



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 20, 2018 – 08:03 PM EST

PDB ID : 6BJY
Title : VSV Nucleocapsid with Polyamide Bound
Authors : Gumpfer, R.H.; Luo, M.
Deposited on : 2017-11-07
Resolution : 3.46 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20030736
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20030736

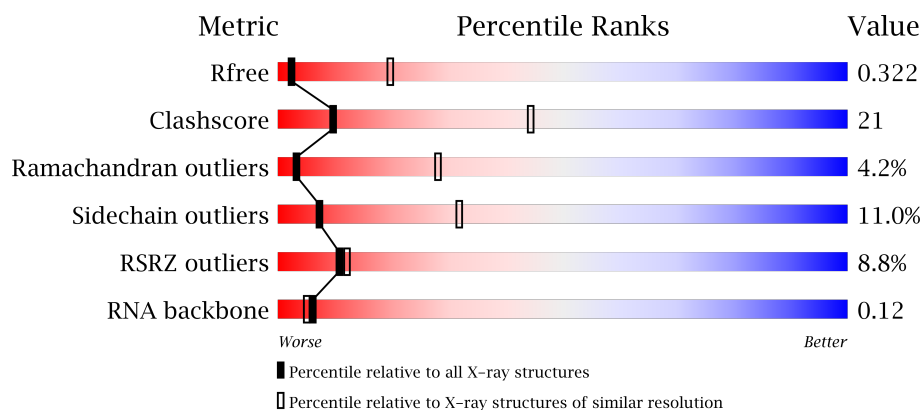
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.46 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}	100719	1135 (3.56-3.36)
Clashscore	112137	1040 (3.52-3.40)
Ramachandran outliers	110173	1009 (3.52-3.40)
Sidechain outliers	110143	1010 (3.52-3.40)
RSRZ outliers	101464	1017 (3.54-3.38)
RNA backbone	2435	1020 (4.02-2.86)

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. 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	R	45	<div> <div>9%</div> <div>7% 24% 56% 13%</div> </div>
2	A	421	<div> <div>11%</div> <div>61% 33% 5%</div> </div>
2	B	421	<div> <div>7%</div> <div>59% 32% 8% .</div> </div>
2	C	421	<div> <div>7%</div> <div>59% 31% 8% .</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	421	<div><div></div><div>6%</div><div>59%</div><div>31%</div><div>9%</div><div></div></div>
2	E	421	<div><div></div><div>13%</div><div>66%</div><div>29%</div><div>5%</div><div></div></div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 17494 atoms, of which 28 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.

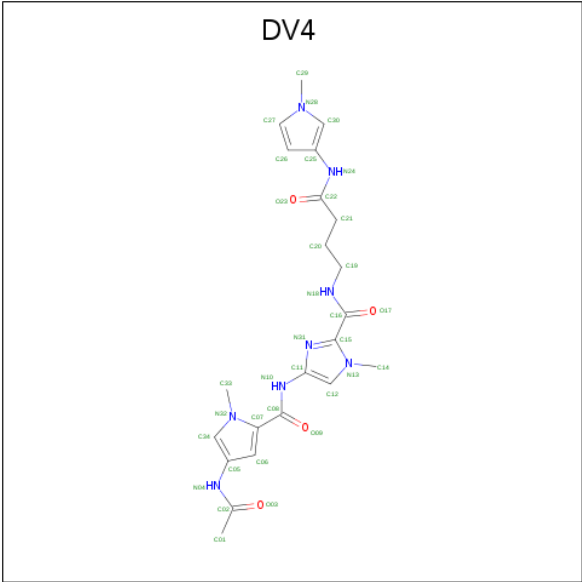
- Molecule 1 is a RNA chain called RNA (45-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	R	45	Total	C	N	O	P	0	0	0
			900	405	90	360	45			

- Molecule 2 is a protein called Nucleoprotein.

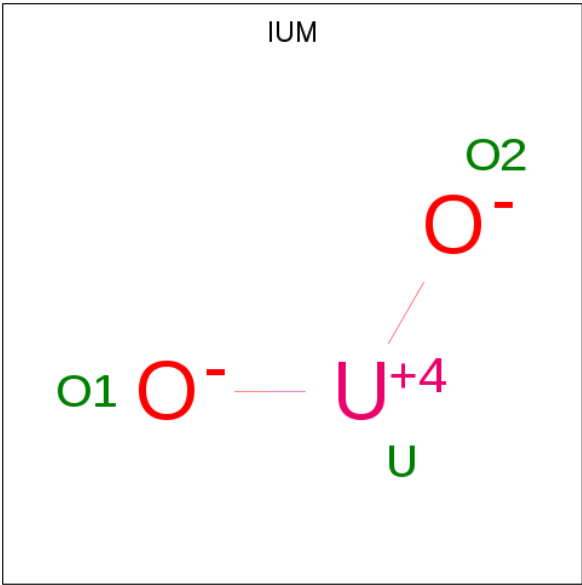
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	421	Total	C	N	O	S	0	0	0
			3327	2118	558	633	18			
2	B	415	Total	C	N	O	S	0	0	0
			3290	2097	552	623	18			
2	C	413	Total	C	N	O	S	0	0	0
			3275	2089	550	618	18			
2	D	416	Total	C	N	O	S	0	0	0
			3298	2103	553	624	18			
2	E	421	Total	C	N	O	S	0	0	0
			3327	2118	558	633	18			

- Molecule 3 is 4-{[4-(acetylamino)-1-methyl-1H-pyrrole-2-carbonyl]amino}-1-methyl-N-{4-[(1-methyl-1H-pyrrol-3-yl)amino]-4-oxobutyl}-1H-imidazole-2-carboxamide (three-letter code: DV4) (formula: C₂₂H₂₈N₈O₄).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	R	1	Total	C	H	N	O	0	0
			62	22	28	8	4		

- Molecule 4 is URANYL (VI) ION (three-letter code: IUM) (formula: O₂U).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	U	0	0
			1	1		
4	A	1	Total	U	0	0
			1	1		
4	A	1	Total	U	0	0
			1	1		

Continued on next page...

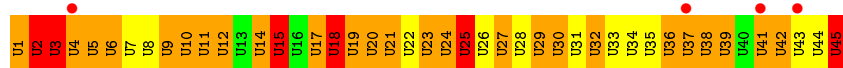
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total 1	U 1	0	0
4	A	1	Total 1	U 1	0	0
4	B	1	Total 1	U 1	0	0
4	B	1	Total 1	U 1	0	0
4	B	1	Total 1	U 1	0	0
4	C	1	Total 1	U 1	0	0
4	C	1	Total 1	U 1	0	0
4	C	1	Total 1	U 1	0	0
4	D	1	Total 1	U 1	0	0
4	D	1	Total 1	U 1	0	0
4	E	1	Total 1	U 1	0	0
4	E	1	Total 1	U 1	0	0

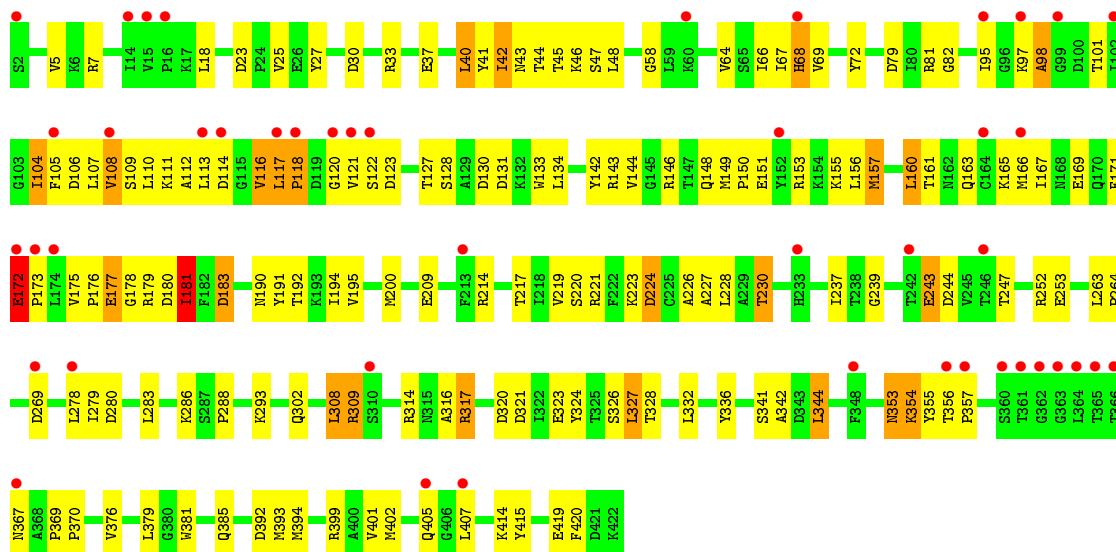
3 Residue-property plots

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.

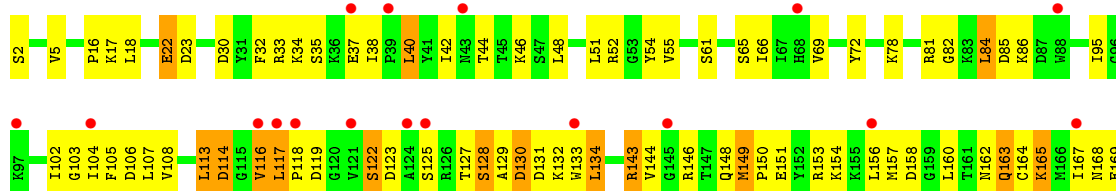
- Molecule 1: RNA (45-MER)

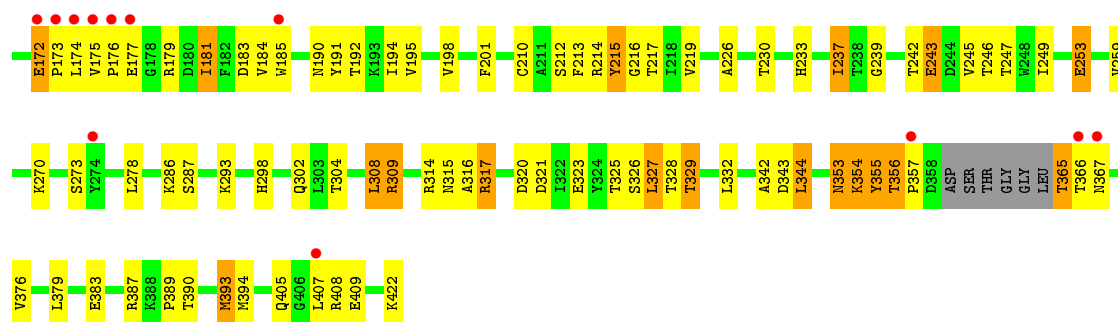


- Molecule 2: Nucleoprotein

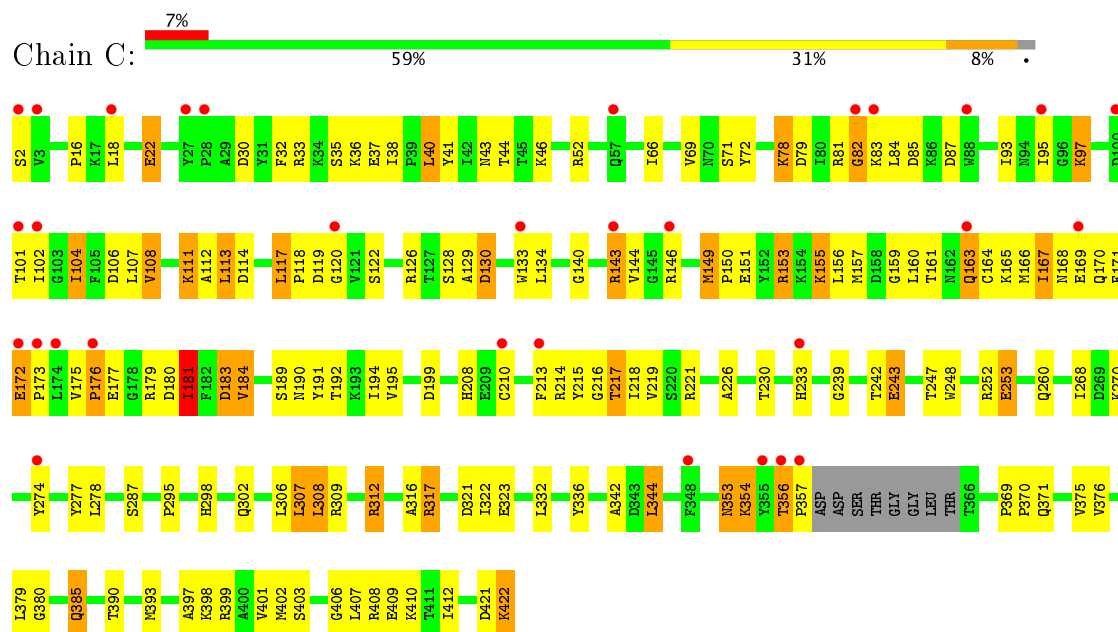


- Molecule 2: Nucleoprotein

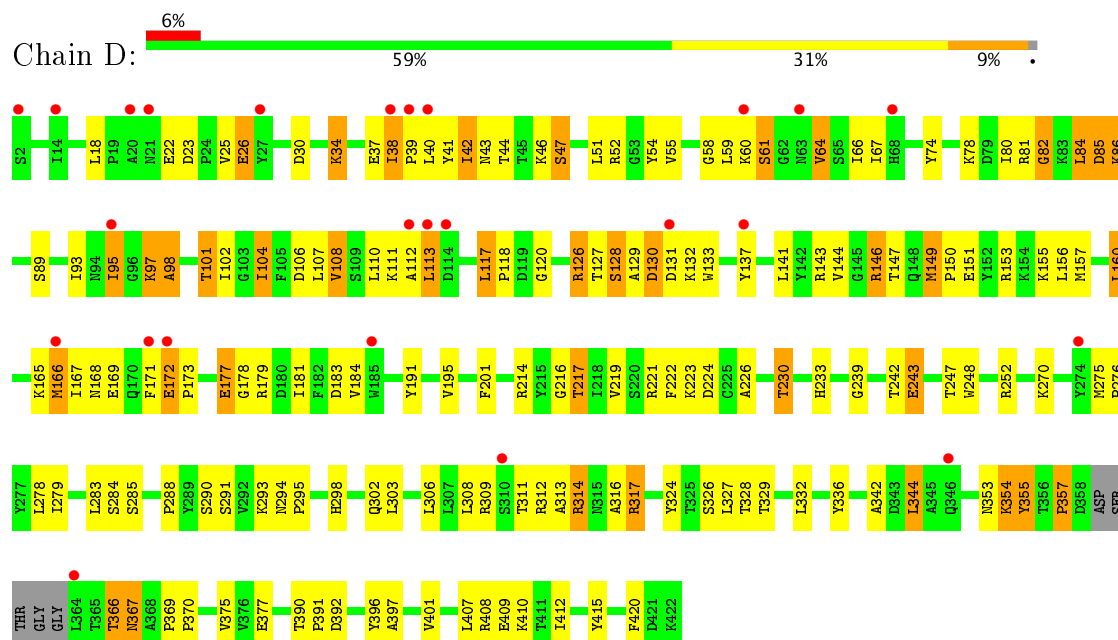




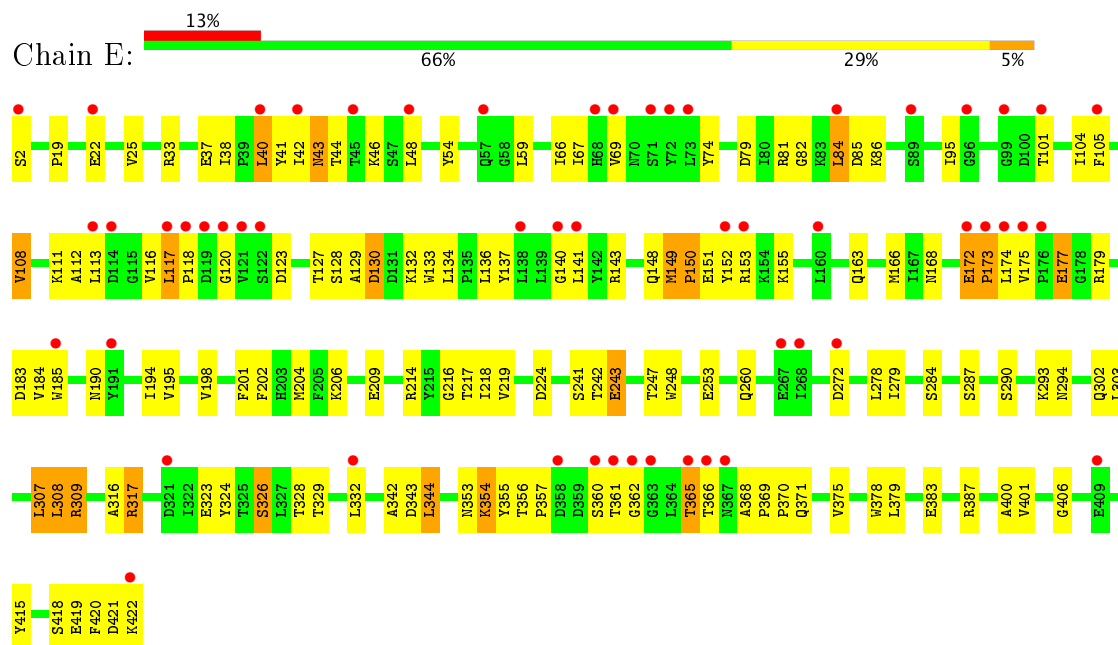
• Molecule 2: Nucleoprotein



• Molecule 2: Nucleoprotein



● Molecule 2: Nucleoprotein



4 Data and refinement statistics

Property	Value	Source
Space group	P 2 ₁ 2 ₁ 2 ₁	Depositor
Cell constants a, b, c, α , β , γ	75.48Å 165.48Å 234.46Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.30 – 3.46 46.30 – 3.46	Depositor EDS
% Data completeness (in resolution range)	92.3 (46.30-3.46) 92.3 (46.30-3.46)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.82 (at 3.48Å)	Xtriage
Refinement program	PHENIX (1.12_2829: ???)	Depositor
R, R_{free}	0.266 , 0.311 0.280 , 0.322	Depositor DCC
R_{free} test set	1998 reflections (5.52%)	DCC
Wilson B-factor (Å ²)	100.3	Xtriage
Anisotropy	0.630	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 47.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.36$, $\langle L^2 \rangle = 0.18$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	17494	wwPDB-VP
Average B, all atoms (Å ²)	121.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.57% 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: DV4, IUM

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	R	0.58	1/989 (0.1%)	1.42	27/1526 (1.8%)
2	A	0.25	0/3403	0.45	0/4607
2	B	0.25	0/3365	0.46	0/4554
2	C	0.25	0/3350	0.48	0/4533
2	D	0.24	0/3373	0.44	0/4565
2	E	0.25	0/3403	0.45	0/4607
All	All	0.28	1/17883 (0.0%)	0.57	27/24392 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	R	1	U	C4-C5	-5.97	1.38	1.43

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	R	1	U	O5'-P-OP1	-15.45	91.79	105.70
1	R	1	U	OP1-P-OP2	-10.34	104.09	119.60
1	R	25	U	N3-C2-O2	-9.23	115.74	122.20
1	R	1	U	N1-C2-O2	-9.09	116.44	122.80
1	R	45	U	C6-N1-C2	8.56	126.14	121.00

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	R	900	0	451	158	0
2	A	3327	0	3287	135	0
2	B	3290	0	3253	135	0
2	C	3275	0	3242	147	0
2	D	3298	0	3264	147	0
2	E	3327	0	3287	115	0
3	R	34	28	0	20	0
4	A	5	0	0	0	0
4	B	3	0	0	0	0
4	C	3	0	0	0	0
4	D	2	0	0	0	0
4	E	2	0	0	0	0
All	All	17466	28	16784	736	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 21.

The worst 5 of 736 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:R:25:U:H1'	3:R:101:DV4:C21	1.59	1.32
1:R:18:U:O2'	1:R:19:U:H3'	1.23	1.29
1:R:25:U:H1'	3:R:101:DV4:C20	1.72	1.20
2:B:37:GLU:HB2	2:B:108:VAL:HG21	1.18	1.15
2:A:165:LYS:HA	2:E:184:VAL:HG22	1.20	1.11

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 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	
2	A	419/421 (100%)	357 (85%)	41 (10%)	21 (5%)	2	23
2	B	411/421 (98%)	355 (86%)	41 (10%)	15 (4%)	4	32
2	C	409/421 (97%)	354 (87%)	39 (10%)	16 (4%)	3	30
2	D	412/421 (98%)	360 (87%)	34 (8%)	18 (4%)	3	27
2	E	419/421 (100%)	356 (85%)	46 (11%)	17 (4%)	3	29
All	All	2070/2105 (98%)	1782 (86%)	201 (10%)	87 (4%)	3	28

5 of 87 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	105	PHE
2	A	113	LEU
2	B	113	LEU
2	B	128	SER
2	C	113	LEU

5.3.2 Protein sidechains ⓘ

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	
2	A	362/362 (100%)	329 (91%)	33 (9%)	11	41
2	B	358/362 (99%)	317 (88%)	41 (12%)	6	29
2	C	356/362 (98%)	306 (86%)	50 (14%)	4	21
2	D	359/362 (99%)	314 (88%)	45 (12%)	5	26
2	E	362/362 (100%)	333 (92%)	29 (8%)	14	48
All	All	1797/1810 (99%)	1599 (89%)	198 (11%)	7	32

5 of 198 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	122	SER
2	C	253	GLU
2	E	149	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	C	143	ARG
2	C	167	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	C	163	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	44/45 (97%)	31 (70%)	5 (11%)

5 of 31 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	2	U
1	R	3	U
1	R	4	U
1	R	5	U
1	R	6	U

All (5) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	R	5	U
1	R	17	U
1	R	20	U
1	R	36	U
1	R	38	U

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 16 ligands modelled in this entry, 15 are modelled with single atom - leaving 1 for Mogul analysis.

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	DV4	R	101	-	31,36,36	2.34	13 (41%)	31,50,50	3.66	15 (48%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	DV4	R	101	-	-	3/14/27/27	0/3/3/3

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	R	101	DV4	C30-N28	-4.00	1.32	1.38
3	R	101	DV4	O17-C16	-2.80	1.17	1.23
3	R	101	DV4	O23-C22	-2.72	1.17	1.23
3	R	101	DV4	O09-C08	-2.66	1.17	1.23
3	R	101	DV4	O03-C02	-2.40	1.17	1.23

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	R	101	DV4	O09-C08-N10	-5.81	110.43	123.69
3	R	101	DV4	C05-N04-C02	-4.04	121.23	128.03
3	R	101	DV4	O17-C16-N18	-3.60	115.15	122.59
3	R	101	DV4	O17-C16-C15	-3.52	115.46	120.37
3	R	101	DV4	C25-N24-C22	-2.78	122.52	127.49

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	R	101	DV4	O09-C08-N10-C11
3	R	101	DV4	C07-C08-N10-C11
3	R	101	DV4	C15-C16-N18-C19

There are no ring outliers.

1 monomer is involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	R	101	DV4	20	0

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	R	45/45 (100%)	0.77	4 (8%) 10 12	104, 149, 197, 237	0
2	A	421/421 (100%)	0.56	45 (10%) 7 8	56, 112, 185, 250	0
2	B	415/421 (98%)	0.41	29 (6%) 17 17	54, 107, 170, 314	0
2	C	413/421 (98%)	0.43	31 (7%) 15 15	53, 110, 178, 258	0
2	D	416/421 (98%)	0.43	25 (6%) 23 21	58, 116, 180, 225	0
2	E	421/421 (100%)	0.72	54 (12%) 4 6	63, 125, 197, 246	0
All	All	2131/2150 (99%)	0.51	188 (8%) 11 12	53, 115, 185, 314	0

The worst 5 of 188 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	361	THR	15.2
2	E	2	SER	11.7
2	C	2	SER	10.6
2	E	366	THR	10.2
2	E	176	PRO	7.5

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
3	DV4	R	101	34/34	0.69	0.32	1.17	78,121,148,163	0
4	IUM	B	501	1/3	0.96	0.21	-0.77	164,164,164,164	0
4	IUM	E	501	1/3	0.94	0.17	-	144,144,144,144	0
4	IUM	E	502	1/3	0.89	0.14	-	203,203,203,203	0
4	IUM	C	502	1/3	0.82	0.23	-	144,144,144,144	0
4	IUM	B	502	1/3	0.84	0.14	-	130,130,130,130	0
4	IUM	A	502	1/3	0.81	0.22	-	104,104,104,104	0
4	IUM	C	501	1/3	0.97	0.16	-	146,146,146,146	0
4	IUM	A	505	1/3	0.91	0.19	-	125,125,125,125	0
4	IUM	D	502	1/3	0.73	0.25	-	131,131,131,131	0
4	IUM	B	503	1/3	0.94	0.14	-	194,194,194,194	0
4	IUM	A	501	1/3	0.97	0.18	-	128,128,128,128	0
4	IUM	A	504	1/3	0.97	0.14	-	175,175,175,175	0
4	IUM	C	503	1/3	0.98	0.16	-	144,144,144,144	0
4	IUM	A	503	1/3	0.86	0.16	-	169,169,169,169	0
4	IUM	D	501	1/3	0.91	0.21	-	150,150,150,150	0

6.5 Other polymers [i](#)

There are no such residues in this entry.