



# wwPDB NMR Structure Validation Summary Report ⓘ

Feb 12, 2017 – 05:25 pm GMT

PDB ID : 1A66  
Title : SOLUTION NMR STRUCTURE OF THE CORE NFATC1/DNA COMPLEX, 18 STRUCTURES  
Authors : Zhou, P.; Sun, L.J.; Doetsch, V.; Wagner, G.; Verdine, G.L.  
Deposited on : 1998-03-06

This is a wwPDB NMR Structure Validation Summary 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

<http://wwpdb.org/validation/2016/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:

Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	trunk28760
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

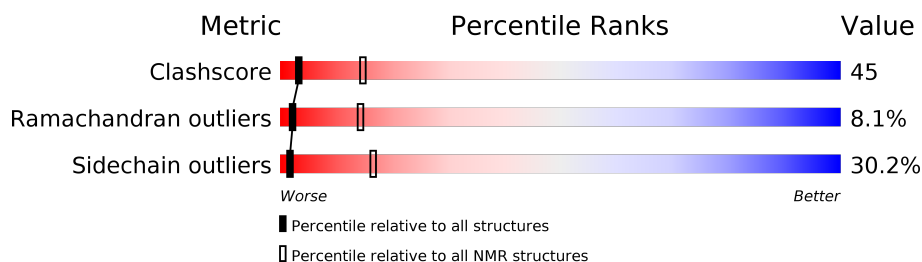
# 1 Overall quality at a glance

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	125131	11601
Ramachandran outliers	121729	10391
Sidechain outliers	121581	10367

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	12	
2	C	12	
3	A	178	

## 2 Ensemble composition and analysis

This entry contains 18 models. Model 3 is the overall representative, medoid model (most similar to other models).

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:40, A:45-A:177 (164)	0.77	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 3 single-model clusters were found.

Cluster number	Models
1	1, 2, 3, 5, 6, 8, 9, 11, 12, 13, 15, 17, 18
2	4, 16
Single-model clusters	7; 10; 14

### 3 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3595 atoms, of which 1704 are hydrogens and 0 are deuteriums.

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

Mol	Chain	Residues	Atoms						Trace
1	B	12	Total	C	H	N	O	P	0
			385	119	136	52	67	11	

- Molecule 2 is a DNA chain called DNA (5'-D(\*CP\*AP\*AP\*TP\*TP\*TP\*TP\*CP\*CP\*TP\*CP\*G)-3').

Mol	Chain	Residues	Atoms						Trace
2	C	12	Total	C	H	N	O	P	0
			376	116	139	37	73	11	

- Molecule 3 is a protein called CORE NFATC1.

Mol	Chain	Residues	Atoms						Trace
3	A	178	Total	C	H	N	O	S	0
			2834	875	1429	267	258	5	

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	ALA	ENGINEERED	UNP O95644
A	2	LYS	LEU	ENGINEERED	UNP O95644
A	28	ARG	HIS	ENGINEERED	UNP O95644

## 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(\*CP\*GP\*AP\*GP\*GP\*AP\*AP\*AP\*AP\*TP\*TP\*G)-3')

Chain B:  8% 92%

C315  
C316  
A317  
G318  
G319  
A320  
A321  
A322  
A323  
T324  
T325  
G326

- Molecule 2: DNA (5'-D(\*CP\*AP\*AP\*TP\*TP\*TP\*TP\*CP\*CP\*TP\*CP\*G)-3')

Chain C:  33% 67%

C340  
A341  
A342  
T343  
T344  
T345  
T346  
C347  
C348  
T349  
C350  
G351

- Molecule 3: CORE NFATC1

Chain A:  25% 58% 9% 8%

M1  
K2  
D3  
W4  
Q5  
L6  
P7  
S8  
H9  
S10  
G11  
P12  
Y13  
E14  
L15  
R16  
V17  
I18  
E19  
V19  
Q20  
P21  
K22  
H25  
R26  
A27  
R28  
Y29  
E30  
T31  
E32  
R35  
G36  
A37  
V38  
K39  
A40  
S41  
A42  
G43  
G44  
H45  
P46  
L47  
V48  
Q49  
L50  
H51  
G52  
Y53  
L54  
E55  
N56  
E57  
P58  
L59  
M60  
L61  
Q62

L63  
F64  
I65  
G66  
T67  
R71  
L72  
L73  
R74  
P75  
H76  
A77  
F78  
Y79  
Q80  
V81  
H82  
R83  
I84  
T85  
G86  
K87  
T88  
V89  
T92  
S93  
H94  
E95  
A96  
I97  
L98  
S99  
N100  
K101  
T102  
V103  
L104  
E105  
I106  
P107  
L108  
L109  
P110  
M114  
V117  
I118  
D119  
C120  
T123  
L124  
K125  
L126  
R127  
N128  
S129

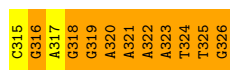
D130  
I131  
E132  
L133  
R134  
K135  
G136  
E137  
T138  
D139  
I140  
G141  
R142  
K143  
N144  
T145  
R146  
V147  
L148  
L149  
V150  
F151  
R152  
V153  
H154  
V155  
P156  
Q157  
P158  
S159  
G160  
R161  
T162  
L163  
S164  
L165  
Q166  
V167  
A168  
S169  
I172  
E173  
C174  
S175  
Q176  
R177  
S178

### 4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 3. Colouring as in section 4.1 above.

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

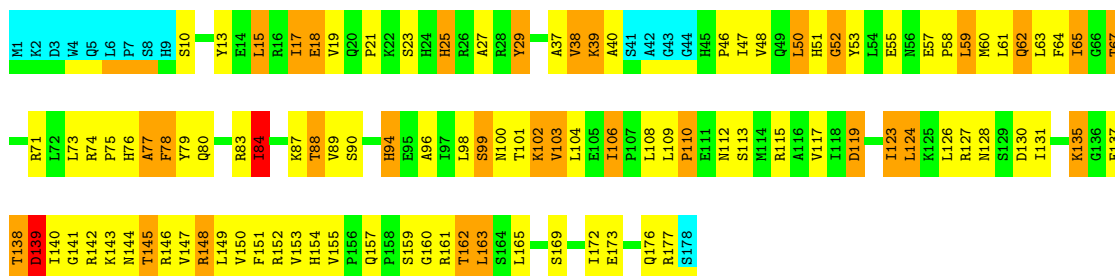
Chain B:  17% 83%



- Molecule 2: DNA (5'-D(\*CP\*AP\*AP\*TP\*TP\*TP\*TP\*CP\*CP\*TP\*CP\*G)-3')



- Molecule 3: CORE NFATC1



## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *TORSION ANGLE DYNAMICS AND SIMULATED ANNEALING*.

Of the 18 calculated structures, 18 were deposited, based on the following criterion: *NOE VIOLATION  $\leq 0.4$  ANGSTROM, DIHEDRAL ANGLE VIOLATION  $\leq 5$  DEGREE*.

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

Software name	Classification	Version
X-PLOR	refinement	3.1
DYANA-1.4	structure solution	
X-PLOR	structure solution	3.1

No chemical shift data was provided. No validations of the models with respect to experimental NMR restraints is performed at this time.

## 6 Model quality

### 6.1 Standard geometry

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	B	1.28±0.02	2±0/281 (0.7±0.1%)	2.45±0.01	25±1/433 (5.8±0.3%)
2	C	1.23±0.01	1±0/263 (0.3±0.2%)	2.26±0.01	17±1/403 (4.3±0.3%)
3	A	1.02±0.01	0±0/1325 (0.0±0.0%)	0.80±0.01	0±0/1794 (0.0±0.0%)
All	All	1.10	51/33642 (0.2%)	1.49	765/47340 (1.6%)

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	B	325	DT	C5-C7	5.54	1.53	1.50	12	17
2	C	345	DT	C5-C7	5.49	1.53	1.50	1	15
1	B	324	DT	C5-C7	5.48	1.53	1.50	3	18
2	C	343	DT	C5-C7	5.11	1.53	1.50	13	1

5 of 59 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	B	326	DG	N7-C8-N9	9.71	117.95	113.10	8	18
1	B	316	DG	N7-C8-N9	9.38	117.79	113.10	16	18
2	C	351	DG	N7-C8-N9	9.23	117.71	113.10	5	18
1	B	319	DG	N7-C8-N9	9.15	117.68	113.10	17	18
1	B	318	DG	N7-C8-N9	8.94	117.57	113.10	11	18

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	249	136	136	15±3
2	C	237	139	139	21±3
3	A	1300	1332	1332	127±13
All	All	32148	28926	28926	2773

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 45.

5 of 1008 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
3:A:61:LEU:HD13	3:A:153:VAL:HG22	1.11	1.11	14	18
3:A:15:LEU:HD13	3:A:153:VAL:HG21	1.09	1.23	10	18
3:A:37:ALA:HB2	3:A:123:ILE:HG22	1.05	1.17	14	6
3:A:79:TYR:CE2	3:A:126:LEU:HD21	1.04	1.87	13	7
3:A:47:ILE:HD12	3:A:117:VAL:HG13	1.00	1.34	17	12

## 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	164/178 (92%)	114±5 (70±3%)	37±5 (22±3%)	13±3 (8±2%)	2	14
All	All	2952/3204 (92%)	2054 (70%)	659 (22%)	239 (8%)	2	14

5 of 49 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
3	A	99	SER	18

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
3	A	110	PRO	18
3	A	77	ALA	16
3	A	18	GLU	15
3	A	52	GLY	14

### 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	145/156 (93%)	101±7 (70±5%)	44±7 (30±5%)	2	16
All	All	2610/2808 (93%)	1822 (70%)	788 (30%)	2	16

5 of 113 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	65	ILE	18
3	A	15	LEU	16
3	A	161	ARG	16
3	A	62	GLN	15
3	A	67	THR	15

### 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 carbohydrates 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