wwPDB X-ray Structure Validation Summary Report (i)

Aug 3, 2023 – 03:27 AM EDT

PDB ID	:	1EQZ
Title	:	X-RAY STRUCTURE OF THE NUCLEOSOME CORE PARTICLE AT 2.5 A RESOLUTION
Authors	:	Hanson, B.L.; Harp, J.M.; Timm, D.E.; Bunick, G.J.
Deposited on	:	2000-04-06
Resolution	:	2.50 Å(reported)

This is a wwPDB X-ray Structure 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/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

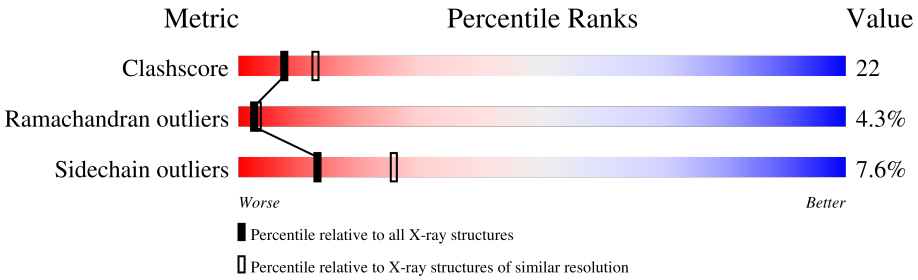
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.34

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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 of the lower 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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	I	146	40% 58% ·
1	J	146	34% 61% 5%
2	A	129	65% 23% 8% · ·
2	E	129	60% 32% 7% ·
3	B	126	58% 19% 9% 14%
3	F	126	57% 24% · 15%
4	C	136	54% 21% · · 21%
4	G	136	57% 23% 8% 12%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	D	103	
5	H	103	

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 13325 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 DNA chain called 146 NUCLEOTIDES LONG DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	I	146	Total	C	N	O	P	0	0	0
			2994	1430	541	877	146			
1	J	146	Total	C	N	O	P	0	0	0
			2994	1430	541	877	146			

- Molecule 2 is a protein called PROTEIN (HISTONE H2A).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	A	125	Total	C	N	O	0	0	0
			950	594	190	166			
2	E	127	Total	C	N	O	0	0	0
			968	606	194	168			

- Molecule 3 is a protein called PROTEIN (HISTONE H2B).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	108	Total	C	N	O	S	0	0	0
			852	534	157	159	2			
3	F	107	Total	C	N	O	S	0	0	0
			845	529	156	158	2			

- Molecule 4 is a protein called PROTEIN (HISTONE H3).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	C	107	Total	C	N	O	S	0	1	0
			875	550	171	151	3			
4	G	120	Total	C	N	O	S	0	3	0
			996	627	197	169	3			

- Molecule 5 is a protein called PROTEIN (HISTONE H4).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	D	89	Total	C	N	O	S	0	1	0
			723	453	148	121	1			
5	H	95	Total	C	N	O	S	0	0	0
			750	471	152	126	1			

- Molecule 6 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	I	1	Total	K	0	0
			1	1		
6	J	6	Total	K	0	0
			6	6		

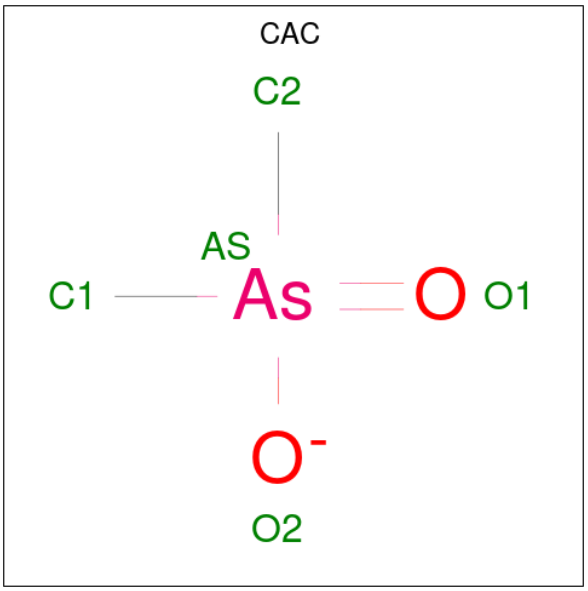
- Molecule 7 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	I	9	Total	Mn	0	0
			9	9		
7	J	6	Total	Mn	0	0
			6	6		

- Molecule 8 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	1	Total	Cl	0	0
			1	1		
8	E	1	Total	Cl	0	0
			1	1		

- Molecule 9 is CACODYLATE ION (three-letter code: CAC) (formula: $C_2H_6AsO_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	G	1	Total	As	C	O	0	0
			5	1	2	2		

- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	I	45	Total	O	0	0
			45	45		
10	J	45	Total	O	0	0
			45	45		
10	A	43	Total	O	0	0
			43	43		
10	B	15	Total	O	0	0
			15	15		
10	C	39	Total	O	0	0
			39	39		
10	D	26	Total	O	0	0
			26	26		
10	E	39	Total	O	0	0
			39	39		
10	F	16	Total	O	0	0
			16	16		
10	G	48	Total	O	0	0
			48	48		
10	H	33	Total	O	0	0
			33	33		

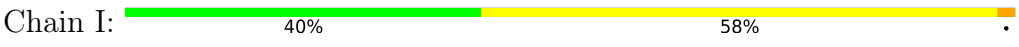
3 Residue-property plots (i)

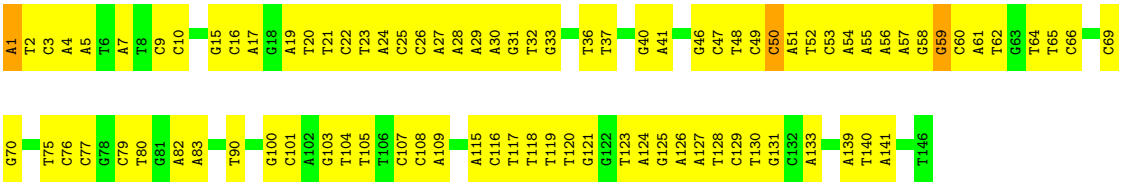
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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.

Note EDS was not executed.

• Molecule 1: 146 NUCLEOTIDES LONG DNA

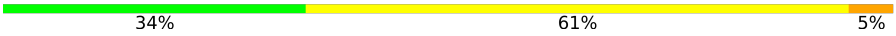
Chain I:

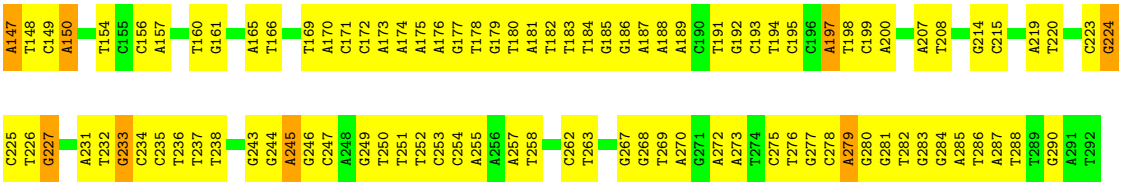




• Molecule 1: 146 NUCLEOTIDES LONG DNA

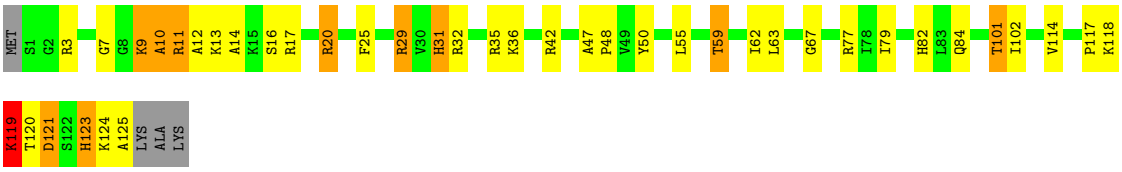
Chain J:





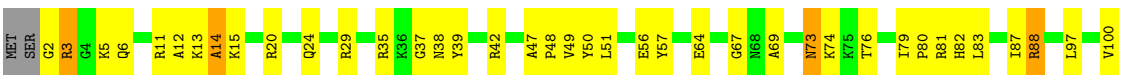
• Molecule 2: PROTEIN (HISTONE H2A)

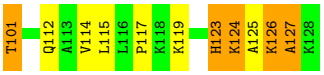
Chain A:



• Molecule 2: PROTEIN (HISTONE H2A)

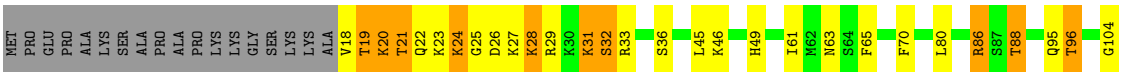
Chain E:

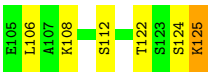




• Molecule 3: PROTEIN (HISTONE H2B)

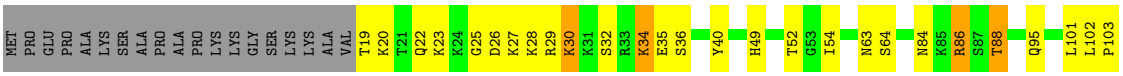
Chain B:

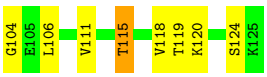




• Molecule 3: PROTEIN (HISTONE H2B)

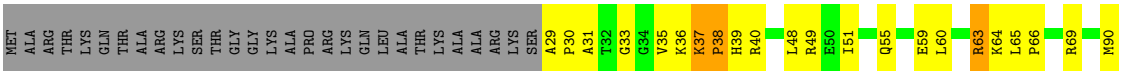
Chain F:

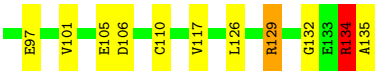




• Molecule 4: PROTEIN (HISTONE H3)

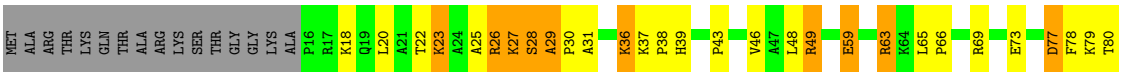
Chain C:

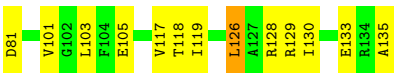




• Molecule 4: PROTEIN (HISTONE H3)

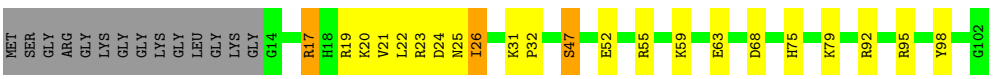
Chain G:





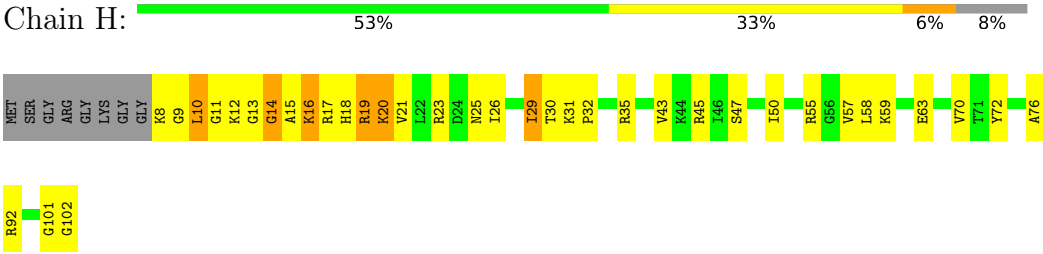
• Molecule 5: PROTEIN (HISTONE H4)

Chain D:



• Molecule 5: PROTEIN (HISTONE H4)

Chain H:



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	105.28Å 109.71Å 181.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.34 – 2.50	Depositor
% Data completeness (in resolution range)	98.6 (25.34-2.50)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.219 , 0.282	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	13325	wwPDB-VP
Average B, all atoms (Å^2)	53.0	wwPDB-VP

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: K, CL, MN, CAC

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	I	0.66	1/3358 (0.0%)	0.84	3/5179 (0.1%)
1	J	0.70	1/3358 (0.0%)	0.87	2/5179 (0.0%)
2	A	0.79	0/961	0.89	0/1289
2	E	0.70	0/979	0.87	2/1310 (0.2%)
3	B	0.84	0/863	0.86	0/1151
3	F	0.76	0/856	0.83	0/1141
4	C	0.79	0/865	0.94	1/1159 (0.1%)
4	G	0.83	0/944	0.94	1/1260 (0.1%)
5	D	0.77	0/708	0.89	0/943
5	H	0.91	0/758	0.98	3/1008 (0.3%)
All	All	0.74	2/13650 (0.0%)	0.88	12/19619 (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
1	I	0	3
1	J	0	8
2	E	0	1
3	F	0	1
All	All	0	13

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	I	1	DA	OP3-P	-6.92	1.52	1.61
1	J	147	DA	OP3-P	-6.91	1.52	1.61

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	88	ARG	NE-CZ-NH2	-8.22	116.19	120.30
2	E	88	ARG	NE-CZ-NH1	7.35	123.98	120.30
4	C	134	ARG	N-CA-C	6.83	129.45	111.00
5	H	23	ARG	NE-CZ-NH2	-6.46	117.07	120.30
1	I	50	DC	C5'-C4'-C3'	-5.71	103.82	114.10

There are no chirality outliers.

5 of 13 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	I	41	DA	Sidechain
1	I	59	DG	Sidechain
1	I	79	DC	Sidechain
1	J	150	DA	Sidechain
1	J	197	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	I	2994	0	1650	124	0
1	J	2994	0	1650	120	0
2	A	950	0	1021	51	0
2	E	968	0	1044	48	0
3	B	852	0	900	52	0
3	F	845	0	891	38	0
4	C	875	0	920	40	0
4	G	996	0	1059	63	0
5	D	723	0	775	22	0
5	H	750	0	809	57	0
6	I	1	0	0	0	0
6	J	6	0	0	0	0
7	I	9	0	0	0	0
7	J	6	0	0	0	0
8	A	1	0	0	0	0
8	E	1	0	0	0	0
9	G	5	0	0	1	0
10	A	43	0	0	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	B	15	0	0	0	0
10	C	39	0	0	1	0
10	D	26	0	0	0	0
10	E	39	0	0	2	0
10	F	16	0	0	0	0
10	G	48	0	0	1	0
10	H	33	0	0	5	0
10	I	45	0	0	0	0
10	J	45	0	0	0	0
All	All	13325	0	10719	523	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 22.

The worst 5 of 523 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:J:188:DA:H2"	1:J:189:DA:H5'	1.26	1.17
2:A:17:ARG:HH22	2:A:31:HIS:CD2	1.65	1.14
1:I:124:DA:H2"	1:I:125:DG:H5"	1.33	1.09
1:J:187:DA:H2"	1:J:188:DA:H5"	1.35	1.07
1:I:103:DG:H2"	1:I:104:DT:H5'	1.34	1.06

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 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
2	A	123/129 (95%)	108 (88%)	9 (7%)	6 (5%)	2 2
2	E	125/129 (97%)	102 (82%)	16 (13%)	7 (6%)	2 1

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
3	B	106/126 (84%)	94 (89%)	6 (6%)	6 (6%)	1 1
3	F	105/126 (83%)	97 (92%)	6 (6%)	2 (2%)	8 13
4	C	106/136 (78%)	93 (88%)	7 (7%)	6 (6%)	1 1
4	G	121/136 (89%)	103 (85%)	14 (12%)	4 (3%)	4 5
5	D	88/103 (85%)	81 (92%)	5 (6%)	2 (2%)	6 10
5	H	93/103 (90%)	82 (88%)	7 (8%)	4 (4%)	2 3
All	All	867/988 (88%)	760 (88%)	70 (8%)	37 (4%)	2 3

5 of 37 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	9	LYS
2	A	10	ALA
2	A	119	LYS
2	A	121	ASP
3	B	21	THR

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 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
2	A	95/98 (97%)	85 (90%)	10 (10%)	7 13
2	E	96/98 (98%)	89 (93%)	7 (7%)	14 27
3	B	94/107 (88%)	84 (89%)	10 (11%)	6 13
3	F	93/107 (87%)	85 (91%)	8 (9%)	10 20
4	C	89/111 (80%)	86 (97%)	3 (3%)	37 63
4	G	97/111 (87%)	89 (92%)	8 (8%)	11 22
5	D	71/79 (90%)	67 (94%)	4 (6%)	21 40
5	H	75/79 (95%)	71 (95%)	4 (5%)	22 43
All	All	710/790 (90%)	656 (92%)	54 (8%)	13 25

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	64	GLU
3	F	34	LYS
4	G	126	LEU
2	E	73	ASN
2	E	101	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
4	C	39	HIS
2	E	73	ASN
4	G	39	HIS
3	F	95	GLN
3	B	95	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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)

Of 25 ligands modelled in this entry, 24 are monoatomic - leaving 1 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	$\#\|Z\| > 2$
9	CAC	G	451	4,3	0,4,4	-	-	0,6,6	-	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	G	451	CAC	1	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.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.