



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 03:35 am BST

PDB ID : 3PU1
Title : Crystal Structure of a vesicular stomatitis virus nucleocapsid-polyG complex
Authors : Luo, M.; Green, T.J.; Rowse, M.
Deposited on : 2010-12-03
Resolution : 3.14 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
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

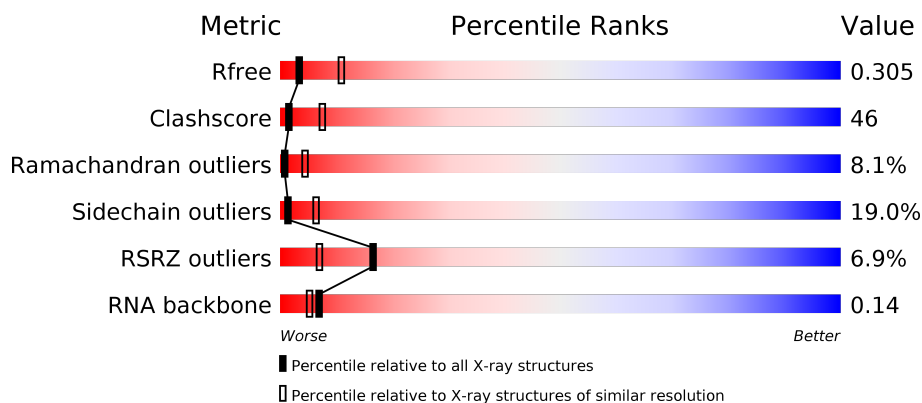
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.14 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	1626 (3.18-3.10)
Clashscore	141614	1735 (3.18-3.10)
Ramachandran outliers	138981	1677 (3.18-3.10)
Sidechain outliers	138945	1677 (3.18-3.10)
RSRZ outliers	127900	1588 (3.18-3.10)
RNA backbone	3102	1000 (3.46-2.82)

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	421	<div> <div>3%</div> <div> <div></div> <div>43%</div> <div>40%</div> <div>15%</div> <div>•</div> </div> </div>
1	B	421	<div> <div>4%</div> <div> <div></div> <div>44%</div> <div>37%</div> <div>14%</div> <div>• •</div> </div> </div>
1	C	421	<div> <div>4%</div> <div> <div></div> <div>40%</div> <div>43%</div> <div>12%</div> <div>• •</div> </div> </div>
1	D	421	<div> <div>5%</div> <div> <div></div> <div>43%</div> <div>39%</div> <div>14%</div> <div>• •</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	E	421	<div><div></div><div>9%</div><div>44%</div><div>41%</div><div>12%</div><div></div></div>
2	R	45	<div><div></div><div>96%</div><div>42%</div><div>56%</div><div></div></div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 17572 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.

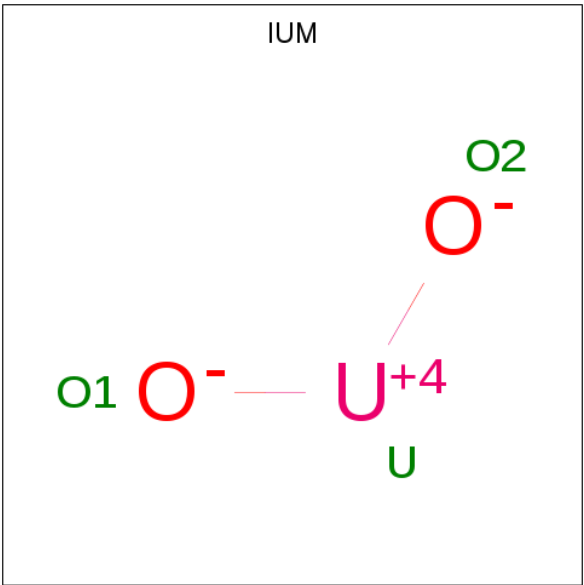
- Molecule 1 is a protein called Nucleoprotein.

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

- Molecule 2 is a RNA chain called RNA (45-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	R	45	Total	C	N	O	P	0	0	0
			1035	450	225	315	45			

- Molecule 3 is URANYL (VI) ION (three-letter code: IUM) (formula: O₂U).



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

Continued on next page...

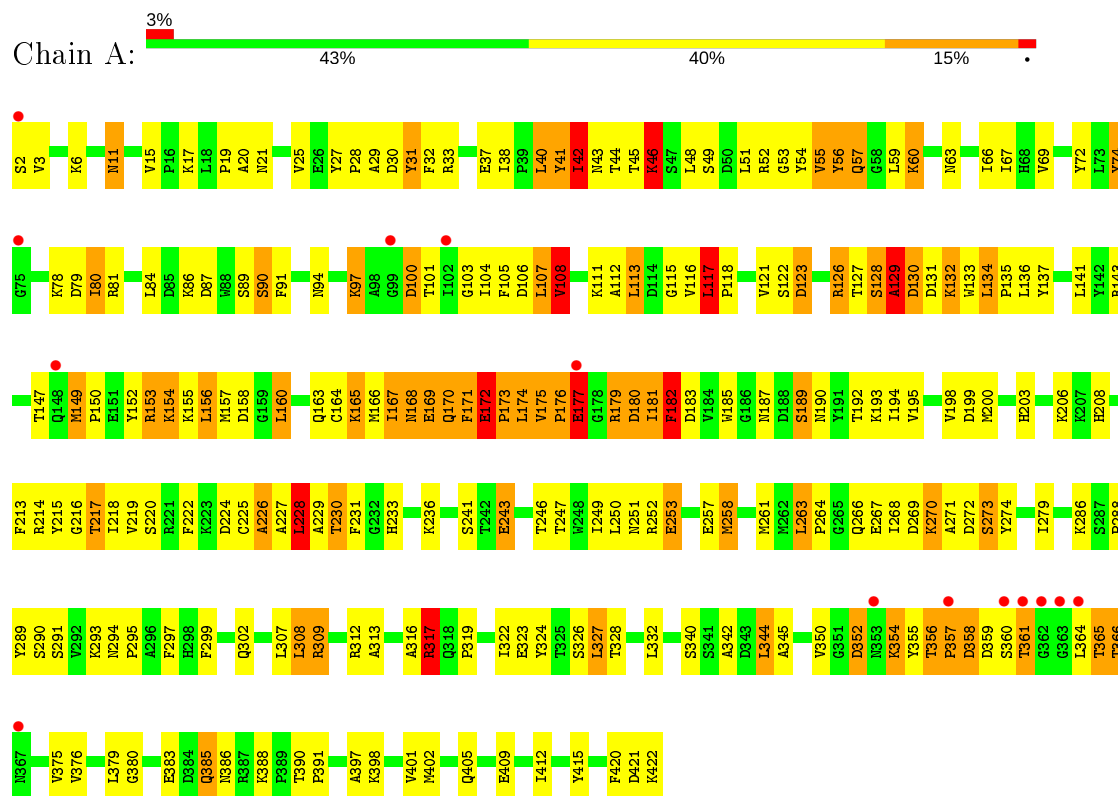
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	1	Total 1	U 1	0	0
3	R	1	Total 1	U 1	0	0
3	R	1	Total 1	U 1	0	0
3	R	1	Total 1	U 1	0	0
3	R	1	Total 1	U 1	0	0
3	R	1	Total 1	U 1	0	0

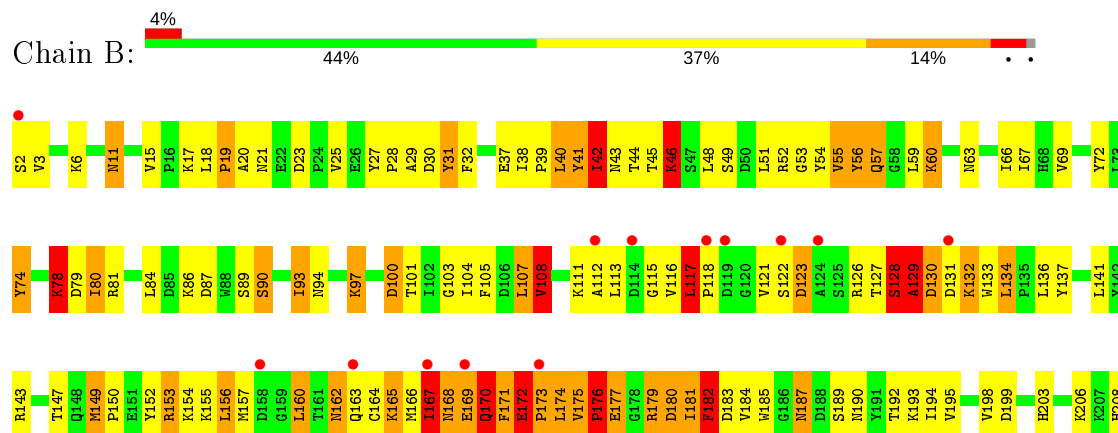
3 Residue-property plots

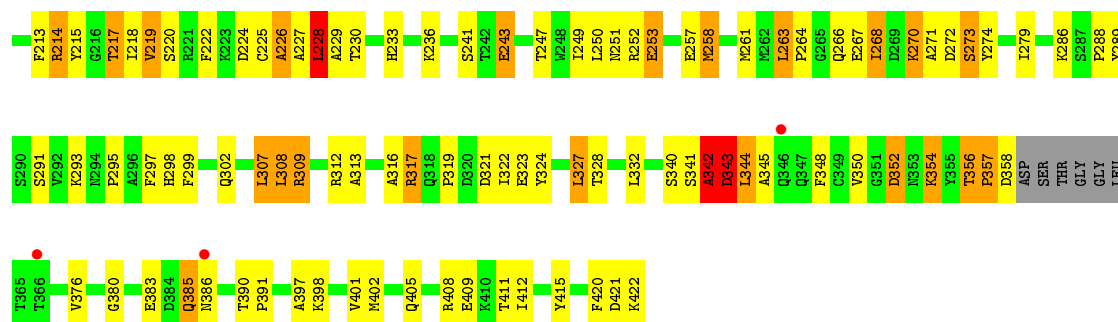
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Nucleoprotein

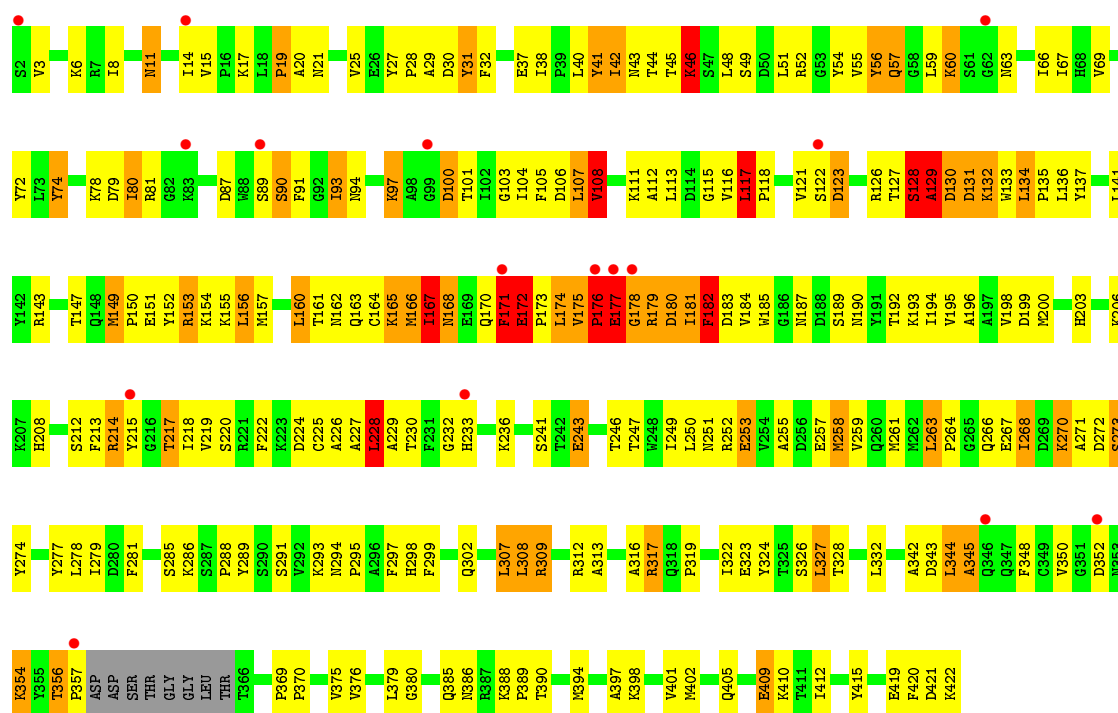


• Molecule 1: Nucleoprotein

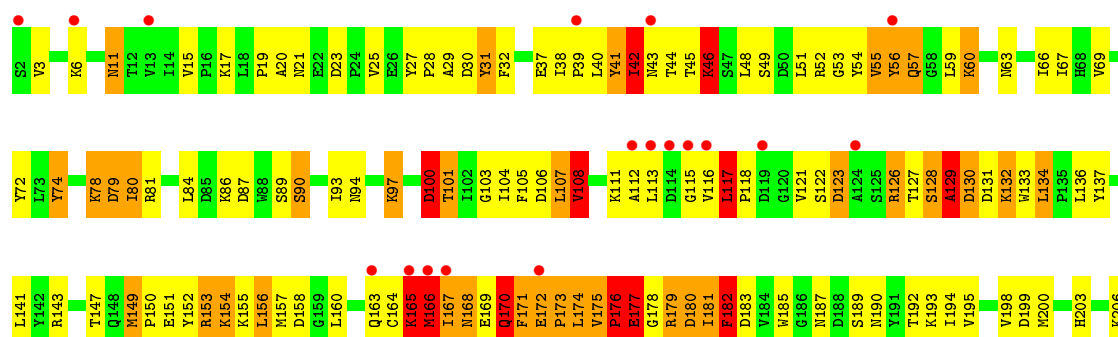


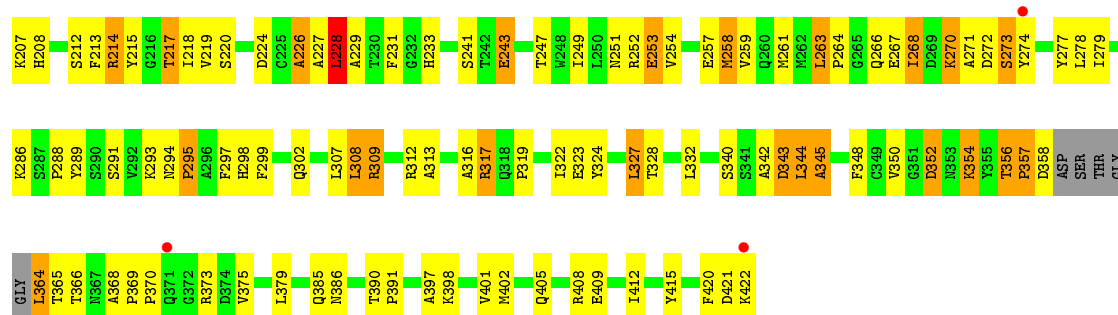


• Molecule 1: Nucleoprotein



• Molecule 1: Nucleoprotein

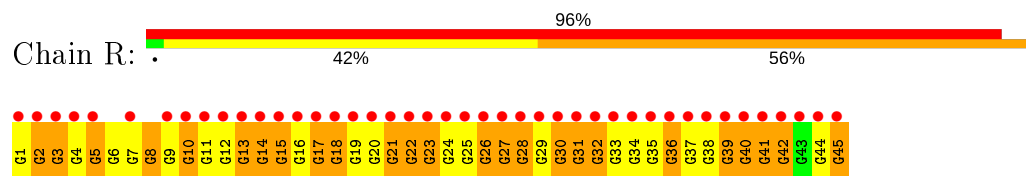




• Molecule 1: Nucleoprotein



• Molecule 2: RNA (45-MER)



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	163.22Å 233.19Å 75.57Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.10 – 3.14 45.14 – 3.14	Depositor EDS
% Data completeness (in resolution range)	79.0 (45.10-3.14) 79.0 (45.14-3.14)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.92 (at 3.12Å)	Xtriage
Refinement program	PHENIX 1.6.1 _357	Depositor
R, R_{free}	0.259 , 0.305 0.255 , 0.305	Depositor DCC
R_{free} test set	2000 reflections (4.75%)	wwPDB-VP
Wilson B-factor (Å ²)	76.3	Xtriage
Anisotropy	0.674	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 48.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	17572	wwPDB-VP
Average B, all atoms (Å ²)	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.80% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 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	A	0.50	0/3403	0.75	6/4607 (0.1%)
1	B	0.50	0/3365	0.67	2/4554 (0.0%)
1	C	0.53	0/3350	0.76	8/4533 (0.2%)
1	D	0.50	0/3373	0.69	4/4565 (0.1%)
1	E	0.48	0/3403	0.71	6/4607 (0.1%)
2	R	0.93	0/1169	1.21	4/1841 (0.2%)
All	All	0.54	0/18063	0.76	30/24707 (0.1%)

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	A	0	5
1	B	0	8
1	C	0	4
1	D	0	6
1	E	0	7
All	All	0	30

There are no bond length outliers.

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	177	GLU	N-CA-CB	-13.66	86.01	110.60
1	E	177	GLU	N-CA-CB	-13.58	86.15	110.60
1	C	171	PHE	CB-CA-C	-10.53	89.33	110.40
1	D	130	ASP	N-CA-CB	-9.40	93.68	110.60
1	C	130	ASP	N-CA-CB	-9.40	93.69	110.60

There are no chirality outliers.

5 of 30 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	107	LEU	Peptide
1	A	129	ALA	Peptide
1	A	169	GLU	Peptide
1	A	317	ARG	Sidechain
1	A	41	TYR	Peptide

5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3327	0	3287	332	0
1	B	3290	0	3253	345	0
1	C	3275	0	3242	326	0
1	D	3298	0	3264	321	0
1	E	3327	0	3287	292	0
2	R	1035	0	496	125	0
3	A	5	0	0	0	0
3	B	3	0	0	0	0
3	C	3	0	0	0	0
3	D	2	0	0	0	0
3	E	2	0	0	0	0
3	R	5	0	0	0	0
All	All	17572	0	16829	1572	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 46.

The worst 5 of 1572 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:C:171:PHE:CG	1:C:171:PHE:O	1.78	1.31
1:C:171:PHE:CD2	1:C:171:PHE:O	1.91	1.23
1:A:326:SER:HB3	1:B:343:ASP:CB	1.72	1.19
1:D:317:ARG:H	1:D:317:ARG:CZ	1.56	1.17

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:166:MET:O	1:D:167:ILE:HG13	1.42	1.17

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	
1	A	419/421 (100%)	319 (76%)	66 (16%)	34 (8%)	1	4
1	B	411/421 (98%)	321 (78%)	53 (13%)	37 (9%)	1	3
1	C	409/421 (97%)	312 (76%)	67 (16%)	30 (7%)	1	5
1	D	412/421 (98%)	317 (77%)	59 (14%)	36 (9%)	1	4
1	E	419/421 (100%)	322 (77%)	66 (16%)	31 (7%)	1	5
All	All	2070/2105 (98%)	1591 (77%)	311 (15%)	168 (8%)	1	4

5 of 168 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	63	ASN
1	A	113	LEU
1	A	117	LEU
1	A	167	ILE
1	A	173	PRO

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all 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	362/362 (100%)	296 (82%)	66 (18%)	1	7
1	B	358/362 (99%)	290 (81%)	68 (19%)	1	6
1	C	356/362 (98%)	289 (81%)	67 (19%)	1	6
1	D	359/362 (99%)	289 (80%)	70 (20%)	1	6
1	E	362/362 (100%)	292 (81%)	70 (19%)	1	6
All	All	1797/1810 (99%)	1456 (81%)	341 (19%)	1	6

5 of 341 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	131	ASP
1	C	308	LEU
1	E	243	GLU
1	C	149	MET
1	C	182	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 48 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	170	GLN
1	C	385	GLN
1	E	347	GLN
1	C	190	ASN
1	C	302	GLN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	R	44/45 (97%)	24 (54%)	3 (6%)

5 of 24 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	R	2	G
2	R	3	G
2	R	5	G
2	R	8	G

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	R	10	G

All (3) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	R	17	G
2	R	18	G
2	R	26	G

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](#)

Of 20 ligands modelled in this entry, 20 are modelled with single atom - 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.

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(Å²)	Q<0.9
1	A	421/421 (100%)	0.21	14 (3%)	46	25	56, 92, 130, 142	0
1	B	415/421 (98%)	0.20	16 (3%)	39	20	57, 91, 130, 141	0
1	C	413/421 (98%)	0.21	16 (3%)	39	20	56, 92, 129, 142	0
1	D	416/421 (98%)	0.31	21 (5%)	28	13	56, 92, 130, 141	0
1	E	421/421 (100%)	0.37	37 (8%)	10	4	57, 93, 131, 147	0
2	R	45/45 (100%)	4.36	43 (95%)	0	0	118, 132, 143, 145	0
All	All	2131/2150 (99%)	0.35	147 (6%)	16	7	56, 93, 131, 147	0

The worst 5 of 147 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2	SER	11.5
2	R	18	G	10.4
1	E	2	SER	9.7
1	E	114	ASP	9.7
2	R	37	G	8.8

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

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(Å ²)	Q<0.9
3	IUM	R	538	1/3	0.76	0.10	234,234,234,234	1
3	IUM	B	525	1/3	0.81	0.08	157,157,157,157	1
3	IUM	R	537	1/3	0.81	0.08	161,161,161,161	1
3	IUM	R	536	1/3	0.84	0.14	212,212,212,212	1
3	IUM	A	521	1/3	0.86	0.08	171,171,171,171	1
3	IUM	R	540	1/3	0.87	0.09	238,238,238,238	1
3	IUM	B	527	1/3	0.88	0.07	168,168,168,168	1
3	IUM	D	531	1/3	0.90	0.05	144,144,144,144	1
3	IUM	E	535	1/3	0.91	0.05	160,160,160,160	1
3	IUM	R	539	1/3	0.92	0.11	156,156,156,156	1
3	IUM	A	523	1/3	0.95	0.07	177,177,177,177	1
3	IUM	A	524	1/3	0.95	0.03	159,159,159,159	1
3	IUM	C	530	1/3	0.97	0.04	151,151,151,151	1
3	IUM	C	528	1/3	0.98	0.07	130,130,130,130	1
3	IUM	E	533	1/3	0.98	0.07	152,152,152,152	1
3	IUM	D	532	1/3	0.99	0.15	119,119,119,119	0
3	IUM	A	534	1/3	0.99	0.14	128,128,128,128	0
3	IUM	B	526	1/3	0.99	0.15	103,103,103,103	1
3	IUM	C	529	1/3	0.99	0.12	108,108,108,108	1
3	IUM	A	522	1/3	0.99	0.15	113,113,113,113	1

6.5 Other polymers

There are no such residues in this entry.