wwPDB NMR Structure Validation Summary Report (i)

Aug 20, 2022 – 08:27 AM EDT

PDB ID	:	1EEK
Title	:	SOLUTION STRUCTURE OF A NONPOLAR, NON HYDROGEN BONDED BASE PAIR SURROGATE IN DNA.
Authors	:	Kool, E.T.; Krugh, T.R.; Guckian, K.M.
Deposited on	:	2000-02-01

This is a wwPDB NMR 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/NMRValidationReportHelp
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 (1)) 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)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.29
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.29

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment was not calculated.

There are no overall percentile quality scores available for this entry.

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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	1	12	100%
2	2	12	100%

2 Ensemble composition and analysis (i)

This entry contains 1 models. Identification of well-defined residues and clustering analysis are not possible.

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 763 atoms, of which 277 are hydrogens and 0 are deuteriums.

- Molecule 1 is a DNA chain called 5'-D(*CP*GP*CP*AP*TP*(DFT)P*GP*TP*TP*AP*CP*C)-3'.

Mol	Chain	Residues	Atoms							Trace
			Total	C	F	H	N	O	P	
1	1	12	377	118	2	138	38	70	11	0

- Molecule 2 is a DNA chain called 5'-D(*GP*GP*TP*AP*AP*CP*(MBZ)P*AP*TP*GP*CP*G)-3'.

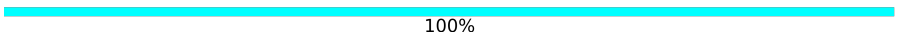
Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
2	2	12	386	121	139	47	68	11	0

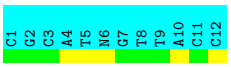
4 Residue-property plots (i)

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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 outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

- Molecule 1: 5'-D(*CP*GP*CP*AP*TP*(DFT)P*GP*TP*TP*AP*CP*C)-3'

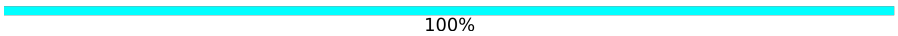
Chain 1:

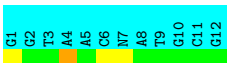




- Molecule 2: 5'-D(*GP*GP*TP*AP*AP*CP*(MBZ)P*AP*TP*GP*CP*G)-3'

Chain 2:





5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *restrained molecular dynamics*.

Of the ? calculated structures, 1 were deposited, based on the following criterion: *?*.

The authors did not provide any information on software used for structure solution, optimization or refinement.

No chemical shift data was provided.

6 Model quality (i)

6.1 Standard geometry (i)

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

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

6.2 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 each 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 averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	1	0	0	0	0
2	2	0	0	0	0
All	All	0	0	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 -.

There are no clashes.

6.3 Torsion angles (i)

6.3.1 Protein backbone (i)

There are no protein molecules in this entry.

6.3.2 Protein sidechains (i)

There are no protein molecules in this entry.

6.3.3 RNA (i)

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains (i)

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds 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 is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
1	DFT	1	6	1	18,21,22	1.35	2 (11%)
2	MBZ	2	7	2	18,23,24	1.41	3 (16%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of 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 angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
1	DFT	1	6	1	26,30,33	1.86	4 (15%)
2	MBZ	2	7	2	21,33,36	0.96	2 (9%)

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
1	DFT	1	6	1	-	0,7,21,22	0,2,2,2
2	MBZ	2	7	2	-	0,3,21,22	0,3,3,3

All bond outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	2	7	MBZ	C5M-C6	3.27	1.57	1.51
1	1	6	DFT	C5-C4	3.10	1.41	1.37
2	2	7	MBZ	C12-C6	2.83	1.43	1.37
2	2	7	MBZ	C8-N7	2.49	1.30	1.34
1	1	6	DFT	C3-C4	2.24	1.41	1.37

5 of 6 angle outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	1	6	DFT	C3-C4-C5	6.09	120.49	124.39
1	1	6	DFT	C6-C5-C4	4.43	118.88	116.06
2	2	7	MBZ	C10-C9-N1	3.05	129.51	132.11
1	1	6	DFT	C6-C1-C2	3.02	118.99	116.48
1	1	6	DFT	C3-C2-C1	2.82	120.46	123.83

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6.5 Carbohydrates (i)

There are no monosaccharides in this entry.

6.6 Ligand geometry (i)

There are no ligands in this entry.

6.7 Other polymers (i)

There are no such molecules in this entry.

6.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

7 Chemical shift validation (i)

No chemical shift data were provided