Full wwPDB NMR Structure Validation Report (i)

Feb 19, 2022 – 10:32 AM EST

PDB ID	:	1QE7
Title	:	SOLUTION STRUCTURE OF A URACIL CONTAINING HAIRPIN DNA
Authors	:	Chary, K.V.R.; Ghosh, M.
Deposited on	:	1999-07-13

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

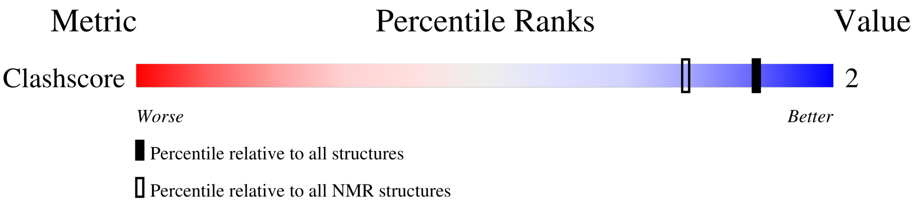
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.26
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.26

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.



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	A	22	45% 36% 18%

2 Ensemble composition and analysis (i)

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 (i)

There is only 1 type of molecule in this entry. The entry contains 567 atoms, of which 204 are hydrogens and 0 are deuteriums.

- Molecule 1 is a DNA chain called URACIL CONTAINING HAIRPIN 22-MER DNA.

Mol	Chain	Residues	Atoms						Trace
1	A	18	Total	C	H	N	O	P	0
			567	175	204	61	110	17	

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: URACIL CONTAINING HAIRPIN 22-MER DNA

Chain A:

4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

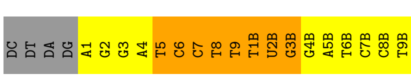
• Molecule 1: URACIL CONTAINING HAIRPIN 22-MER DNA

Chain A:

4.2.2 Score per residue for model 2

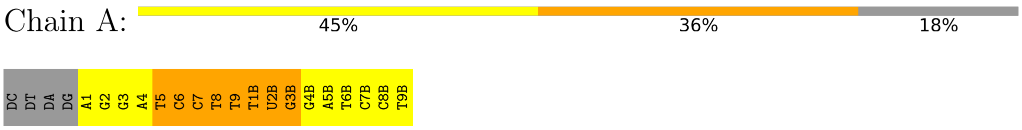
• Molecule 1: URACIL CONTAINING HAIRPIN 22-MER DNA

Chain A:



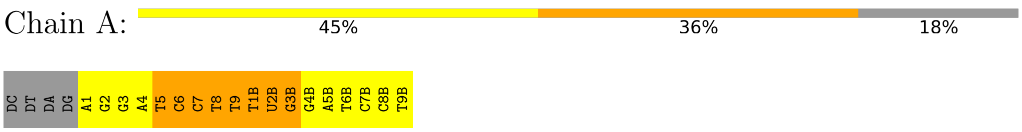
4.2.3 Score per residue for model 3

• Molecule 1: URACIL CONTAINING HAIRPIN 22-MER DNA



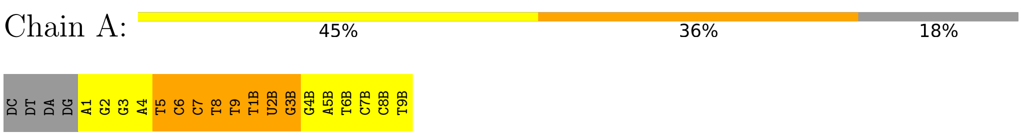
4.2.4 Score per residue for model 4

- Molecule 1: URACIL CONTAINING HAIRPIN 22-MER DNA



4.2.5 Score per residue for model 5

- Molecule 1: URACIL CONTAINING HAIRPIN 22-MER DNA

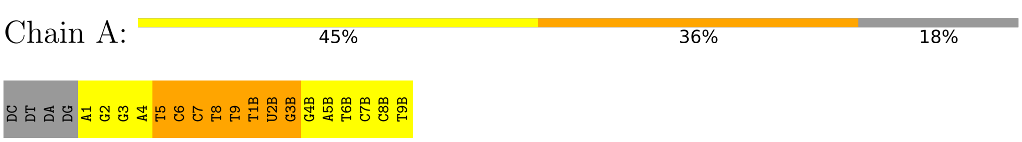


4.2.6 Score per residue for model 6

- Molecule 1: URACIL CONTAINING HAIRPIN 22-MER DNA

4.2.7 Score per residue for model 7

- Molecule 1: URACIL CONTAINING HAIRPIN 22-MER DNA



4.2.8 Score per residue for model 8

- Molecule 1: URACIL CONTAINING HAIRPIN 22-MER DNA

4.2.9 Score per residue for model 9

- Molecule 1: URACIL CONTAINING HAIRPIN 22-MER DNA

4.2.10 Score per residue for model 10

- Molecule 1: URACIL CONTAINING HAIRPIN 22-MER DNA

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *RESTRAINED MOLECULAR DYNAMICS, ENERGY MINIMIZATION*.

Of the 200 calculated structures, 10 were deposited, based on the following criterion: *STRUCTURES WITH ACCEPTABLE COVALENT GEOMETRY,STRUCTURES WITH FAVORABLE NON- BOND ENERGY, STRUCTURES WITH THE LEAST RESTRAINT VIOLATIONS,STRUCTURES WITH THE LOWEST ENERGY*.

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

Software name	Classification	Version
Discover	structure solution	97
Discover	refinement	97

No chemical shift data was provided.

6 Model quality (i)

6.1 Standard geometry (i)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the (average) root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	1.64±0.00	6±0/405 (1.5± 0.0%)	2.54±0.01	47±1/623 (7.6± 0.2%)
All	All	1.64	60/4050 (1.5%)	2.54	473/6230 (7.6%)

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	0.0±0.0	5.8±0.4
All	All	0	58

All 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	A	2(B)	DU	N1-C6	9.14	1.46	1.38	5	10
1	A	1(B)	DT	C5-C7	6.10	1.53	1.50	10	10
1	A	6(B)	DT	C5-C7	5.83	1.53	1.50	8	10
1	A	8	DT	C5-C7	5.78	1.53	1.50	9	10
1	A	9(B)	DT	C5-C7	5.77	1.53	1.50	1	10
1	A	5	DT	C5-C7	5.61	1.53	1.50	4	10

All 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(°)	Ideal(°)	Models	
								Worst	Total
1	A	9	DT	C6-C5-C7	-10.02	116.89	122.90	9	10
1	A	1	DA	N1-C6-N6	-8.26	113.64	118.60	1	10
1	A	9(B)	DT	C6-C5-C7	-8.11	118.03	122.90	2	10
1	A	2(B)	DU	C2-N1-C1'	7.80	127.06	117.70	8	10

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	4	DA	C5-C6-N1	7.64	121.52	117.70	7	10
1	A	2(B)	DU	N1-C2-N3	7.63	119.48	114.90	7	10
1	A	2(B)	DU	N3-C2-O2	-7.55	116.92	122.20	8	10
1	A	2(B)	DU	C5-C6-N1	-7.40	119.00	122.70	9	10
1	A	5(B)	DA	C5-C6-N1	7.26	121.33	117.70	3	10
1	A	5	DT	O4'-C1'-C2'	-7.16	100.17	105.90	3	10
1	A	1	DA	C5-C6-N1	7.15	121.28	117.70	6	10
1	A	1	DA	C4-C5-C6	-6.67	113.67	117.00	1	10
1	A	6(B)	DT	N3-C2-O2	-6.62	118.33	122.30	7	10
1	A	6(B)	DT	C6-C5-C7	-6.59	118.95	122.90	7	10
1	A	4	DA	N1-C6-N6	-6.45	114.73	118.60	6	10
1	A	5(B)	DA	N1-C6-N6	-6.30	114.82	118.60	7	10
1	A	5	DT	C6-C5-C7	-6.28	119.13	122.90	6	10
1	A	1(B)	DT	N3-C2-O2	-6.26	118.54	122.30	8	10
1	A	2(B)	DU	P-O3'-C3'	-6.25	112.20	119.70	8	10
1	A	8(B)	DC	N3-C2-O2	-6.20	117.56	121.90	7	10
1	A	3(B)	DG	C5-C6-N1	6.19	114.59	111.50	5	10
1	A	5	DT	C1'-O4'-C4'	-6.17	103.93	110.10	2	10
1	A	9(B)	DT	N3-C2-O2	-6.16	118.61	122.30	1	10
1	A	4	DA	C4-C5-C6	-6.12	113.94	117.00	6	10
1	A	2	DG	C5-C6-N1	6.05	114.53	111.50	8	10
1	A	5(B)	DA	C4-C5-C6	-6.04	113.98	117.00	8	10
1	A	5	DT	N3-C2-O2	-6.01	118.69	122.30	3	10
1	A	3	DG	C5-C6-N1	5.99	114.50	111.50	5	10
1	A	1	DA	O4'-C1'-N9	5.98	112.19	108.00	1	10
1	A	8	DT	C6-C5-C7	-5.98	119.31	122.90	1	10
1	A	9	DT	N3-C2-O2	-5.97	118.72	122.30	9	10
1	A	4(B)	DG	C5-C6-N1	5.83	114.42	111.50	6	10
1	A	7(B)	DC	N3-C2-O2	-5.83	117.82	121.90	3	10
1	A	2(B)	DU	C6-N1-C1'	-5.78	113.10	121.20	8	10
1	A	8	DT	C5-C6-N1	-5.72	120.27	123.70	1	10
1	A	7	DC	N3-C2-O2	-5.71	117.90	121.90	10	10
1	A	4(B)	DG	O4'-C1'-N9	5.71	112.00	108.00	6	10
1	A	8	DT	N3-C2-O2	-5.67	118.90	122.30	10	10
1	A	7	DC	O4'-C1'-N1	5.66	111.96	108.00	6	3
1	A	2	DG	O4'-C1'-N9	5.63	111.94	108.00	3	8
1	A	6	DC	N3-C2-O2	-5.57	118.00	121.90	4	10
1	A	9(B)	DT	C4-C5-C7	5.55	122.33	119.00	2	9
1	A	4(B)	DG	N1-C6-O6	-5.52	116.59	119.90	8	10
1	A	1(B)	DT	C5-C6-N1	-5.51	120.40	123.70	8	10
1	A	3	DG	N1-C6-O6	-5.35	116.69	119.90	2	10

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	2	DG	O4'-C1'-C2'	-5.30	101.66	105.90	3	1
1	A	2(B)	DU	C6-N1-C2	-5.21	117.88	121.00	8	5
1	A	9	DT	O4'-C4'-C3'	5.15	109.09	106.00	9	4
1	A	3(B)	DG	N1-C6-O6	-5.15	116.81	119.90	1	8
1	A	1(B)	DT	O4'-C1'-N1	5.09	111.56	108.00	5	3
1	A	2	DG	N1-C6-O6	-5.08	116.85	119.90	2	7
1	A	8(B)	DC	N1-C2-O2	5.07	121.94	118.90	7	5

There are no chirality outliers.

All 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	7	DC	Sidechain	10
1	A	9	DT	Sidechain	10
1	A	1(B)	DT	Sidechain	10
1	A	2(B)	DU	Sidechain	10
1	A	3(B)	DG	Sidechain	10
1	A	8	DT	Sidechain	8

6.2 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	363	204	205	1±0
All	All	3630	2040	2050	10

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 2.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:5:DT:C2	1:A:6:DC:C5	0.45	3.05	1	10

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)

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