Full wwPDB X-ray Structure Validation Report ⓘ

Mar 29, 2016 – 10:48 AM EDT

PDB ID	:	5F5F
Title	:	X-ray structure of Roquin ROQ domain in complex with a Selex-derived hexa-loop RNA motif
Authors	:	Janowski, R.; Schlundt, A.; Sattler, M.; Niessing, D.
Deposited on	:	2015-12-04
Resolution	:	3.00 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
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:

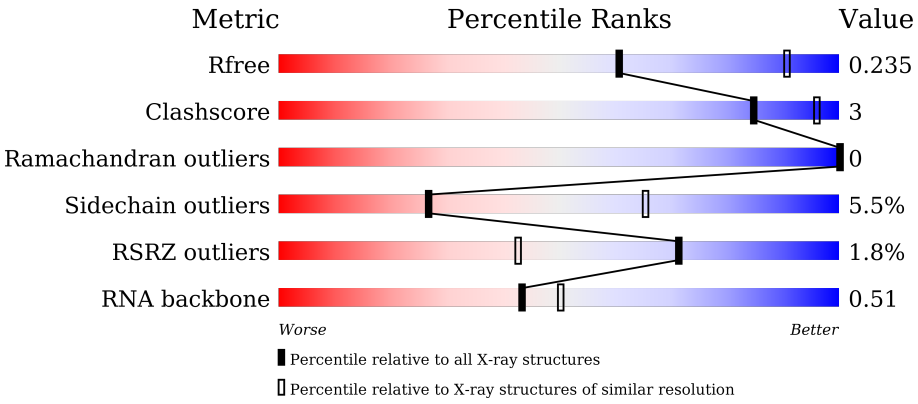
MolProbity	:	4.02b-467
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027107
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027107

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 3.00 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}	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)
RNA backbone	2183	1036 (3.40-2.60)

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 <=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	190	70% 8% • 20%
1	C	190	71% 8% •• 20%
1	E	190	72% 6% • 20%
1	G	190	6% 69% 7% • 22%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	20	75% 15% 10%
2	D	20	75% 15% 10%
2	F	20	65% 30% 5%
2	H	20	75% 25%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6577 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 Roquin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	152	Total	C	N	O	S	0	0	0
			1212	763	219	224	6			
1	C	152	Total	C	N	O	S	0	1	0
			1215	765	219	225	6			
1	E	152	Total	C	N	O	S	0	0	0
			1212	763	219	224	6			
1	G	148	Total	C	N	O	S	0	0	0
			1181	745	213	217	6			

- Molecule 2 is a RNA chain called RNA (5'-R(P*UP*GP*AP*CP*UP*GP*CP*GP*UP*UP*UP*UP*AP*GP*GP*AP*GP*UP*UP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	20	Total	C	N	O	P	0	0	0
			427	190	72	145	20			
2	D	20	Total	C	N	O	P	0	0	0
			427	190	72	145	20			
2	F	20	Total	C	N	O	P	0	0	0
			427	190	72	145	20			
2	H	20	Total	C	N	O	P	0	0	0
			427	190	72	145	20			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	15	Total	O	0	0
			15	15		
3	B	9	Total	O	0	0
			9	9		
3	C	11	Total	O	0	0
			11	11		
3	D	5	Total	O	0	0
			5	5		

Continued on next page...

Continued from previous page...

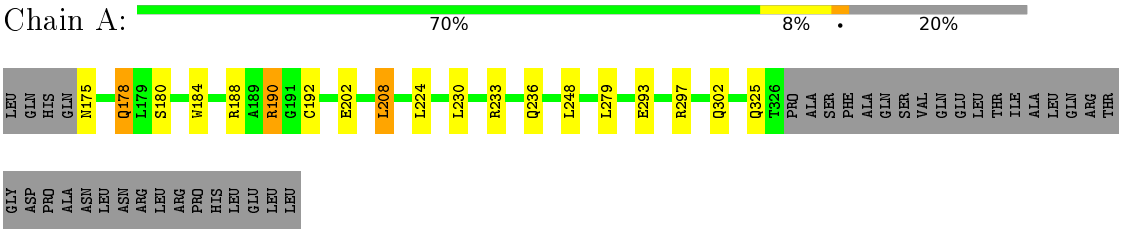
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	E	3	Total O 3 3	0	0
3	F	4	Total O 4 4	0	0
3	H	2	Total O 2 2	0	0

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 errors 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.

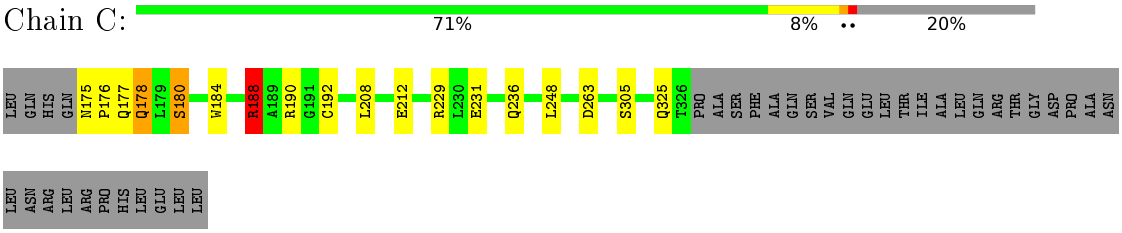
- Molecule 1: Roquin-1

Chain A:



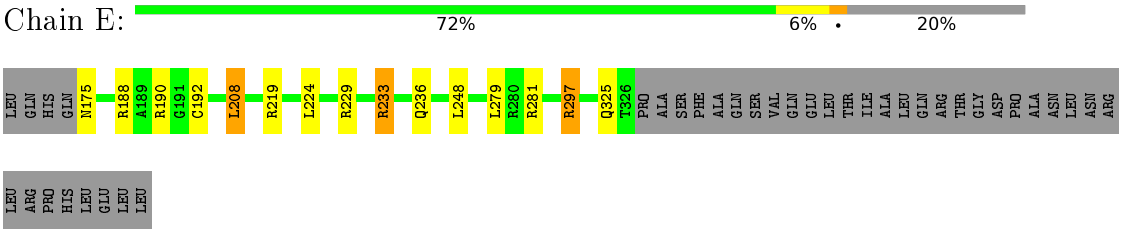
- Molecule 1: Roquin-1

Chain C:



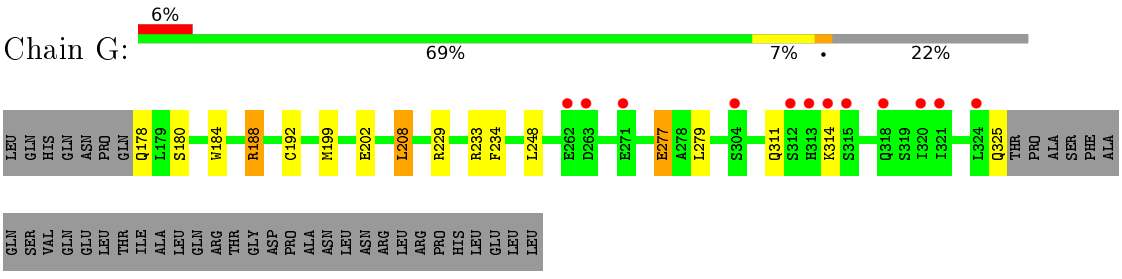
- Molecule 1: Roquin-1

Chain E:



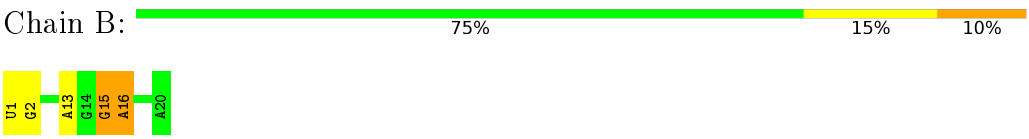
- Molecule 1: Roquin-1

Chain G:



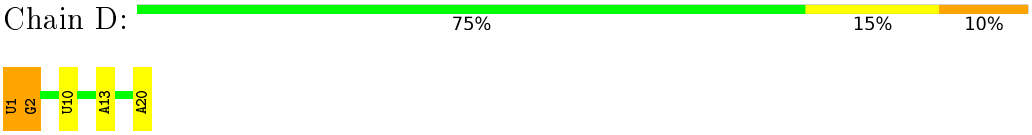
• Molecule 2: RNA (5'-R(P*UP*GP*AP*CP*UP*GP*CP*GP*UP*UP*UP*UP*AP*GP*GP*AP*GP*UP*UP*A)-3')

Chain B:



• Molecule 2: RNA (5'-R(P*UP*GP*AP*CP*UP*GP*CP*GP*UP*UP*UP*UP*AP*GP*GP*AP*GP*UP*UP*A)-3')

Chain D:

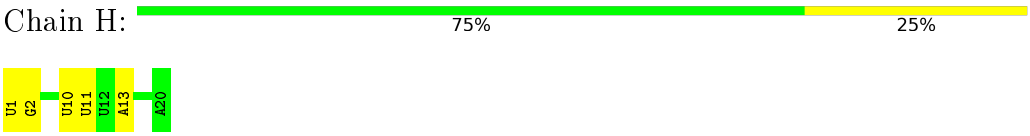


• Molecule 2: RNA (5'-R(P*UP*GP*AP*CP*UP*GP*CP*GP*UP*UP*UP*UP*AP*GP*GP*AP*GP*UP*UP*A)-3')

Chain F:

• Molecule 2: RNA (5'-R(P*UP*GP*AP*CP*UP*GP*CP*GP*UP*UP*UP*UP*AP*GP*GP*AP*GP*UP*UP*A)-3')

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	72.90Å 89.30Å 144.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.00 44.65 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (50.00-3.00) 100.0 (44.65-3.00)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.96 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.8.0123	Depositor
R, R_{free}	0.183 , 0.234 0.187 , 0.235	Depositor DCC
R_{free} test set	952 reflections (5.12%)	DCC
Wilson B-factor (Å^2)	60.4	Xtriage
Anisotropy	0.053	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 39.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Outliers	1 of 19550 reflections (0.005%)	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6577	wwPDB-VP
Average B, all atoms (Å^2)	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 21.12 % of the origin peak, indicating pseudo translational symmetry.The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 7.5026e-03.The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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	A	0.85	0/1232	0.91	1/1660 (0.1%)
1	C	0.87	1/1238 (0.1%)	0.99	5/1668 (0.3%)
1	E	0.79	0/1232	1.02	6/1660 (0.4%)
1	G	0.68	0/1200	0.91	5/1615 (0.3%)
2	B	0.98	2/476 (0.4%)	1.07	3/738 (0.4%)
2	D	0.83	1/476 (0.2%)	0.89	0/738
2	F	0.75	1/476 (0.2%)	0.95	0/738
2	H	0.72	1/476 (0.2%)	0.87	1/738 (0.1%)
All	All	0.81	6/6806 (0.1%)	0.96	21/9555 (0.2%)

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	U	OP3-P	-10.56	1.48	1.61
2	H	1	U	OP3-P	-10.36	1.48	1.61
2	D	1	U	OP3-P	-10.13	1.49	1.61
2	F	1	U	OP3-P	-8.45	1.51	1.61
1	C	231	GLU	CD-OE2	6.79	1.33	1.25
2	B	15	G	O3'-P	-6.63	1.53	1.61

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	16	A	O5'-P-OP2	-12.97	94.02	105.70
1	E	233	ARG	NE-CZ-NH1	8.48	124.54	120.30
1	C	188	ARG	NE-CZ-NH2	7.68	124.14	120.30
1	E	188	ARG	NE-CZ-NH1	-7.65	116.47	120.30
1	E	188	ARG	NE-CZ-NH2	7.47	124.03	120.30
2	B	15	G	C2'-C3'-O3'	7.43	125.84	109.50
1	G	188	ARG	NE-CZ-NH2	7.40	124.00	120.30
1	G	229	ARG	NE-CZ-NH1	6.98	123.79	120.30
1	G	188	ARG	NE-CZ-NH1	-6.44	117.08	120.30
1	C	188	ARG	NE-CZ-NH1	-6.09	117.25	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	G	188	ARG	CD-NE-CZ	5.85	131.79	123.60
1	C	188	ARG	CD-NE-CZ	5.74	131.64	123.60
2	B	16	A	P-O5'-C5'	-5.64	111.87	120.90
1	C	178	GLN	N-CA-C	5.60	126.11	111.00
1	E	190	ARG	NE-CZ-NH1	5.60	123.10	120.30
1	A	224	LEU	CB-CG-CD2	5.58	120.48	111.00
2	H	11	U	O4'-C1'-N1	-5.48	103.81	108.20
1	C	263	ASP	CB-CG-OD1	5.05	122.85	118.30
1	G	277	GLU	CA-CB-CG	5.05	124.50	113.40
1	E	297	ARG	CB-CA-C	5.04	120.49	110.40
1	E	224	LEU	CB-CG-CD2	5.02	119.54	111.00

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	A	1212	0	1232	8	1
1	C	1215	0	1237	4	1
1	E	1212	0	1232	3	0
1	G	1181	0	1204	14	0
2	B	427	0	213	0	0
2	D	427	0	212	12	0
2	F	427	0	211	9	0
2	H	427	0	213	0	0
3	A	15	0	0	1	0
3	B	9	0	0	0	0
3	C	11	0	0	0	0
3	D	5	0	0	0	0
3	E	3	0	0	0	0
3	F	4	0	0	0	0
3	H	2	0	0	0	0
All	All	6577	0	5754	35	1

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 3.

All (35) 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)
2:D:20:A:H2'	1:G:314:LYS:HD3	1.78	0.66
1:A:297:ARG:NH2	3:A:401:HOH:O	2.30	0.62
2:F:15:G:O2'	1:G:277:GLU:OE1	2.14	0.60
2:D:2:G:O2'	2:F:2:G:O2'	1.72	0.59
2:D:1:U:O4	1:G:314:LYS:NZ	2.32	0.58
2:D:20:A:C2'	1:G:314:LYS:HD3	2.37	0.54
2:F:20:A:N6	1:G:311:GLN:HB3	2.24	0.53
1:A:192:CYS:HB3	1:A:248:LEU:HD13	1.91	0.53
2:D:20:A:N3	1:G:314:LYS:NZ	2.53	0.53
1:E:192:CYS:HB3	1:E:248:LEU:HD13	1.92	0.51
1:G:192:CYS:HB3	1:G:248:LEU:HD13	1.92	0.51
1:C:192:CYS:HB3	1:C:248:LEU:HD13	1.93	0.51
1:G:199:MET:HG3	1:G:234:PHE:CD2	2.45	0.50
2:D:2:G:HO2'	2:F:2:G:HO2'	0.53	0.50
2:D:20:A:O2'	1:G:314:LYS:HD3	2.12	0.50
2:D:2:G:H1'	2:F:2:G:H1'	1.93	0.49
1:A:175:ASN:HB3	1:A:178:GLN:HB3	1.94	0.49
2:D:20:A:H2'	1:G:314:LYS:CD	2.42	0.48
2:D:2:G:HO2'	2:F:2:G:C2'	2.13	0.47
2:D:2:G:O2'	2:F:2:G:C2'	2.59	0.47
1:A:190:ARG:HD2	1:A:190:ARG:HA	1.52	0.46
1:C:184:TRP:O	1:C:188:ARG:HG2	2.16	0.46
1:A:184:TRP:O	1:A:188:ARG:HG2	2.15	0.46
1:G:184:TRP:O	1:G:188:ARG:HG2	2.16	0.45
1:G:202:GLU:OE1	1:G:233:ARG:NH2	2.49	0.45
2:D:20:A:H2'	1:G:314:LYS:CE	2.48	0.44
1:G:208:LEU:HD12	1:G:279:LEU:HD23	1.99	0.44
1:A:208:LEU:HD12	1:A:279:LEU:HD23	2.00	0.43
1:C:175:ASN:HB3	1:C:176:PRO:HD3	2.01	0.43
1:E:208:LEU:HD12	1:E:279:LEU:HD23	2.00	0.43
1:A:230:LEU:HA	1:A:230:LEU:HD12	1.93	0.42
1:E:219:ARG:NH1	2:F:8:G:N7	2.68	0.42
1:A:202:GLU:OE1	1:A:233:ARG:NH2	2.53	0.41
1:C:177:GLN:HA	1:C:180:SER:OG	2.21	0.41
2:F:14:G:H2'	2:F:15:G:O4'	2.21	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:297:ARG:NH1	1:C:212:GLU:OE1[3_755]	2.19	0.01

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	150/190 (79%)	149 (99%)	1 (1%)	0	100	100
1	C	151/190 (80%)	151 (100%)	0	0	100	100
1	E	150/190 (79%)	149 (99%)	1 (1%)	0	100	100
1	G	146/190 (77%)	146 (100%)	0	0	100	100
All	All	597/760 (79%)	595 (100%)	2 (0%)	0	100	100

There are no Ramachandran outliers to report.

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	133/166 (80%)	125 (94%)	8 (6%)	24	62
1	C	134/166 (81%)	125 (93%)	9 (7%)	20	57
1	E	133/166 (80%)	125 (94%)	8 (6%)	24	62
1	G	129/166 (78%)	125 (97%)	4 (3%)	47	83
All	All	529/664 (80%)	500 (94%)	29 (6%)	27	65

All (29) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	178	GLN
1	A	180	SER
1	A	190	ARG
1	A	208	LEU
1	A	236	GLN
1	A	293	GLU
1	A	302	GLN
1	A	325	GLN
1	C	178	GLN
1	C	180	SER
1	C	188	ARG
1	C	190	ARG
1	C	208	LEU
1	C	229	ARG
1	C	236	GLN
1	C	305	SER
1	C	325	GLN
1	E	175	ASN
1	E	208	LEU
1	E	229	ARG
1	E	233	ARG
1	E	236	GLN
1	E	281	ARG
1	E	297	ARG
1	E	325	GLN
1	G	178	GLN
1	G	180	SER
1	G	208	LEU
1	G	325	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	175	ASN
1	A	178	GLN
1	A	318	GLN
1	C	228	GLN
1	C	236	GLN
1	C	318	GLN
1	C	325	GLN
1	E	325	GLN
1	G	178	GLN
1	G	228	GLN

5.3.3 RNA

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	19/20 (95%)	4 (21%)	1 (5%)
2	D	19/20 (95%)	3 (15%)	0
2	F	19/20 (95%)	2 (10%)	0
2	H	19/20 (95%)	3 (15%)	0
All	All	76/80 (95%)	12 (15%)	1 (1%)

All (12) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	2	G
2	B	13	A
2	B	15	G
2	B	16	A
2	D	2	G
2	D	10	U
2	D	13	A
2	F	2	G
2	F	13	A
2	H	2	G
2	H	10	U
2	H	13	A

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	B	15	G

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 (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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	A	152/190 (80%)	-0.43	0 100 100	34, 50, 84, 140	0
1	C	152/190 (80%)	-0.44	0 100 100	35, 51, 91, 124	0
1	E	152/190 (80%)	-0.38	0 100 100	45, 66, 99, 119	0
1	G	148/190 (77%)	0.48	12 (8%) 15 5	63, 90, 124, 142	0
2	B	20/20 (100%)	-0.62	0 100 100	39, 48, 84, 91	0
2	D	20/20 (100%)	-0.51	0 100 100	42, 51, 116, 128	1 (5%)
2	F	20/20 (100%)	-0.25	0 100 100	55, 67, 113, 119	1 (5%)
2	H	20/20 (100%)	0.51	0 100 100	67, 79, 99, 120	0
All	All	684/840 (81%)	-0.20	12 (1%) 71 43	34, 63, 114, 142	2 (0%)

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	262	GLU	4.4
1	G	271	GLU	3.6
1	G	312	SER	3.3
1	G	314	LYS	2.9
1	G	263	ASP	2.8
1	G	304	SER	2.6
1	G	315	SER	2.5
1	G	324	LEU	2.4
1	G	320	ILE	2.2
1	G	321	ILE	2.2
1	G	313	HIS	2.0
1	G	318	GLN	2.0

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

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

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.