



# Full wwPDB/EMDatabank EM Map/Model Validation Report ⓘ

Feb 19, 2018 – 09:11 am GMT

PDB ID : 5G2X  
EMDB ID: : EMD-3331  
Title : Structure a of Group II Intron Complexed with its Reverse Transcriptase  
Authors : Qu, G.; Kaushal, P.S.; Wang, J.; Shigematsu, H.; Piazza, C.L.; Agrawal, R.K.;  
Belfort, M.; Wang, H.W.  
Deposited on : 2016-04-16  
Resolution : 3.80 Å(reported)

This is a Full wwPDB/EMDatabank EM Map/Model Validation Report  
for a publicly released PDB/EMDB 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.

---

MolProbity : 4.02b-467  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30686

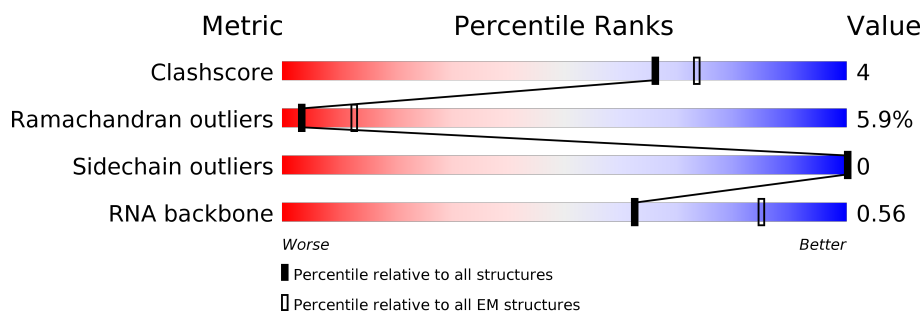
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.80 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)
Clashscore	136279	1886
Ramachandran outliers	132675	1663
Sidechain outliers	132484	1531
RNA backbone	3744	458

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  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	704	 67% 28% . .
2	B	12	 67% 33%
3	C	599	 62% 18% . 19%

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 19072 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 GROUP II INTRON.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	692	Total	C	N	O	P	0	0
			14825	6637	2735	4762	691		

- Molecule 2 is a RNA chain called 5'-R(\*CP\*AP\*CP\*AP\*UP\*CP\*CP\*AP\*UP\*AP\*AP\*CP)-3'.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	12	Total	C	N	O	P	0	0
			250	113	44	81	12		

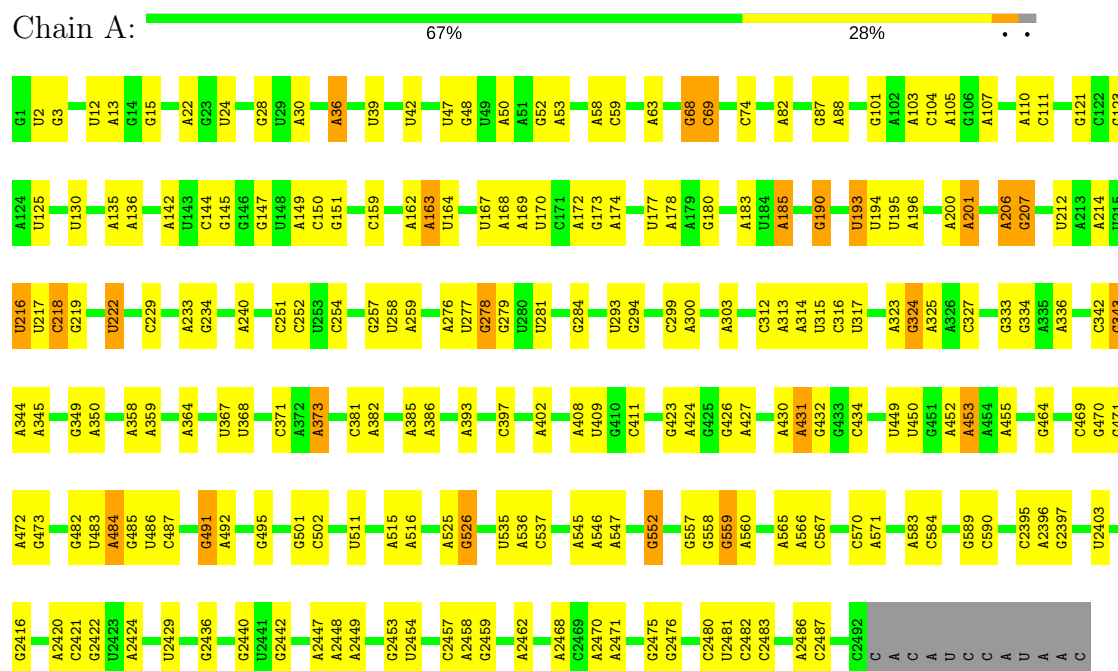
- Molecule 3 is a protein called GROUP II INTRON-ENCODED PROTEIN LTRA.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	487	Total	C	N	O	S	0	1
			3997	2570	688	720	19		

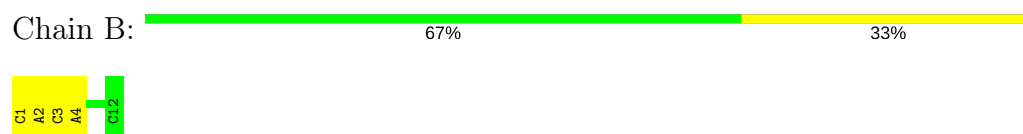
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

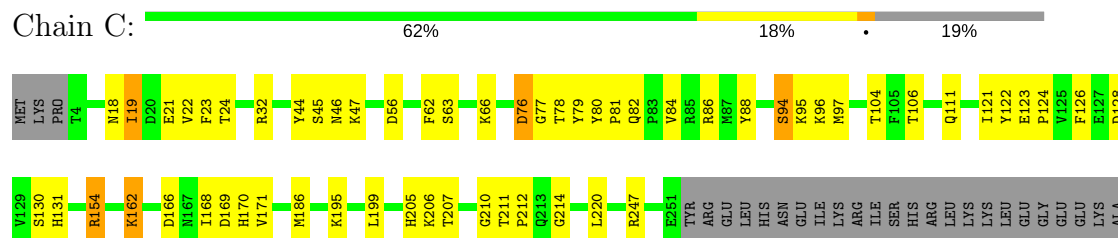
#### • Molecule 1: GROUP II INTRON



#### • Molecule 2: 5'-R>(\*CP\*AP\*CP\*AP\*UP\*CP\*CP\*AP\*UP\*AP\*AP\*CP)-3'



#### • Molecule 3: GROUP II INTRON-ENCODED PROTEIN LTRA





## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	450296	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	INDIVIDUAL PARTICLES	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	Not provided	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	22500	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.11	0/16614	0.65	0/25893
2	B	0.11	0/278	0.65	0/429
3	C	0.25	0/4080	0.42	0/5482
All	All	0.15	0/20972	0.62	0/31804

There are no bond length outliers.

There are no bond angle outliers.

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	14825	0	7458	58	0
2	B	250	0	131	10	0
3	C	3997	0	4065	65	0
All	All	19072	0	11654	124	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 4.

All (124) 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)
2:B:1:C:H2'	2:B:2:A:H8	1.19	1.03

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:1:C:H2'	2:B:2:A:C8	2.01	0.95
2:B:1:C:P	3:C:380:LEU:HD22	2.11	0.90
2:B:1:C:O2'	2:B:2:A:H5'	1.89	0.72
3:C:18:ASN:HB2	3:C:19:ILE:HB	1.73	0.69
3:C:421:ILE:HG12	3:C:470:LYS:HA	1.72	0.69
3:C:483:LYS:HB3	3:C:541:LYS:HD2	1.77	0.66
1:A:22:A:N7	1:A:373:A:N6	2.43	0.65
1:A:15:G:N2	1:A:385:A:N1	2.45	0.65
1:A:2447:A:H61	1:A:2475:G:H1	1.45	0.65
2:B:1:C:O2'	2:B:2:A:C5'	2.46	0.63
1:A:2442:G:H22	1:A:2480:C:H42	1.45	0.63
3:C:154:ARG:HB2	3:C:363:VAL:HB	1.82	0.62
1:A:234:G:H1	1:A:254:C:H42	1.46	0.61
3:C:18:ASN:H	3:C:19:ILE:C	2.02	0.61
3:C:398:ILE:HG23	3:C:403:ILE:HB	1.83	0.61
3:C:84:VAL:HB	3:C:199:LEU:HA	1.83	0.59
1:A:557:G:H4'	3:C:186:MET:H	1.67	0.59
3:C:416:HIS:H	3:C:417:ARG:HA	1.67	0.58
3:C:168:ILE:HG22	3:C:170:HIS:H	1.70	0.57
1:A:411:C:H42	1:A:423:G:H1	1.51	0.57
1:A:323:A:O2'	1:A:324:G:N2	2.37	0.57
3:C:416:HIS:N	3:C:417:ARG:HA	2.21	0.56
1:A:2403:U:H3	1:A:2424:A:H61	1.51	0.56
3:C:199:LEU:HD11	3:C:205:HIS:HB2	1.88	0.56
3:C:63:SER:HB2	3:C:66:LYS:HB3	1.88	0.56
1:A:583:A:O2'	3:C:195:LYS:NZ	2.39	0.56
1:A:121:G:N2	1:A:284:G:O2'	2.39	0.55
3:C:46:ASN:N	3:C:47:LYS:HA	2.22	0.55
1:A:434:C:OP1	1:A:455:A:N6	2.40	0.55
2:B:1:C:O2'	2:B:2:A:O4'	2.22	0.55
1:A:214:A:N7	1:A:218:C:N4	2.56	0.54
1:A:24:U:OP1	1:A:36:A:N6	2.41	0.54
1:A:432:G:H1	1:A:449:U:H3	1.55	0.53
3:C:360:ASP:HB2	3:C:390:LEU:HD23	1.89	0.53
1:A:324:G:H4'	3:C:247:ARG:HH12	1.74	0.53
1:A:565:A:O5'	3:C:32:ARG:NH1	2.41	0.53
3:C:527:TYR:CG	3:C:528:GLY:HA2	2.44	0.53
3:C:44:TYR:N	3:C:45:SER:HA	2.25	0.53
1:A:559:G:N2	1:A:571:A:N1	2.58	0.52
3:C:421:ILE:HG23	3:C:470:LYS:HG3	1.89	0.52
3:C:302:LYS:HB3	3:C:313:SER:HB2	1.92	0.52

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:222:U:O4	2:B:4:A:N6	2.43	0.51
1:A:430:A:H3'	1:A:431:A:H8	1.76	0.51
1:A:193:U:O4'	1:A:342:C:N4	2.44	0.50
2:B:1:C:P	3:C:380:LEU:CD2	2.94	0.50
2:B:1:C:OP1	3:C:380:LEU:HD22	2.11	0.50
1:A:312:C:H2'	1:A:313:A:C8	2.47	0.50
1:A:313:A:H2'	1:A:314:A:C8	2.47	0.49
1:A:185:A:N6	1:A:201:A:OP1	2.45	0.49
3:C:97:MET:SD	3:C:162:LYS:NZ	2.73	0.49
3:C:86:ARG:HH11	3:C:88:TYR:HE2	1.61	0.49
3:C:550:ASP:HB2	3:C:555:TYR:HE1	1.78	0.49
1:A:190:G:O6	1:A:343:G:O2'	2.30	0.48
3:C:169:ASP:O	3:C:171:VAL:N	2.46	0.48
3:C:316:GLY:HA2	3:C:317:SER:HA	1.53	0.48
3:C:375:VAL:N	3:C:376:LYS:HB2	2.29	0.48
3:C:487:GLY:HA2	3:C:488:SER:HA	1.65	0.47
1:A:552:G:H1	1:A:584:C:H42	1.62	0.47
3:C:84:VAL:HG13	3:C:212:PRO:HD2	1.96	0.47
1:A:2459:G:N2	1:A:2462:A:OP2	2.46	0.47
3:C:332:ILE:O	3:C:337:LYS:N	2.45	0.47
3:C:76:ASP:HA	3:C:77:GLY:HA2	1.57	0.47
1:A:483:U:O2'	1:A:484:A:N7	2.40	0.46
3:C:210:GLY:HA3	3:C:212:PRO:HD3	1.97	0.46
3:C:436:LEU:HD11	3:C:459:MET:HE2	1.97	0.46
1:A:183:A:H61	1:A:350:A:H61	1.64	0.46
3:C:104:THR:HG22	3:C:106:THR:H	1.80	0.46
1:A:2440:G:N2	1:A:2483:C:O2	2.49	0.46
3:C:349:HIS:ND1	3:C:349:HIS:O	2.47	0.46
3:C:123:GLU:HA	3:C:124:PRO:HA	1.67	0.46
3:C:211:THR:OG1	3:C:214:GLY:O	2.33	0.46
3:C:356:PHE:O	3:C:358:GLY:N	2.50	0.45
1:A:2436:G:H1	1:A:2487:C:H42	1.63	0.45
3:C:122:TYR:HA	3:C:126:PHE:HB3	1.99	0.45
3:C:407:LYS:HA	3:C:408:LYS:HB2	1.98	0.45
1:A:281:U:H4'	1:A:2416:G:H21	1.82	0.44
1:A:47:U:H2'	1:A:48:G:H8	1.82	0.44
1:A:162:A:H5''	1:A:163:A:C8	2.53	0.44
1:A:430:A:H3'	1:A:431:A:C8	2.53	0.44
3:C:390:LEU:HD11	3:C:393:LYS:HG2	2.00	0.44
3:C:210:GLY:HA3	3:C:211:THR:HA	1.58	0.43
3:C:415:VAL:HA	3:C:416:HIS:HA	1.53	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:23:PHE:HA	3:C:24:THR:HA	1.61	0.43
1:A:206:A:H2'	1:A:207:G:H4'	1.99	0.43
1:A:2448:A:H2'	1:A:2449:A:H8	1.83	0.43
3:C:556:GLU:O	3:C:558:HIS:ND1	2.48	0.43
1:A:30:A:H61	1:A:42:U:H3	1.66	0.43
1:A:491:G:H5''	1:A:492:A:H8	1.84	0.43
1:A:397:C:H42	1:A:464:G:H1	1.65	0.42
3:C:389:PRO:HD3	3:C:446:ALA:HB1	2.01	0.42
3:C:484:ASP:HB2	3:C:485:GLY:HA2	2.00	0.42
3:C:170:HIS:ND1	3:C:206:LYS:O	2.51	0.42
1:A:525:A:H3'	1:A:526:G:H8	1.83	0.42
1:A:545:A:H61	1:A:590:C:H42	1.68	0.42
1:A:2457:C:H2'	1:A:2458:A:C8	2.54	0.42
1:A:570:C:H2'	1:A:571:A:H8	1.84	0.42
3:C:332:ILE:HG23	3:C:336:LEU:HD12	2.00	0.42
1:A:173:G:H2'	1:A:174:A:C8	2.55	0.42
3:C:78:THR:HA	3:C:79:TYR:HA	1.53	0.41
3:C:401:LYS:HE2	3:C:403:ILE:HD11	2.01	0.41
1:A:278:G:H1'	3:C:405:ILE:HD12	2.02	0.41
1:A:349:G:H2'	1:A:350:A:C8	2.55	0.41
1:A:216:U:OP2	1:A:313:A:N6	2.50	0.41
1:A:314:A:H2'	1:A:315:U:C6	2.56	0.41
3:C:342:GLU:C	3:C:344:LYS:H	2.24	0.41
3:C:414:PRO:HG3	3:C:461:TYR:HB2	2.03	0.41
3:C:375:VAL:HA	3:C:377:LYS:N	2.36	0.41
1:A:167:U:H5''	1:A:168:A:H5'	2.03	0.41
2:B:3:C:H2'	2:B:4:A:C8	2.56	0.41
1:A:2440:G:O6	1:A:2482:C:N4	2.46	0.41
1:A:172:A:H2'	1:A:173:G:C8	2.56	0.41
1:A:2475:G:H2'	1:A:2476:G:C8	2.56	0.41
1:A:450:U:O2'	1:A:453:A:N6	2.53	0.41
1:A:123:G:O6	1:A:159:C:N4	2.52	0.40
1:A:349:G:H2'	1:A:350:A:H8	1.87	0.40
3:C:111:GLN:HE22	3:C:220:LEU:HD11	1.86	0.40
3:C:195:LYS:HB3	3:C:206:LYS:HD3	2.03	0.40
3:C:390:LEU:HD13	3:C:392:ASP:HB2	2.03	0.40
1:A:68:G:HO2'	1:A:69:C:P	2.45	0.40
3:C:94:SER:O	3:C:96:LYS:N	2.54	0.40
1:A:381:C:H2'	1:A:382:A:C8	2.57	0.40
3:C:121:ILE:HG22	3:C:126:PHE:HB2	2.03	0.40
3:C:62:PHE:HB3	3:C:66:LYS:HG2	2.03	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	477/599 (80%)	355 (74%)	94 (20%)	28 (6%)	2	25

All (28) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	22	VAL
3	C	81	PRO
3	C	95	LYS
3	C	349	HIS
3	C	357	LEU
3	C	414	PRO
3	C	56	ASP
3	C	76	ASP
3	C	154	ARG
3	C	207	THR
3	C	406	GLN
3	C	415	VAL
3	C	537	ARG
3	C	21	GLU
3	C	94	SER
3	C	128	ASP
3	C	130	SER
3	C	131	HIS
3	C	162	LYS
3	C	166	ASP
3	C	344	LYS
3	C	346	LEU
3	C	358	GLY
3	C	496	LYS
3	C	82	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	C	314	VAL
3	C	19	ILE
3	C	80	TYR

### 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	
3	C	441/545 (81%)	441 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
3	C	111	GLN
3	C	140	HIS
3	C	213	GLN
3	C	453	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	688/704 (97%)	146 (21%)	7 (1%)
2	B	11/12 (91%)	0	0
All	All	699/716 (97%)	146 (20%)	7 (1%)

All (146) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	2	U
1	A	3	G
1	A	12	U
1	A	13	A

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	28	G
1	A	36	A
1	A	39	U
1	A	50	A
1	A	52	G
1	A	53	A
1	A	58	A
1	A	59	C
1	A	63	A
1	A	69	C
1	A	74	C
1	A	82	A
1	A	87	G
1	A	88	A
1	A	101	G
1	A	104	C
1	A	105	A
1	A	107	A
1	A	110	A
1	A	111	C
1	A	125	U
1	A	130	U
1	A	135	A
1	A	136	A
1	A	142	A
1	A	144	C
1	A	145	G
1	A	147	G
1	A	149	A
1	A	150	C
1	A	151	G
1	A	163	A
1	A	164	U
1	A	169	A
1	A	170	U
1	A	177	U
1	A	178	A
1	A	180	G
1	A	185	A
1	A	190	G
1	A	193	U
1	A	194	U

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	195	U
1	A	196	A
1	A	200	A
1	A	201	A
1	A	206	A
1	A	207	G
1	A	212	U
1	A	216	U
1	A	217	U
1	A	218	C
1	A	219	G
1	A	222	U
1	A	229	C
1	A	233	A
1	A	240	A
1	A	251	C
1	A	252	C
1	A	257	G
1	A	258	U
1	A	259	A
1	A	276	A
1	A	277	U
1	A	278	G
1	A	279	G
1	A	293	U
1	A	294	G
1	A	299	C
1	A	300	A
1	A	303	A
1	A	316	C
1	A	317	U
1	A	324	G
1	A	325	A
1	A	327	C
1	A	333	G
1	A	334	G
1	A	336	A
1	A	343	G
1	A	344	A
1	A	345	A
1	A	358	A
1	A	359	A

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	364	A
1	A	367	U
1	A	368	U
1	A	371	C
1	A	373	A
1	A	386	A
1	A	393	A
1	A	402	A
1	A	408	A
1	A	409	U
1	A	424	A
1	A	426	G
1	A	427	A
1	A	431	A
1	A	452	A
1	A	453	A
1	A	469	C
1	A	470	G
1	A	471	G
1	A	472	A
1	A	473	G
1	A	482	G
1	A	484	A
1	A	485	G
1	A	486	U
1	A	487	C
1	A	491	G
1	A	495	G
1	A	501	G
1	A	502	C
1	A	511	U
1	A	515	A
1	A	516	A
1	A	526	G
1	A	535	U
1	A	536	A
1	A	537	C
1	A	546	A
1	A	547	A
1	A	552	G
1	A	558	G
1	A	560	A

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	566	A
1	A	567	C
1	A	589	G
1	A	2396	A
1	A	2397	G
1	A	2420	A
1	A	2421	C
1	A	2422	G
1	A	2429	U
1	A	2453	G
1	A	2454	U
1	A	2468	A
1	A	2470	A
1	A	2471	A
1	A	2481	U
1	A	2486	A

All (7) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	68	G
1	A	103	A
1	A	149	A
1	A	177	U
1	A	546	A
1	A	559	G
1	A	2395	C

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

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.



## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	3

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	598:G	O3'	2388:A	P	22.02
1	A	572:A	O3'	579:A	P	19.41
1	A	414:A	O3'	420:A	P	15.60