wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 07:56 AM GMT

PDB ID	:	3CF5
Title	:	Thiopeptide antibiotic Thiostrepton bound to the large ribosomal subunit of Deinococcus radiodurans
Authors	:	Harms, J.M.; Wilson, D.N.; Schluenzen, F.; Connell, S.R.; Stachelhaus, T.; Zaborowska, Z.; Spahn, C.M.T.; Fucini, P.
Deposited on	:	2008-03-02
Resolution	:	3.30 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 (1)) were used in the production of this report:

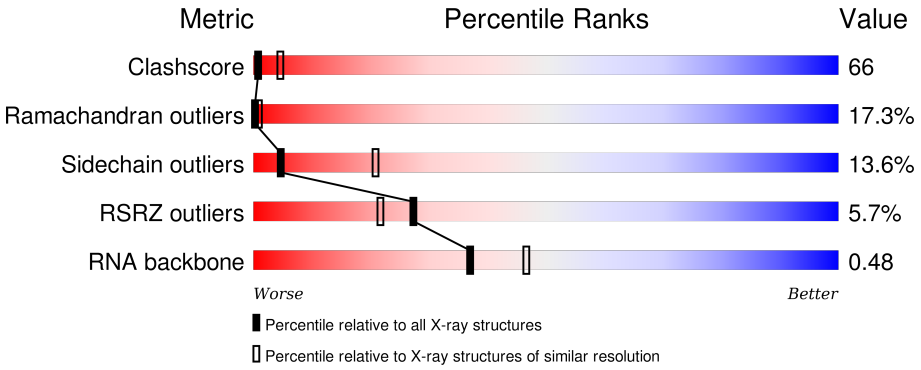
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	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 3.30 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))
Clashscore	102246	1058 (3.38-3.22)
Ramachandran outliers	100387	1038 (3.38-3.22)
Sidechain outliers	100360	1037 (3.38-3.22)
RSRZ outliers	91569	2070 (3.40-3.20)
RNA backbone	2183	1005 (3.82-2.78)

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	1	55	73%, 96%, ·
2	2	47	98%, 98%, ·
3	3	66	89%, 94%, · 5%
4	4	37	19%, 8%, 73%, 19%
5	5	19	5%, 68%, 32%

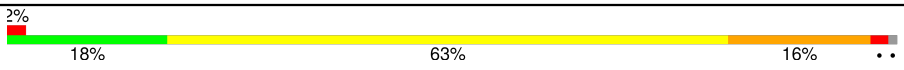
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	A	274	2% 17% 54% 15% 12%
7	B	211	27% 52% 16%
8	C	205	2% 8% 62% 22%
9	D	180	6% 9% 69% 18%
10	E	185	2% 17% 58% 17% 8%
11	F	144	30% 32% 53% 15%
12	G	174	2% 9% 47% 22% 18%
13	H	134	22% 63% 13%
14	I	156	8% 13% 49% 26% 10%
15	J	142	% 15% 57% 18% 5%
16	K	116	24% 53% 19%
17	L	114	4% 8% 54% 25% 5% 9%
18	M	166	16% 31% 15% 35%
19	N	118	13% 64% 20%
20	O	100	14% 51% 29% 6%
21	P	134	19% 61% 13% 5%
22	Q	95	% 8% 58% 29%
23	R	115	4% 8% 59% 24%
24	S	237	6% 8% 51% 12% 26%
25	T	91	10% 22% 54% 15% 8%
26	U	81	4% 6% 52% 23% 7% 11%
27	V	67	16% 60% 21%
28	W	55	16% 62% 18%
29	X	2880	2% 11% 51% 22% 9% 7%
30	Y	60	2% 27% 53% 12% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
31	Z	123	

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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
32	MG	X	2888	-	-	-	X
32	MG	X	2903	-	-	-	X
32	MG	X	2905	-	-	-	X
32	MG	X	2907	-	-	-	X
32	MG	X	2909	-	-	-	X
32	MG	Z	124	-	-	-	X

2 Entry composition (i)

There are 32 unique types of molecules in this entry. The entry contains 84475 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 50S RIBOSOMAL PROTEIN L33.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
1	1	53	Total C 53 53	0	0	53

- Molecule 2 is a protein called 50S RIBOSOMAL PROTEIN L34.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
2	2	46	Total C 46 46	0	0	46

- Molecule 3 is a protein called 50S RIBOSOMAL PROTEIN L35.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
3	3	63	Total C 63 63	0	0	63

- Molecule 4 is a protein called 50S RIBOSOMAL PROTEIN L36.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
4	4	37	Total C N O S 297 179 66 47 5	0	0	0

- Molecule 5 is a protein called THIOSTREPTON.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
5	5	19	Total C N O S 114 72 19 18 5	0	0	1

- Molecule 6 is a protein called 50S RIBOSOMAL PROTEIN L2.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
6	A	240	Total C N O S 1826 1137 366 321 2	0	0	0

- Molecule 7 is a protein called 50S RIBOSOMAL PROTEIN L3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	B	205	Total	C	N	O	S	0	0	0
			1539	965	295	271	8			

- Molecule 8 is a protein called 50S RIBOSOMAL PROTEIN L4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	C	197	Total	C	N	O	S	0	0	0
			1506	935	287	282	2			

- Molecule 9 is a protein called 50S RIBOSOMAL PROTEIN L5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	D	177	Total	C	N	O	S	0	0	0
			1400	892	247	254	7			

- Molecule 10 is a protein called 50S RIBOSOMAL PROTEIN L6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	E	171	Total	C	N	O	S	0	0	0
			1286	812	237	236	1			

- Molecule 11 is a protein called 50S RIBOSOMAL PROTEIN L11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	F	144	Total	C	N	O	S	0	0	0
			1044	663	179	197	5			

- Molecule 12 is a protein called 50S RIBOSOMAL PROTEIN L13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	G	142	Total	C	N	O	S	0	0	0
			1114	704	209	198	3			

- Molecule 13 is a protein called 50S RIBOSOMAL PROTEIN L14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	H	134	Total	C	N	O	S	0	0	0
			997	614	198	180	5			

- Molecule 14 is a protein called 50S RIBOSOMAL PROTEIN L15.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
14	I	141	Total	C	N	O	0	0	0
			1067	655	216	196			

- Molecule 15 is a protein called 50S RIBOSOMAL PROTEIN L16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	J	136	Total	C	N	O	S	0	0	0
			1090	696	202	185	7			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
J	1	MET	-	INITIATING METHIONINE	UNP Q9RXJ5

- Molecule 16 is a protein called 50S RIBOSOMAL PROTEIN L17.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
16	K	113	Total	C	N	O	S	0	0	0
			878	541	178	157	2			

- Molecule 17 is a protein called 50S RIBOSOMAL PROTEIN L18.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
17	L	104	Total	C	N	O	0	0	0
			779	476	161	142			

- Molecule 18 is a protein called 50S RIBOSOMAL PROTEIN L19.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
18	M	108	Total	C	N	O	0	0	0
			871	543	172	156			

- Molecule 19 is a protein called 50S RIBOSOMAL PROTEIN L20.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
19	N	117	Total	C	N	O	S	0	0	0
			978	608	210	159	1			

- Molecule 20 is a protein called 50S RIBOSOMAL PROTEIN L21.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
20	O	94	Total	C	N	O	0	0	0
			741	465	139	137			

- Molecule 21 is a protein called 50S RIBOSOMAL PROTEIN L22.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
21	P	127	Total	C	N	O	S	0	0	0
			1014	639	199	174	2			

- Molecule 22 is a protein called 50S RIBOSOMAL PROTEIN L23.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
22	Q	93	Total	C	N	O	S	0	0	0
			726	458	136	130	2			

- Molecule 23 is a protein called 50S RIBOSOMAL PROTEIN L24.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
23	R	110	Total	C	N	O	S	0	0	0
			825	513	160	151	1			

- Molecule 24 is a protein called 50S RIBOSOMAL PROTEIN L25.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
24	S	175	Total	C	N	O	S	0	0	0
			1345	849	236	254	6			

- Molecule 25 is a protein called 50S RIBOSOMAL PROTEIN L27.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
25	T	84	Total	C	N	O	S	0	0	0
			625	393	122	109	1			

- Molecule 26 is a protein called 50S RIBOSOMAL PROTEIN L28.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
26	U	72	Total	C	N	O	0	0	0
			552	341	116	95			

- Molecule 27 is a protein called 50S RIBOSOMAL PROTEIN L29.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
27	V	66	Total	C	N	O	S	0	0	0
			533	327	107	96	3			

- Molecule 28 is a protein called 50S RIBOSOMAL PROTEIN L30.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
28	W	55	Total	C	N	O	S	0	0	0
			424	264	82	76	2			

- Molecule 29 is a RNA chain called RRNA-23S RIBOSOMAL RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
29	X	2686	Total	C	N	O	P	0	0	0
			57651	25718	10642	18606	2685			

- Molecule 30 is a protein called 50S RIBOSOMAL PROTEIN L32.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
30	Y	58	Total	C	N	O	S	0	0	0
			457	281	94	77	5			

- Molecule 31 is a RNA chain called RRNA-5S RIBOSOMAL RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
31	Z	122	Total	C	N	O	P	0	0	0
			2598	1161	476	840	121			

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

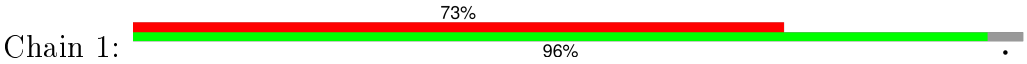
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
32	X	30	Total	Mg	0	0
			30	30		
32	Z	5	Total	Mg	0	0
			5	5		
32	M	1	Total	Mg	0	0
			1	1		

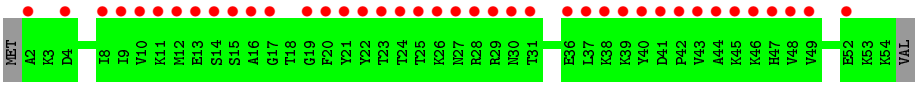
3 Residue-property plots (i)

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: 50S RIBOSOMAL PROTEIN L33

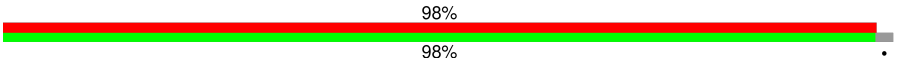
Chain 1:

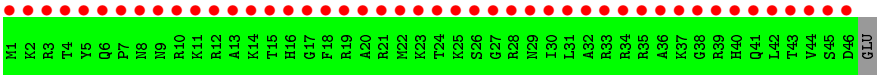




- Molecule 2: 50S RIBOSOMAL PROTEIN L34

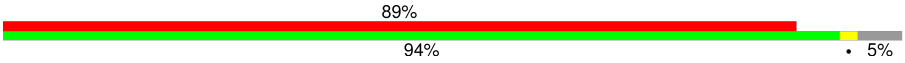
Chain 2:

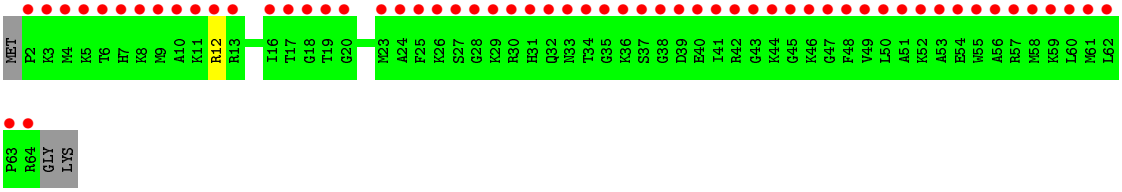




- Molecule 3: 50S RIBOSOMAL PROTEIN L35

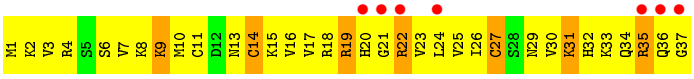
Chain 3:





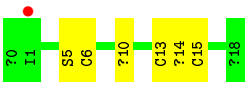
- Molecule 4: 50S RIBOSOMAL PROTEIN L36

Chain 4:



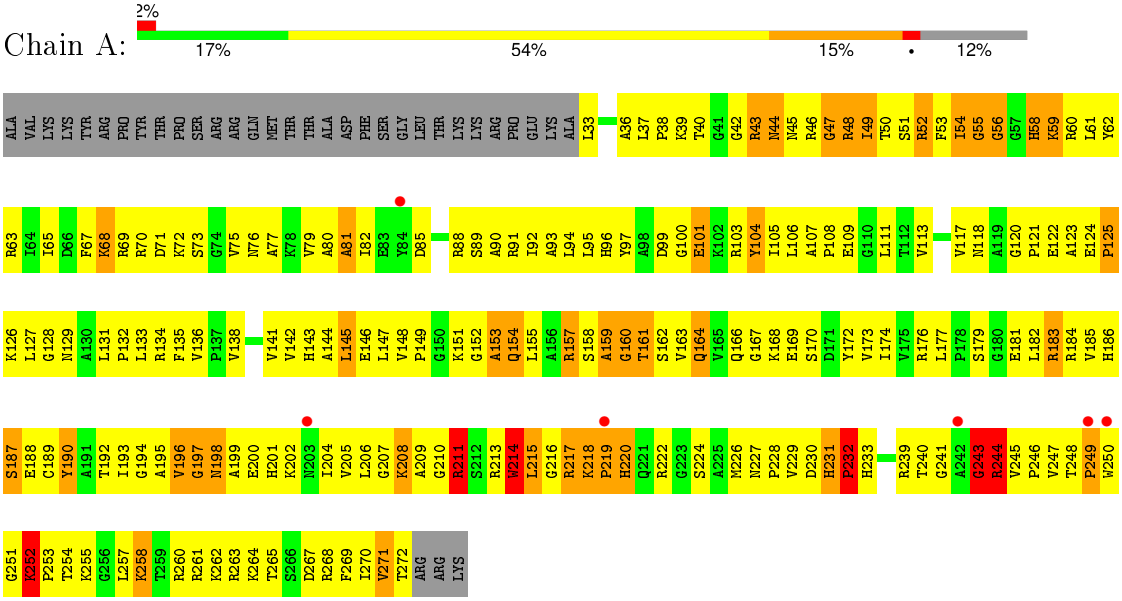
- Molecule 5: THIOSTREPTON

Chain 5:



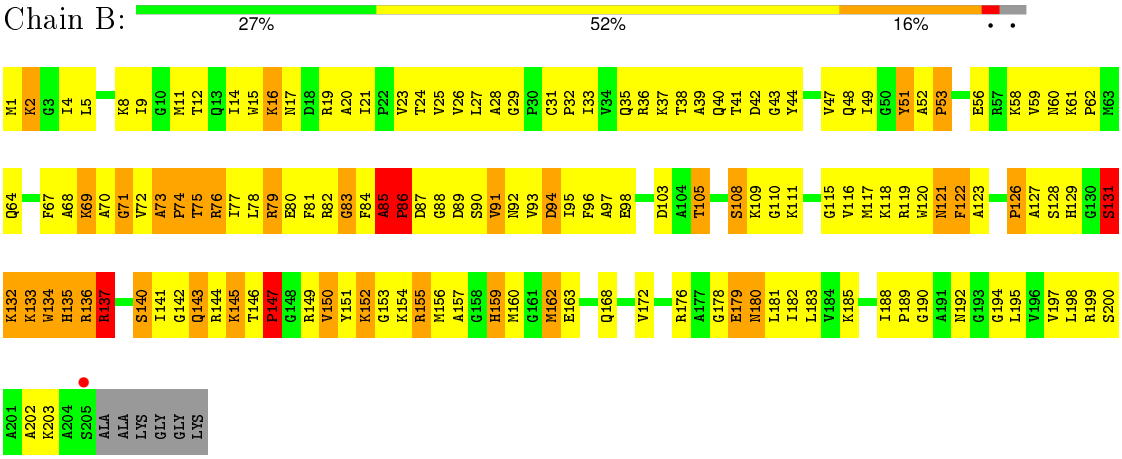
- Molecule 6: 50S RIBOSOMAL PROTEIN L2

Chain A:



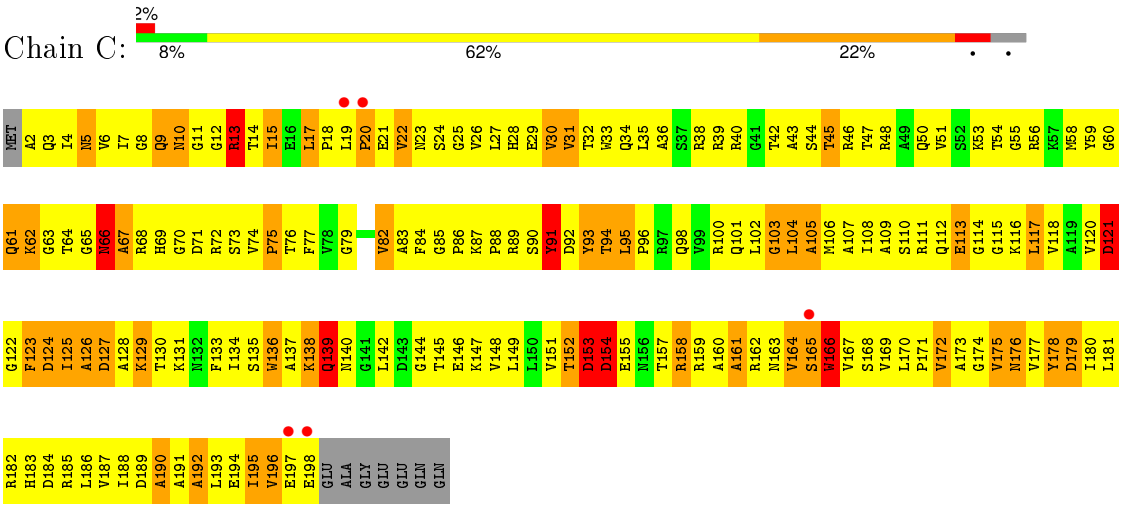
• Molecule 7: 50S RIBOSOMAL PROTEIN L3

Chain B:



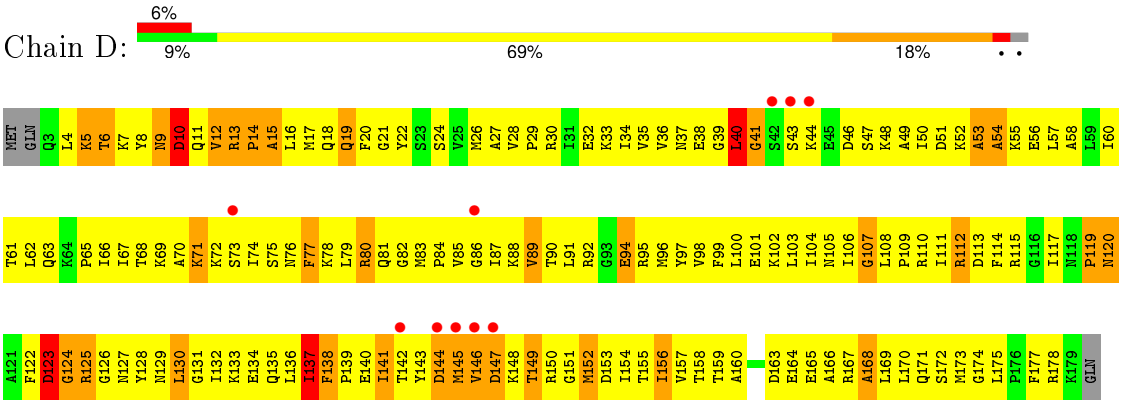
• Molecule 8: 50S RIBOSOMAL PROTEIN L4

Chain C:



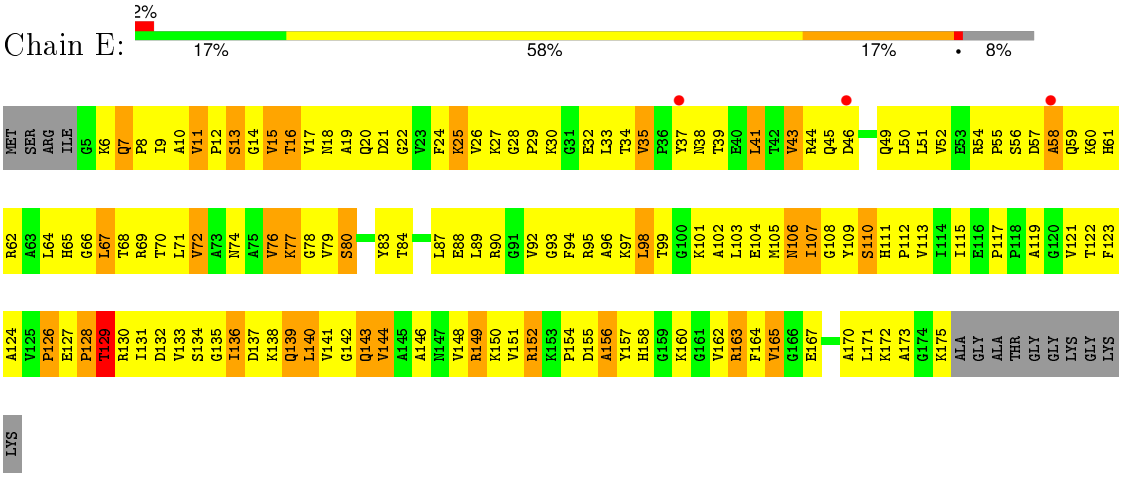
• Molecule 9: 50S RIBOSOMAL PROTEIN L5

Chain D:



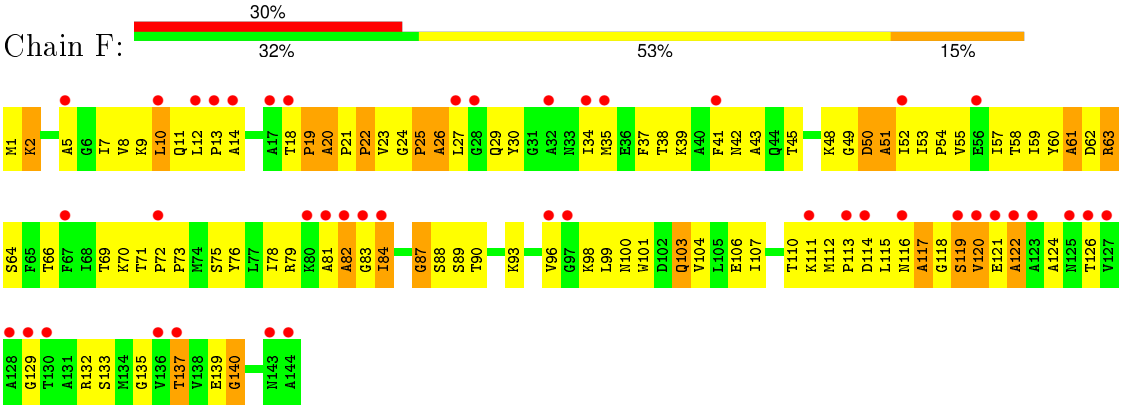
• Molecule 10: 50S RIBOSOMAL PROTEIN L6

Chain E:



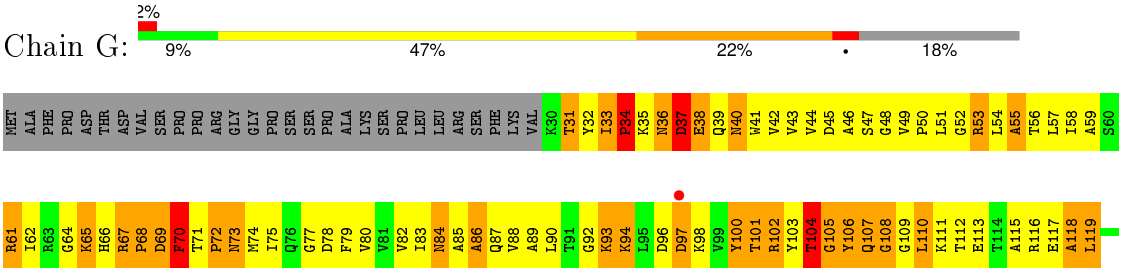
• Molecule 11: 50S RIBOSOMAL PROTEIN L11

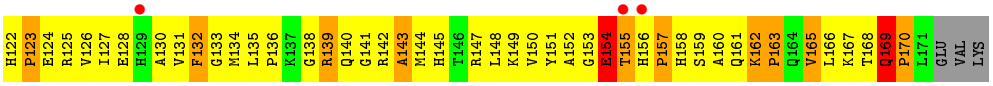
Chain F:



• Molecule 12: 50S RIBOSOMAL PROTEIN L13

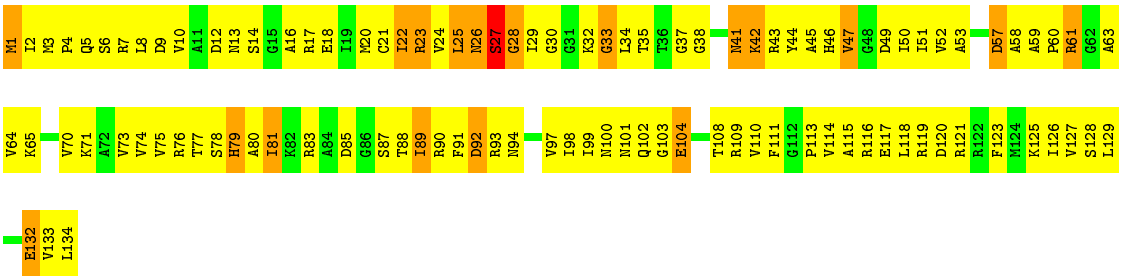
Chain G:





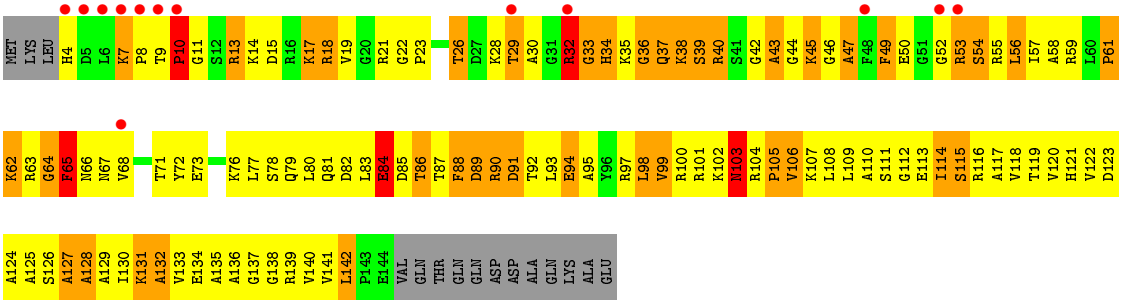
- Molecule 13: 50S RIBOSOMAL PROTEIN L14

Chain H:



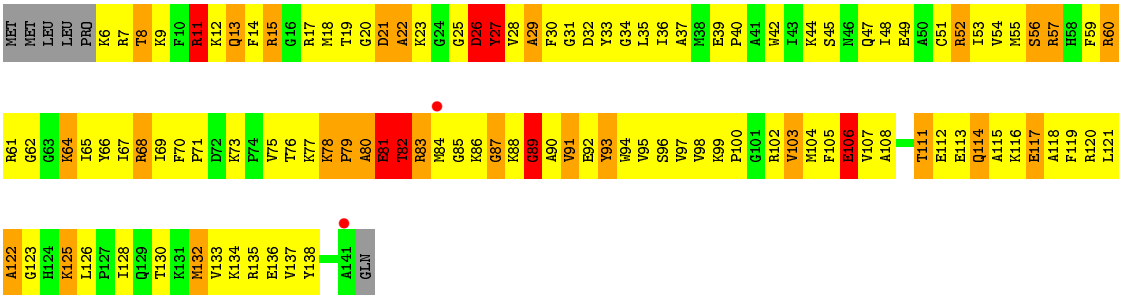
- Molecule 14: 50S RIBOSOMAL PROTEIN L15

Chain I:



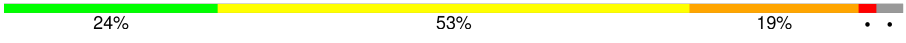
- Molecule 15: 50S RIBOSOMAL PROTEIN L16

Chain J:



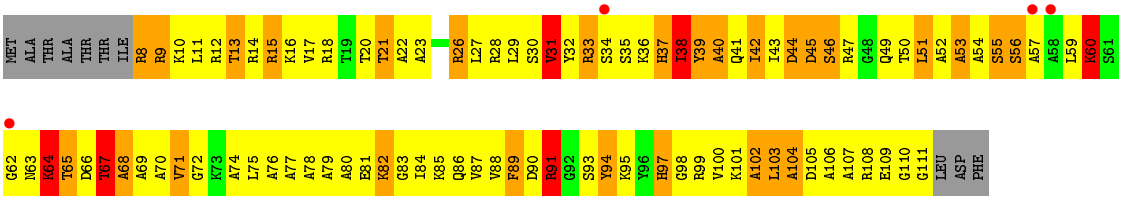
- Molecule 16: 50S RIBOSOMAL PROTEIN L17

Chain K:



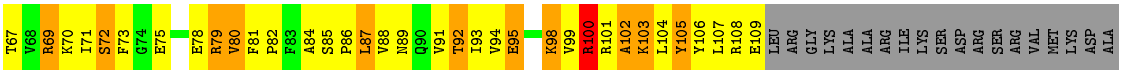
• Molecule 17: 50S RIBOSOMAL PROTEIN L18

Chain L:



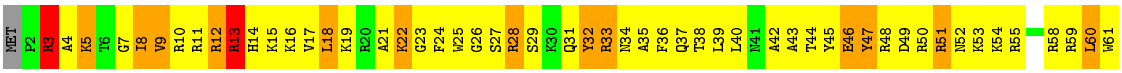
• Molecule 18: 50S RIBOSOMAL PROTEIN L19

Chain M:



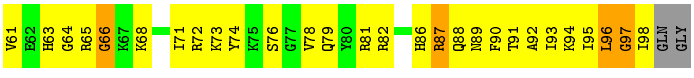
• Molecule 19: 50S RIBOSOMAL PROTEIN L20

Chain N:



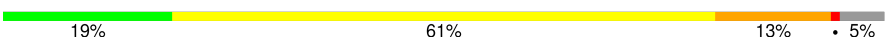
• Molecule 20: 50S RIBOSOMAL PROTEIN L21

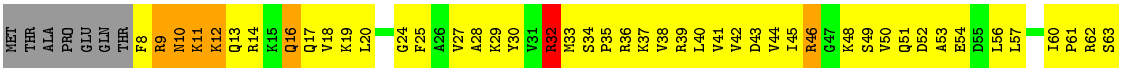
Chain O:

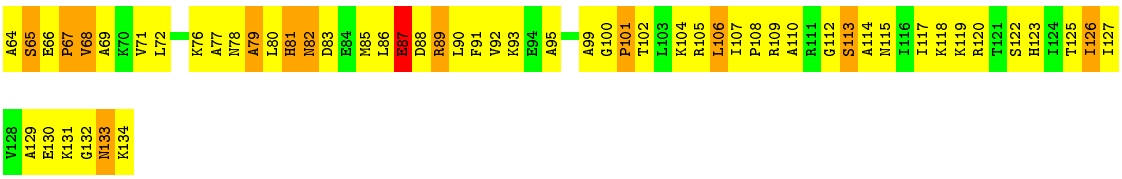


• Molecule 21: 50S RIBOSOMAL PROTEIN L22

Chain P:

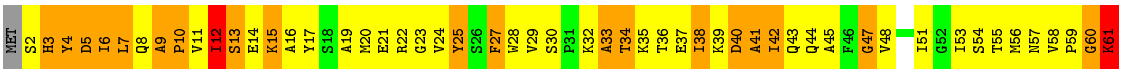


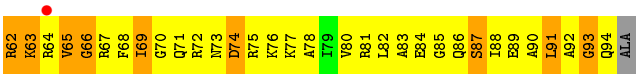




• Molecule 22: 50S RIBOSOMAL PROTEIN L23

Chain Q:



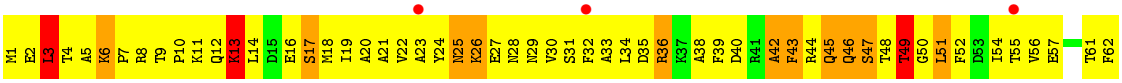


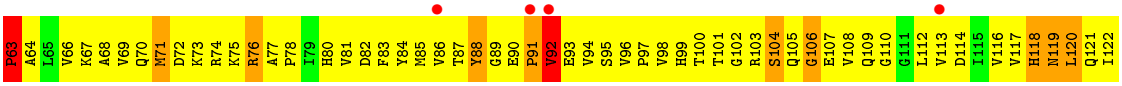
• Molecule 23: 50S RIBOSOMAL PROTEIN L24

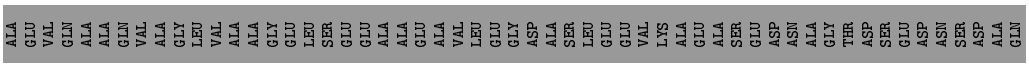
Chain R:

• Molecule 24: 50S RIBOSOMAL PROTEIN L25

Chain S:

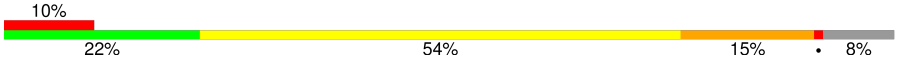


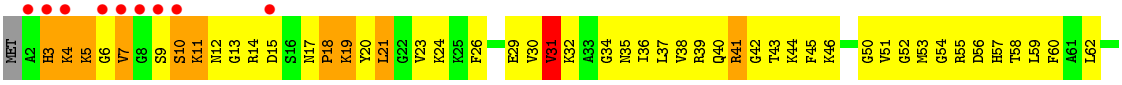


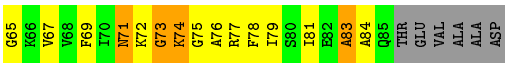


• Molecule 25: 50S RIBOSOMAL PROTEIN L27

Chain T:

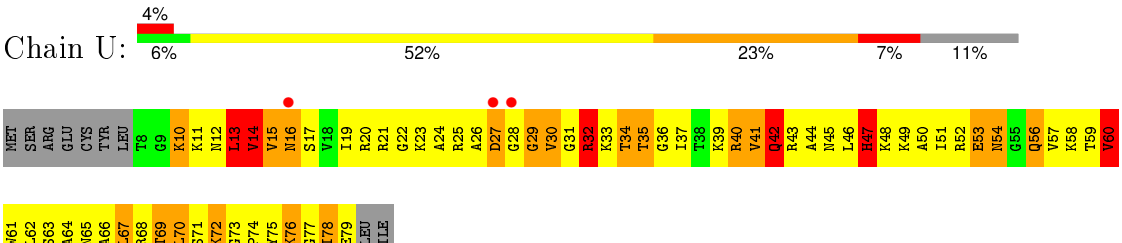






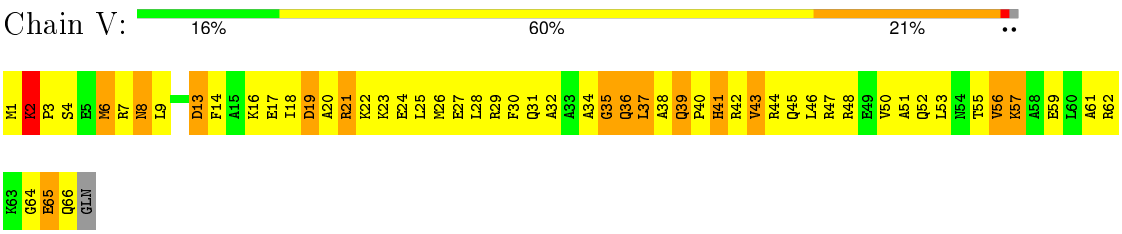
- Molecule 26: 50S RIBOSOMAL PROTEIN L28

Chain U:



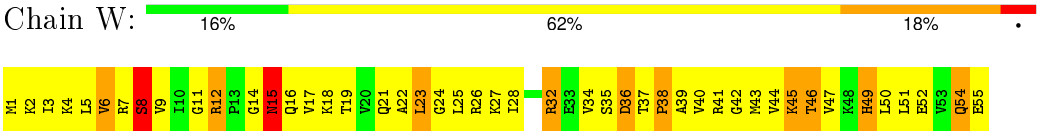
- Molecule 27: 50S RIBOSOMAL PROTEIN L29

Chain V:



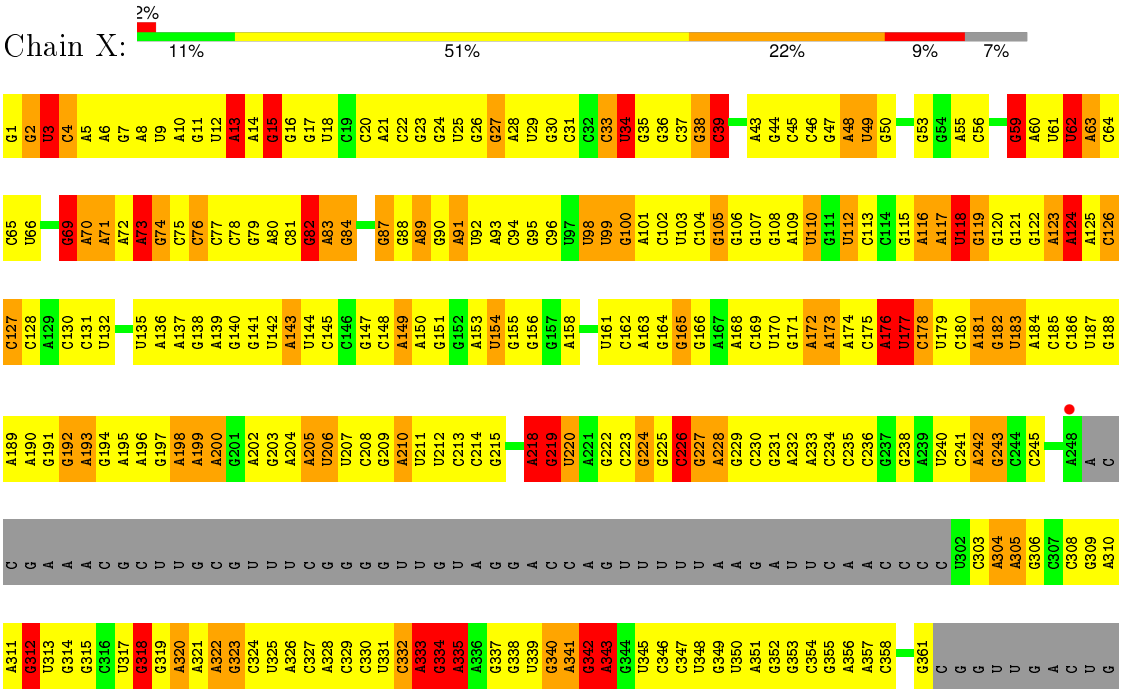
- Molecule 28: 50S RIBOSOMAL PROTEIN L30

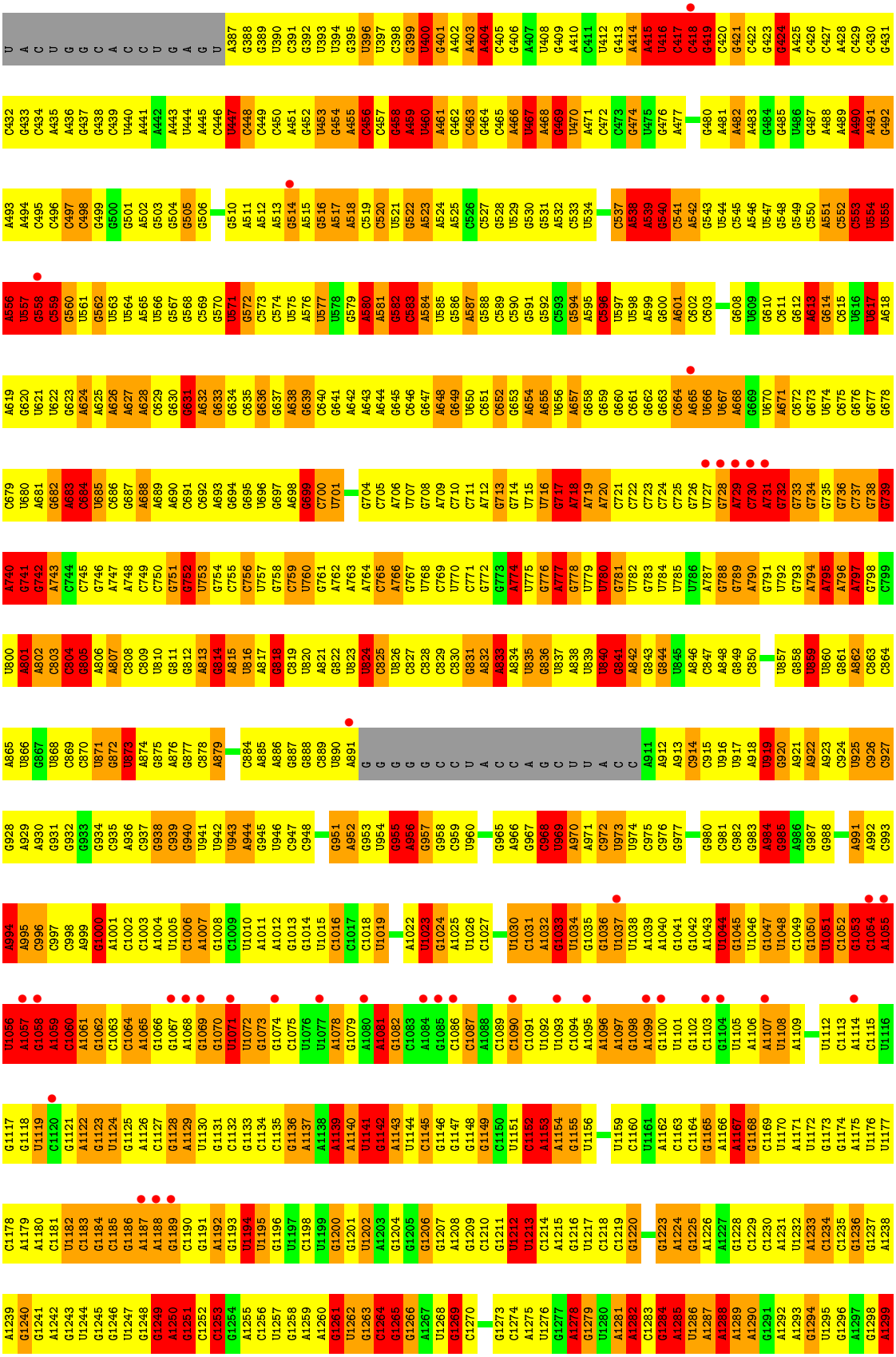
Chain W:

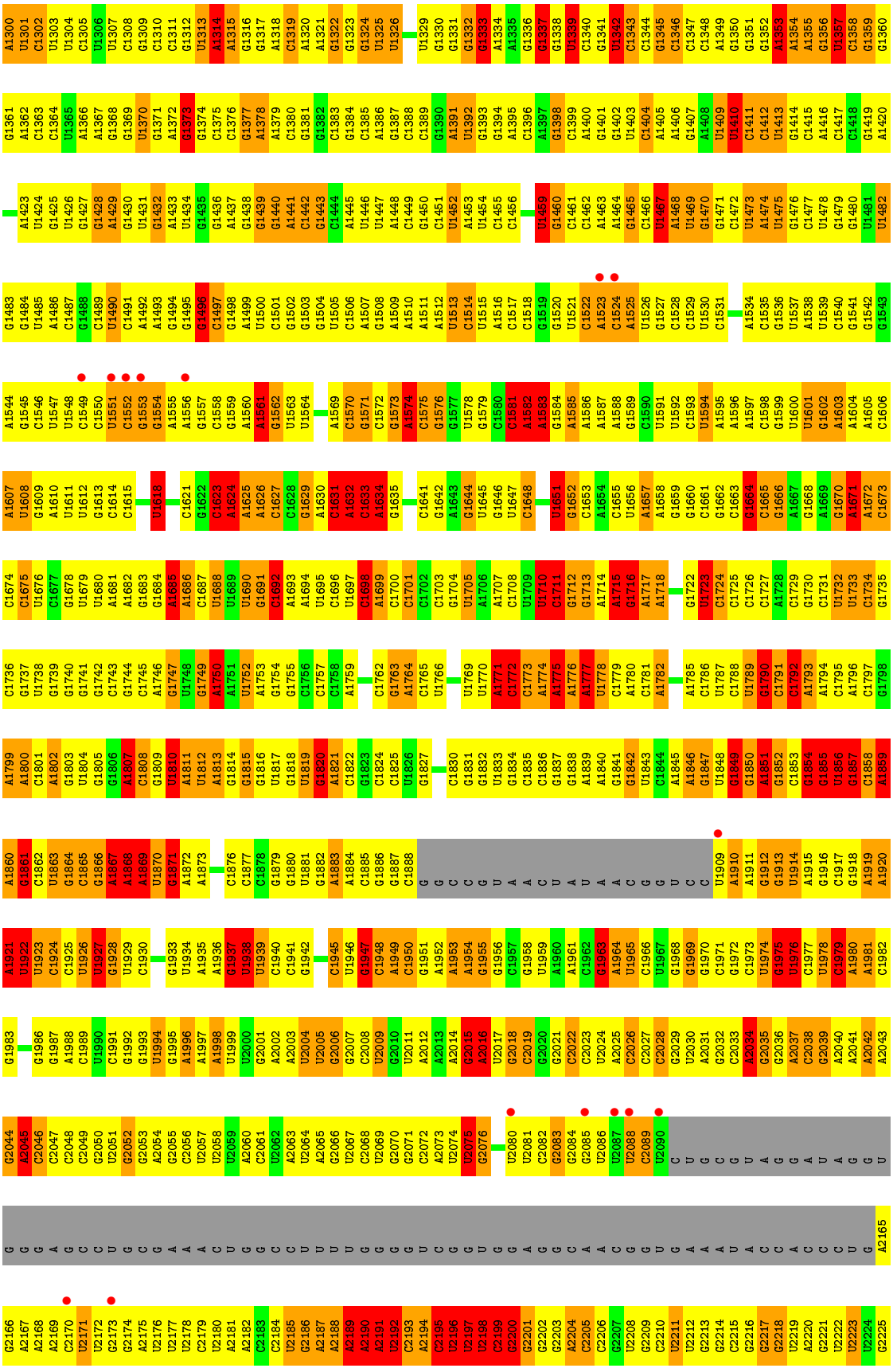


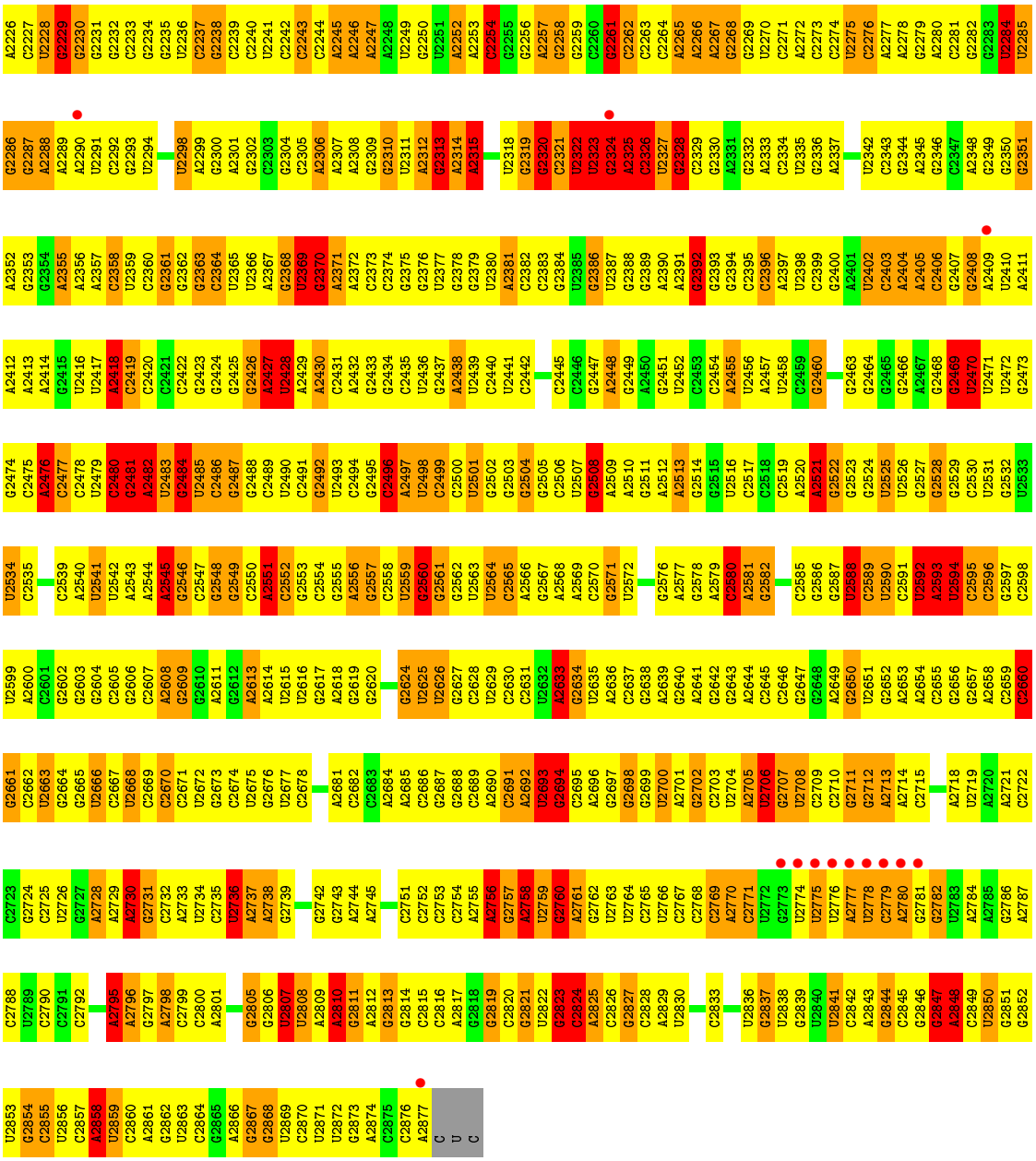
- Molecule 29: RRNA-23S RIBOSOMAL RNA

Chain X:



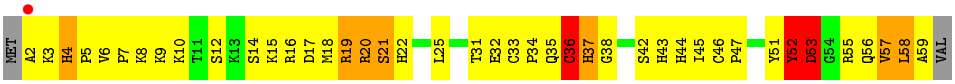






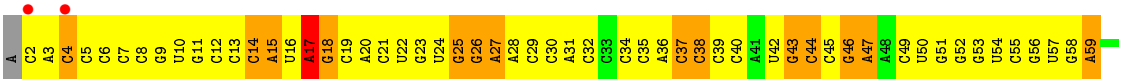
- Molecule 30: 50S RIBOSOMAL PROTEIN L32

Chain Y:



- Molecule 31: RRNA-5S RIBOSOMAL RNA

Chain Z:




C62
A63
C64
A65
G66
C67
A68
G69
C70
G71
C72
C73
A74
A75
U76
G77
A78
U79
A80
C81
U82
C83
G84
G85
G93
G94
U95
C96
C97
G100
A101
A102
A103
U106
C107
G108
G109
U110
C111
A112
G113
C114
G115
C116
G117
G118
G119
G120
G121
U122
U123

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	169.90Å 408.90Å 694.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 3.30 29.76 – 3.31	Depositor EDS
% Data completeness (in resolution range)	94.1 (30.00-3.30) 93.2 (29.76-3.31)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.14 (at 3.31Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.276 , 0.318 0.241 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å^2)	67.5	Xtriage
Anisotropy	0.110	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.18 , 60.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.42$, $< L^2 > = 0.25$	Xtriage
Outliers	1 of 332876 reflections (0.000%)	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	84475	wwPDB-VP
Average B, all atoms (Å^2)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.08% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: MG, DCY, DHA, QUA, BB9, NH2, MH6, DBU, TS9

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	#$\|Z\|>5$
4	4	0.46	0/298	0.67	0/390
5	5	1.46	0/31	1.18	0/38
6	A	0.55	0/1862	0.85	2/2510 (0.1%)
7	B	0.77	0/1567	1.04	4/2105 (0.2%)
8	C	0.63	0/1529	0.91	0/2070
9	D	0.48	0/1419	0.71	0/1903
10	E	0.48	0/1308	0.80	1/1771 (0.1%)
11	F	0.50	0/1063	0.71	0/1440
12	G	0.69	0/1138	1.00	3/1539 (0.2%)
13	H	0.79	0/1007	0.96	1/1352 (0.1%)
14	I	0.65	0/1081	0.94	3/1448 (0.2%)
15	J	0.67	0/1113	0.91	2/1486 (0.1%)
16	K	0.87	0/886	1.06	2/1188 (0.2%)
17	L	0.52	0/785	0.86	0/1048
18	M	0.73	0/884	1.20	6/1186 (0.5%)
19	N	0.63	0/994	0.89	0/1323
20	O	0.61	0/750	0.90	0/1000
21	P	0.77	0/1027	0.93	1/1373 (0.1%)
22	Q	0.67	0/737	0.98	4/988 (0.4%)
23	R	0.55	0/835	0.95	2/1121 (0.2%)
24	S	0.50	0/1370	0.75	0/1862
25	T	0.56	0/633	0.83	1/838 (0.1%)
26	U	0.58	0/556	0.95	1/741 (0.1%)
27	V	0.44	0/537	0.67	0/714
28	W	0.56	0/426	0.84	0/568
29	X	0.88	59/64561 (0.1%)	1.05	497/100708 (0.5%)
30	Y	0.70	0/469	1.11	2/629 (0.3%)
31	Z	0.55	0/2904	0.76	0/4525
All	All	0.81	59/91770 (0.1%)	1.01	532/137864 (0.4%)

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
6	A	0	1
19	N	0	2
22	Q	0	1
29	X	2	257
30	Y	0	1
31	Z	0	4
All	All	2	266

The worst 5 of 59 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
29	X	1856	U	C4'-C3'	-9.30	1.43	1.53
29	X	1856	U	O3'-P	-8.64	1.50	1.61
29	X	1056	U	P-O5'	8.52	1.68	1.59
29	X	1855	G	O3'-P	-8.11	1.51	1.61
29	X	551	A	O3'-P	-8.05	1.51	1.61

The worst 5 of 532 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
29	X	1055	A	N9-C1'-C2'	-29.46	75.70	114.00
29	X	2324	G	N9-C1'-C2'	22.22	142.88	114.00
29	X	557	U	N1-C1'-C2'	19.61	139.50	114.00
29	X	417	C	N1-C1'-C2'	18.73	138.35	114.00
18	M	28	ARG	C-N-CD	-18.52	79.85	120.60

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
29	X	1278	A	C1'
29	X	2592	U	C1'

5 of 266 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
6	A	104	TYR	Sidechain
19	N	32	TYR	Sidechain
19	N	76	TYR	Sidechain
22	Q	25	TYR	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
29	X	12	U	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	1	53	0	0	0	0
2	2	46	0	0	0	0
3	3	63	0	0	1	0
4	4	297	0	330	62	0
5	5	114	0	79	5	0
6	A	1826	0	1885	451	0
7	B	1539	0	1600	303	0
8	C	1506	0	1525	371	0
9	D	1400	0	1481	373	0
10	E	1286	0	1336	264	0
11	F	1044	0	1088	176	0
12	G	1114	0	1144	310	0
13	H	997	0	1046	194	0
14	I	1067	0	1103	301	0
15	J	1090	0	1125	273	0
16	K	878	0	930	135	0
17	L	779	0	820	231	0
18	M	871	0	894	208	0
19	N	978	0	1020	239	0
20	O	741	0	756	186	0
21	P	1014	0	1096	181	0
22	Q	726	0	753	150	0
23	R	825	0	881	266	0
24	S	1345	0	1372	303	0
25	T	625	0	655	111	0
26	U	552	0	604	207	0
27	V	533	0	558	109	0
28	W	424	0	470	83	0
29	X	57651	0	29049	4301	0
30	Y	457	0	462	86	0
31	Z	2598	0	1328	185	0
32	M	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	X	30	0	0	0	0
32	Z	5	0	0	0	0
All	All	84475	0	55390	9214	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 66.

The worst 5 of 9214 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)
7:B:116:VAL:N	7:B:136:ARG:HE	1.23	1.30
29:X:1053:G:H2'	29:X:1054:C:C6	1.70	1.26
29:X:2196:U:H2'	29:X:2197:U:O4'	1.31	1.23
29:X:2736:U:O2'	29:X:2737:A:H5"	1.36	1.21
29:X:2496:C:O2'	29:X:2497:A:H3'	1.40	1.19

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
4	4	35/37 (95%)	20 (57%)	10 (29%)	5 (14%)	0 1
5	5	5/19 (26%)	4 (80%)	1 (20%)	0	100 100
6	A	238/274 (87%)	154 (65%)	50 (21%)	34 (14%)	0 1
7	B	203/211 (96%)	148 (73%)	32 (16%)	23 (11%)	0 3
8	C	195/205 (95%)	97 (50%)	54 (28%)	44 (23%)	0 0
9	D	175/180 (97%)	95 (54%)	48 (27%)	32 (18%)	0 1
10	E	169/185 (91%)	100 (59%)	38 (22%)	31 (18%)	0 1
11	F	142/144 (99%)	94 (66%)	29 (20%)	19 (13%)	0 2

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
12	G	140/174 (80%)	76 (54%)	34 (24%)	30 (21%)	0 0
13	H	132/134 (98%)	105 (80%)	18 (14%)	9 (7%)	1 12
14	I	139/156 (89%)	59 (42%)	45 (32%)	35 (25%)	0 0
15	J	134/142 (94%)	82 (61%)	31 (23%)	21 (16%)	0 1
16	K	111/116 (96%)	74 (67%)	25 (22%)	12 (11%)	0 4
17	L	102/114 (90%)	59 (58%)	19 (19%)	24 (24%)	0 0
18	M	106/166 (64%)	70 (66%)	23 (22%)	13 (12%)	0 2
19	N	115/118 (98%)	57 (50%)	40 (35%)	18 (16%)	0 1
20	O	92/100 (92%)	57 (62%)	10 (11%)	25 (27%)	0 0
21	P	125/134 (93%)	89 (71%)	21 (17%)	15 (12%)	0 3
22	Q	91/95 (96%)	39 (43%)	28 (31%)	24 (26%)	0 0
23	R	108/115 (94%)	62 (57%)	27 (25%)	19 (18%)	0 1
24	S	173/237 (73%)	93 (54%)	46 (27%)	34 (20%)	0 1
25	T	82/91 (90%)	47 (57%)	19 (23%)	16 (20%)	0 1
26	U	70/81 (86%)	35 (50%)	16 (23%)	19 (27%)	0 0
27	V	64/67 (96%)	35 (55%)	16 (25%)	13 (20%)	0 1
28	W	53/55 (96%)	38 (72%)	9 (17%)	6 (11%)	0 3
30	Y	56/60 (93%)	40 (71%)	9 (16%)	7 (12%)	0 2
All	All	3055/3410 (90%)	1829 (60%)	698 (23%)	528 (17%)	0 1

5 of 528 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
6	A	59	LYS
6	A	145	LEU
6	A	168	LYS
6	A	217	ARG
6	A	220	HIS

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
4	4	35/35 (100%)	32 (91%)	3 (9%)	13 45
5	5	3/4 (75%)	3 (100%)	0	100 100
6	A	185/215 (86%)	161 (87%)	24 (13%)	5 22
7	B	155/157 (99%)	132 (85%)	23 (15%)	4 17
8	C	157/163 (96%)	131 (83%)	26 (17%)	3 13
9	D	153/156 (98%)	138 (90%)	15 (10%)	10 37
10	E	136/144 (94%)	128 (94%)	8 (6%)	24 63
11	F	107/107 (100%)	100 (94%)	7 (6%)	21 59
12	G	118/146 (81%)	96 (81%)	22 (19%)	2 9
13	H	103/103 (100%)	88 (85%)	15 (15%)	4 18
14	I	108/121 (89%)	91 (84%)	17 (16%)	3 15
15	J	110/116 (95%)	89 (81%)	21 (19%)	2 8
16	K	90/93 (97%)	76 (84%)	14 (16%)	3 16
17	L	74/82 (90%)	54 (73%)	20 (27%)	0 2
18	M	94/134 (70%)	72 (77%)	22 (23%)	1 4
19	N	96/97 (99%)	83 (86%)	13 (14%)	5 21
20	O	75/79 (95%)	70 (93%)	5 (7%)	20 58
21	P	109/115 (95%)	100 (92%)	9 (8%)	14 47
22	Q	75/76 (99%)	67 (89%)	8 (11%)	8 32
23	R	91/96 (95%)	72 (79%)	19 (21%)	1 6
24	S	149/192 (78%)	137 (92%)	12 (8%)	15 48
25	T	62/67 (92%)	57 (92%)	5 (8%)	15 48
26	U	57/66 (86%)	44 (77%)	13 (23%)	1 4
27	V	54/55 (98%)	48 (89%)	6 (11%)	8 31
28	W	48/48 (100%)	38 (79%)	10 (21%)	1 6
30	Y	51/53 (96%)	48 (94%)	3 (6%)	24 63
All	All	2495/2720 (92%)	2155 (86%)	340 (14%)	5 21

5 of 340 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
14	I	65	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
16	K	83	VAL
26	U	47	HIS
14	I	103	ASN
15	J	82	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 78 such sidechains are listed below:

Mol	Chain	Res	Type
13	H	41	ASN
16	K	13	ASN
27	V	45	GLN
13	H	79	HIS
14	I	66	ASN

5.3.3 RNA

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
29	X	2680/2880 (93%)	688 (25%)	313 (11%)
31	Z	121/123 (98%)	24 (19%)	1 (0%)
All	All	2801/3003 (93%)	712 (25%)	314 (11%)

5 of 712 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
29	X	2	G
29	X	4	C
29	X	13	A
29	X	14	A
29	X	27	G

5 of 314 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
29	X	1261	G
29	X	1552	C
29	X	2660	C
29	X	1278	A
29	X	1345	G

5.4 Non-standard residues in protein, DNA, RNA chains (i)

11 non-standard protein/DNA/RNA residues 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	$\#\|Z\| > 2$
5	TS9	5	10	5	6,8,10	0.89	0	6,12,15	1.16	0
5	BB9	5	11	5	3,5,6	1.35	0	1,5,7	2.19	1 (100%)
5	BB9	5	13	5	0,4,6	0.00	-	0,4,7	0.00	-
5	MH6	5	14	5	3,3,6	1.61	1 (33%)	1,3,7	0.51	0
5	BB9	5	15	5	3,5,6	3.63	1 (33%)	1,5,7	3.43	1 (100%)
5	DHA	5	16	5	4,4,5	4.09	2 (50%)	3,4,6	4.29	1 (33%)
5	DHA	5	17	5	4,4,5	1.86	1 (25%)	3,4,6	2.19	2 (66%)
5	DHA	5	3	5	4,4,5	2.07	2 (50%)	3,4,6	1.71	1 (33%)
5	BB9	5	6	5	3,5,6	1.81	1 (33%)	1,5,7	2.21	1 (100%)
5	DBU	5	8	5	2,4,6	0.78	0	0,4,7	0.00	-
5	DCY	5	9	5	4,5,6	0.33	0	3,5,7	3.07	2 (66%)

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
5	TS9	5	10	5	-	0/9/12/16	0/0/0/0
5	BB9	5	11	5	-	0/0/4/6	0/0/0/0
5	BB9	5	13	5	-	0/0/2/6	0/0/0/0
5	MH6	5	14	5	-	0/0/0/6	0/0/0/0
5	BB9	5	15	5	-	0/0/4/6	0/0/0/0
5	DHA	5	16	5	-	0/0/2/4	0/0/0/0
5	DHA	5	17	5	-	0/0/2/4	0/0/0/0
5	DHA	5	3	5	-	0/0/2/4	0/0/0/0
5	BB9	5	6	5	-	0/0/4/6	0/0/0/0
5	DBU	5	8	5	-	0/1/2/6	0/0/0/0
5	DCY	5	9	5	-	0/1/4/6	0/0/0/0

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	5	15	BB9	O-C	-6.26	1.07	1.22
5	5	16	DHA	C-CA	-2.72	1.40	1.45
5	5	14	MH6	C-CA	2.17	1.53	1.49
5	5	3	DHA	C-CA	2.48	1.48	1.45
5	5	3	DHA	CA-N	2.82	1.42	1.34

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	5	16	DHA	O-C-CA	-7.39	111.08	125.35
5	5	9	DCY	CA-CB-SG	-4.49	104.20	114.48
5	5	9	DCY	O-C-CA	-2.86	118.03	125.49
5	5	17	DHA	O-C-CA	-2.81	119.92	125.35
5	5	3	DHA	O-C-CA	-2.35	120.82	125.35

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	5	10	TS9	1	0
5	5	13	BB9	1	0
5	5	14	MH6	1	0
5	5	15	BB9	1	0
5	5	6	BB9	1	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 36 ligands modelled in this entry, 36 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.

No monomer is involved in short contacts.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	1	53/55 (96%)	3.55	40 (75%) 0 0	48, 56, 77, 82	0
2	2	46/47 (97%)	5.68	46 (100%) 0 0	9, 29, 38, 40	0
3	3	63/66 (95%)	5.20	59 (93%) 0 0	23, 41, 51, 57	0
4	4	37/37 (100%)	1.21	7 (18%) 2 1	60, 69, 77, 81	0
5	5	6/19 (31%)	1.14	1 (16%) 2 2	79, 83, 86, 86	0
6	A	240/274 (87%)	-0.09	6 (2%) 61 54	25, 63, 77, 84	0
7	B	205/211 (97%)	-0.68	1 (0%) 91 90	3, 22, 49, 63	0
8	C	197/205 (96%)	-0.29	5 (2%) 61 54	8, 51, 73, 83	0
9	D	177/180 (98%)	0.06	10 (5%) 28 22	60, 75, 85, 91	0
10	E	171/185 (92%)	-0.33	3 (1%) 71 65	44, 66, 79, 88	0
11	F	144/144 (100%)	1.53	43 (29%) 1 1	74, 89, 98, 102	0
12	G	142/174 (81%)	-0.27	4 (2%) 56 50	22, 43, 67, 72	0
13	H	134/134 (100%)	-0.77	0 100 100	3, 16, 37, 45	0
14	I	141/156 (90%)	0.28	13 (9%) 11 9	22, 62, 77, 85	0
15	J	136/142 (95%)	-0.30	2 (1%) 76 71	27, 51, 73, 80	0
16	K	113/116 (97%)	-0.83	0 100 100	3, 9, 24, 34	0
17	L	104/114 (91%)	-0.04	4 (3%) 44 37	43, 62, 72, 75	0
18	M	108/166 (65%)	-0.73	0 100 100	4, 19, 43, 64	0
19	N	117/118 (99%)	-0.56	0 100 100	4, 40, 62, 73	0
20	O	94/100 (94%)	-0.42	0 100 100	18, 53, 71, 81	0
21	P	127/134 (94%)	-0.71	0 100 100	4, 18, 53, 76	0
22	Q	93/95 (97%)	-0.39	1 (1%) 82 78	32, 50, 69, 80	0
23	R	110/115 (95%)	-0.13	5 (4%) 37 30	36, 54, 80, 87	0
24	S	175/237 (73%)	0.26	14 (8%) 15 12	61, 71, 82, 87	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
25	T	84/91 (92%)	0.04	9 (10%) 8 6	35, 51, 80, 90	0
26	U	72/81 (88%)	0.11	3 (4%) 40 33	45, 61, 72, 78	0
27	V	66/67 (98%)	-0.48	0 100 100	49, 61, 81, 88	0
28	W	55/55 (100%)	-0.41	0 100 100	23, 41, 61, 78	0
29	X	2686/2880 (93%)	-0.36	67 (2%) 61 54	4, 41, 116, 151	0
30	Y	58/60 (96%)	-0.51	1 (1%) 73 67	4, 17, 44, 52	0
31	Z	122/123 (99%)	-0.09	3 (2%) 61 54	30, 75, 102, 129	0
All	All	6076/6581 (92%)	-0.13	347 (5%) 27 22	3, 49, 95, 151	0

The worst 5 of 347 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	3	39	ASP	16.2
2	2	26	SER	9.9
2	2	4	THR	9.8
1	1	2	ALA	9.8
3	3	31	HIS	9.7

6.2 Non-standard residues in protein, DNA, RNA chains (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. 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^{th} 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	LLDF	B-factors(Å^2)	Q<0.9
5	BB9	5	15	6/7	0.78	0.48	-	88,88,88,88	0
5	TS9	5	10	9/11	0.87	0.38	-	87,88,89,89	0
5	BB9	5	6	6/7	0.89	0.18	-	82,84,85,86	0
5	BB9	5	11	6/7	0.86	0.23	-	85,87,87,88	0
5	DHA	5	3	5/6	0.77	0.47	-	82,83,84,85	0
5	DHA	5	16	5/6	0.52	0.55	-	83,83,85,86	0
5	BB9	5	13	5/7	0.83	0.22	-	85,86,86,87	0
5	DHA	5	17	5/6	0.55	0.39	-	76,77,78,80	2
5	MH6	5	14	4/7	0.85	0.30	-	86,86,87,87	0
5	DCY	5	9	6/7	0.81	0.18	-	87,87,87,87	0
5	DBU	5	8	5/7	0.93	0.16	-	85,86,87,87	0

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. 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^{th} 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	LLDF	B-factors(Å^2)	Q<0.9
32	MG	X	2907	1/1	0.96	0.73	63.93	17,17,17,17	0
32	MG	X	2905	1/1	0.97	0.50	9.56	13,13,13,13	0
32	MG	X	2888	1/1	0.95	0.30	8.43	3,3,3,3	0
32	MG	X	2903	1/1	0.95	0.30	6.38	3,3,3,3	0
32	MG	Z	124	1/1	0.95	0.31	6.24	26,26,26,26	0
32	MG	X	2909	1/1	0.96	0.21	5.32	3,3,3,3	0
32	MG	X	2895	1/1	0.98	0.25	1.39	3,3,3,3	0
32	MG	X	2910	1/1	0.90	0.38	-	19,19,19,19	0
32	MG	X	2885	1/1	0.93	0.41	-	56,56,56,56	0
32	MG	X	2901	1/1	0.98	0.08	-	60,60,60,60	0
32	MG	X	2887	1/1	0.98	0.18	-	3,3,3,3	0
32	MG	X	2881	1/1	0.95	0.25	-	59,59,59,59	0
32	MG	X	2899	1/1	0.99	0.53	-	3,3,3,3	0
32	MG	X	2882	1/1	0.97	0.36	-	12,12,12,12	0
32	MG	X	2906	1/1	0.98	0.19	-	58,58,58,58	0
32	MG	X	2883	1/1	0.99	0.10	-	49,49,49,49	0
32	MG	X	2897	1/1	0.98	0.47	-	3,3,3,3	0
32	MG	X	2884	1/1	0.83	0.79	-	55,55,55,55	0
32	MG	X	2904	1/1	0.97	0.31	-	6,6,6,6	0
32	MG	X	2900	1/1	0.96	0.26	-	3,3,3,3	0
32	MG	Z	126	1/1	0.95	0.34	-	25,25,25,25	0
32	MG	X	2893	1/1	0.96	0.15	-	13,13,13,13	0
32	MG	M	167	1/1	0.98	0.54	-	3,3,3,3	0
32	MG	X	2894	1/1	0.99	0.40	-	15,15,15,15	0
32	MG	X	2898	1/1	0.90	0.54	-	19,19,19,19	0
32	MG	X	2902	1/1	0.96	0.38	-	24,24,24,24	0
32	MG	Z	128	1/1	0.95	0.09	-	41,41,41,41	0
32	MG	X	2889	1/1	0.95	0.77	-	3,3,3,3	0
32	MG	X	2890	1/1	0.91	0.34	-	49,49,49,49	0
32	MG	X	2886	1/1	0.91	0.26	-	41,41,41,41	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
32	MG	Z	127	1/1	0.95	0.17	-	12,12,12,12	0
32	MG	X	2892	1/1	0.95	0.16	-	22,22,22,22	0
32	MG	X	2896	1/1	0.95	0.27	-	3,3,3,3	0
32	MG	X	2908	1/1	0.97	0.11	-	3,3,3,3	0
32	MG	X	2891	1/1	0.99	0.41	-	12,12,12,12	0
32	MG	Z	125	1/1	0.98	0.29	-	9,9,9,9	0

6.5 Other polymers

There are no such residues in this entry.