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

wwPDB NMR Structure Validation Summary Report (i)

May 14, 2024 – 12:20 am BST

PDB ID	:	7B71
BMRB ID	:	34580
Title	:	Single modified phosphoryl guanidine DNA duplex, Sp diastereomer
Authors	:	Lomzov, A.A.; Shernuykov, A.V.; Apukhtina, V.S.; Pyshnyi, D.V.
Deposited on	:	2020-12-09

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 (i)) were used in the production of this report:

MolProbity	:	FAILED
Mogul	:	1.8.4, CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance ⓘ

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

The overall completeness of chemical shifts assignment is 44%.

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

The sequence quality summary graphics cannot be shown.

2 Ensemble composition and analysis ⓘ

This entry contains 10 models. This entry does not contain polypeptide chains, therefore identification of well-defined residues and clustering analysis are not possible. All residues are included in the validation scores.

3 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 647 atoms, of which 236 are hydrogens and 0 are deuteriums.

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

Mol	Chain	Residues	Atoms						Trace
1	A	10	Total	C	H	N	O	P	0
			312	95	113	37	58	9	

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

Mol	Chain	Residues	Atoms						Trace
2	B	10	Total	C	H	N	O	P	0
			335	102	123	44	57	9	

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

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: DNA (5'-D(*CP*AP*CP*GP*CP*CP*GP*CP*TP*G)-3')

Chain A: 100%

There are no outlier residues in this chain.

• Molecule 2: DNA (5'-D(*CP*AP*GP*CP*GP*GP*CP*GP*(SGT)P*G)-3')

Chain B: 90%   10%

C11 T19 G20

4.2 Residue scores for the representative (author defined) model from the NMR ensemble

The representative model is number 1. Colouring as in section 4.1 above.

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

Chain A: 100%

There are no outlier residues in this chain.

• Molecule 2: DNA (5'-D(*CP*AP*GP*CP*GP*GP*CP*GP*(SGT)P*G)-3')

Chain B: 90%   10%

C11 T19 G20

5 Refinement protocol and experimental data overview ⓘ

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

Of the 10 calculated structures, 10 were deposited, based on the following criterion: *all calculated structures submitted.*

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

Software name	Classification	Version
Amber	refinement	16
Amber	structure calculation	16

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	2
Total number of shifts	285
Number of shifts mapped to atoms	251
Number of unparsed shifts	0
Number of shifts with mapping errors	34
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	44%

6 Model quality ⓘ

6.1 Standard geometry ⓘ

MolProbity failed to run properly - this section will have to be empty.

6.2 Too-close contacts ⓘ

MolProbity failed to run properly - this section will have to be empty.

6.3 Torsion angles ⓘ

6.3.1 Protein backbone ⓘ

MolProbity failed to run properly - this section will have to be empty.

6.3.2 Protein sidechains ⓘ

MolProbity failed to run properly - this section will have to be empty.

6.3.3 RNA ⓘ

MolProbity failed to run properly - this section will have to be empty.

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

MolProbity failed to run properly - this section will have to be empty.

6.5 Carbohydrates ⓘ

MolProbity failed to run properly - this section will have to be empty.

6.6 Ligand geometry ⓘ

MolProbity failed to run properly - this section will have to be empty.

6.7 Other polymers ⓘ

MolProbity failed to run properly - this section will have to be empty.

6.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

7 Chemical shift validation ⓘ

The completeness of assignment taking into account all chemical shift lists is 44% for the well-defined parts and 44% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *LD2_ID_SGT_D2O_30.5C_AB.csdep*

7.1.1 Bookkeeping ⓘ

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	195
Number of shifts mapped to atoms	165
Number of unparsed shifts	0
Number of shifts with mapping errors	30
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. First 5 (of 30) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	19	T0Q	C31	49.558	0.000	1
1	B	19	T0Q	C32	35.638	0.000	1
1	B	19	T0Q	C33	35.638	0.000	1
1	B	19	T0Q	C34	49.558	0.000	1
1	B	19	T0Q	C4	168.871	0.000	1
1	B	19	T0Q	C5	113.64	0.000	1
1	B	19	T0Q	C6	138.431	0.021	1
1	B	19	T0Q	CN	162.527	0.000	1
1	B	19	T0Q	H1'	5.924	0.003	1
1	B	19	T0Q	H11	3.556	0.000	1
1	B	19	T0Q	H12	3.556	0.000	1
1	B	19	T0Q	H2'	1.879	0.001	1
1	B	19	T0Q	H2"	2.35	0.001	1
1	B	19	T0Q	H21	2.945	0.001	1

Continued on next page...

Continued from previous page...

| List ID | Chain | Res | Type | Atom | Shift Data | | |
					Value	Uncertainty	Ambiguity
1	B	19	T0Q	H22	2.945	0.001	1
1	B	19	T0Q	H23	2.945	0.001	1
1	B	19	T0Q	H3'	4.836	0.001	1
1	B	19	T0Q	H31	2.945	0.001	1
1	B	19	T0Q	H32	2.945	0.001	1
1	B	19	T0Q	H33	2.945	0.001	1
1	B	19	T0Q	H4'	4.173	0.001	1
1	B	19	T0Q	H41	3.556	0.000	1
1	B	19	T0Q	H42	3.556	0.000	1
1	B	19	T0Q	H5'	4.235	0.001	2
1	B	19	T0Q	H5"	4.354	0.001	2
1	B	19	T0Q	H6	7.17	0.001	1
1	B	19	T0Q	H71	1.581	0.001	2
1	B	19	T0Q	H72	1.581	0.001	2
1	B	19	T0Q	H73	1.581	0.001	2
1	B	19	T0Q	P	-1.217	0.001	1

7.1.2   Chemical shift referencing ⓘ

No chemical shift referencing corrections were calculated (not enough data).

7.1.3   Completeness of resonance assignments ⓘ

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 38%, i.e. 147 atoms were assigned a chemical shift out of a possible 386. 0 out of 0 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Sugar	118/228 (52%)	115/133 (86%)	3/95 (3%)	0/0 (—%)
Base	29/158 (18%)	29/101 (29%)	0/30 (0%)	0/27 (0%)
Overall	147/386 (38%)	144/234 (62%)	3/125 (2%)	0/27 (0%)

7.1.4   Statistically unusual chemical shifts ⓘ

There are no statistically unusual chemical shifts.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

7.1.5 Random Coil Index (RCI) plots ⓘ

No *random coil index*(RCI) plot could be generated from the current chemical shift list. RCI is only applicable to proteins

7.2 Chemical shift list 2

File name: working_cs.cif

Chemical shift list name: *D_1292111843_cs_P1.str.V1*

7.2.1 Bookkeeping ⓘ

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	90
Number of shifts mapped to atoms	86
Number of unparsed shifts	0
Number of shifts with mapping errors	4
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. All 4 occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
2	B	19	T0Q	H3	14.01	0.002	1
2	B	19	T0Q	H71	1.588	0.000	2
2	B	19	T0Q	H72	1.588	0.000	2
2	B	19	T0Q	H73	1.588	0.000	2

7.2.2 Chemical shift referencing ⓘ

No chemical shift referencing corrections were calculated (not enough data).

7.2.3 Completeness of resonance assignments ⓘ

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 22%, i.e. 86 atoms were assigned a chemical

WORLDWIDE
PDB
PROTEIN DATA BANK

shift out of a possible 386. 0 out of 0 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Sugar	37/228 (16%)	37/133 (28%)	0/95 (0%)	0/0 (—%)
Base	49/158 (31%)	49/101 (49%)	0/30 (0%)	0/27 (0%)
Overall	86/386 (22%)	86/234 (37%)	0/125 (0%)	0/27 (0%)

7.2.4   Statistically unusual chemical shifts ⓘ

There are no statistically unusual chemical shifts.

7.2.5   Random Coil Index (RCI) plots ⓘ

No *random coil index*(RCI) plot could be generated from the current chemical shift list. RCI is only applicable to proteins

8 NMR restraints analysis ⓘ

No restraints data found

9 Distance violation analysis ⓘ

No distance restraints data found

10 Dihedral-angle violation analysis ⓘ

No dihedral-angle restraints found