wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 06:21 PM GMT

PDB ID	:	4KZY
Title	:	Rabbit 40S ribosomal subunit in complex with eIF1 and eIF1A.
Authors	:	Lomakin, I.B.; Steitz, T.A.
Deposited on	:	2013-05-30
Resolution	:	7.01 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:

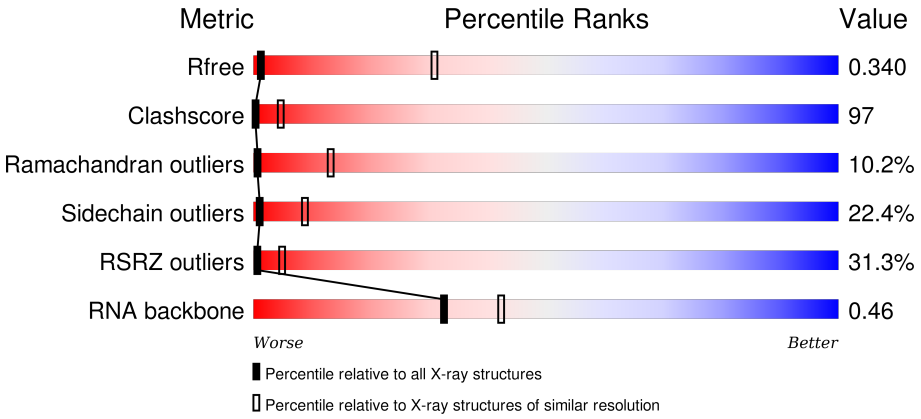
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 7.01 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1014 (9.50-3.66)
Clashscore	102246	1063 (10.00-3.70)
Ramachandran outliers	100387	1035 (9.50-3.66)
Sidechain outliers	100360	1005 (9.50-3.66)
RSRZ outliers	91569	1013 (9.50-3.66)
RNA backbone	2183	1105 (10.00-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	295	7%, 9%, 39%, 19%, ·, 29%
2	B	264	42%, 14%, 45%, 19%, ·, 19%
3	C	278	15%, 14%, 46%, 18%, ·, 19%
4	D	243	71%, 21%, 45%, 23%, 5%, 7%

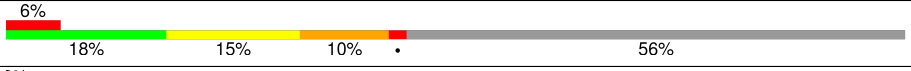
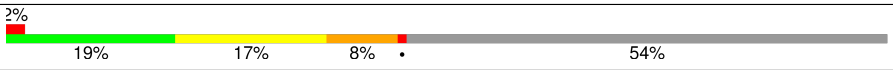
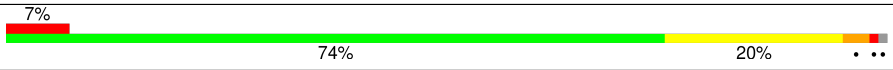
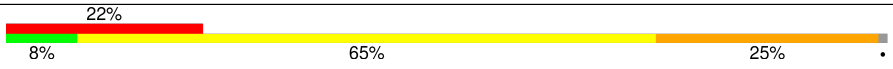
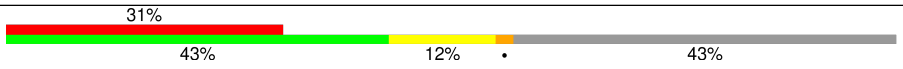
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	263	33% 23% 48% 25%
6	F	204	48% 21% 49% 17% 6% 6%
7	G	249	10% 21% 48% 22% 5%
8	H	194	24% 16% 48% 24% 10%
9	I	208	48% 30% 45% 15% 9%
10	J	194	25% 14% 48% 22% 9% 6%
11	K	165	41% 5% 27% 15% 12% 41%
12	L	158	39% 27% 42% 23% 8%
13	M	132	7% 21% 43% 23% 7% 6%
14	N	151	18% 29% 47% 21%
15	O	151	17% 15% 50% 23% 10%
16	P	145	30% 12% 40% 26% 10% 12%
17	Q	146	41% 18% 53% 20% 5%
18	R	135	18% 19% 54% 11% 10% 7%
19	S	152	24% 18% 43% 18% 10% 10%
20	T	145	20% 54% 19% 5%
21	U	119	81% 15% 36% 27% 9% 13%
22	V	83	16% 14% 43% 28% 13%
23	W	130	74% 29% 57% 12%
24	X	143	70% 26% 48% 20% 6%
25	Y	133	12% 14% 47% 24% 9% 5%
26	Z	125	54% 11% 31% 11% 6% 40%
27	a	115	39% 62% 23% 7% 7%
28	b	84	24% 65% 25% 10%
29	c	69	35% 65% 20% 7% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
30	d	56	
31	e	133	
32	f	156	
33	g	317	
34	i	1863	
35	l	113	
36	n	144	

2 Entry composition

There are 36 unique types of molecules in this entry. The entry contains 78412 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 40S Ribosomal Protein SA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	208	Total	C	N	O	S	0	0	0
			1642	1045	289	300	8			

- Molecule 2 is a protein called 40S Ribosomal Protein S3A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	215	Total	C	N	O	S	0	0	0
			1741	1107	309	310	15			

- Molecule 3 is a protein called 40S Ribosomal Protein S2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	226	Total	C	N	O	S	0	0	0
			1742	1127	300	306	9			

- Molecule 4 is a protein called 40S Ribosomal Protein S3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	227	Total	C	N	O	S	0	0	0
			1764	1124	317	315	8			

- Molecule 5 is a protein called 40S Ribosomal Protein S4X.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	263	Total	C	N	O	S	0	0	0
			2083	1329	385	359	10			

- Molecule 6 is a protein called 40S Ribosomal Protein S5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	191	Total	C	N	O	S	0	0	0
			1509	943	286	273	7			

- Molecule 7 is a protein called 40S Ribosomal Protein S6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	237	Total	C	N	O	S	0	0	0
			1923	1200	387	329	7			

- Molecule 8 is a protein called 40S Ribosomal Protein S7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	190	Total	C	N	O	S	0	0	0
			1530	975	281	273	1			

- Molecule 9 is a protein called 40S Ribosomal Protein S8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	206	Total	C	N	O	S	0	0	0
			1679	1054	329	291	5			

- Molecule 10 is a protein called 40S Ribosomal Protein S9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	182	Total	C	N	O	S	0	0	0
			1498	952	300	244	2			

- Molecule 11 is a protein called 40S Ribosomal Protein S10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	98	Total	C	N	O	S	0	0	0
			827	539	148	134	6			

- Molecule 12 is a protein called 40S Ribosomal Protein S11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	158	Total	C	N	O	S	0	0	0
			1296	827	241	221	7			

- Molecule 13 is a protein called 40S Ribosomal Protein S12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	124	Total	C	N	O	S	0	0	0
			951	594	169	179	9			

- Molecule 14 is a protein called 40S Ribosomal Protein S13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	N	150	Total	C	N	O	S	0	0	0
			1208	773	229	205	1			

- Molecule 15 is a protein called 40S Ribosomal Protein S14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	O	136	Total	C	N	O	S	0	0	0
			1016	621	199	190	6			

- Molecule 16 is a protein called 40S Ribosomal Protein S15.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
16	P	127	Total	C	N	O	S	0	0	0
			1060	673	201	179	7			

- Molecule 17 is a protein called 40S Ribosomal Protein S16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
17	Q	141	Total	C	N	O	S	0	0	0
			1124	715	212	194	3			

- Molecule 18 is a protein called 40S Ribosomal Protein S17.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
18	R	126	Total	C	N	O	S	0	0	0
			1019	639	188	187	5			

- Molecule 19 is a protein called 40S Ribosomal Protein S18.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
19	S	137	Total	C	N	O	S	0	0	0
			1139	714	231	193	1			

- Molecule 20 is a protein called 40S Ribosomal Protein S19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
20	T	141	Total	C	N	O	S	0	0	0
			1112	701	213	195	3			

- Molecule 21 is a protein called 40S Ribosomal Protein S20.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
21	U	104	Total	C	N	O	S	0	0	0
			822	514	156	148	4			

- Molecule 22 is a protein called 40S Ribosomal Protein S21.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
22	V	82	Total	C	N	O	S	0	0	0
			619	378	117	119	5			

- Molecule 23 is a protein called 40S Ribosomal Protein S15A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
23	W	129	Total	C	N	O	S	0	0	0
			1034	659	193	176	6			

- Molecule 24 is a protein called 40S Ribosomal Protein S23.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
24	X	142	Total	C	N	O	S	0	0	0
			1106	698	220	184	4			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
X	1	MET	ALA	SEE REMARK 999	UNP G1SZ47

- Molecule 25 is a protein called 40S Ribosomal Protein S24.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
25	Y	126	Total	C	N	O	S	0	0	0
			1021	645	198	173	5			

- Molecule 26 is a protein called 40S Ribosomal Protein S25.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
26	Z	75	Total	C	N	O	S	0	0	0
			598	382	111	104	1			

- Molecule 27 is a protein called 40S Ribosomal Protein S26.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
27	a	107	Total	C	N	O	S	0	0	0
			844	527	173	138	6			

- Molecule 28 is a protein called 40S Ribosomal Protein S27.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
28	b	84	Total	C	N	O	S	0	0	0
			659	413	122	116	8			

- Molecule 29 is a protein called 40S Ribosomal Protein S28.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
29	c	64	Total	C	N	O	S	0	0	0
			506	308	102	94	2			

- Molecule 30 is a protein called 40S Ribosomal Protein S29.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
30	d	53	Total	C	N	O	S	0	0	0
			445	278	90	72	5			

- Molecule 31 is a protein called 40S Ribosomal Protein S30.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
31	e	59	Total	C	N	O	S	0	0	0
			468	290	102	75	1			

- Molecule 32 is a protein called 40S Ribosomal Protein S27A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
32	f	71	Total	C	N	O	S	0	0	0
			581	367	109	98	7			

- Molecule 33 is a protein called 40S Ribosomal Protein RACK1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
33	g	313	Total	C	N	O	S	0	0	0
			2436	1535	424	465	12			

- Molecule 34 is a RNA chain called 18S Ribosomal RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
34	i	1840	Total	C	N	O	P	0	0	0
			38071	16944	6695	12593	1839			

- Molecule 35 is a protein called human initiation factor eIF1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
35	l	85	Total	C	N	O	S	0	0	0
			691	438	125	126	2			

- Molecule 36 is a protein called human initiation factor eIF1A.

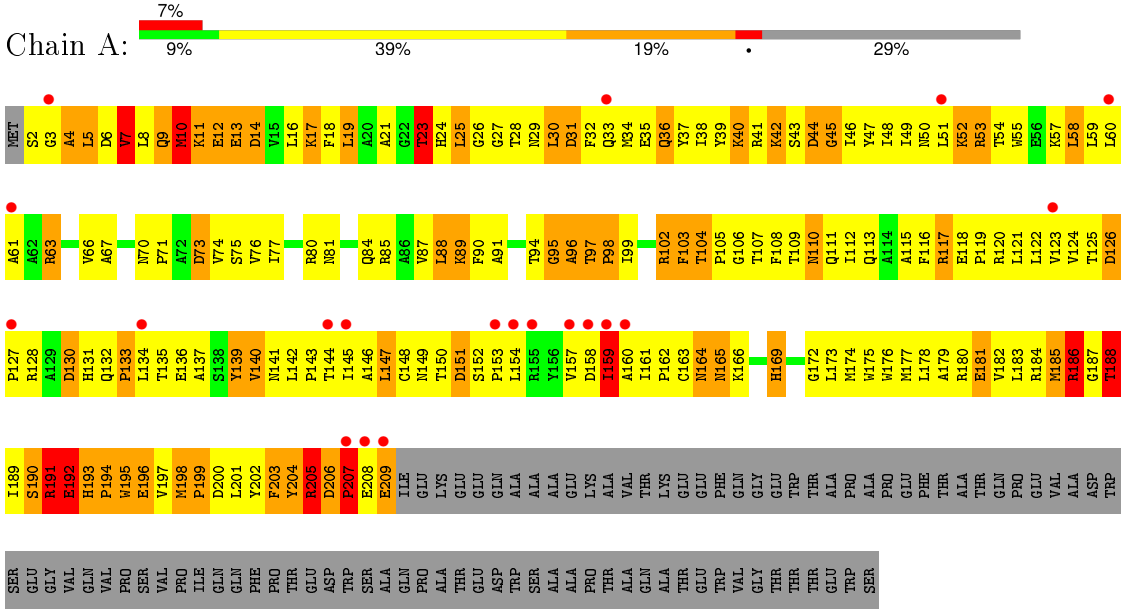
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
36	n	82	Total	C	N	O	S	0	0	0
			648	409	118	117	4			

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.

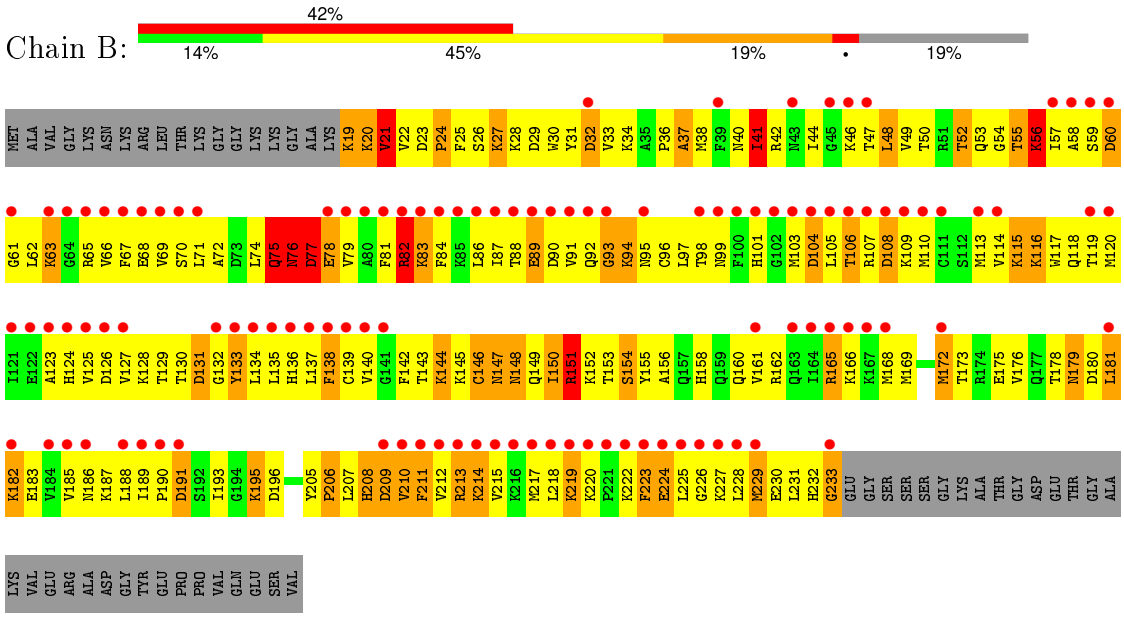
- Molecule 1: 40S Ribosomal Protein SA

Chain A:



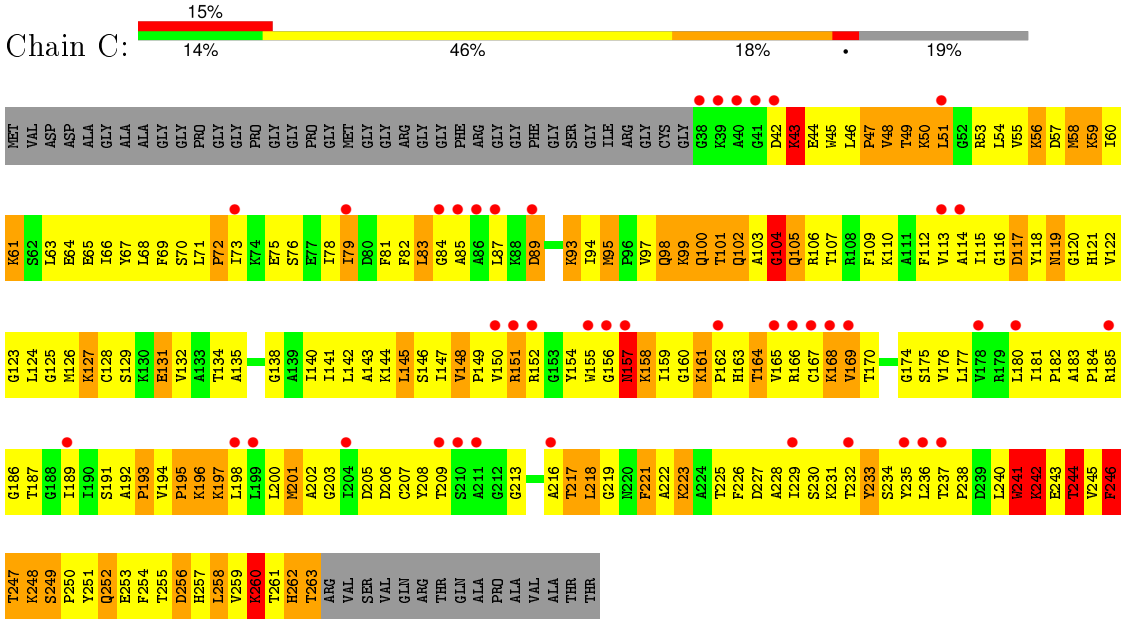
- Molecule 2: 40S Ribosomal Protein S3A

Chain B:



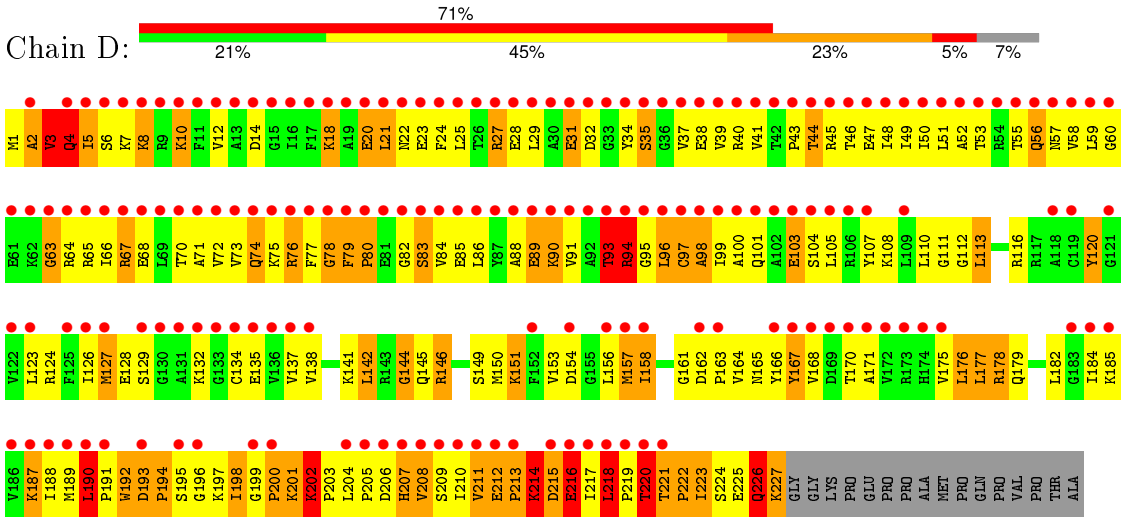
- Molecule 3: 40S Ribosomal Protein S2

Chain C:



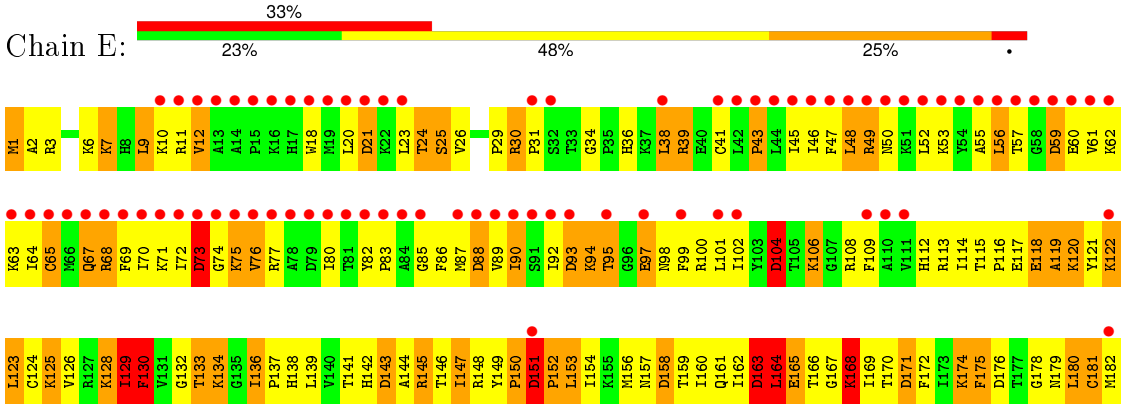
- Molecule 4: 40S Ribosomal Protein S3

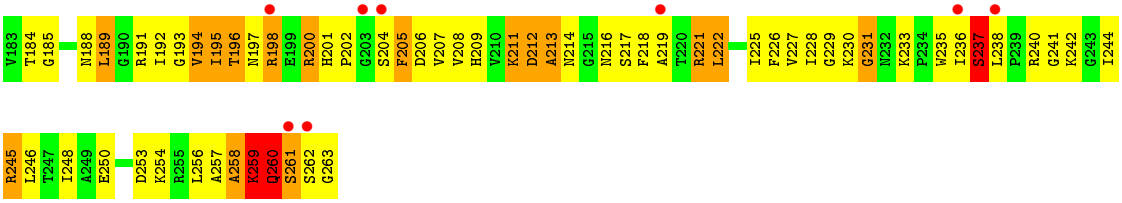
Chain D:



- Molecule 5: 40S Ribosomal Protein S4X

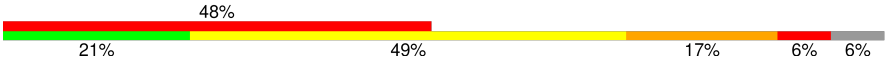
Chain E:

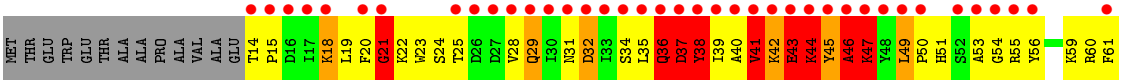


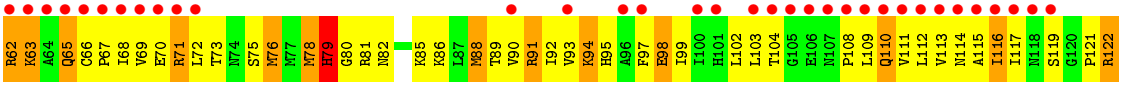


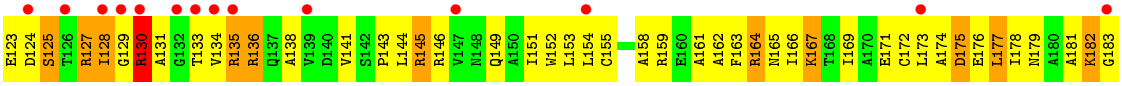
• Molecule 6: 40S Ribosomal Protein S5

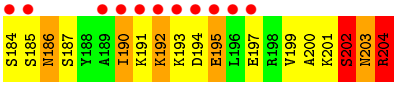
Chain F:





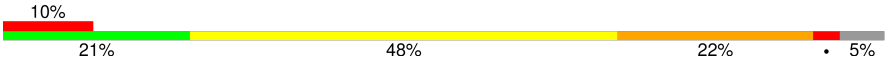


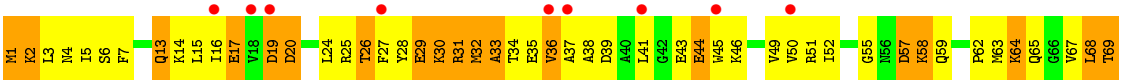




• Molecule 7: 40S Ribosomal Protein S6

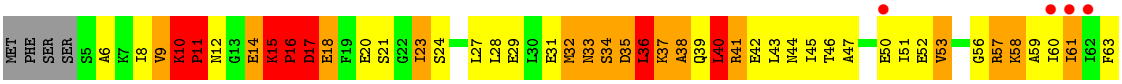
Chain G:

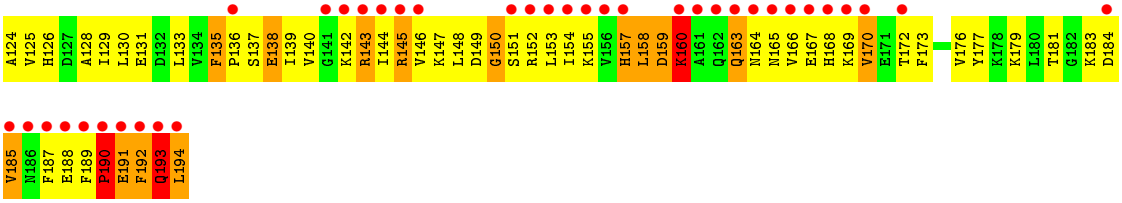




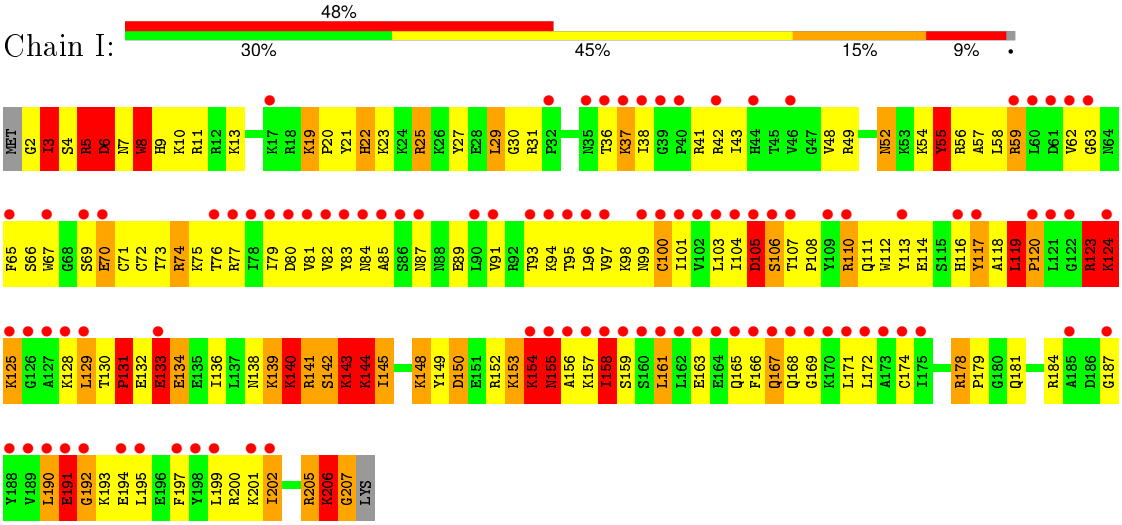
• Molecule 8: 40S Ribosomal Protein S7

Chain H:

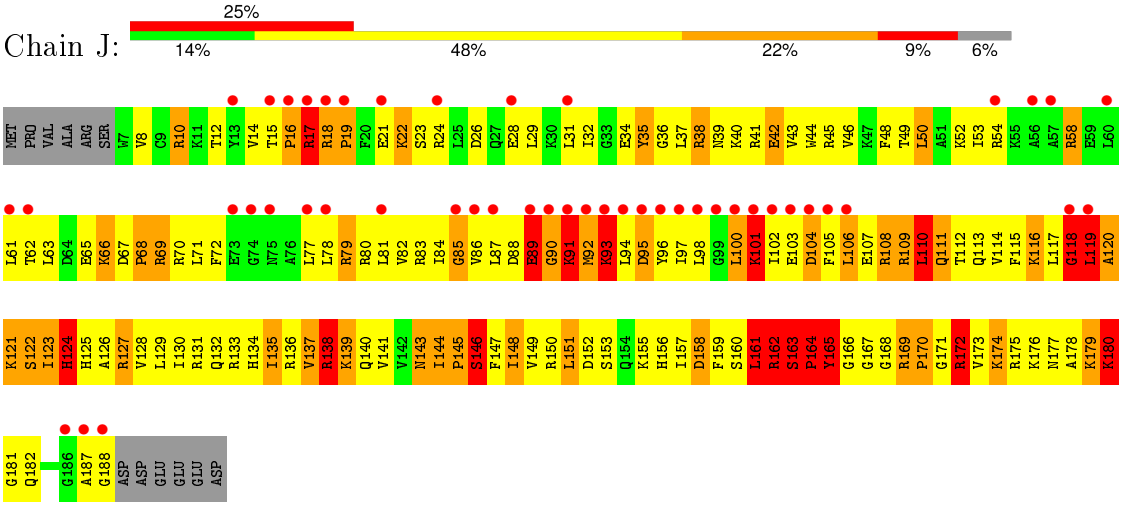




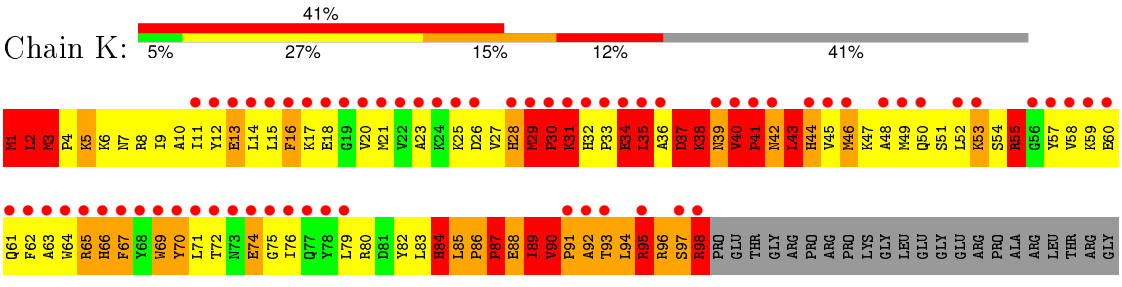
- Molecule 9: 40S Ribosomal Protein S8



- Molecule 10: 40S Ribosomal Protein S9

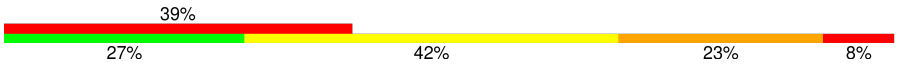


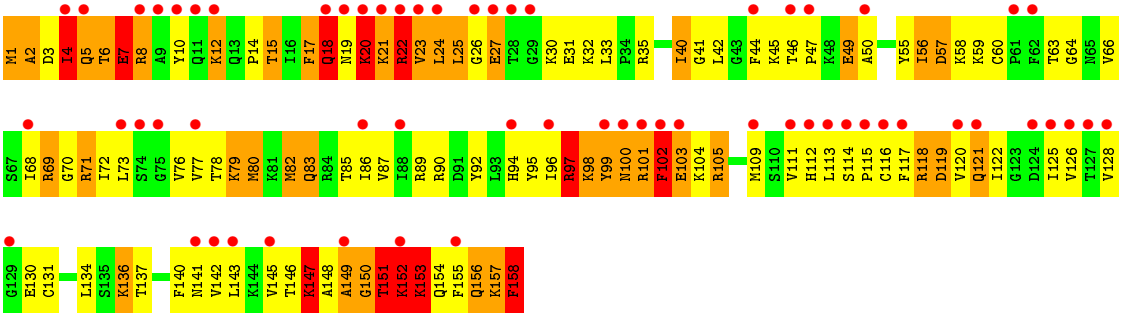
- Molecule 11: 40S Ribosomal Protein S10



- Molecule 12: 40S Ribosomal Protein S11

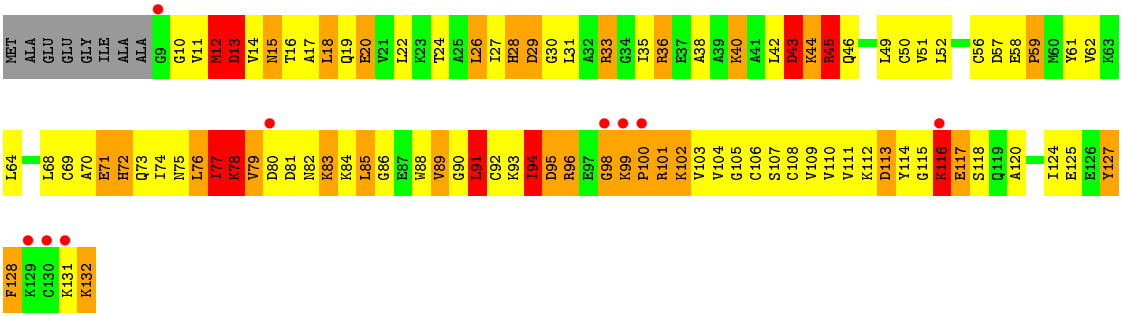
Chain L:





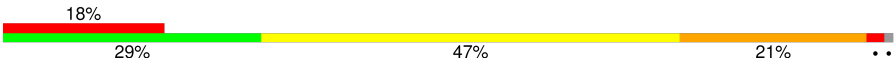
- Molecule 13: 40S Ribosomal Protein S12

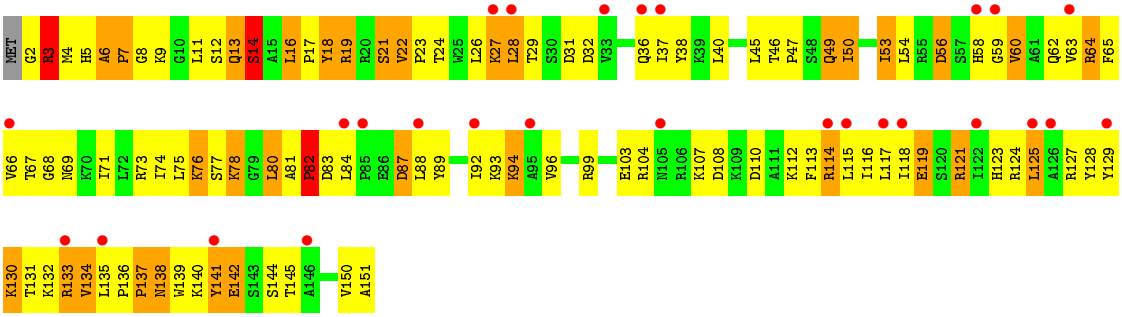
Chain M:



- Molecule 14: 40S Ribosomal Protein S13

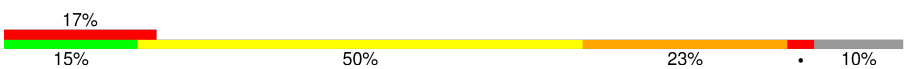
Chain N:

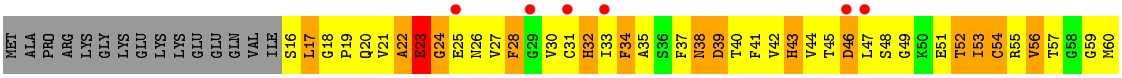


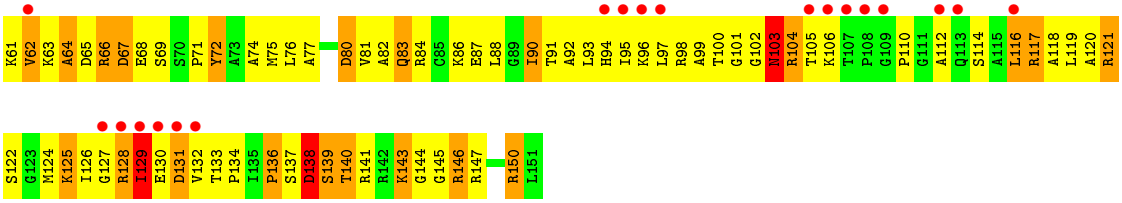


- Molecule 15: 40S Ribosomal Protein S14

Chain O:

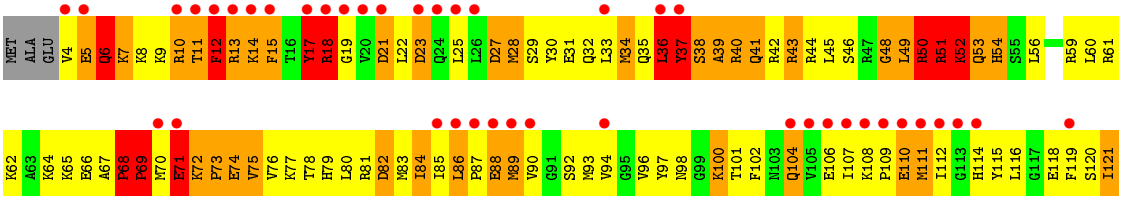


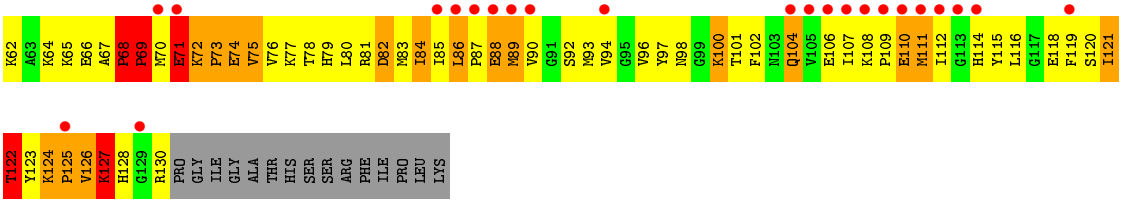




• Molecule 16: 40S Ribosomal Protein S15

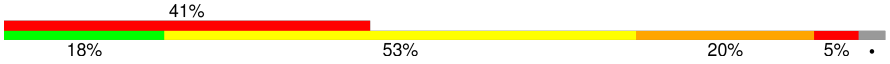
Chain P:

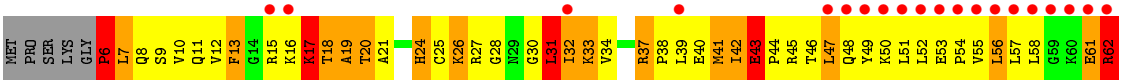


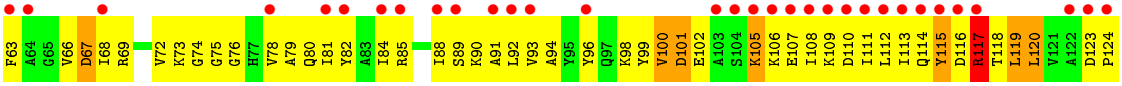


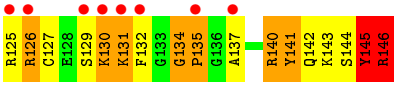
• Molecule 17: 40S Ribosomal Protein S16

Chain Q:



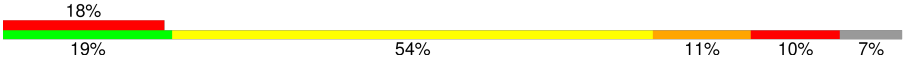


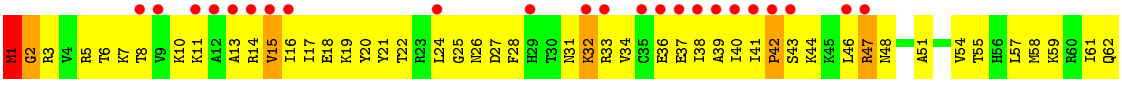


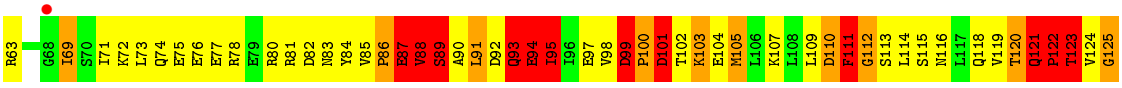


• Molecule 18: 40S Ribosomal Protein S17

Chain R:

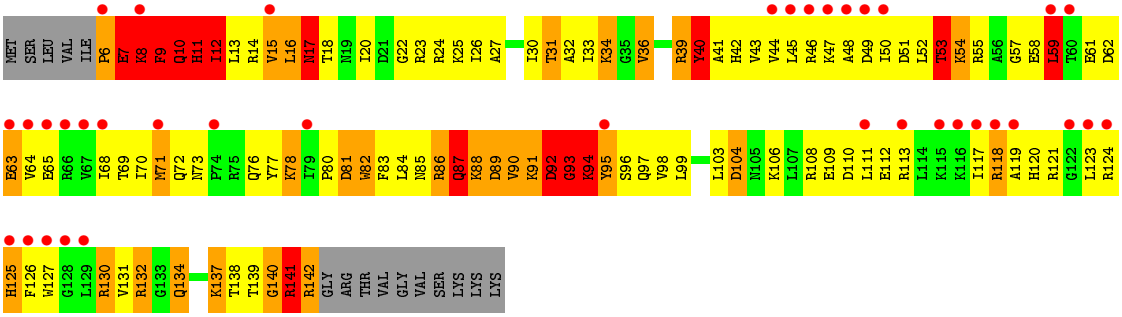






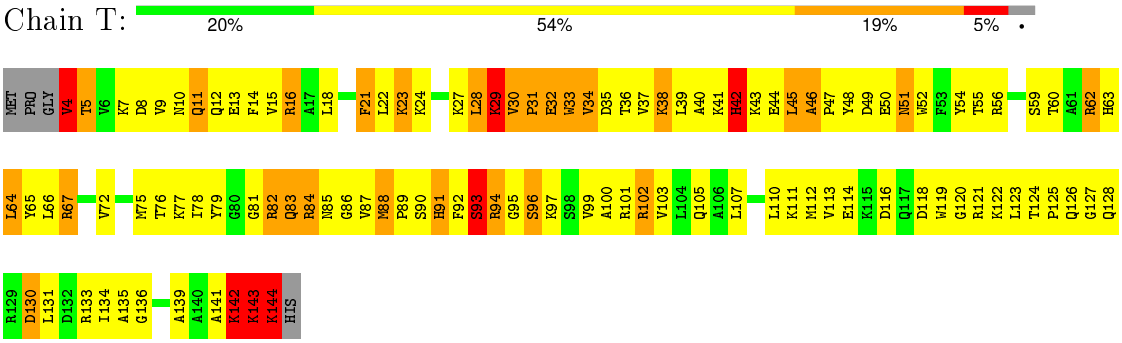
• Molecule 19: 40S Ribosomal Protein S18

Chain S:



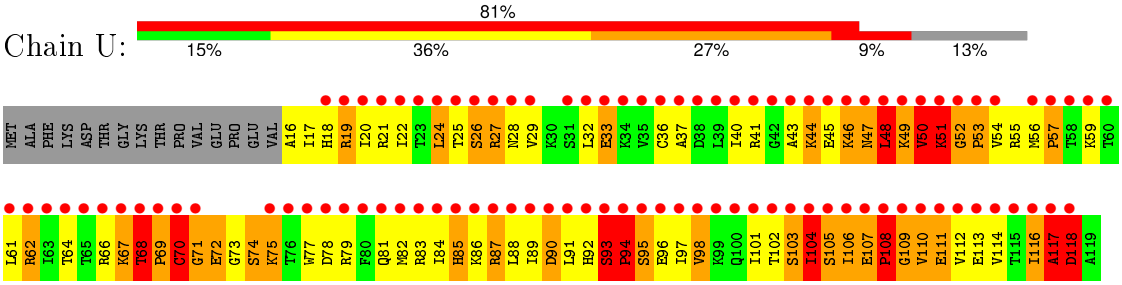
• Molecule 20: 40S Ribosomal Protein S19

Chain T:



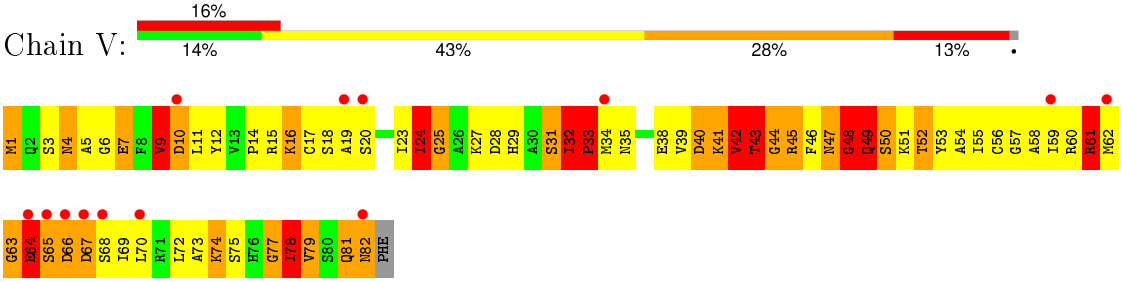
• Molecule 21: 40S Ribosomal Protein S20

Chain U:



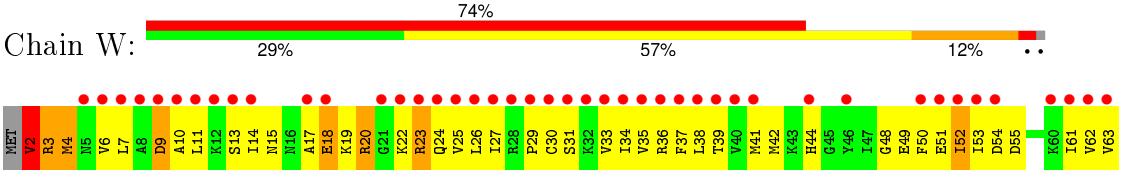
• Molecule 22: 40S Ribosomal Protein S21

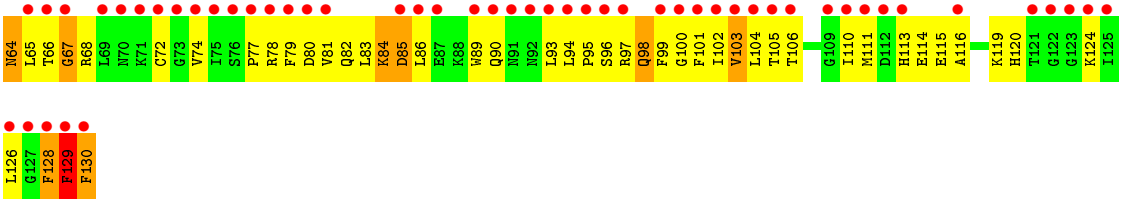
Chain V:



• Molecule 23: 40S Ribosomal Protein S15A

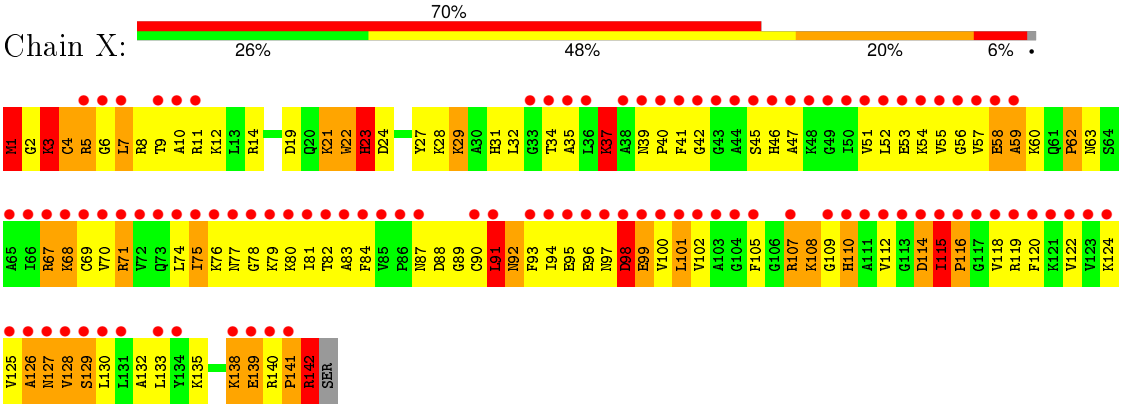
Chain W:





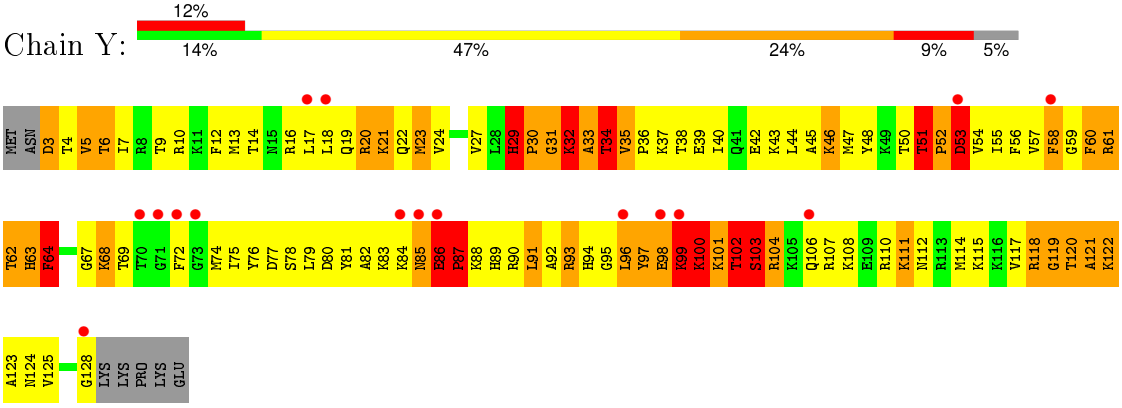
• Molecule 24: 40S Ribosomal Protein S23

Chain X:



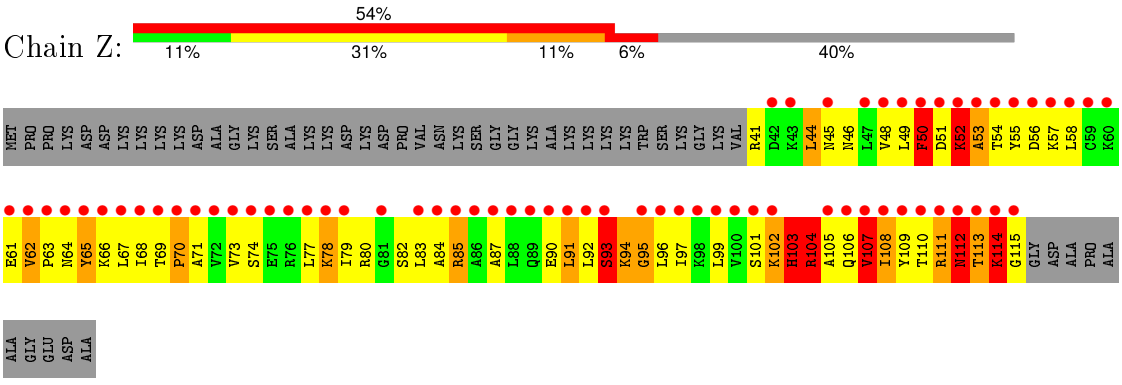
• Molecule 25: 40S Ribosomal Protein S24

Chain Y:



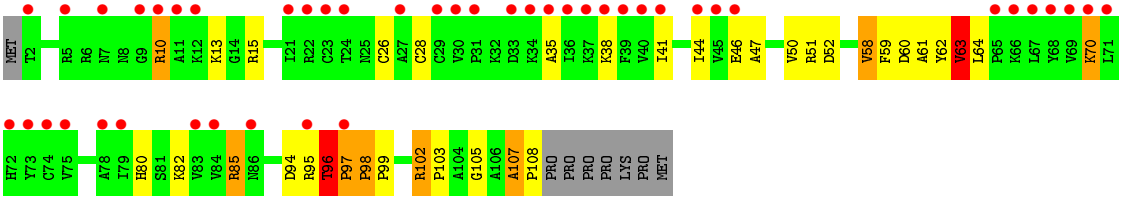
• Molecule 26: 40S Ribosomal Protein S25

Chain Z:

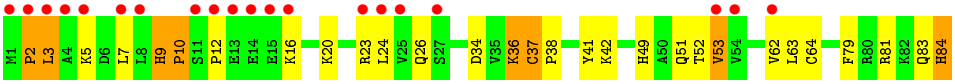


• Molecule 27: 40S Ribosomal Protein S26

Chain a:

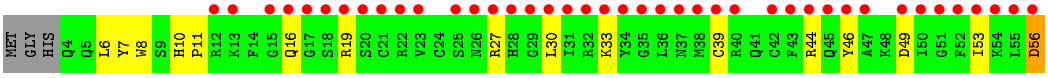


• Molecule 28: 40S Ribosomal Protein S27

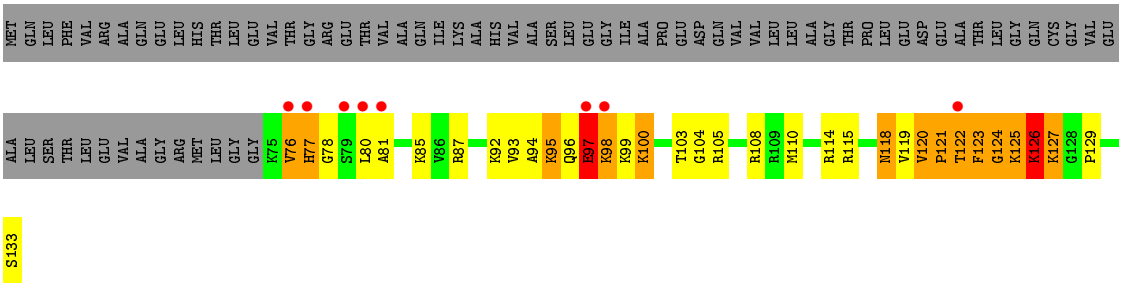


• Molecule 29: 40S Ribosomal Protein S28

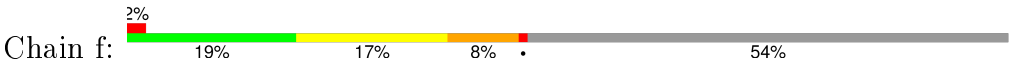
• Molecule 30: 40S Ribosomal Protein S29

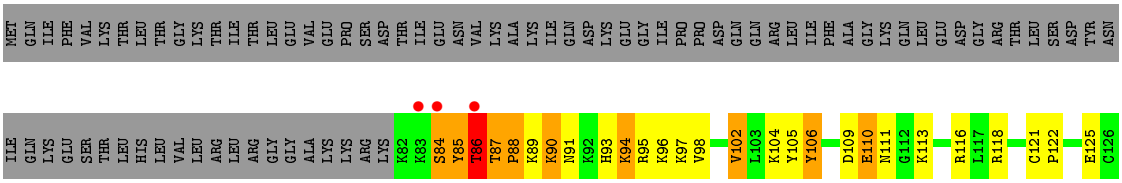


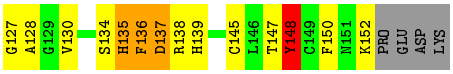
• Molecule 31: 40S Ribosomal Protein S30



• Molecule 32: 40S Ribosomal Protein S27A

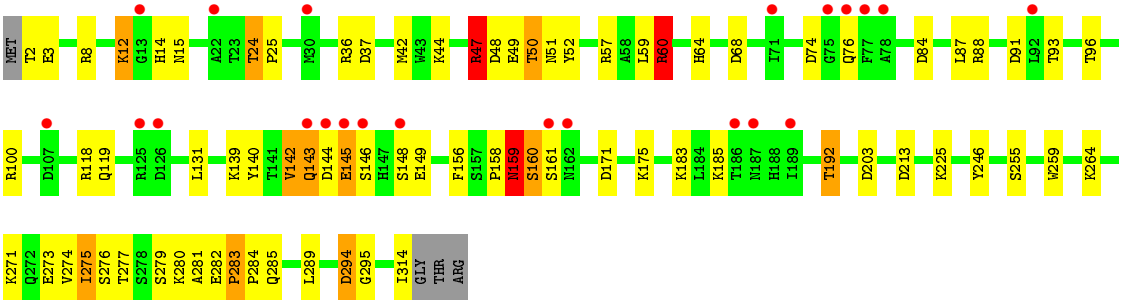






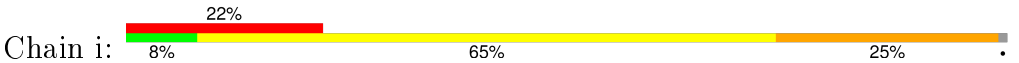
- Molecule 33: 40S Ribosomal Protein RACK1

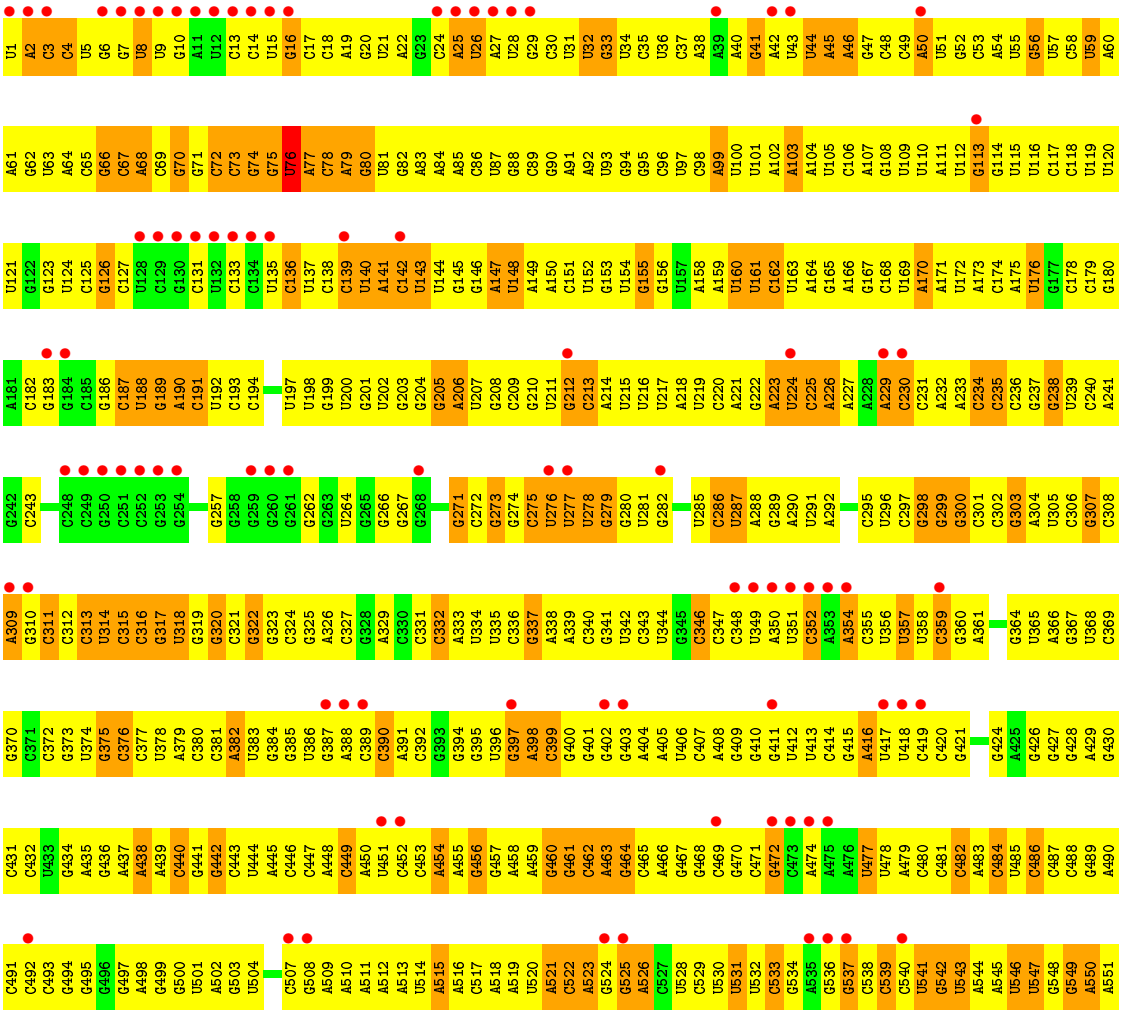
Chain g:

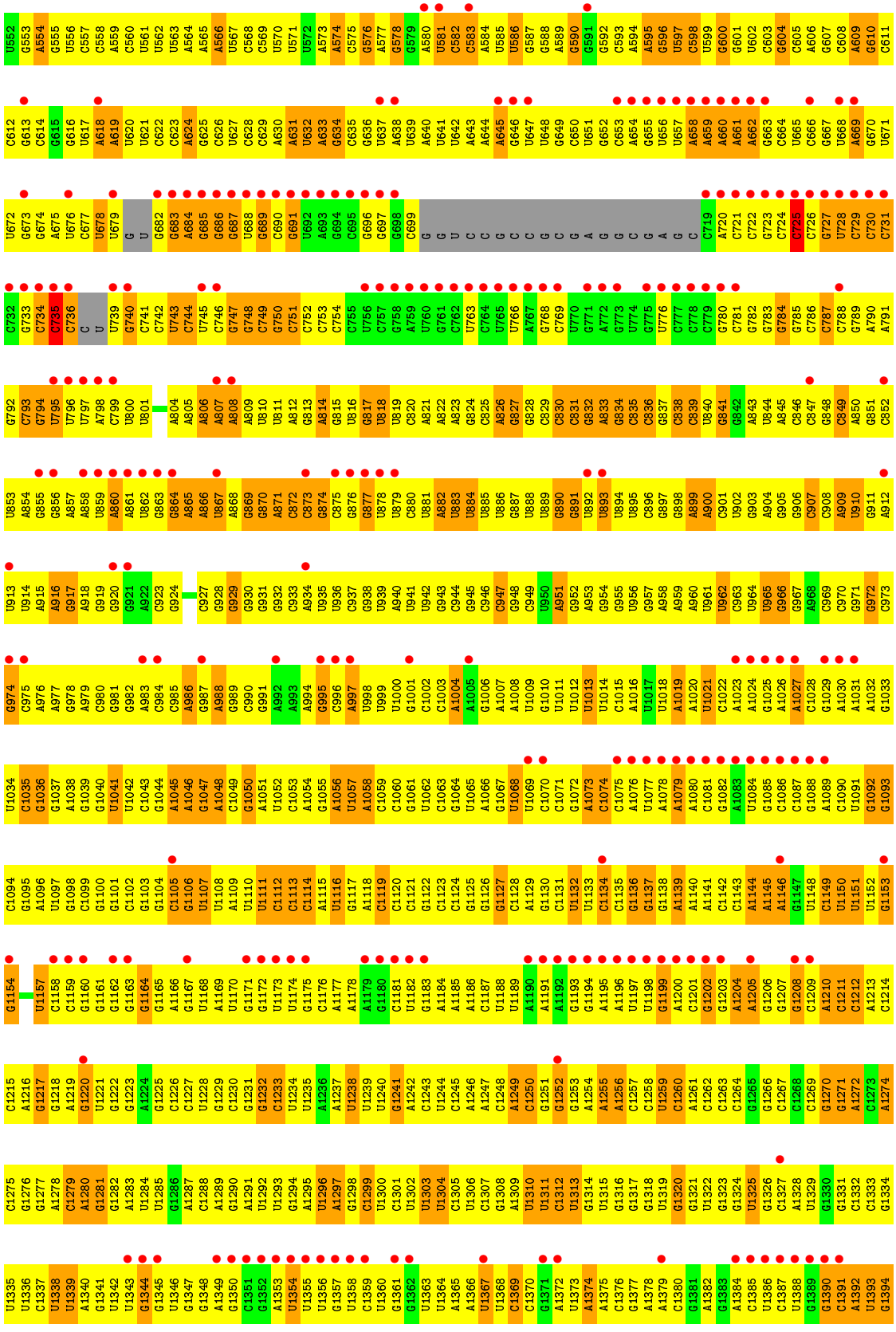


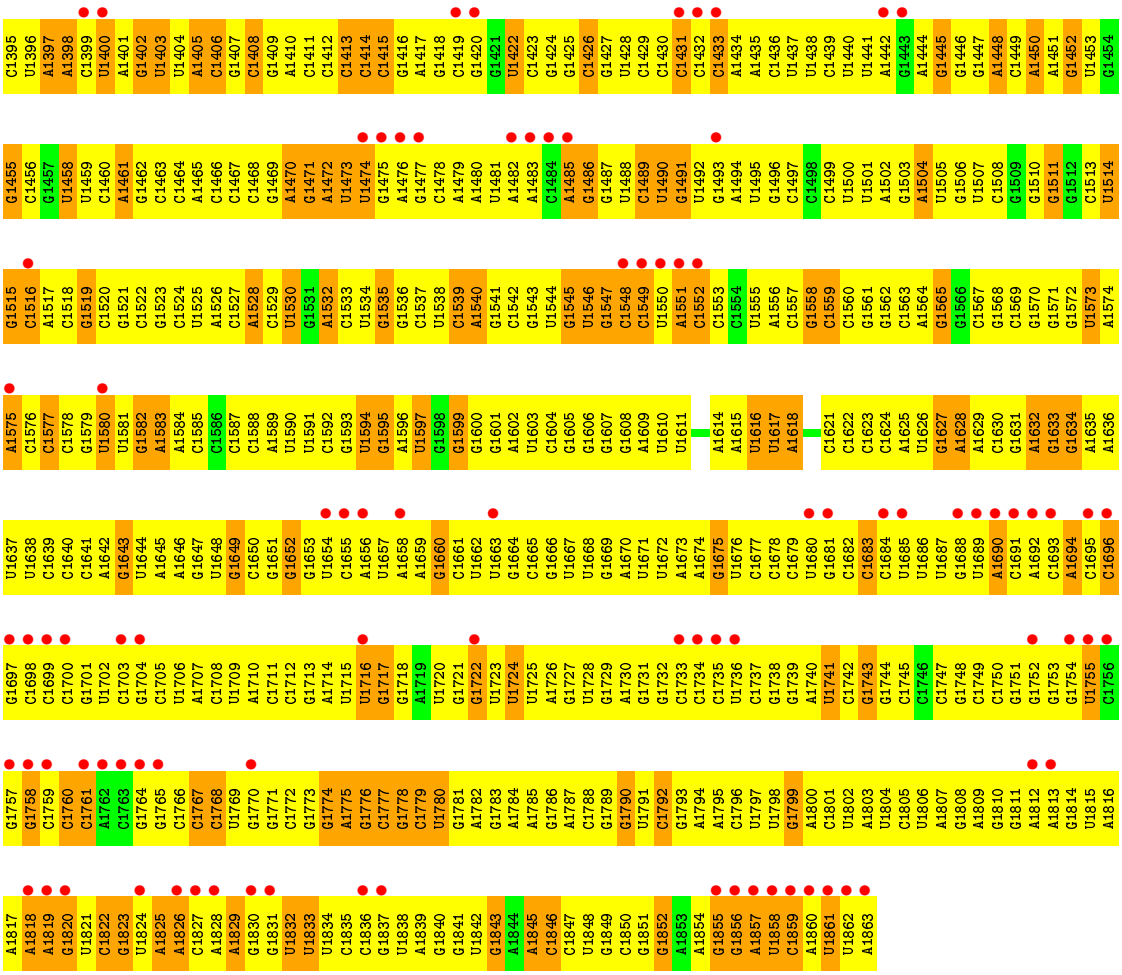
- Molecule 34: 18S Ribosomal RNA

Chain i:



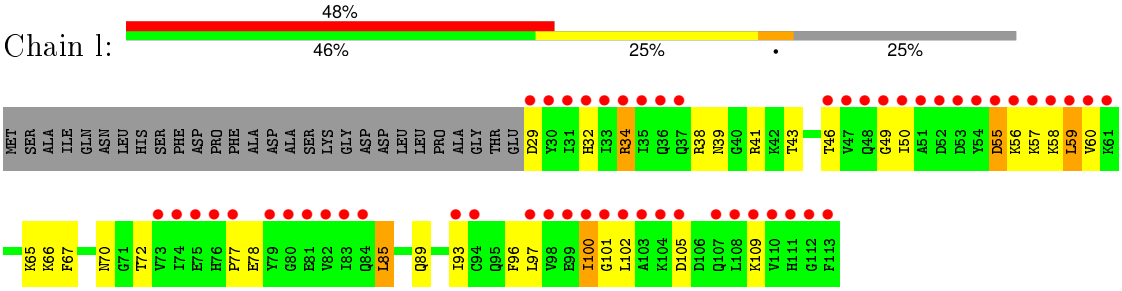






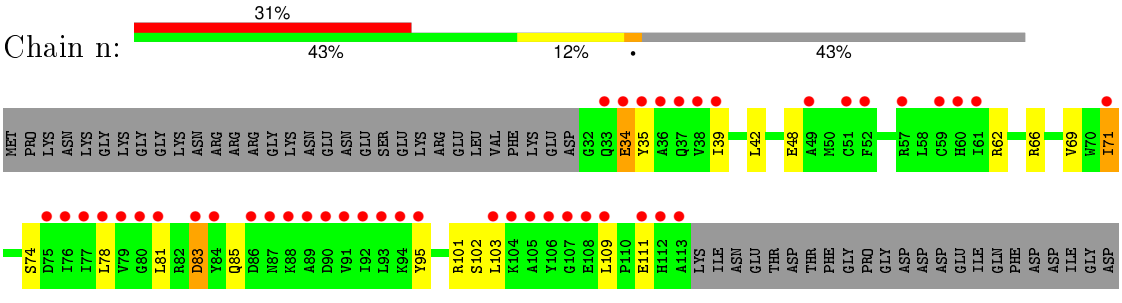
- Molecule 35: human initiation factor eIF1

Chain l:



- Molecule 36: human initiation factor eIF1A

Chain n:



ASP
ASP
GLU
ASP
ILE
ASP
ASP
ILE

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	296.90Å 296.90Å 478.31Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	113.24 – 7.01 113.24 – 7.01	Depositor EDS
% Data completeness (in resolution range)	98.0 (113.24-7.01) 98.1 (113.24-7.01)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.05 (at 6.73Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, R_{free}	0.366 , 0.348 0.347 , 0.340	Depositor DCC
R_{free} test set	1920 reflections (5.03%)	DCC
Wilson B-factor (Å^2)	494.7	Xtriage
Anisotropy	0.354	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.21 , 83.5	EDS
Estimated twinning fraction	0.094 for -h,-k,l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.36$, $< L^2 > = 0.20$	Xtriage
Outliers	1 of 38261 reflections (0.003%)	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	78412	wwPDB-VP
Average B, all atoms (Å^2)	235.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.79% 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

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	0.76	3/1679 (0.2%)	1.05	17/2283 (0.7%)
2	B	0.79	6/1769 (0.3%)	1.08	20/2367 (0.8%)
3	C	0.97	7/1778 (0.4%)	1.19	19/2399 (0.8%)
4	D	1.03	6/1792 (0.3%)	1.30	22/2412 (0.9%)
5	E	0.76	4/2125 (0.2%)	0.98	23/2856 (0.8%)
6	F	0.99	5/1531 (0.3%)	1.21	17/2059 (0.8%)
7	G	0.97	15/1946 (0.8%)	1.23	25/2590 (1.0%)
8	H	1.09	7/1553 (0.5%)	2.20	29/2079 (1.4%)
9	I	1.11	7/1708 (0.4%)	1.51	34/2278 (1.5%)
10	J	1.27	19/1522 (1.2%)	1.51	42/2031 (2.1%)
11	K	1.21	7/851 (0.8%)	1.78	32/1147 (2.8%)
12	L	1.10	6/1319 (0.5%)	1.40	17/1761 (1.0%)
13	M	1.00	3/961 (0.3%)	1.23	7/1288 (0.5%)
14	N	0.83	3/1232 (0.2%)	1.01	13/1656 (0.8%)
15	O	0.61	0/1029	1.05	12/1380 (0.9%)
16	P	0.75	1/1079 (0.1%)	1.43	32/1437 (2.2%)
17	Q	0.70	3/1142 (0.3%)	1.11	15/1528 (1.0%)
18	R	1.23	10/1031 (1.0%)	1.64	30/1383 (2.2%)
19	S	1.21	11/1157 (1.0%)	1.61	36/1548 (2.3%)
20	T	0.95	3/1132 (0.3%)	1.26	13/1517 (0.9%)
21	U	0.96	1/832 (0.1%)	1.59	29/1117 (2.6%)
22	V	0.75	1/626 (0.2%)	1.40	15/839 (1.8%)
23	W	0.85	4/1051 (0.4%)	0.86	9/1406 (0.6%)
24	X	0.99	9/1124 (0.8%)	1.25	21/1500 (1.4%)
25	Y	0.93	3/1038 (0.3%)	1.42	22/1380 (1.6%)
26	Z	1.04	6/604 (1.0%)	1.35	16/810 (2.0%)
27	a	0.89	5/860 (0.6%)	1.60	21/1156 (1.8%)
28	b	1.03	2/673 (0.3%)	1.36	12/902 (1.3%)
29	c	0.80	1/508 (0.2%)	1.17	8/680 (1.2%)
30	d	0.90	2/455 (0.4%)	0.79	3/603 (0.5%)
31	e	1.48	5/472 (1.1%)	1.43	11/620 (1.8%)
32	f	1.10	4/593 (0.7%)	1.49	14/786 (1.8%)
33	g	0.92	1/2493 (0.0%)	1.29	25/3394 (0.7%)
34	i	2.41	1879/42474 (4.4%)	2.22	2609/66043 (4.0%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
35	l	1.18	5/700 (0.7%)	1.29	8/933 (0.9%)
36	n	0.40	0/657	0.38	0/881
All	All	1.85	2054/83496 (2.5%)	1.87	3278/121049 (2.7%)

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	0	11
2	B	0	4
3	C	1	5
4	D	0	5
5	E	1	2
6	F	0	3
7	G	0	1
8	H	0	10
9	I	0	8
10	J	1	11
11	K	0	11
12	L	0	7
13	M	0	1
14	N	0	4
15	O	0	1
16	P	0	10
17	Q	0	4
18	R	1	5
19	S	1	10
20	T	1	6
21	U	0	8
22	V	0	9
23	W	0	2
24	X	0	4
25	Y	1	6
26	Z	0	6
27	a	0	2
28	b	0	3
31	e	0	5
32	f	0	6
33	g	0	13
34	i	6	0
All	All	13	183

The worst 5 of 2054 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
34	i	1322	U	C2'-C1'	-25.25	1.25	1.53
34	i	66	G	C2'-C1'	-24.54	1.26	1.53
34	i	858	A	C2'-C1'	-23.80	1.27	1.53
34	i	652	G	C2'-C1'	-23.70	1.27	1.53
34	i	1307	C	C2'-C1'	-22.32	1.28	1.53

The worst 5 of 3278 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
8	H	109	ARG	NE-CZ-NH2	-53.45	93.58	120.30
8	H	109	ARG	NE-CZ-NH1	42.64	141.62	120.30
34	i	1774	G	P-O3'-C3'	38.29	165.65	119.70
34	i	1114	C	O4'-C1'-N1	35.27	136.41	108.20
34	i	582	C	O4'-C1'-N1	32.57	134.25	108.20

5 of 13 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	C	157	ASN	CA
5	E	171	ASP	CA
10	J	138	ARG	CA
18	R	3	ARG	CA
19	S	92	ASP	CA

5 of 183 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	146	ALA	Mainchain
1	A	23	THR	Mainchain
1	A	4	ALA	Peptide
1	A	63	ARG	Sidechain
1	A	97	THR	Mainchain

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	1642	0	1640	607	0
2	B	1741	0	1809	557	1
3	C	1742	0	1830	590	0
4	D	1764	0	1857	600	0
5	E	2083	0	2189	525	0
6	F	1509	0	1557	475	0
7	G	1923	0	2084	503	10
8	H	1530	0	1624	477	0
9	I	1679	0	1762	432	4
10	J	1498	0	1598	540	0
11	K	827	0	853	353	18
12	L	1296	0	1370	397	0
13	M	951	0	972	247	0
14	N	1208	0	1294	258	0
15	O	1016	0	1036	285	0
16	P	1060	0	1120	486	1
17	Q	1124	0	1193	445	0
18	R	1019	0	1070	394	0
19	S	1139	0	1187	434	10
20	T	1112	0	1149	392	0
21	U	822	0	886	209	0
22	V	619	0	620	284	0
23	W	1034	0	1079	259	0
24	X	1106	0	1177	309	0
25	Y	1021	0	1083	491	0
26	Z	598	0	652	208	0
27	a	844	0	895	0	0
28	b	659	0	680	0	0
29	c	506	0	536	0	0
30	d	445	0	441	0	0
31	e	468	0	514	0	46
32	f	581	0	598	0	0
33	g	2436	0	2388	0	0
34	i	38071	0	19021	0	110
35	l	691	0	702	0	1
36	n	648	0	654	0	0
All	All	78412	0	61120	9868	121

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 97.

The worst 5 of 9868 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)
17:Q:135:PRO:HD3	17:Q:141:TYR:CE1	1.16	1.68
3:C:50:LYS:HD2	3:C:251:TYR:CE1	1.20	1.67
19:S:42:HIS:CD2	20:T:45:LEU:HD11	1.21	1.67
4:D:132:LYS:CB	4:D:191:PRO:HG3	1.23	1.66
3:C:197:LYS:HA	3:C:200:LEU:CD2	1.22	1.66

The worst 5 of 121 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
11:K:97:SER:O	34:i:76:U:OP1[3_454]	0.42	1.78
31:e:125:LYS:CD	34:i:1761:C:P[3_454]	0.61	1.59
34:i:137:U:OP2	34:i:532:U:N3[2_565]	0.64	1.56
31:e:125:LYS:CE	34:i:1761:C:O5'[3_454]	0.72	1.48
31:e:125:LYS:CG	34:i:1761:C:OP2[3_454]	0.77	1.43

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	206/295 (70%)	156 (76%)	23 (11%)	27 (13%)	0 7
2	B	213/264 (81%)	174 (82%)	24 (11%)	15 (7%)	1 22
3	C	224/278 (81%)	200 (89%)	13 (6%)	11 (5%)	3 31
4	D	225/243 (93%)	180 (80%)	23 (10%)	22 (10%)	1 14
5	E	261/263 (99%)	210 (80%)	27 (10%)	24 (9%)	1 16
6	F	189/204 (93%)	162 (86%)	15 (8%)	12 (6%)	2 25
7	G	235/249 (94%)	202 (86%)	18 (8%)	15 (6%)	2 25
8	H	188/194 (97%)	146 (78%)	11 (6%)	31 (16%)	0 5
9	I	204/208 (98%)	169 (83%)	13 (6%)	22 (11%)	0 11
10	J	180/194 (93%)	138 (77%)	18 (10%)	24 (13%)	0 7

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
11	K	96/165 (58%)	67 (70%)	11 (12%)	18 (19%)	0 3
12	L	156/158 (99%)	132 (85%)	10 (6%)	14 (9%)	1 17
13	M	122/132 (92%)	85 (70%)	16 (13%)	21 (17%)	0 4
14	N	148/151 (98%)	124 (84%)	18 (12%)	6 (4%)	3 35
15	O	134/151 (89%)	101 (75%)	14 (10%)	19 (14%)	0 6
16	P	125/145 (86%)	92 (74%)	16 (13%)	17 (14%)	0 6
17	Q	139/146 (95%)	109 (78%)	20 (14%)	10 (7%)	1 22
18	R	124/135 (92%)	96 (77%)	14 (11%)	14 (11%)	0 10
19	S	135/152 (89%)	107 (79%)	19 (14%)	9 (7%)	1 24
20	T	139/145 (96%)	119 (86%)	10 (7%)	10 (7%)	1 22
21	U	102/119 (86%)	76 (74%)	10 (10%)	16 (16%)	0 5
22	V	80/83 (96%)	55 (69%)	11 (14%)	14 (18%)	0 4
23	W	127/130 (98%)	111 (87%)	14 (11%)	2 (2%)	12 56
24	X	140/143 (98%)	121 (86%)	11 (8%)	8 (6%)	2 27
25	Y	124/133 (93%)	91 (73%)	15 (12%)	18 (14%)	0 6
26	Z	73/125 (58%)	52 (71%)	12 (16%)	9 (12%)	0 8
27	a	105/115 (91%)	72 (69%)	14 (13%)	19 (18%)	0 4
28	b	82/84 (98%)	57 (70%)	14 (17%)	11 (13%)	0 7
29	c	62/69 (90%)	44 (71%)	13 (21%)	5 (8%)	1 19
30	d	51/56 (91%)	46 (90%)	3 (6%)	2 (4%)	4 36
31	e	57/133 (43%)	37 (65%)	7 (12%)	13 (23%)	0 2
32	f	69/156 (44%)	38 (55%)	13 (19%)	18 (26%)	0 1
33	g	311/317 (98%)	271 (87%)	23 (7%)	17 (6%)	2 29
35	l	82/113 (73%)	49 (60%)	22 (27%)	11 (13%)	0 7
36	n	80/144 (56%)	61 (76%)	15 (19%)	4 (5%)	3 31
All	All	4988/5792 (86%)	3950 (79%)	530 (11%)	508 (10%)	1 13

5 of 508 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	9	GLN
1	A	31	ASP
1	A	45	GLY

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	103	PHE
1	A	164	ASN

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	174/244 (71%)	140 (80%)	34 (20%)	2 12
2	B	196/231 (85%)	155 (79%)	41 (21%)	1 10
3	C	187/215 (87%)	147 (79%)	40 (21%)	1 9
4	D	190/202 (94%)	144 (76%)	46 (24%)	1 6
5	E	225/225 (100%)	173 (77%)	52 (23%)	1 7
6	F	161/170 (95%)	117 (73%)	44 (27%)	0 4
7	G	207/218 (95%)	158 (76%)	49 (24%)	1 7
8	H	170/174 (98%)	124 (73%)	46 (27%)	0 4
9	I	177/179 (99%)	142 (80%)	35 (20%)	1 12
10	J	157/168 (94%)	128 (82%)	29 (18%)	2 14
11	K	89/136 (65%)	61 (68%)	28 (32%)	0 2
12	L	142/142 (100%)	105 (74%)	37 (26%)	0 5
13	M	102/108 (94%)	79 (78%)	23 (22%)	1 8
14	N	130/131 (99%)	103 (79%)	27 (21%)	1 10
15	O	106/119 (89%)	87 (82%)	19 (18%)	2 15
16	P	116/130 (89%)	84 (72%)	32 (28%)	0 4
17	Q	117/121 (97%)	89 (76%)	28 (24%)	1 7
18	R	114/121 (94%)	90 (79%)	24 (21%)	1 9
19	S	119/132 (90%)	95 (80%)	24 (20%)	1 11
20	T	113/116 (97%)	87 (77%)	26 (23%)	1 7
21	U	94/107 (88%)	74 (79%)	20 (21%)	1 9
22	V	67/68 (98%)	50 (75%)	17 (25%)	1 6

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
23	W	112/113 (99%)	98 (88%)	14 (12%)	6 30
24	X	114/115 (99%)	91 (80%)	23 (20%)	1 11
25	Y	108/115 (94%)	85 (79%)	23 (21%)	1 9
26	Z	66/103 (64%)	53 (80%)	13 (20%)	1 12
27	a	91/99 (92%)	76 (84%)	15 (16%)	3 19
28	b	76/76 (100%)	63 (83%)	13 (17%)	2 17
29	c	57/62 (92%)	46 (81%)	11 (19%)	2 13
30	d	47/49 (96%)	35 (74%)	12 (26%)	1 6
31	e	48/106 (45%)	25 (52%)	23 (48%)	0 0
32	f	64/140 (46%)	43 (67%)	21 (33%)	0 2
33	g	272/275 (99%)	223 (82%)	49 (18%)	2 15
35	l	74/96 (77%)	56 (76%)	18 (24%)	1 6
36	n	66/123 (54%)	47 (71%)	19 (29%)	0 3
All	All	4348/4929 (88%)	3373 (78%)	975 (22%)	1 9

5 of 975 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
11	K	95	ARG
15	O	129	ILE
33	g	64	HIS
12	L	30	LYS
13	M	78	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 116 such sidechains are listed below:

Mol	Chain	Res	Type
13	M	19	GLN
16	P	103	ASN
33	g	162	ASN
13	M	75	ASN
14	N	101	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
34	i	1735/1863 (93%)	503 (28%)	0

5 of 503 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
34	i	2	A
34	i	3	C
34	i	4	C
34	i	8	U
34	i	16	G

There are no RNA pucker outliers to report.

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

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

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

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	A	208/295 (70%)	0.28	20 (9%) 10 14	209, 304, 369, 385	0
2	B	215/264 (81%)	2.82	111 (51%) 0 4	170, 238, 290, 301	0
3	C	226/278 (81%)	0.95	43 (19%) 2 7	113, 192, 306, 337	0
4	D	227/243 (93%)	5.83	172 (75%) 0 2	186, 248, 328, 357	0
5	E	263/263 (100%)	2.06	88 (33%) 0 5	98, 185, 238, 251	0
6	F	191/204 (93%)	2.33	98 (51%) 0 4	228, 288, 316, 326	0
7	G	237/249 (95%)	0.34	24 (10%) 9 14	126, 225, 339, 361	0
8	H	190/194 (97%)	1.03	46 (24%) 1 6	177, 322, 371, 383	0
9	I	206/208 (99%)	2.91	99 (48%) 0 4	80, 224, 300, 314	0
10	J	182/194 (93%)	1.27	48 (26%) 1 5	119, 186, 238, 281	0
11	K	98/165 (59%)	5.02	67 (68%) 0 3	256, 329, 366, 374	0
12	L	158/158 (100%)	2.00	61 (38%) 0 4	89, 163, 283, 295	0
13	M	124/132 (93%)	0.26	9 (7%) 18 21	295, 428, 439, 441	0
14	N	150/151 (99%)	0.85	27 (18%) 2 7	111, 167, 276, 299	0
15	O	136/151 (90%)	1.00	25 (18%) 2 7	119, 235, 306, 339	0
16	P	127/145 (87%)	1.25	43 (33%) 0 4	274, 351, 387, 394	0
17	Q	141/146 (96%)	2.04	60 (42%) 0 4	198, 305, 333, 340	0
18	R	126/135 (93%)	0.80	24 (19%) 2 7	208, 271, 378, 382	0
19	S	137/152 (90%)	1.17	37 (27%) 1 5	253, 328, 349, 358	0
20	T	141/145 (97%)	-0.01	0 100 100	273, 331, 358, 363	0
21	U	104/119 (87%)	7.31	96 (92%) 0 1	197, 304, 340, 357	0
22	V	82/83 (98%)	0.51	13 (15%) 3 8	196, 246, 356, 365	0
23	W	129/130 (99%)	4.38	96 (74%) 0 2	116, 169, 214, 229	0
24	X	142/143 (99%)	4.78	100 (70%) 0 3	74, 101, 124, 133	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
25	Y	126/133 (94%)	0.65	16 (12%) 5 11	134, 185, 219, 236	0
26	Z	75/125 (60%)	4.28	67 (89%) 0 2	295, 323, 347, 353	0
27	a	107/115 (93%)	2.43	45 (42%) 0 4	115, 166, 278, 294	0
28	b	84/84 (100%)	1.27	20 (23%) 1 6	186, 244, 315, 338	0
29	c	64/69 (92%)	1.74	24 (37%) 0 4	210, 264, 310, 317	0
30	d	53/56 (94%)	5.46	41 (77%) 0 2	215, 248, 337, 357	0
31	e	59/133 (44%)	0.30	8 (13%) 4 10	112, 169, 214, 226	0
32	f	71/156 (45%)	-0.59	3 (4%) 40 38	243, 417, 429, 432	0
33	g	313/317 (98%)	0.34	22 (7%) 19 22	282, 330, 361, 377	0
34	i	1840/1863 (98%)	1.26	410 (22%) 1 6	70, 205, 399, 456	0
35	l	85/113 (75%)	3.16	54 (63%) 0 3	270, 272, 274, 274	0
36	n	82/144 (56%)	2.62	44 (53%) 0 3	257, 260, 262, 263	0
All	All	6899/7655 (90%)	1.77	2161 (31%) 1 5	70, 246, 379, 456	0

The worst 5 of 2161 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
34	i	721	C	33.5
34	i	722	C	31.6
34	i	250	G	29.7
34	i	720	A	29.3
21	U	36	CYS	26.7

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

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

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.