



# Full wwPDB NMR Structure Validation Report ⓘ

Mar 5, 2022 – 03:48 PM EST

PDB ID : 2JX1  
Title : Structure of the fifth zinc finger of Myelin Transcription Factor 1 in complex with RARE DNA  
Authors : Gamsjaeger, R.; Swanton, M.K.; Kobus, F.J.; Lehtomaki, E.; Lowry, J.A.; Kwan, A.H.; Matthews, J.M.; Mackay, J.P.  
Deposited on : 2007-11-01

This is a Full wwPDB NMR Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) 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 : 2.27  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.27

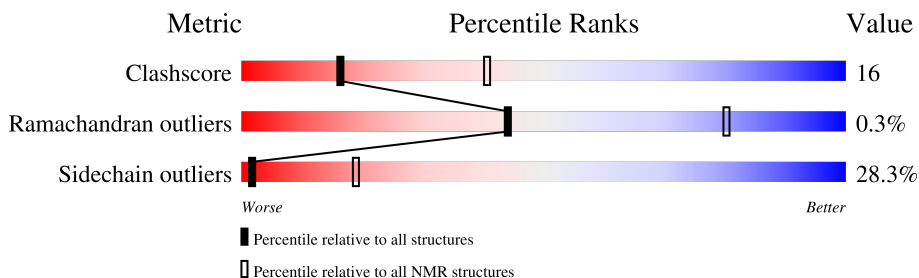
# 1 Overall quality at a glance

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
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

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  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	B	13	
2	C	13	
3	A	31	

## 2 Ensemble composition and analysis

This entry contains 10 models. Model 3 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	Residue range (total)	Backbone RMSD (Å)	Medoid model
1	A:10-A:39 (30)	0.28	3

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 2 clusters and 2 single-model clusters were found.

Cluster number	Models
1	1, 4, 5, 6, 8
2	2, 3, 7
Single-model clusters	9; 10

### 3 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 1249 atoms, of which 503 are hydrogens and 0 are deuteriums.

- Molecule 1 is a DNA chain called DNA (5'-D(\*DAP\*DCP\*DCP\*DGP\*DAP\*DAP\*DAP\*DGP\*DTP\*DTP\*DCP\*DAP\*DC)-3').

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
1	B	13	409	126	147	51	73	12	0

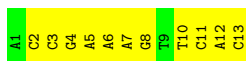
- Molecule 2 is a DNA chain called DNA (5'-D(\*DGP\*DTP\*DGP\*DAP\*DAP\*DCP\*DTP\*DTP\*DTP\*DCP\*DGP\*DGP\*DT)-3').

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
2	C	13	415	128	150	46	79	12	0

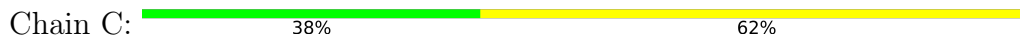
- Molecule 3 is a protein called Myelin transcription factor 1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
3	A	31	425	129	206	43	44	3	0

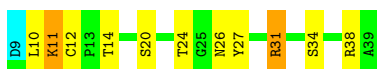




- Molecule 2: DNA (5'-D(\*DGP\*DTP\*DGP\*DAP\*DAP\*DCP\*DTP\*DTP\*DTP\*DCP\*DGP\*DG\*DT)-3')

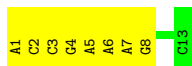


- Molecule 3: Myelin transcription factor 1

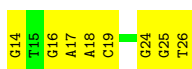


#### 4.2.2 Score per residue for model 2

- Molecule 1: DNA (5'-D(\*DAP\*DCP\*DCP\*DGP\*DAP\*DAP\*DAP\*DGP\*DTP\*DTP\*DCP\*DAP\*DC)-3')



- Molecule 2: DNA (5'-D(\*DGP\*DTP\*DGP\*DAP\*DAP\*DCP\*DTP\*DTP\*DTP\*DCP\*DGP\*DG\*DT)-3')



- Molecule 3: Myelin transcription factor 1



#### 4.2.3 Score per residue for model 3 (medoid)

- Molecule 1: DNA (5'-D(\*DAP\*DCP\*DCP\*DGP\*DAP\*DAP\*DAP\*DGP\*DTP\*DTP\*DCP\*DAP\*DC)-3')

Chain B:  54% 46%



- Molecule 2: DNA (5'-D(\*DGP\*DTP\*DGP\*DAP\*DAP\*DCP\*DTP\*DTP\*DTP\*DCP\*DGP\*DG\*DT)-3')

Chain C:  15% 85%



- Molecule 3: Myelin transcription factor 1

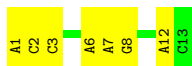
Chain A:  74% 23%



#### 4.2.4 Score per residue for model 4

- Molecule 1: DNA (5'-D(\*DAP\*DCP\*DCP\*DGP\*DAP\*DAP\*DAP\*DGP\*DTP\*DTP\*DCP\*DA\*DC)-3')

Chain B:  46% 54%



- Molecule 2: DNA (5'-D(\*DGP\*DTP\*DGP\*DAP\*DAP\*DCP\*DTP\*DTP\*DTP\*DCP\*DGP\*DG\*DT)-3')

Chain C:  46% 54%



- Molecule 3: Myelin transcription factor 1

Chain A:  68% 29%



### 4.2.5 Score per residue for model 5

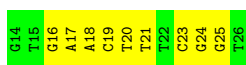
- Molecule 1: DNA (5'-D(\*DAP\*DCP\*DCP\*DGP\*DAP\*DAP\*DAP\*DGP\*DTP\*DTP\*DCP\*DAP\*DC)-3')

Chain B:  31% 69%



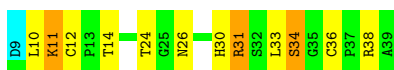
- Molecule 2: DNA (5'-D(\*DGP\*DTP\*DGP\*DAP\*DAP\*DCP\*DTP\*DTP\*DTP\*DCP\*DGP\*DGP\*DT)-3')

Chain C:  31% 69%



- Molecule 3: Myelin transcription factor 1

Chain A:  58% 29% 10% .



### 4.2.6 Score per residue for model 6

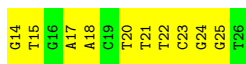
- Molecule 1: DNA (5'-D(\*DAP\*DCP\*DCP\*DGP\*DAP\*DAP\*DAP\*DGP\*DTP\*DTP\*DCP\*DAP\*DC)-3')

Chain B:  54% 46%



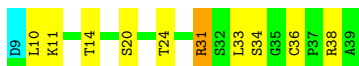
- Molecule 2: DNA (5'-D(\*DGP\*DTP\*DGP\*DAP\*DAP\*DCP\*DTP\*DTP\*DTP\*DCP\*DGP\*DGP\*DT)-3')

Chain C:  23% 77%



- Molecule 3: Myelin transcription factor 1

Chain A:  65% 29% . .

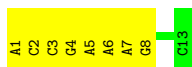




#### 4.2.7 Score per residue for model 7

- Molecule 1: DNA (5'-D(\*DAP\*DCP\*DCP\*DGP\*DAP\*DAP\*DAP\*DGP\*DTP\*DTP\*DCP\*DAP\*DC)-3')

Chain B:  38% 62%



- Molecule 2: DNA (5'-D(\*DGP\*DTP\*DGP\*DAP\*DAP\*DCP\*DTP\*DTP\*DTP\*DCP\*DGP\*DIGP\*DT)-3')

Chain C:  54% 46%



- Molecule 3: Myelin transcription factor 1

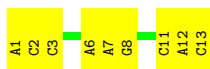
Chain A:  68% 26% . .



#### 4.2.8 Score per residue for model 8

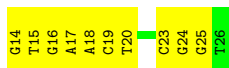
- Molecule 1: DNA (5'-D(\*DAP\*DCP\*DCP\*DGP\*DAP\*DAP\*DAP\*DGP\*DTP\*DTP\*DCP\*DAP\*DC)-3')

Chain B:  31% 69%



- Molecule 2: DNA (5'-D(\*DGP\*DTP\*DGP\*DAP\*DAP\*DCP\*DTP\*DTP\*DTP\*DCP\*DGP\*DIGP\*DT)-3')

Chain C:  23% 77%




- Molecule 3: Myelin transcription factor 1

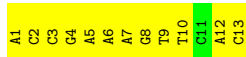
Chain A:  68% 26% . .



### 4.2.9 Score per residue for model 9

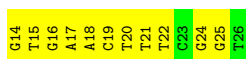
- Molecule 1: DNA (5'-D(\*DAP\*DCP\*DCP\*DGP\*DAP\*DAP\*DAP\*DGP\*DTP\*DTP\*DCP\*DAP\*DC)-3')

Chain B:  8% 92%

  
A1 C2 C3 G4 A5 A6 A7 G8 T9 T10 C11 A12 C13

- Molecule 2: DNA (5'-D(\*DGP\*DTP\*DGP\*DAP\*DAP\*DCP\*DTP\*DTP\*DTP\*DCP\*DGP\*DG\*DT)-3')

Chain C:  15% 85%

  
G14 T15 G16 A17 A18 C19 T20 T21 T22 C23 G24 G25 T26


- Molecule 3: Myelin transcription factor 1

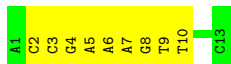
Chain A:  55% 35% 6% .

  
D9 L10 K11 C12 P13 T14 G19 S20 T23 T24 S29 L33 S34 G35 C36 P37 R38 A39

### 4.2.10 Score per residue for model 10

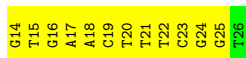
- Molecule 1: DNA (5'-D(\*DAP\*DCP\*DCP\*DGP\*DAP\*DAP\*DAP\*DGP\*DTP\*DTP\*DCP\*DAP\*DC)-3')

Chain B:  31% 69%

  
A1 C2 G4 A5 A6 A7 G8 T10 C13

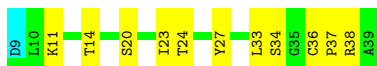
- Molecule 2: DNA (5'-D(\*DGP\*DTP\*DGP\*DAP\*DAP\*DCP\*DTP\*DTP\*DTP\*DCP\*DGP\*DG\*DT)-3')

Chain C:  8% 92%

  
G14 T15 G16 A17 A18 C19 T20 T21 T22 C23 G24 G25 T26

- Molecule 3: Myelin transcription factor 1

Chain A:  61% 35% .

  
D9 L10 K11 T14 S20 T23 T24 Y27 L33 S34 C36 P37 R38 A39

## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *HADDOCK docking, simulated annealing*.

Of the 10 calculated structures, 10 were deposited, based on the following criterion: *structures with the lowest energy*.

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

Software name	Classification	Version
CNS	refinement	

No chemical shift data was provided.

## 6 Model quality i

### 6.1 Standard geometry i

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
1	B	262	147	147	10±3
2	C	265	150	150	11±3
3	A	211	200	200	3±1
All	All	7380	4970	4970	195

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 16.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:B:3:DC:O2	2:C:24:DG:N2	0.97	1.97	8	10
1:B:2:DC:O2	2:C:25:DG:N2	0.97	1.97	3	10
2:C:14:DG:HO5'	2:C:14:DG:H8	0.79	1.18	6	7
1:B:1:DA:H8	1:B:1:DA:HO5'	0.77	1.20	9	4
1:B:1:DA:HO5'	1:B:1:DA:H8	0.71	1.26	6	3
1:B:7:DA:H2''	1:B:8:DG:N7	0.67	2.03	1	10
3:A:33:LEU:HA	3:A:36:CYS:SG	0.66	2.31	9	5
2:C:17:DA:C4	2:C:18:DA:N7	0.59	2.71	9	8
1:B:1:DA:H8	1:B:1:DA:O5'	0.57	1.82	8	5
2:C:19:DC:N4	3:A:27:TYR:CZ	0.56	2.73	2	7
2:C:16:DG:C5	2:C:17:DA:C6	0.55	2.94	9	5
2:C:20:DT:H2''	2:C:21:DT:H71	0.54	1.78	6	2

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:C:19:DC:N4	3:A:27:TYR:CE2	0.51	2.77	7	4
1:B:6:DA:C6	1:B:7:DA:N6	0.51	2.79	6	7
1:B:6:DA:H2''	1:B:7:DA:O5'	0.49	2.07	2	3
1:B:1:DA:O5'	1:B:1:DA:C8	0.49	2.64	8	2
3:A:11:LYS:HD2	3:A:12:CYS:N	0.49	2.22	9	5
2:C:16:DG:H2''	2:C:17:DA:C8	0.48	2.43	5	2
2:C:16:DG:C6	2:C:17:DA:C6	0.48	3.02	2	3
2:C:17:DA:C5	2:C:18:DA:N7	0.48	2.81	9	1
1:B:4:DG:C5	1:B:5:DA:C6	0.48	3.02	7	1
3:A:19:GLY:O	3:A:29:SER:HB3	0.47	2.08	9	1
1:B:10:DT:H2''	1:B:11:DC:C5	0.47	2.45	1	1
3:A:10:LEU:O	3:A:31:ARG:HD3	0.47	2.09	1	4
2:C:16:DG:C2	2:C:17:DA:C2	0.47	3.03	5	1
2:C:23:DC:H2''	2:C:24:DG:C8	0.47	2.45	5	3
1:B:1:DA:C8	1:B:1:DA:O5'	0.47	2.68	5	2
1:B:5:DA:C4	1:B:6:DA:C6	0.46	3.03	2	1
1:B:8:DG:O5'	1:B:8:DG:C8	0.46	2.69	9	2
2:C:21:DT:H2''	2:C:22:DT:H71	0.45	1.87	4	4
1:B:6:DA:N6	1:B:7:DA:N6	0.45	2.65	2	3
2:C:21:DT:C2'	2:C:22:DT:H71	0.45	2.42	1	2
2:C:17:DA:C6	2:C:18:DA:N6	0.45	2.84	8	3
2:C:22:DT:H2''	2:C:23:DC:OP2	0.45	2.12	7	3
2:C:14:DG:H8	2:C:14:DG:O5'	0.45	1.94	10	2
1:B:4:DG:C6	1:B:5:DA:C6	0.45	3.05	7	4
3:A:17:CYS:SG	3:A:36:CYS:HB3	0.45	2.51	7	1
1:B:9:DT:C6	1:B:10:DT:H72	0.45	2.47	9	1
3:A:23:ILE:HD11	3:A:37:PRO:HA	0.45	1.89	10	2
1:B:3:DC:H1'	1:B:4:DG:N7	0.44	2.27	9	1
1:B:2:DC:C2	1:B:3:DC:C4	0.44	3.05	2	1
1:B:4:DG:C2	1:B:5:DA:C2	0.44	3.05	10	1
2:C:14:DG:O5'	2:C:14:DG:C8	0.44	2.71	10	1
1:B:12:DA:C4	1:B:13:DC:C5	0.44	3.06	1	1
1:B:4:DG:C2	2:C:24:DG:C2	0.44	3.06	7	2
2:C:19:DC:O2	2:C:20:DT:C2	0.44	2.71	9	3
2:C:20:DT:C2'	2:C:21:DT:H71	0.44	2.43	6	1
2:C:25:DG:H2''	2:C:26:DT:H71	0.43	1.89	2	1
2:C:15:DT:C4	2:C:16:DG:O6	0.43	2.70	9	1
1:B:9:DT:H1'	1:B:10:DT:H5'	0.43	1.89	10	1
1:B:4:DG:N1	2:C:24:DG:C2	0.43	2.86	10	4
1:B:1:DA:C2	1:B:2:DC:C2	0.43	3.07	8	1
2:C:25:DG:C2'	2:C:26:DT:H72	0.43	2.43	1	1

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:C:14:DG:H2''	2:C:15:DT:H71	0.43	1.90	6	2
1:B:5:DA:C5	1:B:6:DA:N6	0.43	2.87	2	2
2:C:21:DT:H2'	2:C:22:DT:H72	0.42	1.92	10	1
2:C:23:DC:H2''	2:C:24:DG:OP2	0.42	2.14	3	2
1:B:1:DA:H2''	1:B:2:DC:C6	0.42	2.50	6	1
2:C:24:DG:H2''	2:C:25:DG:OP2	0.42	2.15	9	1
1:B:8:DG:C8	1:B:8:DG:O5'	0.41	2.73	7	2
1:B:12:DA:H2''	1:B:13:DC:OP2	0.41	2.14	9	1
1:B:6:DA:H2''	1:B:7:DA:C8	0.41	2.50	9	1
2:C:18:DA:H2''	2:C:19:DC:OP2	0.41	2.16	9	1
1:B:11:DC:H2''	1:B:12:DA:OP2	0.41	2.15	5	1
1:B:11:DC:C2	1:B:12:DA:C5	0.41	3.08	8	1
1:B:12:DA:C4	1:B:13:DC:C4	0.41	3.09	8	1
2:C:21:DT:C2'	2:C:22:DT:H72	0.41	2.45	10	1
1:B:1:DA:C4	1:B:2:DC:C4	0.41	3.09	2	1
2:C:15:DT:H2''	2:C:16:DG:C8	0.41	2.50	8	1
3:A:10:LEU:O	3:A:31:ARG:HD2	0.41	2.16	2	1
2:C:21:DT:H2''	2:C:22:DT:C6	0.41	2.51	6	1
1:B:1:DA:C2	1:B:2:DC:N3	0.41	2.89	8	1
2:C:20:DT:C2'	2:C:21:DT:H72	0.41	2.45	10	1
2:C:14:DG:H2''	2:C:15:DT:C5	0.41	2.51	10	2
1:B:1:DA:O5'	1:B:1:DA:H8	0.41	1.99	7	1
1:B:6:DA:C5	1:B:7:DA:N6	0.41	2.90	9	1
1:B:5:DA:H2''	1:B:6:DA:N7	0.40	2.32	2	1
1:B:12:DA:C2	2:C:16:DG:C2	0.40	3.10	4	1
2:C:20:DT:H2''	2:C:21:DT:C7	0.40	2.47	6	1

## 6.3 Torsion angles [i](#)

### 6.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the backbone conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
3	A	29/31 (94%)	27±1 (93±4%)	2±1 (7±4%)	0±0 (0±1%)	44 80
All	All	290/310 (94%)	270 (93%)	19 (7%)	1 (0%)	44 80

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
3	A	34	SER	1

### 6.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the sidechain conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	A	23/24 (96%)	16±1 (72±4%)	6±1 (28±4%)	2	19
All	All	230/240 (96%)	165 (72%)	65 (28%)	2	19

All 10 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
3	A	11	LYS	10
3	A	14	THR	10
3	A	24	THR	10
3	A	34	SER	10
3	A	38	ARG	9
3	A	20	SER	8
3	A	31	ARG	4
3	A	26	ASN	2
3	A	30	HIS	1
3	A	12	CYS	1

### 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

No chemical shift data were provided