

Full wwPDB NMR Structure Validation Report (i)

Feb 12, 2022 - 08:04 PM EST

PDB ID	:	1GJ1
Title	:	NMR structure of d(CCAAAGXACTGGG), X is a 3'phosphoglycolate, 5'phos-
		phate gapped lesion
Authors	:	Junker, HD.; Hoehn, S.T.; Bunt, R.C.; Marathius, V.; Chen, J.; Turner, C.J.;
		Stubbe, J.
Deposited on	:	2002-10-14

This is a Full wwPDB NMR Structure Validation 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 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	:	2.26
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.26

Clashscore

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.

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.

Metrie	С	Percent	ile Ranks	Value
Clashscore				0
	Worse			Better
	Perc	entile relative to all structures		
	Perc	entile relative to all NMR structures		
				1
Metri	c	Whole archive	NMR archive	

(# Entries)

158937

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%

(#Entries)

12864

Mol	Chain	Length	Quality of chain
1	А	6	100%
2	В	6	100%
3	С	13	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 4 unique types of molecules in this entry. The entry contains 802 atoms, of which 286 are hydrogens and 0 are deuteriums.

• Molecule 1 is a DNA chain called 5'-D(*CP*CP*AP*AP*AP*G)-3'.

Mol	Chain	Residues	Atoms						Trace
1	٨	6	Total	С	Η	Ν	0	Р	0
1	А	0	187	58	67	26	31	5	0

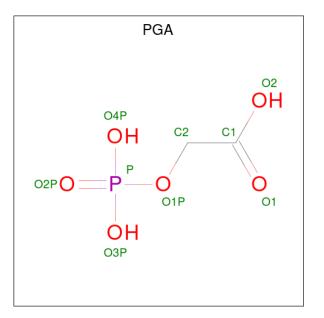
• Molecule 2 is a DNA chain called 5'-D(P*AP*CP*TP*GP*GP*G)-3'.

	Mol	Chain	Residues	Atoms					Trace	
ſ	9	D	6	Total	С	Η	Ν	0	Р	0
	Z	D	0	195	59	68	25	37	6	0

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

Mol	Chain	Residues		Atoms					Trace
3	С	12	Total	С	Н	Ν	Ο	Р	0
0	U	15	410	126	149	45	78	12	0

• Molecule 4 is 2-PHOSPHOGLYCOLIC ACID (three-letter code: PGA) (formula: $C_2H_5O_6P$).



Mol	Chain	Residues		Atoms			
4	٨	1	Total	С	Η	Ο	Р
4	A	1	10	2	2	5	1



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*CP*AP*AP*AP*G)-3'

Chain A:	100%
C1 A3 A5 C2 A5 C2 C2 A5 C2 C2 A5 C2 C2 A5 C2 C2 C2 C2 A5 C2 C2 C2 C2 C2 C2 C2 C2 C2 C2 C2 C2 C2	
• Molecule 2: 5'-D(P*AP*CP*TP	*GP*GP*G)-3'
Chain B:	100%
A8 710 611 612 612 613	
• Molecule 3: 5'-D(*CP*CP*CP*A	AP*GP*TP*AP*CP*TP*TP*TP*GP*G)-3'
Chain C:	100%
28 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: CONJUGATE GRADIENT MINIMIZA-TION.

Of the 10 calculated structures, 1 were deposited, based on the following criterion: *MINIMUM RMSD AND NMR CONSTRAINTS VIOLATION*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
X-PLOR	refinement	3.851
X-PLOR	structure solution	3.851

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: PGA

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
All	All	8	2	0	-

The all-atom clash score is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clash score 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)

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

6.5 Carbohydrates (i)

There are no monosaccharides in this entry.

6.6 Ligand geometry (i)

1 ligand is 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	Tuno	Chain	Dog	Link		nd lengt	
WIOI	Type	Ullaili	nes	LINK	Counts	RMSZ	#Z>2
4	PGA	А	7	1	0,7,8	0.00	-

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.

Mo	Type	Chain	Bos	Link	$\begin{array}{ c c c } \hline \textbf{Bond angles} \\ \hline \text{Counts} & \text{RMSZ} & \#\text{Z}{>}2 \end{array}$		
IVIO	Type	Ullalli	nes	LIIIK	Counts	RMSZ	#Z>2
4	PGA	A	7	1	0,8,11	0.00	-

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
4	PGA	А	7	1	-	$0,\!1,\!5,\!6$	-



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.

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

