wwPDB/EMDataBank EM Map/Model Validation Summary Report (i)

Jan 30, 2017 – 02:31 PM EST

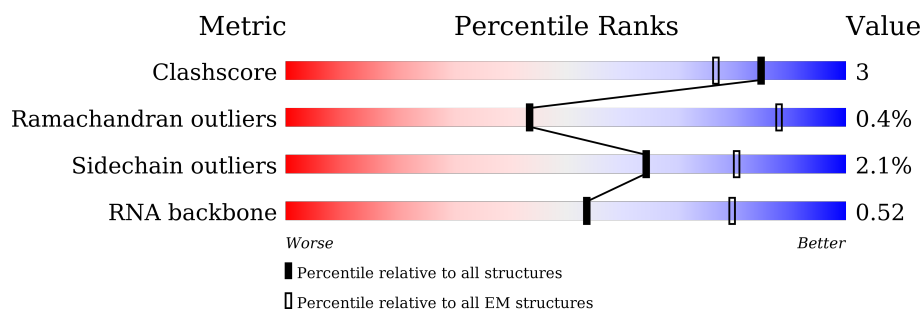|		
PDB ID	:	5U4J
EMDB ID:	:	EMD-8506
Title	:	Structural Basis of Co-translational Quality Control by ArfA and RF2 Bound to Ribosome
Authors	:	Zeng, F.; Chen, Y.; Remis, J.; Shekhar, M.; Phillips, J.C.; Tajkhorshid, E.; Jin, H.
Deposited on	:	2016-12-04
Resolution	:	3.70 Å(reported)

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.70 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	Percentile Ranks	Value
Clashscore		3
Ramachandran outliers		0.4%
Sidechain outliers		2.1%
RNA backbone		0.52

Worse   *Better*

■ Percentile relative to all structures

▯ Percentile relative to all EM structures

Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686
RNA backbone	3027	244

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain
1	a	1533	15% 5% 80%
2	A	2904	100%
3	c	233	9% 91%
4	d	206	32% 68%
5	e	167	80% • 19%
6	z	18	28% 6% 67%
7	x	77	19% 81%
8	l	124	98% ••

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	v	383	34% · 66%
10	w	57	72% 11% 18%

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 11360 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 16S rRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	a	307	Total	C	N	O	P	0	0
			6591	2944	1198	2142	307		

- Molecule 2 is a RNA chain called 23S rRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	A	11	Total	C	N	O	P	0	0
			233	105	42	75	11		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	887	A	U	conflict	GB 42756

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

Mol	Chain	Residues	Atoms				AltConf	Trace
3	c	21	Total	C	N	O	0	0
			169	105	32	32		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
4	d	65	Total	C	N	O	S	0	0
			523	327	106	89	1		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
5	e	136	Total	C	N	O	S	0	0
			1001	625	190	182	4		

- Molecule 6 is a RNA chain called mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	z	6	Total 131	C 59	N 27	O 39	P 6	0	0

- Molecule 7 is a RNA chain called P-site tRNA fMet.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	x	15	Total 315	C 141	N 54	O 105	P 15	0	0

- Molecule 8 is a protein called 30S ribosomal protein S12.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	l	123	Total 955	C 590	N 196	O 165	S 4	0	0

- Molecule 9 is a protein called Peptide chain release factor 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	v	132	Total 1054	C 659	N 182	O 208	S 5	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
v	-17	HIS	-	expression tag	UNP P07012
v	-16	HIS	-	expression tag	UNP P07012
v	-15	HIS	-	expression tag	UNP P07012
v	-14	HIS	-	expression tag	UNP P07012
v	-13	HIS	-	expression tag	UNP P07012
v	-12	HIS	-	expression tag	UNP P07012
v	-11	SER	-	expression tag	UNP P07012
v	-10	ALA	-	expression tag	UNP P07012
v	-9	ALA	-	expression tag	UNP P07012
v	-8	LEU	-	expression tag	UNP P07012
v	-7	GLU	-	expression tag	UNP P07012
v	-6	VAL	-	expression tag	UNP P07012
v	-5	LEU	-	expression tag	UNP P07012
v	-4	PHE	-	expression tag	UNP P07012
v	-3	GLN	-	expression tag	UNP P07012
v	-2	GLY	-	expression tag	UNP P07012
v	-1	PRO	-	expression tag	UNP P07012
v	0	GLY	-	expression tag	UNP P07012

- Molecule 10 is a protein called Alternative ribosome-rescue factor A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	w	47	388	239	82	66	1	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
w	-1	GLY	-	expression tag	UNP P36675
w	0	SER	-	expression tag	UNP P36675

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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: 16S rRNA

• Molecule 2: 23S rRNA

Chain A:

Modified/highlighted residues labeled in sequence: 1MG, PSU, 5MU, PSU, 2MG, G1910, A1913, C1914, C1920

(RNA sequence blocks displayed vertically, with modified nucleotides labeled: 5MU, 5MC, G7M, OMG, 2MG, PSU, OMC, 2MA, PSU, DMU, PSU, PSU)

• Molecule 3: 30S ribosomal protein S3

Chain c: 9%   91%

(Protein sequence displayed vertically)

MET GLY GLN LYS VAL HIS PRO ASN GLY ILE ARG LEU GLY ILE VAL LYS PRO TRP ASN SER THR TRP PHE ALA ASN THR LYS GLU PHE ALA ASP ASN LEU ASP SER ASP PHE LYS VAL ARG GLN TYR LEU THR LYS GLU LEU ALA LYS ALA SER VAL SER ARG ILE VAL ILE GLU ARG PRO

ALA LYS SER ILE ARG VAL THR ILE HIS THR ALA ARG PRO GLY ILE VAL ILE GLY LYS LYS GLY GLU ASP VAL GLU LYS LEU ARG LYS VAL VAL ALA ASP ILE ALA GLY VAL PRO ALA GLN ILE ASN ILE ALA GLU VAL ARG LYS PRO GLU LEU ASP ALA LYS LEU VAL ALA ASP SER ILE

THR SER GLN LEU GLU ARG VAL MET PHE ARG ARG ALA MET LYS ARG ALA VAL GLN ASN ALA MET ARG LEU GLY ALA LYS GLY ILE **K150** **E170** GLY ARG VAL PRO LEU HIS THR LEU ARG ALA ILE ASP THR TYR ASN THR SER GLU ALA HIS THR THR TYR GLY VAL ILE GLY VAL

LYS VAL TRP ILE PHE LYS GLY GLU ILE LEU GLY GLY MET ALA ALA VAL GLU GLN PRO GLU LYS PRO ALA ALA GLN PRO LYS LYS GLN ARG LYS GLY ARG LYS

- Molecule 4: 30S ribosomal protein S4

Chain d: 32%   68%

MET **A2** **G66** VAL LEU GLU ARG GLN PHE ARG ASN TYR LYS GLU ALA ALA ARG LEU LYS GLY ASN THR GLY GLU ASN LEU LEU ALA LEU LEU GLU GLY ARG LEU ASP ASN VAL VAL TYR ARG MET GLY PHE GLY ALA THR ARG ALA GLU ALA ARG GLN LEU VAL SER HIS LYS ALA

ILE MET VAL ASN GLY ARG VAL VAL ASN ILE ALA SER TYR GLN VAL SER PRO ASN ASP VAL VAL SER ILE ARG GLU LYS ALA LYS LYS GLN SER ARG VAL LYS ALA ALA LEU GLU LEU ALA GLU GLN ARG GLU LYS PRO THR TRP LEU GLU VAL ASP ALA GLY LYS MET GLU GLY THR PHE

LYS ARG LYS PRO GLU ARG SER ASP LEU SER ALA ASP ILE ASN GLU HIS LEU ILE VAL GLU LEU TYR SER LYS

- Molecule 5: 30S ribosomal protein S5

Chain e: 80%   ·   19%

MET ALA HIS ILE GLU LYS GLN ALA GLY **E10** **V25** **G44** **L124** **E145** ASN MET ASN SER PRO GLU MET VAL ALA ALA LYS ARG GLY LYS SER VAL GLU LEU ILE LEU GLY LYS

- Molecule 6: mRNA

Chain z: 28%   6%   67%

G G C A A G G A G G U A **A43** **A46** **U47** **G48**

- Molecule 7: P-site tRNA fMet

Chain x: 19%   81%

C G G G G U G G A G C A G C U G G U A G C U C **U28** **C42** G A A A G U C G U C G G U C A A U C C G G C C C C C G C A A C C A

- Molecule 8: 30S ribosomal protein S12

Chain l: 98%

MET **A2** **K44** **V52** **A124**

- Molecule 9: Peptide chain release factor 2

Chain v: 34%   ·   66%

HIS HIS HIS HIS HIS HIS SER ALA ALA LEU GLU VAL LEU PHE GLN GLY PRO GLY MET PHE GLU ILE ASN PRO VAL ASN ASN ARG ILE GLN ASP LEU THR GLU ARG SER ASP VAL LEU ARG GLY TYR LEU ASP TYR ASP ALA LYS LYS GLU ARG GLU GLU VAL ASN ALA GLU LEU GLU

GLN PRO ASP VAL TRP ASN GLU PRO GLU ARG ALA GLN ALA LEU GLY LYS GLU ARG SER SER LEU GLU ALA VAL VAL ASP THR LEU ASP GLN MET LYS GLN LEU GLU GLU ASP VAL SER GLY LEU LEU GLU LEU ALA VAL GLU ASP ASP GLU GLU THR PHE ASN GLU ALA VAL ALA GLU
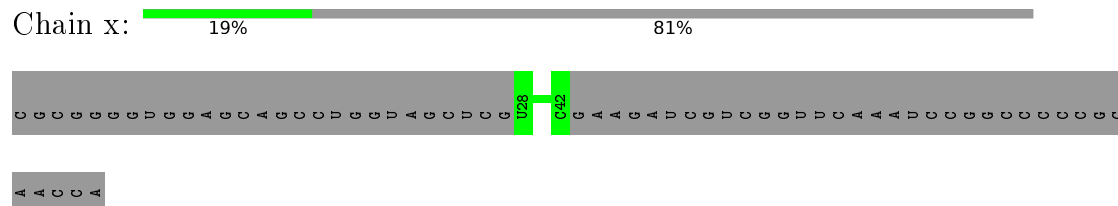
LEU ASP ALA LEU GLU GLU LYS LEU ALA GLN LEU GLU GLU PHE ARG ARG MET PHE SER GLU GLU TYR ASP SER ALA **D127** V179 V226 ASP ASP ASP ILE ILE ASP GLU ILE LEU ASN PRO ALA ASP LEU ARG ILE ASP VAL TYR ARG THR SER GLY ALA GLY GLN HIS VAL LEU ASN ARG THR

GLU SER ALA VAL ARG ILE THR HIS ILE PRO THR GLY ILE VAL THR GLN CYS GLN ASN ASP ARG GLN SER ARG GLN HIS LYS ASN LYS ASP GLN ALA LEU MET LYS GLN MET LYS SER ALA LYS LEU **Y296** W319 V327 LEU ASP ASP SER ARG ILE LYS ASP LEU ARG THR GLY VAL GLU THR ARG ASN

THR GLU ALA VAL LEU ASP GLY SER LEU ASP GLN PHE ILE GLU ALA SER LEU LYS ALA GLY LEU

● Molecule 10: Alternative ribosome-rescue factor A

Chain w: [====72%====][11%][18%]

GLY SER MET **S2** K8 D13 E17 L20 V29 E30 **R48** GLY ASN TRP GLU GLU ALA SER GLY

4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	155440	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	JEOL 3200FSC	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/Å^2)$	Not provided	Depositor
Minimum defocus (nm)	700	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	83822	Depositor
Image detector	Not provided	Depositor

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: MA6, G7M, 2MG, 5MC, UR3, 4OC, PSU

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	#$|Z|$ >2	RMSZ	#$|Z|$ >2
1	a	0.23	0/7167	0.64	3/11159 (0.0%)
10	w	0.34	0/394	0.69	2/519 (0.4%)
2	A	0.18	0/216	0.63	0/334
3	c	0.37	0/171	0.51	0/228
4	d	0.37	0/530	0.55	0/704
5	e	0.36	0/1013	0.56	0/1365
6	z	0.19	0/147	0.59	0/227
7	x	0.20	0/350	0.63	0/542
8	l	0.35	0/969	0.59	0/1300
9	v	0.39	0/1075	0.52	0/1443
All	All	0.28	0/12032	0.62	5/17821 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	a	527	G7M	O3'-P-O5'	9.93	122.87	104.00
10	w	8	LYS	CB-CA-C	5.74	121.87	110.40
1	a	528	C	OP1-P-OP2	-5.62	111.17	119.60
10	w	20	LEU	CA-CB-CG	5.23	127.33	115.30
1	a	1070	U	C2'-C3'-O3'	5.02	121.73	113.70

There are no chirality outliers.

There are no planarity outliers.

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	6591	0	3340	0	0
2	A	233	0	119	0	0
3	c	169	0	165	0	0
4	d	523	0	558	0	0
5	e	1001	0	1039	0	0
6	z	131	0	66	0	0
7	x	315	0	163	0	0
8	l	955	0	1016	0	0
9	v	1054	0	1004	0	0
10	w	388	0	400	0	0
All	All	11360	0	7870	0	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 3.

There are no clashes within the asymmetric unit.

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 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	
3	c	19/233 (8%)	18 (95%)	1 (5%)	0	100	100
4	d	63/206 (31%)	62 (98%)	1 (2%)	0	100	100
5	e	134/167 (80%)	123 (92%)	10 (8%)	1 (1%)	26	72
8	l	121/124 (98%)	114 (94%)	6 (5%)	1 (1%)	24	70
9	v	128/383 (33%)	121 (94%)	7 (6%)	0	100	100
10	w	45/57 (79%)	41 (91%)	4 (9%)	0	100	100
All	All	510/1170 (44%)	479 (94%)	29 (6%)	2 (0%)	43	80

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
8	l	44	LYS
5	e	44	GLY

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 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	
3	c	16/190 (8%)	16 (100%)	0	100	100
4	d	54/173 (31%)	54 (100%)	0	100	100
5	e	102/126 (81%)	100 (98%)	2 (2%)	63	87
8	l	103/104 (99%)	102 (99%)	1 (1%)	82	92
9	v	110/325 (34%)	108 (98%)	2 (2%)	66	89
10	w	40/46 (87%)	36 (90%)	4 (10%)	9	44
All	All	425/964 (44%)	416 (98%)	9 (2%)	64	86

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	v	319	TRP
10	w	30	GLU
10	w	17	GLU
8	l	52	VAL
10	w	13	ASP

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

Mol	Chain	Res	Type
5	e	132	ASN
9	v	214	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	a	297/1533 (19%)	68 (22%)	0
2	A	10/2904 (0%)	2 (20%)	1 (10%)
6	z	5/18 (27%)	1 (20%)	0
7	x	14/77 (18%)	0	0
All	All	326/4532 (7%)	71 (21%)	1 (0%)

5 of 71 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	a	7	A
1	a	9	G
1	a	13	U
1	a	14	U
1	a	22	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	A	1913	A

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

10 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
2	PSU	A	1911	2	15,21,22	1.09	1 (6%)	16,30,33	2.31	5 (31%)
2	PSU	A	1917	2	15,21,22	1.11	1 (6%)	16,30,33	2.38	5 (31%)
1	4OC	a	1402	1	15,23,24	0.70	0	21,32,35	2.03	5 (23%)
1	5MC	a	1407	1	14,22,23	1.58	2 (14%)	17,32,35	1.10	1 (5%)
1	UR3	a	1498	1	13,22,23	1.07	1 (7%)	18,32,35	1.07	1 (5%)
1	2MG	a	1516	1	18,26,27	1.39	2 (11%)	21,38,41	2.44	8 (38%)
1	MA6	a	1518	1	18,26,27	1.26	1 (5%)	15,38,41	2.34	4 (26%)
1	MA6	a	1519	1	18,26,27	1.23	1 (5%)	15,38,41	2.51	3 (20%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
1	PSU	a	516	1	15,21,22	1.10	1 (6%)	16,30,33	2.39	5 (31%)
1	G7M	a	527	1	18,26,27	1.04	1 (5%)	21,39,42	2.27	5 (23%)

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
2	PSU	A	1911	2	-	0/7/25/26	0/2/2/2
2	PSU	A	1917	2	-	0/7/25/26	0/2/2/2
1	4OC	a	1402	1	-	0/7/29/30	0/2/2/2
1	5MC	a	1407	1	-	0/3/25/26	0/2/2/2
1	UR3	a	1498	1	-	0/3/25/26	0/2/2/2
1	2MG	a	1516	1	-	0/5/27/28	0/3/3/3
1	MA6	a	1518	1	-	0/7/29/30	0/3/3/3
1	MA6	a	1519	1	-	0/7/29/30	0/3/3/3
1	PSU	a	516	1	-	0/7/25/26	0/2/2/2
1	G7M	a	527	1	-	0/3/25/26	0/3/3/3

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	a	516	PSU	C5-C1'	-3.04	1.49	1.52
2	A	1917	PSU	C5-C1'	-2.98	1.49	1.52
2	A	1911	PSU	C5-C1'	-2.95	1.49	1.52
1	a	1407	5MC	O4'-C1'	2.06	1.44	1.41
1	a	1498	UR3	C4-N3	3.07	1.42	1.38

The worst 5 of 42 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	a	1519	MA6	N3-C2-N1	-7.04	123.34	128.87
1	a	1518	MA6	N3-C2-N1	-6.72	123.59	128.87
1	a	527	G7M	C5-C6-N1	-4.88	117.14	123.52
1	a	527	G7M	C1'-N9-C4	-4.84	121.41	126.81
1	a	1516	2MG	C5-C6-N1	-4.03	118.26	123.52

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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.