Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 03:23 pm BST

PDB ID	:	4WRT
Title	:	Crystal structure of Influenza B polymerase with bound vRNA promoter (form FluB2)
Authors	:	Reich, S.; Guilligay, D.; Pflug, A.; Cusack, S.
Deposited on	:	2014-10-25
Resolution	:	2.70 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

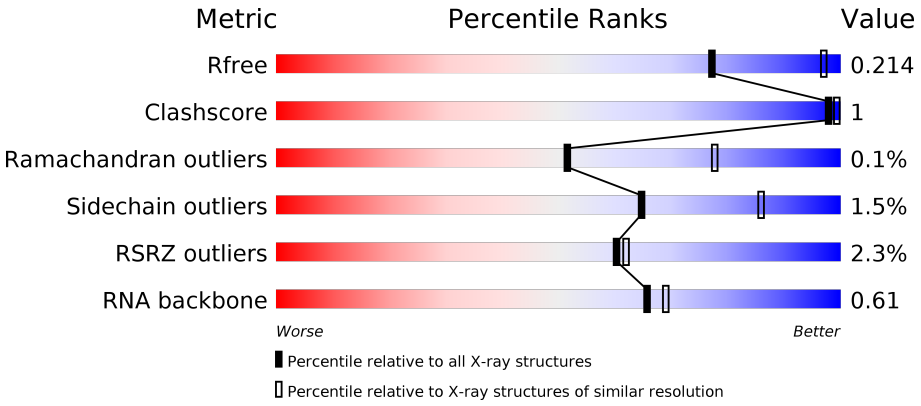
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)
RNA backbone	3102	1159 (3.00-2.40)

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 <=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	18	17%, 78%, 6%, 17%
2	V	14	71%, 29%
3	A	751	3%, 92%, 6%
4	B	772	2%, 93%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	C	798	

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 14972 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 RNA chain called Influenza virus polymerase vRNA promoter 3' end.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	R	15	Total	C	N	O	P	0	0	0
			308	138	45	110	15			

- Molecule 2 is a RNA chain called Influenza virus polymerase vRNA promoter 5' end.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	V	14	Total	C	N	O	P	0	0	0
			307	137	62	94	14			

- Molecule 3 is a protein called PA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	709	Total	C	N	O	S	0	0	0
			5691	3618	954	1079	40			

There are 25 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	GLY	-	expression tag	UNP Q5V8Z9
A	-12	SER	-	expression tag	UNP Q5V8Z9
A	-11	HIS	-	expression tag	UNP Q5V8Z9
A	-10	HIS	-	expression tag	UNP Q5V8Z9
A	-9	HIS	-	expression tag	UNP Q5V8Z9
A	-8	HIS	-	expression tag	UNP Q5V8Z9
A	-7	HIS	-	expression tag	UNP Q5V8Z9
A	-6	HIS	-	expression tag	UNP Q5V8Z9
A	-5	HIS	-	expression tag	UNP Q5V8Z9
A	-4	HIS	-	expression tag	UNP Q5V8Z9
A	-3	GLY	-	expression tag	UNP Q5V8Z9
A	-2	SER	-	expression tag	UNP Q5V8Z9
A	-1	GLY	-	expression tag	UNP Q5V8Z9
A	0	SER	-	expression tag	UNP Q5V8Z9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	727	GLY	-	expression tag	UNP Q5V8Z9
A	728	SER	-	expression tag	UNP Q5V8Z9
A	729	GLY	-	expression tag	UNP Q5V8Z9
A	730	SER	-	expression tag	UNP Q5V8Z9
A	731	GLY	-	expression tag	UNP Q5V8Z9
A	732	GLU	-	expression tag	UNP Q5V8Z9
A	733	ASN	-	expression tag	UNP Q5V8Z9
A	734	LEU	-	expression tag	UNP Q5V8Z9
A	735	TYR	-	expression tag	UNP Q5V8Z9
A	736	PHE	-	expression tag	UNP Q5V8Z9
A	737	GLN	-	expression tag	UNP Q5V8Z9

- Molecule 4 is a protein called RNA-directed RNA polymerase catalytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	742	Total	C	N	O	S	0	0	0
			5828	3677	1009	1090	52			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-8	GLY	-	expression tag	UNP Q5V8Y6
B	-7	SER	-	expression tag	UNP Q5V8Y6
B	-6	GLY	-	expression tag	UNP Q5V8Y6
B	-5	SER	-	expression tag	UNP Q5V8Y6
B	-4	GLY	-	expression tag	UNP Q5V8Y6
B	-3	SER	-	expression tag	UNP Q5V8Y6
B	-2	GLY	-	expression tag	UNP Q5V8Y6
B	-1	SER	-	expression tag	UNP Q5V8Y6
B	0	GLY	-	expression tag	UNP Q5V8Y6
B	753	GLY	-	expression tag	UNP Q5V8Y6
B	754	SER	-	expression tag	UNP Q5V8Y6
B	755	GLY	-	expression tag	UNP Q5V8Y6
B	756	SER	-	expression tag	UNP Q5V8Y6
B	757	GLY	-	expression tag	UNP Q5V8Y6
B	758	GLU	-	expression tag	UNP Q5V8Y6
B	759	ASN	-	expression tag	UNP Q5V8Y6
B	760	LEU	-	expression tag	UNP Q5V8Y6
B	761	TYR	-	expression tag	UNP Q5V8Y6
B	762	PHE	-	expression tag	UNP Q5V8Y6
B	763	GLN	-	expression tag	UNP Q5V8Y6

- Molecule 5 is a protein called PB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	C	250	Total	C	N	O	S	0	0	0
			2029	1296	360	359	14			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-8	GLY	-	expression tag	UNP Q5V8X3
C	-7	SER	-	expression tag	UNP Q5V8X3
C	-6	GLY	-	expression tag	UNP Q5V8X3
C	-5	SER	-	expression tag	UNP Q5V8X3
C	-4	GLY	-	expression tag	UNP Q5V8X3
C	-3	SER	-	expression tag	UNP Q5V8X3
C	-2	GLY	-	expression tag	UNP Q5V8X3
C	-1	SER	-	expression tag	UNP Q5V8X3
C	0	GLY	-	expression tag	UNP Q5V8X3
C	771	GLY	-	expression tag	UNP Q5V8X3
C	772	TRP	-	expression tag	UNP Q5V8X3
C	773	SER	-	expression tag	UNP Q5V8X3
C	774	HIS	-	expression tag	UNP Q5V8X3
C	775	PRO	-	expression tag	UNP Q5V8X3
C	776	GLN	-	expression tag	UNP Q5V8X3
C	777	PHE	-	expression tag	UNP Q5V8X3
C	778	GLU	-	expression tag	UNP Q5V8X3
C	779	LYS	-	expression tag	UNP Q5V8X3
C	780	GLY	-	expression tag	UNP Q5V8X3
C	781	SER	-	expression tag	UNP Q5V8X3
C	782	GLY	-	expression tag	UNP Q5V8X3
C	783	SER	-	expression tag	UNP Q5V8X3
C	784	GLU	-	expression tag	UNP Q5V8X3
C	785	ASN	-	expression tag	UNP Q5V8X3
C	786	LEU	-	expression tag	UNP Q5V8X3
C	787	TYR	-	expression tag	UNP Q5V8X3
C	788	PHE	-	expression tag	UNP Q5V8X3
C	789	GLN	-	expression tag	UNP Q5V8X3

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	R	17	Total	O	0	0
			17	17		
6	V	46	Total	O	0	0
			46	46		

Continued on next page...

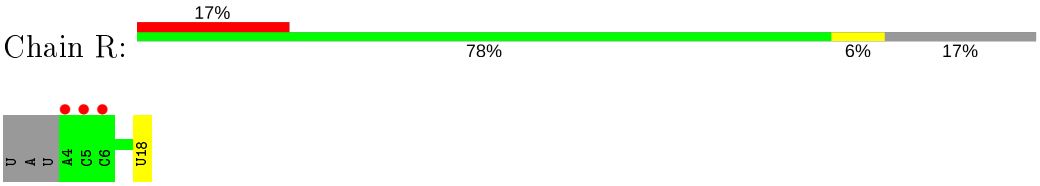
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	303	Total O 303 303	0	0
6	B	338	Total O 338 338	0	0
6	C	105	Total O 105 105	0	0

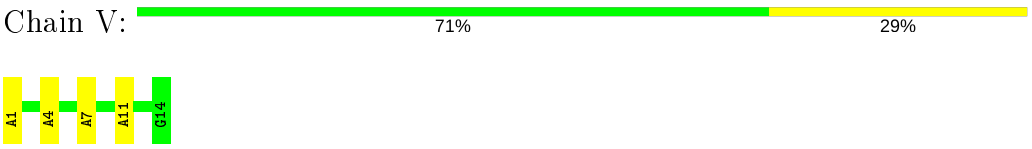
3 Residue-property plots (i)

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.

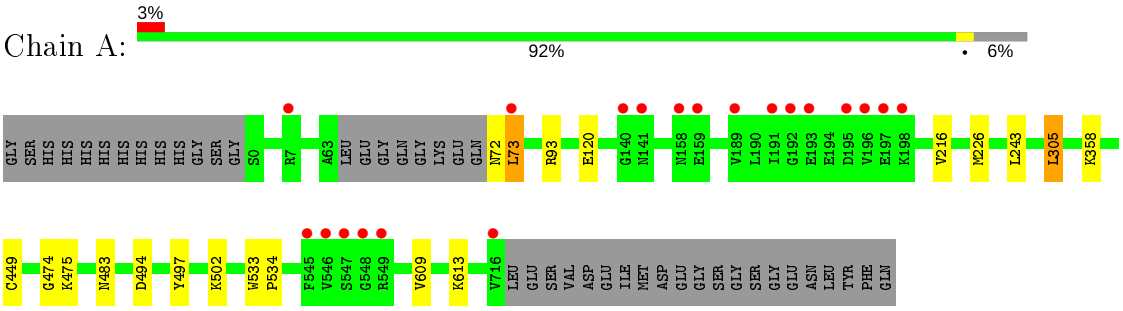
- Molecule 1: Influenza virus polymerase vRNA promoter 3' end



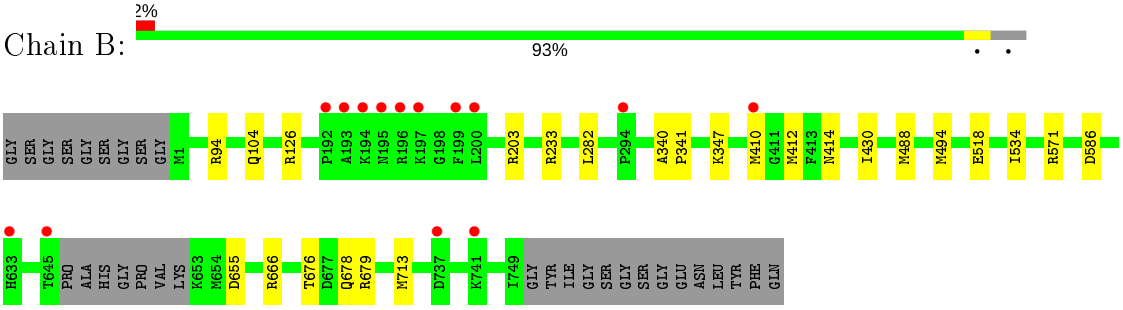
- Molecule 2: Influenza virus polymerase vRNA promoter 5' end



- Molecule 3: PA

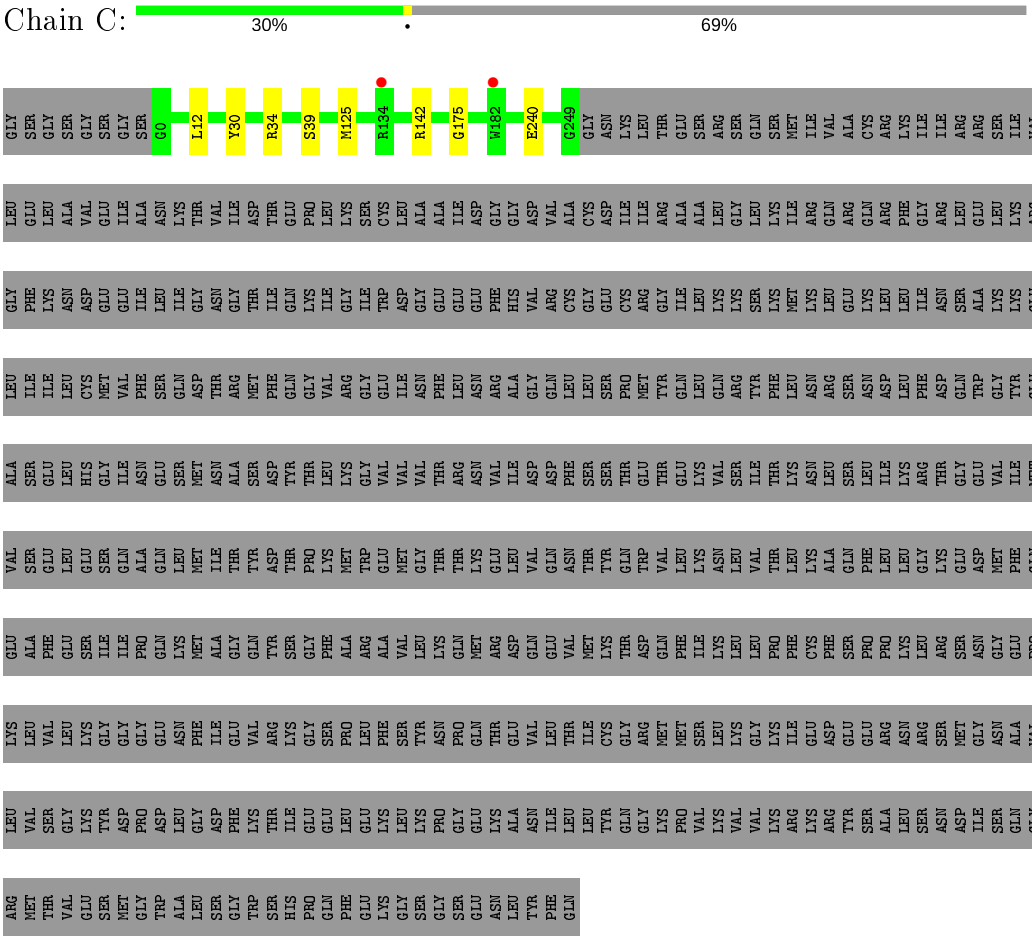


- Molecule 4: RNA-directed RNA polymerase catalytic subunit



- Molecule 5: PB2

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, α, β, γ	207.37Å 207.37Å 345.69Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.66 – 2.70 49.66 – 2.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.66-2.70) 100.0 (49.66-2.70)	Depositor EDS
R_{merge}	0.25	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.65 (at 2.69Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1760)	Depositor
R, R_{free}	0.173 , 0.211 0.178 , 0.214	Depositor DCC
R_{free} test set	5931 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å^2)	40.6	Xtriage
Anisotropy	0.064	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 46.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	14972	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.66% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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

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	0.18	0/340	0.74	0/525
2	V	0.60	1/345 (0.3%)	0.70	0/535
3	A	0.23	0/5807	0.39	0/7831
4	B	0.23	0/5940	0.40	0/8005
5	C	0.23	0/2071	0.41	0/2793
All	All	0.24	1/14503 (0.0%)	0.42	0/19689

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	V	1	A	OP3-P	-10.55	1.48	1.61

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	R	308	0	160	1	0
2	V	307	0	153	0	0
3	A	5691	0	5670	12	0
4	B	5828	0	5849	8	0
5	C	2029	0	2092	2	0
6	A	303	0	0	0	0
6	B	338	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	C	105	0	0	0	0
6	R	17	0	0	0	0
6	V	46	0	0	0	0
All	All	14972	0	13924	19	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 1.

All (19) 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)
3:A:358:LYS:NZ	3:A:474:GLY:O	2.35	0.59
1:R:18:U:OP1	4:B:203:ARG:NH2	2.36	0.58
3:A:305:LEU:HD12	3:A:305:LEU:C	2.27	0.56
4:B:586:ASP:N	4:B:586:ASP:OD1	2.43	0.52
4:B:340:ALA:HB3	4:B:341:PRO:HD3	1.94	0.50
4:B:518:GLU:OE2	4:B:571:ARG:NH1	2.40	0.47
3:A:483:ASN:HB2	3:A:497:TYR:HE1	1.80	0.46
3:A:305:LEU:HD13	3:A:494:ASP:HB3	1.98	0.46
3:A:305:LEU:O	3:A:305:LEU:HD12	2.17	0.45
4:B:676:THR:O	4:B:679:ARG:NH1	2.50	0.45
3:A:93:ARG:NH1	5:C:175:GLY:O	2.44	0.44
3:A:475:LYS:O	3:A:502:LYS:NZ	2.51	0.44
4:B:282:LEU:HD12	4:B:494:MET:HE1	1.99	0.43
3:A:72:ASN:OD1	3:A:73:LEU:N	2.45	0.43
4:B:534:ILE:O	5:C:142:ARG:NH2	2.46	0.43
3:A:609:VAL:HG12	3:A:609:VAL:O	2.20	0.42
3:A:216:VAL:HG12	3:A:226:MET:CE	2.50	0.41
3:A:533:TRP:N	3:A:534:PRO:CD	2.84	0.41
3:A:243:LEU:CD1	4:B:430:ILE:HD13	2.50	0.41

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
3	A	705/751 (94%)	685 (97%)	20 (3%)	0	100 100
4	B	738/772 (96%)	718 (97%)	19 (3%)	1 (0%)	51 78
5	C	248/798 (31%)	245 (99%)	3 (1%)	0	100 100
All	All	1691/2321 (73%)	1648 (98%)	42 (2%)	1 (0%)	51 78

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	B	410	MET

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
3	A	630/664 (95%)	625 (99%)	5 (1%)	81 93
4	B	638/657 (97%)	626 (98%)	12 (2%)	57 82
5	C	219/694 (32%)	213 (97%)	6 (3%)	44 74
All	All	1487/2015 (74%)	1464 (98%)	23 (2%)	65 86

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	73	LEU
3	A	120	GLU
3	A	305	LEU
3	A	449	CYS
3	A	613	LYS
4	B	94	ARG
4	B	104	GLN
4	B	126	ARG
4	B	233	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	B	347	LYS
4	B	412	MET
4	B	414	ASN
4	B	488	MET
4	B	655	ASP
4	B	666	ARG
4	B	678	GLN
4	B	713	MET
5	C	12	LEU
5	C	30	TYR
5	C	34	ARG
5	C	39	SER
5	C	125	MET
5	C	240	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	14/18 (77%)	0	0
2	V	13/14 (92%)	3 (23%)	0
All	All	27/32 (84%)	3 (11%)	0

All (3) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	V	4	A
2	V	7	A
2	V	11	A

There are no RNA pucker outliers to report.

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

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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, 95^{th} 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	15/18 (83%)	0.35	3 (20%) 1 0	33, 47, 134, 155	0
2	V	14/14 (100%)	-0.92	0 100 100	28, 32, 61, 75	0
3	A	709/751 (94%)	-0.39	20 (2%) 53 54	19, 37, 88, 129	0
4	B	742/772 (96%)	-0.51	14 (1%) 66 69	20, 36, 77, 131	0
5	C	250/798 (31%)	-0.54	2 (0%) 86 87	23, 40, 71, 97	0
All	All	1730/2353 (73%)	-0.46	39 (2%) 60 62	19, 37, 82, 155	0

All (39) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	B	196	ARG	6.4
3	A	546	VAL	5.7
1	R	4	A	5.1
4	B	194	LYS	4.1
4	B	645	THR	3.8
1	R	5	C	3.6
4	B	195	ASN	3.6
3	A	191	ILE	3.5
3	A	159	GLU	3.4
5	C	182	TRP	3.4
3	A	192	GLY	3.4
4	B	199	PHE	3.4
4	B	200	LEU	3.4
3	A	549	ARG	3.3
3	A	196	VAL	3.2
4	B	633	HIS	3.2
3	A	195	ASP	3.0
4	B	197	LYS	2.9
3	A	7	ARG	2.9
3	A	716	VAL	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	A	548	GLY	2.8
3	A	193	GLU	2.8
3	A	189	VAL	2.7
4	B	192	PRO	2.7
3	A	158	ASN	2.5
3	A	73	LEU	2.5
3	A	545	PHE	2.5
3	A	547	SER	2.4
3	A	141	ASN	2.4
3	A	197	GLU	2.4
3	A	140	GLY	2.3
5	C	134	ARG	2.3
4	B	737	ASP	2.3
1	R	6	C	2.2
3	A	198	LYS	2.2
4	B	193	ALA	2.1
4	B	294	PRO	2.1
4	B	741	LYS	2.1
4	B	410	MET	2.0

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.