



# wwPDB X-ray Structure Validation Summary Report

Oct 5, 2023 – 11:06 PM EDT

PDB ID : 6MPI  
Title : Structure of the *Thermus thermophilus* 30S ribosomal subunit complexed with a 2-thiocytidine (s2C32) and inosine (I34) modified anticodon stem loop (ASL) of *Escherichia coli* transfer RNA Arginine 1 (TRNAARG1) bound to an mRNA with an CGU-codon in the A-site and paromomycin  
Authors : Cantara, W.A.; DeMirci, H.; Agris, P.F.  
Deposited on : 2018-10-07  
Resolution : 3.33 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the  symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references](#) ①) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.35.1  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.35.1

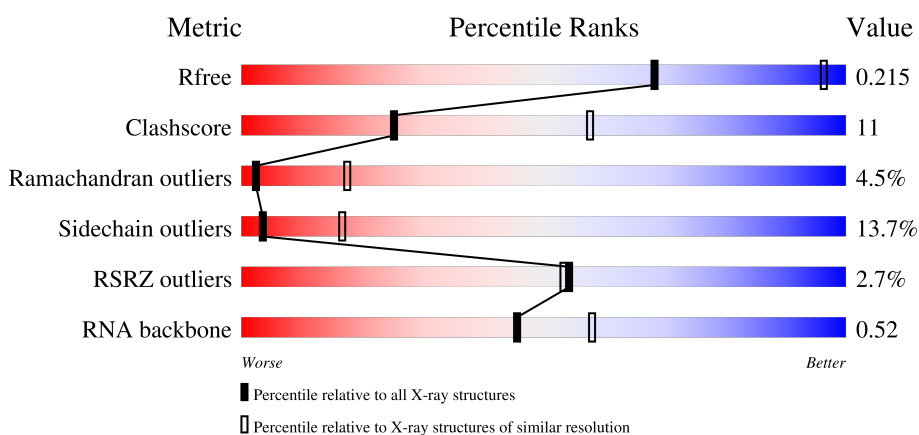
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.33 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.










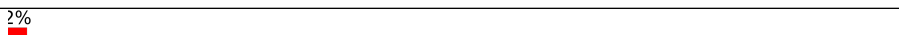
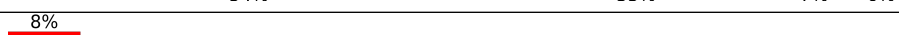
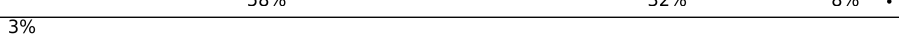
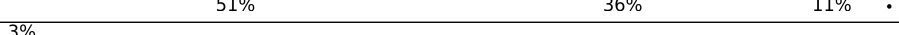






Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1060 (3.38-3.30)
Clashscore	141614	1111 (3.38-3.30)
Ramachandran outliers	138981	1090 (3.38-3.30)
Sidechain outliers	138945	1089 (3.38-3.30)
RSRZ outliers	127900	1028 (3.38-3.30)
RNA backbone	3102	1129 (3.78-2.90)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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  $\leq 5\%$ . 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	1507	
2	B	256	
3	C	239	
4	D	209	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	162	
6	F	101	
7	G	156	
8	H	138	
9	I	128	
10	J	105	
11	K	129	
12	L	132	
13	M	126	
14	N	61	
15	O	89	
16	P	88	
17	Q	105	
18	R	88	
19	S	93	
20	T	106	
21	V	27	
22	X	17	
23	W	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
22	RSP	X	32	-	-	-	X
24	MG	A	1603	-	-	-	X
24	MG	A	1605	-	-	-	X
24	MG	A	1606	-	-	-	X
24	MG	A	1609	-	-	-	X

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
24	MG	A	1624	-	-	-	X
24	MG	A	1634	-	-	-	X
24	MG	A	1650	-	-	-	X
24	MG	A	1666	-	-	-	X
24	MG	A	1667	-	-	-	X
24	MG	A	1669	-	-	-	X
24	MG	A	1670	-	-	-	X
24	MG	A	1671	-	-	-	X
24	MG	A	1676	-	-	-	X
24	MG	A	1680	-	-	-	X
24	MG	A	1691	-	-	-	X
24	MG	N	101	-	-	-	X

## 2 Entry composition

There are 26 unique types of molecules in this entry. The entry contains 52062 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 rRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	A	1507	32380	14414	5990	10470	1506	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	G	deletion	GB 55771382
A	?	-	G	deletion	GB 55771382
A	?	-	C	deletion	GB 55771382
A	?	-	G	deletion	GB 55771382
A	?	-	A	deletion	GB 55771382
A	?	-	C	deletion	GB 55771382
A	1540	U	-	insertion	GB 55771382
A	1541	U	-	insertion	GB 55771382
A	1542	U	-	insertion	GB 55771382
A	1543	C	-	insertion	GB 55771382
A	1544	U	-	insertion	GB 55771382

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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
			Total	C	N	O	S			
10	J	98	792	498	156	137	1	0	0	0

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

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

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
13	M	125	997	617	207	171	2	0	0	0

- Molecule 14 is a protein called 30S ribosomal protein S14 type Z.

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

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

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

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

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

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



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

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Q	96	GLN	GLU	conflict	UNP P0DOY7

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

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

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	60	GLY	ALA	conflict	UNP Q5SLQ0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
19	S	80	Total	C	N	O	S	0	0	0
			647	414	119	112	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
			762	469	162	129	2			

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

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

- Molecule 22 is a RNA chain called tRNA ASL Escherichia coli Arg1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
22	X	11	Total	C	N	O	P	S	0	0	0
			232	105	43	73	10	1			

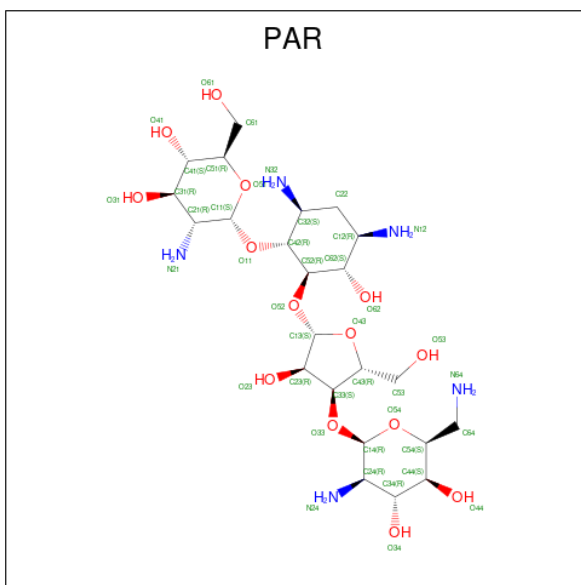
- Molecule 23 is a RNA chain called mRNA A-site fragment.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
			Total	C	N	O				P
23	W	4	82	38	15	26	3	0	0	0

- Molecule 24 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
24	A	93	Total	Mg	0	0
			93	93		
24	J	1	Total	Mg	0	0
			1	1		
24	N	1	Total	Mg	0	0
			1	1		
24	W	1	Total	Mg	0	0
			1	1		

- Molecule 25 is PAROMOMYCIN (three-letter code: PAR) (formula: C<sub>23</sub>H<sub>45</sub>N<sub>5</sub>O<sub>14</sub>).



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

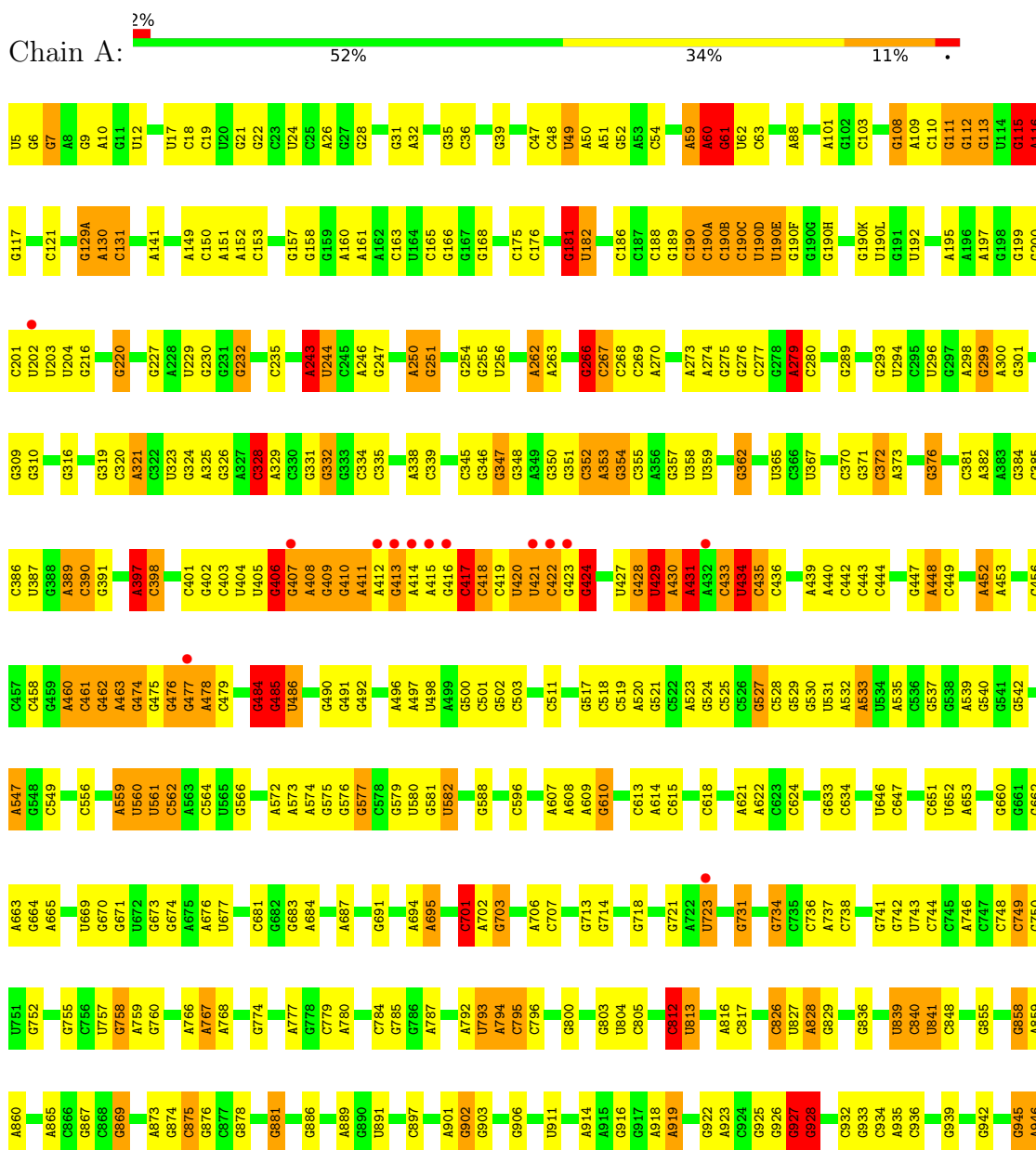
- Molecule 26 is ZINC ION (three-letter code: ZN) (formula: Zn).

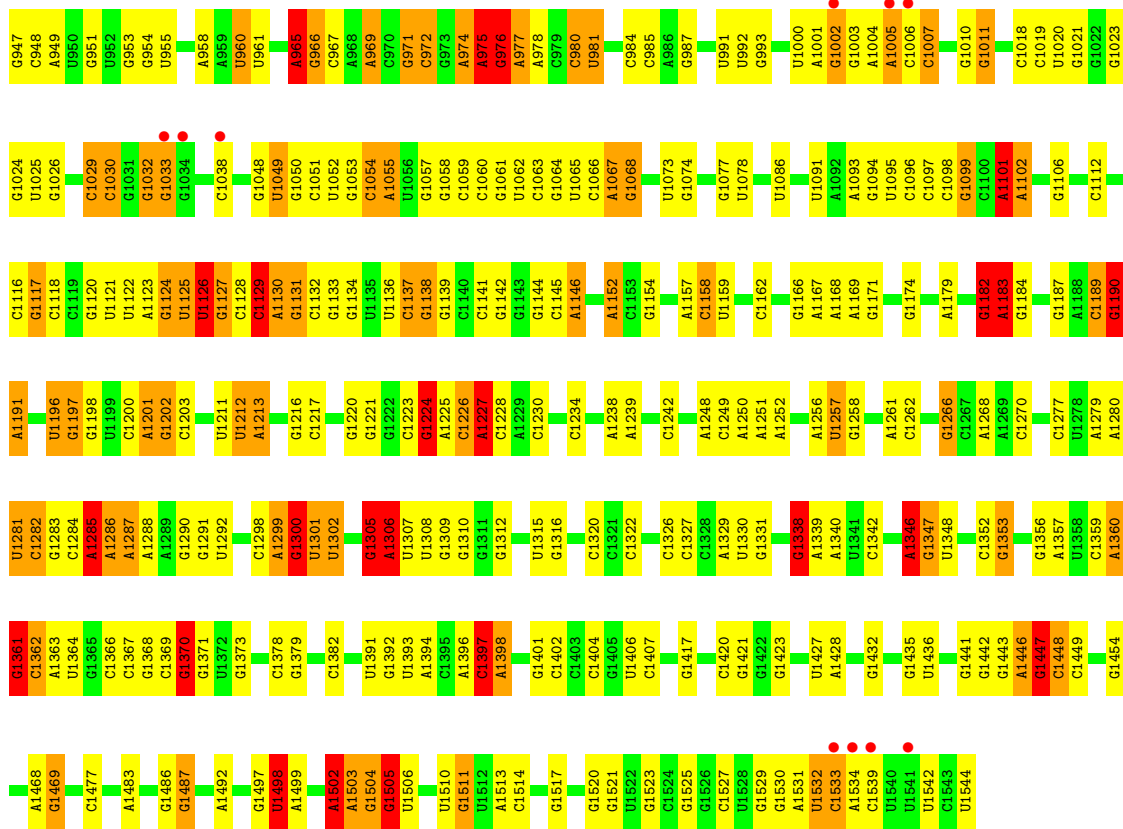
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
26	D	1	Total 1	Zn 1	0	0

### 3 Residue-property plots i

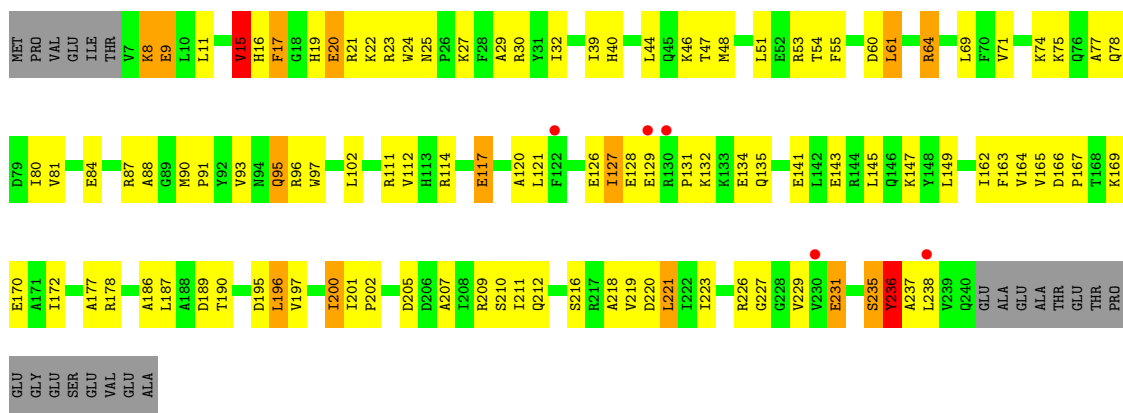
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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 rRNA



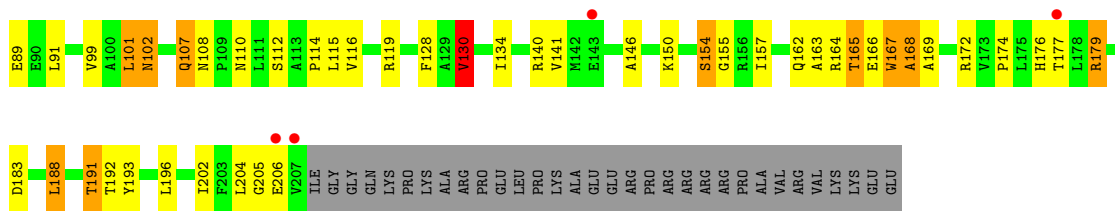


● Molecule 2: 30S ribosomal protein S2

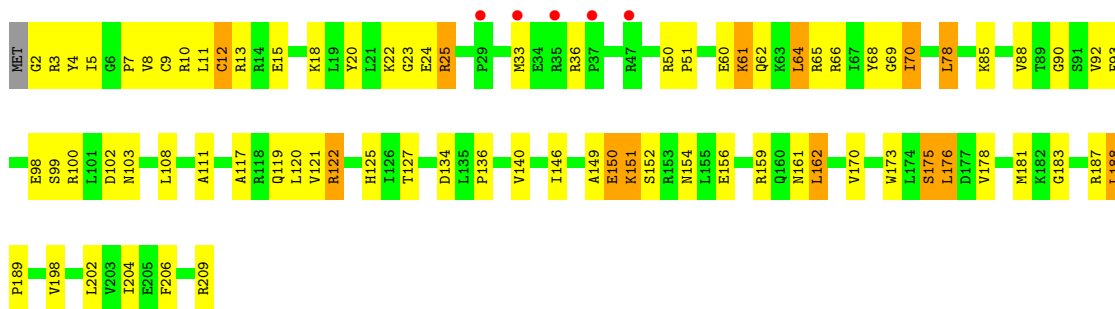


● Molecule 3: 30S ribosomal protein S3

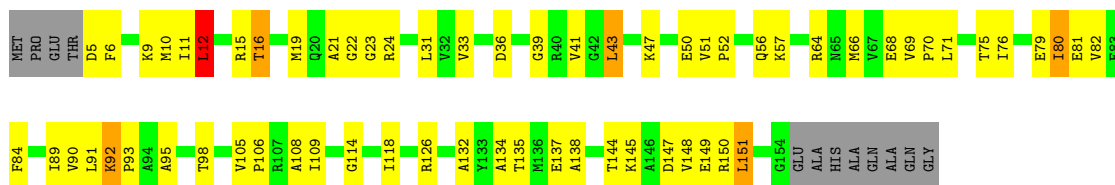




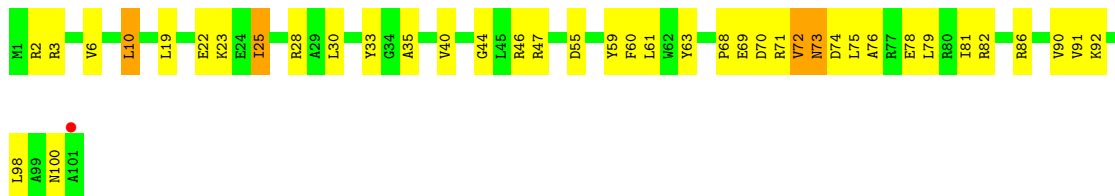
• Molecule 4: 30S ribosomal protein S4



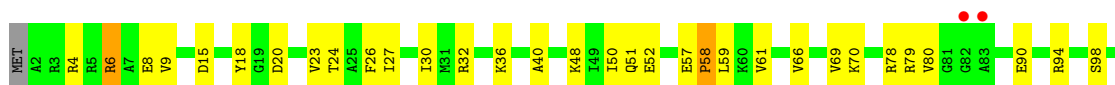
• Molecule 5: 30S ribosomal protein S5



• Molecule 6: 30S ribosomal protein S6

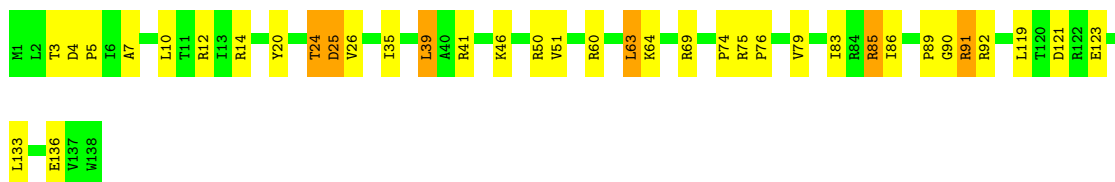


• Molecule 7: 30S ribosomal protein S7

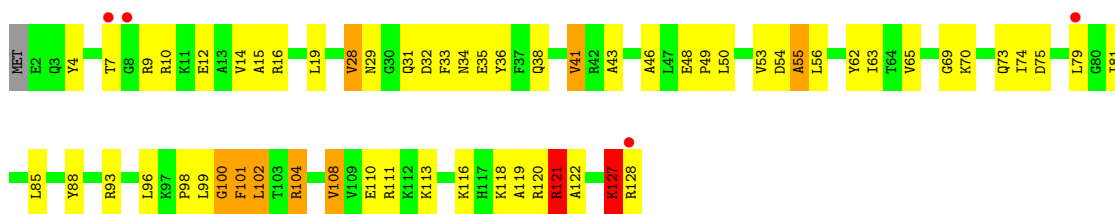




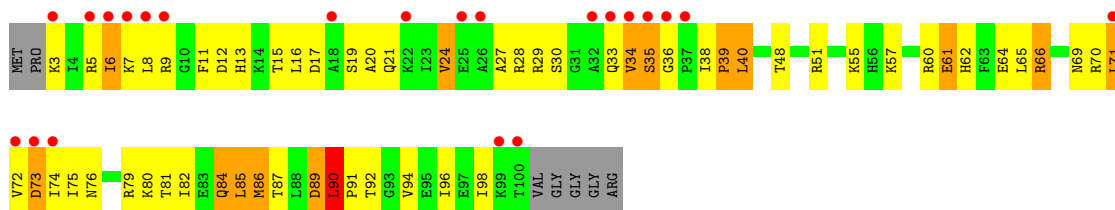
- Molecule 8: 30S ribosomal protein S8



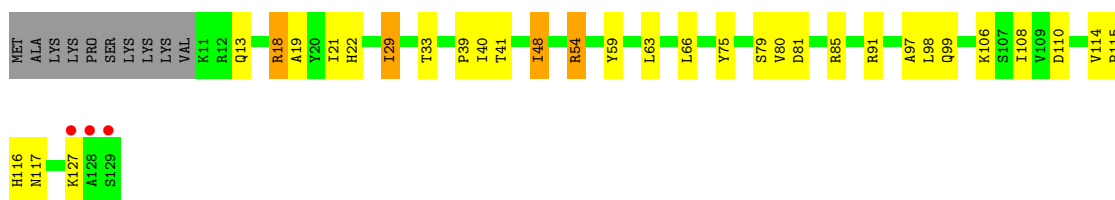
- Molecule 9: 30S ribosomal protein S9



- Molecule 10: 30S ribosomal protein S10

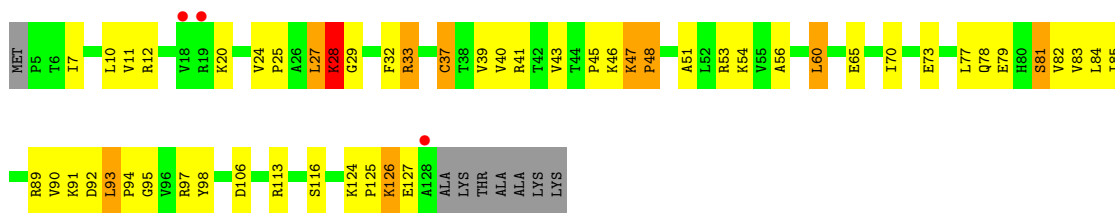


- Molecule 11: 30S ribosomal protein S11

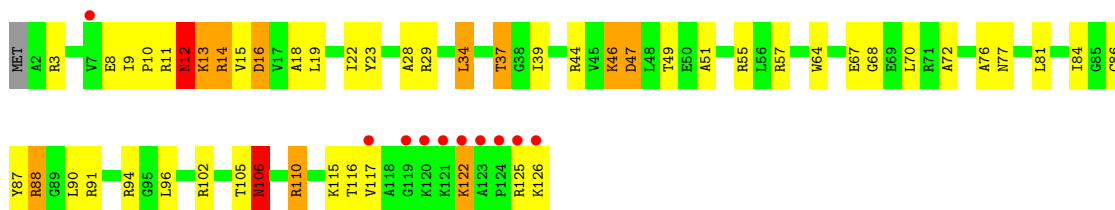


- Molecule 12: 30S ribosomal protein S12





- Molecule 13: 30S ribosomal protein S13



- Molecule 14: 30S ribosomal protein S14 type Z



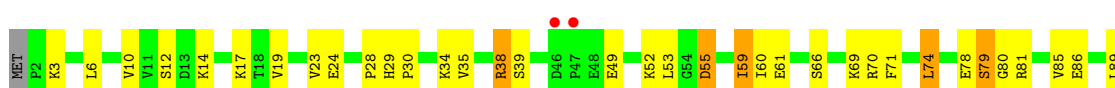
- Molecule 15: 30S ribosomal protein S15



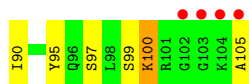
- Molecule 16: 30S ribosomal protein S16



- Molecule 17: 30S ribosomal protein S17



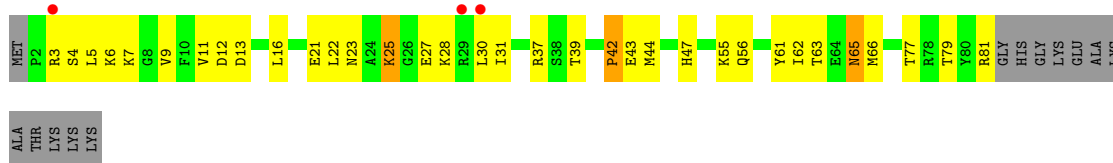




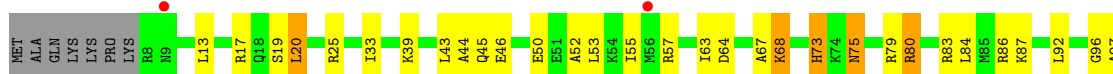
- Molecule 18: 30S ribosomal protein S18



- Molecule 19: 30S ribosomal protein S19



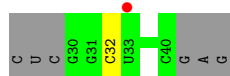
- Molecule 20: 30S ribosomal protein S20



- Molecule 21: 30S ribosomal protein Thx



- Molecule 22: tRNA ASL Escherichia coli Arg1



- Molecule 23: mRNA A-site fragment



C1	C2	U3	A4	A	A
----	----	----	----	---	---

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	403.96Å 403.96Å 176.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.77 – 3.33 48.21 – 3.33	Depositor EDS
% Data completeness (in resolution range)	86.7 (49.77-3.33) 83.0 (48.21-3.33)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.34 (at 3.33Å)	Xtrriage
Refinement program	PHENIX 1.8.4_1496	Depositor
R, $R_{free}$	0.190 , 0.216 0.192 , 0.215	Depositor DCC
$R_{free}$ test set	1739 reflections (0.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	115.8	Xtrriage
Anisotropy	0.101	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 89.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	52062	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	124.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.97% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality i

### 5.1 Standard geometry i

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

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	0.79	28/36244 (0.1%)	1.39	355/56567 (0.6%)
2	B	0.40	0/1935	0.59	0/2609
3	C	0.45	0/1636	0.61	0/2205
4	D	0.45	0/1733	0.61	1/2318 (0.0%)
5	E	0.55	0/1162	0.70	1/1564 (0.1%)
6	F	0.43	0/856	0.59	0/1154
7	G	0.39	0/1276	0.55	0/1709
8	H	0.57	0/1136	0.72	0/1527
9	I	0.42	0/1029	0.62	0/1378
10	J	0.43	0/805	0.63	0/1082
11	K	0.46	0/900	0.61	0/1213
12	L	0.56	0/986	0.74	0/1320
13	M	0.41	0/1008	0.60	0/1347
14	N	0.47	0/501	0.65	0/664
15	O	0.44	0/745	0.61	0/992
16	P	0.56	0/716	0.72	0/963
17	Q	0.53	0/870	0.68	0/1159
18	R	0.44	0/603	0.62	0/799
19	S	0.36	0/661	0.59	0/890
20	T	0.51	0/764	0.69	1/1006 (0.1%)
21	V	0.47	0/212	0.65	0/277
22	X	0.68	0/235	1.27	0/360
23	W	0.75	0/91	1.09	0/140
All	All	0.69	28/56104 (0.0%)	1.20	358/83243 (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
8	H	0	1

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	758	G	N1-C2	11.00	1.46	1.37
1	A	758	G	C5-C4	10.29	1.45	1.38
1	A	758	G	C6-N1	9.39	1.46	1.39
1	A	758	G	C5-C6	-9.13	1.33	1.42
1	A	758	G	N9-C8	7.90	1.43	1.37

The worst 5 of 358 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	758	G	N1-C6-O6	34.05	140.33	119.90
1	A	758	G	C6-C5-N7	-23.82	116.11	130.40
1	A	758	G	C5-N7-C8	-22.43	93.09	104.30
1	A	758	G	N7-C8-N9	21.87	124.03	113.10
1	A	758	G	C4-C5-N7	21.09	119.24	110.80

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
8	H	90	GLY	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	32380	0	16346	436	0
2	B	1900	0	1951	62	0
3	C	1612	0	1677	56	0
4	D	1703	0	1763	45	0
5	E	1146	0	1207	38	0
6	F	843	0	857	27	0
7	G	1257	0	1296	29	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	H	1116	0	1177	25	0
9	I	1011	0	1043	46	0
10	J	792	0	835	43	0
11	K	885	0	904	21	0
12	L	970	0	1057	41	0
13	M	997	0	1072	37	0
14	N	492	0	529	28	0
15	O	734	0	771	12	0
16	P	700	0	720	28	0
17	Q	857	0	930	22	0
18	R	597	0	668	16	0
19	S	647	0	673	11	0
20	T	762	0	856	18	0
21	V	208	0	221	6	0
22	X	232	0	121	0	0
23	W	82	0	45	1	0
24	A	93	0	0	0	0
24	J	1	0	0	0	0
24	N	1	0	0	0	0
24	W	1	0	0	0	0
25	A	42	0	45	0	0
26	D	1	0	0	0	0
All	All	52062	0	36764	935	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 11.

The worst 5 of 935 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)
4:D:125:HIS:HD1	4:D:152:SER:HG	1.13	0.95
2:B:77:ALA:HB2	2:B:211:ILE:HD13	1.52	0.91
1:A:410:G:OP1	1:A:431:A:N6	2.07	0.87
8:H:86:ILE:HD11	8:H:136:GLU:HG3	1.57	0.86
1:A:542:G:OP1	4:D:10:ARG:NH2	2.10	0.84

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	232/256 (91%)	184 (79%)	34 (15%)	14 (6%)	1	12
3	C	204/239 (85%)	166 (81%)	25 (12%)	13 (6%)	1	10
4	D	206/209 (99%)	182 (88%)	17 (8%)	7 (3%)	3	24
5	E	148/162 (91%)	134 (90%)	12 (8%)	2 (1%)	11	41
6	F	99/101 (98%)	89 (90%)	9 (9%)	1 (1%)	15	49
7	G	153/156 (98%)	137 (90%)	12 (8%)	4 (3%)	5	29
8	H	136/138 (99%)	129 (95%)	7 (5%)	0	100	100
9	I	125/128 (98%)	102 (82%)	11 (9%)	12 (10%)	0	4
10	J	96/105 (91%)	68 (71%)	15 (16%)	13 (14%)	0	1
11	K	117/129 (91%)	104 (89%)	12 (10%)	1 (1%)	17	51
12	L	122/132 (92%)	100 (82%)	16 (13%)	6 (5%)	2	16
13	M	123/126 (98%)	98 (80%)	20 (16%)	5 (4%)	3	19
14	N	58/61 (95%)	46 (79%)	9 (16%)	3 (5%)	2	14
15	O	86/89 (97%)	76 (88%)	9 (10%)	1 (1%)	13	45
16	P	81/88 (92%)	69 (85%)	9 (11%)	3 (4%)	3	22
17	Q	102/105 (97%)	86 (84%)	11 (11%)	5 (5%)	2	16
18	R	71/88 (81%)	57 (80%)	11 (16%)	3 (4%)	3	19
19	S	78/93 (84%)	56 (72%)	15 (19%)	7 (9%)	1	5
20	T	97/106 (92%)	78 (80%)	12 (12%)	7 (7%)	1	8
21	V	22/27 (82%)	18 (82%)	4 (18%)	0	100	100
All	All	2356/2538 (93%)	1979 (84%)	270 (12%)	107 (4%)	2	17

5 of 107 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	91	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
4	D	23	GLY
6	F	72	VAL
9	I	33	PHE
10	J	34	VAL

### 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	202/220 (92%)	174 (86%)	28 (14%)	3 16
3	C	160/188 (85%)	132 (82%)	28 (18%)	2 8
4	D	180/181 (99%)	158 (88%)	22 (12%)	5 20
5	E	115/123 (94%)	97 (84%)	18 (16%)	2 12
6	F	90/90 (100%)	78 (87%)	12 (13%)	4 17
7	G	126/127 (99%)	108 (86%)	18 (14%)	3 15
8	H	119/119 (100%)	107 (90%)	12 (10%)	7 29
9	I	98/99 (99%)	89 (91%)	9 (9%)	9 33
10	J	87/92 (95%)	71 (82%)	16 (18%)	1 7
11	K	90/99 (91%)	79 (88%)	11 (12%)	5 20
12	L	104/109 (95%)	91 (88%)	13 (12%)	4 19
13	M	100/101 (99%)	83 (83%)	17 (17%)	2 9
14	N	49/50 (98%)	41 (84%)	8 (16%)	2 10
15	O	79/80 (99%)	64 (81%)	15 (19%)	1 6
16	P	72/74 (97%)	64 (89%)	8 (11%)	6 25
17	Q	96/97 (99%)	86 (90%)	10 (10%)	7 28
18	R	64/77 (83%)	60 (94%)	4 (6%)	18 49
19	S	71/80 (89%)	58 (82%)	13 (18%)	1 7
20	T	75/82 (92%)	66 (88%)	9 (12%)	5 21
21	V	19/22 (86%)	16 (84%)	3 (16%)	2 12

*Continued on next page...*



Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1996/2110 (95%)	1722 (86%)	274 (14%)	<b>3</b> <b>17</b>

5 of 274 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
16	P	62	VAL
17	Q	52	LYS
19	S	79	THR
6	F	25	ILE
5	E	150	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	16	HIS
3	C	6	HIS
3	C	139	GLN
16	P	82	GLN
19	S	47	HIS

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1506/1507 (99%)	299 (19%)	46 (3%)
22	X	9/17 (52%)	0	0
23	W	3/6 (50%)	0	0
All	All	1518/1530 (99%)	299 (19%)	46 (3%)

5 of 299 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	7	G
1	A	9	G
1	A	32	A
1	A	39	G
1	A	47	C

5 of 46 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	1126	U
1	A	1224	G
1	A	1145	C
1	A	1190	G
1	A	1285	A

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

1 non-standard protein/DNA/RNA residue is 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
22	RSP	X	32	22	17,21,22	0.78	0	22,30,33	0.69	1 (4%)

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
22	RSP	X	32	22	-	2/7/25/26	0/2/2/2

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
22	X	32	RSP	S2-C2-N3	-2.01	118.01	121.49

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
22	X	32	RSP	O4'-C4'-C5'-O5'
22	X	32	RSP	C3'-C4'-C5'-O5'

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 98 ligands modelled in this entry, 97 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
25	PAR	A	1693	-	45,45,45	1.56	10 (22%)	64,67,67	1.65	13 (20%)

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
25	PAR	A	1693	-	-	3/18/94/94	0/4/4/4

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
25	A	1693	PAR	C64-C54	5.13	1.59	1.52
25	A	1693	PAR	C52-C42	3.39	1.59	1.52
25	A	1693	PAR	C13-C23	2.74	1.56	1.52
25	A	1693	PAR	O54-C14	2.68	1.48	1.41
25	A	1693	PAR	O43-C13	2.30	1.45	1.41

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
25	A	1693	PAR	O33-C14-C24	6.20	118.90	108.22
25	A	1693	PAR	O52-C13-C23	3.38	114.96	107.96
25	A	1693	PAR	O54-C54-C44	3.10	115.33	109.69
25	A	1693	PAR	C34-C24-N24	-2.98	104.95	111.05
25	A	1693	PAR	O34-C34-C44	-2.86	103.74	110.35

There are no chirality outliers.

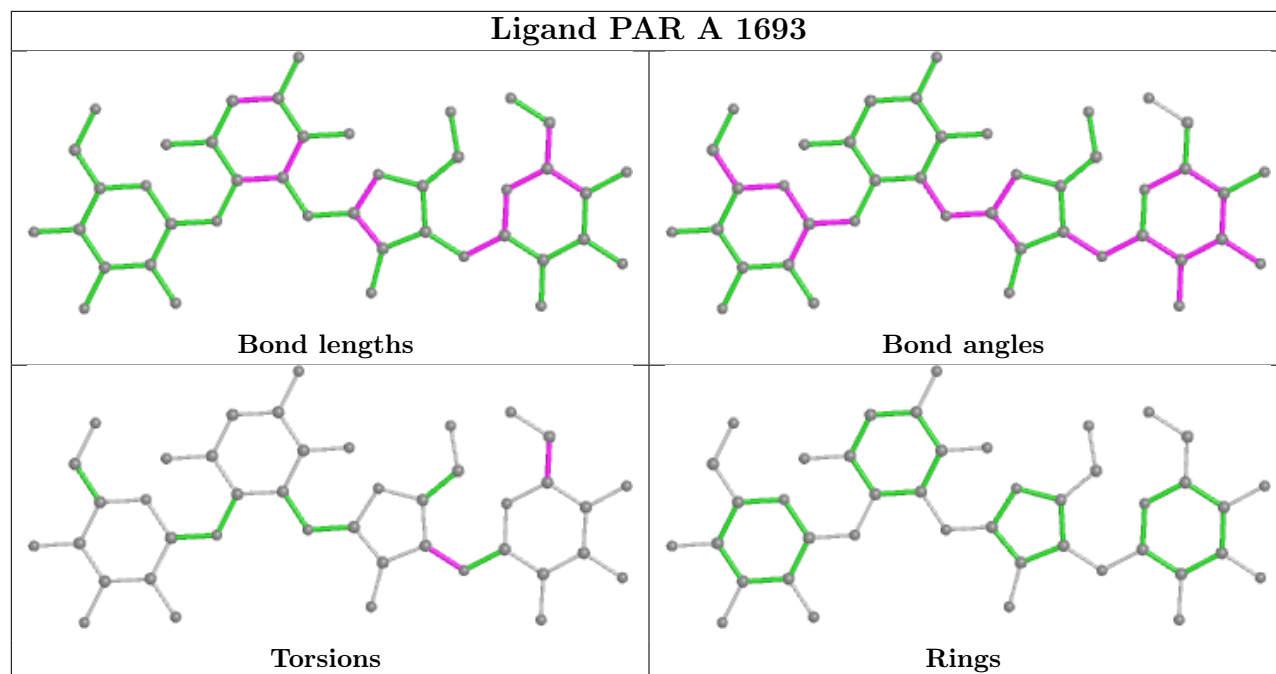
All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
25	A	1693	PAR	C44-C54-C64-N64
25	A	1693	PAR	O54-C54-C64-N64
25	A	1693	PAR	C23-C33-O33-C14

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	1507/1507 (100%)	0.05	23 (1%) 73 72	69, 107, 204, 419	0
2	B	234/256 (91%)	0.09	5 (2%) 63 63	91, 147, 211, 278	0
3	C	206/239 (86%)	0.01	4 (1%) 66 66	87, 138, 201, 253	0
4	D	208/209 (99%)	0.05	5 (2%) 59 57	78, 120, 175, 210	0
5	E	150/162 (92%)	-0.10	0 100 100	71, 102, 147, 189	0
6	F	101/101 (100%)	-0.15	1 (0%) 82 83	98, 137, 178, 233	0
7	G	155/156 (99%)	-0.11	2 (1%) 77 78	89, 132, 192, 229	0
8	H	138/138 (100%)	-0.13	0 100 100	55, 96, 139, 190	0
9	I	127/128 (99%)	0.36	4 (3%) 49 48	78, 142, 183, 208	0
10	J	98/105 (93%)	0.94	22 (22%) 0 0	89, 170, 259, 294	0
11	K	119/129 (92%)	0.18	3 (2%) 57 56	69, 110, 169, 283	0
12	L	124/132 (93%)	0.17	3 (2%) 59 57	61, 110, 166, 216	0
13	M	125/126 (99%)	0.76	10 (8%) 12 12	98, 139, 217, 295	0
14	N	60/61 (98%)	0.06	2 (3%) 46 45	94, 125, 176, 244	0
15	O	88/89 (98%)	0.14	3 (3%) 45 44	72, 115, 177, 236	0
16	P	83/88 (94%)	0.33	1 (1%) 79 79	71, 99, 121, 215	0
17	Q	104/105 (99%)	0.68	6 (5%) 23 24	71, 107, 185, 241	0
18	R	73/88 (82%)	0.17	3 (4%) 37 36	89, 119, 198, 283	0
19	S	80/93 (86%)	0.36	3 (3%) 40 39	103, 161, 199, 248	0
20	T	99/106 (93%)	0.20	2 (2%) 65 64	80, 106, 172, 200	0
21	V	24/27 (88%)	0.78	2 (8%) 11 11	98, 129, 165, 197	0
22	X	9/17 (52%)	0.76	1 (11%) 5 5	121, 149, 208, 277	0
23	W	4/6 (66%)	0.61	0 100 100	112, 114, 131, 149	0
All	All	3916/4068 (96%)	0.14	105 (2%) 54 53	55, 118, 201, 419	0

The worst 5 of 105 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	416	G	19.8
13	M	123	ALA	18.7
1	A	413	G	16.4
17	Q	103	GLY	14.1
1	A	422	C	14.0

## 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. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
22	RSP	X	32	20/21	0.72	0.44	238,239,242,242	0

## 6.3 Carbohydrates [i](#)

There are no monosaccharides 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. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
24	MG	A	1669	1/1	0.28	1.44	98,98,98,98	0
24	MG	A	1604	1/1	0.43	0.21	97,97,97,97	0
24	MG	A	1671	1/1	0.47	0.57	121,121,121,121	0
24	MG	A	1680	1/1	0.54	1.19	110,110,110,110	0
24	MG	A	1609	1/1	0.59	0.93	96,96,96,96	0
24	MG	A	1650	1/1	0.62	0.67	113,113,113,113	0
24	MG	A	1624	1/1	0.68	1.09	87,87,87,87	0
24	MG	A	1667	1/1	0.69	0.91	109,109,109,109	0
24	MG	A	1607	1/1	0.69	0.37	106,106,106,106	0
24	MG	A	1670	1/1	0.71	0.68	99,99,99,99	0
24	MG	A	1603	1/1	0.72	0.53	102,102,102,102	0

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
24	MG	A	1691	1/1	0.74	1.64	98,98,98,98	0
24	MG	A	1690	1/1	0.75	0.29	97,97,97,97	0
24	MG	A	1605	1/1	0.75	1.38	81,81,81,81	0
24	MG	A	1666	1/1	0.76	0.92	78,78,78,78	0
24	MG	A	1634	1/1	0.76	1.18	96,96,96,96	0
24	MG	A	1676	1/1	0.77	0.58	87,87,87,87	0
24	MG	N	101	1/1	0.78	0.43	94,94,94,94	0
24	MG	A	1606	1/1	0.79	0.46	106,106,106,106	0
24	MG	A	1682	1/1	0.80	0.31	60,60,60,60	0
24	MG	A	1647	1/1	0.80	0.30	67,67,67,67	0
24	MG	A	1692	1/1	0.82	0.65	98,98,98,98	0
24	MG	A	1675	1/1	0.82	0.69	98,98,98,98	0
24	MG	A	1672	1/1	0.83	0.32	84,84,84,84	0
24	MG	A	1674	1/1	0.84	0.60	87,87,87,87	0
24	MG	A	1613	1/1	0.84	0.72	106,106,106,106	0
24	MG	A	1640	1/1	0.84	0.13	93,93,93,93	0
24	MG	A	1631	1/1	0.86	1.09	85,85,85,85	0
24	MG	A	1665	1/1	0.86	0.42	101,101,101,101	0
24	MG	A	1668	1/1	0.86	0.85	111,111,111,111	0
24	MG	J	201	1/1	0.86	0.37	82,82,82,82	0
24	MG	A	1685	1/1	0.86	0.28	84,84,84,84	0
24	MG	A	1614	1/1	0.87	1.35	80,80,80,80	0
24	MG	A	1663	1/1	0.87	0.10	96,96,96,96	0
24	MG	A	1628	1/1	0.87	0.84	68,68,68,68	0
24	MG	A	1646	1/1	0.88	0.93	72,72,72,72	0
24	MG	A	1679	1/1	0.88	0.94	99,99,99,99	0
24	MG	A	1661	1/1	0.89	0.18	58,58,58,58	0
24	MG	A	1625	1/1	0.90	0.52	62,62,62,62	0
24	MG	A	1645	1/1	0.90	0.57	71,71,71,71	0
24	MG	A	1601	1/1	0.90	0.69	100,100,100,100	0
24	MG	A	1687	1/1	0.91	0.59	99,99,99,99	0
24	MG	A	1626	1/1	0.91	0.13	103,103,103,103	0
24	MG	A	1681	1/1	0.91	0.76	106,106,106,106	0
24	MG	A	1673	1/1	0.91	0.15	57,57,57,57	0
24	MG	A	1621	1/1	0.91	0.49	78,78,78,78	0
24	MG	A	1686	1/1	0.91	0.90	87,87,87,87	0
24	MG	A	1653	1/1	0.92	0.48	62,62,62,62	0
24	MG	A	1654	1/1	0.92	0.72	66,66,66,66	0
24	MG	A	1637	1/1	0.92	0.33	83,83,83,83	0
24	MG	A	1610	1/1	0.92	0.30	108,108,108,108	0
24	MG	W	101	1/1	0.92	0.54	105,105,105,105	0
24	MG	A	1684	1/1	0.93	0.26	94,94,94,94	0

Continued on next page...



*Continued from previous page...*

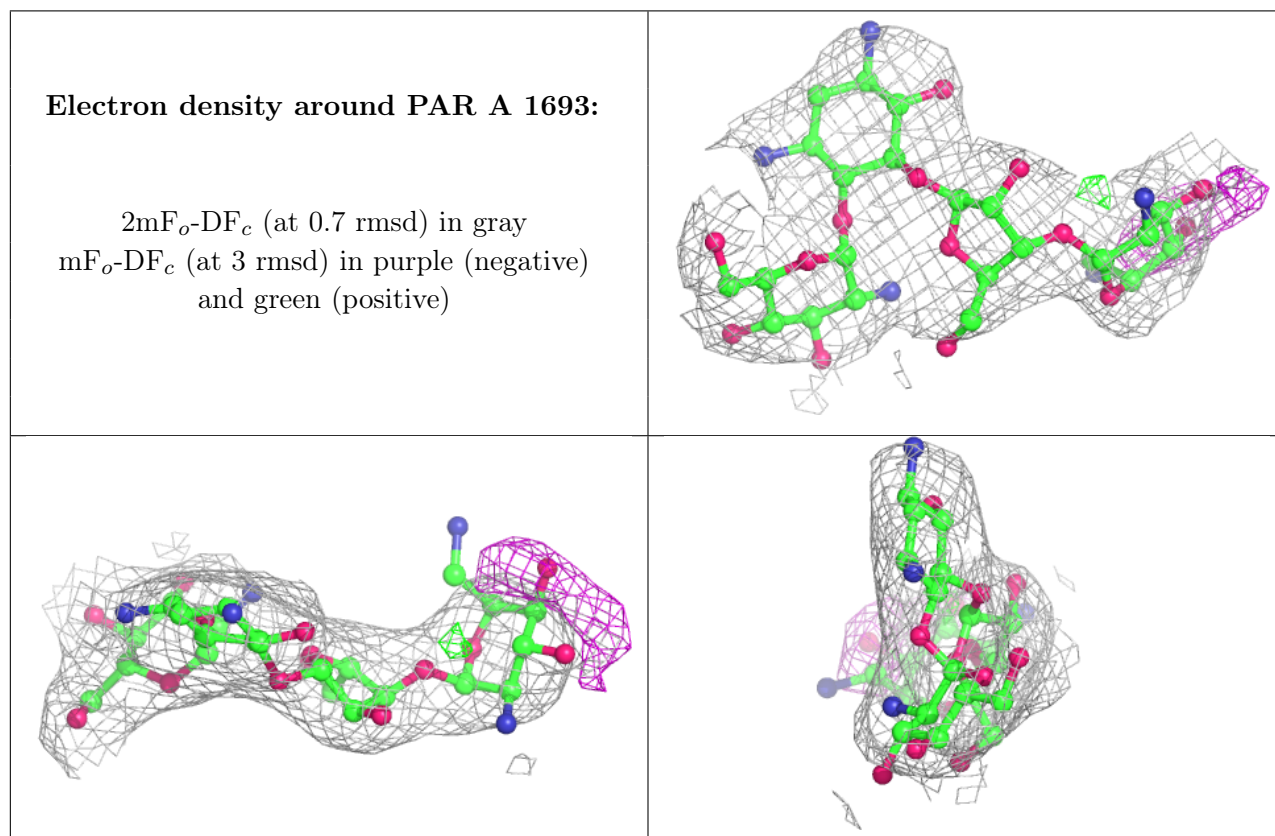
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
24	MG	A	1608	1/1	0.93	0.22	85,85,85,85	0
24	MG	A	1638	1/1	0.93	0.71	62,62,62,62	0
24	MG	A	1615	1/1	0.93	0.31	94,94,94,94	0
24	MG	A	1643	1/1	0.93	0.29	80,80,80,80	0
24	MG	A	1677	1/1	0.93	0.28	92,92,92,92	0
24	MG	A	1659	1/1	0.93	0.40	108,108,108,108	0
24	MG	A	1660	1/1	0.93	0.73	107,107,107,107	0
24	MG	A	1633	1/1	0.93	0.43	70,70,70,70	0
24	MG	A	1627	1/1	0.93	0.36	57,57,57,57	0
25	PAR	A	1693	42/42	0.93	0.23	90,95,99,100	0
24	MG	A	1611	1/1	0.94	0.42	76,76,76,76	0
24	MG	A	1632	1/1	0.94	0.67	66,66,66,66	0
24	MG	A	1619	1/1	0.94	0.57	83,83,83,83	0
24	MG	A	1664	1/1	0.94	0.19	88,88,88,88	0
24	MG	A	1689	1/1	0.95	0.26	123,123,123,123	0
24	MG	A	1657	1/1	0.95	0.37	65,65,65,65	0
24	MG	A	1623	1/1	0.95	0.51	56,56,56,56	0
24	MG	A	1629	1/1	0.95	0.44	69,69,69,69	0
24	MG	A	1651	1/1	0.96	0.37	78,78,78,78	0
24	MG	A	1635	1/1	0.96	0.46	76,76,76,76	0
24	MG	A	1644	1/1	0.96	0.19	93,93,93,93	0
24	MG	A	1655	1/1	0.96	0.19	90,90,90,90	0
24	MG	A	1656	1/1	0.96	0.39	92,92,92,92	0
24	MG	A	1678	1/1	0.96	0.25	77,77,77,77	0
24	MG	A	1617	1/1	0.96	0.54	66,66,66,66	0
24	MG	A	1658	1/1	0.96	1.14	86,86,86,86	0
24	MG	A	1618	1/1	0.96	0.47	68,68,68,68	0
24	MG	A	1616	1/1	0.96	0.39	54,54,54,54	0
24	MG	A	1641	1/1	0.96	0.38	77,77,77,77	0
24	MG	A	1639	1/1	0.97	0.29	61,61,61,61	0
24	MG	A	1602	1/1	0.97	0.56	83,83,83,83	0
24	MG	A	1636	1/1	0.97	0.36	76,76,76,76	0
24	MG	A	1662	1/1	0.97	0.35	74,74,74,74	0
24	MG	A	1652	1/1	0.97	0.15	65,65,65,65	0
24	MG	A	1642	1/1	0.97	0.16	91,91,91,91	0
24	MG	A	1630	1/1	0.98	0.30	85,85,85,85	0
24	MG	A	1612	1/1	0.98	0.64	83,83,83,83	0
24	MG	A	1694	1/1	0.98	0.62	83,83,83,83	0
24	MG	A	1622	1/1	0.98	0.34	69,69,69,69	0
24	MG	A	1688	1/1	0.98	0.49	83,83,83,83	0
24	MG	A	1683	1/1	0.98	0.29	76,76,76,76	0
24	MG	A	1620	1/1	0.98	0.35	58,58,58,58	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
24	MG	A	1648	1/1	0.99	0.17	66,66,66,66	0
24	MG	A	1649	1/1	0.99	0.24	70,70,70,70	0
26	ZN	D	301	1/1	0.99	0.24	118,118,118,118	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.