wwPDB NMR Structure Validation Summary Report (i)

May 28, 2020 – 11:36 pm BST

PDB ID	:	2LWH
Title	:	NMR Structure of the Self-Complementary 10 mer DNA Duplex 5'-GGATATATCC-3' in Complex with Netropsin
Authors	:	Rettig, M.; Germann, M.W.; Wilson, W.; Wang, S.
Deposited on	:	2012-07-31

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 following versions of software and data (see references (1)) were used in the production of this report:

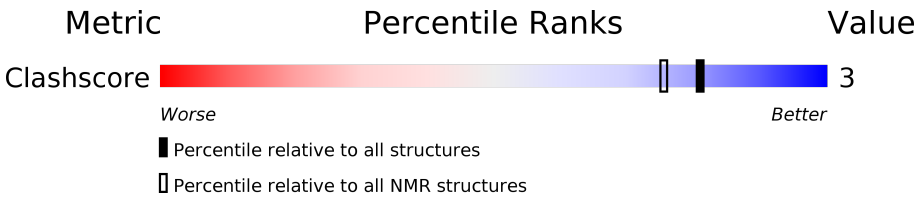
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 is 40%.

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)	NMR archive (#Entries)
Clashscore	158937	12864

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	10	60% 40%
1	2	10	70% 20% 10%

2 Ensemble composition and analysis (i)

This entry contains 11 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.

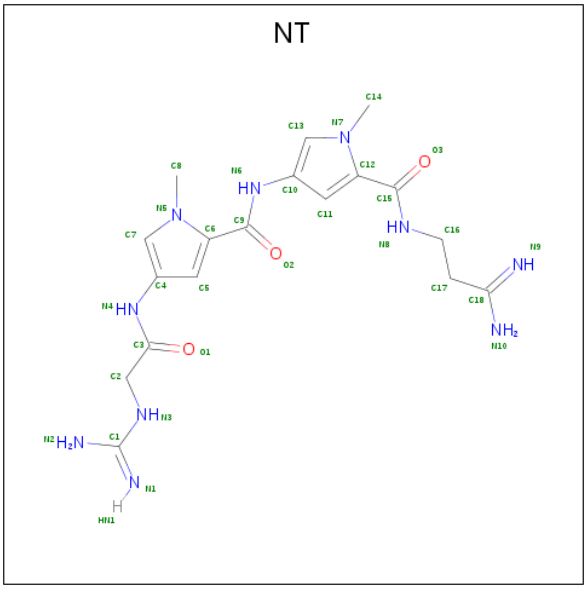
3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 689 atoms, of which 254 are hydrogens and 0 are deuteriums.

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

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
1	1	10	315	98	113	37	58	9	0
1	2	10	315	98	113	37	58	9	0

- Molecule 2 is NETROPSIN (three-letter code: NT) (formula: $C_{18}H_{26}N_{10}O_3$).



Mol	Chain	Residues	Atoms				
			Total	C	H	N	O
2	1	1	59	18	28	10	3

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA 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(*GP*GP*AP*TP*AP*TP*AP*TP*CP*C)-3')

Chain 1:

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

Chain 2:

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(*GP*GP*AP*TP*AP*TP*AP*TP*CP*C)-3')

Chain 1:

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

Chain 2:

5 Refinement protocol and experimental data overview (i)

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

Of the 9500 calculated structures, 11 were deposited, based on the following criterion: *target function.*

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

Software name	Classification	Version
AMBER	structure solution	9
AMBER	refinement	9

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 6 of this report.

Chemical shift file(s)	input_cs.cif
Number of chemical shift lists	1
Total number of shifts	198
Number of shifts mapped to atoms	198
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	40%

No validations of the models with respect to experimental NMR restraints is performed at this time.

COVALENT-GEOMETRY INFOmissingINFO

5.1 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	202	113	115	2±1
1	2	202	113	115	2±1
All	All	4785	2794	2794	20

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 3.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:1:5:DA:C2	1:2:17:DA:C2	0.51	2.99	6	9
1:1:7:DA:C2	1:2:15:DA:C2	0.49	3.00	3	9
1:2:15:DA:C2	1:2:16:DT:C2	0.40	3.10	6	1
1:1:5:DA:C2	1:1:6:DT:C2	0.40	3.09	9	1

5.2 Torsion angles (i)

5.2.1 Protein backbone (i)

There are no protein molecules in this entry.

5.2.2 Protein sidechains (i)

There are no protein molecules in this entry.

5.2.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.4 Carbohydrates (i)

There are no carbohydrates in this entry.

LIGAND-GEOMETRY INFOmissingINFO

5.5 Other polymers (i)

There are no such molecules in this entry.

5.6 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Chemical shift validation (i)

The completeness of assignment taking into account all chemical shift lists is 40% for the well-defined parts and 40% for the entire structure.

6.1 Chemical shift list 1

File name: input_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

6.1.1 Bookkeeping (i)

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	198
Number of shifts mapped to atoms	198
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

6.1.2 Chemical shift referencing (i)

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

6.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 40%, i.e. 157 atoms were assigned a chemical shift out of a possible 392. 0 out of 0 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}**H**	^{13}**C**	^{15}**N**
Backbone	0/0 (—%)	0/0 (—%)	0/0 (—%)	0/0 (—%)
Sidechain	0/0 (—%)	0/0 (—%)	0/0 (—%)	0/0 (—%)
Aromatic	0/0 (—%)	0/0 (—%)	0/0 (—%)	0/0 (—%)
Overall	157/392 (40%)	157/232 (68%)	0/136 (0%)	0/24 (0%)

6.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

6.1.5 Random Coil Index (RCI) plots (i)

No *random coil index* (RCI) plot could be generated from the current chemical shift list (assigned_chem_shift_list_1). RCI is only applicable to proteins.