



wwPDB EM Validation Summary Report ⓘ

Sep 18, 2021 – 08:04 am BST

PDB ID : 7OQC
EMDB ID : EMD-13029
Title : The U1 part of *Saccharomyces cerevisiae* spliceosomal pre-A complex (delta BS-A ACT1)
Authors : Zhang, Z.; Rigo, N.; Dybkov, O.; Fourmann, J.; Will, C.L.; Kumar, V.; Urlaub, H.; Stark, H.; Luehrmann, R.
Deposited on : 2021-06-03
Resolution : 4.10 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

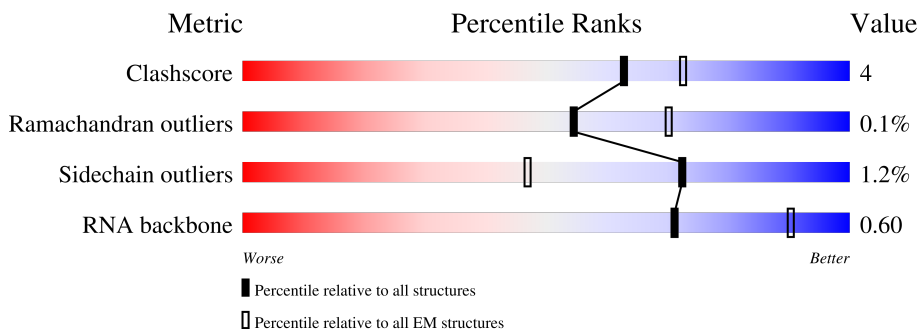
EMDB validation analysis : **FAILED**
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.23.1

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 4.10 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.













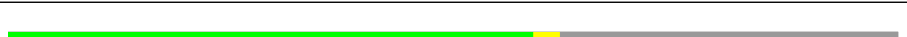
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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 ≥ 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	F	523	47% (green), 49% (grey), 4% (yellow), 0% (orange), 0% (red)
2	I	371	16% (green), 83% (grey), 1% (yellow), 0% (orange), 0% (red)
3	E	544	87% (green), 12% (yellow), 1% (orange), 0% (red), 0% (grey)
4	J	620	16% (green), 83% (grey), 1% (yellow), 0% (orange), 0% (red)
5	1	568	56% (green), 37% (yellow), 5% (orange), 0% (red), 0% (grey)
6	G	492	41% (green), 7% (yellow), 51% (grey), 0% (orange), 0% (red)
7	A	298	38% (green), 6% (yellow), 56% (grey), 0% (orange), 0% (red)

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	C	231	 77% 8% 16%
9	b	196	 62% 38%
10	d	101	 90% 8%
11	e	94	 80% 18%
12	f	86	 85% 15%
13	g	77	 91% 6%
14	h	146	 73% 27%
15	i	110	 89% 10%
16	H	261	 71% 26%
17	D	629	 88% 8%
18	B	300	 59% 38%

2 Entry composition i

There are 18 unique types of molecules in this entry. The entry contains 35050 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 protein called Protein NAM8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	F	267	Total	C	N	O	S	0	0
			1758	1091	316	341	10		

- Molecule 2 is a RNA chain called ACT1 pre-mRNA (delta BS-A),ACT1 pre-mRNA (delta BS-A),ACT1 pre-mRNA (delta BS-A),ACT1 pre-mRNA (delta BS-A),ACT1 pre-mRNA (delta BS-A),ACT1 pre-mRNA (delta BS-A),ACT1 pre-mRNA (delta BS-A),ACT1 pre-mRNA (delta BS-A),ACT1 pre-mRNA (delta BS-A).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	I	63	Total	C	N	O	P	0	0
			847	359	33	392	63		

- Molecule 3 is a protein called U1 small nuclear ribonucleoprotein component PRP42.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	E	544	Total	C	N	O	S	0	0
			4561	2990	723	828	20		

- Molecule 4 is a protein called U1 small nuclear ribonucleoprotein component SNU71.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	J	105	Total	C	N	O	S	0	0
			687	427	127	132	1		

- Molecule 5 is a RNA chain called U1 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
5	1	558	Total	C	N	O	P	0	0
			11822	5287	2003	3974	558		

- Molecule 6 is a protein called 56 kDa U1 small nuclear ribonucleoprotein component.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	G	239	1954	1267	321	354	12	0	0

- Molecule 7 is a protein called U1 small nuclear ribonucleoprotein A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	A	132	1058	674	193	187	4	0	0

- Molecule 8 is a protein called U1 small nuclear ribonucleoprotein C.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	C	195	1570	976	301	288	5	0	0

- Molecule 9 is a protein called Small nuclear ribonucleoprotein-associated protein B.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
9	b	121	972	613	183	173	3	0	0

- Molecule 10 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	d	93	714	453	125	133	3	0	0

- Molecule 11 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
11	e	77	600	395	96	106	3	0	0

- Molecule 12 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
12	f	73	585	376	102	106	1	0	0

- Molecule 13 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	g	72	Total	C	N	O	S	0	0
			556	352	97	105	2		

- Molecule 14 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	h	107	Total	C	N	O	S	0	0
			771	487	138	143	3		

- Molecule 15 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	i	99	Total	C	N	O	S	0	0
			805	514	148	139	4		

- Molecule 16 is a protein called Protein LUC7.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	H	192	Total	C	N	O	S	0	0
			1201	740	226	228	7		


- Molecule 17 is a protein called Pre-mRNA-processing factor 39.

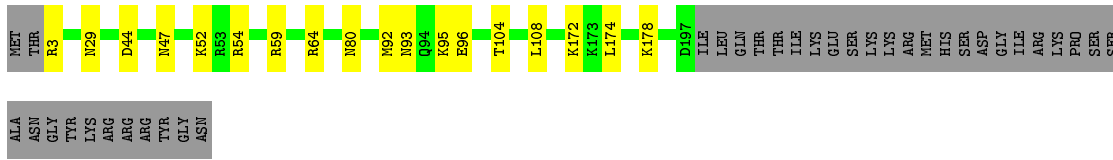
Mol	Chain	Residues	Atoms					AltConf	Trace
17	D	576	Total	C	N	O	S	0	0
			3530	2204	642	680	4		

- Molecule 18 is a protein called U1 small nuclear ribonucleoprotein 70 kDa homolog.

Mol	Chain	Residues	Atoms				AltConf	Trace
18	B	186	Total	C	N	O	0	0
			1059	647	206	206		

- Molecule 8: U1 small nuclear ribonucleoprotein C

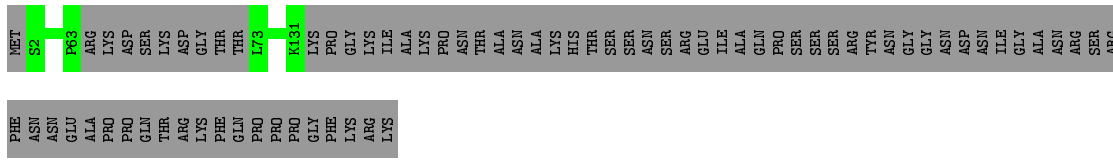
Chain C:  77% 8% 16%



ALA, ASN, GLY, TYR, LYS, ARG, ARG, ARG, TYR, GLY, ASN

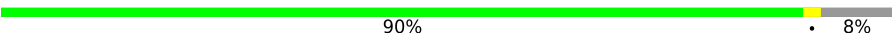
- Molecule 9: Small nuclear ribonucleoprotein-associated protein B

Chain b:  62% 38%




PHE, ASN, ASN, GLU, ALA, PRO, GLN, THR, ARG, LYS, PHE, GLN, PRO, PRO, PHE, GLY, ARG, LYS

- Molecule 10: Small nuclear ribonucleoprotein Sm D3

Chain d:  90% 8%




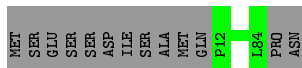
- Molecule 11: Small nuclear ribonucleoprotein E

Chain e:  80% 18%




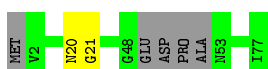
- Molecule 12: Small nuclear ribonucleoprotein F

Chain f:  85% 15%



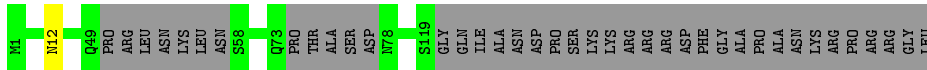
- Molecule 13: Small nuclear ribonucleoprotein G

Chain g:  91% 6%

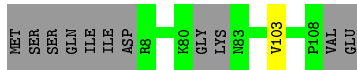
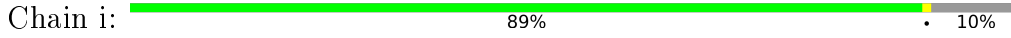


- Molecule 14: Small nuclear ribonucleoprotein Sm D1

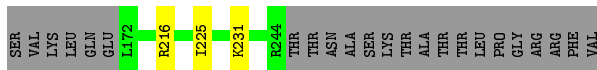
Chain h:  73% 27%



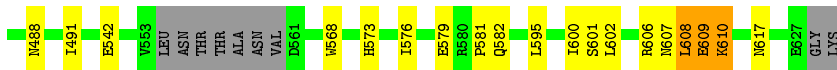
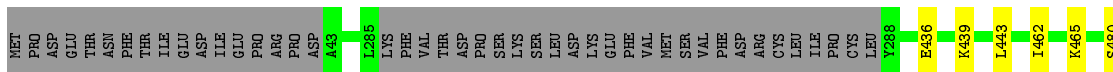
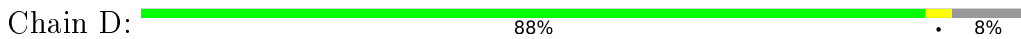
- Molecule 15: Small nuclear ribonucleoprotein Sm D2



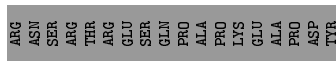
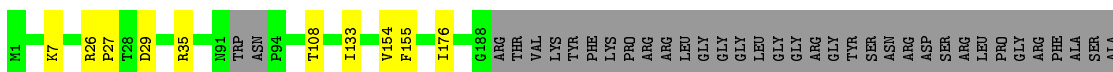
- Molecule 16: Protein LUC7



- Molecule 17: Pre-mRNA-processing factor 39



- Molecule 18: U1 small nuclear ribonucleoprotein 70 kDa homolog



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	217460	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	44	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	FEI FALCON III (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 > 5$	RMSZ	# $ Z > 5$
1	F	0.27	0/1784	0.52	0/2436
2	I	0.23	0/234	0.96	1/362 (0.3%)
3	E	0.27	0/4676	0.51	1/6320 (0.0%)
4	J	0.26	0/695	0.48	1/948 (0.1%)
5	1	0.42	1/13201 (0.0%)	1.01	30/20553 (0.1%)
6	G	0.27	0/1996	0.51	0/2682
7	A	0.26	0/1072	0.57	0/1437
8	C	0.25	0/1601	0.51	0/2154
9	b	0.26	0/978	0.61	0/1306
10	d	0.28	0/726	0.60	1/984 (0.1%)
11	e	0.28	0/610	0.58	0/826
12	f	0.28	0/597	0.56	0/807
13	g	0.26	0/559	0.57	0/751
14	h	0.25	0/776	0.52	0/1053
15	i	0.25	0/818	0.51	0/1099
16	H	0.24	0/1212	0.40	0/1652
17	D	0.25	0/3570	0.42	0/4924
18	B	0.25	0/1071	0.45	0/1482
All	All	0.33	1/36176 (0.0%)	0.75	34/51776 (0.1%)

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
7	A	0	1
13	g	0	1
18	B	0	1
All	All	0	3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	1	325	A	O3'-P	-33.85	1.20	1.61

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	1	442	U	OP2-P-O3'	-10.46	82.18	105.20
5	1	442	U	OP1-P-O3'	-10.44	82.24	105.20
3	E	310	ASP	CB-CG-OD1	9.75	127.08	118.30
4	J	274	PRO	CA-N-CD	-8.53	99.56	111.50
5	1	289	U	N3-C2-O2	-7.13	117.21	122.20

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
7	A	12	ARG	Peptide
18	B	176	ILE	Peptide
13	g	20	ASN	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	F	1758	0	1443	14	0
2	I	847	0	532	11	0
3	E	4561	0	4549	43	0
4	J	687	0	534	12	0
5	1	11822	0	5940	91	0
6	G	1954	0	1961	23	0
7	A	1058	0	1118	10	0
8	C	1570	0	1555	10	0
9	b	972	0	1048	0	0
10	d	714	0	738	0	0
11	e	600	0	634	0	0
12	f	585	0	587	0	0
13	g	556	0	583	0	0
14	h	771	0	751	0	0
15	i	805	0	834	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
16	H	1201	0	902	4	0
17	D	3530	0	2469	24	0
18	B	1059	0	651	5	0
All	All	35050	0	26829	206	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.

The worst 5 of 206 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)
5:1:384:U:H3	5:1:434:G:H1	1.01	0.97
5:1:380:G:H1	5:1:438:U:H3	1.07	0.96
4:J:273:LEU:N	4:J:274:PRO:HD3	1.89	0.87
5:1:389:G:H1	5:1:430:U:H3	0.86	0.85
4:J:272:ASP:C	4:J:274:PRO:HD3	1.97	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 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	
1	F	259/523 (50%)	251 (97%)	8 (3%)	0	100	100
3	E	542/544 (100%)	521 (96%)	21 (4%)	0	100	100
4	J	101/620 (16%)	92 (91%)	8 (8%)	1 (1%)	15	52
6	G	235/492 (48%)	222 (94%)	13 (6%)	0	100	100
7	A	126/298 (42%)	116 (92%)	10 (8%)	0	100	100
8	C	193/231 (84%)	183 (95%)	10 (5%)	0	100	100
9	b	117/196 (60%)	110 (94%)	7 (6%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
10	d	91/101 (90%)	87 (96%)	4 (4%)	0	100	100
11	e	73/94 (78%)	67 (92%)	5 (7%)	1 (1%)	11	45
12	f	71/86 (83%)	69 (97%)	2 (3%)	0	100	100
13	g	68/77 (88%)	62 (91%)	5 (7%)	1 (2%)	10	44
14	h	101/146 (69%)	98 (97%)	3 (3%)	0	100	100
15	i	95/110 (86%)	91 (96%)	4 (4%)	0	100	100
16	H	186/261 (71%)	180 (97%)	6 (3%)	0	100	100
17	D	570/629 (91%)	554 (97%)	16 (3%)	0	100	100
18	B	182/300 (61%)	169 (93%)	13 (7%)	0	100	100
All	All	3010/4708 (64%)	2872 (95%)	135 (4%)	3 (0%)	54	84

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	J	274	PRO
11	e	34	GLN
13	g	21	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	
1	F	138/451 (31%)	137 (99%)	1 (1%)	84	90
3	E	508/519 (98%)	504 (99%)	4 (1%)	81	88
4	J	49/568 (9%)	47 (96%)	2 (4%)	30	57
6	G	218/448 (49%)	218 (100%)	0	100	100
7	A	117/273 (43%)	114 (97%)	3 (3%)	46	67
8	C	171/214 (80%)	166 (97%)	5 (3%)	42	64
9	b	108/176 (61%)	108 (100%)	0	100	100
10	d	81/89 (91%)	80 (99%)	1 (1%)	71	83

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
11	e	68/83 (82%)	67 (98%)	1 (2%)	65	79
12	f	65/77 (84%)	65 (100%)	0	100	100
13	g	62/66 (94%)	62 (100%)	0	100	100
14	h	75/129 (58%)	74 (99%)	1 (1%)	69	81
15	i	90/103 (87%)	89 (99%)	1 (1%)	73	84
16	H	72/234 (31%)	72 (100%)	0	100	100
17	D	193/603 (32%)	189 (98%)	4 (2%)	53	72
18	B	40/265 (15%)	39 (98%)	1 (2%)	47	68
All	All	2055/4298 (48%)	2031 (99%)	24 (1%)	72	83

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
8	C	172	LYS
14	h	12	ASN
11	e	10	MET
15	i	103	VAL
4	J	51	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
17	D	488	ASN
17	D	573	HIS
3	E	474	GLN
3	E	459	ASN
17	D	582	GLN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	I	9/371 (2%)	0	0
5	1	556/568 (97%)	117 (21%)	9 (1%)
All	All	565/939 (60%)	117 (20%)	9 (1%)

5 of 117 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
5	1	11	U
5	1	12	A
5	1	40	A
5	1	41	C
5	1	55	G

5 of 9 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
5	1	399	A
5	1	505	U
5	1	151	C
5	1	152	G
5	1	258	U

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 monosaccharides 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	I	2
5	1	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	I	8:C	O3'	103:N	P	54.06
1	I	128:N	O3'	130:N	P	17.87
1	1	325:A	O3'	326:G	P	1.20