wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 06:09 PM GMT

PDB ID	:	4KHP
Title	:	Structure of the Thermus thermophilus 30S ribosomal subunit in complex with de-6-MSA-pactamycin
Authors	:	Tourigny, D.S.; Fernandez, I.S.; Kelley, A.C.; Ramakrishnan, V.
Deposited on	:	2013-05-01
Resolution	:	3.10 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:

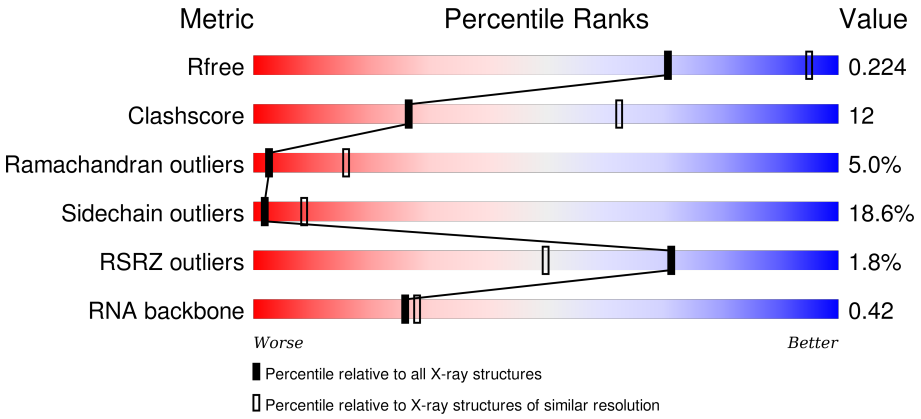
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 3.10 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.



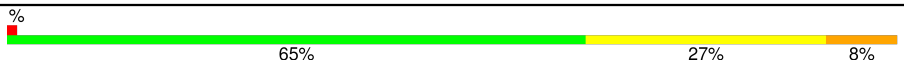
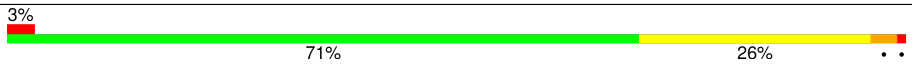
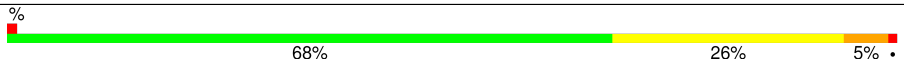
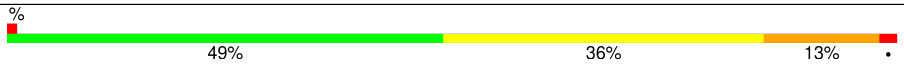
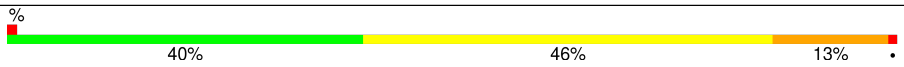
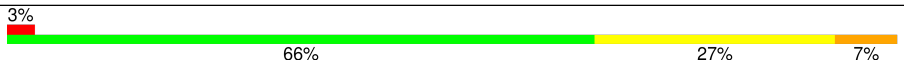
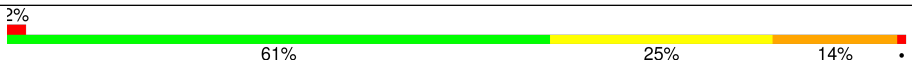
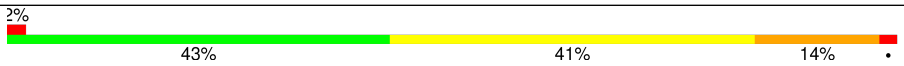
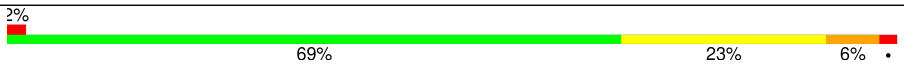
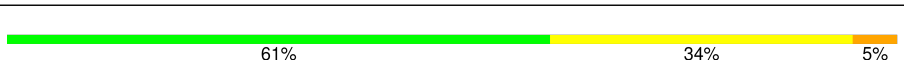
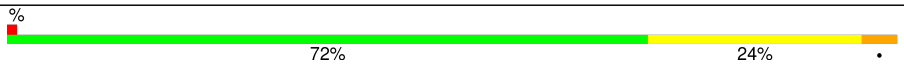
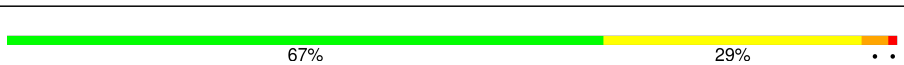
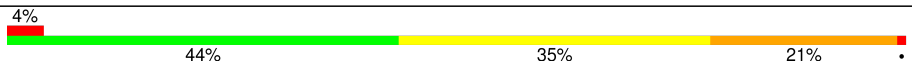
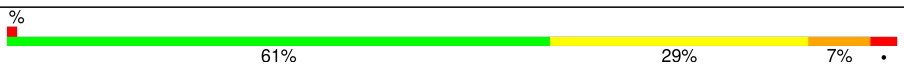
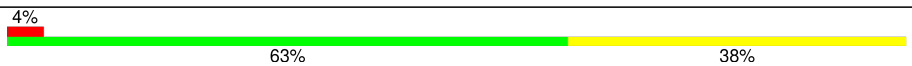
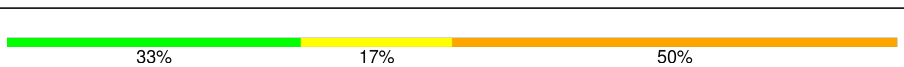
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1114 (3.14-3.06)
Clashscore	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)
RSRZ outliers	91569	1119 (3.14-3.06)
RNA backbone	2183	1010 (3.52-2.68)

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	1506	2% 51% 35% 12% •
2	B	234	3% 47% 39% 12% •
3	C	206	62% 31% 5% •
4	D	208	63% 28% 7% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	150	
6	F	101	
7	G	155	
8	H	138	
9	I	127	
10	J	98	
11	K	119	
12	L	125	
13	M	120	
14	N	60	
15	O	88	
16	P	83	
17	Q	99	
18	R	70	
19	S	78	
20	T	99	
21	U	24	
22	X	6	

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
23	PAR	A	1602	-	-	-	X
23	PAR	A	1603	-	-	-	X
23	PAR	A	1605	-	-	-	X
23	PAR	A	1606	-	-	-	X
23	PAR	A	1607	-	-	-	X
25	MG	A	1609	-	-	-	X
25	MG	A	1610	-	-	-	X

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
25	MG	A	1614	-	-	-	X
25	MG	A	1617	-	-	-	X
25	MG	A	1620	-	-	-	X
25	MG	A	1622	-	-	-	X
25	MG	A	1632	-	-	-	X
25	MG	A	1633	-	-	-	X
25	MG	A	1634	-	-	-	X
25	MG	A	1636	-	-	-	X
25	MG	A	1639	-	-	-	X
25	MG	A	1645	-	-	-	X
25	MG	A	1647	-	-	-	X
25	MG	A	1651	-	-	-	X
25	MG	A	1652	-	-	-	X
25	MG	A	1654	-	-	-	X
25	MG	A	1659	-	-	-	X
25	MG	A	1661	-	-	-	X
25	MG	A	1667	-	-	-	X
25	MG	A	1682	-	-	-	X
25	MG	A	1685	-	-	-	X
25	MG	A	1703	-	-	-	X
25	MG	D	301	-	-	-	X
25	MG	E	201	-	-	-	X

2 Entry composition

There are 25 unique types of molecules in this entry. The entry contains 52032 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 16S Ribosomal RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1506	Total	C	N	O	P	0	0	0
			32368	14408	5997	10458	1505			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	79	A	G	CONFLICT	GB 55771382

- Molecule 2 is a protein called 30S Ribosomal protein S2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	234	Total	C	N	O	S	0	0	0
			1900	1213	341	341	5			

- Molecule 3 is a protein called 30S Ribosomal protein S3.

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

- Molecule 4 is a protein called 30S Ribosomal protein S4.

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

- Molecule 5 is a protein called 30S Ribosomal protein S5.

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

- Molecule 6 is a protein called 30S Ribosomal protein S6.

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

- Molecule 7 is a protein called 30S Ribosomal protein S7.

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

- Molecule 8 is a protein called 30S Ribosomal protein S8.

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

- Molecule 9 is a protein called 30S Ribosomal protein S9.

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

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	58	ARG	HIS	CONFLICT	UNP P80374

- Molecule 10 is a protein called 30S Ribosomal protein S10.

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

- Molecule 11 is a protein called 30S Ribosomal protein S11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	119	Total	C	N	O	S	0	0	0
			885	549	168	165	3			

- Molecule 12 is a protein called 30S Ribosomal protein S12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	125	Total	C	N	O	S	0	0	0
			975	614	196	164	1			

- Molecule 13 is a protein called 30S Ribosomal protein S13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	120	Total	C	N	O	S	0	0	0
			955	591	197	165	2			

- Molecule 14 is a protein called 30S Ribosomal protein S14.

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

- Molecule 15 is a protein called 30S Ribosomal protein S15.

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

- Molecule 16 is a protein called 30S Ribosomal protein S16.

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

- Molecule 17 is a protein called 30S Ribosomal protein S17.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
17	Q	99	Total	C	N	O	S	0	0	0
			823	528	151	142	2			

- Molecule 18 is a protein called 30S Ribosomal protein S18.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
18	R	70	Total	C	N	O	0	0	0
			574	367	112	95			

- Molecule 19 is a protein called 30S Ribosomal protein S19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
19	S	78	Total	C	N	O	S	0	0	0
			629	403	114	110	2			

- Molecule 20 is a protein called 30S Ribosomal protein S20.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
20	T	99	Total	C	N	O	S	0	0	0
			763	470	162	129	2			

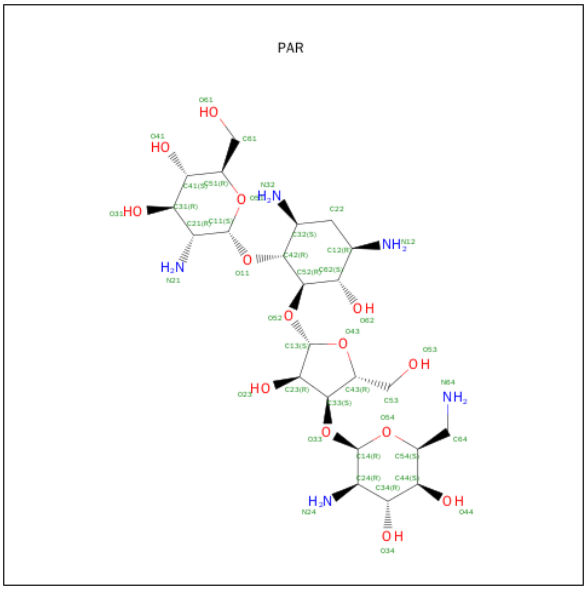
- Molecule 21 is a protein called 30S Ribosomal protein THX.

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

- Molecule 22 is a RNA chain called Fragment of messenger RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
22	X	6	Total	C	N	O	P	0	0	0
			117	54	14	44	5			

- Molecule 23 is PAROMOMYCIN (three-letter code: PAR) (formula: $C_{23}H_{45}N_5O_{14}$).



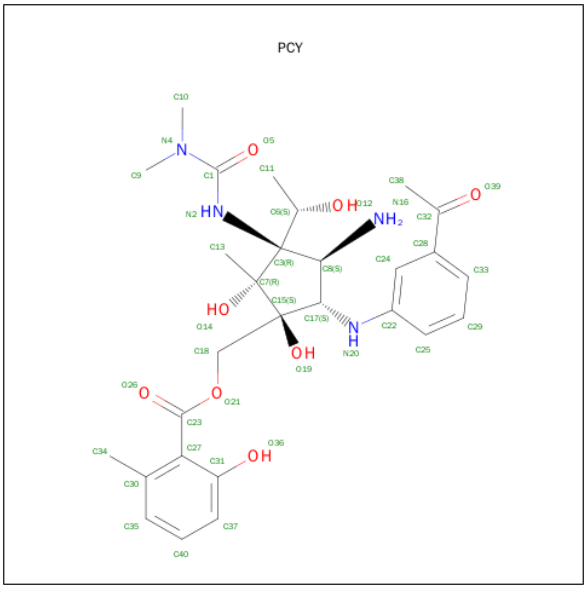
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
23	A	1	Total	C	N	O	0	0
			42	23	5	14		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
23	A	1	Total C N O 42 23 5 14	0	0
23	A	1	Total C N O 42 23 5 14	0	0
23	A	1	Total C N O 42 23 5 14	0	0
23	A	1	Total C N O 42 23 5 14	0	0
23	A	1	Total C N O 42 23 5 14	0	0
23	A	1	Total C N O 42 23 5 14	0	0

- Molecule 24 is DE-6-MSA-PACTAMYCIN (three-letter code: PCY) (formula: $C_{28}H_{38}N_4O_8$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
24	A	1	Total C N O 30 20 4 6	0	0

- Molecule 25 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
25	S	1	Total Mg 1 1	0	0
25	A	99	Total Mg 99 99	0	0

Continued on next page...

Continued from previous page...

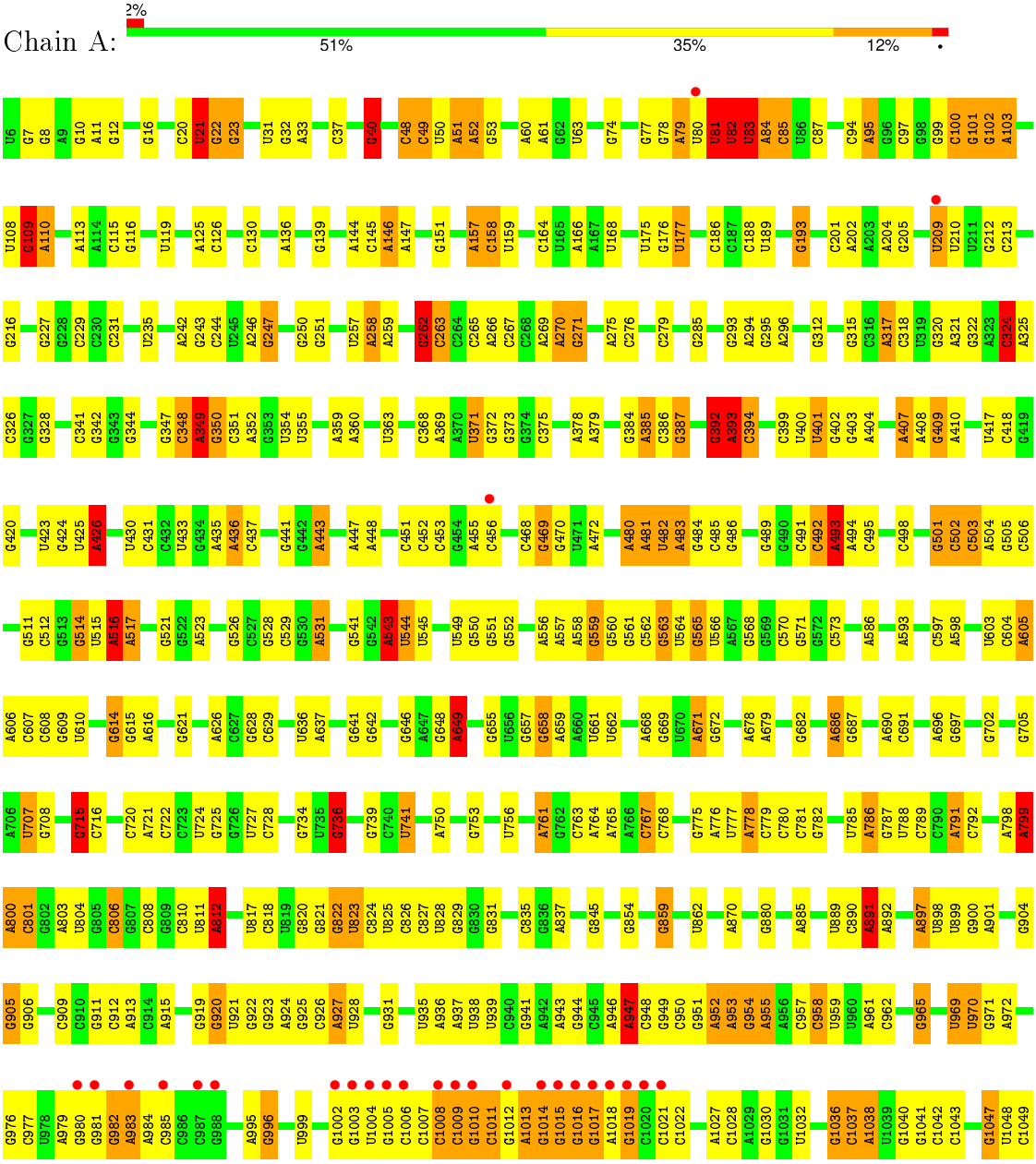
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
25	D	1	Total Mg 1 1	0	0
25	K	1	Total Mg 1 1	0	0
25	E	1	Total Mg 1 1	0	0

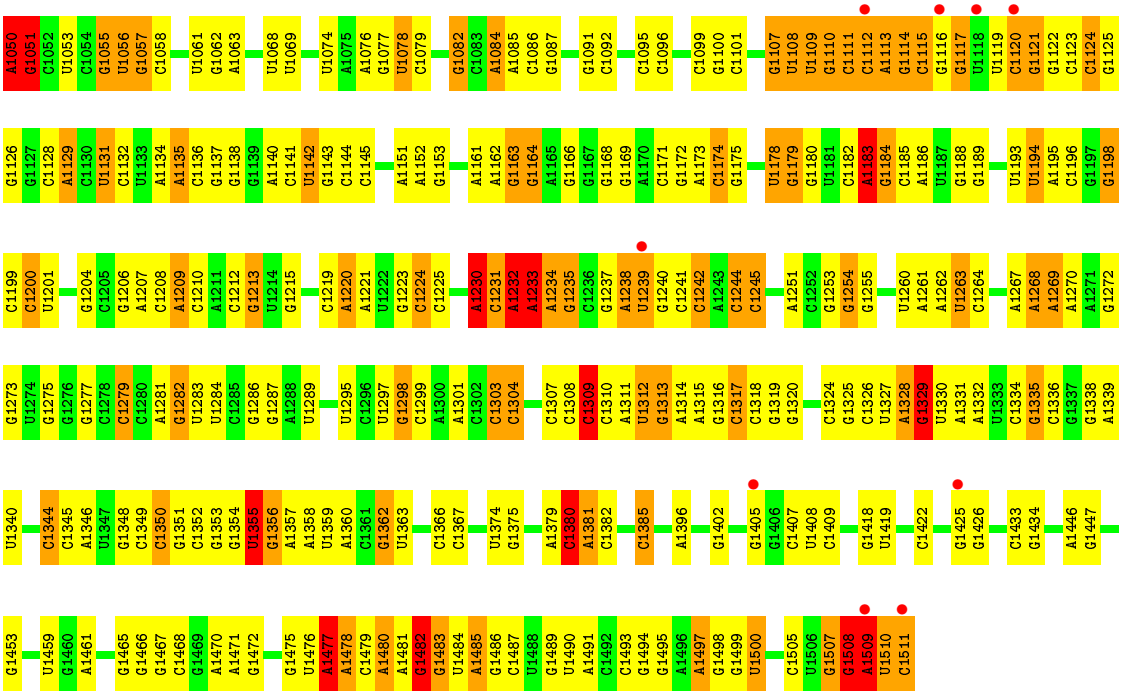
3 Residue-property plots (i)

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.

- Molecule 1: 16S Ribosomal RNA

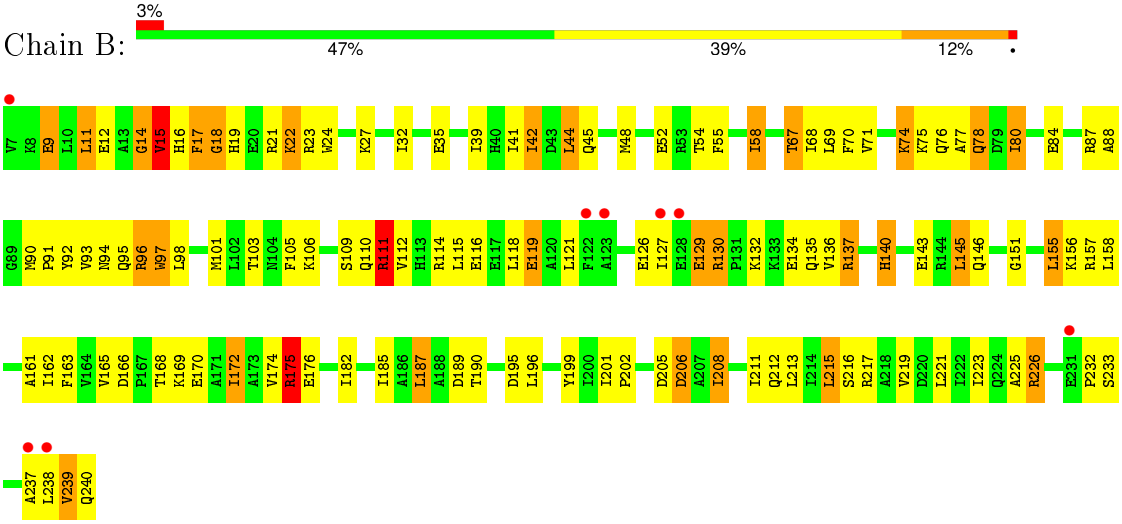
Chain A:





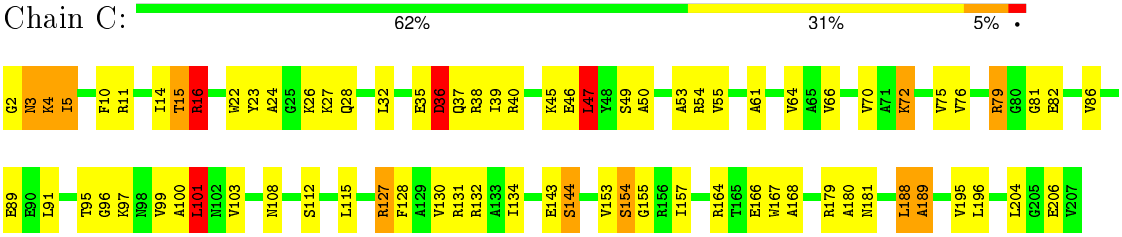
- Molecule 2: 30S Ribosomal protein S2

Chain B:



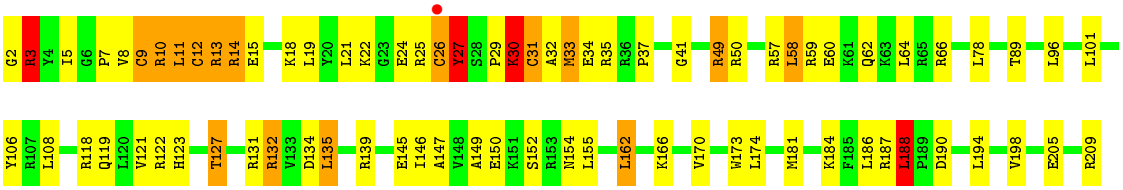
- Molecule 3: 30S Ribosomal protein S3

Chain C:

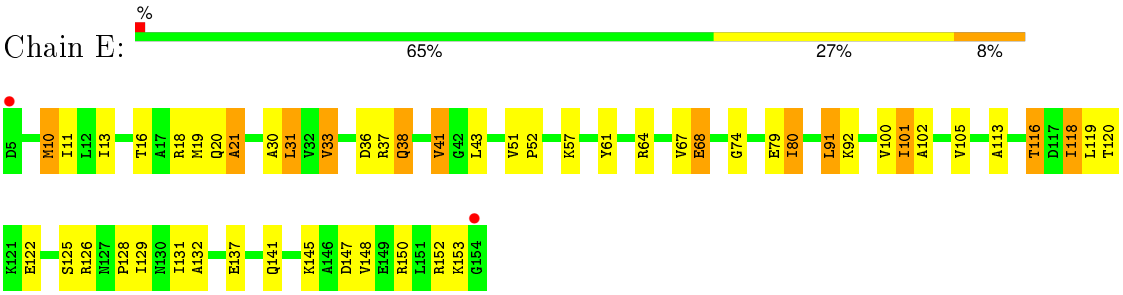


- Molecule 4: 30S Ribosomal protein S4

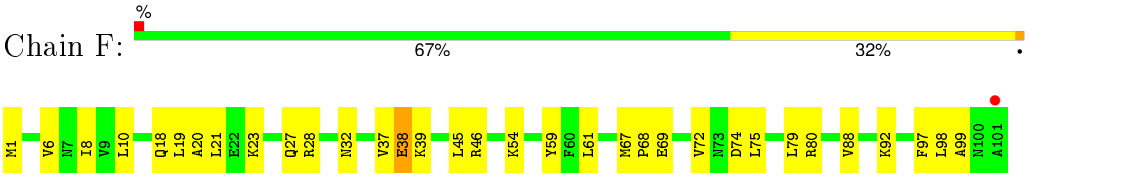
Chain D:



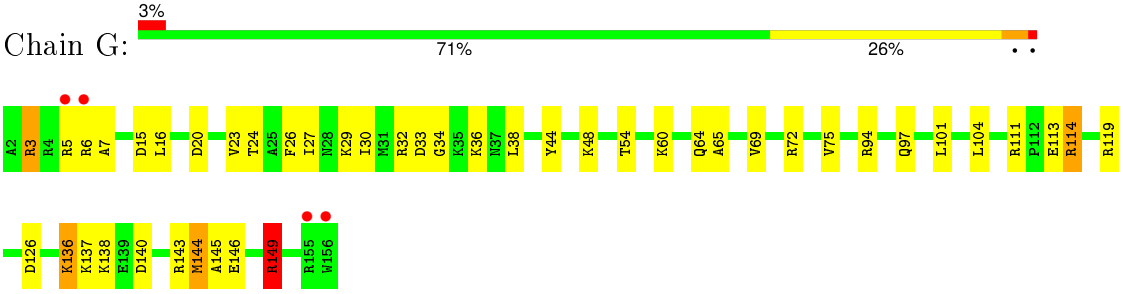
- Molecule 5: 30S Ribosomal protein S5



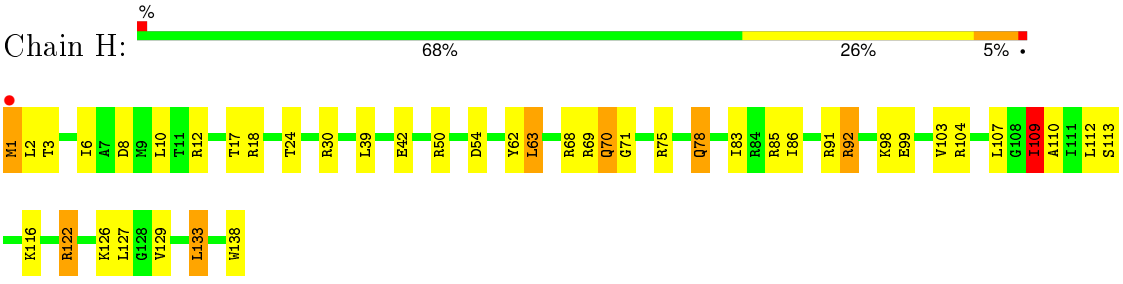
- Molecule 6: 30S Ribosomal protein S6



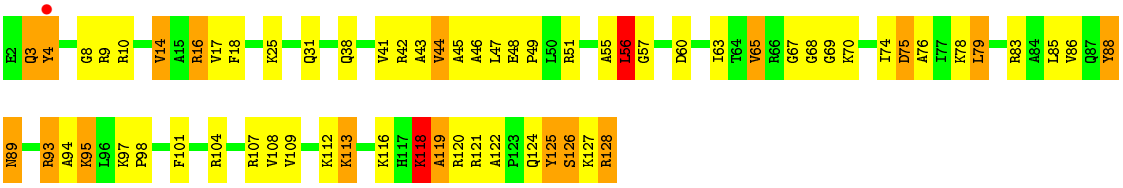
- Molecule 7: 30S Ribosomal protein S7



- Molecule 8: 30S Ribosomal protein S8

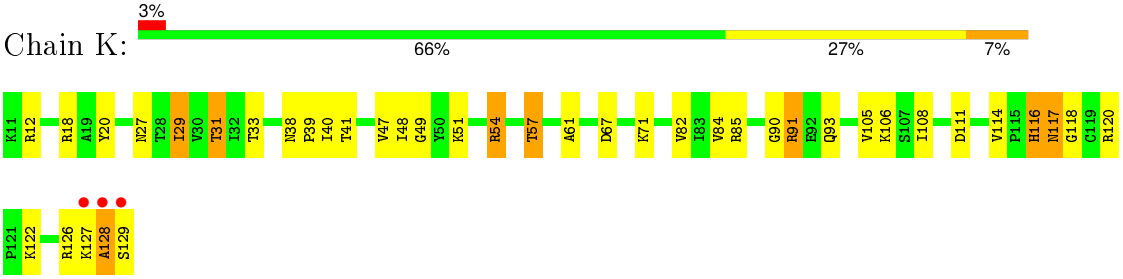


- Molecule 9: 30S Ribosomal protein S9

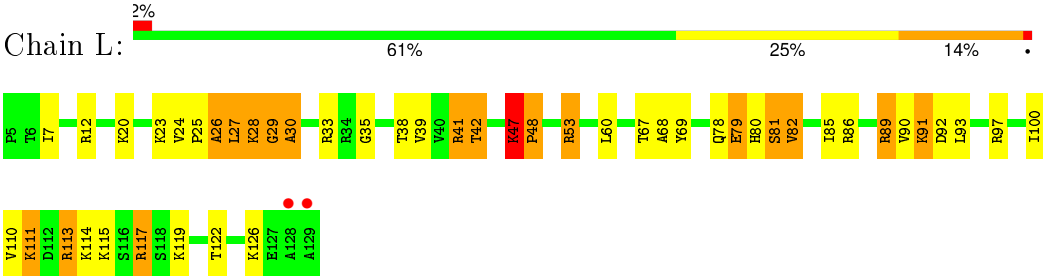


• Molecule 10: 30S Ribosomal protein S10

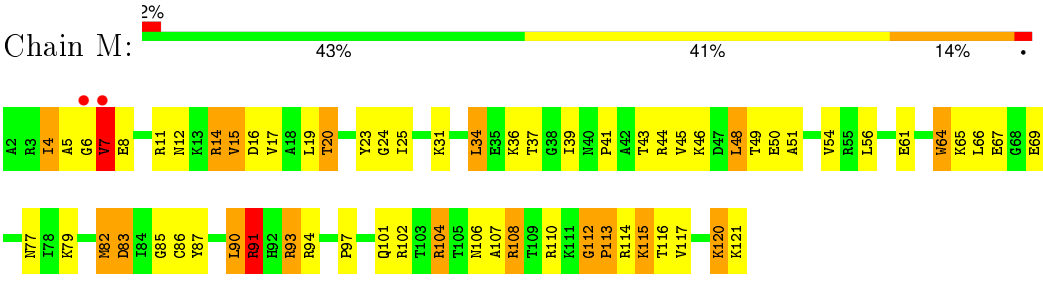
• Molecule 11: 30S Ribosomal protein S11



• Molecule 12: 30S Ribosomal protein S12



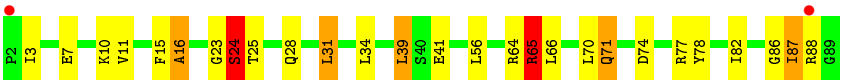
• Molecule 13: 30S Ribosomal protein S13



• Molecule 14: 30S Ribosomal protein S14

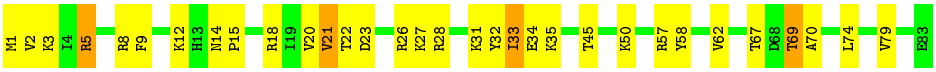
• Molecule 15: 30S Ribosomal protein S15

Chain O:



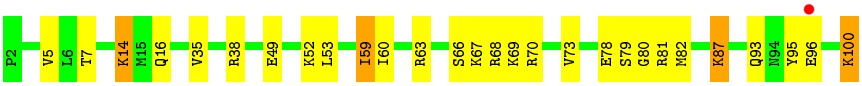
• Molecule 16: 30S Ribosomal protein S16

Chain P:



• Molecule 17: 30S Ribosomal protein S17

Chain Q:

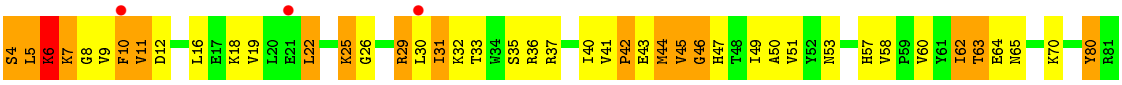


• Molecule 18: 30S Ribosomal protein S18

Chain R:

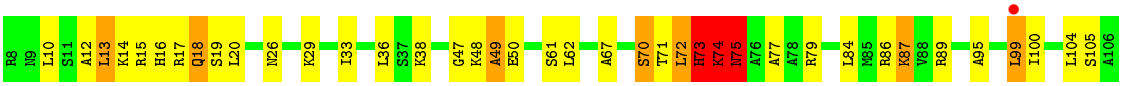
• Molecule 19: 30S Ribosomal protein S19

Chain S:



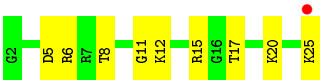
• Molecule 20: 30S Ribosomal protein S20

Chain T:



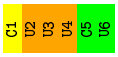
• Molecule 21: 30S Ribosomal protein THX

Chain U:



- Molecule 22: Fragment of messenger RNA

Chain X:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	401.00Å 401.00Å 176.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.67 – 3.10 29.67 – 3.10	Depositor EDS
% Data completeness (in resolution range)	91.3 (29.67-3.10) 91.5 (29.67-3.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.91 (at 3.11Å)	Xtriage
Refinement program	REFMAC 5.8.0031	Depositor
R, R_{free}	0.185 , 0.226 0.187 , 0.224	Depositor DCC
R_{free} test set	11726 reflections (5.26%)	DCC
Wilson B-factor (Å^2)	75.3	Xtriage
Anisotropy	0.012	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 71.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Outliers	0 of 234683 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	52032	wwPDB-VP
Average B, all atoms (Å^2)	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.90% 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: PAR, PCY, MG

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	0.55	14/36233 (0.0%)	0.94	127/56552 (0.2%)
2	B	0.59	0/1935	0.89	2/2609 (0.1%)
3	C	0.64	0/1636	0.86	1/2205 (0.0%)
4	D	0.68	2/1733 (0.1%)	0.96	3/2318 (0.1%)
5	E	0.72	0/1162	0.96	0/1564
6	F	0.50	0/856	0.76	0/1154
7	G	0.51	0/1276	0.79	1/1709 (0.1%)
8	H	0.73	0/1136	0.97	1/1527 (0.1%)
9	I	0.52	0/1029	0.85	0/1378
10	J	0.59	0/807	0.93	2/1085 (0.2%)
11	K	0.58	0/900	0.85	1/1213 (0.1%)
12	L	0.70	0/991	1.00	4/1327 (0.3%)
13	M	0.56	0/965	0.93	2/1292 (0.2%)
14	N	0.74	0/501	1.10	2/664 (0.3%)
15	O	0.56	0/745	0.87	1/992 (0.1%)
16	P	0.67	0/716	0.89	1/963 (0.1%)
17	Q	0.64	0/836	0.86	0/1117
18	R	0.53	0/579	0.83	0/768
19	S	0.55	0/642	0.80	0/865
20	T	0.63	0/765	0.95	0/1007
21	U	0.61	0/212	0.86	0/277
22	X	0.58	0/128	0.85	0/196
All	All	0.58	16/55783 (0.0%)	0.92	148/82782 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
4	D	0	2

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
8	H	0	1
11	K	0	1
12	L	0	2
14	N	0	1
19	S	0	1
20	T	0	3
All	All	0	11

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	82	U	O3'-P	9.75	1.72	1.61
1	A	83	U	O3'-P	8.72	1.71	1.61
1	A	799	A	O3'-P	-6.81	1.52	1.61
1	A	1057	G	O3'-P	-6.41	1.53	1.61
4	D	12	CYS	CA-CB	6.18	1.67	1.53

The worst 5 of 148 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1477	A	O5'-P-OP1	-12.26	94.67	105.70
1	A	767	C	O5'-P-OP2	-12.24	94.68	105.70
1	A	812	A	O5'-P-OP2	-11.96	94.93	105.70
1	A	492	C	C2'-C3'-O3'	10.89	133.46	109.50
1	A	399	C	O5'-P-OP2	-9.77	96.91	105.70

There are no chirality outliers.

5 of 11 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	D	30	LYS	Peptide
4	D	31	CYS	Peptide
8	H	109	ILE	Peptide
11	K	116	HIS	Peptide
12	L	26	ALA	Peptide

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	32368	0	16339	525	0
2	B	1900	0	1951	85	0
3	C	1612	0	1677	40	0
4	D	1703	0	1763	62	0
5	E	1146	0	1207	43	0
6	F	843	0	857	20	0
7	G	1257	0	1296	22	0
8	H	1116	0	1177	24	0
9	I	1011	0	1043	49	0
10	J	794	0	840	46	0
11	K	885	0	904	21	0
12	L	975	0	1062	36	0
13	M	955	0	1021	35	0
14	N	492	0	529	30	0
15	O	734	0	771	13	0
16	P	700	0	720	24	0
17	Q	823	0	891	16	0
18	R	574	0	644	14	0
19	S	629	0	652	36	0
20	T	763	0	861	24	0
21	U	208	0	221	4	0
22	X	117	0	64	2	0
23	A	294	0	315	40	0
24	A	30	0	31	5	0
25	A	99	0	0	0	0
25	D	1	0	0	0	0
25	E	1	0	0	0	0
25	K	1	0	0	0	0
25	S	1	0	0	0	0
All	All	52032	0	36836	1080	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 12.

The worst 5 of 1080 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)
1:A:1231:C:H2'	1:A:1232:A:C8	1.70	1.18
23:A:1606:PAR:H33	23:A:1606:PAR:N24	1.24	1.16
1:A:1287:G:N2	1:A:1313:G:O2'	1.80	1.13

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
23:A:1606:PAR:N24	23:A:1606:PAR:C33	2.18	1.06
5:E:101:ILE:HD11	5:E:119:LEU:HD23	1.31	1.06

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	B	232/234 (99%)	180 (78%)	39 (17%)	13 (6%)	2 13
3	C	204/206 (99%)	156 (76%)	33 (16%)	15 (7%)	1 7
4	D	206/208 (99%)	172 (84%)	23 (11%)	11 (5%)	2 14
5	E	148/150 (99%)	127 (86%)	20 (14%)	1 (1%)	26 65
6	F	99/101 (98%)	85 (86%)	12 (12%)	2 (2%)	9 38
7	G	153/155 (99%)	128 (84%)	22 (14%)	3 (2%)	9 38
8	H	136/138 (99%)	122 (90%)	13 (10%)	1 (1%)	26 65
9	I	125/127 (98%)	99 (79%)	15 (12%)	11 (9%)	1 5
10	J	96/98 (98%)	73 (76%)	18 (19%)	5 (5%)	2 15
11	K	117/119 (98%)	102 (87%)	8 (7%)	7 (6%)	2 11
12	L	123/125 (98%)	98 (80%)	16 (13%)	9 (7%)	1 7
13	M	118/120 (98%)	88 (75%)	17 (14%)	13 (11%)	0 3
14	N	58/60 (97%)	40 (69%)	12 (21%)	6 (10%)	1 4
15	O	86/88 (98%)	76 (88%)	7 (8%)	3 (4%)	4 24
16	P	81/83 (98%)	73 (90%)	8 (10%)	0	100 100
17	Q	97/99 (98%)	89 (92%)	6 (6%)	2 (2%)	9 37
18	R	68/70 (97%)	60 (88%)	7 (10%)	1 (2%)	13 46
19	S	76/78 (97%)	55 (72%)	15 (20%)	6 (8%)	1 6

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
20	T	97/99 (98%)	81 (84%)	7 (7%)	9 (9%)	1 4
21	U	22/24 (92%)	20 (91%)	2 (9%)	0	100 100
All	All	2342/2382 (98%)	1924 (82%)	300 (13%)	118 (5%)	3 16

5 of 118 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	15	VAL
2	B	17	PHE
2	B	18	GLY
2	B	44	LEU
2	B	226	ARG

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	B	202/202 (100%)	154 (76%)	48 (24%)	1 3
3	C	160/160 (100%)	130 (81%)	30 (19%)	2 8
4	D	180/180 (100%)	149 (83%)	31 (17%)	2 11
5	E	115/115 (100%)	97 (84%)	18 (16%)	3 13
6	F	90/90 (100%)	84 (93%)	6 (7%)	20 56
7	G	126/126 (100%)	106 (84%)	20 (16%)	3 13
8	H	119/119 (100%)	97 (82%)	22 (18%)	2 9
9	I	98/98 (100%)	74 (76%)	24 (24%)	1 3
10	J	88/88 (100%)	71 (81%)	17 (19%)	2 8
11	K	90/90 (100%)	74 (82%)	16 (18%)	2 10
12	L	104/104 (100%)	88 (85%)	16 (15%)	3 14
13	M	96/96 (100%)	69 (72%)	27 (28%)	0 1
14	N	49/49 (100%)	36 (74%)	13 (26%)	0 2
15	O	79/79 (100%)	67 (85%)	12 (15%)	3 14

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
16	P	72/72 (100%)	62 (86%)	10 (14%)	4 19
17	Q	94/94 (100%)	81 (86%)	13 (14%)	4 19
18	R	61/61 (100%)	50 (82%)	11 (18%)	2 10
19	S	69/69 (100%)	50 (72%)	19 (28%)	0 1
20	T	76/76 (100%)	63 (83%)	13 (17%)	2 11
21	U	19/19 (100%)	15 (79%)	4 (21%)	1 6
All	All	1987/1987 (100%)	1617 (81%)	370 (19%)	2 9

5 of 370 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
8	H	92	ARG
10	J	70	ARG
19	S	29	ARG
8	H	122	ARG
9	I	85	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
8	H	70	GLN
10	J	68	HIS
20	T	18	GLN
9	I	3	GLN
9	I	31	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1505/1506 (99%)	339 (22%)	66 (4%)
22	X	5/6 (83%)	3 (60%)	0
All	All	1510/1512 (99%)	342 (22%)	66 (4%)

5 of 342 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	7	G
1	A	10	G

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	21	U
1	A	23	G
1	A	32	G

5 of 66 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	671	A
1	A	909	C
1	A	1309	C
1	A	686	A
1	A	824	C

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

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

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 111 ligands modelled in this entry, 103 are monoatomic - leaving 8 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	$\#\|Z\| > 2$
23	PAR	A	1601	-	45,45,45	0.79	1 (2%)	59,67,67	1.82	12 (20%)
23	PAR	A	1602	-	45,45,45	0.82	1 (2%)	59,67,67	1.63	14 (23%)
23	PAR	A	1603	-	45,45,45	0.79	2 (4%)	59,67,67	1.97	10 (16%)
23	PAR	A	1604	-	45,45,45	0.87	2 (4%)	59,67,67	1.81	12 (20%)
23	PAR	A	1605	-	45,45,45	0.83	0	59,67,67	1.90	12 (20%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
23	PAR	A	1606	-	45,45,45	0.75	0	59,67,67	1.85	15 (25%)
23	PAR	A	1607	-	45,45,45	1.13	4 (8%)	59,67,67	2.01	19 (32%)
24	PCY	A	1608	-	26,31,42	1.32	2 (7%)	24,49,65	1.47	4 (16%)

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
23	PAR	A	1601	-	-	0/18/94/94	0/4/4/4
23	PAR	A	1602	-	-	0/18/94/94	0/4/4/4
23	PAR	A	1603	-	-	1/18/94/94	0/4/4/4
23	PAR	A	1604	-	-	0/18/94/94	0/4/4/4
23	PAR	A	1605	-	-	1/18/94/94	0/4/4/4
23	PAR	A	1606	-	-	1/18/94/94	0/4/4/4
23	PAR	A	1607	-	-	0/18/94/94	1/4/4/4
24	PCY	A	1608	-	-	0/26/60/67	0/2/2/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
24	A	1608	PCY	C28-C32	-4.04	1.42	1.49
23	A	1607	PAR	C31-C21	-3.90	1.48	1.53
24	A	1608	PCY	C22-N20	-3.29	1.33	1.39
23	A	1604	PAR	C13-C23	-3.27	1.48	1.52
23	A	1601	PAR	C13-C23	-2.48	1.49	1.52

The worst 5 of 98 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
23	A	1603	PAR	O34-C34-C24	-5.92	100.31	110.31
23	A	1604	PAR	O54-C54-C64	-5.37	95.61	106.10
23	A	1603	PAR	C14-O33-C33	-5.10	104.69	118.01
23	A	1607	PAR	O53-C53-C43	-4.87	95.25	111.33
23	A	1607	PAR	O11-C11-C21	-4.66	99.33	107.96

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
23	A	1603	PAR	C52-O52-C13-C23
23	A	1605	PAR	C52-O52-C13-C23
23	A	1606	PAR	C33-O33-C14-C24

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
23	A	1607	PAR	C12-C22-C32-C42-C52-C62

7 monomers are involved in 45 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
23	A	1601	PAR	3	0
23	A	1603	PAR	6	0
23	A	1604	PAR	9	0
23	A	1605	PAR	1	0
23	A	1606	PAR	9	0
23	A	1607	PAR	12	0
24	A	1608	PCY	5	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

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	1506/1506 (100%)	-0.23	35 (2%) 64 40	39, 69, 142, 245	0
2	B	234/234 (100%)	-0.08	8 (3%) 49 24	53, 98, 168, 194	0
3	C	206/206 (100%)	-0.35	0 100 100	55, 95, 139, 157	0
4	D	208/208 (100%)	-0.44	1 (0%) 91 83	53, 80, 113, 134	0
5	E	150/150 (100%)	-0.47	2 (1%) 79 62	41, 65, 98, 136	0
6	F	101/101 (100%)	-0.27	1 (0%) 84 69	72, 101, 131, 172	0
7	G	155/155 (100%)	-0.32	4 (2%) 59 35	65, 96, 145, 196	0
8	H	138/138 (100%)	-0.49	1 (0%) 89 78	40, 61, 91, 131	0
9	I	127/127 (100%)	-0.27	1 (0%) 87 75	51, 103, 140, 219	0
10	J	98/98 (100%)	-0.03	1 (1%) 84 69	61, 126, 174, 199	0
11	K	119/119 (100%)	-0.23	3 (2%) 61 37	53, 78, 113, 164	0
12	L	125/125 (100%)	-0.38	2 (1%) 74 55	39, 76, 112, 179	0
13	M	120/120 (100%)	-0.18	2 (1%) 73 52	60, 97, 144, 183	0
14	N	60/60 (100%)	-0.36	0 100 100	62, 85, 113, 131	0
15	O	88/88 (100%)	-0.37	2 (2%) 64 40	55, 82, 121, 155	0
16	P	83/83 (100%)	-0.54	0 100 100	50, 67, 99, 140	0
17	Q	99/99 (100%)	-0.49	1 (1%) 84 69	42, 72, 106, 142	0
18	R	70/70 (100%)	-0.36	0 100 100	61, 81, 123, 157	0
19	S	78/78 (100%)	0.03	3 (3%) 44 21	72, 114, 168, 186	0
20	T	99/99 (100%)	-0.40	1 (1%) 84 69	51, 78, 136, 162	0
21	U	24/24 (100%)	0.13	1 (4%) 40 19	67, 82, 111, 134	0
22	X	6/6 (100%)	0.03	0 100 100	59, 75, 131, 138	0
All	All	3894/3894 (100%)	-0.28	69 (1%) 71 50	39, 79, 145, 245	0

The worst 5 of 69 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	7	VAL	6.4
11	K	129	SER	6.3
1	A	1012	G	5.7
11	K	128	ALA	5.6
19	S	30	LEU	5.4

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

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

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. 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
25	MG	A	1622	1/1	0.93	0.45	28.56	38,38,38,38	0
25	MG	A	1661	1/1	0.93	1.01	25.06	60,60,60,60	0
25	MG	A	1651	1/1	0.98	0.55	23.99	42,42,42,42	0
25	MG	A	1617	1/1	0.98	0.43	22.86	35,35,35,35	0
25	MG	A	1610	1/1	0.99	0.57	19.68	42,42,42,42	0
25	MG	A	1659	1/1	0.88	0.42	19.49	36,36,36,36	1
25	MG	A	1632	1/1	0.92	0.71	19.38	56,56,56,56	0
25	MG	A	1703	1/1	0.77	0.51	18.13	88,88,88,88	0
25	MG	D	301	1/1	0.81	0.47	17.00	93,93,93,93	0
25	MG	A	1620	1/1	0.95	0.30	16.34	34,34,34,34	0
25	MG	A	1634	1/1	0.98	0.31	14.72	44,44,44,44	0
25	MG	A	1614	1/1	0.98	0.45	13.87	31,31,31,31	0
25	MG	A	1609	1/1	0.94	0.34	13.66	45,45,45,45	0
25	MG	A	1639	1/1	0.94	0.34	13.30	47,47,47,47	0
25	MG	A	1647	1/1	0.98	0.32	10.80	34,34,34,34	0
25	MG	E	201	1/1	0.95	0.41	8.11	39,39,39,39	0
25	MG	A	1645	1/1	0.94	0.23	6.35	34,34,34,34	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
25	MG	A	1633	1/1	0.95	0.60	5.97	52,52,52,52	0
25	MG	A	1685	1/1	0.84	0.26	5.66	46,46,46,46	0
25	MG	A	1654	1/1	0.97	0.23	5.62	26,26,26,26	0
25	MG	A	1682	1/1	0.93	0.21	5.59	52,52,52,52	0
25	MG	A	1636	1/1	0.90	0.20	4.84	46,46,46,46	0
23	PAR	A	1605	42/42	0.91	0.25	4.79	58,75,85,89	42
23	PAR	A	1607	42/42	0.93	0.25	4.58	11,15,17,19	42
23	PAR	A	1602	42/42	0.96	0.19	4.53	32,41,50,61	42
25	MG	A	1667	1/1	0.81	0.29	4.07	39,39,39,39	0
23	PAR	A	1606	42/42	0.86	0.27	3.71	38,44,55,56	42
23	PAR	A	1603	42/42	0.90	0.21	3.52	53,84,97,117	42
25	MG	A	1652	1/1	0.97	0.25	3.51	49,49,49,49	0
25	MG	A	1683	1/1	0.86	0.20	1.93	44,44,44,44	1
25	MG	A	1699	1/1	0.90	0.31	1.79	55,55,55,55	0
25	MG	A	1644	1/1	0.95	0.17	1.18	41,41,41,41	0
23	PAR	A	1604	42/42	0.94	0.22	1.08	64,79,104,137	0
25	MG	A	1655	1/1	0.93	0.18	1.08	39,39,39,39	1
24	PCY	A	1608	30/40	0.96	0.21	0.65	50,59,68,76	0
25	MG	A	1657	1/1	0.97	0.17	0.61	23,23,23,23	0
25	MG	A	1700	1/1	0.93	0.17	0.35	52,52,52,52	0
23	PAR	A	1601	42/42	0.96	0.15	0.18	52,59,74,83	0
25	MG	A	1640	1/1	0.97	0.17	-0.08	36,36,36,36	0
25	MG	A	1674	1/1	0.90	0.13	-1.36	49,49,49,49	0
25	MG	A	1650	1/1	0.99	0.12	-1.94	37,37,37,37	0
25	MG	A	1623	1/1	0.94	0.10	-2.01	57,57,57,57	0
25	MG	A	1677	1/1	0.81	0.59	-	58,58,58,58	0
25	MG	A	1681	1/1	0.99	0.37	-	36,36,36,36	0
25	MG	A	1665	1/1	0.88	0.12	-	40,40,40,40	0
25	MG	A	1690	1/1	0.85	0.16	-	54,54,54,54	0
25	MG	A	1689	1/1	0.89	0.34	-	52,52,52,52	0
25	MG	A	1664	1/1	0.95	0.24	-	61,61,61,61	0
25	MG	A	1613	1/1	0.97	0.16	-	56,56,56,56	1
25	MG	A	1638	1/1	0.97	0.26	-	40,40,40,40	0
25	MG	A	1627	1/1	0.92	0.47	-	50,50,50,50	0
25	MG	A	1615	1/1	0.96	0.68	-	38,38,38,38	0
25	MG	A	1629	1/1	0.96	0.37	-	42,42,42,42	0
25	MG	A	1706	1/1	0.82	0.49	-	71,71,71,71	0
25	MG	A	1691	1/1	0.71	0.34	-	69,69,69,69	0
25	MG	A	1680	1/1	0.96	0.25	-	36,36,36,36	1
25	MG	A	1676	1/1	0.82	0.27	-	51,51,51,51	0
25	MG	A	1669	1/1	0.94	0.63	-	57,57,57,57	0
25	MG	A	1686	1/1	0.90	0.39	-	42,42,42,42	1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
25	MG	A	1692	1/1	0.91	0.37	-	65,65,65,65	0
25	MG	A	1616	1/1	0.98	0.67	-	40,40,40,40	0
25	MG	A	1630	1/1	0.92	0.44	-	33,33,33,33	1
25	MG	A	1628	1/1	0.96	0.34	-	39,39,39,39	0
25	MG	A	1642	1/1	0.91	0.21	-	43,43,43,43	0
25	MG	A	1684	1/1	0.43	0.63	-	67,67,67,67	1
25	MG	A	1658	1/1	0.85	0.41	-	41,41,41,41	1
25	MG	A	1646	1/1	0.99	0.26	-	34,34,34,34	0
25	MG	A	1611	1/1	0.93	0.26	-	37,37,37,37	0
25	MG	A	1693	1/1	0.92	0.29	-	52,52,52,52	0
25	MG	A	1688	1/1	0.79	0.36	-	49,49,49,49	0
25	MG	A	1673	1/1	0.96	0.13	-	41,41,41,41	1
25	MG	A	1625	1/1	0.95	0.28	-	50,50,50,50	0
25	MG	A	1696	1/1	0.88	0.25	-	54,54,54,54	0
25	MG	A	1695	1/1	0.93	0.34	-	64,64,64,64	0
25	MG	A	1641	1/1	0.94	0.17	-	42,42,42,42	0
25	MG	S	101	1/1	0.94	0.10	-	58,58,58,58	0
25	MG	A	1694	1/1	0.85	0.73	-	61,61,61,61	0
25	MG	A	1619	1/1	0.94	0.13	-	25,25,25,25	0
25	MG	A	1656	1/1	0.98	0.37	-	50,50,50,50	0
25	MG	A	1653	1/1	0.97	0.50	-	30,30,30,30	0
25	MG	A	1621	1/1	0.94	0.24	-	45,45,45,45	0
25	MG	A	1702	1/1	0.82	0.14	-	51,51,51,51	0
25	MG	A	1698	1/1	0.97	0.44	-	53,53,53,53	0
25	MG	A	1626	1/1	0.97	0.34	-	31,31,31,31	0
25	MG	A	1687	1/1	0.94	0.45	-	43,43,43,43	1
25	MG	A	1662	1/1	0.90	0.68	-	48,48,48,48	0
25	MG	A	1705	1/1	0.90	0.30	-	61,61,61,61	0
25	MG	A	1678	1/1	0.90	0.17	-	54,54,54,54	0
25	MG	A	1618	1/1	0.87	0.49	-	56,56,56,56	0
25	MG	A	1707	1/1	0.77	0.43	-	67,67,67,67	0
25	MG	A	1666	1/1	0.96	0.36	-	47,47,47,47	0
25	MG	A	1631	1/1	0.97	0.19	-	36,36,36,36	0
25	MG	A	1660	1/1	0.93	0.48	-	61,61,61,61	0
25	MG	A	1635	1/1	0.98	0.12	-	42,42,42,42	0
25	MG	A	1697	1/1	0.67	0.38	-	63,63,63,63	0
25	MG	A	1671	1/1	0.85	0.48	-	57,57,57,57	1
25	MG	A	1648	1/1	0.96	0.43	-	43,43,43,43	0
25	MG	A	1643	1/1	0.91	0.45	-	59,59,59,59	0
25	MG	A	1704	1/1	0.96	0.08	-	53,53,53,53	0
25	MG	A	1668	1/1	0.97	0.47	-	39,39,39,39	1
25	MG	A	1637	1/1	0.89	0.27	-	59,59,59,59	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
25	MG	A	1670	1/1	0.74	0.50	-	49,49,49,49	1
25	MG	A	1675	1/1	0.87	0.38	-	63,63,63,63	0
25	MG	A	1612	1/1	0.97	0.73	-	46,46,46,46	0
25	MG	A	1679	1/1	0.95	0.25	-	61,61,61,61	0
25	MG	A	1701	1/1	0.79	0.60	-	68,68,68,68	0
25	MG	A	1624	1/1	0.87	0.09	-	38,38,38,38	0
25	MG	A	1649	1/1	0.98	0.14	-	53,53,53,53	1
25	MG	A	1663	1/1	0.89	0.75	-	64,64,64,64	0
25	MG	A	1672	1/1	0.94	0.46	-	65,65,65,65	0
25	MG	K	201	1/1	0.96	0.59	-	52,52,52,52	0

6.5 Other polymers (i)

There are no such residues in this entry.