



wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 01:40 PM BST

PDB ID : 2Y7H
EMDB ID: : EMD-1534
Title : Atomic model of the DNA-bound methylase complex from the Type I restrict ion-modification enzyme EcoKI (M2S1). Based on fitting into EM map 1534.
Authors : Kennaway, C.K.; Obarska-Kosinska, A.; White, J.H.; Tuszynska, I.; Cooper, L.P.; Bujnicki, J.M.; Trinick, J.; Dryden, D.T.F.
Deposited on : 2011-01-31
Resolution : 18.00 Å(reported)
Based on PDB ID : 1S7Z, 1YF2, 2AR0

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>

MolProbity : 4.02b-467
Mogul : 1.7.1 (RC1), CSD as537be (2016)
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : trunk27241

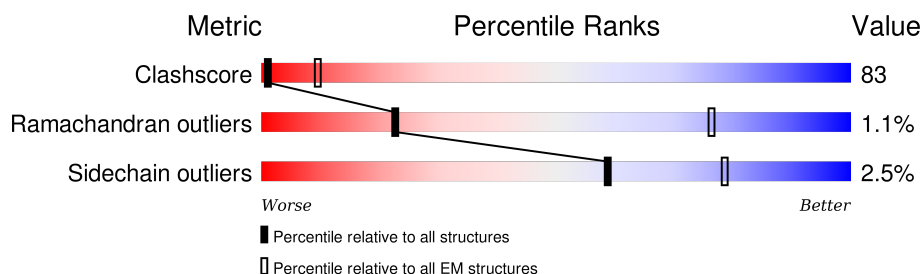
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 18.00 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	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	464	24% 71% 5%
2	B	529	38% 59% .
2	C	529	37% 61% .
3	D	20	25% 70% 5%
4	E	20	30% 70%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	SAM	B	530	-	-	X	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	SAM	C	530	-	-	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 12846 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 TYPE-1 RESTRICTION ENZYME ECOKI SPECIFICITY PROTEIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	464	Total	C	N	O	S	0	0
			3622	2298	644	671	9		

- Molecule 2 is a protein called TYPE I RESTRICTION ENZYME ECOKI M PROTEIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	529	Total	C	N	O	S	0	0
			4175	2612	730	816	17		
2	C	529	Total	C	N	O	S	0	0
			4175	2612	730	816	17		

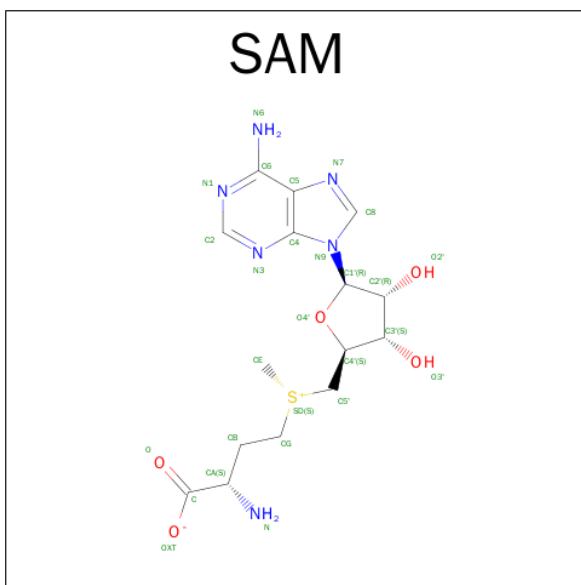
- Molecule 3 is a DNA chain called 5'-D(*GP*TP*TP*CP*AP*AP*CP*GP*TP*CP*GP*A P*CP*GP *TP*GP*CP*AP*AP*C)-3'.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	20	Total	C	N	O	P	0	0
			409	194	76	119	20		

- Molecule 4 is a DNA chain called 5'-D(*GP*TP*TP*GP*CP*AP*CP*GP*TP*CP*GP*A P*CP*GP *TP*TP*GP*AP*AP*C)-3'.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	E	20	Total	C	N	O	P	0	0
			411	195	75	121	20		

- Molecule 5 is S-ADENOSYLMETHIONINE (three-letter code: SAM) (formula: C₁₅H₂₂N₆O₅S).

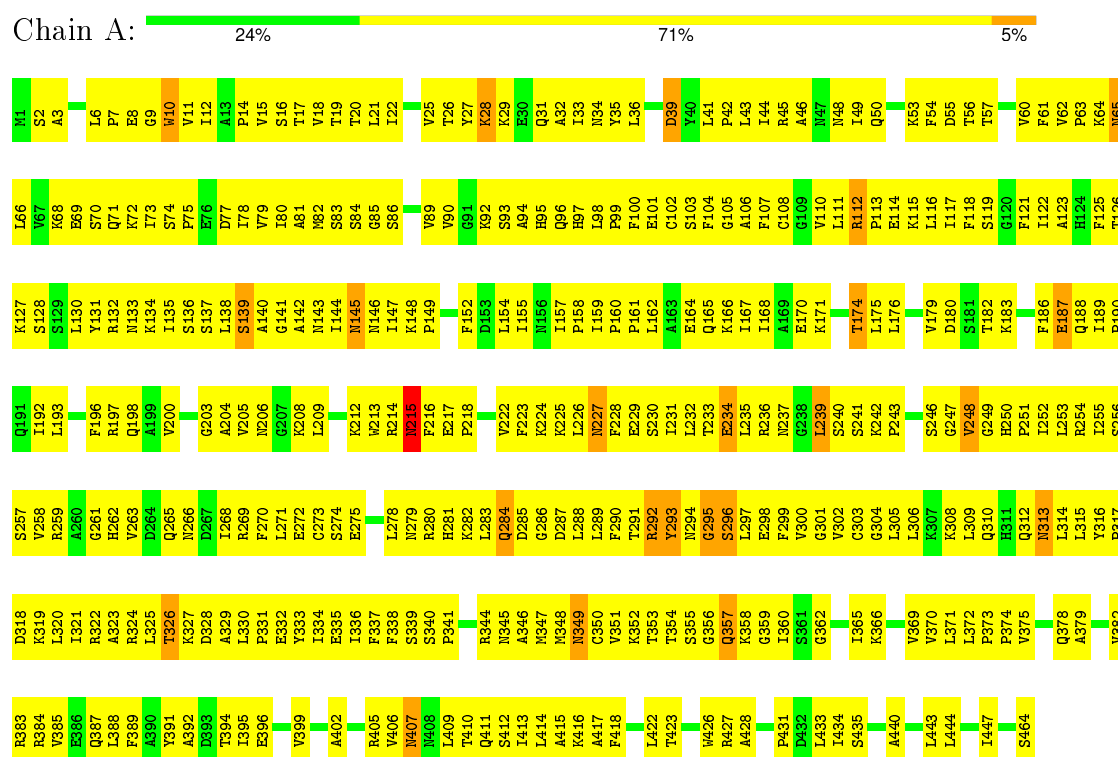


Mol	Chain	Residues	Atoms					AltConf
5	B	1	Total 27	C 15	N 6	O 5	S 1	0
5	C	1	Total 27	C 15	N 6	O 5	S 1	0

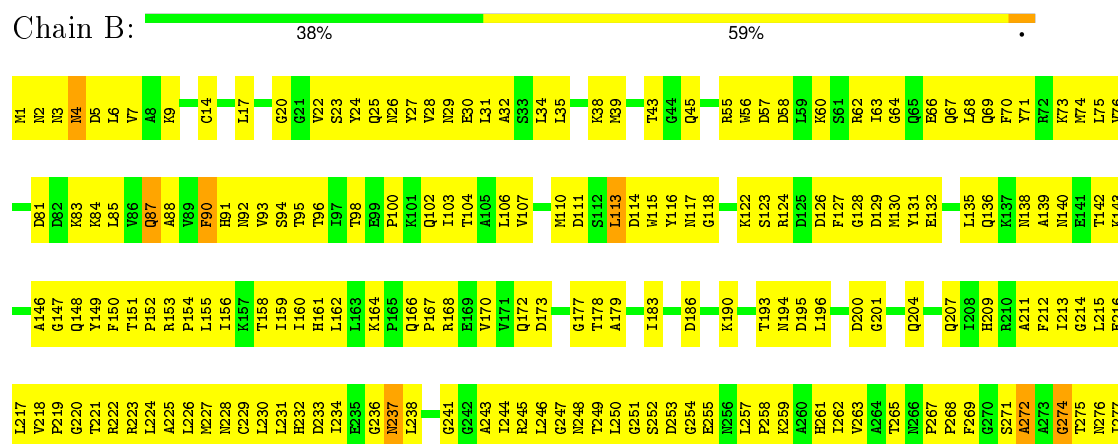
3 Residue-property plots

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: TYPE-1 RESTRICTION ENZYME ECOKI SPECIFICITY PROTEIN

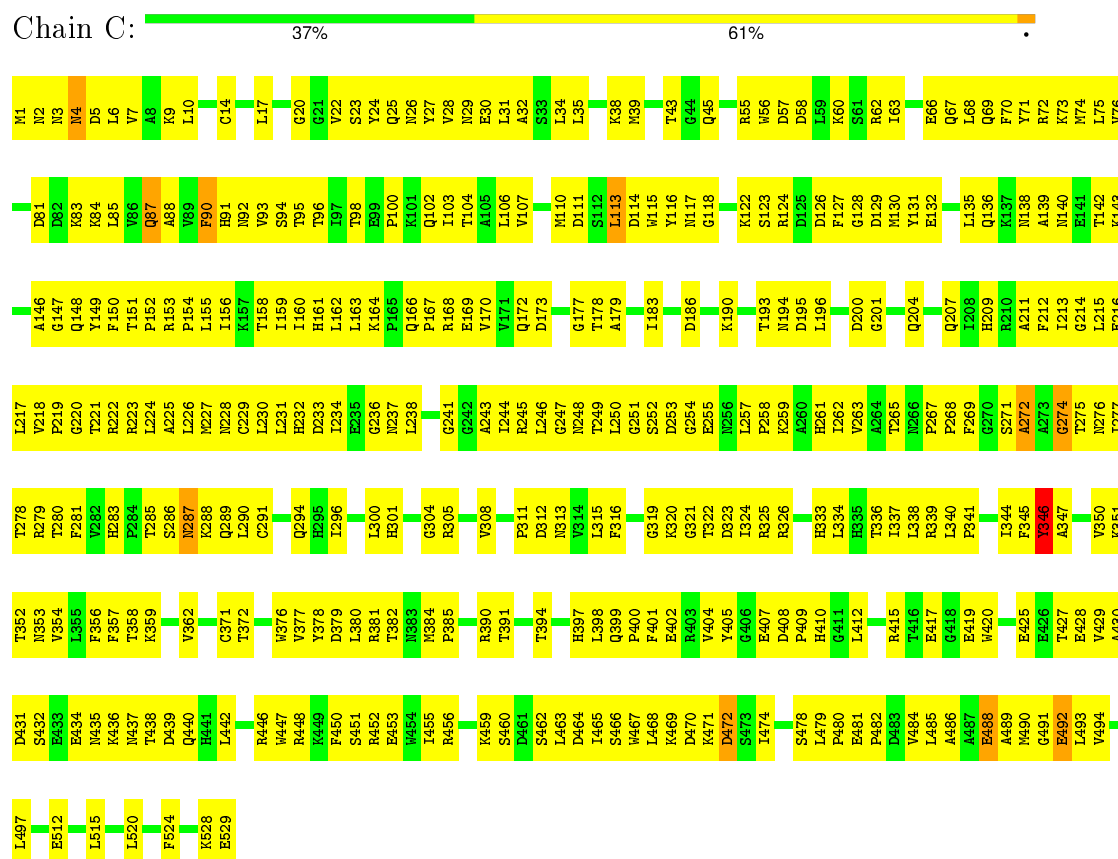


• Molecule 2: TYPE I RESTRICTION ENZYME ECOKI M PROTEIN





• Molecule 2: TYPE I RESTRICTION ENZYME ECOKI M PROTEIN



• Molecule 3: 5'-D(*GP*TP*TP*CP*AP*AP*CP*GP*TP*CP*GP*AP*CP*GP *TP*GP*CP*AP*AP*C)-3'



• Molecule 4: 5'-D(*GP*TP*TP*GP*CP*AP*CP*GP*TP*CP*GP*AP*CP*GP *TP*TP*GP*AP*AP*C)-3'



G1	T2	T3	C4	C5	A6	C7	C8	T9	C10	G11	A12	C13	G14	T15	T16	G17	A18	A19	C20
----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	FILTERED AT FIRST ZERO	Depositor
Microscope	OTHER	Depositor
Voltage (kV)	80	Depositor
Electron dose ($e^-/\text{\AA}^2$)	25	Depositor
Minimum defocus (nm)	275	Depositor
Maximum defocus (nm)	870	Depositor
Magnification	40000	Depositor
Image detector	KODAK SO163 FILM	Depositor

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: SAM

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	0.98	2/3685 (0.1%)	1.05	3/4968 (0.1%)
2	B	0.94	0/4262	1.05	3/5773 (0.1%)
2	C	0.94	0/4262	1.02	2/5773 (0.0%)
3	D	3.89	80/458 (17.5%)	5.67	167/704 (23.7%)
4	E	3.95	79/460 (17.2%)	5.69	171/708 (24.2%)
All	All	1.39	161/13127 (1.2%)	1.88	346/17926 (1.9%)

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
1	A	0	1
2	B	0	1
3	D	1	1
All	All	1	3

The worst 5 of 161 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	292	ARG	C-N	15.10	1.68	1.34
3	D	14	DG	N3-C4	13.28	1.44	1.35
3	D	11	DG	N3-C4	13.10	1.44	1.35
3	D	16	DG	N3-C4	13.07	1.44	1.35
4	E	1	DG	N3-C4	13.03	1.44	1.35

The worst 5 of 346 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
4	E	4	DG	C4-C5-N7	-24.73	100.91	110.80
3	D	1	DG	C4-C5-N7	-24.21	101.12	110.80
4	E	8	DG	C4-C5-N7	-24.19	101.12	110.80
4	E	17	DG	C4-C5-N7	-24.18	101.13	110.80
4	E	11	DG	C4-C5-N7	-24.05	101.18	110.80

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	D	15	DT	C4'

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	293	TYR	Sidechain
2	B	470	ASP	Mainchain
3	D	18	DA	Sidechain

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	3622	0	3738	1029	0
2	B	4175	0	4050	560	0
2	C	4175	0	4050	582	0
3	D	409	0	225	91	0
4	E	411	0	226	80	0
5	B	27	0	22	28	0
5	C	27	0	22	30	0
All	All	12846	0	12333	2091	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 83.

The worst 5 of 2091 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)
1:A:292:ARG:C	1:A:293:TYR:N	1.68	1.43
1:A:209:LEU:HB2	1:A:216:PHE:CZ	1.60	1.36
1:A:3:ALA:O	2:B:488:GLU:HG3	1.22	1.28
1:A:412:SER:HB2	2:B:495:GLN:CG	1.68	1.23
2:C:512:GLU:HB2	2:C:515:LEU:HD23	1.20	1.17

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	A	462/464 (100%)	433 (94%)	22 (5%)	7 (2%)	13	57
2	B	527/529 (100%)	503 (95%)	18 (3%)	6 (1%)	17	63
2	C	527/529 (100%)	503 (95%)	20 (4%)	4 (1%)	24	69
All	All	1516/1522 (100%)	1439 (95%)	60 (4%)	17 (1%)	23	63

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	139	SER
1	A	296	SER
1	A	326	THR
1	A	248	VAL
2	B	472	ASP

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	A	398/398 (100%)	383 (96%)	15 (4%)	40	73
2	B	452/452 (100%)	443 (98%)	9 (2%)	63	85
2	C	452/452 (100%)	443 (98%)	9 (2%)	63	85
All	All	1302/1302 (100%)	1269 (98%)	33 (2%)	59	81

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	407	ASN
2	B	287	ASN
2	C	453	GLU
2	B	87	GLN
2	B	113	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 65 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	140	ASN
2	B	294	GLN
2	C	397	HIS
2	B	232	HIS
2	B	261	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry

2 ligands 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
5	SAM	B	530	-	23,29,29	1.18	2 (8%)	15,42,42	2.86	2 (13%)
5	SAM	C	530	-	23,29,29	1.17	2 (8%)	15,42,42	2.88	2 (13%)

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
5	SAM	B	530	-	-	0/8/33/33	0/3/3/3
5	SAM	C	530	-	-	0/8/33/33	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	530	SAM	C2-N1	2.67	1.39	1.33
5	B	530	SAM	C2-N1	2.70	1.39	1.33
5	B	530	SAM	C2-N3	3.73	1.38	1.32
5	C	530	SAM	C2-N3	3.74	1.38	1.32

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	530	SAM	N3-C2-N1	-10.47	120.65	128.87
5	B	530	SAM	N3-C2-N1	-10.42	120.69	128.87
5	B	530	SAM	O2'-C2'-C3'	2.11	118.67	111.86
5	C	530	SAM	O2'-C2'-C3'	2.13	118.75	111.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 58 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	530	SAM	28	0
5	C	530	SAM	30	0

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.