



Full wwPDB X-ray Structure Validation Report i

Feb 28, 2014 – 06:19 PM GMT

PDB ID : 4IKP
Title : Crystal structure of coactivator-associated arginine methyltransferase 1
Authors : Dong, A.; Dombrowski, L.; He, H.; Ibanez, G.; Wernimont, A.; Zheng, W.; Bountra, C.; Arrowsmith, C.H.; Edwards, A.M.; Brown, P.J.; Min, J.; Luo, M.; Wu, H.; Structural Genomics Consortium (SGC)
Deposited on : 2012-12-27
Resolution : 2.00 Å (reported)

This is a full wwPDB validation report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

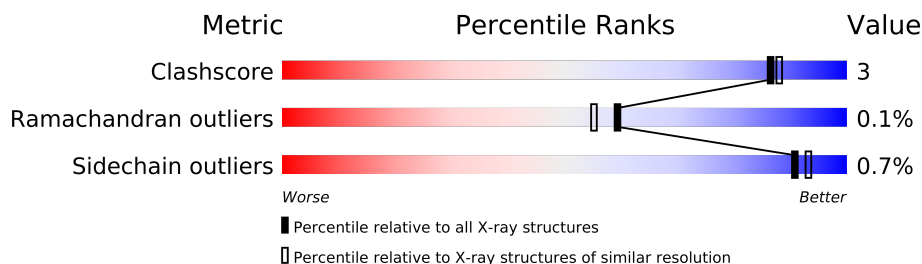
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	FAILED
Percentile statistics	:	21963
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.00 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)	Similar resolution (#Entries, resolution range(Å))
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	341	
1	B	341	
1	C	341	
1	D	341	

2 Entry composition i

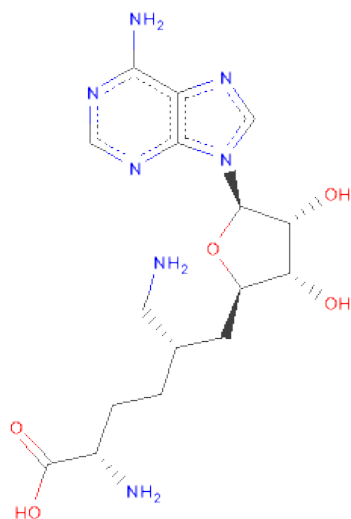
There are 5 unique types of molecules in this entry. The entry contains 11705 atoms, of which 0 are hydrogen and 0 are deuterium.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 Histone-arginine methyltransferase CARM1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	334	Total	C	N	O	S	0	12	0
			2721	1761	446	498	16			
1	B	334	Total	C	N	O	S	0	7	0
			2672	1739	431	487	15			
1	C	334	Total	C	N	O	S	0	12	0
			2708	1750	445	496	17			
1	D	334	Total	C	N	O	S	0	6	0
			2669	1732	434	488	15			

- Molecule 2 is (2S,5S)-2,6-DIAMINO-5-{[(2R,3S,4R,5R)-5-(6-AMINO-9H-PURIN-9-YL)-3,4-DIHYDROXYTETRAHYDROFURAN-2-YL]METHYL}HEXANOICACID (three-letter code: 4IK) (formula: C₁₆H₂₅N₇O₅).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	1
			33	20	8	5		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	B	1	Total	C	N	O	0	0
			28	16	7	5		
2	C	1	Total	C	N	O	0	0
			28	16	7	5		
2	D	1	Total	C	N	O	0	0
			28	16	7	5		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		
3	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is UNKNOWN ATOM OR ION (three-letter code: UNX) (formula: X).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	15	Total	X	0	0
			15	15		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	10	Total 10	X 10	0	0
4	D	6	Total 6	X 6	0	0
4	C	9	Total 9	X 9	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	184	Total 186	O 186	0	2
5	B	215	Total 217	O 217	0	2
5	C	179	Total 180	O 180	0	1
5	D	165	Total 165	O 165	0	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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). 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.

Note EDS failed to run properly.

- Molecule 1: Histone-arginine methyltransferase CARM1

Chain A: 



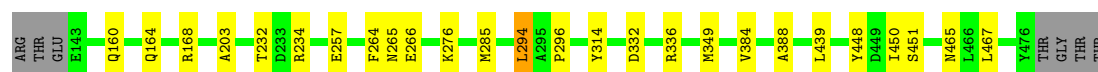
- Molecule 1: Histone-arginine methyltransferase CARM1

Chain B: 



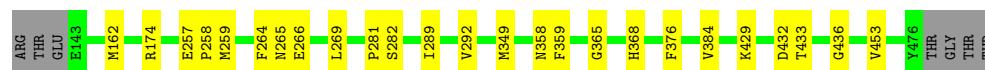
- Molecule 1: Histone-arginine methyltransferase CARM1

Chain C: 



- Molecule 1: Histone-arginine methyltransferase CARM1

Chain D: 



4 Data and refinement statistics

EDS failed to run properly - this section will therefore be incomplete.

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	75.06Å 98.84Å 206.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.11 – 2.00	Depositor
% Data completeness (in resolution range)	99.2 (48.11-2.00)	Depositor
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.72 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.203 , 0.231	Depositor
Wilson B-factor (Å ²)	31.4	Xtriage
Anisotropy	0.135	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 103958 reflections (0.001%)	Xtriage
Total number of atoms	11705	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 55.17 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 3.2849e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: UNX, GOL, 4IK

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	0.37	0/2808	0.53	0/3811
1	B	0.40	0/2757	0.56	0/3745
1	C	0.36	0/2795	0.54	0/3797
1	D	0.36	0/2745	0.53	0/3729
All	All	0.37	0/11105	0.54	0/15082

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Close contacts

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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2721	0	2642	19	0
1	B	2672	0	2584	7	0
1	C	2708	0	2613	16	0
1	D	2669	0	2575	14	0
2	A	33	0	24	0	0
2	B	28	0	24	1	0
2	C	28	0	24	0	0
2	D	28	0	24	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	6	0	8	0	0
3	C	18	0	24	0	0
3	D	6	0	8	0	0
4	A	10	0	0	0	0
4	B	15	0	0	0	0
4	C	9	0	0	0	0
4	D	6	0	0	0	0
5	A	186	0	0	2	0
5	B	217	0	0	1	0
5	C	180	0	0	2	0
5	D	165	0	0	1	0
All	All	11705	0	10550	55	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 3.

All (55) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:162:MET:SD	2:D:501:4IK:HE3	2.29	0.72
1:C:332:ASP:O	1:C:336:ARG:HG3	1.94	0.67
1:C:232[A]:THR:HG22	5:C:727:HOH:O	1.98	0.63
1:C:349[A]:MET:CE	1:C:384:VAL:HG22	2.32	0.60
1:A:217:THR:O	1:A:220[B]:GLN:HG2	2.03	0.58
1:A:440:LEU:HD23	1:A:450:ILE:HG12	1.85	0.58
1:C:265:ASN:O	1:C:266:GLU:HB2	2.04	0.57
1:A:219:ALA:O	1:A:223[B]:GLU:HG2	2.05	0.57
1:C:276:LYS:HD3	1:C:285[A]:MET:SD	2.45	0.57
1:A:327[A]:ARG:NH2	1:C:314:TYR:OH	2.39	0.56
1:A:265:ASN:O	1:A:266:GLU:HB2	2.08	0.54
1:A:349[A]:MET:HE3	1:A:384:VAL:HG22	1.90	0.54
1:B:162:MET:SD	2:B:501:4IK:HE3	2.49	0.52
1:A:349[A]:MET:CE	1:A:384:VAL:HG22	2.40	0.51
1:D:264:PHE:CE1	1:D:292:VAL:HG21	2.45	0.51
1:B:174[A]:ARG:HD3	5:B:622:HOH:O	2.12	0.49
1:C:168:ARG:NH1	1:C:257:GLU:OE1	2.45	0.49
1:D:365:GLY:HA2	1:D:368:HIS:CE1	2.48	0.49
1:D:258:PRO:HG2	1:D:269:LEU:HD23	1.94	0.48
1:A:189:LEU:HD13	1:A:247:LEU:HD21	1.95	0.48
1:C:439:LEU:HB3	1:C:451:SER:HB3	1.96	0.48
1:A:150:PHE:CE2	1:A:215:ALA:HB3	2.49	0.47
1:C:465:ASN:HB3	5:C:639:HOH:O	2.14	0.47
1:C:294:LEU:HA	1:C:388:ALA:O	2.14	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:168:ARG:NH1	1:A:257:GLU:OE1	2.48	0.46
1:D:258:PRO:HD2	1:D:259:MET:SD	2.56	0.46
1:D:376:PHE:O	1:D:433:THR:HA	2.15	0.46
1:A:450:ILE:HB	1:A:467:LEU:HB2	1.98	0.46
1:D:281[A]:PRO:O	1:D:282[A]:SER:CB	2.64	0.46
1:D:174:ARG:NH2	5:D:672:HOH:O	2.49	0.45
1:A:289:ILE:HG22	1:A:358:ASN:HA	1.98	0.45
1:C:450:ILE:HB	1:C:467:LEU:HB2	1.98	0.45
1:C:264:PHE:HB3	1:C:448:TYR:CE2	2.51	0.45
1:D:429:LYS:HE3	1:D:432:ASP:OD2	2.16	0.45
1:D:436:GLY:HA3	1:D:453:VAL:O	2.17	0.45
1:C:203:ALA:O	1:C:234:ARG:NH2	2.49	0.45
1:C:349[A]:MET:HE3	1:C:384:VAL:HG22	1.97	0.44
1:A:249:GLU:HG2	5:A:639:HOH:O	2.17	0.44
1:C:296:PRO:HB2	1:C:349[A]:MET:HE3	1.99	0.44
1:D:349[B]:MET:SD	1:D:384:VAL:HG22	2.58	0.43
1:A:324:SER:O	1:A:327[A]:ARG:HG2	2.18	0.43
1:C:160:GLN:O	1:C:164:GLN:HG3	2.18	0.43
1:B:440:LEU:HD22	1:B:450:ILE:HG12	2.01	0.43
1:B:272:TYR:OH	1:B:285[B]:MET:HB2	2.19	0.42
1:B:254:ILE:HG13	1:B:279:LEU:HD13	2.01	0.42
1:A:465:ASN:HB3	5:A:602:HOH:O	2.20	0.42
1:A:376:PHE:O	1:A:433:THR:HA	2.20	0.42
1:A:450:ILE:HD11	1:A:469:LEU:HB2	2.02	0.42
1:A:175:ALA:HA	1:A:402:VAL:HG21	2.02	0.42
1:D:265:ASN:O	1:D:266:GLU:HB2	2.19	0.42
1:B:377:HIS:ND1	1:B:433:THR:HG22	2.35	0.42
1:D:289:ILE:HG22	1:D:358:ASN:HA	2.03	0.41
1:A:294:LEU:HA	1:A:388:ALA:O	2.20	0.41
1:B:376:PHE:O	1:B:433:THR:HA	2.21	0.40
1:D:269:LEU:HD22	1:D:359:PHE:CE1	2.57	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 X-ray entries followed by that with respect to entries of similar resolution. 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	344/341 (101%)	334 (97%)	10 (3%)	0	100	100
1	B	339/341 (99%)	328 (97%)	10 (3%)	1 (0%)	50	44
1	C	344/341 (101%)	334 (97%)	10 (3%)	0	100	100
1	D	338/341 (99%)	326 (96%)	12 (4%)	0	100	100
All	All	1365/1364 (100%)	1322 (97%)	42 (3%)	1 (0%)	59	55

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	258	PRO

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 X-ray entries followed by that with respect to entries of similar resolution. 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	290/296 (98%)	287 (99%)	3 (1%)	85	88
1	B	281/296 (95%)	278 (99%)	3 (1%)	84	86
1	C	288/296 (97%)	287 (100%)	1 (0%)	96	97
1	D	281/296 (95%)	280 (100%)	1 (0%)	95	97
All	All	1140/1184 (96%)	1132 (99%)	8 (1%)	91	93

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	168	ARG
1	A	209	LYS
1	A	294	LEU
1	B	257	GLU
1	B	294	LEU
1	B	346	ARG
1	C	294	LEU
1	D	257	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

Of 50 ligands modelled in this entry, 40 are unknown - leaving 10 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	$\# Z > 2$
2	4IK	A	501[A]	-	2,4,30	0.34	0	1,4,43	0.33	0
2	4IK	A	501[B]	-	2,4,30	0.35	0	1,4,43	0.34	0
3	GOL	A	502	-	5,5,5	0.18	0	5,5,5	0.32	0
2	4IK	B	501	-	30,30,30	1.41	5 (16%)	43,43,43	2.49	11 (25%)
2	4IK	C	501	-	30,30,30	1.51	4 (13%)	43,43,43	2.50	10 (23%)
3	GOL	C	502	-	5,5,5	0.22	0	5,5,5	0.26	0
3	GOL	C	503	-	5,5,5	0.22	0	5,5,5	0.46	0
3	GOL	C	504	-	5,5,5	0.23	0	5,5,5	0.34	0
2	4IK	D	501	-	30,30,30	1.49	5 (16%)	43,43,43	2.55	9 (20%)
3	GOL	D	502	-	5,5,5	0.24	0	5,5,5	0.39	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
2	4IK	A	501[A]	-	-	0/2/2/35	0/0/0/3
2	4IK	A	501[B]	-	-	0/2/2/35	0/0/0/3
3	GOL	A	502	-	-	0/4/4/4	0/0/0/0
2	4IK	B	501	-	-	0/19/35/35	0/1/3/3
2	4IK	C	501	-	-	0/19/35/35	0/1/3/3
3	GOL	C	502	-	-	0/4/4/4	0/0/0/0
3	GOL	C	503	-	-	0/4/4/4	0/0/0/0
3	GOL	C	504	-	-	0/4/4/4	0/0/0/0
2	4IK	D	501	-	-	0/19/35/35	0/1/3/3
3	GOL	D	502	-	-	0/4/4/4	0/0/0/0

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	501	4IK	C2-N3	4.35	1.40	1.32
2	C	501	4IK	C2-N3	4.33	1.40	1.32
2	B	501	4IK	C2-N3	3.30	1.38	1.32
2	C	501	4IK	C2-N1	3.27	1.40	1.33
2	D	501	4IK	C2-N1	3.26	1.40	1.33
2	B	501	4IK	C2-N1	3.09	1.40	1.33
2	D	501	4IK	C4-N9	-3.07	1.33	1.37
2	B	501	4IK	C5-C4	-2.97	1.33	1.40
2	B	501	4IK	C4-N9	-2.82	1.33	1.37
2	C	501	4IK	C4-N9	-2.78	1.33	1.37
2	C	501	4IK	C5-C4	-2.60	1.34	1.40
2	D	501	4IK	C5-C4	-2.58	1.34	1.40
2	D	501	4IK	C2'-C1'	-2.34	1.50	1.53
2	B	501	4IK	C8-N9	-2.12	1.33	1.36

All (30) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	501	4IK	N3-C2-N1	-13.53	117.40	128.71
2	C	501	4IK	N3-C2-N1	-12.95	117.88	128.71
2	B	501	4IK	N3-C2-N1	-12.44	118.31	128.71
2	B	501	4IK	O4'-C1'-N9	-5.54	103.29	108.44
2	C	501	4IK	O4'-C1'-N9	-4.90	103.89	108.44
2	D	501	4IK	O4'-C1'-N9	-4.72	104.05	108.44
2	D	501	4IK	N3-C4-N9	3.29	131.38	125.43
2	C	501	4IK	N3-C4-N9	3.24	131.29	125.43
2	B	501	4IK	C4-C5-N7	-3.14	106.83	109.52
2	C	501	4IK	C5'-C4'-C3'	-2.86	112.20	115.51
2	B	501	4IK	O4'-C4'-C5'	-2.80	104.17	109.64

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	501	4IK	O4'-C1'-C2'	-2.80	102.48	106.77
2	B	501	4IK	N3-C4-N9	2.73	130.37	125.43
2	D	501	4IK	C5'-C4'-C3'	-2.73	112.35	115.51
2	C	501	4IK	CD-C5'-C4'	-2.66	111.65	115.29
2	D	501	4IK	CD-C5'-C4'	-2.58	111.75	115.29
2	B	501	4IK	O4'-C1'-C2'	-2.56	102.84	106.77
2	C	501	4IK	C4-C5-N7	-2.42	107.45	109.52
2	C	501	4IK	C5-C4-N3	-2.40	120.48	125.70
2	B	501	4IK	N7-C8-N9	-2.37	107.64	114.36
2	B	501	4IK	C5-C4-N3	-2.33	120.62	125.70
2	D	501	4IK	N7-C8-N9	-2.30	107.87	114.36
2	C	501	4IK	N7-C8-N9	-2.26	107.95	114.36
2	B	501	4IK	C5'-C4'-C3'	-2.17	113.00	115.51
2	D	501	4IK	C5-C4-N3	-2.14	121.04	125.70
2	D	501	4IK	C8-N9-C4	2.11	108.51	106.90
2	B	501	4IK	CB-CG-CD	-2.11	109.95	115.91
2	C	501	4IK	C2-N3-C4	2.09	119.96	114.01
2	D	501	4IK	C2-N3-C4	2.08	119.94	114.01
2	B	501	4IK	C2-N3-C4	2.01	119.73	114.01

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

EDS failed to run properly - this section will therefore be empty.

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

EDS failed to run properly - this section will therefore be empty.

6.3 Carbohydrates ⓘ

EDS failed to run properly - this section will therefore be empty.

6.4 Ligands ⓘ

EDS failed to run properly - this section will therefore be empty.

6.5 Other polymers ⓘ

EDS failed to run properly - this section will therefore be empty.