



wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 10:00 am BST

PDB ID : 2I82
Title : Crystal structure of pseudouridine synthase RluA: indirect sequence readout through protein-induced RNA structure
Authors : Hoang, C.
Deposited on : 2006-08-31
Resolution : 2.05 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

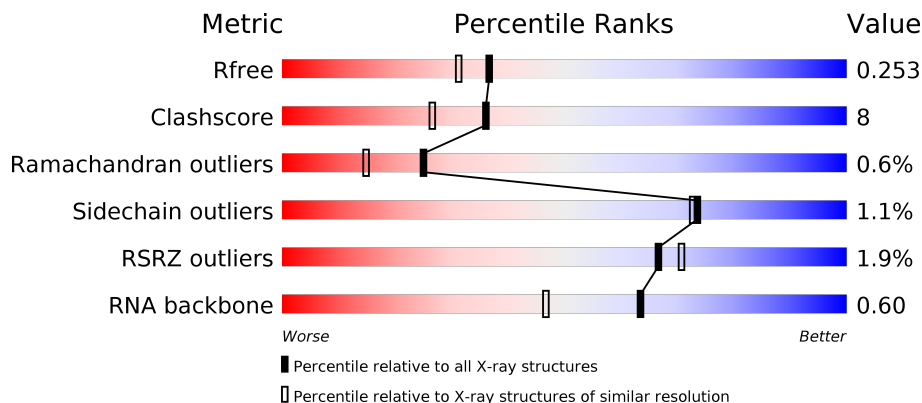
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.05 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	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)
RSRZ outliers	127900	1672 (2.04-2.04)
RNA backbone	3102	1018 (2.50-1.58)

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	E	21	
1	F	21	
1	G	21	
1	H	21	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	A	217	 84% 15%
2	B	217	 % 89% 9% •
2	C	217	 5% 73% 24% ••
2	D	217	 2% 72% 27%

2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 8854 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 5'-R(*GP*AP*GP*GP*GP*GP*AP*UP*UP*GP*AP*AP*AP*AP*UP*CP*CP*CP*CP*UP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	21	Total	C	N	O	P	0	0	0
			439	197	81	141	20			
1	F	21	Total	C	N	O	P	0	0	0
			439	197	81	141	20			
1	G	21	Total	C	N	O	P	0	0	0
			439	197	81	141	20			
1	H	21	Total	C	N	O	P	0	0	0
			429	192	76	141	20			

- Molecule 2 is a protein called Ribosomal large subunit pseudouridine synthase A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
2	A	217	Total	C	N	O	S	Se	0	2	0
			1748	1108	317	313	2	8			
2	B	214	Total	C	N	O	S	Se	0	3	0
			1703	1085	307	301	3	7			
2	C	215	Total	C	N	O	S	Se	0	0	0
			1684	1072	302	301	2	7			
2	D	216	Total	C	N	O	S	Se	0	1	0
			1698	1079	305	305	2	7			

There are 32 discrepancies between the modelled and reference sequences:

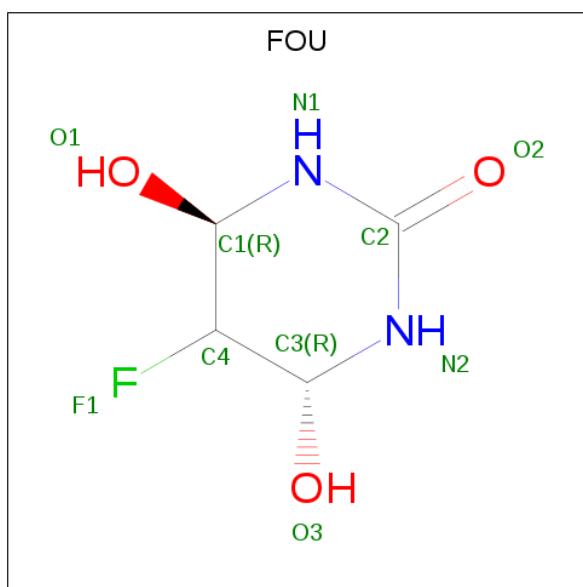
Chain	Residue	Modelled	Actual	Comment	Reference
A	3	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
A	24	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
A	47	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
A	65	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
A	168	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
A	189	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
A	199	MSE	MET	MODIFIED RESIDUE	UNP P0AA37

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	211	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
B	303	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
B	324	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
B	347	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
B	365	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
B	468	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
B	489	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
B	499	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
B	511	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
C	603	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
C	624	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
C	647	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
C	665	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
C	768	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
C	789	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
C	799	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
C	811	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
D	903	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
D	924	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
D	947	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
D	965	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
D	1068	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
D	1089	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
D	1099	MSE	MET	MODIFIED RESIDUE	UNP P0AA37
D	1111	MSE	MET	MODIFIED RESIDUE	UNP P0AA37

- Molecule 3 is (5S,6R)-5-FLUORO-6-HYDROXYDIHYDROPYRIMIDINE-2,4(1H,3H)-DIONE (three-letter code: FOU) (formula: C₄H₇N₂O₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	F	N			O
3	E	1	10	4	1	2	3	0	0
3	G	1	10	4	1	2	3	0	0

- Molecule 4 is water.

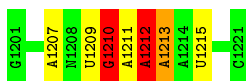
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	E	15	Total	O	0	0
			15	15		
4	F	28	Total	O	0	0
			28	28		
4	G	6	Total	O	0	0
			6	6		
4	H	10	Total	O	0	0
			10	10		
4	A	79	Total	O	0	0
			79	79		
4	B	66	Total	O	0	0
			66	66		
4	C	26	Total	O	0	0
			26	26		
4	D	25	Total	O	0	0
			25	25		

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: 5'-R(*GP*AP*GP*GP*GP*GP*AP*UP*UP*GP*AP*AP*AP*AP*UP*CP*CP*CP*CP*UP*C)-3'

Chain E: 



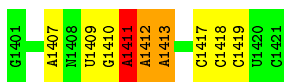
- Molecule 1: 5'-R(*GP*AP*GP*GP*GP*GP*AP*UP*UP*GP*AP*AP*AP*AP*UP*CP*CP*CP*CP*UP*C)-3'

Chain F: 



- Molecule 1: 5'-R(*GP*AP*GP*GP*GP*GP*AP*UP*UP*GP*AP*AP*AP*AP*UP*CP*CP*CP*CP*UP*C)-3'

Chain G: 




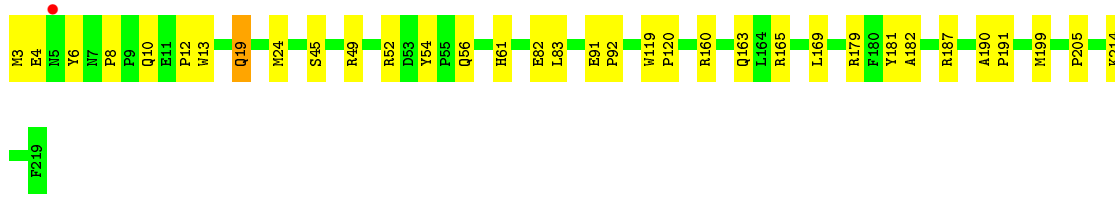
- Molecule 1: 5'-R(*GP*AP*GP*GP*GP*GP*AP*UP*UP*GP*AP*AP*AP*AP*UP*CP*CP*CP*CP*UP*C)-3'

Chain H: 



- Molecule 2: Ribosomal large subunit pseudouridine synthase A

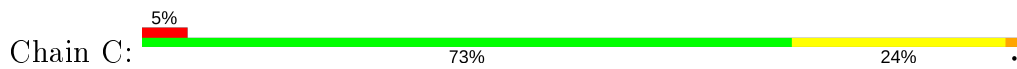
Chain A: 



• Molecule 2: Ribosomal large subunit pseudouridine synthase A



• Molecule 2: Ribosomal large subunit pseudouridine synthase A



• Molecule 2: Ribosomal large subunit pseudouridine synthase A



• Molecule 2: Ribosomal large subunit pseudouridine synthase A



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	320.74Å 51.69Å 81.23Å 90.00° 90.81° 90.00°	Depositor
Resolution (Å)	19.83 – 2.05 19.83 – 2.05	Depositor EDS
% Data completeness (in resolution range)	97.8 (19.83-2.05) 94.1 (19.83-2.05)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.26 (at 2.06Å)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.234 , 0.263 0.224 , 0.253	Depositor DCC
R_{free} test set	16072 reflections (9.92%)	wwPDB-VP
Wilson B-factor (Å ²)	35.2	Xtrriage
Anisotropy	0.179	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 28.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.41$, $\langle L^2 \rangle = 0.24$	Xtrriage
Estimated twinning fraction	0.109 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	8854	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.35% 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: FOU

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	E	0.35	0/477	0.96	4/740 (0.5%)
1	F	0.40	0/477	1.00	5/740 (0.7%)
1	G	0.24	0/477	0.90	1/740 (0.1%)
1	H	0.27	0/465	0.92	2/721 (0.3%)
2	A	0.36	0/1787	0.62	0/2417
2	B	0.35	0/1742	0.63	0/2361
2	C	0.29	0/1723	0.60	0/2338
2	D	0.30	0/1735	0.60	0/2351
All	All	0.33	0/8883	0.71	12/12408 (0.1%)

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	1411	A	C2'-C3'-O3'	7.38	125.74	109.50
1	E	1212	A	N9-C1'-C2'	6.31	122.21	114.00
1	E	1210	G	N9-C1'-C2'	6.22	122.09	114.00
1	F	1307	A	O4'-C1'-N9	6.11	113.09	108.20
1	F	1309	U	O5'-P-OP1	-5.71	100.56	105.70

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	E	439	0	226	4	0
1	F	439	0	227	3	0
1	G	439	0	225	5	0
1	H	429	0	221	8	0
2	A	1748	0	1743	25	0
2	B	1703	0	1689	14	0
2	C	1684	0	1657	42	0
2	D	1698	0	1683	40	0
3	E	10	0	5	1	0
3	G	10	0	5	0	0
4	A	79	0	0	4	0
4	B	66	0	0	2	0
4	C	26	0	0	1	0
4	D	25	0	0	1	0
4	E	15	0	0	1	0
4	F	28	0	0	0	0
4	G	6	0	0	0	0
4	H	10	0	0	1	0
All	All	8854	0	7681	130	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 8.

The worst 5 of 130 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:C:799:MSE:HE2	2:C:814:LYS:HE2	1.39	1.03
2:D:1099:MSE:HE2	2:D:1114:LYS:HE2	1.49	0.93
2:A:199:MSE:HG3	2:A:214:LYS:HG2	1.56	0.87
2:D:1022:ARG:CB	2:D:1023:PRO:HD3	2.10	0.82
2:B:347:MSE:O	2:B:351:GLN:HG3	1.86	0.76

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	217/217 (100%)	209 (96%)	7 (3%)	1 (0%)	29	18
2	B	215/217 (99%)	209 (97%)	6 (3%)	0	100	100
2	C	213/217 (98%)	196 (92%)	14 (7%)	3 (1%)	11	3
2	D	215/217 (99%)	202 (94%)	12 (6%)	1 (0%)	29	18
All	All	860/868 (99%)	816 (95%)	39 (4%)	5 (1%)	25	15

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	4	GLU
2	C	732	GLY
2	D	1022	ARG
2	C	730	GLU
2	C	733	LYS

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	
2	A	187/179 (104%)	185 (99%)	2 (1%)	73	73
2	B	179/179 (100%)	178 (99%)	1 (1%)	86	87
2	C	176/179 (98%)	172 (98%)	4 (2%)	50	44
2	D	179/179 (100%)	178 (99%)	1 (1%)	86	87
All	All	721/716 (101%)	713 (99%)	8 (1%)	73	73

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	649	ARG
2	D	961	HIS
2	C	736	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	361	HIS
2	C	661	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	322	HIS
2	B	436	GLN
2	D	1021	ASN
2	B	319	GLN
2	D	942	HIS

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	E	19/21 (90%)	5 (26%)	0
1	F	19/21 (90%)	5 (26%)	1 (5%)
1	G	19/21 (90%)	5 (26%)	2 (10%)
1	H	19/21 (90%)	6 (31%)	1 (5%)
All	All	76/84 (90%)	21 (27%)	4 (5%)

5 of 21 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	E	1209	U
1	E	1210	G
1	E	1211	A
1	E	1212	A
1	E	1213	A

All (4) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	F	1307	A
1	G	1407	A
1	G	1411	A
1	H	1507	A

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](#)

2 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	FOU	E	3001	1	8,10,10	2.78	3 (37%)	4,14,14	1.20	1 (25%)
3	FOU	G	3003	1	8,10,10	2.55	3 (37%)	4,14,14	0.44	0

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	FOU	E	3001	1	-	-	0/1/1/1
3	FOU	G	3003	1	-	-	0/1/1/1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	3001	FOU	C2-N2	4.74	1.44	1.34
3	E	3001	FOU	C2-N1	4.61	1.44	1.34
3	G	3003	FOU	C2-N1	4.48	1.44	1.34
3	G	3003	FOU	C2-N2	4.08	1.43	1.34
3	E	3001	FOU	O3-C3	-3.96	1.22	1.40

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
3	E	3001	FOU	F1-C4-C1	-2.27	106.72	108.81

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	3001	FOU	1	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	E	20/21 (95%)	-0.76	0 100 100	26, 38, 49, 56	0
1	F	20/21 (95%)	-0.80	0 100 100	22, 39, 44, 48	0
1	G	20/21 (95%)	-0.45	0 100 100	44, 60, 68, 70	0
1	H	20/21 (95%)	-0.40	0 100 100	43, 59, 75, 83	0
2	A	209/217 (96%)	-0.08	1 (0%) 91 92	21, 31, 50, 77	0
2	B	207/217 (95%)	-0.03	2 (0%) 82 84	17, 33, 49, 61	0
2	C	208/217 (95%)	0.47	10 (4%) 30 33	28, 49, 65, 73	0
2	D	209/217 (96%)	0.31	4 (1%) 66 71	27, 44, 61, 70	0
All	All	913/952 (95%)	0.10	17 (1%) 66 71	17, 40, 62, 83	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	957	ALA	4.0
2	C	727	VAL	3.5
2	C	772	GLY	2.7
2	D	938	ARG	2.6
2	C	640	GLU	2.6

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	FOU	G	3003	10/10	0.72	0.27	63,66,69,69	0
3	FOU	E	3001	10/10	0.80	0.26	54,56,59,63	0

6.5 Other polymers [i](#)

There are no such residues in this entry.