wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 09:08 PM GMT

PDB ID	:	1N8R
Title	:	Structure of large ribosomal subunit in complex with virginiamycin M
Authors	:	Hansen, J.L.; Moore, P.B.; Steitz, T.A.
Deposited on	:	2002-11-21
Resolution	:	3.00 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

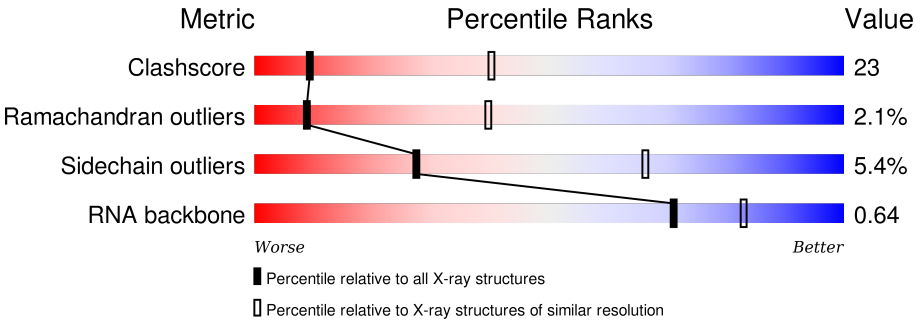
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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))
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	2922	48% 37% 8% • 6%
2	B	122	43% 43% 10% •
3	C	239	50% 43% 6% •
4	D	337	49% 45% 7%
5	E	246	52% 43% •
6	F	176	26% 45% 7% • 20%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	G	177	54% 42%
8	H	119	52% 45%
9	I	348	5% 92%
10	J	167	33% 54% 6% 7%
11	K	145	54% 37% 7%
12	L	132	60% 39%
13	M	164	46% 41% 12%
14	N	194	37% 59% 5%
15	O	186	40% 54% 6%
16	P	115	72% 27%
17	Q	148	58% 36%
18	R	95	66% 31%
19	S	154	60% 34%
20	T	84	68% 26%
21	U	119	57% 40%
22	V	66	32% 44% 5% 20%
23	W	70	50% 41% 7%
24	X	154	43% 51% 6%
25	Y	91	42% 41% 8% 10%
26	Z	240	38% 19% 41%
27	1	73	23% 70% 7%
28	2	56	63% 38%
29	3	48	54% 40%
30	4	92	29% 65% 5%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
35	CL	M	8510	-	-	X	-
35	CL	N	8518	-	-	X	-

2 Entry composition (i)

There are 37 unique types of molecules in this entry. The entry contains 98569 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 23S ribosomal RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	A	2754	59017	26346	10878	19048	2745	0	0	0

- Molecule 2 is a RNA chain called 5S ribosomal RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	122	2600	1160	472	847	121	0	0	0

- Molecule 3 is a protein called 50S ribosomal protein L2P.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	237	1754	1072	352	325	5	0	0	0

- Molecule 4 is a protein called 50S ribosomal protein L3P.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	337	2624	1616	493	510	5	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	?	-	PRO	DELETION	UNP P20279
D	310	ARG	PHE	CONFLICT	UNP P20279

- Molecule 5 is a protein called 50S ribosomal protein L4E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	246	1858	1131	344	382	1	0	0	0

- Molecule 6 is a protein called 50S ribosomal protein L5P.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	140	Total	C	N	O	S	0	0	0
			1094	685	195	210	4			

- Molecule 7 is a protein called 50S ribosomal protein L6P.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	172	Total	C	N	O	S	0	0	0
			1357	840	224	289	4			

- Molecule 8 is a protein called 50S ribosomal protein L7Ae.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	119	Total	C	N	O	S	0	0	0
			885	552	141	191	1			

- Molecule 9 is a protein called Acidic ribosomal protein P0 homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	29	Total	C	N	O	S	0	0	0
			240	149	39	51	1			

- Molecule 10 is a protein called 50S ribosomal protein L10e.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	156	Total	C	N	O	S	0	0	0
			1215	766	233	212	4			

- Molecule 11 is a protein called 50S ribosomal protein L13P.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	142	Total	C	N	O	S	0	0	0
			1119	696	199	221	3			

- Molecule 12 is a protein called 50S ribosomal protein L14P.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	132	Total	C	N	O	S	0	0	0
			993	609	189	191	4			

- Molecule 13 is a protein called 50S ribosomal protein L15P.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
13	M	145	Total	C	N	O	0	0	0
			1114	668	222	224			

- Molecule 14 is a protein called 50S ribosomal protein L15E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	N	194	Total	C	N	O	S	0	0	0
			1605	988	346	266	5			

- Molecule 15 is a protein called 50S ribosomal protein L18P.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	O	186	Total	C	N	O	S	0	0	0
			1444	895	262	285	2			

- Molecule 16 is a protein called 50S ribosomal protein L18E.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
16	P	115	Total	C	N	O	0	0	0
			864	529	161	174			

- Molecule 17 is a protein called 50S ribosomal protein L19E.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
17	Q	143	Total	C	N	O	0	0	0
			1133	680	230	223			

- Molecule 18 is a protein called 50S ribosomal protein L21e.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
18	R	95	Total	C	N	O	0	0	0
			734	450	141	143			

- Molecule 19 is a protein called 50S ribosomal protein L22P.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
19	S	150	Total	C	N	O	S	0	0	0
			1149	713	209	223	4			

- Molecule 20 is a protein called 50S ribosomal protein L23P.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
20	T	81	Total	C	N	O	S	0	0	0
			641	389	111	138	3			

- Molecule 21 is a protein called 50S ribosomal protein L24P.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
21	U	119	Total	C	N	O	0	0	0
			949	568	180	201			

- Molecule 22 is a protein called 50S ribosomal protein L24E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
22	V	53	Total	C	N	O	S	0	0	0
			410	244	75	86	5			

- Molecule 23 is a protein called 50S ribosomal protein L29P.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
23	W	65	Total	C	N	O	S	0	0	0
			499	304	94	100	1			

- Molecule 24 is a protein called 50S ribosomal protein L30P.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
24	X	154	Total	C	N	O	S	0	0	0
			1195	737	209	243	6			

- Molecule 25 is a protein called 50S ribosomal protein L31E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
25	Y	82	Total	C	N	O	S	0	0	0
			654	402	129	122	1			

- Molecule 26 is a protein called 50S ribosomal protein L32E.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
26	Z	142	Total	C	N	O	0	0	0
			1130	686	228	216			

- Molecule 27 is a protein called 50S ribosomal protein L37AE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
27	1	73	Total	C	N	O	S	0	0	0
			563	359	111	86	7			

- Molecule 28 is a protein called 50S ribosomal protein L37e.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
28	2	56	Total	C	N	O	S	0	0	0
			430	258	86	82	4			

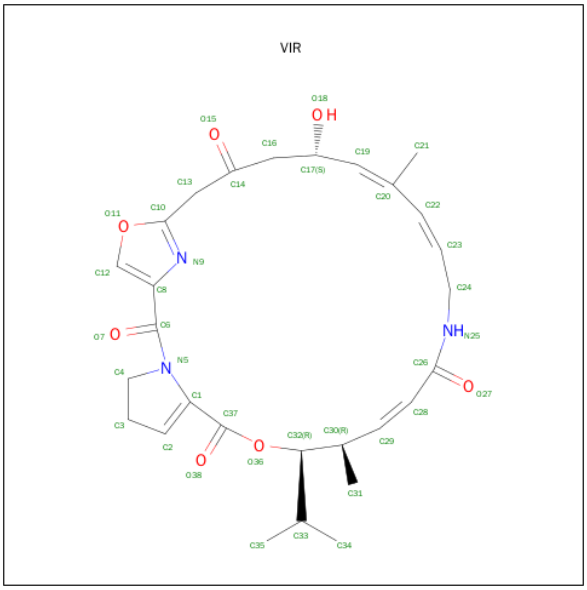
- Molecule 29 is a protein called 50S ribosomal protein L39e.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
29	3	46	Total	C	N	O	S	0	0	0
			393	238	86	68	1			

- Molecule 30 is a protein called 50S ribosomal protein L44E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
30	4	92	Total	C	N	O	S	0	0	0
			755	458	153	137	7			

- Molecule 31 is VIRGINIAMYCIN M1 (three-letter code: VIR) (formula: $C_{28}H_{35}N_3O_7$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
31	A	1	Total	C	N	O	0	0
			38	28	3	7		

- Molecule 32 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
32	D	1	Total Mg 1 1	0	0
32	B	1	Total Mg 1 1	0	0
32	C	2	Total Mg 2 2	0	0
32	Z	1	Total Mg 1 1	0	0
32	A	109	Total Mg 109 109	0	0
32	4	1	Total Mg 1 1	0	0
32	U	1	Total Mg 1 1	0	0
32	L	1	Total Mg 1 1	0	0

- Molecule 33 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
33	A	1	Total K 1 1	0	0

- Molecule 34 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
34	J	2	Total Na 2 2	0	0
34	K	1	Total Na 1 1	0	0
34	E	1	Total Na 1 1	0	0
34	B	2	Total Na 2 2	0	0
34	C	1	Total Na 1 1	0	0
34	A	71	Total Na 71 71	0	0
34	T	1	Total Na 1 1	0	0
34	N	1	Total Na 1 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
34	U	1	Total Na 1 1	0	0
34	4	1	Total Na 1 1	0	0
34	R	1	Total Na 1 1	0	0
34	S	2	Total Na 2 2	0	0
34	M	1	Total Na 1 1	0	0

- Molecule 35 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
35	P	1	Total Cl 1 1	0	0
35	D	1	Total Cl 1 1	0	0
35	K	3	Total Cl 3 3	0	0
35	C	1	Total Cl 1 1	0	0
35	Z	2	Total Cl 2 2	0	0
35	A	8	Total Cl 8 8	0	0
35	4	1	Total Cl 1 1	0	0
35	N	1	Total Cl 1 1	0	0
35	O	1	Total Cl 1 1	0	0
35	L	1	Total Cl 1 1	0	0
35	S	1	Total Cl 1 1	0	0
35	M	1	Total Cl 1 1	0	0

- Molecule 36 is CADMIUM ION (three-letter code: CD) (formula: Cd).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
36	P	1	Total Cd 1 1	0	0
36	2	1	Total Cd 1 1	0	0
36	1	1	Total Cd 1 1	0	0
36	4	1	Total Cd 1 1	0	0
36	V	1	Total Cd 1 1	0	0

- Molecule 37 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
37	A	5881	Total O 5881 5881	0	0
37	B	146	Total O 146 146	0	0
37	C	135	Total O 135 135	0	0
37	D	141	Total O 141 141	0	0
37	E	178	Total O 178 178	0	0
37	F	49	Total O 49 49	0	0
37	G	43	Total O 43 43	0	0
37	H	30	Total O 30 30	0	0
37	I	21	Total O 21 21	0	0
37	J	76	Total O 76 76	0	0
37	K	55	Total O 55 55	0	0
37	L	64	Total O 64 64	0	0
37	M	85	Total O 85 85	0	0
37	N	141	Total O 141 141	0	0
37	O	67	Total O 67 67	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
37	P	45	Total O 45 45	0	0
37	Q	72	Total O 72 72	0	0
37	R	57	Total O 57 57	0	0
37	S	87	Total O 87 87	0	0
37	T	34	Total O 34 34	0	0
37	U	33	Total O 33 33	0	0
37	V	27	Total O 27 27	0	0
37	W	16	Total O 16 16	0	0
37	X	68	Total O 68 68	0	0
37	Y	27	Total O 27 27	0	0
37	Z	100	Total O 100 100	0	0
37	1	35	Total O 35 35	0	0
37	2	57	Total O 57 57	0	0
37	3	40	Total O 40 40	0	0
37	4	72	Total O 72 72	0	0

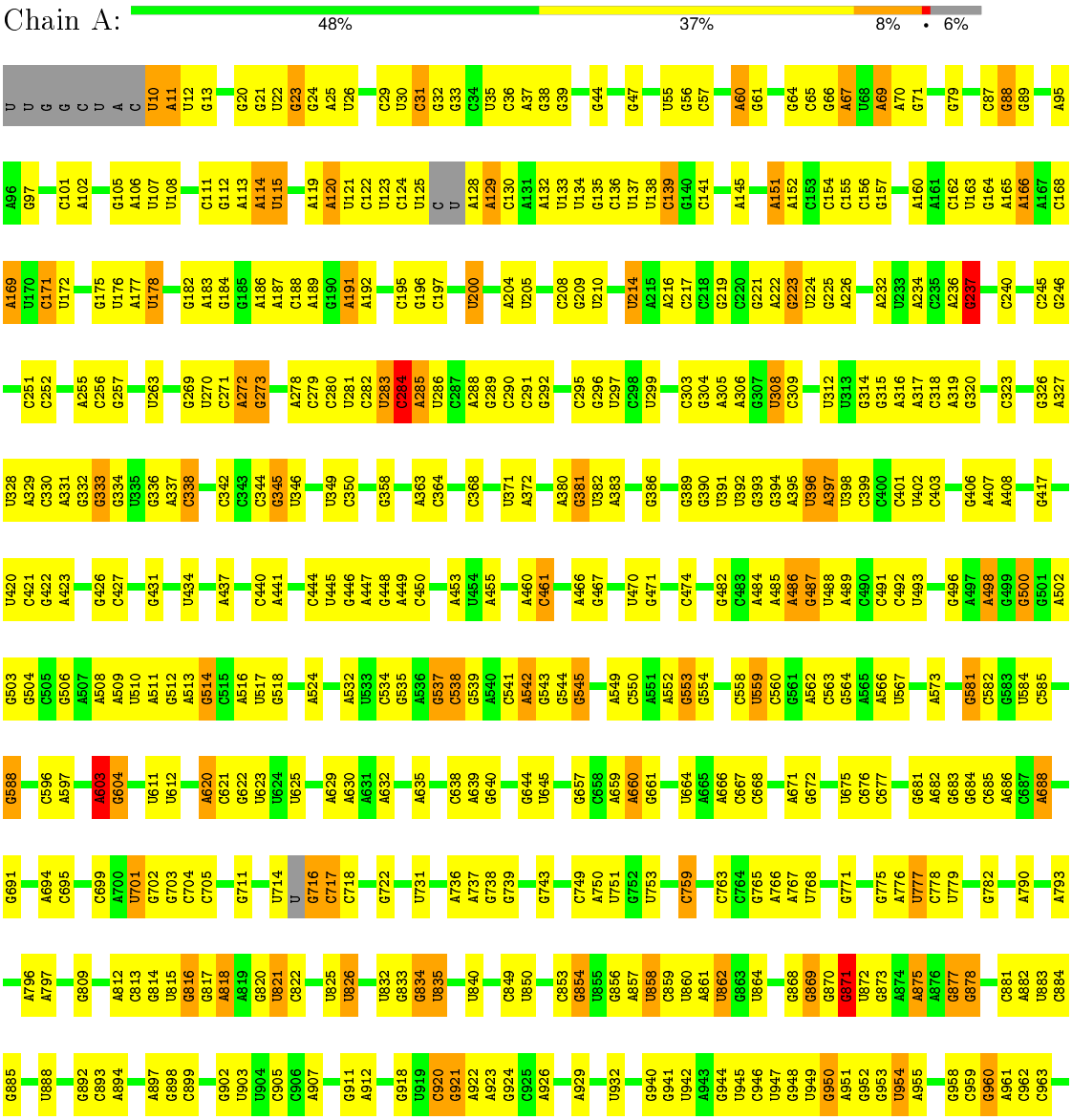
3 Residue-property plots (i)

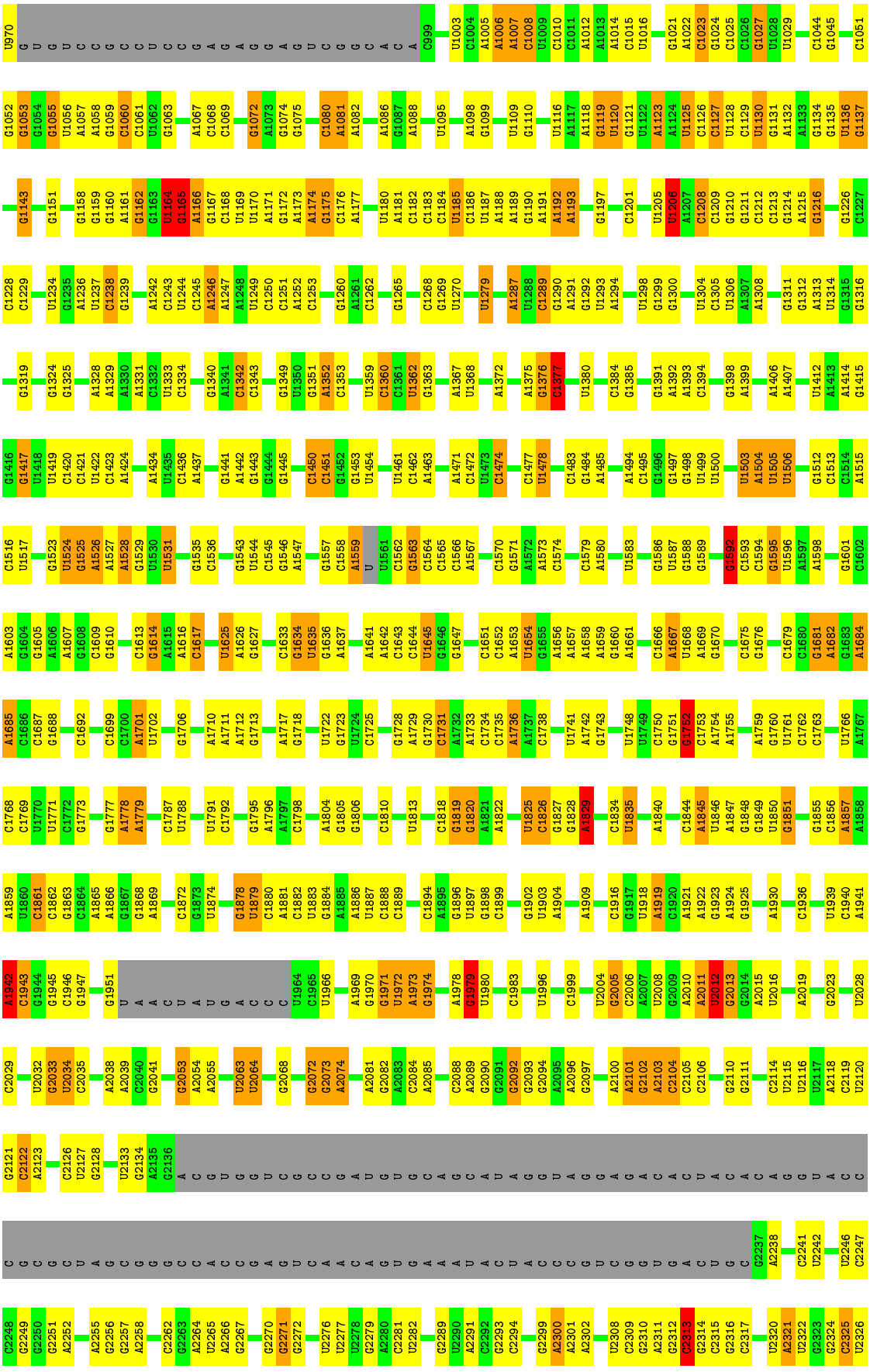
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according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

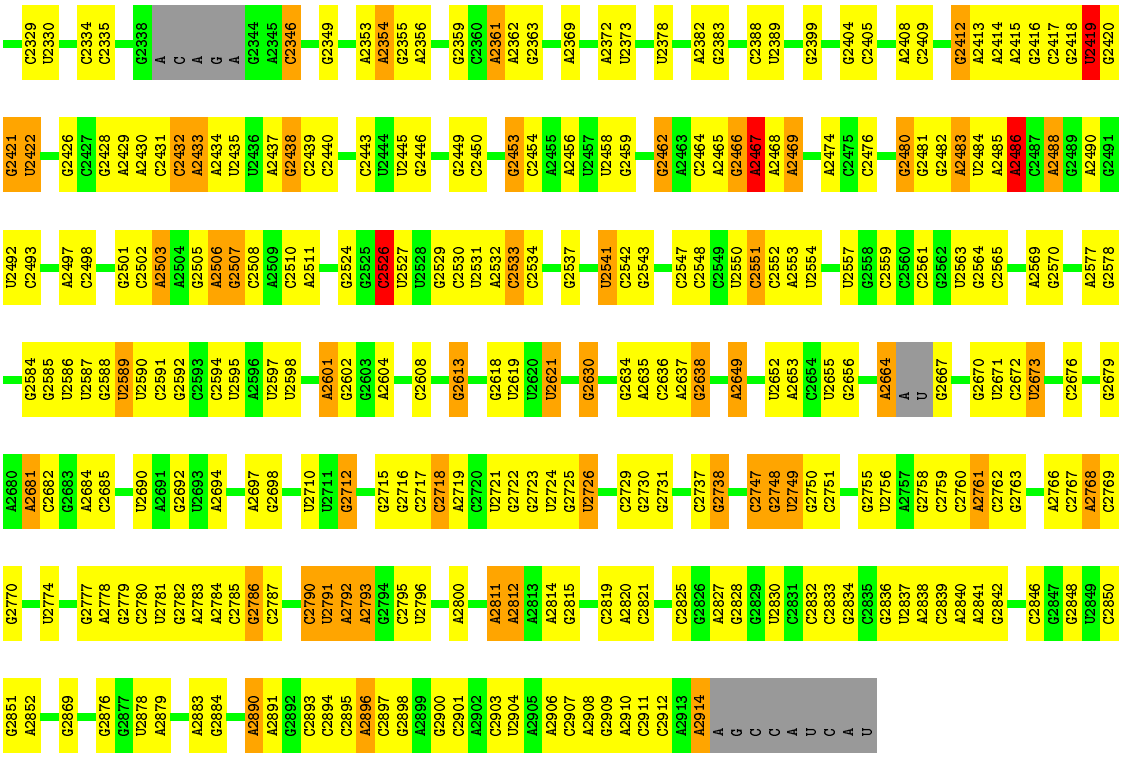
Note EDS was not executed.

- Molecule 1: 23S ribosomal RNA

Chain A:

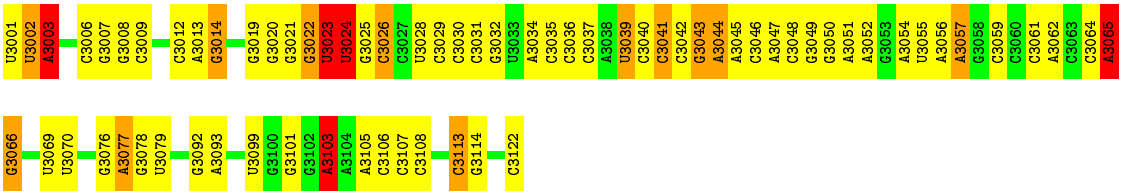






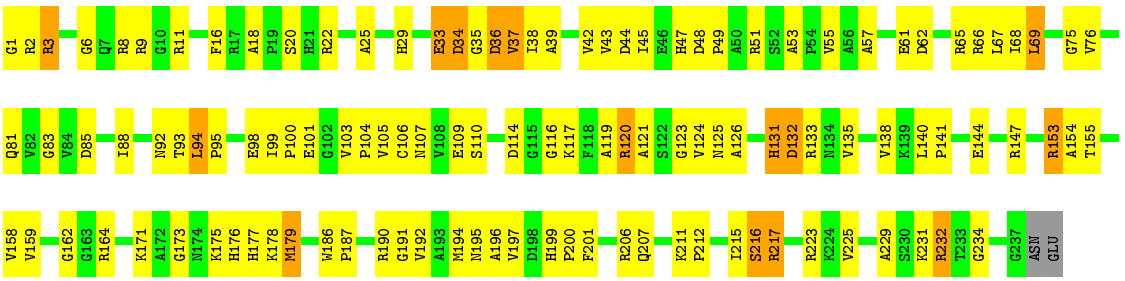
• Molecule 2: 5S ribosomal RNA

Chain B:



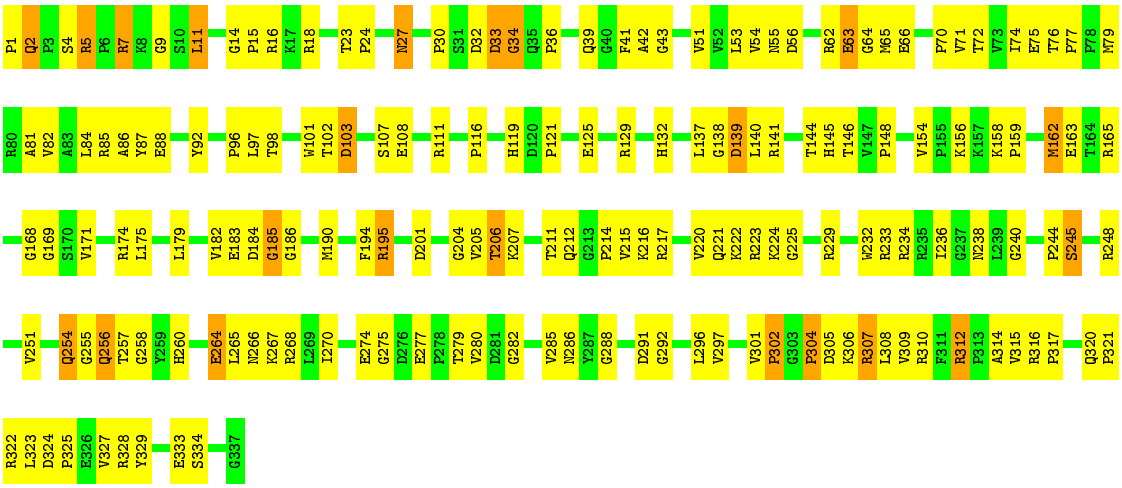
• Molecule 3: 50S ribosomal protein L2P

Chain C:



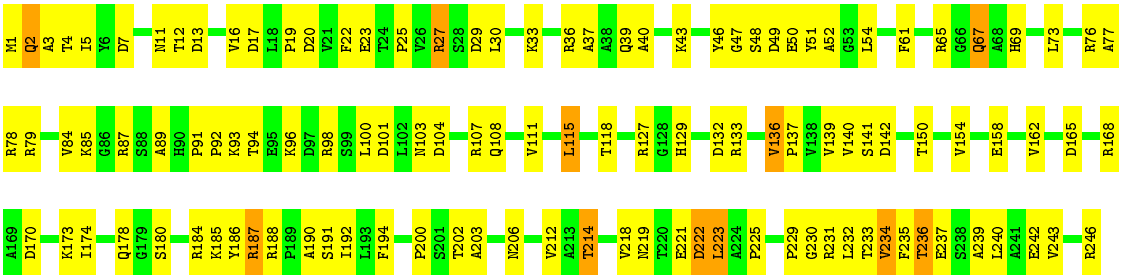
• Molecule 4: 50S ribosomal protein L3P

Chain D:



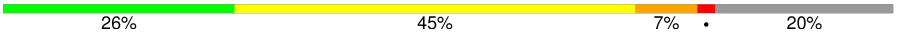
- Molecule 5: 50S ribosomal protein L4E

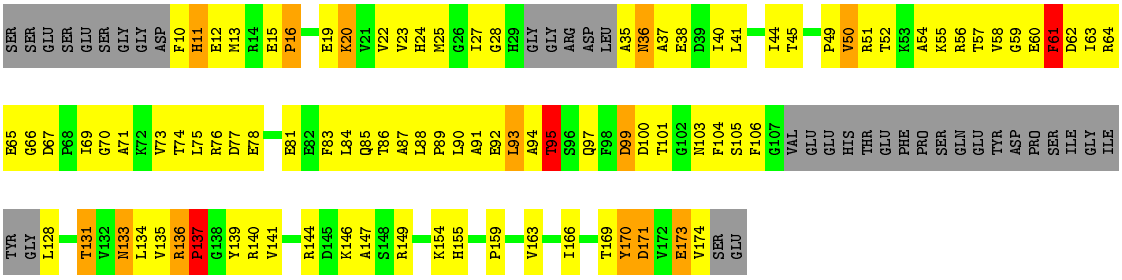
Chain E:



- Molecule 6: 50S ribosomal protein L5P

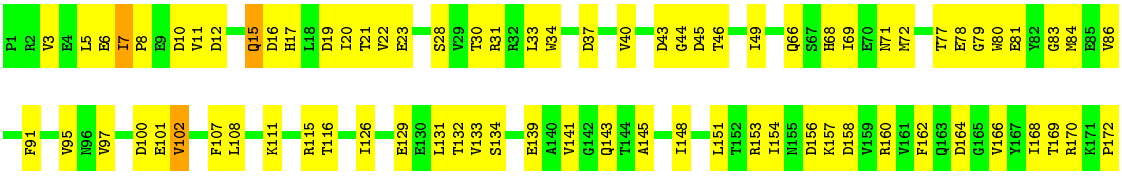
Chain F:





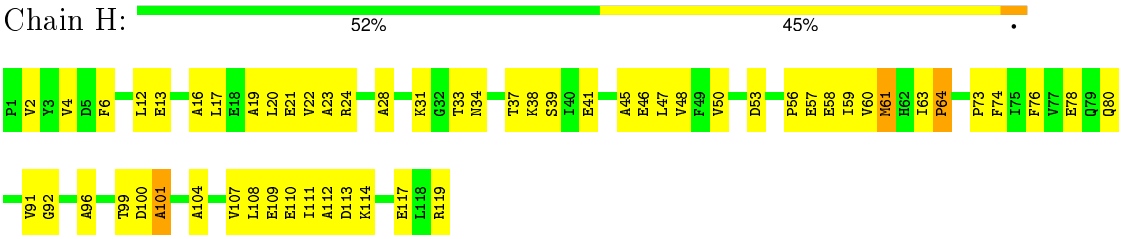
- Molecule 7: 50S ribosomal protein L6P

Chain G:



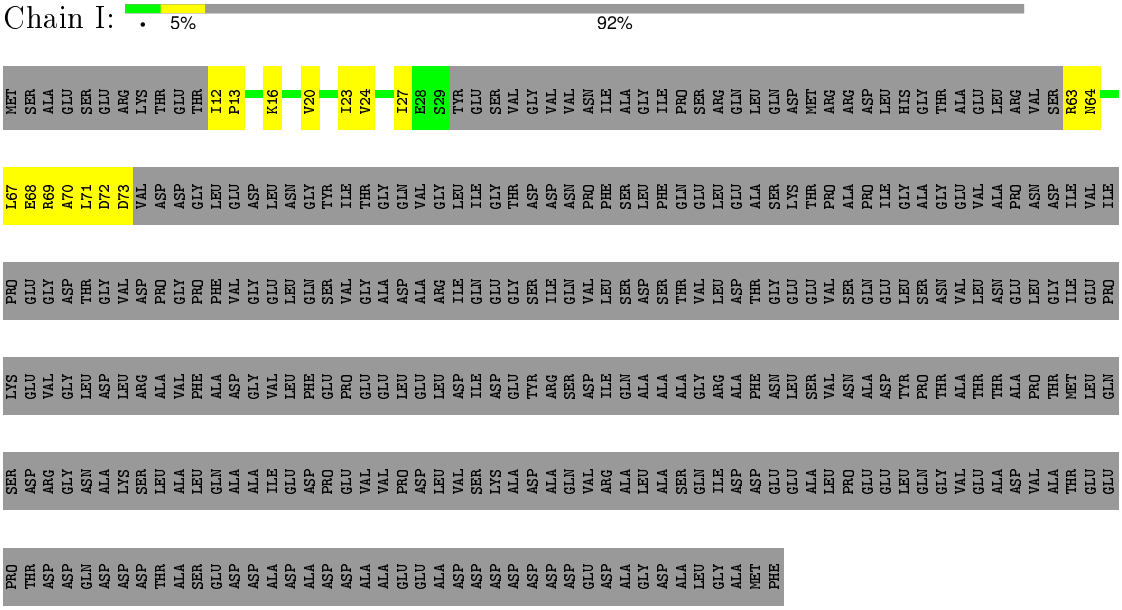
• Molecule 8: 50S ribosomal protein L7Ae

Chain H:



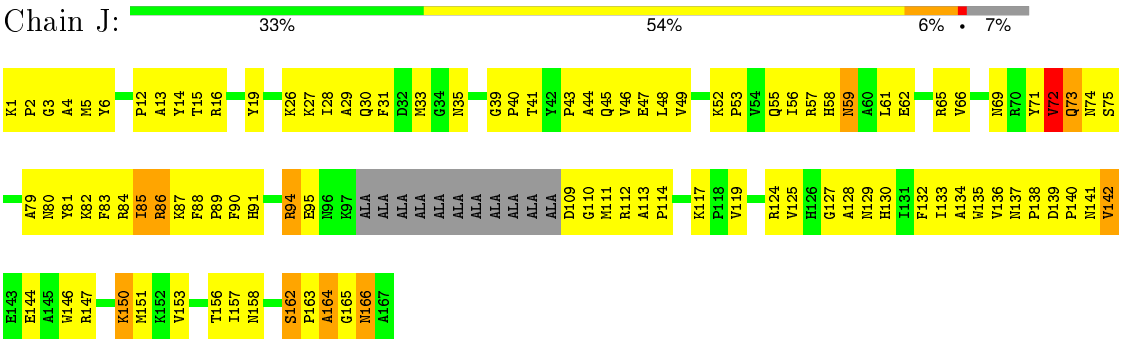
• Molecule 9: Acidic ribosomal protein P0 homolog

Chain I:



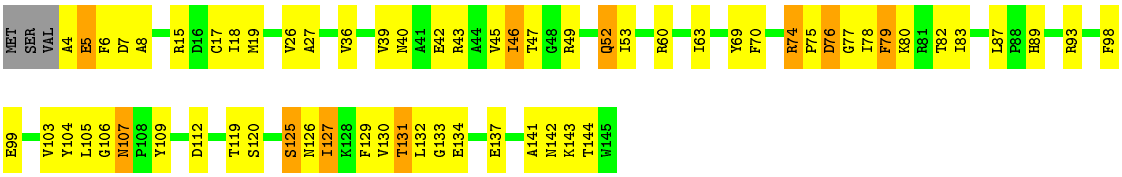
• Molecule 10: 50S ribosomal protein L10e

Chain J:



• Molecule 11: 50S ribosomal protein L13P

Chain K:

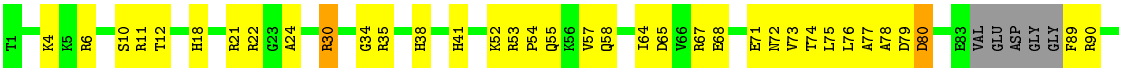


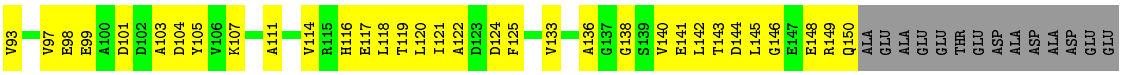
• Molecule 12: 50S ribosomal protein L14P

Chain L:

• Molecule 13: 50S ribosomal protein L15P

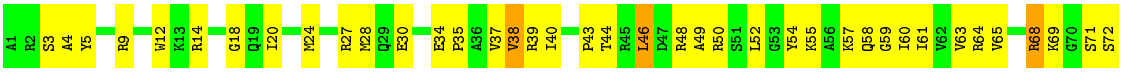
Chain M:

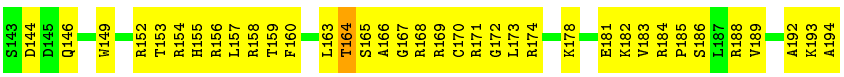




• Molecule 14: 50S ribosomal protein L15E

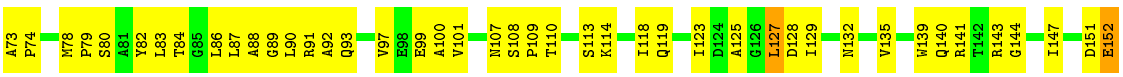
Chain N:

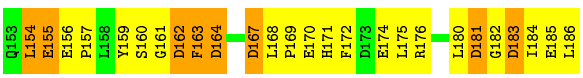




• Molecule 15: 50S ribosomal protein L18P

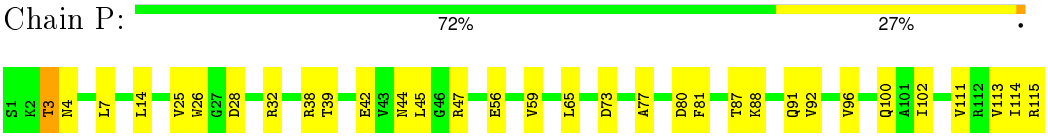
Chain O:





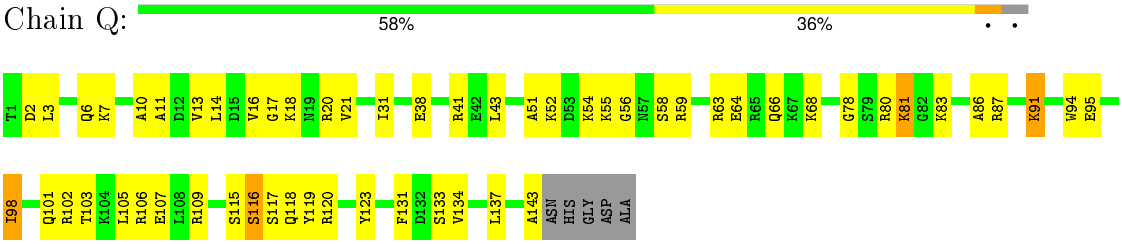
- Molecule 16: 50S ribosomal protein L18E

Chain P:



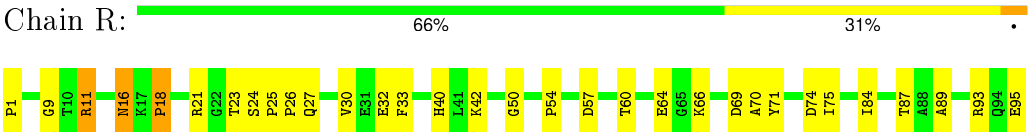
- Molecule 17: 50S ribosomal protein L19E

Chain Q:



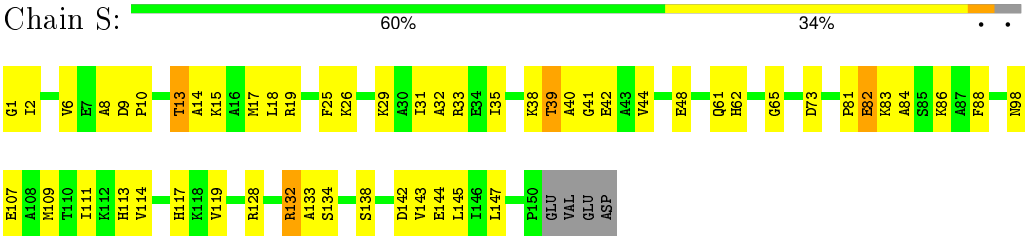
- Molecule 18: 50S ribosomal protein L21e

Chain R:



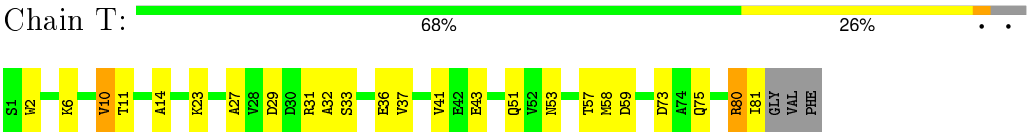
- Molecule 19: 50S ribosomal protein L22P

Chain S:



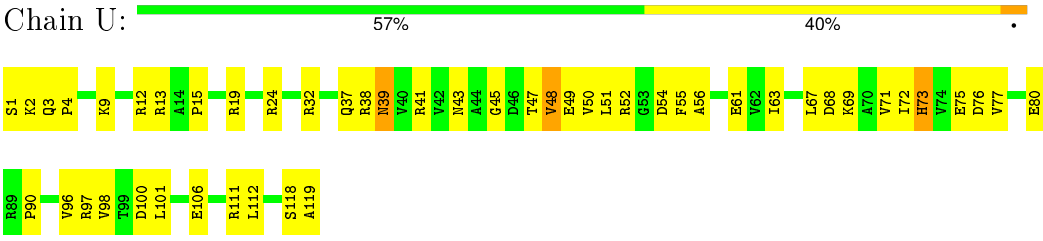
- Molecule 20: 50S ribosomal protein L23P

Chain T:



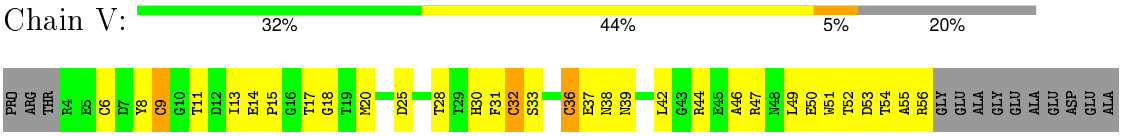
- Molecule 21: 50S ribosomal protein L24P

Chain U:



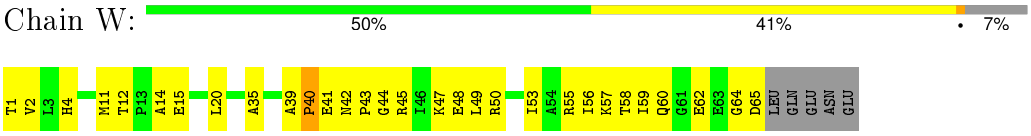
- Molecule 22: 50S ribosomal protein L24E

Chain V:



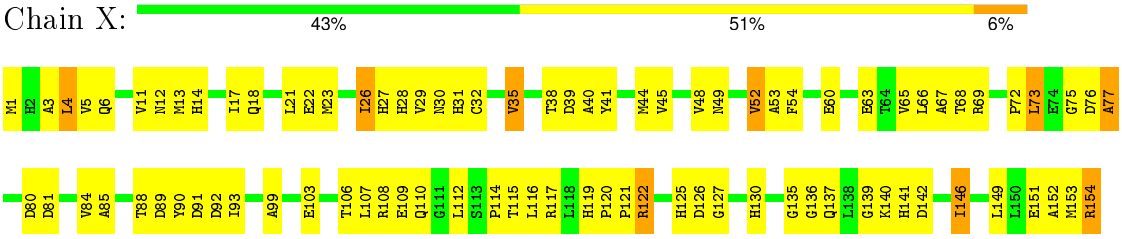
• Molecule 23: 50S ribosomal protein L29P

Chain W:



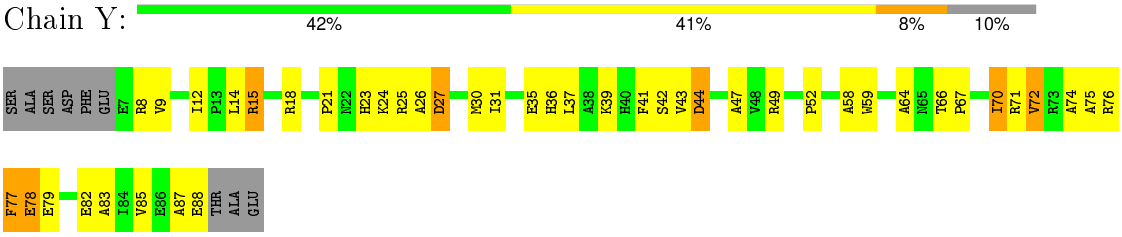
• Molecule 24: 50S ribosomal protein L30P

Chain X:



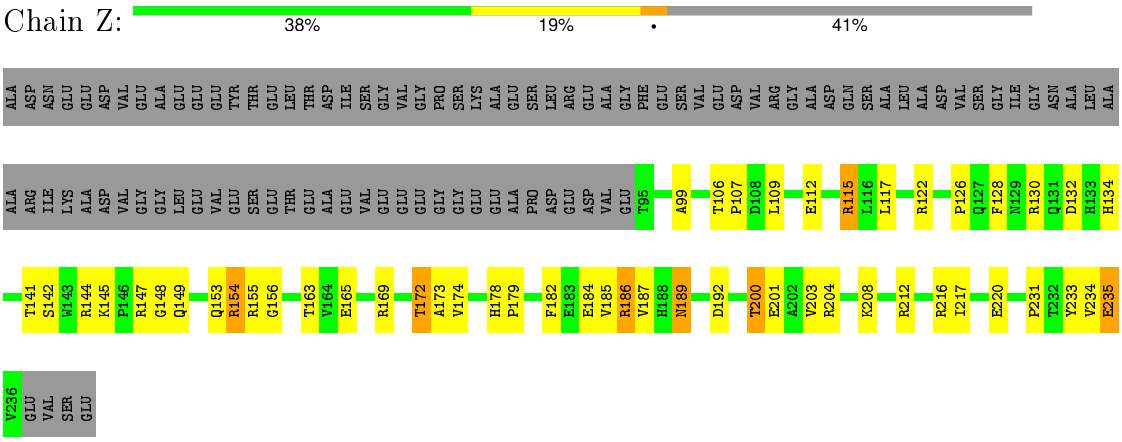
• Molecule 25: 50S ribosomal protein L31E

Chain Y:



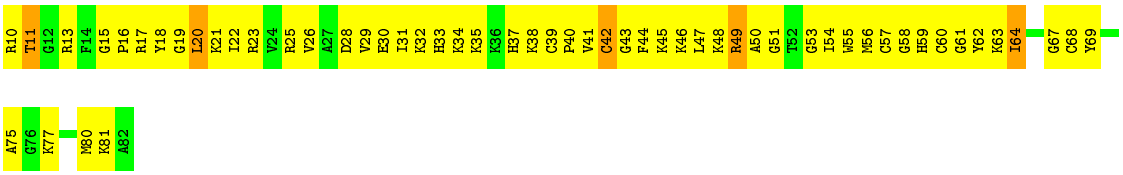
• Molecule 26: 50S ribosomal protein L32E

Chain Z:



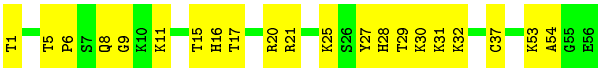
• Molecule 27: 50S ribosomal protein L37AE

Chain 1:



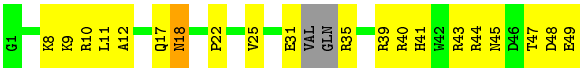
- Molecule 28: 50S ribosomal protein L37e

Chain 2:



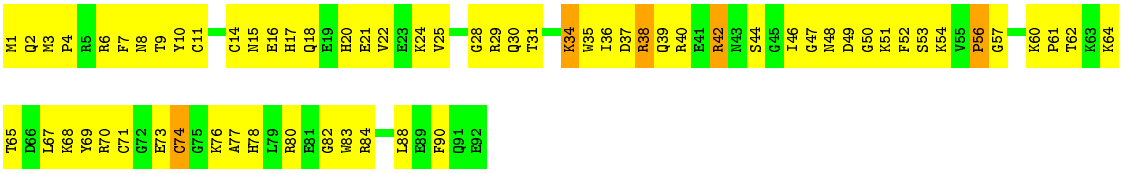
- Molecule 29: 50S ribosomal protein L39e

Chain 3:



- Molecule 30: 50S ribosomal protein L44E

Chain 4:



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	212.90Å 300.47Å 575.18Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.00	Depositor
% Data completeness (in resolution range)	94.6 (20.00-3.00)	Depositor
R_{merge}	0.14	Depositor
R_{sym}	0.14	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.201 , 0.242	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	98569	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, CL, NA, K, CD, VIR

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	0.57	9/66076 (0.0%)	0.76	37/103052 (0.0%)
2	B	0.81	10/2905 (0.3%)	0.91	17/4528 (0.4%)
3	C	0.44	0/1787	0.75	0/2409
4	D	0.41	0/2689	0.71	0/3652
5	E	0.48	0/1883	0.74	0/2551
6	F	0.40	0/1111	0.64	0/1498
7	G	0.44	0/1382	0.66	0/1880
8	H	0.40	0/896	0.63	0/1219
9	I	0.33	0/241	0.53	0/324
10	J	0.46	0/1246	0.83	2/1686 (0.1%)
11	K	0.46	0/1135	0.70	0/1530
12	L	0.43	0/1003	0.75	0/1351
13	M	0.42	0/1126	0.75	0/1504
14	N	0.63	0/1633	0.86	1/2180 (0.0%)
15	O	0.42	0/1473	0.71	0/1999
16	P	0.45	0/873	0.71	0/1181
17	Q	0.44	0/1143	0.64	0/1521
18	R	0.45	0/748	0.79	0/1005
19	S	0.45	0/1172	0.76	0/1578
20	T	0.41	0/648	0.65	0/875
21	U	0.38	0/957	0.70	0/1289
22	V	0.82	0/417	0.81	1/562 (0.2%)
23	W	0.40	0/502	0.61	0/675
24	X	0.50	0/1218	0.73	0/1655
25	Y	0.44	0/664	0.70	0/895
26	Z	0.46	0/1146	0.72	0/1536
27	1	0.77	0/575	0.82	0/763
28	2	0.54	0/437	0.78	0/578
29	3	0.42	0/398	0.63	0/527
30	4	0.98	0/771	0.81	0/1024
All	All	0.56	19/98255 (0.0%)	0.76	58/147027 (0.0%)

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	157
2	B	0	4
All	All	1	161

The worst 5 of 19 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	2488	A	O5'-C5'	-8.71	1.28	1.42
2	B	3025	G	C2'-O2'	-7.62	1.31	1.41
2	B	3003	A	O5'-C5'	7.52	1.56	1.44
2	B	3024	U	O5'-C5'	7.15	1.55	1.44
2	B	3023	U	C4'-O4'	7.08	1.54	1.45

The worst 5 of 58 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1164	U	OP2-P-O3'	-18.66	64.14	105.20
1	A	1164	U	OP1-P-O3'	-18.19	65.19	105.20
1	A	1979	G	C2'-C3'-O3'	9.72	130.90	109.50
1	A	1563	G	C2'-C3'-O3'	9.58	130.58	109.50
1	A	1165	G	O5'-P-OP1	-8.38	98.16	105.70

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	1563	G	C3'

5 of 161 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	22	U	Sidechain
1	A	23	G	Sidechain
1	A	26	U	Sidechain
1	A	32	G	Sidechain
1	A	33	G	Sidechain

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	59017	0	29801	1306	0
2	B	2600	0	1326	85	0
3	C	1754	0	1763	144	0
4	D	2624	0	2533	195	0
5	E	1858	0	1816	146	0
6	F	1094	0	1085	135	0
7	G	1357	0	1266	79	0
8	H	885	0	854	73	0
9	I	240	0	231	23	0
10	J	1215	0	1215	168	0
11	K	1119	0	1098	70	0
12	L	993	0	1027	60	0
13	M	1114	0	1072	79	0
14	N	1605	0	1676	212	0
15	O	1444	0	1401	152	0
16	P	864	0	873	29	0
17	Q	1133	0	1127	68	0
18	R	734	0	729	27	0
19	S	1149	0	1122	68	0
20	T	641	0	605	26	0
21	U	949	0	923	52	0
22	V	410	0	368	48	0
23	W	499	0	511	30	0
24	X	1195	0	1137	115	0
25	Y	654	0	653	48	0
26	Z	1130	0	1133	62	0
27	1	563	0	601	83	0
28	2	430	0	426	30	0
29	3	393	0	406	22	0
30	4	755	0	732	89	0
31	A	38	0	34	3	0
32	4	1	0	0	0	0
32	A	109	0	0	0	0
32	B	1	0	0	0	0
32	C	2	0	0	0	0
32	D	1	0	0	0	0
32	L	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
32	U	1	0	0	0	0
32	Z	1	0	0	0	0
33	A	1	0	0	0	0
34	4	1	0	0	0	0
34	A	71	0	0	0	0
34	B	2	0	0	0	0
34	C	1	0	0	0	0
34	E	1	0	0	0	0
34	J	2	0	0	0	0
34	K	1	0	0	0	0
34	M	1	0	0	0	0
34	N	1	0	0	0	0
34	R	1	0	0	0	0
34	S	2	0	0	0	0
34	T	1	0	0	0	0
34	U	1	0	0	0	0
35	4	1	0	0	1	0
35	A	8	0	0	0	0
35	C	1	0	0	0	0
35	D	1	0	0	0	0
35	K	3	0	0	1	0
35	L	1	0	0	0	0
35	M	1	0	0	2	0
35	N	1	0	0	2	0
35	O	1	0	0	1	0
35	P	1	0	0	0	0
35	S	1	0	0	0	0
35	Z	2	0	0	0	0
36	1	1	0	0	0	0
36	2	1	0	0	0	0
36	4	1	0	0	0	0
36	P	1	0	0	0	0
36	V	1	0	0	0	0
37	1	35	0	0	15	0
37	2	57	0	0	1	0
37	3	40	0	0	6	0
37	4	72	0	0	11	0
37	A	5881	0	0	302	0
37	B	146	0	0	21	0
37	C	135	0	0	15	0
37	D	141	0	0	33	0
37	E	178	0	0	46	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
37	F	49	0	0	17	0
37	G	43	0	0	10	0
37	H	30	0	0	11	0
37	I	21	0	0	7	0
37	J	76	0	0	23	0
37	K	55	0	0	6	0
37	L	64	0	0	16	0
37	M	85	0	0	21	0
37	N	141	0	0	35	0
37	O	67	0	0	20	0
37	P	45	0	0	10	0
37	Q	72	0	0	10	0
37	R	57	0	0	3	0
37	S	87	0	0	10	0
37	T	34	0	0	5	0
37	U	33	0	0	6	0
37	V	27	0	0	6	0
37	W	16	0	0	2	0
37	X	68	0	0	11	0
37	Y	27	0	0	5	0
37	Z	100	0	0	14	0
All	All	98569	0	59544	3401	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 23.

The worst 5 of 3401 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)
10:J:86:ARG:NH1	10:J:133:ILE:HG13	1.52	1.20
10:J:165:GLY:HA3	37:J:8397:HOH:O	1.39	1.18
27:1:39:CYS:SG	27:1:47:LEU:HD21	1.84	1.17
14:N:87:MET:HG2	30:4:46:ILE:HG21	1.25	1.15
21:U:71:VAL:HG11	21:U:90:PRO:HB3	1.31	1.11

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	C	235/239 (98%)	200 (85%)	29 (12%)	6 (3%)	7 33
4	D	335/337 (99%)	302 (90%)	22 (7%)	11 (3%)	5 26
5	E	244/246 (99%)	222 (91%)	22 (9%)	0	100 100
6	F	134/176 (76%)	95 (71%)	27 (20%)	12 (9%)	1 4
7	G	170/177 (96%)	157 (92%)	12 (7%)	1 (1%)	30 72
8	H	117/119 (98%)	104 (89%)	10 (8%)	3 (3%)	7 33
9	I	25/348 (7%)	23 (92%)	1 (4%)	1 (4%)	4 21
10	J	152/167 (91%)	130 (86%)	17 (11%)	5 (3%)	5 26
11	K	140/145 (97%)	127 (91%)	8 (6%)	5 (4%)	4 24
12	L	130/132 (98%)	120 (92%)	8 (6%)	2 (2%)	13 50
13	M	141/164 (86%)	120 (85%)	20 (14%)	1 (1%)	26 70
14	N	192/194 (99%)	165 (86%)	24 (12%)	3 (2%)	12 48
15	O	184/186 (99%)	164 (89%)	12 (6%)	8 (4%)	3 19
16	P	113/115 (98%)	109 (96%)	4 (4%)	0	100 100
17	Q	141/148 (95%)	135 (96%)	5 (4%)	1 (1%)	26 70
18	R	93/95 (98%)	86 (92%)	3 (3%)	4 (4%)	3 19
19	S	148/154 (96%)	134 (90%)	13 (9%)	1 (1%)	26 70
20	T	79/84 (94%)	73 (92%)	6 (8%)	0	100 100
21	U	117/119 (98%)	110 (94%)	7 (6%)	0	100 100
22	V	51/66 (77%)	48 (94%)	3 (6%)	0	100 100
23	W	63/70 (90%)	58 (92%)	3 (5%)	2 (3%)	5 27
24	X	152/154 (99%)	142 (93%)	8 (5%)	2 (1%)	15 53
25	Y	80/91 (88%)	71 (89%)	5 (6%)	4 (5%)	3 15
26	Z	140/240 (58%)	133 (95%)	7 (5%)	0	100 100
27	1	71/73 (97%)	61 (86%)	7 (10%)	3 (4%)	3 20

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
28	2	54/56 (96%)	52 (96%)	2 (4%)	0	100 100
29	3	42/48 (88%)	41 (98%)	1 (2%)	0	100 100
30	4	90/92 (98%)	83 (92%)	5 (6%)	2 (2%)	8 38
All	All	3633/4235 (86%)	3265 (90%)	291 (8%)	77 (2%)	9 40

5 of 77 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	139	ASP
6	F	93	LEU
6	F	95	THR
6	F	173	GLU
8	H	101	ALA

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	C	179/181 (99%)	166 (93%)	13 (7%)	17 52
4	D	282/282 (100%)	264 (94%)	18 (6%)	22 59
5	E	193/193 (100%)	178 (92%)	15 (8%)	16 49
6	F	117/147 (80%)	106 (91%)	11 (9%)	11 39
7	G	152/155 (98%)	148 (97%)	4 (3%)	54 85
8	H	92/92 (100%)	91 (99%)	1 (1%)	80 94
9	I	27/283 (10%)	27 (100%)	0	100 100
10	J	122/122 (100%)	111 (91%)	11 (9%)	12 41
11	K	118/121 (98%)	107 (91%)	11 (9%)	11 39
12	L	106/106 (100%)	102 (96%)	4 (4%)	40 78
13	M	112/126 (89%)	108 (96%)	4 (4%)	42 79
14	N	166/166 (100%)	158 (95%)	8 (5%)	31 71
15	O	149/149 (100%)	144 (97%)	5 (3%)	44 81

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
16	P	93/93 (100%)	91 (98%)	2 (2%)	60 88
17	Q	113/116 (97%)	109 (96%)	4 (4%)	43 80
18	R	79/79 (100%)	75 (95%)	4 (5%)	29 69
19	S	117/121 (97%)	113 (97%)	4 (3%)	44 81
20	T	71/73 (97%)	69 (97%)	2 (3%)	51 84
21	U	105/105 (100%)	102 (97%)	3 (3%)	50 84
22	V	44/52 (85%)	42 (96%)	2 (4%)	34 74
23	W	51/56 (91%)	51 (100%)	0	100 100
24	X	130/130 (100%)	121 (93%)	9 (7%)	19 56
25	Y	66/73 (90%)	61 (92%)	5 (8%)	16 51
26	Z	120/195 (62%)	110 (92%)	10 (8%)	14 46
27	1	56/56 (100%)	49 (88%)	7 (12%)	6 24
28	2	46/46 (100%)	46 (100%)	0	100 100
29	3	42/44 (96%)	41 (98%)	1 (2%)	57 87
30	4	79/79 (100%)	73 (92%)	6 (8%)	16 51
All	All	3027/3441 (88%)	2863 (95%)	164 (5%)	27 66

5 of 164 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
11	K	52	GLN
13	M	117	GLU
27	1	32	LYS
11	K	76	ASP
11	K	131	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 95 such sidechains are listed below:

Mol	Chain	Res	Type
13	M	116	HIS
18	R	40	HIS
29	3	18	ASN
14	N	58	GLN
15	O	153	GLN

5.3.3 RNA

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	2747/2922 (94%)	248 (9%)	38 (1%)
2	B	121/122 (99%)	14 (11%)	4 (3%)
All	All	2868/3044 (94%)	262 (9%)	42 (1%)

5 of 262 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	11	A
1	A	31	C
1	A	60	A
1	A	67	A
1	A	69	A

5 of 42 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	1352	A
1	A	1667	A
2	B	3023	U
1	A	1377	C
1	A	1474	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 232 ligands modelled in this entry, 231 are monoatomic - leaving 1 for Mogul analysis.

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	$\#\|Z\| > 2$
31	VIR	A	9403	-	34,40,40	2.38	15 (44%)	37,55,55	2.17	10 (27%)

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
31	VIR	A	9403	-	-	0/42/58/58	0/1/3/3

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
31	A	9403	VIR	C28-C29	-6.50	1.15	1.32
31	A	9403	VIR	C17-C19	-3.06	1.46	1.50
31	A	9403	VIR	C13-C14	-3.00	1.46	1.51
31	A	9403	VIR	C1-C37	-2.62	1.38	1.47
31	A	9403	VIR	C16-C17	-2.49	1.50	1.54

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
31	A	9403	VIR	C28-C26-N25	-4.96	103.44	114.87
31	A	9403	VIR	C8-C6-N5	-3.21	110.64	118.55
31	A	9403	VIR	C24-N25-C26	-2.29	119.17	121.76
31	A	9403	VIR	C12-C8-C6	-2.20	122.14	129.47
31	A	9403	VIR	C31-C30-C32	2.18	115.25	111.08

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
31	A	9403	VIR	3	0

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

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

6.4 Ligands

EDS was not executed - this section will therefore be empty.

6.5 Other polymers

EDS was not executed - this section will therefore be empty.