



## Full wwPDB EM Validation Report ⓘ

Mar 6, 2025 – 06:31 pm GMT

PDB ID : 8P2H  
EMDB ID : EMD-17365  
Title : Staphylococcus aureus 70S ribosome with elongation factor G locked with fusidic acid with a tRNA in pe/E chimeric state  
Authors : Gonzalez-Lopez, A.; Selmer, M.  
Deposited on : 2023-05-16  
Resolution : 2.49 Å (reported)  
Based on initial models : 7NHM, 2XEX, 7SSD

This is a Full wwPDB EM Validation 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/EMValidationReportHelp>  
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 ⓘ](#)) were used in the production of this report:

EMDB validation analysis : **FAILED**  
Mogul : 1.8.4, CSD as541be (2020)  
MolProbity : 4.02b-467  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
MapQ : **FAILED**  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.41

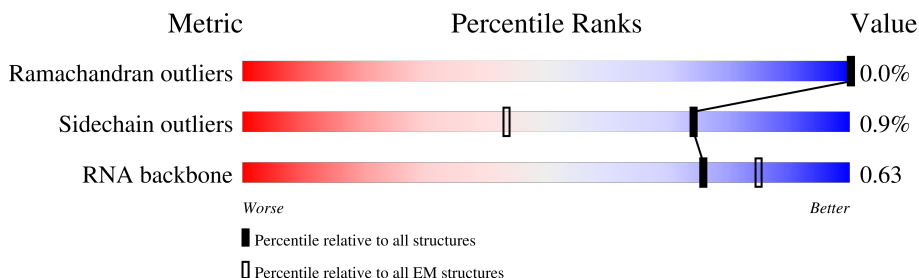
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*



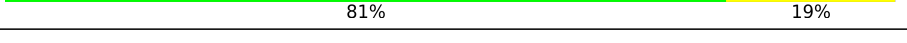
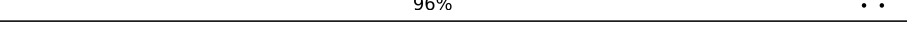
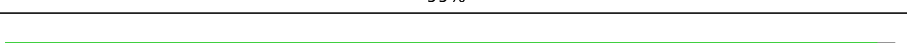

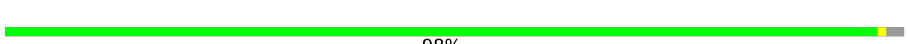
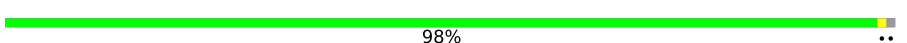

The reported resolution of this entry is 2.49 Å.

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.



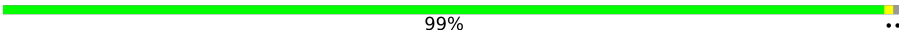
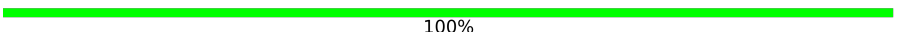
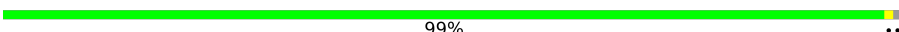

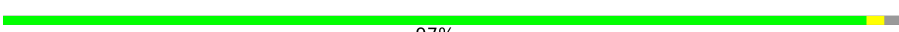







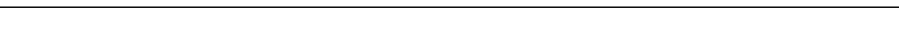

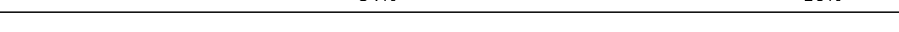
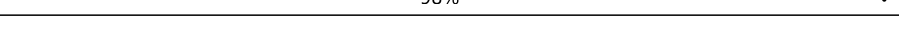
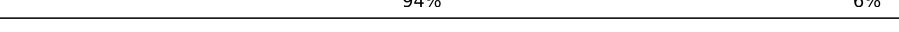
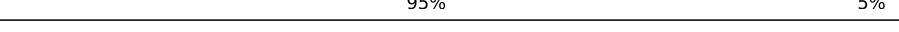

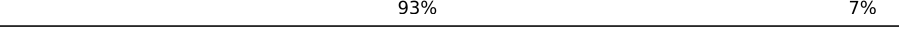
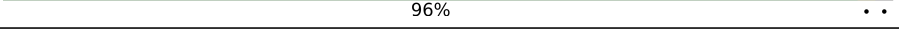

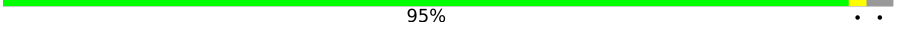
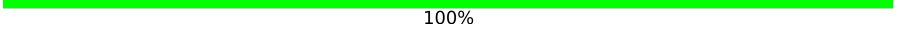

Metric	Whole archive (#Entries)	EM structures (#Entries)
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415
RNA backbone	6643	2191

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	2923	 86%11%.
2	B	115	 84%14%.
3	D	77	 81%19%.
4	E	693	 96%..
5	G	277	 99%.
6	H	220	 98%.
7	I	207	 99%.
8	J	179	 98%..
9	K	178	 98%..

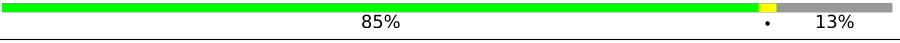
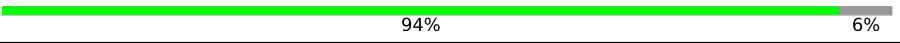
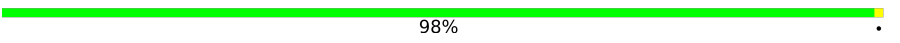
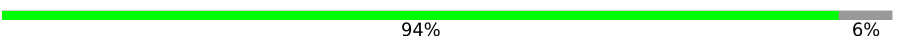
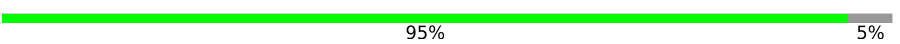

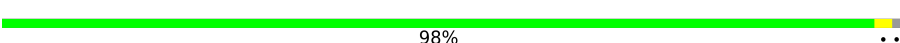
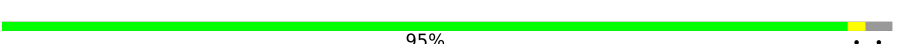
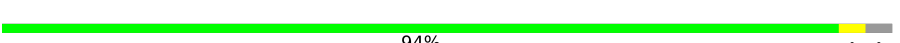

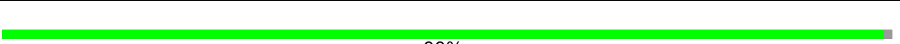
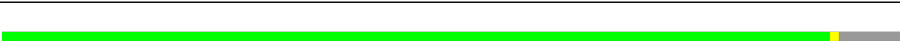



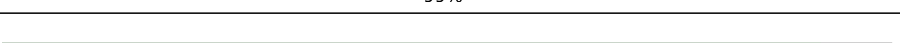
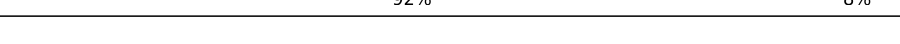
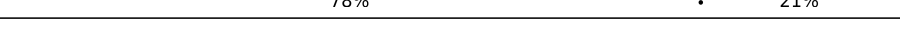

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	M	145	 99% ..
11	N	122	 100%
12	O	146	 99% ..
13	P	144	 94% • 6%
14	Q	122	 97% ..
15	R	119	 100%
16	S	116	 97% ..
17	T	118	 98% •
18	U	102	 98% ..
19	V	117	 93% • 5%
20	W	91	 98% •
21	X	105	 97% •
22	Y	217	 43% 57%
23	Z	94	 84% 16%
24	1	62	 98% •
25	2	69	 94% 6%
26	3	59	 95% 5%
27	4	84	 89% • 8%
28	5	57	 93% 7%
29	6	49	 96% ..
30	7	45	 93% • •
31	8	66	 95% • •
32	9	37	 100%
33	a	1555	 86% 13% •
34	b	24	 58% 8% 33%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
35	c	255	
36	d	217	
37	e	200	
38	f	166	
39	g	98	
40	h	156	
41	i	132	
42	j	132	
43	k	102	
44	l	129	
45	m	137	
46	n	121	
47	o	61	
48	p	89	
49	q	91	
50	r	87	
51	s	80	
52	t	92	
53	u	83	

## 2 Entry composition

There are 59 unique types of molecules in this entry. The entry contains 146530 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a RNA chain called 23S ribosomal RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	2844	Total	C	N	O	P	3	0
			61041	27258	11160	19776	2847		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	48	G	A	conflict	GB CP000253
A	1584	U	A	conflict	GB CP000253
A	2261	G	A	conflict	GB CP000253

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	113	Total	C	N	O	P	0	0
			2408	1076	431	788	113		

- Molecule 3 is a RNA chain called tRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	77	Total	C	N	O	P	0	0
			1640	732	297	535	76		

- Molecule 4 is a protein called Elongation factor G.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	E	669	Total	C	N	O	S	0	0
			5177	3248	868	1032	29		

- Molecule 5 is a protein called 50S ribosomal protein L2.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	G	273	Total	C	N	O	S	0	0
			2085	1297	413	370	5		

- Molecule 6 is a protein called 50S ribosomal protein L3.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	H	216	Total	C	N	O	S	0	0
			1637	1024	301	307	5		

- Molecule 7 is a protein called 50S ribosomal protein L4.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	I	204	Total	C	N	O	S	0	0
			1564	981	286	295	2		

- Molecule 8 is a protein called 50S ribosomal protein L5.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	J	176	Total	C	N	O	S	0	0
			1392	885	238	262	7		

- Molecule 9 is a protein called 50S ribosomal protein L6.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	K	176	Total	C	N	O	S	0	0
			1372	852	251	266	3		

- Molecule 10 is a protein called 50S ribosomal protein L13.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	M	144	Total	C	N	O	S	0	0
			1146	715	210	218	3		

- Molecule 11 is a protein called 50S ribosomal protein L14.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	N	122	Total	C	N	O	S	0	0
			920	572	174	170	4		

- Molecule 12 is a protein called 50S ribosomal protein L15.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	O	145	Total	C	N	O	S	0	0
			1090	674	214	201	1		

- Molecule 13 is a protein called 50S ribosomal protein L16.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	P	136	Total	C	N	O	S	0	0
			1089	698	206	181	4		

- Molecule 14 is a protein called 50S ribosomal protein L17.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	Q	120	Total	C	N	O	S	0	0
			952	584	182	185	1		

- Molecule 15 is a protein called 50S ribosomal protein L18.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	R	119	Total	C	N	O	S	0	0
			922	574	174	173	1		

- Molecule 16 is a protein called 50S ribosomal protein L19.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	S	114	Total	C	N	O	S	0	0
			922	580	185	157			

- Molecule 17 is a protein called 50S ribosomal protein L20.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	T	116	Total	C	N	O	S	0	0
			943	593	189	157	4		

- Molecule 18 is a protein called 50S ribosomal protein L21.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	U	101	Total	C	N	O	S	0	0
			793	503	141	148	1		

- Molecule 19 is a protein called 50S ribosomal protein L22.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	V	111	Total	C	N	O	S	0	0
			853	532	163	155	3		

- Molecule 20 is a protein called 50S ribosomal protein L23.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	W	89	Total	C	N	O	S	0	0
			725	457	130	134	4		

- Molecule 21 is a protein called 50S ribosomal protein L24.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	X	102	Total	C	N	O	S	0	0
			787	497	144	144	2		

- Molecule 22 is a protein called 50S ribosomal protein L25.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	Y	94	Total	C	N	O	S	0	0
			738	471	131	134	2		

- Molecule 23 is a protein called 50S ribosomal protein L27.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	Z	79	Total	C	N	O		0	0
			604	372	117	115			

- Molecule 24 is a protein called 50S ribosomal protein L28.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	1	61	Total	C	N	O	S	0	0
			481	298	104	78	1		

- Molecule 25 is a protein called 50S ribosomal protein L29.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	2	65	Total	C	N	O	S	0	0
			535	329	100	105	1		

- Molecule 26 is a protein called 50S ribosomal protein L30.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	3	56	Total	C	N	O		0	0
			436	271	82	83			

- Molecule 27 is a protein called 50S ribosomal protein L31 type B.



Mol	Chain	Residues	Atoms					AltConf	Trace
27	4	77	Total	C	N	O	S	0	0
			631	400	108	120	3		

- Molecule 28 is a protein called Large ribosomal subunit protein bL32.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	5	53	Total	C	N	O	S	0	0
			422	256	86	75	5		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
5	?	-	ARG	deletion	UNP Q2FZF1

- Molecule 29 is a protein called Large ribosomal subunit protein bL33A.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	6	48	Total	C	N	O	S	0	0
			402	245	79	73	5		

- Molecule 30 is a protein called 50S ribosomal protein L34.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	7	43	Total	C	N	O	S	0	0
			367	225	89	52	1		

- Molecule 31 is a protein called 50S ribosomal protein L35.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	8	64	Total	C	N	O	S	0	0
			521	324	113	82	2		

- Molecule 32 is a protein called 50S ribosomal protein L36.

Mol	Chain	Residues	Atoms					AltConf	Trace
32	9	37	Total	C	N	O	S	0	0
			296	186	60	45	5		

- Molecule 33 is a RNA chain called 16S ribosomal RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
33	a	1534	Total	C	N	O	P	0	0
			32869	14681	5998	10656	1534		

- Molecule 34 is a RNA chain called mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
34	b	16	Total	C	N	O	P	0	0
			354	158	74	106	16		

- Molecule 35 is a protein called 30S ribosomal protein S2.

Mol	Chain	Residues	Atoms					AltConf	Trace
35	c	221	Total	C	N	O	S	0	0
			1781	1136	310	328	7		

- Molecule 36 is a protein called 30S ribosomal protein S3.

Mol	Chain	Residues	Atoms					AltConf	Trace
36	d	204	Total	C	N	O	S	0	0
			1612	1015	302	293	2		

- Molecule 37 is a protein called 30S ribosomal protein S4.

Mol	Chain	Residues	Atoms					AltConf	Trace
37	e	199	Total	C	N	O	S	0	0
			1617	1020	302	293	2		

- Molecule 38 is a protein called 30S ribosomal protein S5.

Mol	Chain	Residues	Atoms					AltConf	Trace
38	f	156	Total	C	N	O	S	0	0
			1160	730	213	215	2		

- Molecule 39 is a protein called 30S ribosomal protein S6.

Mol	Chain	Residues	Atoms					AltConf	Trace
39	g	93	Total	C	N	O	S	0	0
			773	489	136	146	2		

- Molecule 40 is a protein called 30S ribosomal protein S7.

Mol	Chain	Residues	Atoms					AltConf	Trace
40	h	145	Total	C	N	O	S	0	0
			1155	721	219	211	4		

- Molecule 41 is a protein called 30S ribosomal protein S8.

Mol	Chain	Residues	Atoms					AltConf	Trace
41	i	131	Total	C	N	O	S	0	0
			1032	652	183	193	4		

- Molecule 42 is a protein called 30S ribosomal protein S9.

Mol	Chain	Residues	Atoms					AltConf	Trace
42	j	128	Total	C	N	O	S	0	0
			1016	629	203	183	1		

- Molecule 43 is a protein called Small ribosomal subunit protein uS10.

Mol	Chain	Residues	Atoms					AltConf	Trace
43	k	99	Total	C	N	O	S	0	0
			792	499	145	147	1		

- Molecule 44 is a protein called 30S ribosomal protein S11.

Mol	Chain	Residues	Atoms					AltConf	Trace
44	l	118	Total	C	N	O	S	0	0
			876	542	166	165	3		

- Molecule 45 is a protein called 30S ribosomal protein S12.

Mol	Chain	Residues	Atoms					AltConf	Trace
45	m	135	Total	C	N	O	S	0	0
			1058	658	214	184	2		

- Molecule 46 is a protein called 30S ribosomal protein S13.

Mol	Chain	Residues	Atoms					AltConf	Trace
46	n	113	Total	C	N	O	S	0	0
			902	554	179	168	1		

- Molecule 47 is a protein called 30S ribosomal protein S14 type Z.

Mol	Chain	Residues	Atoms					AltConf	Trace
47	o	60	Total	C	N	O	S	0	0
			501	317	100	79	5		

- Molecule 48 is a protein called 30S ribosomal protein S15.

Mol	Chain	Residues	Atoms					AltConf	Trace
48	p	86	Total	C	N	O	S	0	0
			721	445	148	127	1		

- Molecule 49 is a protein called 30S ribosomal protein S16.

Mol	Chain	Residues	Atoms					AltConf	Trace
49	q	90	Total	C	N	O	S	0	0
			712	448	132	131	1		

- Molecule 50 is a protein called 30S ribosomal protein S17.

Mol	Chain	Residues	Atoms					AltConf	Trace
50	r	80	Total	C	N	O	S	0	0
			662	419	120	122	1		

- Molecule 51 is a protein called 30S ribosomal protein S18.

Mol	Chain	Residues	Atoms					AltConf	Trace
51	s	63	Total	C	N	O	S	0	0
			516	330	96	87	3		

- Molecule 52 is a protein called 30S ribosomal protein S19.

Mol	Chain	Residues	Atoms					AltConf	Trace
52	t	80	Total	C	N	O	S	0	0
			651	419	117	113	2		

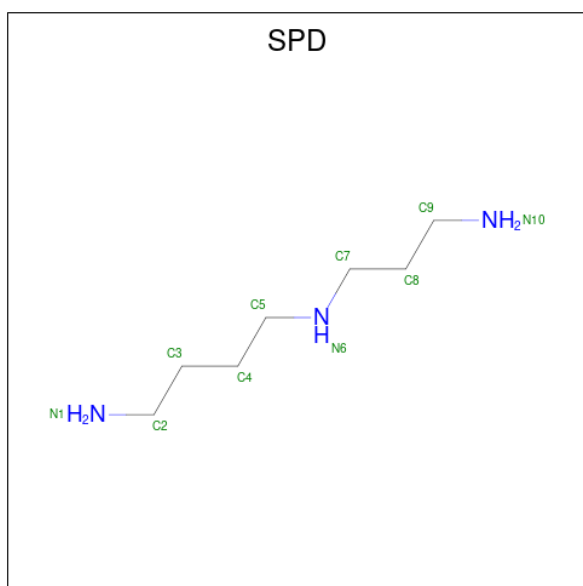
- Molecule 53 is a protein called 30S ribosomal protein S20.

Mol	Chain	Residues	Atoms					AltConf	Trace
53	u	80	Total	C	N	O	S	0	0
			606	367	119	118	2		

- Molecule 54 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

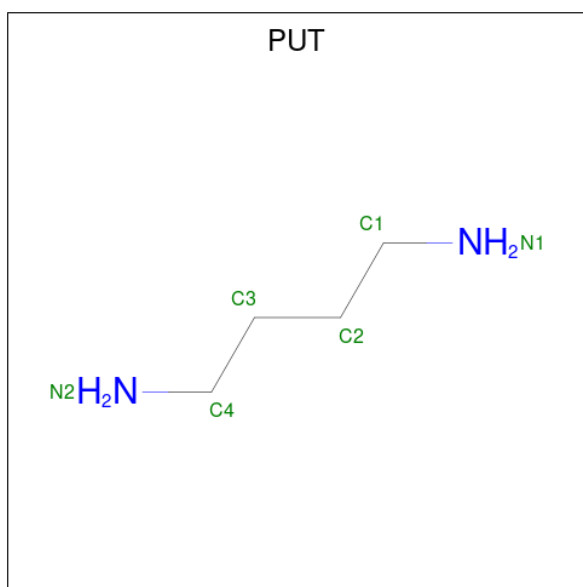
Mol	Chain	Residues	Atoms		AltConf
54	A	142	Total	Mg	0
			142	142	
54	B	1	Total	Mg	0
			1	1	
54	E	1	Total	Mg	0
			1	1	
54	G	2	Total	Mg	0
			2	2	
54	a	23	Total	Mg	0
			23	23	

- Molecule 55 is SPERMIDINE (three-letter code: SPD) (formula:  $C_7H_{19}N_3$ ).



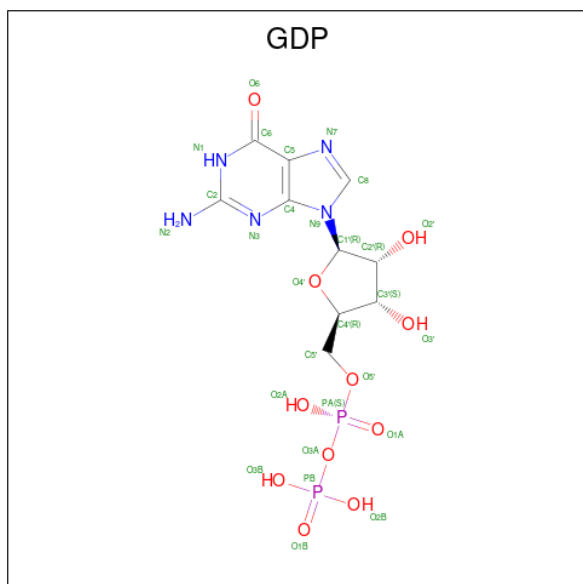
Mol	Chain	Residues	Atoms			AltConf
55	A	1	Total	C	N	0
			10	7	3	
55	A	1	Total	C	N	0
			10	7	3	
55	A	1	Total	C	N	0
			10	7	3	

- Molecule 56 is 1,4-DIAMINOBUTANE (three-letter code: PUT) (formula:  $C_4H_{12}N_2$ ).



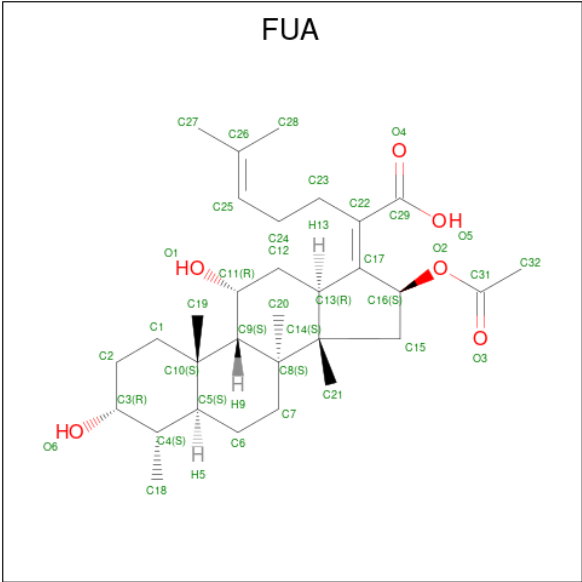
Mol	Chain	Residues	Atoms			AltConf
56	A	1	Total	C	N	0
			6	4	2	

- Molecule 57 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula:  $C_{10}H_{15}N_5O_{11}P_2$ ).



Mol	Chain	Residues	Atoms					AltConf
57	E	1	Total	C	N	O	P	0
			28	10	5	11	2	

- Molecule 58 is FUSIDIC ACID (three-letter code: FUA) (formula:  $C_{31}H_{48}O_6$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			AltConf
58	E	1	Total	C	O	0
			37	31	6	

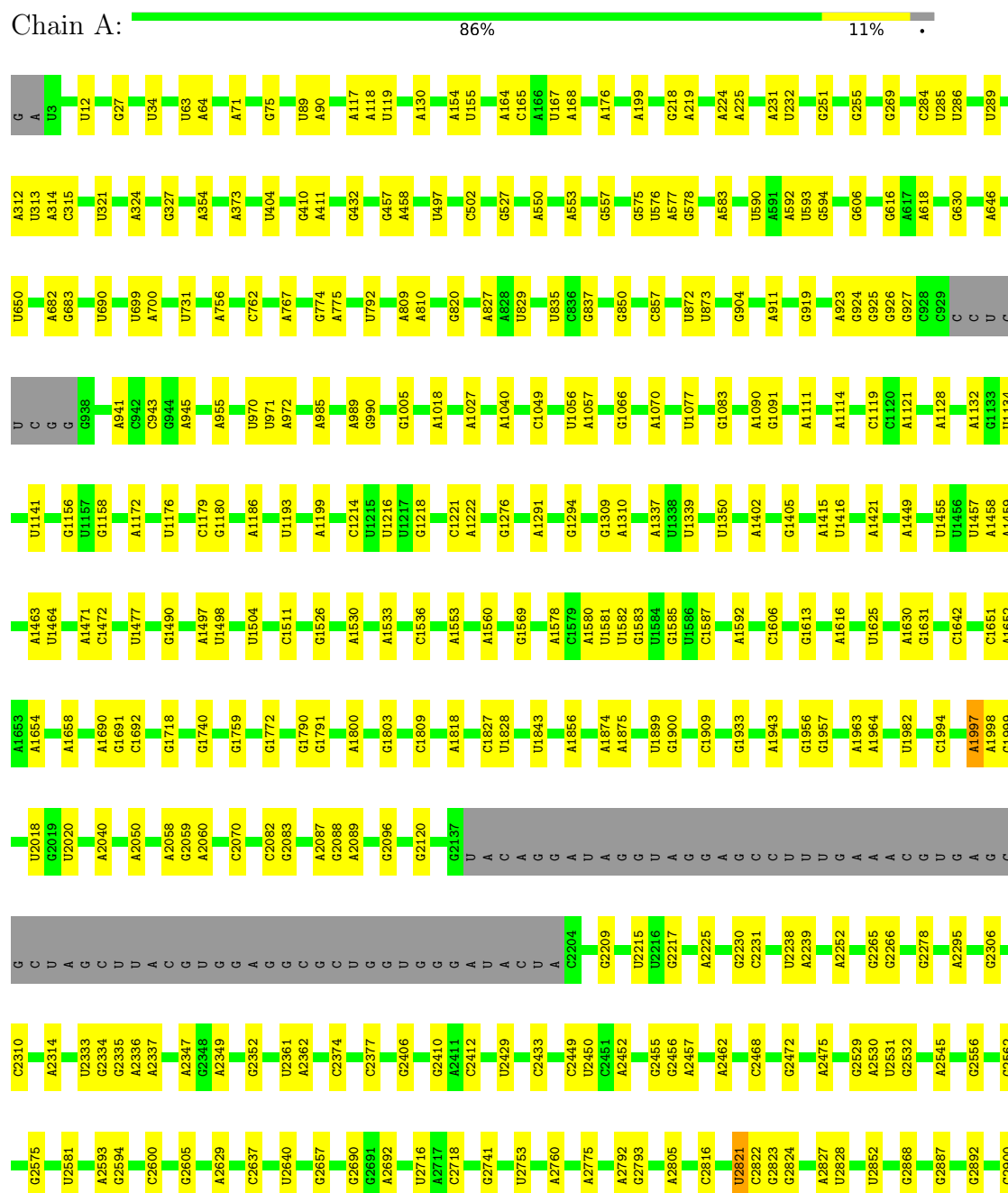
- Molecule 59 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
59	5	1	Total	Zn	0
			1	1	
59	9	1	Total	Zn	0
			1	1	
59	o	1	Total	Zn	0
			1	1	

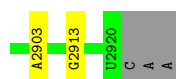
### 3 Residue-property plots [i](#)

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. 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. 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: 23S ribosomal RNA







- Molecule 2: 5S ribosomal RNA

Chain B: 84% 14%



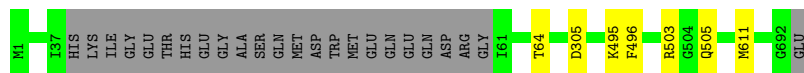
- Molecule 3: tRNA

Chain D: 81% 19%



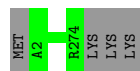
- Molecule 4: Elongation factor G

Chain E: 96%



- Molecule 5: 50S ribosomal protein L2

Chain G: 99%



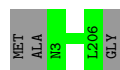
- Molecule 6: 50S ribosomal protein L3

Chain H: 98%



- Molecule 7: 50S ribosomal protein L4

Chain I: 99%



- Molecule 8: 50S ribosomal protein L5

Chain J: 98%



- Molecule 9: 50S ribosomal protein L6

Chain K: 98% ..



- Molecule 10: 50S ribosomal protein L13

Chain M: 99% ..



- Molecule 11: 50S ribosomal protein L14

Chain N: 100%

There are no outlier residues recorded for this chain.

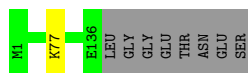
- Molecule 12: 50S ribosomal protein L15

Chain O: 99% ..



- Molecule 13: 50S ribosomal protein L16

Chain P: 94% • 6%



- Molecule 14: 50S ribosomal protein L17

Chain Q: 97% ..



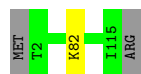
- Molecule 15: 50S ribosomal protein L18

Chain R: 100%

There are no outlier residues recorded for this chain.

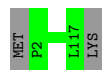
- Molecule 16: 50S ribosomal protein L19

Chain S:  97% ..



- Molecule 17: 50S ribosomal protein L20

Chain T:  98% .



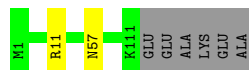
- Molecule 18: 50S ribosomal protein L21

Chain U:  98% ..



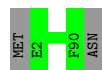
- Molecule 19: 50S ribosomal protein L22

Chain V:  93% • 5%



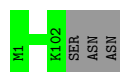
- Molecule 20: 50S ribosomal protein L23

Chain W:  98% .



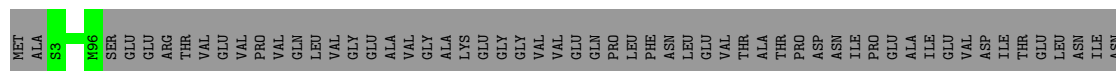
- Molecule 21: 50S ribosomal protein L24

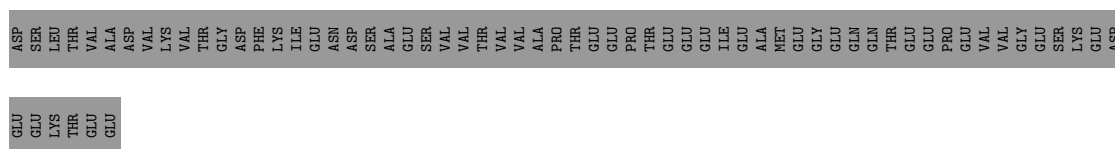
Chain X:  97% .



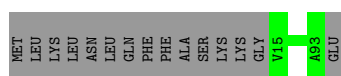
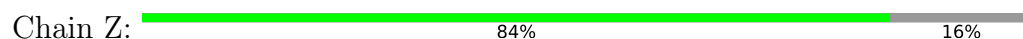
- Molecule 22: 50S ribosomal protein L25

Chain Y:  43% 57%





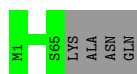
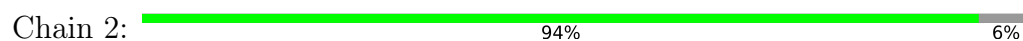
- Molecule 23: 50S ribosomal protein L27



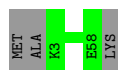
- Molecule 24: 50S ribosomal protein L28



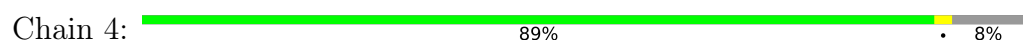
- Molecule 25: 50S ribosomal protein L29



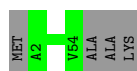
- Molecule 26: 50S ribosomal protein L30



- Molecule 27: 50S ribosomal protein L31 type B



- Molecule 28: Large ribosomal subunit protein bL32



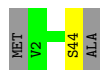
- Molecule 29: Large ribosomal subunit protein bL33A





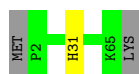
- Molecule 30: 50S ribosomal protein L34

Chain 7: 93%



- Molecule 31: 50S ribosomal protein L35

Chain 8: 95%



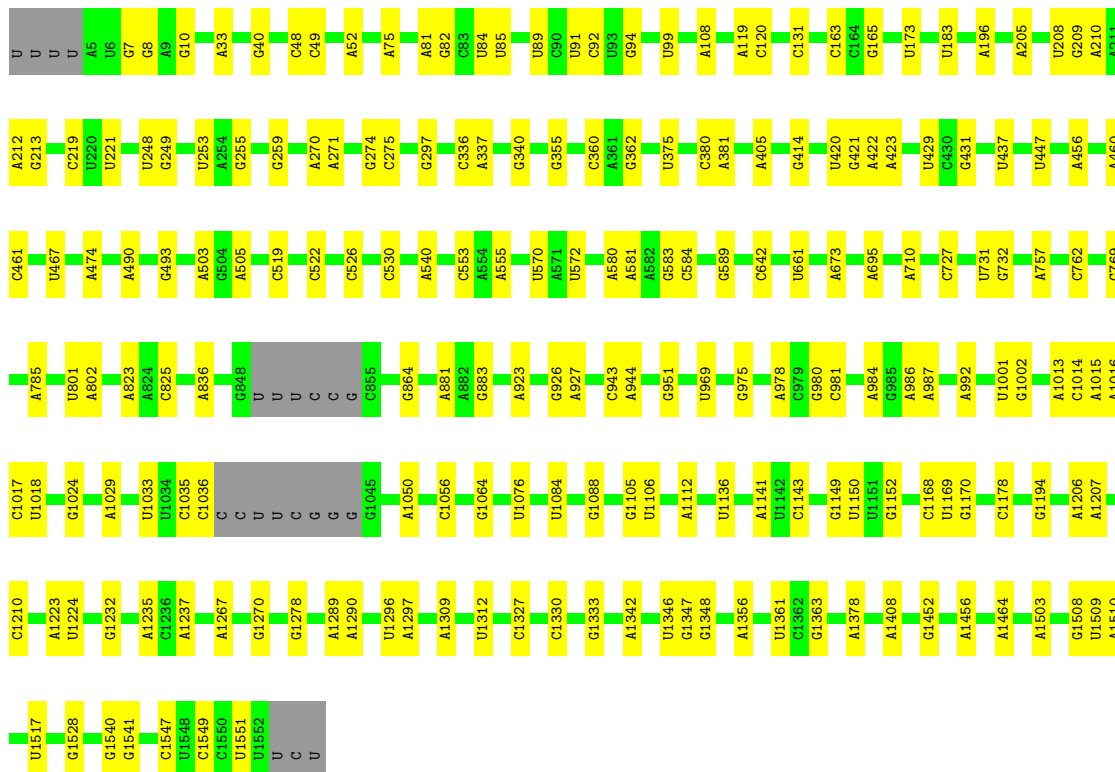
- Molecule 32: 50S ribosomal protein L36

Chain 9: 100%

There are no outlier residues recorded for this chain.

- Molecule 33: 16S ribosomal RNA

Chain a: 86% 13%




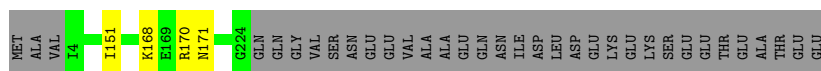
- Molecule 34: mRNA

Chain b:  58% 8% 33%



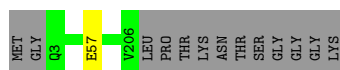
- Molecule 35: 30S ribosomal protein S2

Chain c:  85% 13%



- Molecule 36: 30S ribosomal protein S3

Chain d:  94% 6%



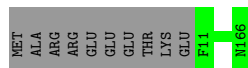
- Molecule 37: 30S ribosomal protein S4

Chain e:  98%



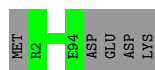
- Molecule 38: 30S ribosomal protein S5

Chain f:  94% 6%



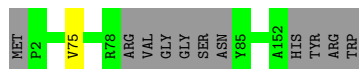
- Molecule 39: 30S ribosomal protein S6

Chain g:  95% 5%



- Molecule 40: 30S ribosomal protein S7

Chain h:  92% 7%



- Molecule 41: 30S ribosomal protein S8

Chain i:  98% ..



- Molecule 42: 30S ribosomal protein S9

Chain j:  95% ..




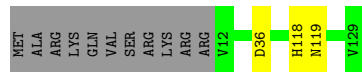
- Molecule 43: Small ribosomal subunit protein uS10

Chain k:  94% ..



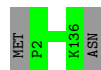
- Molecule 44: 30S ribosomal protein S11

Chain l:  89% 9%



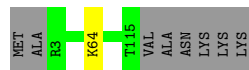
- Molecule 45: 30S ribosomal protein S12

Chain m:  99% .



- Molecule 46: 30S ribosomal protein S13

Chain n:  93% 7%



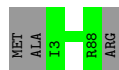
- Molecule 47: 30S ribosomal protein S14 type Z

Chain o:  98% .



- Molecule 48: 30S ribosomal protein S15

Chain p:  97% .




- Molecule 49: 30S ribosomal protein S16

Chain q:  99% .




- Molecule 50: 30S ribosomal protein S17

Chain r:  92% 8%




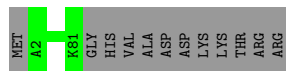
- Molecule 51: 30S ribosomal protein S18

Chain s:  78% . 21%



- Molecule 52: 30S ribosomal protein S19

Chain t:  87% 13%



- Molecule 53: 30S ribosomal protein S20

Chain u:  95% . .





## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	178679	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	30.83	Depositor
Minimum defocus (nm)	300	Depositor
Maximum defocus (nm)	2400	Depositor
Magnification	105000	Depositor
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GDP, PUT, SPD, 4OC, MA6, 5MC, ZN, UR3, FUA, MG, 2MG, G7M, OMG, H2U, 5MU, 3TD, 2MA

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	0.67	0/68163	0.84	9/106299 (0.0%)
2	B	0.49	0/2692	0.81	0/4193
3	D	0.37	0/1832	0.79	0/2855
4	E	0.26	0/5267	0.48	0/7123
5	G	0.34	0/2120	0.60	0/2847
6	H	0.34	0/1661	0.56	0/2227
7	I	0.31	0/1587	0.55	0/2143
8	J	0.25	0/1410	0.49	0/1893
9	K	0.28	0/1390	0.54	0/1870
10	M	0.35	0/1168	0.55	0/1573
11	N	0.32	0/927	0.60	0/1243
12	O	0.32	0/1104	0.57	0/1471
13	P	0.35	0/1113	0.59	0/1493
14	Q	0.32	0/956	0.60	0/1277
15	R	0.28	0/931	0.53	0/1244
16	S	0.31	0/934	0.63	0/1249
17	T	0.33	0/955	0.60	0/1265
18	U	0.34	0/803	0.55	0/1073
19	V	0.32	0/861	0.60	0/1159
20	W	0.32	0/733	0.58	0/978
21	X	0.29	0/796	0.49	0/1063
22	Y	0.29	0/746	0.53	0/1000
23	Z	0.33	0/610	0.61	0/811
24	1	0.30	0/487	0.59	0/649
25	2	0.26	0/536	0.56	0/713
26	3	0.29	0/438	0.56	0/590
27	4	0.27	0/647	0.48	0/868
28	5	0.33	0/429	0.61	0/571
29	6	0.31	0/407	0.58	0/545
30	7	0.34	0/371	0.71	0/484
31	8	0.30	0/526	0.64	0/690

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
32	9	0.32	0/299	0.56	0/393
33	a	0.40	0/36619	0.77	3/57097 (0.0%)
34	b	0.30	0/399	0.77	0/622
35	c	0.25	0/1808	0.47	0/2426
36	d	0.25	0/1634	0.53	0/2195
37	e	0.26	0/1647	0.52	0/2211
38	f	0.27	0/1174	0.54	0/1583
39	g	0.25	0/784	0.51	0/1052
40	h	0.23	0/1170	0.51	0/1572
41	i	0.26	0/1044	0.55	0/1401
42	j	0.25	0/1032	0.56	0/1386
43	k	0.26	0/804	0.54	0/1083
44	l	0.24	0/891	0.52	0/1203
45	m	0.28	0/1075	0.58	0/1439
46	n	0.23	0/909	0.57	0/1218
47	o	0.28	0/511	0.56	0/678
48	p	0.24	0/730	0.56	0/975
49	q	0.26	0/723	0.53	0/971
50	r	0.26	0/670	0.51	0/895
51	s	0.26	0/525	0.56	0/704
52	t	0.26	0/668	0.47	0/896
53	u	0.24	0/606	0.46	0/810
All	All	0.51	0/158322	0.75	12/236269 (0.0%)

There are no bond length outliers.

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	557	G	O4'-C1'-N9	9.24	115.59	108.20
1	A	2868	G	O4'-C1'-N9	6.12	113.09	108.20
1	A	2816	C	O4'-C1'-N1	5.82	112.86	108.20
33	a	1033	U	C2-N1-C1'	5.74	124.59	117.70
1	A	1997	A	O5'-P-OP1	-5.72	100.56	105.70
1	A	1350	U	C2-N1-C1'	5.48	124.27	117.70
1	A	12	U	N3-C2-O2	-5.17	118.58	122.20
1	A	12	U	C2-N1-C1'	5.16	123.89	117.70
33	a	583	G	O4'-C1'-N9	-5.09	104.13	108.20
33	a	762	C	C2-N1-C1'	5.07	124.38	118.80
1	A	27	G	O4'-C1'-N9	5.06	112.25	108.20
1	A	2821	U	N3-C2-O2	-5.01	118.69	122.20

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts [i](#)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

## 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 PDB entries followed by that with respect to all EM entries.

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	E	665/693 (96%)	644 (97%)	21 (3%)	0	100	100
5	G	271/277 (98%)	266 (98%)	5 (2%)	0	100	100
6	H	214/220 (97%)	204 (95%)	10 (5%)	0	100	100
7	I	202/207 (98%)	199 (98%)	3 (2%)	0	100	100
8	J	174/179 (97%)	170 (98%)	4 (2%)	0	100	100
9	K	174/178 (98%)	169 (97%)	5 (3%)	0	100	100
10	M	142/145 (98%)	138 (97%)	4 (3%)	0	100	100
11	N	120/122 (98%)	117 (98%)	3 (2%)	0	100	100
12	O	143/146 (98%)	138 (96%)	5 (4%)	0	100	100
13	P	134/144 (93%)	131 (98%)	3 (2%)	0	100	100
14	Q	118/122 (97%)	115 (98%)	3 (2%)	0	100	100
15	R	117/119 (98%)	113 (97%)	4 (3%)	0	100	100
16	S	112/116 (97%)	110 (98%)	2 (2%)	0	100	100
17	T	114/118 (97%)	112 (98%)	2 (2%)	0	100	100
18	U	99/102 (97%)	98 (99%)	1 (1%)	0	100	100
19	V	109/117 (93%)	108 (99%)	1 (1%)	0	100	100
20	W	87/91 (96%)	86 (99%)	1 (1%)	0	100	100
21	X	100/105 (95%)	96 (96%)	4 (4%)	0	100	100
22	Y	92/217 (42%)	91 (99%)	1 (1%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
23	Z	77/94 (82%)	75 (97%)	2 (3%)	0	100	100
24	1	59/62 (95%)	59 (100%)	0	0	100	100
25	2	63/69 (91%)	63 (100%)	0	0	100	100
26	3	54/59 (92%)	52 (96%)	2 (4%)	0	100	100
27	4	75/84 (89%)	72 (96%)	3 (4%)	0	100	100
28	5	51/57 (90%)	49 (96%)	2 (4%)	0	100	100
29	6	46/49 (94%)	43 (94%)	3 (6%)	0	100	100
30	7	41/45 (91%)	41 (100%)	0	0	100	100
31	8	62/66 (94%)	61 (98%)	1 (2%)	0	100	100
32	9	35/37 (95%)	35 (100%)	0	0	100	100
35	c	219/255 (86%)	211 (96%)	8 (4%)	0	100	100
36	d	202/217 (93%)	194 (96%)	8 (4%)	0	100	100
37	e	197/200 (98%)	193 (98%)	4 (2%)	0	100	100
38	f	154/166 (93%)	148 (96%)	6 (4%)	0	100	100
39	g	91/98 (93%)	85 (93%)	6 (7%)	0	100	100
40	h	141/156 (90%)	140 (99%)	1 (1%)	0	100	100
41	i	129/132 (98%)	121 (94%)	8 (6%)	0	100	100
42	j	126/132 (96%)	123 (98%)	3 (2%)	0	100	100
43	k	97/102 (95%)	93 (96%)	4 (4%)	0	100	100
44	l	116/129 (90%)	109 (94%)	6 (5%)	1 (1%)	14	28
45	m	133/137 (97%)	129 (97%)	4 (3%)	0	100	100
46	n	111/121 (92%)	105 (95%)	6 (5%)	0	100	100
47	o	58/61 (95%)	57 (98%)	1 (2%)	0	100	100
48	p	84/89 (94%)	81 (96%)	3 (4%)	0	100	100
49	q	88/91 (97%)	86 (98%)	2 (2%)	0	100	100
50	r	78/87 (90%)	76 (97%)	2 (3%)	0	100	100
51	s	61/80 (76%)	60 (98%)	1 (2%)	0	100	100
52	t	78/92 (85%)	76 (97%)	2 (3%)	0	100	100
53	u	78/83 (94%)	77 (99%)	1 (1%)	0	100	100
All	All	5991/6468 (93%)	5819 (97%)	171 (3%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
44	1	118	HIS

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	E	559/579 (96%)	552 (99%)	7 (1%)	65	85
5	G	220/224 (98%)	220 (100%)	0	100	100
6	H	174/177 (98%)	173 (99%)	1 (1%)	84	94
7	I	168/169 (99%)	168 (100%)	0	100	100
8	J	155/158 (98%)	154 (99%)	1 (1%)	84	94
9	K	153/155 (99%)	151 (99%)	2 (1%)	65	85
10	M	123/123 (100%)	122 (99%)	1 (1%)	79	91
11	N	100/100 (100%)	100 (100%)	0	100	100
12	O	111/112 (99%)	110 (99%)	1 (1%)	75	90
13	P	113/119 (95%)	112 (99%)	1 (1%)	75	90
14	Q	101/102 (99%)	99 (98%)	2 (2%)	50	75
15	R	95/95 (100%)	95 (100%)	0	100	100
16	S	100/102 (98%)	99 (99%)	1 (1%)	73	88
17	T	96/98 (98%)	96 (100%)	0	100	100
18	U	86/86 (100%)	85 (99%)	1 (1%)	67	86
19	V	90/94 (96%)	88 (98%)	2 (2%)	47	73
20	W	80/82 (98%)	80 (100%)	0	100	100
21	X	87/90 (97%)	87 (100%)	0	100	100
22	Y	83/190 (44%)	83 (100%)	0	100	100
23	Z	62/75 (83%)	62 (100%)	0	100	100
24	1	51/52 (98%)	51 (100%)	0	100	100
25	2	59/62 (95%)	59 (100%)	0	100	100
26	3	51/53 (96%)	51 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
27	4	69/75 (92%)	67 (97%)	2 (3%)	37	64
28	5	48/50 (96%)	48 (100%)	0	100	100
29	6	46/47 (98%)	45 (98%)	1 (2%)	47	73
30	7	39/40 (98%)	38 (97%)	1 (3%)	41	68
31	8	55/57 (96%)	54 (98%)	1 (2%)	54	78
32	9	35/35 (100%)	35 (100%)	0	100	100
35	c	192/221 (87%)	188 (98%)	4 (2%)	48	74
36	d	166/175 (95%)	165 (99%)	1 (1%)	84	94
37	e	174/175 (99%)	172 (99%)	2 (1%)	70	87
38	f	122/131 (93%)	122 (100%)	0	100	100
39	g	81/86 (94%)	81 (100%)	0	100	100
40	h	123/132 (93%)	122 (99%)	1 (1%)	79	91
41	i	112/113 (99%)	110 (98%)	2 (2%)	54	78
42	j	106/109 (97%)	103 (97%)	3 (3%)	38	65
43	k	89/91 (98%)	86 (97%)	3 (3%)	32	58
44	l	94/104 (90%)	92 (98%)	2 (2%)	48	74
45	m	117/119 (98%)	117 (100%)	0	100	100
46	n	98/104 (94%)	97 (99%)	1 (1%)	73	88
47	o	52/53 (98%)	52 (100%)	0	100	100
48	p	79/81 (98%)	79 (100%)	0	100	100
49	q	76/77 (99%)	76 (100%)	0	100	100
50	r	75/82 (92%)	75 (100%)	0	100	100
51	s	56/68 (82%)	55 (98%)	1 (2%)	54	78
52	t	70/80 (88%)	70 (100%)	0	100	100
53	u	67/69 (97%)	66 (98%)	1 (2%)	60	82
All	All	5158/5471 (94%)	5112 (99%)	46 (1%)	74	90

All (46) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	E	64	THR
4	E	305	ASP
4	E	495	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
4	E	496	PHE
4	E	503	ARG
4	E	505	GLN
4	E	611	MET
6	H	159	ASP
8	J	134	GLU
9	K	25	THR
9	K	173	GLU
10	M	142	GLU
12	O	80	ASP
13	P	77	LYS
14	Q	29	ARG
14	Q	75	ASP
16	S	82	LYS
18	U	57	THR
19	V	11	ARG
19	V	57	ASN
27	4	22	LYS
27	4	39	ASP
29	6	1	MET
30	7	44	SER
31	8	31	HIS
35	c	151	ILE
35	c	168	LYS
35	c	170	ARG
35	c	171	ASN
36	d	57	GLU
37	e	3	ARG
37	e	65	GLU
40	h	75	VAL
41	i	25	LEU
41	i	77	LEU
42	j	70	HIS
42	j	93	ASP
42	j	108	ARG
43	k	4	GLN
43	k	14	ASP
43	k	72	ARG
44	l	36	ASP
44	l	119	ASN
46	n	64	LYS
51	s	61	THR

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
53	u	28	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
43	k	4	GLN
53	u	69	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	2837/2923 (97%)	315 (11%)	17 (0%)
2	B	112/115 (97%)	16 (14%)	0
3	D	76/77 (98%)	15 (19%)	0
33	a	1528/1555 (98%)	193 (12%)	0
34	b	15/24 (62%)	2 (13%)	0
All	All	4568/4694 (97%)	541 (11%)	17 (0%)

All (541) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	34	U
1	A	63	U
1	A	64	A
1	A	71	A
1	A	75	G
1	A	89	U
1	A	90	A
1	A	117	A
1	A	118	A
1	A	119	U
1	A	155	U
1	A	164	A
1	A	165	C
1	A	167	U
1	A	168	A
1	A	176	A
1	A	199	A
1	A	218	G
1	A	219	A

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	224	A
1	A	225	A
1	A	231	A
1	A	232	U
1	A	251	G
1	A	255	G
1	A	269	G
1	A	284	C
1	A	285	U
1	A	286	U
1	A	289	U
1	A	312	A
1	A	313	U
1	A	314	A
1	A	315	C
1	A	321	U
1	A	324	A
1	A	327	G
1	A	354	A
1	A	373	A
1	A	404	U
1	A	410	G
1	A	411	A
1	A	432	G
1	A	457	G
1	A	458	A
1	A	497	U
1	A	502	C
1	A	527	G
1	A	550	A
1	A	553	A
1	A	575	G
1	A	576	U
1	A	577	A
1	A	578	G
1	A	583	A
1	A	590	U
1	A	592	A
1	A	593	U
1	A	594	G
1	A	606	G
1	A	616	G

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	618	A
1	A	630	G
1	A	646	A
1	A	650	U
1	A	682	A
1	A	683	G
1	A	690	U
1	A	699	U
1	A	700	A
1	A	731	U
1	A	762	C
1	A	767	A
1	A	774	G
1	A	775	A
1	A	792	5MU
1	A	810	A
1	A	820	G
1	A	827	A
1	A	829	U
1	A	835	U
1	A	837	G
1	A	850	G
1	A	857	C
1	A	872	U
1	A	873	U
1	A	904	G
1	A	911	A
1	A	919	G
1	A	923	A
1	A	924	G
1	A	925	G
1	A	927	G
1	A	941	A
1	A	943	C
1	A	945	A
1	A	955	A
1	A	970	U
1	A	971	U
1	A	972	A
1	A	985	A
1	A	989	A
1	A	990	G

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	1005	G
1	A	1018	A
1	A	1027	A
1	A	1040	A
1	A	1049	C
1	A	1056	U
1	A	1057	A
1	A	1066	G
1	A	1070	A
1	A	1077	U
1	A	1090	A
1	A	1091	G
1	A	1111	A
1	A	1114	A
1	A	1119	C
1	A	1121	A
1	A	1128	A
1	A	1132	A
1	A	1134	U
1	A	1141	U
1	A	1156	G
1	A	1158	G
1	A	1172	A
1	A	1176	U
1	A	1179	C
1	A	1180	G
1	A	1186	A
1	A	1199	A
1	A	1214	C
1	A	1216	U
1	A	1218	G
1	A	1221	C
1	A	1222	A
1	A	1276	G
1	A	1291	A
1	A	1294	G
1	A	1309	G
1	A	1310	A
1	A	1337	A
1	A	1339	U
1	A	1402	A
1	A	1405	G

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	1415	A
1	A	1416	U
1	A	1421	A
1	A	1449	A
1	A	1455	U
1	A	1457	U
1	A	1458	A
1	A	1459	A
1	A	1463	A
1	A	1464	U
1	A	1471	A
1	A	1472	C
1	A	1477	U
1	A	1490	G
1	A	1497	A
1	A	1498	U
1	A	1504	U
1	A	1511	C
1	A	1526	G
1	A	1530	A
1	A	1533	A
1	A	1536	C
1	A	1553	A
1	A	1560	A
1	A	1569	G
1	A	1578	A
1	A	1580	A
1	A	1581	U
1	A	1582	U
1	A	1583	G
1	A	1585	G
1	A	1587	C
1	A	1592	A
1	A	1606	C
1	A	1613	G
1	A	1616	A
1	A	1625	U
1	A	1630	A
1	A	1631	G
1	A	1642	C
1	A	1651	C
1	A	1652	A

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	1654	A
1	A	1658	A
1	A	1690	A
1	A	1691	G
1	A	1692	C
1	A	1718	G
1	A	1740	G
1	A	1759	G
1	A	1772	G
1	A	1790	G
1	A	1791	G
1	A	1800	A
1	A	1803	G
1	A	1809	C
1	A	1818	A
1	A	1827	C
1	A	1828	U
1	A	1843	U
1	A	1856	A
1	A	1875	A
1	A	1899	U
1	A	1900	G
1	A	1909	C
1	A	1933	G
1	A	1943	A
1	A	1956	G
1	A	1957	G
1	A	1963	A
1	A	1964	A
1	A	1982	U
1	A	1994	C
1	A	1997	A
1	A	1998	A
1	A	1999	G
1	A	2018	U
1	A	2020	U
1	A	2040	A
1	A	2050	A
1	A	2058	A
1	A	2059	G
1	A	2060	A
1	A	2070	C

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	2082	C
1	A	2083	G
1	A	2087	A
1	A	2088	G
1	A	2089	A
1	A	2096	G
1	A	2120	G
1	A	2209	G
1	A	2215	U
1	A	2217	G
1	A	2225	A
1	A	2230	G
1	A	2231	C
1	A	2238	U
1	A	2239	A
1	A	2252	A
1	A	2265	G
1	A	2266	G
1	A	2295	A
1	A	2306	G
1	A	2310	C
1	A	2314	A
1	A	2333	U
1	A	2334	G
1	A	2335	G
1	A	2336	A
1	A	2337	A
1	A	2347	A
1	A	2349	A
1	A	2352	G
1	A	2361	U
1	A	2362	A
1	A	2374	C
1	A	2377	C
1	A	2406	G
1	A	2410	G
1	A	2412	C
1	A	2429	U
1	A	2433	C
1	A	2449	C
1	A	2450	U
1	A	2452	A

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	2455	G
1	A	2456	G
1	A	2457	A
1	A	2462	A
1	A	2468	C
1	A	2475	A
1	A	2529	G
1	A	2531	U
1	A	2532	G
1	A	2545	A
1	A	2556	G
1	A	2562	G
1	A	2575	G
1	A	2581	U
1	A	2593	A
1	A	2594	G
1	A	2600	C
1	A	2605	G
1	A	2629	A
1	A	2637	C
1	A	2640	U
1	A	2657	G
1	A	2690	G
1	A	2692	A
1	A	2716	U
1	A	2718	C
1	A	2741	G
1	A	2753	U
1	A	2760	A
1	A	2775	A
1	A	2792	A
1	A	2793	G
1	A	2805	A
1	A	2821	U
1	A	2823	G
1	A	2824	G
1	A	2827	A
1	A	2828	U
1	A	2852	U
1	A	2887	G
1	A	2892	G
1	A	2900	C

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	2903	A
1	A	2913	G
2	B	10	U
2	B	11	A
2	B	23	U
2	B	24	C
2	B	26	C
2	B	30	U
2	B	33	U
2	B	39	G
2	B	40	C
2	B	42	G
2	B	43	A
2	B	54	U
2	B	55	A
2	B	87	C
2	B	88	G
2	B	106	G
3	D	8	U
3	D	9	G
3	D	16	C
3	D	17(A)	U
3	D	19	G
3	D	20	U
3	D	21	A
3	D	22	G
3	D	43	A
3	D	46	G
3	D	47	U
3	D	52	G
3	D	58	A
3	D	61	C
3	D	76	A
33	a	7	G
33	a	8	G
33	a	10	G
33	a	33	A
33	a	40	G
33	a	48	C
33	a	49	C
33	a	52	A
33	a	75	A

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
33	a	81	A
33	a	82	G
33	a	84	U
33	a	85	U
33	a	89	U
33	a	91	U
33	a	92	C
33	a	94	G
33	a	99	U
33	a	108	A
33	a	119	A
33	a	120	C
33	a	131	C
33	a	163	C
33	a	165	G
33	a	173	U
33	a	183	U
33	a	196	A
33	a	205	A
33	a	208	U
33	a	209	G
33	a	210	A
33	a	212	A
33	a	213	G
33	a	219	C
33	a	221	U
33	a	248	U
33	a	249	G
33	a	253	U
33	a	255	G
33	a	259	G
33	a	270	A
33	a	271	A
33	a	274	G
33	a	275	C
33	a	297	G
33	a	336	C
33	a	337	A
33	a	340	G
33	a	355	G
33	a	360	C
33	a	362	G

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
33	a	375	U
33	a	380	C
33	a	381	A
33	a	405	A
33	a	414	G
33	a	420	U
33	a	421	G
33	a	422	A
33	a	423	A
33	a	429	U
33	a	431	G
33	a	437	U
33	a	447	U
33	a	456	A
33	a	460	A
33	a	461	C
33	a	467	U
33	a	474	A
33	a	490	A
33	a	493	G
33	a	503	A
33	a	505	A
33	a	519	C
33	a	522	C
33	a	526	C
33	a	530	C
33	a	540	A
33	a	553	C
33	a	555	A
33	a	570	U
33	a	572	U
33	a	580	A
33	a	581	A
33	a	584	C
33	a	589	G
33	a	642	C
33	a	661	U
33	a	673	A
33	a	695	A
33	a	710	A
33	a	727	C
33	a	731	U

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
33	a	732	G
33	a	757	A
33	a	769	G
33	a	785	A
33	a	801	U
33	a	802	A
33	a	823	A
33	a	825	C
33	a	836	A
33	a	864	G
33	a	881	A
33	a	883	G
33	a	923	A
33	a	926	G
33	a	927	A
33	a	943	C
33	a	944	A
33	a	951	G
33	a	969	U
33	a	978	A
33	a	980	G
33	a	981	C
33	a	984	A
33	a	986	A
33	a	987	A
33	a	992	A
33	a	1001	U
33	a	1002	G
33	a	1013	A
33	a	1014	C
33	a	1015	A
33	a	1016	A
33	a	1017	C
33	a	1018	U
33	a	1024	G
33	a	1029	A
33	a	1035	C
33	a	1036	C
33	a	1050	A
33	a	1056	C
33	a	1064	G
33	a	1076	U

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
33	a	1084	U
33	a	1088	G
33	a	1105	G
33	a	1106	U
33	a	1112	A
33	a	1136	U
33	a	1141	A
33	a	1143	C
33	a	1149	G
33	a	1150	U
33	a	1152	G
33	a	1168	C
33	a	1169	U
33	a	1170	G
33	a	1178	C
33	a	1194	G
33	a	1206	A
33	a	1207	A
33	a	1210	C
33	a	1223	A
33	a	1224	U
33	a	1232	G
33	a	1235	A
33	a	1237	A
33	a	1267	A
33	a	1270	G
33	a	1278	G
33	a	1289	A
33	a	1290	A
33	a	1296	U
33	a	1297	A
33	a	1309	A
33	a	1312	U
33	a	1327	C
33	a	1330	C
33	a	1333	G
33	a	1342	A
33	a	1346	U
33	a	1347	G
33	a	1348	G
33	a	1356	A
33	a	1361	U

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
33	a	1363	G
33	a	1378	A
33	a	1408	A
33	a	1452	G
33	a	1456	A
33	a	1464	A
33	a	1503	A
33	a	1508	G
33	a	1510	A
33	a	1517	U
33	a	1528	G
33	a	1540	G
33	a	1541	G
33	a	1547	C
33	a	1549	C
33	a	1551	U
34	b	8	A
34	b	16	A

All (17) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	130	A
1	A	154	A
1	A	232	U
1	A	312	A
1	A	756	A
1	A	809	A
1	A	926	G
1	A	1083	G
1	A	1179	C
1	A	1193	U
1	A	1291	A
1	A	1874	A
1	A	2336	A
1	A	2449	C
1	A	2457	A
1	A	2822	C
1	A	2827	A

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

15 non-standard protein/DNA/RNA residues are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	5MU	A	792	1	19,22,23	0.33	0	28,32,35	0.36	0
33	4OC	a	1412	33	20,23,24	0.33	0	26,32,35	0.50	0
1	H2U	A	2476	1	18,21,22	0.35	0	21,30,33	0.44	0
33	MA6	a	1530	33	18,26,27	0.76	0	19,38,41	0.69	0
33	UR3	a	1509	33	19,22,23	0.37	0	26,32,35	0.68	1 (3%)
1	OMG	A	2278	1,54	18,26,27	1.08	3 (16%)	19,38,41	1.05	1 (5%)
33	2MG	a	975	33	18,26,27	0.99	3 (16%)	16,38,41	0.67	0
1	3TD	A	1942[A]	1	19,22,23	0.57	0	21,32,35	0.57	0
33	5MC	a	976	33	18,22,23	0.34	0	26,32,35	0.41	0
33	MA6	a	1529	33	18,26,27	0.79	0	19,38,41	0.67	0
33	G7M	a	535	33	20,26,27	0.64	0	17,39,42	0.39	0
1	5MU	A	1966	1	19,22,23	0.34	0	28,32,35	0.44	0
1	2MG	A	2472	1	18,26,27	1.05	3 (16%)	16,38,41	0.62	0
1	2MA	A	2530	1,54	19,25,26	1.09	2 (10%)	21,37,40	3.24	5 (23%)
1	3TD	A	1942[B]	1	19,22,23	0.53	0	21,32,35	0.56	0

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
1	5MU	A	792	1	-	0/7/25/26	0/2/2/2
33	4OC	a	1412	33	-	0/9/29/30	0/2/2/2
1	H2U	A	2476	1	-	0/7/38/39	0/2/2/2
33	MA6	a	1530	33	-	5/7/29/30	0/3/3/3
33	UR3	a	1509	33	-	2/7/25/26	0/2/2/2
1	OMG	A	2278	1,54	-	0/5/27/28	0/3/3/3
33	2MG	a	975	33	-	0/5/27/28	0/3/3/3
1	3TD	A	1942[A]	1	-	0/7/25/26	0/2/2/2

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
33	5MC	a	976	33	-	0/7/25/26	0/2/2/2
33	MA6	a	1529	33	-	0/7/29/30	0/3/3/3
33	G7M	a	535	33	-	2/3/25/26	0/3/3/3
1	5MU	A	1966	1	-	0/7/25/26	0/2/2/2
1	2MG	A	2472	1	-	0/5/27/28	0/3/3/3
1	2MA	A	2530	1,54	-	2/3/25/26	0/3/3/3
1	3TD	A	1942[B]	1	-	2/7/25/26	0/2/2/2

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
33	a	975	2MG	C5-C6	-2.45	1.42	1.47
1	A	2278	OMG	C5-C6	-2.44	1.42	1.47
1	A	2472	2MG	C8-N7	-2.39	1.31	1.35
1	A	2530	2MA	C6-N6	-2.38	1.25	1.34
1	A	2278	OMG	C8-N7	-2.32	1.31	1.35
1	A	2278	OMG	C5-C4	-2.31	1.37	1.43
1	A	2472	2MG	C5-C4	-2.21	1.37	1.43
1	A	2472	2MG	C5-C6	-2.13	1.43	1.47
1	A	2530	2MA	C8-N7	-2.09	1.31	1.34
33	a	975	2MG	C5-C4	-2.07	1.37	1.43
33	a	975	2MG	C8-N7	-2.03	1.31	1.35

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	2530	2MA	C5-C6-N1	-13.04	112.45	121.01
1	A	2530	2MA	C2-N1-C6	4.23	124.68	118.08
1	A	2530	2MA	C2-N3-C4	-3.88	112.37	115.52
1	A	2278	OMG	O3'-C3'-C2'	3.27	120.46	111.17
1	A	2530	2MA	C5-C6-N6	2.55	124.22	120.35
33	a	1509	UR3	C2'-C1'-N1	2.38	119.94	113.22
1	A	2530	2MA	N6-C6-N1	2.30	123.32	117.07

There are no chirality outliers.

All (13) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	1942[B]	3TD	C3'-C4'-C5'-O5'
33	a	1530	MA6	C5-C6-N6-C10

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	Atoms
1	A	1942[B]	3TD	O4'-C4'-C5'-O5'
33	a	1530	MA6	O4'-C4'-C5'-O5'
33	a	1530	MA6	C3'-C4'-C5'-O5'
33	a	1530	MA6	C5-C6-N6-C9
33	a	1509	UR3	O4'-C4'-C5'-O5'
1	A	2530	2MA	C4'-C5'-O5'-P
1	A	2530	2MA	O4'-C4'-C5'-O5'
33	a	535	G7M	C3'-C4'-C5'-O5'
33	a	535	G7M	C4'-C5'-O5'-P
33	a	1530	MA6	C4'-C5'-O5'-P
33	a	1509	UR3	C3'-C4'-C5'-O5'

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 178 ligands modelled in this entry, 172 are monoatomic - leaving 6 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	$\# Z  > 2$
56	PUT	A	3109	-	5,5,5	0.11	0	4,4,4	0.06	0
55	SPD	A	3106	-	9,9,9	0.11	0	8,8,8	0.19	0
55	SPD	A	3107	-	9,9,9	0.11	0	8,8,8	0.20	0
57	GDP	E	701	54	24,30,30	0.89	3 (12%)	30,47,47	0.77	1 (3%)
55	SPD	A	3108	-	9,9,9	0.12	0	8,8,8	0.15	0
58	FUA	E	703	-	39,40,40	0.16	0	49,64,64	0.53	0

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
56	PUT	A	3109	-	-	2/3/3/3	-
55	SPD	A	3106	-	-	2/7/7/7	-
55	SPD	A	3107	-	-	3/7/7/7	-
57	GDP	E	701	54	-	2/12/32/32	0/3/3/3
55	SPD	A	3108	-	-	2/7/7/7	-
58	FUA	E	703	-	-	2/15/92/92	0/4/4/4

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
57	E	701	GDP	C5-C6	-2.20	1.42	1.47
57	E	701	GDP	C5-C4	-2.02	1.37	1.43
57	E	701	GDP	C8-N7	-2.00	1.31	1.35

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
57	E	701	GDP	PA-O3A-PB	2.89	142.74	132.83

There are no chirality outliers.

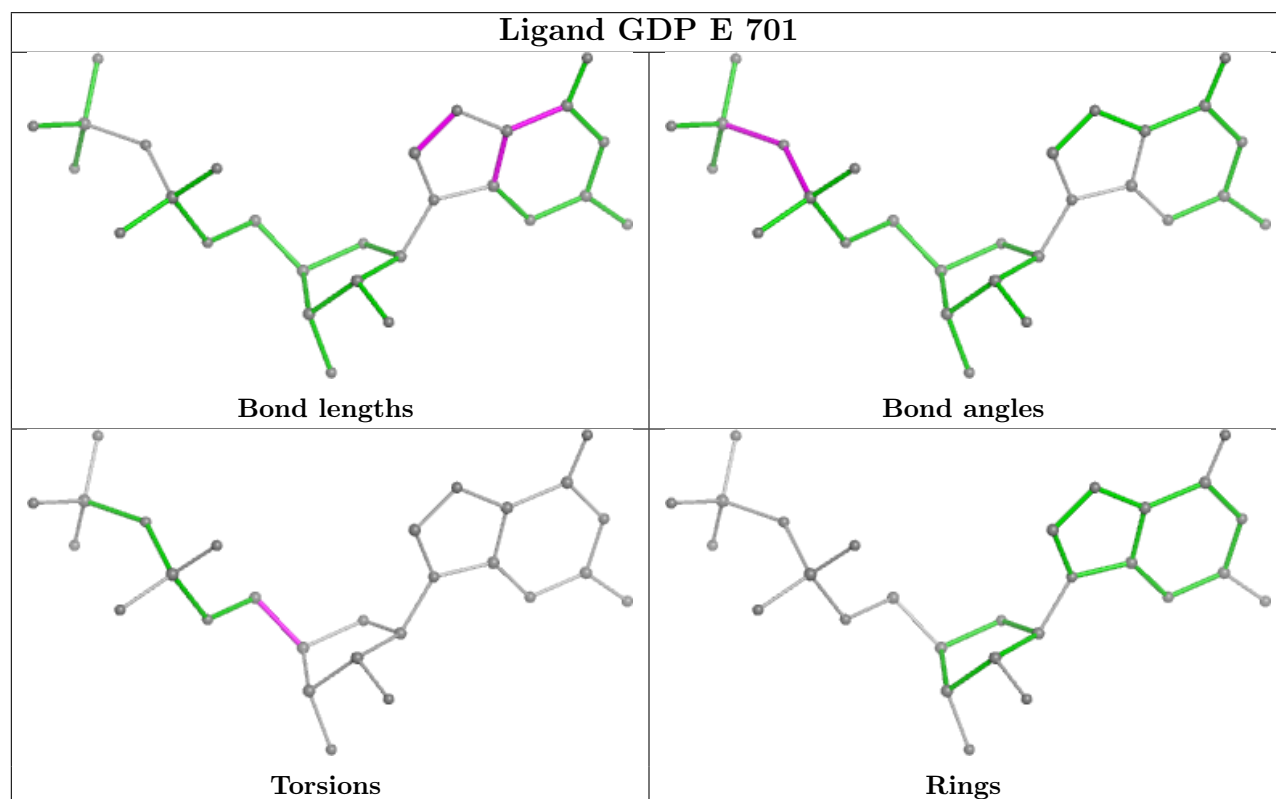
All (13) torsion outliers are listed below:

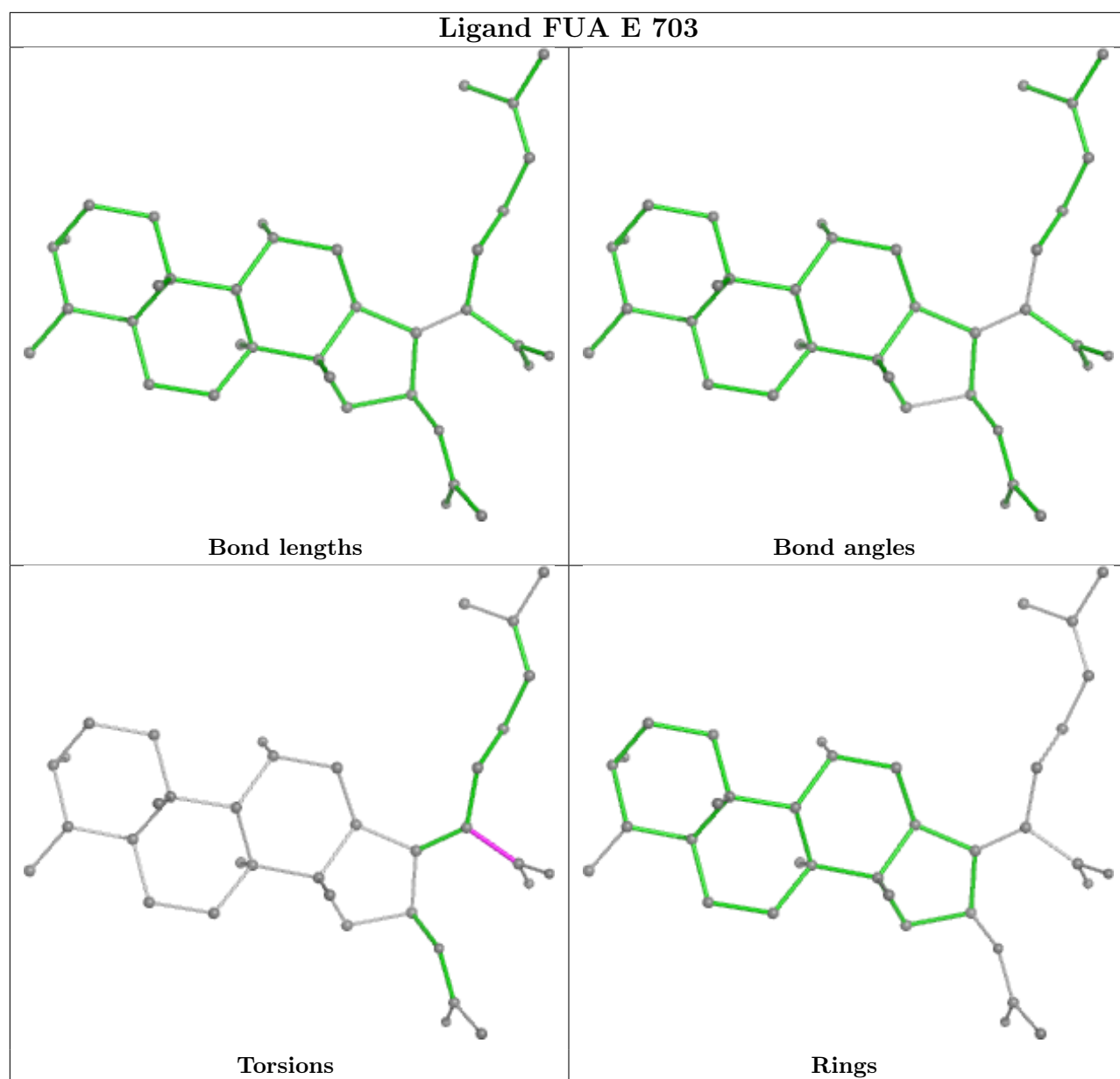
Mol	Chain	Res	Type	Atoms
58	E	703	FUA	C23-C22-C29-O4
58	E	703	FUA	C23-C22-C29-O5
57	E	701	GDP	O4'-C4'-C5'-O5'
55	A	3108	SPD	C3-C4-C5-N6
55	A	3107	SPD	N6-C7-C8-C9
55	A	3106	SPD	C4-C5-N6-C7
56	A	3109	PUT	C2-C3-C4-N2
56	A	3109	PUT	C1-C2-C3-C4
55	A	3107	SPD	C3-C4-C5-N6
57	E	701	GDP	C3'-C4'-C5'-O5'
55	A	3107	SPD	C8-C7-N6-C5
55	A	3108	SPD	C8-C7-N6-C5
55	A	3106	SPD	C8-C7-N6-C5

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.