wwPDB NMR Structure Validation Summary Report

Nov 6, 2023 – 05:21 PM EST

PDB ID	:	2KVJ
Title	:	NMR and MD solution structure of a Gamma-Methylated PNA duplex
Authors	:	He, W.; Crawford, M.J.; Rapireddy, S.; Madrid, M.; Gil, R.R.; Ly, D.H.; Achim, C.
Deposited on	:	2010-03-15

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:

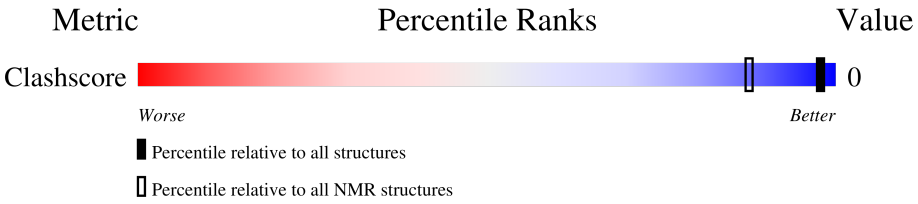
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 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.



Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864

Molprobity failed to run

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.

ENTRY-COMPOSITION INFOmissingINFO

3 Residue-property plots (i)

3.1 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.

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

The representative model is number 1. Colouring as in section 3.1 above.

4 Refinement protocol and experimental data overview

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

Of the 11 calculated structures, 11 were deposited, based on the following criterion: *model 1 is the average of 10 NMR-determined structures.*

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

Software name	Classification	Version
Amber	structure solution	10
Amber	refinement	10

No chemical shift data was provided.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 40A, 40G, 40T, 40C, ACE, NME

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

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	Chirality	Planarity
1	A	1.0±0.0	13.2±1.2
1	B	0.0±0.0	12.4±0.8
All	All	11	281

There are no bond-length outliers.

There are no bond-angle outliers.

All unique chiral outliers are listed below.

Mol	Chain	Res	Type	Atoms	Models (Total)
1	A	2	40G	CA	11

5 of 16 unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	2	40G	Mainchain,Peptide	11
1	A	5	40A	Mainchain,Peptide	11
1	A	6	40T	Peptide,Mainchain	11
1	A	7	40G	Mainchain,Peptide	11
1	A	8	40C	Mainchain,Peptide	11

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 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	A	169	128	52	0±1
All	All	3718	2816	1144	2

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 0.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:1:ACE:H1	1:A:2:40G:C2M	0.99	1.86	1	1
1:A:1:ACE:CH3	1:A:2:40G:C2M	0.83	2.45	1	1

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

There are no ligands in this entry.

5.7 Other polymers (i)

20 such molecules 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	40G	B	13	1	18,23,24	1.40±0.02	5±0 (30±2%)
1	40A	A	5	1	18,22,23	0.92±0.01	2±0 (8±2%)
1	40C	B	19	1	18,19,20	1.35±0.07	4±1 (21±2%)
1	40G	A	2	1	18,23,24	1.78±0.34	6±0 (33±1%)
1	40C	A	4	1	18,19,20	1.32±0.09	4±0 (21±1%)
1	40C	A	9	1	18,19,20	1.37±0.12	4±1 (21±2%)
1	40C	A	8	1	18,19,20	1.25±0.02	3±0 (18±2%)
1	ACE	A	1	-	1,2,2	0.52±0.75	0±0 (9±28%)
1	NME	B	20	1	0,1,1	0.00±0.00	-
1	ACE	B	11	-	1,2,2	0.48±0.81	0±0 (9±28%)
1	NME	A	10	-	0,1,1	0.00±0.00	-
1	40C	B	14	1	18,19,20	1.28±0.02	4±0 (21±1%)
1	40T	A	6	1	19,20,21	1.08±0.02	1±0 (7±2%)
1	40G	B	17	1	18,23,24	1.31±0.01	4±0 (24±2%)
1	40C	B	18	1	18,19,20	1.25±0.03	3±0 (19±2%)
1	40G	B	12	1	18,23,24	1.73±0.35	6±1 (34±3%)
1	40G	A	3	1	18,23,24	1.39±0.04	5±1 (28±3%)
1	40A	B	15	1	18,22,23	0.92±0.01	1±0 (7±2%)
1	40T	B	16	1	19,20,21	1.07±0.01	1±0 (7±2%)
1	40G	A	7	1	18,23,24	1.33±0.01	4±0 (24±2%)

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	40G	B	13	1	16,32,34	1.63±0.02	3±0 (18±0%)
1	40A	A	5	1	14,30,32	1.71±0.02	4±0 (29±2%)
1	40C	B	19	1	19,25,27	2.22±0.11	10±2 (53±8%)
1	40G	A	2	1	16,32,34	1.82±0.10	4±1 (26±4%)
1	40C	A	4	1	19,25,27	2.09±0.03	9±1 (45±3%)
1	40C	A	9	1	19,25,27	2.23±0.12	10±2 (51±9%)
1	40C	A	8	1	19,25,27	2.01±0.03	8±1 (42±4%)
1	ACE	A	1	-	1,1,1	0.76±0.08	0±0 (0±0%)
1	NME	B	20	1	-	-	-
1	ACE	B	11	-	1,1,1	0.66±0.16	0±0 (0±0%)
1	NME	A	10	-	-	-	-
1	40C	B	14	1	19,25,27	2.06±0.03	8±1 (44±4%)
1	40T	A	6	1	22,27,29	1.87±0.03	7±1 (33±2%)
1	40G	B	17	1	16,32,34	1.72±0.04	3±0 (19±2%)
1	40C	B	18	1	19,25,27	2.01±0.03	8±1 (43±3%)
1	40G	B	12	1	16,32,34	1.81±0.14	4±1 (27±8%)
1	40G	A	3	1	16,32,34	1.65±0.04	3±1 (20±5%)
1	40A	B	15	1	14,30,32	1.69±0.03	4±1 (29±3%)
1	40T	B	16	1	22,27,29	1.87±0.03	8±0 (35±2%)
1	40G	A	7	1	16,32,34	1.71±0.03	3±0 (19±1%)

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	40G	A	2	1	1±0,1,3,4	0±0,14,15,16	0±0,2,2,2
1	40C	A	9	1	-	0±0,14,15,16	0±0,1,1,1
1	40G	B	17	1	-	0±0,14,15,16	0±0,2,2,2
1	40A	B	15	1	-	0±0,14,15,16	0±0,2,2,2
1	40G	A	3	1	-	0±0,14,15,16	0±0,2,2,2
1	40A	A	5	1	-	0±0,14,15,16	0±0,2,2,2
1	40C	A	8	1	-	0±0,14,15,16	0±0,1,1,1
1	40C	B	18	1	-	0±0,14,15,16	0±0,1,1,1
1	40T	A	6	1	-	0±0,14,15,16	0±0,1,1,1
1	40G	B	12	1	-	0±0,14,15,16	0±0,2,2,2
1	40G	B	13	1	-	0±0,14,15,16	0±0,2,2,2
1	40C	B	14	1	-	0±0,14,15,16	0±0,1,1,1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	40G	A	7	1	-	0±0,14,15,16	0±0,2,2,2
1	40C	B	19	1	-	0±0,14,15,16	0±0,1,1,1
1	40C	A	4	1	-	0±0,14,15,16	0±0,1,1,1
1	40T	B	16	1	-	0±0,14,15,16	0±0,1,1,1

5 of 76 unique bond outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	B	12	40G	CA-N	9.41	1.23	1.49	10	11
1	A	2	40G	CA-N	7.96	1.27	1.49	9	11
1	B	12	40G	C3'-N4'	4.84	1.37	1.47	1	1
1	A	4	40C	CA-N	4.18	1.37	1.49	1	10
1	A	4	40C	C8'-N1	3.62	1.50	1.46	7	11

5 of 127 unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	B	17	40G	O6-C6-N1	4.51	115.32	120.65	7	11
1	B	12	40G	C2M-CA-N	4.46	89.71	108.81	10	1
1	B	12	40G	C8'-C7'-N4'	4.43	122.63	117.07	1	4
1	A	7	40G	O6-C6-N1	4.26	115.62	120.65	10	11
1	B	19	40C	O2-C2-N3	4.23	115.44	122.33	4	11

All unique chiral outliers are listed below.

Mol	Chain	Res	Type	Atoms	Models (Total)
1	A	2	40G	CA	11

There are no torsion outliers.

There are no ring outliers.

POLYMER-BREAKS INFOmissingINFO

6 Chemical shift validation

No chemical shift data were provided