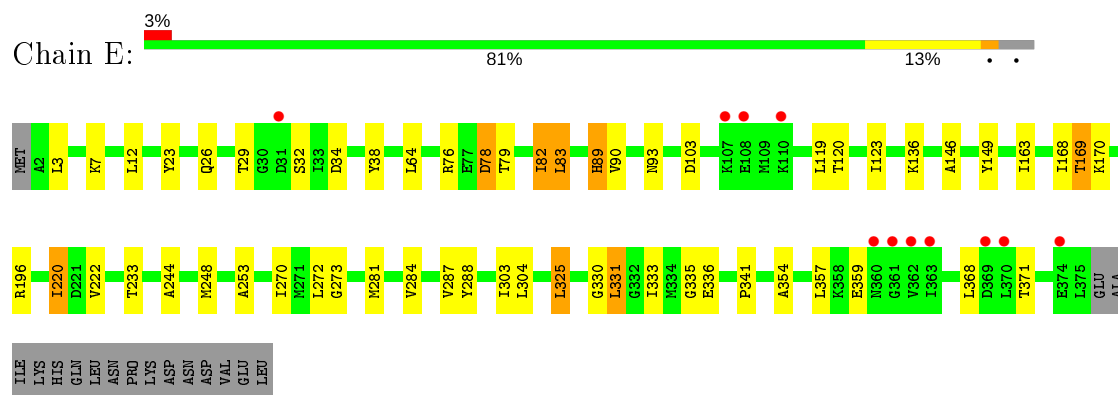
- Molecule 1: NUCLEOPROTEIN



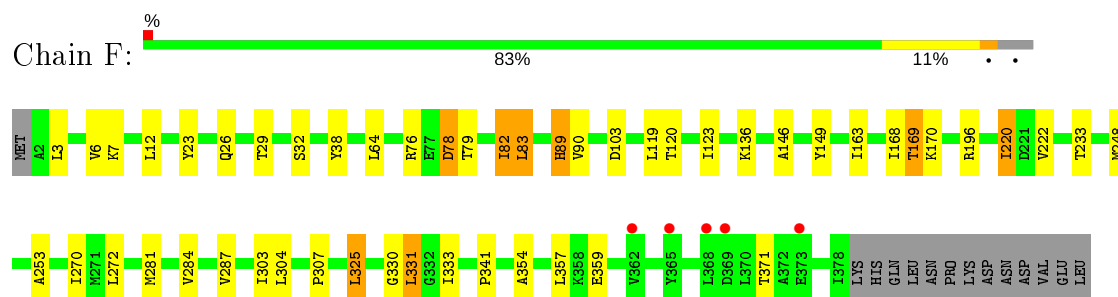
- Molecule 1: NUCLEOPROTEIN



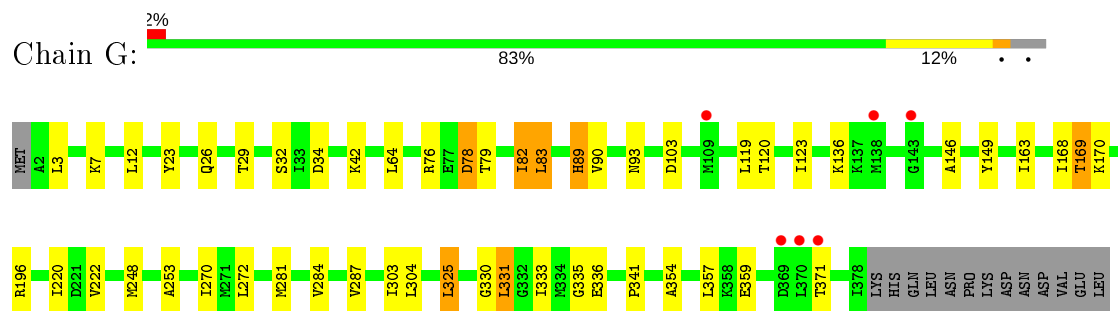
- Molecule 1: NUCLEOPROTEIN



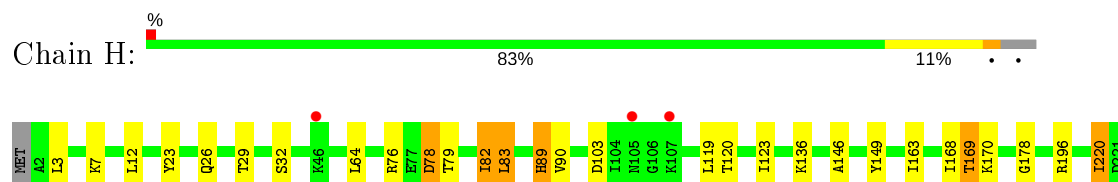
- Molecule 1: NUCLEOPROTEIN



- Molecule 1: NUCLEOPROTEIN

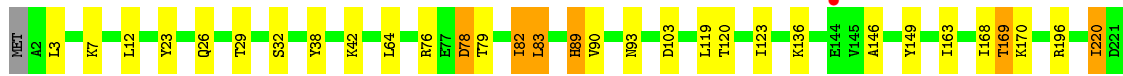
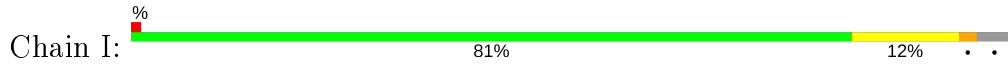


- Molecule 1: NUCLEOPROTEIN

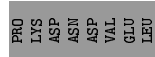
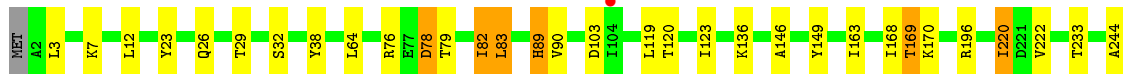
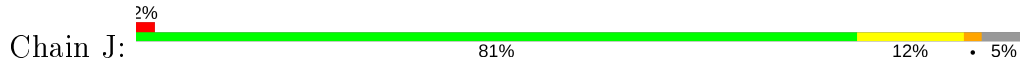




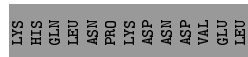
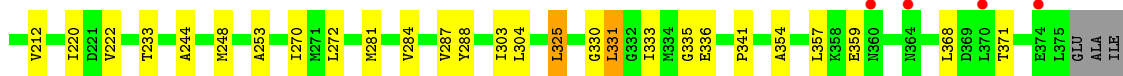
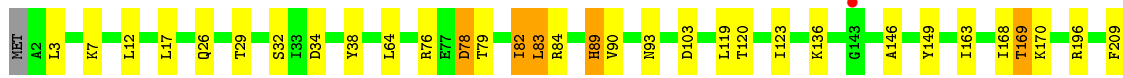
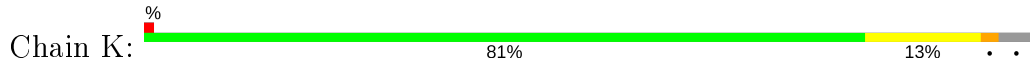
• Molecule 1: NUCLEOPROTEIN



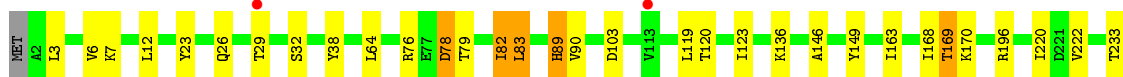
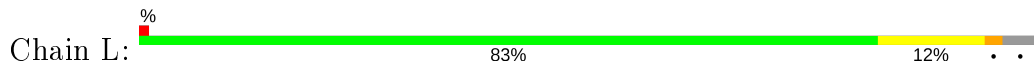
• Molecule 1: NUCLEOPROTEIN



• Molecule 1: NUCLEOPROTEIN

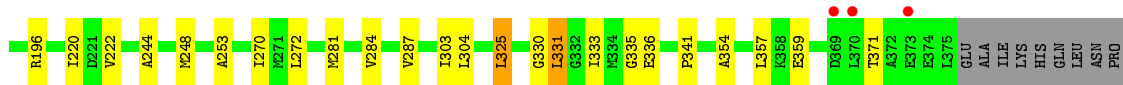
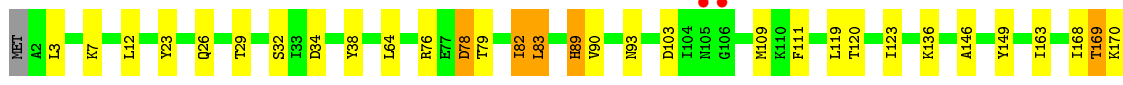
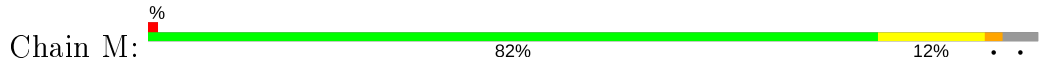


• Molecule 1: NUCLEOPROTEIN



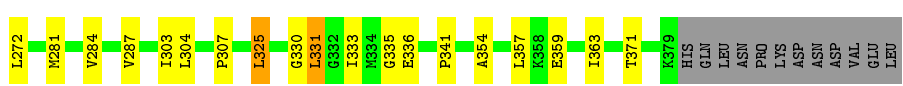
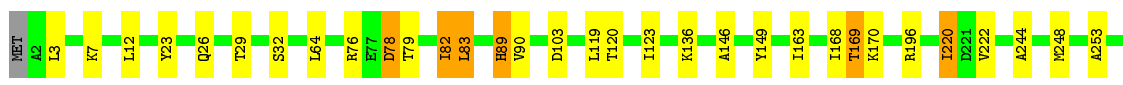
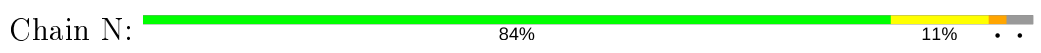


• Molecule 1: NUCLEOPROTEIN

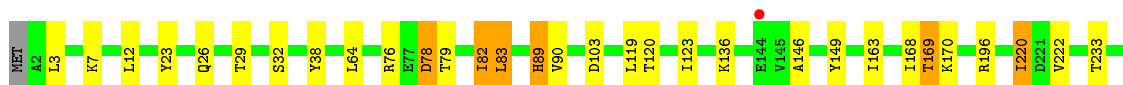
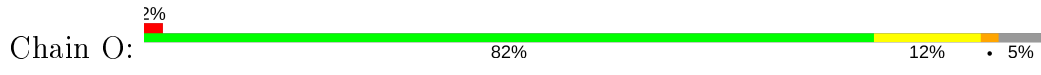


LYS
ASP
ASN
ASP
VAL
GLU
LEU

• Molecule 1: NUCLEOPROTEIN

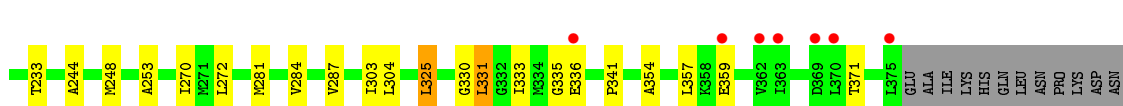
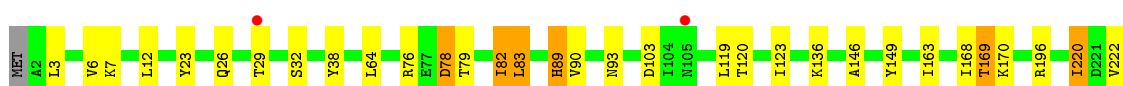
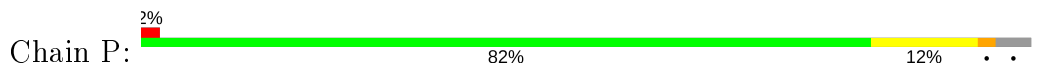


• Molecule 1: NUCLEOPROTEIN



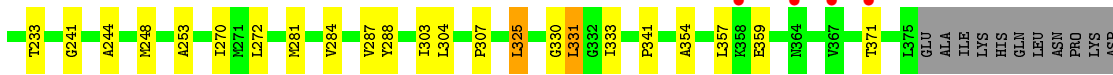
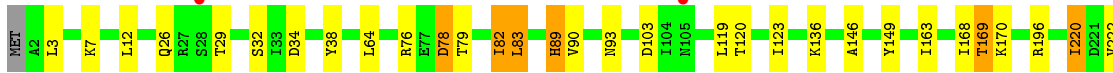
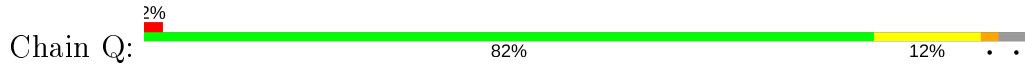
ASN
ASP
VAL
GLU
LEU

• Molecule 1: NUCLEOPROTEIN



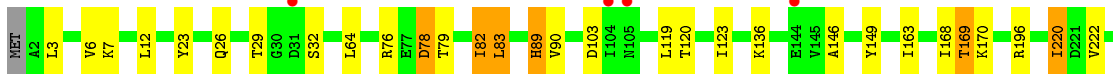
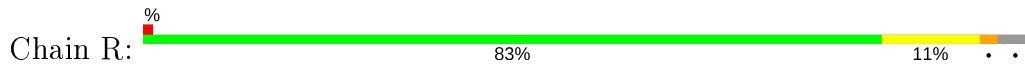
ASP
VAL
GLU
LEU

• Molecule 1: NUCLEOPROTEIN

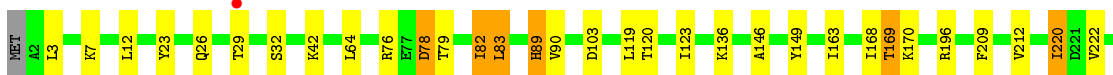
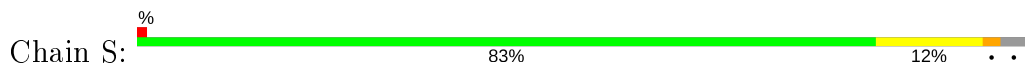


ASN
ASP
VAL
GLU
LEU

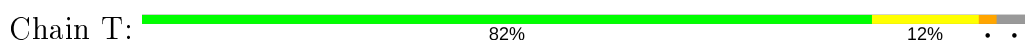
• Molecule 1: NUCLEOPROTEIN



• Molecule 1: NUCLEOPROTEIN



• Molecule 1: NUCLEOPROTEIN



• Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain a:  29% 71%

C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain b:  29% 71%

C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain c:  29% 71%

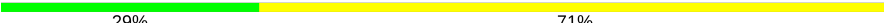
C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain d:  29% 71%

C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain e:  29% 71%

C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain f:  29% 71%

C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain g:  29% 71%

C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain h:  29% 71%

C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain i:  29% 71%

C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain j:  29% 71%

C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain k:  29% 71%

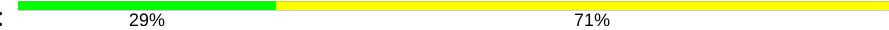
C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain l:  29% 71%


C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain m:  29% 71%

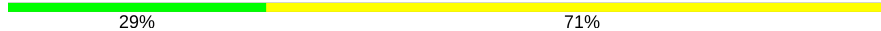
C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain n:  29% 71%

C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain o:  29% 71%


C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain p:  29% 71%

C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain q:  29% 71%


C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain r:  29% 71%

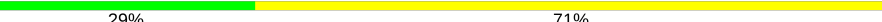
C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain s:  29% 71%

C1001
C1002
C1003
C1004
C1005
C1006
C1007

- Molecule 2: RNA (5'-R(*CP*CP*CP*CP*CP*C)-3')

Chain t:  29% 71%

C1001
C1002
C1003
C1004
C1005
C1006
C1007

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	218.20Å 220.99Å 218.10Å 90.00° 93.09° 90.00°	Depositor
Resolution (Å)	79.18 – 3.29 79.18 – 3.30	Depositor EDS
% Data completeness (in resolution range)	96.7 (79.18-3.29) 96.7 (79.18-3.30)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.03 (at 3.33Å)	Xtrriage
Refinement program	BUSTER-TNT 2.7.0	Depositor
R, R_{free}	0.205 , 0.226 0.215 , 0.223	Depositor DCC
R_{free} test set	5093 reflections (3.39%)	wwPDB-VP
Wilson B-factor (Å ²)	68.3	Xtrriage
Anisotropy	0.465	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 40.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.26$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	61928	wwPDB-VP
Average B, all atoms (Å ²)	71.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.48% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: BO4

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	0.43	0/2960	0.67	0/3988
1	B	0.42	0/2969	0.67	0/4000
1	C	0.43	0/2982	0.68	0/4018
1	D	0.44	0/2960	0.69	1/3988 (0.0%)
1	E	0.43	0/2960	0.68	0/3988
1	F	0.43	0/2982	0.67	0/4018
1	G	0.43	0/2982	0.68	0/4018
1	H	0.43	0/2969	0.68	0/4000
1	I	0.43	0/2960	0.67	0/3988
1	J	0.43	0/2929	0.67	0/3946
1	K	0.43	0/2960	0.68	0/3988
1	L	0.43	0/2982	0.67	0/4018
1	M	0.43	0/2960	0.68	0/3988
1	N	0.43	0/2991	0.68	0/4029
1	O	0.43	0/2952	0.67	0/3977
1	P	0.43	0/2960	0.67	0/3988
1	Q	0.43	0/2960	0.67	0/3988
1	R	0.43	0/2974	0.68	0/4007
1	S	0.43	0/2991	0.68	0/4029
1	T	0.43	0/2960	0.67	0/3988
2	a	0.50	0/153	2.44	9/234 (3.8%)
2	b	0.50	0/153	2.45	9/234 (3.8%)
2	c	0.51	0/153	2.44	9/234 (3.8%)
2	d	0.55	0/153	2.45	9/234 (3.8%)
2	e	0.53	0/153	2.46	9/234 (3.8%)
2	f	0.51	0/153	2.44	9/234 (3.8%)
2	g	0.51	0/153	2.45	9/234 (3.8%)
2	h	0.51	0/153	2.44	9/234 (3.8%)
2	i	0.51	0/153	2.45	9/234 (3.8%)
2	j	0.50	0/153	2.45	9/234 (3.8%)
2	k	0.54	0/153	2.45	9/234 (3.8%)
2	l	0.50	0/153	2.44	9/234 (3.8%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
2	m	0.50	0/153	2.44	9/234 (3.8%)
2	n	0.51	0/153	2.45	9/234 (3.8%)
2	o	0.49	0/153	2.44	9/234 (3.8%)
2	p	0.52	0/153	2.45	9/234 (3.8%)
2	q	0.48	0/153	2.44	9/234 (3.8%)
2	r	0.48	0/153	2.44	9/234 (3.8%)
2	s	0.49	0/153	2.44	9/234 (3.8%)
2	t	0.55	0/153	2.45	9/234 (3.8%)
All	All	0.43	0/62403	0.87	181/84632 (0.2%)

There are no bond length outliers.

The worst 5 of 181 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	p	1002	C	C2-N1-C1'	-16.88	100.23	118.80
2	b	1002	C	C2-N1-C1'	-16.88	100.23	118.80
2	o	1002	C	C2-N1-C1'	-16.88	100.24	118.80
2	j	1002	C	C2-N1-C1'	-16.86	100.26	118.80
2	i	1002	C	C2-N1-C1'	-16.86	100.26	118.80

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	2912	0	2943	25	0
1	B	2921	0	2949	24	0
1	C	2934	0	2965	27	0
1	D	2912	0	2943	29	0
1	E	2912	0	2943	27	0
1	F	2934	0	2965	24	0
1	G	2934	0	2965	21	0
1	H	2921	0	2949	21	0
1	I	2912	0	2943	24	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	J	2881	0	2915	25	0
1	K	2912	0	2943	27	0
1	L	2934	0	2965	23	0
1	M	2912	0	2943	25	0
1	N	2943	0	2978	23	0
1	O	2904	0	2932	24	0
1	P	2912	0	2943	23	1
1	Q	2912	0	2943	25	0
1	R	2926	0	2954	24	0
1	S	2943	0	2978	24	0
1	T	2912	0	2943	23	0
2	a	140	0	77	0	0
2	b	140	0	77	0	0
2	c	140	0	77	0	0
2	d	140	0	77	0	0
2	e	140	0	77	0	0
2	f	140	0	77	0	0
2	g	140	0	77	0	0
2	h	140	0	77	0	0
2	i	140	0	77	0	0
2	j	140	0	77	0	0
2	k	140	0	77	0	0
2	l	140	0	77	0	0
2	m	140	0	77	0	0
2	n	140	0	77	0	0
2	o	140	0	77	0	0
2	p	140	0	77	0	0
2	q	140	0	77	0	0
2	r	140	0	77	0	0
2	s	140	0	77	0	0
2	t	140	0	77	0	0
3	A	5	0	4	0	0
3	B	5	0	4	0	0
3	C	5	0	4	0	0
3	D	5	0	4	0	0
3	H	5	0	4	0	0
3	I	5	0	4	0	1
3	K	5	0	4	0	0
3	L	5	0	4	0	0
3	M	5	0	4	0	0
3	N	5	0	4	0	1
3	O	5	0	4	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	T	5	0	4	0	0
4	A	29	0	0	0	0
4	B	30	0	0	1	0
4	C	38	0	0	0	0
4	D	45	0	0	1	0
4	E	33	0	0	0	0
4	F	39	0	0	1	0
4	G	25	0	0	0	0
4	H	57	0	0	1	0
4	I	48	0	0	0	0
4	J	20	0	0	0	0
4	K	41	0	0	1	0
4	L	23	0	0	0	0
4	M	22	0	0	0	0
4	N	40	0	0	0	0
4	O	34	0	0	0	0
4	P	20	0	0	0	0
4	Q	22	0	0	0	0
4	R	40	0	0	1	0
4	S	37	0	0	0	0
4	T	42	0	0	1	0
All	All	61928	0	60590	422	2

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 4.

The worst 5 of 422 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:D:364:ASN:OD1	1:E:273:GLY:O	2.09	0.70
1:C:375:LEU:HD22	1:D:353:TYR:HE1	1.58	0.68
1:C:82:ILE:HD11	1:C:222:VAL:HA	1.77	0.67
1:N:82:ILE:HD11	1:N:222:VAL:HA	1.76	0.67
1:Q:82:ILE:HD11	1:Q:222:VAL:HA	1.77	0.67

All (2) 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:I:93:ASN:OD1	3:I:2001:BO4:O2[2_557]	2.07	0.13

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:P:93:ASN:OD1	3:N:2001:BO4:O2[2_656]	2.12	0.08

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	372/391 (95%)	356 (96%)	15 (4%)	1 (0%)	41 71
1	B	373/391 (95%)	357 (96%)	15 (4%)	1 (0%)	41 71
1	C	375/391 (96%)	359 (96%)	15 (4%)	1 (0%)	41 71
1	D	372/391 (95%)	356 (96%)	15 (4%)	1 (0%)	41 71
1	E	372/391 (95%)	356 (96%)	15 (4%)	1 (0%)	41 71
1	F	375/391 (96%)	359 (96%)	15 (4%)	1 (0%)	41 71
1	G	375/391 (96%)	358 (96%)	16 (4%)	1 (0%)	41 71
1	H	373/391 (95%)	357 (96%)	15 (4%)	1 (0%)	41 71
1	I	372/391 (95%)	356 (96%)	15 (4%)	1 (0%)	41 71
1	J	368/391 (94%)	353 (96%)	14 (4%)	1 (0%)	41 71
1	K	372/391 (95%)	356 (96%)	15 (4%)	1 (0%)	41 71
1	L	375/391 (96%)	359 (96%)	15 (4%)	1 (0%)	41 71
1	M	372/391 (95%)	356 (96%)	15 (4%)	1 (0%)	41 71
1	N	376/391 (96%)	360 (96%)	15 (4%)	1 (0%)	41 71
1	O	371/391 (95%)	355 (96%)	15 (4%)	1 (0%)	41 71
1	P	372/391 (95%)	356 (96%)	15 (4%)	1 (0%)	41 71
1	Q	372/391 (95%)	356 (96%)	15 (4%)	1 (0%)	41 71
1	R	374/391 (96%)	358 (96%)	15 (4%)	1 (0%)	41 71
1	S	376/391 (96%)	360 (96%)	15 (4%)	1 (0%)	41 71
1	T	372/391 (95%)	356 (96%)	15 (4%)	1 (0%)	41 71

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	7459/7820 (95%)	7139 (96%)	300 (4%)	20 (0%)	41 71

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	76	ARG
1	B	76	ARG
1	C	76	ARG
1	D	76	ARG
1	E	76	ARG

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	312/328 (95%)	283 (91%)	29 (9%)	9 30
1	B	313/328 (95%)	283 (90%)	30 (10%)	8 29
1	C	314/328 (96%)	286 (91%)	28 (9%)	9 32
1	D	312/328 (95%)	281 (90%)	31 (10%)	8 28
1	E	312/328 (95%)	283 (91%)	29 (9%)	9 30
1	F	314/328 (96%)	285 (91%)	29 (9%)	9 31
1	G	314/328 (96%)	285 (91%)	29 (9%)	9 31
1	H	313/328 (95%)	284 (91%)	29 (9%)	9 30
1	I	312/328 (95%)	282 (90%)	30 (10%)	8 29
1	J	309/328 (94%)	280 (91%)	29 (9%)	8 30
1	K	312/328 (95%)	283 (91%)	29 (9%)	9 30
1	L	314/328 (96%)	285 (91%)	29 (9%)	9 31
1	M	312/328 (95%)	284 (91%)	28 (9%)	9 32
1	N	315/328 (96%)	287 (91%)	28 (9%)	9 32
1	O	311/328 (95%)	282 (91%)	29 (9%)	9 30
1	P	312/328 (95%)	283 (91%)	29 (9%)	9 30

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	Q	312/328 (95%)	283 (91%)	29 (9%)	9	30
1	R	313/328 (95%)	285 (91%)	28 (9%)	9	32
1	S	315/328 (96%)	286 (91%)	29 (9%)	9	31
1	T	312/328 (95%)	283 (91%)	29 (9%)	9	30
All	All	6253/6560 (95%)	5673 (91%)	580 (9%)	9	30

5 of 580 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	359	GLU
1	L	12	LEU
1	S	136	LYS
1	J	79	THR
1	K	7	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
1	J	187	ASN
1	L	89	HIS
1	S	360	ASN
1	K	89	HIS
1	L	187	ASN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	a	6/7 (85%)	0	0
2	b	6/7 (85%)	0	0
2	c	6/7 (85%)	0	0
2	d	6/7 (85%)	0	0
2	e	6/7 (85%)	0	0
2	f	6/7 (85%)	0	0
2	g	6/7 (85%)	0	0
2	h	6/7 (85%)	0	0
2	i	6/7 (85%)	0	0
2	j	6/7 (85%)	0	0
2	k	6/7 (85%)	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	l	6/7 (85%)	0	0
2	m	6/7 (85%)	0	0
2	n	6/7 (85%)	0	0
2	o	6/7 (85%)	0	0
2	p	6/7 (85%)	0	0
2	q	6/7 (85%)	0	0
2	r	6/7 (85%)	0	0
2	s	6/7 (85%)	0	0
2	t	6/7 (85%)	0	0
All	All	120/140 (85%)	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 [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](#)

12 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
3	BO4	M	2001	-	4,4,4	0.43	0	6,6,6	0.48	0
3	BO4	I	2001	-	4,4,4	0.47	0	6,6,6	0.61	0
3	BO4	D	2001	-	4,4,4	0.49	0	6,6,6	0.46	0
3	BO4	B	2001	-	4,4,4	0.41	0	6,6,6	0.49	0
3	BO4	L	2001	-	4,4,4	0.49	0	6,6,6	0.40	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	BO4	N	2001	-	4,4,4	0.46	0	6,6,6	0.55	0
3	BO4	H	2001	-	4,4,4	0.41	0	6,6,6	0.49	0
3	BO4	C	2001	-	4,4,4	0.46	0	6,6,6	0.49	0
3	BO4	T	2001	-	4,4,4	0.45	0	6,6,6	0.44	0
3	BO4	O	2001	-	4,4,4	0.46	0	6,6,6	0.45	0
3	BO4	A	2001	-	4,4,4	0.47	0	6,6,6	0.46	0
3	BO4	K	2001	-	4,4,4	0.49	0	6,6,6	0.52	0

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.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	I	2001	BO4	0	1
3	N	2001	BO4	0	1

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled '#RSRZ > 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q < 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	374/391 (95%)	-0.08	5 (1%) 77 77	43, 71, 125, 162	0
1	B	375/391 (95%)	0.05	12 (3%) 47 46	43, 68, 121, 166	0
1	C	377/391 (96%)	-0.07	1 (0%) 94 94	36, 62, 118, 140	0
1	D	374/391 (95%)	-0.14	2 (0%) 91 91	36, 62, 123, 154	0
1	E	374/391 (95%)	-0.04	11 (2%) 51 50	33, 60, 118, 157	0
1	F	377/391 (96%)	-0.07	5 (1%) 77 77	41, 67, 126, 153	0
1	G	377/391 (96%)	-0.04	6 (1%) 72 70	44, 68, 121, 146	0
1	H	375/391 (95%)	-0.02	5 (1%) 77 77	37, 63, 119, 154	0
1	I	374/391 (95%)	-0.09	4 (1%) 80 81	40, 70, 120, 159	0
1	J	370/391 (94%)	0.01	8 (2%) 62 60	36, 67, 116, 158	0
1	K	374/391 (95%)	-0.08	5 (1%) 77 77	35, 59, 115, 154	0
1	L	377/391 (96%)	-0.03	5 (1%) 77 77	41, 68, 126, 154	0
1	M	374/391 (95%)	-0.06	5 (1%) 77 77	44, 68, 119, 151	0
1	N	378/391 (96%)	-0.10	0 100 100	35, 63, 124, 153	0
1	O	373/391 (95%)	-0.04	6 (1%) 72 70	45, 69, 127, 161	0
1	P	374/391 (95%)	0.01	9 (2%) 59 56	43, 67, 124, 168	0
1	Q	374/391 (95%)	-0.05	6 (1%) 72 70	49, 71, 127, 141	0
1	R	376/391 (96%)	-0.06	5 (1%) 77 77	37, 71, 121, 147	0
1	S	378/391 (96%)	-0.08	3 (0%) 86 86	40, 62, 112, 142	0
1	T	374/391 (95%)	-0.14	1 (0%) 94 94	40, 62, 116, 145	0
2	a	7/7 (100%)	-0.26	0 100 100	72, 74, 76, 76	0
2	b	7/7 (100%)	-0.33	0 100 100	72, 72, 80, 81	0
2	c	7/7 (100%)	-0.26	0 100 100	62, 64, 69, 74	0
2	d	7/7 (100%)	-0.39	0 100 100	58, 61, 63, 65	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
2	e	7/7 (100%)	-0.12	0 100 100	48, 51, 55, 59	0
2	f	7/7 (100%)	-0.32	0 100 100	65, 67, 74, 79	0
2	g	7/7 (100%)	-0.15	0 100 100	68, 70, 83, 83	0
2	h	7/7 (100%)	-0.19	0 100 100	63, 67, 71, 73	0
2	i	7/7 (100%)	-0.32	0 100 100	63, 68, 72, 75	0
2	j	7/7 (100%)	-0.30	0 100 100	63, 65, 75, 76	0
2	k	7/7 (100%)	-0.17	0 100 100	51, 56, 58, 62	0
2	l	7/7 (100%)	-0.36	0 100 100	65, 70, 72, 73	0
2	m	7/7 (100%)	-0.12	0 100 100	72, 73, 75, 75	0
2	n	7/7 (100%)	-0.21	0 100 100	54, 61, 71, 72	0
2	o	7/7 (100%)	-0.21	0 100 100	68, 70, 74, 75	0
2	p	7/7 (100%)	-0.29	0 100 100	59, 66, 76, 78	0
2	q	7/7 (100%)	-0.26	0 100 100	70, 77, 81, 83	0
2	r	7/7 (100%)	-0.24	0 100 100	72, 75, 77, 78	0
2	s	7/7 (100%)	-0.03	0 100 100	64, 66, 72, 72	0
2	t	7/7 (100%)	-0.23	0 100 100	65, 65, 69, 71	0
All	All	7639/7960 (95%)	-0.06	104 (1%) 75 75	33, 66, 122, 168	0

The worst 5 of 104 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	108	GLU	5.5
1	E	361	GLY	5.2
1	H	105	ASN	5.2
1	E	362	VAL	5.0
1	R	104	ILE	4.5

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q<0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
3	BO4	M	2001	5/5	0.76	0.23	86,87,88,89	0
3	BO4	O	2001	5/5	0.86	0.14	123,124,124,124	5
3	BO4	I	2001	5/5	0.87	0.17	92,92,93,93	5
3	BO4	K	2001	5/5	0.87	0.24	75,75,76,76	0
3	BO4	N	2001	5/5	0.88	0.20	89,90,90,90	0
3	BO4	C	2001	5/5	0.88	0.20	91,91,92,93	0
3	BO4	B	2001	5/5	0.88	0.17	94,94,95,96	0
3	BO4	L	2001	5/5	0.88	0.20	100,101,101,101	0
3	BO4	H	2001	5/5	0.90	0.18	76,77,78,78	0
3	BO4	A	2001	5/5	0.92	0.16	91,92,93,94	0
3	BO4	T	2001	5/5	0.95	0.18	90,90,90,90	5
3	BO4	D	2001	5/5	0.95	0.13	113,114,114,114	5

6.5 Other polymers [i](#)

There are no such residues in this entry.