



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 12:20 PM GMT

PDB ID : 2E5L  
Title : A snapshot of the 30S ribosomal subunit capturing mRNA via the Shine-Dalgarno interaction  
Authors : Kaminishi, T.; Wilson, D.N.; Takemoto, C.; Harms, J.M.; Kawazoe, M.; Schlunzen, F.; Hanawa-Suetsugu, K.; Shirouzu, M.; Fucini, P.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)  
Deposited on : 2006-12-21  
Resolution : 3.30 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

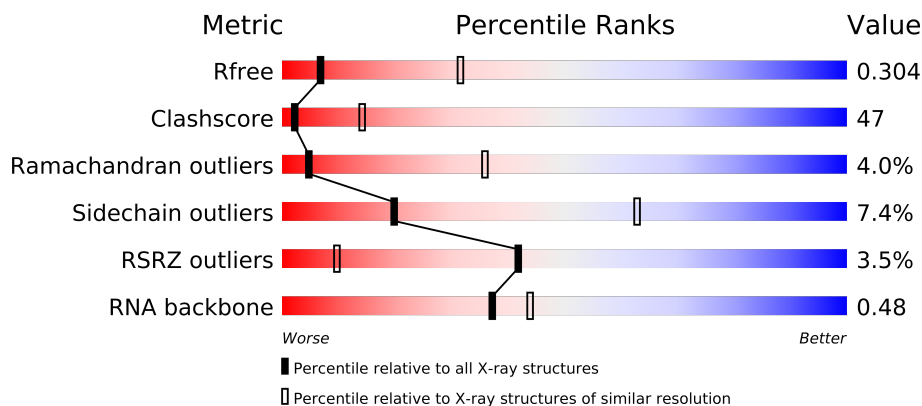
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 3.30 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))
$R_{free}$	66092	1341 (3.40-3.20)
Clashscore	79885	1696 (3.40-3.20)
Ramachandran outliers	78287	1664 (3.40-3.20)
Sidechain outliers	78261	1662 (3.40-3.20)
RSRZ outliers	66119	1342 (3.40-3.20)
RNA backbone	1838	1042 (3.90-2.70)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	1520	
2	1	6	
2	2	6	
3	B	227	
4	C	238	
5	D	208	
6	E	161	
7	F	101	
8	G	155	
9	H	138	
10	I	128	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
11	J	104	
12	K	128	
13	L	131	
14	M	125	
15	N	60	
16	O	88	
17	P	88	
18	Q	104	
19	R	87	
20	S	92	
21	T	105	
22	V	26	

## 2 Entry composition

There are 23 unique types of molecules in this entry. The entry contains 51895 atoms, of which 0 are hydrogen and 0 are deuterium.

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 16S ribosomal RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1517	Total	C	N	O	P	0	0	0
			32594	14508	6027	10542	1517			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	1	6	Total	C	N	O	P	0	0	0
			131	60	30	36	5			
2	2	4	Total	C	N	O	P	0	0	0
			86	40	20	23	3			

- Molecule 3 is a protein called 30S ribosomal protein S2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	222	Total	C	N	O	S	0	0	0
			1811	1154	328	324	5			

- Molecule 4 is a protein called 30S ribosomal protein S3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	C	206	Total	C	N	O	S	0	0	0
			1612	1016	314	281	1			

- Molecule 5 is a protein called 30S ribosomal protein S4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	D	208	Total	C	N	O	S	0	0	0
			1703	1066	339	291	7			

- Molecule 6 is a protein called 30S ribosomal protein S5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	E	150	Total	C	N	O	S	0	0	0
			1146	724	217	201	4			

- Molecule 7 is a protein called 30S ribosomal protein S6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	F	101	Total	C	N	O	S	0	0	0
			843	531	155	154	3			

- Molecule 8 is a protein called 30S ribosomal protein S7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	G	155	Total	C	N	O	S	0	0	0
			1257	781	252	218	6			

- Molecule 9 is a protein called 30S ribosomal protein S8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	H	138	Total	C	N	O	S	0	0	0
			1116	705	215	193	3			

- Molecule 10 is a protein called 30S ribosomal protein S9.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
10	I	127	Total	C	N	O	0	0	0
			1011	639	198	174			

- Molecule 11 is a protein called 30S ribosomal protein S10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	J	98	Total	C	N	O	S	0	0	0
			794	499	156	138	1			

- Molecule 12 is a protein called 30S ribosomal protein S11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	K	115	Total	C	N	O	S	0	0	0
			853	531	160	159	3			

- Molecule 13 is a protein called 30S ribosomal protein S12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	L	124	Total	C	N	O	S	0	0	0
			970	611	195	163	1			

- Molecule 14 is a protein called 30S ribosomal protein S13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	M	122	Total	C	N	O	S	0	0	0
			969	600	200	167	2			

- Molecule 15 is a protein called 30S ribosomal protein S14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	N	60	Total	C	N	O	S	0	0	0
			492	312	104	72	4			

- Molecule 16 is a protein called 30S ribosomal protein S15.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
16	O	88	Total	C	N	O	S	0	0	0
			734	459	147	126	2			

- Molecule 17 is a protein called 30S ribosomal protein S16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
17	P	83	Total	C	N	O	S	0	0	0
			700	443	139	117	1			

- Molecule 18 is a protein called 30S ribosomal protein S17.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
18	Q	104	Total	C	N	O	S	0	0	0
			857	547	161	147	2			

- Molecule 19 is a protein called 30S ribosomal protein S18.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
19	R	73	Total	C	N	O		0	0	0
			597	380	118	99				

- Molecule 20 is a protein called 30S ribosomal protein S19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
20	S	80	Total	C	N	O	S	0	0	0
			647	414	119	112	2			

- Molecule 21 is a protein called 30S ribosomal protein S20.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
21	T	99	Total	C	N	O	S	0	0	0
			762	469	162	129	2			

- Molecule 22 is a protein called 30S ribosomal protein Thx.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
22	V	24	Total	C	N	O	0	0	0
			208	128	50	30			

- Molecule 23 is ZINC ION (three-letter code: ZN) (formula: Zn).

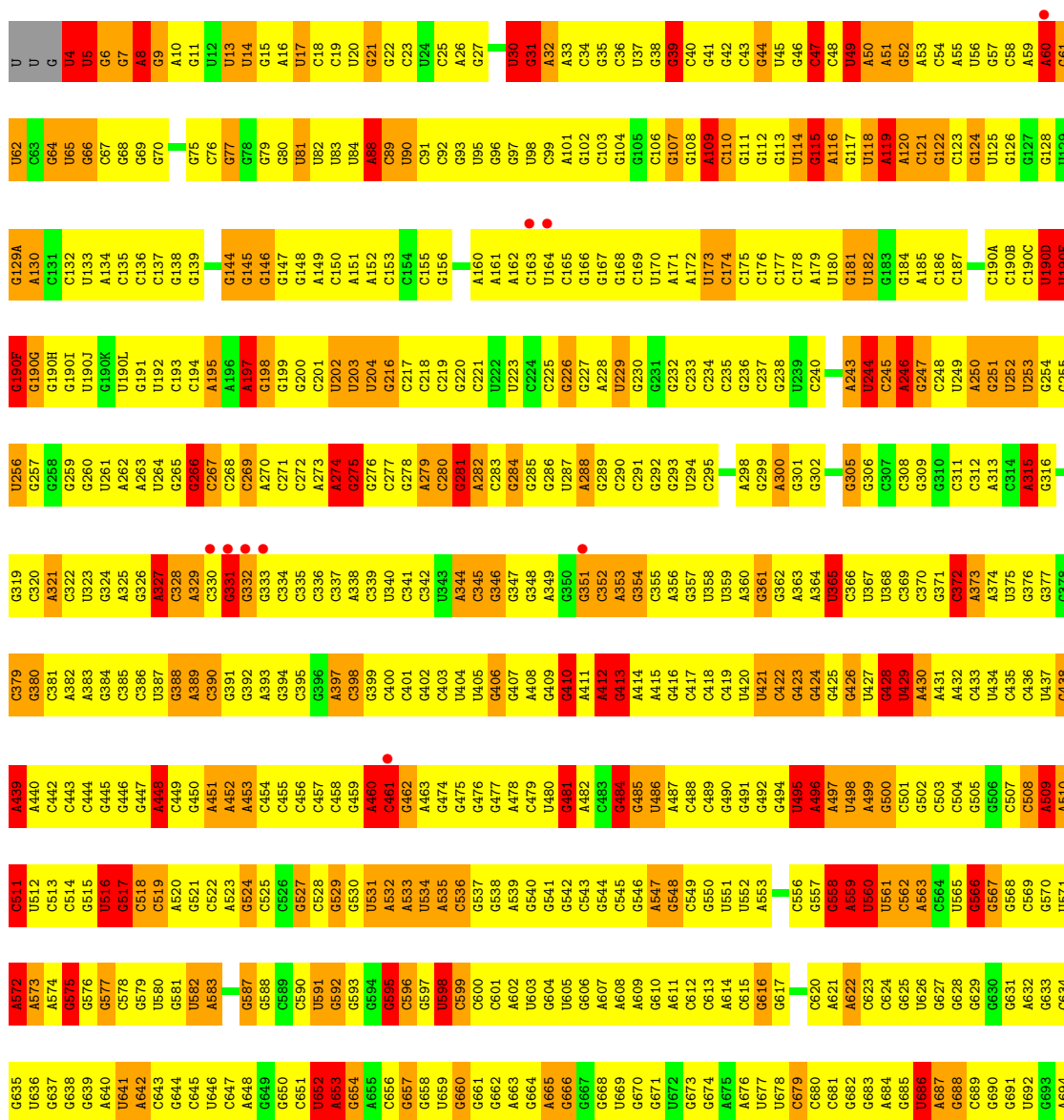
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
23	D	1	Total	Zn	0	0
			1	1		
23	N	1	Total	Zn	0	0
			1	1		

### 3 Residue-property plots

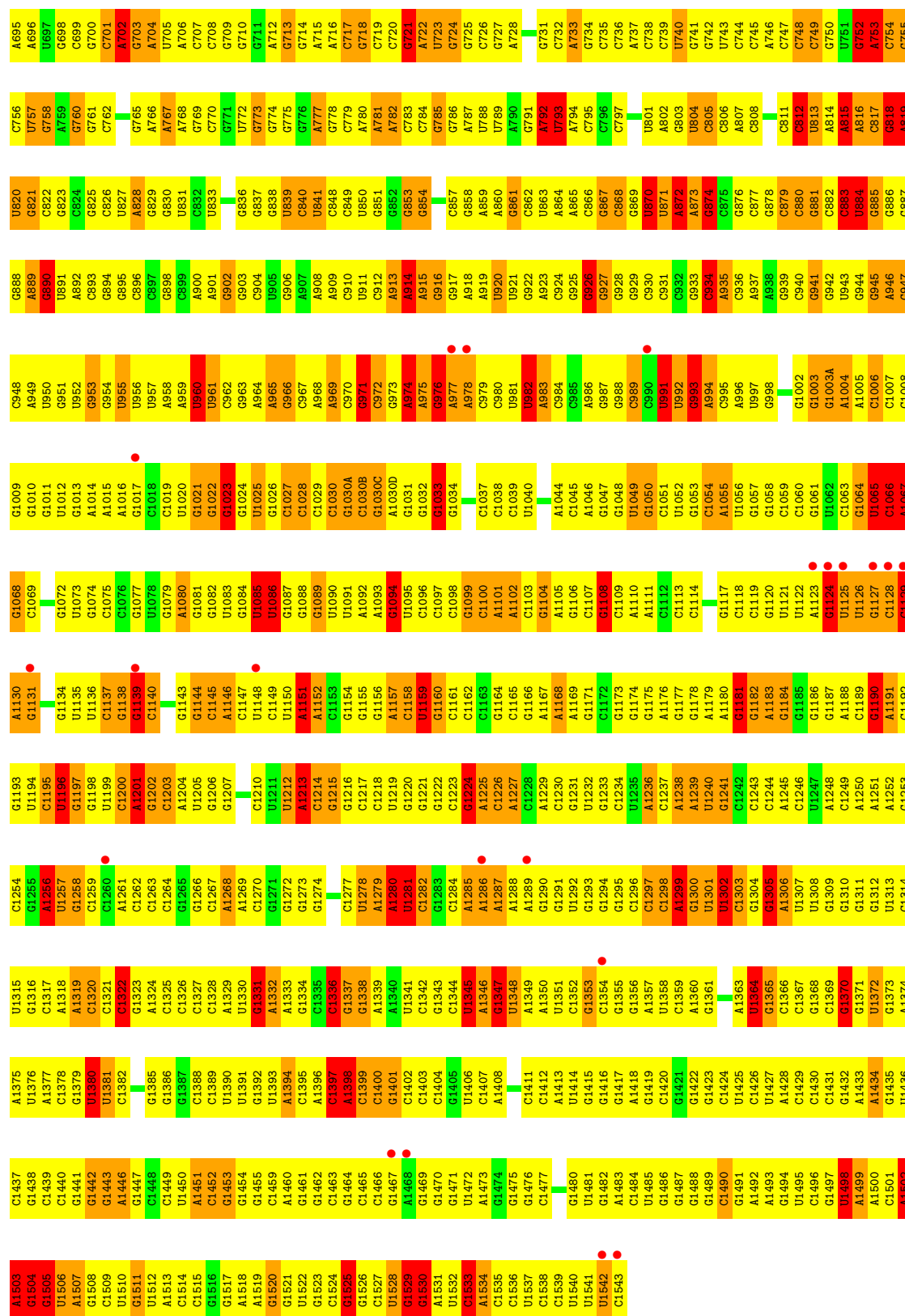
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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.

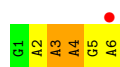
#### • Molecule 1: 16S ribosomal RNA

Chain A:



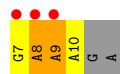






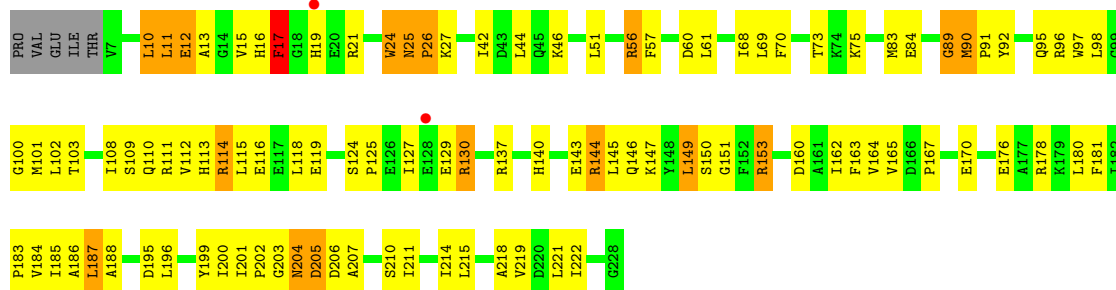
- Molecule 2: 5'-R(\*GP\*AP\*AP\*AP\*GP\*A)-3'

Chain 2:



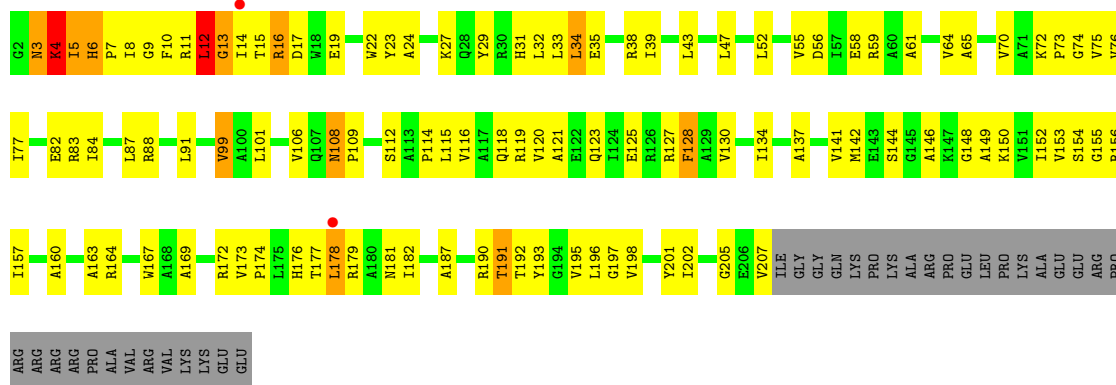
- Molecule 3: 30S ribosomal protein S2

Chain B:



- Molecule 4: 30S ribosomal protein S3

Chain C:

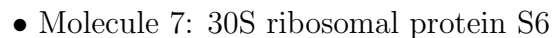


- Molecule 5: 30S ribosomal protein S4

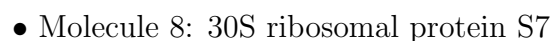
Chain D:



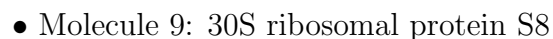
- Molecule 6: 30S ribosomal protein S5



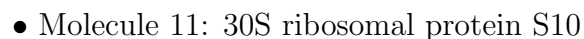
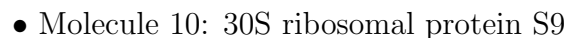
Age Group	Percentage
18-24	15%
25-34	20%
35-44	25%
45-54	20%
55-64	15%
65-74	10%
75-84	5%
85+	5%

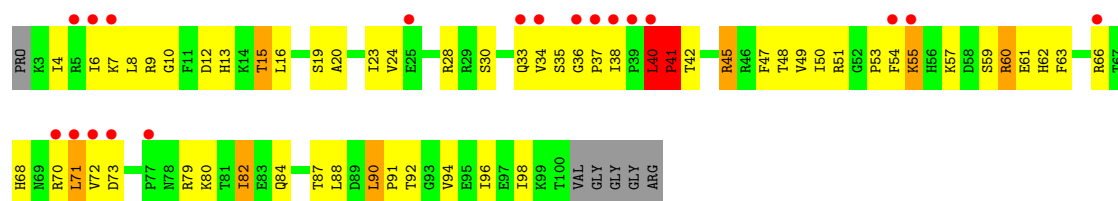


\_\_\_\_\_



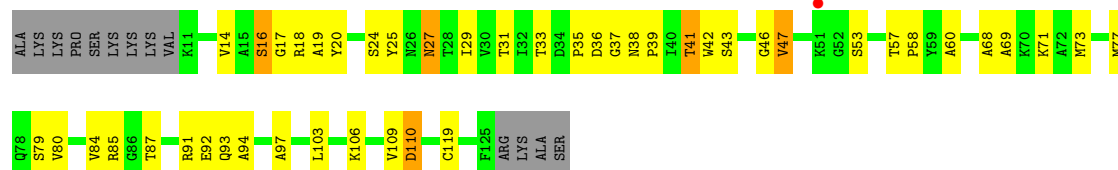
Age Group	Percentage
18-24	2%
25-34	55%
35-44	35%
45-54	8%
55-64	1%
65-74	1%
75-84	1%
85+	1%





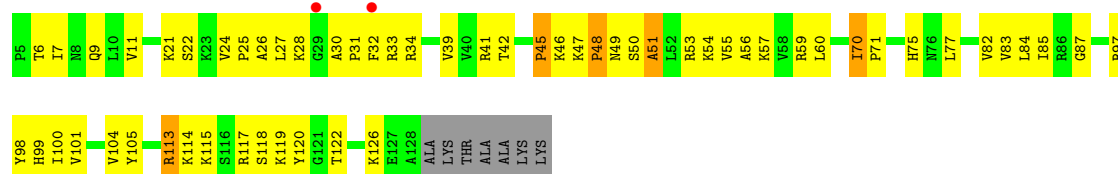
• Molecule 12: 30S ribosomal protein S11

Chain K:



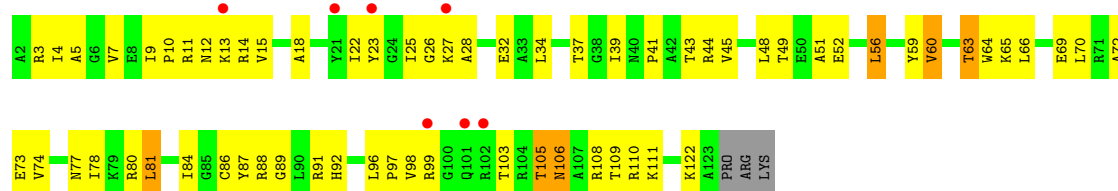
• Molecule 13: 30S ribosomal protein S12

Chain L:



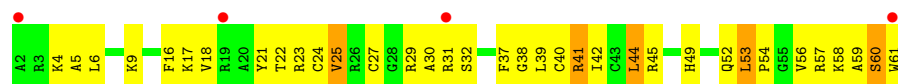
• Molecule 14: 30S ribosomal protein S13

Chain M:



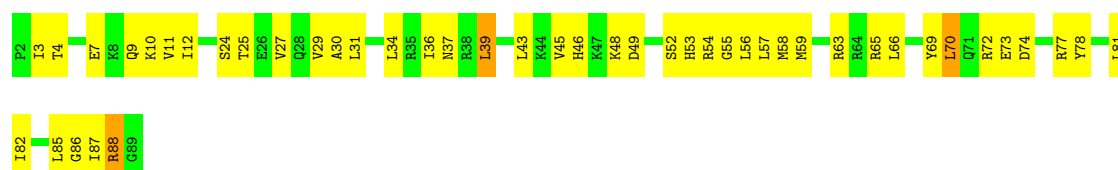
• Molecule 15: 30S ribosomal protein S14

Chain N:



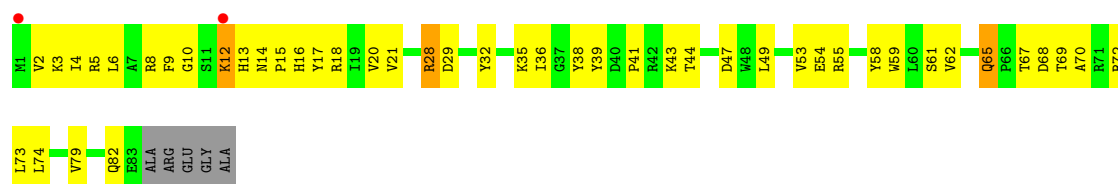
• Molecule 16: 30S ribosomal protein S15

Chain O:



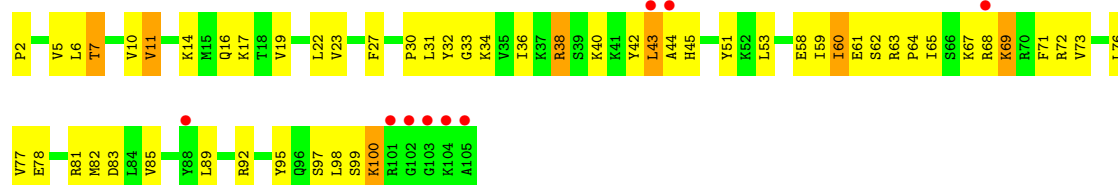
- Molecule 17: 30S ribosomal protein S16

Chain P:



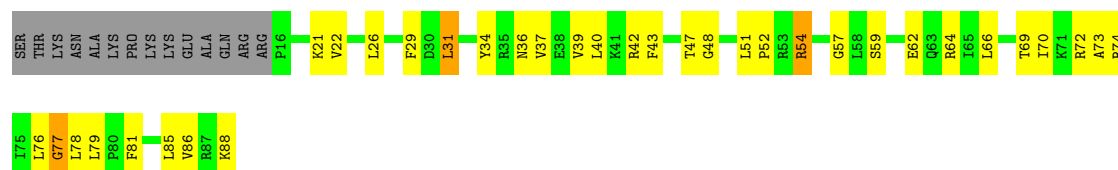
- Molecule 18: 30S ribosomal protein S17

Chain Q:



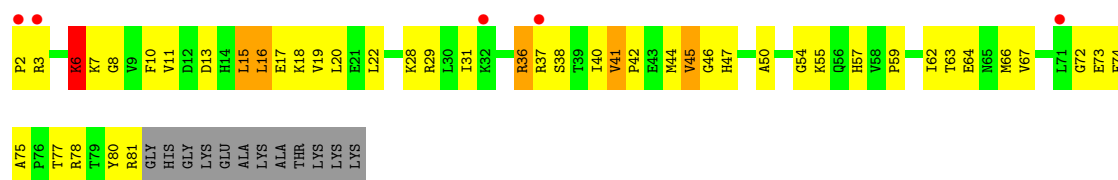
- Molecule 19: 30S ribosomal protein S18

Chain R:



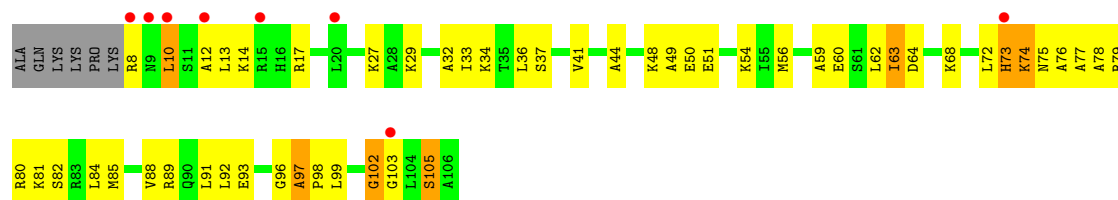
- Molecule 20: 30S ribosomal protein S19

Chain S:



- Molecule 21: 30S ribosomal protein S20

Chain T:



- Molecule 22: 30S ribosomal protein Thx

Chain V:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	411.79Å 411.79Å 173.16Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	148.83 – 3.30 148.83 – 3.30	Depositor EDS
% Data completeness (in resolution range)	97.0 (148.83-3.30) 97.1 (148.83-3.30)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.23 (at 3.33Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.259 , 0.301 0.271 , 0.304	Depositor DCC
$R_{free}$ test set	10897 reflections (5.34%)	DCC
Wilson B-factor (Å <sup>2</sup> )	96.4	Xtriage
Anisotropy	0.310	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.18 , 47.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.38$ , $\langle L^2 \rangle = 0.21$	Xtriage
Outliers	0 of 214963 reflections	Xtriage
$F_o, F_c$ correlation	0.84	EDS
Total number of atoms	51895	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	104.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.54% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section:  
ZN

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	0.97	19/36482 (0.1%)	0.99	145/56937 (0.3%)
2	1	0.51	0/148	0.76	0/230
2	2	0.42	0/97	0.73	0/150
3	B	0.67	0/1843	0.92	5/2479 (0.2%)
4	C	0.63	0/1636	0.89	2/2205 (0.1%)
5	D	0.80	3/1733 (0.2%)	0.96	8/2318 (0.3%)
6	E	0.82	0/1162	0.95	2/1564 (0.1%)
7	F	0.52	0/856	0.78	0/1154
8	G	0.50	0/1276	0.67	0/1709
9	H	0.76	0/1136	1.00	2/1527 (0.1%)
10	I	0.53	0/1029	0.78	0/1378
11	J	0.57	0/807	0.89	3/1085 (0.3%)
12	K	0.53	0/868	0.79	0/1173
13	L	0.62	0/986	0.85	0/1320
14	M	0.53	0/979	0.78	0/1310
15	N	0.66	0/501	0.93	1/664 (0.2%)
16	O	0.61	0/745	0.87	0/992
17	P	0.62	0/716	0.83	0/963
18	Q	0.74	0/870	0.92	1/1159 (0.1%)
19	R	0.59	0/603	0.86	0/799
20	S	0.51	0/661	0.82	0/890
21	T	0.49	0/764	0.73	0/1006
22	V	0.56	0/212	0.72	0/277
All	All	0.86	22/56110 (0.0%)	0.95	169/83289 (0.2%)

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	1	127

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	12	CYS	CB-SG	9.54	1.98	1.82
5	D	12	CYS	CA-CB	8.62	1.73	1.53
1	A	1108	G	C5-C6	7.41	1.49	1.42
1	A	660	G	C5-C6	-6.55	1.35	1.42
1	A	361	G	C5-C6	-6.41	1.35	1.42

The worst 5 of 169 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	12	CYS	CA-CB-SG	14.39	139.90	114.00
1	A	511	C	N1-C1'-C2'	12.07	129.70	114.00
1	A	934	C	N1-C1'-C2'	9.87	126.84	114.00
1	A	246	A	N9-C1'-C2'	9.69	126.60	114.00
1	A	1151	A	N9-C1'-C2'	9.47	126.31	114.00

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	511	C	C1'

5 of 127 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	30	U	Sidechain
1	A	39	G	Sidechain
1	A	47	C	Sidechain
1	A	49	U	Sidechain
1	A	60	A	Sidechain

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	32594	0	16454	3166	0
2	1	131	0	68	14	0
2	2	86	0	46	9	0
3	B	1811	0	1861	95	0
4	C	1612	0	1677	130	0
5	D	1703	0	1763	117	0
6	E	1146	0	1207	93	0
7	F	843	0	857	36	0
8	G	1257	0	1296	81	0
9	H	1116	0	1177	98	0
10	I	1011	0	1043	104	0
11	J	794	0	840	80	0
12	K	853	0	868	54	0
13	L	970	0	1057	75	0
14	M	969	0	1039	78	0
15	N	492	0	529	52	0
16	O	734	0	771	46	0
17	P	700	0	720	52	0
18	Q	857	0	930	53	0
19	R	597	0	668	43	0
20	S	647	0	673	61	0
21	T	762	0	859	48	0
22	V	208	0	221	14	0
23	D	1	0	0	0	0
23	N	1	0	0	0	0
All	All	51895	0	36624	4155	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 47.

The worst 5 of 4155 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:1027:C:C2'	1:A:1028:C:H5''	1.48	1.40
1:A:390:C:H4'	17:P:28:ARG:NH2	1.46	1.28
1:A:1027:C:H2'	1:A:1028:C:C5'	1.65	1.25
1:A:839:U:H5'	1:A:840:C:C5	1.71	1.24
1:A:243:A:H4'	1:A:244:U:C5'	1.65	1.24

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	B	220/227 (97%)	169 (77%)	39 (18%)	12 (6%)	3	25
4	C	204/238 (86%)	149 (73%)	42 (21%)	13 (6%)	2	20
5	D	206/208 (99%)	165 (80%)	31 (15%)	10 (5%)	3	29
6	E	148/161 (92%)	113 (76%)	30 (20%)	5 (3%)	6	42
7	F	99/101 (98%)	83 (84%)	14 (14%)	2 (2%)	11	58
8	G	153/155 (99%)	129 (84%)	23 (15%)	1 (1%)	30	81
9	H	136/138 (99%)	113 (83%)	21 (15%)	2 (2%)	15	65
10	I	125/128 (98%)	94 (75%)	25 (20%)	6 (5%)	4	30
11	J	96/104 (92%)	75 (78%)	14 (15%)	7 (7%)	2	16
12	K	113/128 (88%)	88 (78%)	22 (20%)	3 (3%)	8	49
13	L	122/131 (93%)	96 (79%)	21 (17%)	5 (4%)	4	35
14	M	120/125 (96%)	89 (74%)	26 (22%)	5 (4%)	4	34
15	N	58/60 (97%)	45 (78%)	13 (22%)	0	100	100
16	O	86/88 (98%)	70 (81%)	14 (16%)	2 (2%)	10	54
17	P	81/88 (92%)	64 (79%)	16 (20%)	1 (1%)	19	71
18	Q	102/104 (98%)	86 (84%)	11 (11%)	5 (5%)	3	29
19	R	71/87 (82%)	57 (80%)	13 (18%)	1 (1%)	16	67
20	S	78/92 (85%)	63 (81%)	11 (14%)	4 (5%)	3	28
21	T	97/105 (92%)	72 (74%)	17 (18%)	8 (8%)	1	13
22	V	22/26 (85%)	19 (86%)	1 (4%)	2 (9%)	1	10
All	All	2337/2494 (94%)	1839 (79%)	404 (17%)	94 (4%)	5	36

5 of 94 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	B	12	GLU
3	B	21	ARG
3	B	24	TRP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	B	130	ARG
4	C	4	LYS

### 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	
3	B	191/196 (97%)	170 (89%)	21 (11%)	9	38
4	C	160/187 (86%)	146 (91%)	14 (9%)	14	52
5	D	180/180 (100%)	163 (91%)	17 (9%)	13	47
6	E	115/122 (94%)	103 (90%)	12 (10%)	10	41
7	F	90/90 (100%)	87 (97%)	3 (3%)	50	87
8	G	126/126 (100%)	122 (97%)	4 (3%)	51	88
9	H	119/119 (100%)	110 (92%)	9 (8%)	19	61
10	I	98/99 (99%)	91 (93%)	7 (7%)	21	65
11	J	88/91 (97%)	82 (93%)	6 (7%)	22	67
12	K	87/98 (89%)	80 (92%)	7 (8%)	17	58
13	L	104/108 (96%)	100 (96%)	4 (4%)	44	84
14	M	97/100 (97%)	90 (93%)	7 (7%)	21	64
15	N	49/49 (100%)	44 (90%)	5 (10%)	11	42
16	O	79/79 (100%)	71 (90%)	8 (10%)	11	42
17	P	72/74 (97%)	67 (93%)	5 (7%)	22	67
18	Q	96/96 (100%)	89 (93%)	7 (7%)	20	63
19	R	64/76 (84%)	62 (97%)	2 (3%)	52	88
20	S	71/79 (90%)	65 (92%)	6 (8%)	15	54
21	T	76/81 (94%)	73 (96%)	3 (4%)	43	84
22	V	19/21 (90%)	19 (100%)	0	100	100
All	All	1981/2071 (96%)	1834 (93%)	147 (7%)	20	63

5 of 147 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	F	86	ARG
10	I	60	ASP
18	Q	100	LYS
8	G	16	LEU
9	H	91	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
8	G	106	GLN
10	I	73	GLN
18	Q	16	GLN
9	H	15	ASN
9	H	82	HIS

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1517/1520 (99%)	331 (21%)	187 (12%)
2	1	5/6 (83%)	1 (20%)	1 (20%)
2	2	3/6 (50%)	2 (66%)	0
All	All	1525/1532 (99%)	334 (21%)	188 (12%)

5 of 334 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	5	U
1	A	6	G
1	A	7	G
1	A	8	A
1	A	9	G

5 of 188 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	703	G
1	A	884	U
1	A	1397	C
1	A	721	G
1	A	817	C

## 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 2 ligands modelled in this entry, 2 are monoatomic - 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.

## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	1517/1520 (99%)	-0.09	30 (1%) 62 19	44, 89, 178, 199	0
2	1	6/6 (100%)	1.36	1 (16%) 2 1	199, 199, 199, 199	0
2	2	4/6 (66%)	2.91	3 (75%) 0 0	185, 193, 195, 198	0
3	B	222/227 (97%)	0.09	2 (0%) 81 37	46, 104, 169, 199	0
4	C	206/238 (86%)	-0.03	2 (0%) 79 33	49, 107, 172, 198	0
5	D	208/208 (100%)	0.04	1 (0%) 88 51	32, 90, 156, 199	0
6	E	150/161 (93%)	0.17	2 (1%) 74 27	32, 72, 151, 195	0
7	F	101/101 (100%)	-0.24	0 100 100	63, 116, 167, 182	0
8	G	155/155 (100%)	-0.15	0 100 100	70, 133, 184, 199	0
9	H	138/138 (100%)	0.07	2 (1%) 72 25	31, 72, 145, 181	0
10	I	127/128 (99%)	1.09	24 (18%) 2 1	55, 147, 191, 199	0
11	J	98/104 (94%)	1.02	19 (19%) 2 1	64, 138, 198, 199	0
12	K	115/128 (89%)	-0.05	1 (0%) 81 37	59, 111, 172, 190	0
13	L	124/131 (94%)	0.29	2 (1%) 68 22	46, 104, 165, 199	0
14	M	122/125 (97%)	0.31	7 (5%) 23 5	71, 127, 180, 198	0
15	N	60/60 (100%)	0.56	4 (6%) 17 4	56, 89, 158, 190	0
16	O	88/88 (100%)	0.01	0 100 100	45, 100, 167, 185	0
17	P	83/88 (94%)	0.09	2 (2%) 56 15	38, 91, 146, 185	0
18	Q	104/104 (100%)	0.66	9 (8%) 10 3	49, 90, 172, 199	0
19	R	73/87 (83%)	-0.03	0 100 100	46, 103, 175, 199	0
20	S	80/92 (86%)	0.53	5 (6%) 19 5	74, 136, 187, 199	0
21	T	99/105 (94%)	0.33	8 (8%) 12 3	69, 122, 182, 199	0
22	V	24/26 (92%)	2.53	13 (54%) 0 0	72, 121, 168, 199	0
All	All	3904/4026 (96%)	0.11	137 (3%) 42 10	31, 99, 178, 199	0

The worst 5 of 137 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
20	S	3	ARG	8.7
18	Q	103	GLY	8.2
18	Q	104	LYS	7.1
1	A	1124	G	6.5
18	Q	102	GLY	5.7

## 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 ⓘ

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, 95<sup>th</sup> 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
23	ZN	N	62	1/1	0.10	-	87,87,87,87	0
23	ZN	D	210	1/1	0.34	-	85,85,85,85	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.