W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

Full wwPDB EM Validation Report (i)

Aug 7, 2023 – 11:08 AM EDT

PDB ID	:	8SKW
Title	:	MicroED structure of d(CGCGCG)2 Z-DNA
Authors	:	Haymaker, A.; Bardin, A.A.; Martynowycz, M.W.; Nannenga, B.L.
Deposited on	:	2023-04-20
Resolution	:	1.10 Å(reported)

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

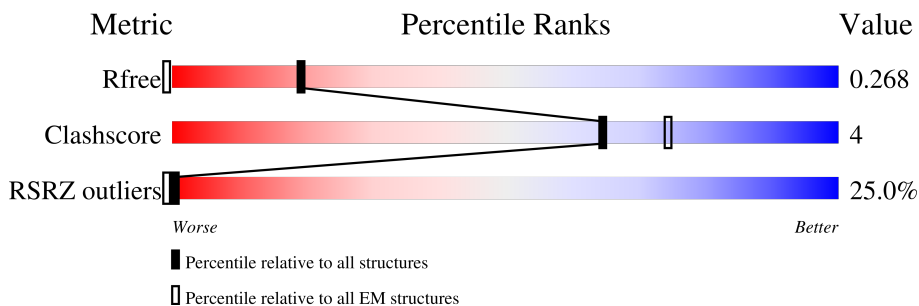|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
ELECTRON CRYSTALLOGRAPHY

The reported resolution of this entry is 1.10 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	Percentile Ranks	Value
Rfree		0.268
Clashscore		4
RSRZ outliers		25.0%

Worse   *Better*

▌ Percentile relative to all structures

▯ Percentile relative to all EM structures

Metric	Whole archive (#Entries)	EM structures (#Entries)
R_{free}	130704	0
Clashscore	158937	4297
RSRZ outliers	127900	0

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 <=5%

| Mol	Chain	Length	Quality of chain
1	A	6	33% / 50% / 17% / 33%
1	B	6	17% / 50% / 50%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 505 atoms, of which 183 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 DNA chain called DNA (5'-D(*CP*GP*CP*GP*CP*G)-3').

Mol	Chain	Residues	Atoms						AltConf	Trace
1	B	6	Total 188	C 57	H 68	N 24	O 34	P 5	0	0
1	A	6	Total 246	C 75	H 91	N 30	O 44	P 6	2	0

- Molecule 2 is SPERMINE (three-letter code: SPM) (formula: $C_{10}H_{26}N_4$).
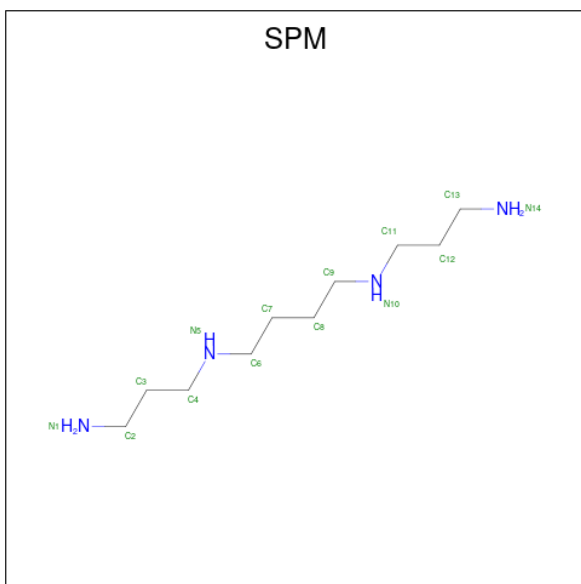
Mol	Chain	Residues	Atoms				AltConf
2	A	1	Total 38	C 10	H 24	N 4	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		AltConf
3	B	16	Total 16	O 16	0
3	A	17	Total 17	O 17	0

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.

• Molecule 1: DNA (5'-D(*CP*GP*CP*GP*CP*G)-3')
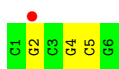
• Molecule 1: DNA (5'-D(*CP*GP*CP*GP*CP*G)-3')

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	18.28Å 32.00Å 41.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.38 – 1.10 25.38 – 1.10	Depositor EDS
% Data completeness (in resolution range)	98.7 (25.38-1.10) 96.0 (25.38-1.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.07 (at 1.10Å)	Xtriage
Refinement program	phenix.refine 1.20.1_4487, PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.210 , 0.226 0.246 , 0.268	Depositor DCC
R_{free} test set	514 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	9.3	Xtriage
Anisotropy	0.328	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 33.8	EDS
L-test for twinning[2]	$< |L| > = 0.43, < L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	505	wwPDB-VP
Average B, all atoms (Å^2)	10.0	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 23.29 % 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 4.8043e-03.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5   Model quality ⓘ

5.1   Standard geometry ⓘ

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

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	1.79	3/173 (1.7%)	1.43	2/265 (0.8%)
1	B	1.90	2/134 (1.5%)	1.53	2/205 (1.0%)
All	All	1.84	5/307 (1.6%)	1.47	4/470 (0.9%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	2	DG	C3'-O3'	6.72	1.52	1.44
1	A	1[A]	DC	C1'-N1	5.94	1.56	1.49
1	A	1[B]	DC	C1'-N1	5.94	1.56	1.49
1	B	4	DG	C6-N1	-5.88	1.35	1.39
1	A	2	DG	N9-C4	5.14	1.42	1.38

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	5	DC	O5'-P-OP1	6.46	118.45	110.70
1	B	4	DG	C8-N9-C4	-6.45	103.82	106.40
1	A	1[A]	DC	O4'-C1'-N1	6.12	112.29	108.00
1	A	1[B]	DC	O4'-C1'-N1	6.12	112.29	108.00

There are no chirality outliers.

There are no planarity outliers.

5.2   Too-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 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	155	91	91	2	0
1	B	120	68	68	0	0
2	A	14	24	26	0	0
3	A	17	0	0	0	0
3	B	16	0	0	0	0
All	All	322	183	185	2	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 (2) 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:1[B]:DC:O4'	1:A:2:DG:C2	2.73	0.42
1:A:2:DG:H5'	1:A:3[B]:DC:OP2	2.19	0.42

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

There are no protein molecules in this entry.

5.3.2 Protein sidechains ⓘ

There are no protein molecules in this entry.

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

1 ligand is 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	#$|Z| > 2$
2	SPM	A	101	-	13,13,13	0.44	0	12,12,12	0.65	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	SPM	A	101	-	-	3/11/11/11	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	101	SPM	N10-C11-C12-C13
2	A	101	SPM	N5-C6-C7-C8
2	A	101	SPM	C8-C9-N10-C11

There are no ring outliers.

No monomer is involved in short contacts.

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.