



# wwPDB NMR Structure Validation Summary Report ⓘ

Feb 20, 2018 – 03:09 am GMT

PDB ID : 1GCC  
Title : SOLUTION NMR STRUCTURE OF THE COMPLEX OF GCC-BOX BINDING DOMAIN OF ATERF1 AND GCC-BOX DNA, MINIMIZED AVERAGE STRUCTURE  
Authors : Yamasaki, K.; Allen, M.D.; Ohme-Takagi, M.; Tateno, M.; Suzuki, M.  
Deposited on : 1998-03-13

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

<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:

Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	trunk30686
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk30686

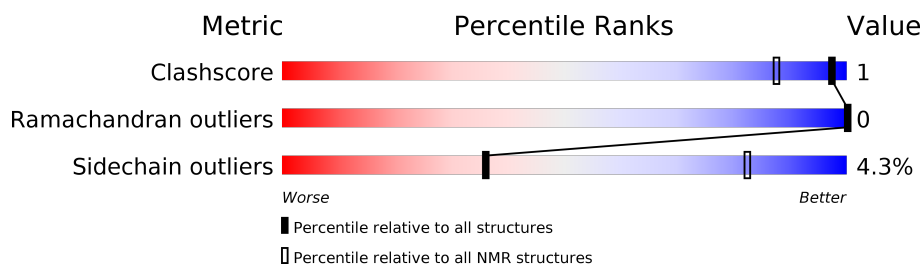
# 1 Overall quality at a glance

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	136279	12091
Ramachandran outliers	132675	10835
Sidechain outliers	132484	10811

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	11	
2	C	11	
3	A	63	

## 2 Ensemble composition and analysis ⓘ

This entry contains 1 models. Identification of well-defined residues and clustering analysis are not possible.

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

There are 3 unique types of molecules in this entry. The entry contains 1719 atoms, of which 761 are hydrogens and 0 are deuteriums.

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

Mol	Chain	Residues	Atoms						Trace
1	B	11	Total	C	H	N	O	P	0
			343	105	123	42	63	10	

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

Mol	Chain	Residues	Atoms						Trace
2	C	11	Total	C	H	N	O	P	0
			349	107	124	43	65	10	

- Molecule 3 is a protein called ETHYLENE RESPONSIVE ELEMENT BINDING FACTOR 1.

Mol	Chain	Residues	Atoms						Trace
3	A	63	Total	C	H	N	O	S	0
			1027	325	514	103	84	1	

## 4 Residue-property plots [i](#)

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

Chain B: 



- Molecule 2: DNA (5'-D(\*GP\*CP\*TP\*GP\*GP\*CP\*GP\*GP\*CP\*TP\*A)-3')

Chain C: 



- Molecule 3: ETHYLENE RESPONSIVE ELEMENT BINDING FACTOR 1

Chain A: 



## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *SIMULATED ANNEALING PROTOCOL IN X-PLOR 3.1 WAS CARRIED OUT TO OBTAIN 25 STRUCTURES. MINIMIZED MEAN STRUCTURE OF THEM WAS FURTHER REVISED BY A 300 PS RESTRAINED MOLECULAR DYNAMICAL CALCULATION IN THE PRESENCE OF APPROXIMATELY 5000 TIP3P WATER MOLECULES, USING CORNELL FORCE FIELD IN AMBER 4.1. THE MEAN COORDINATE OF 200 STRUCTURES OBTAINED AFTER THE EQUILIBRIUM (100 PS) WAS ENERGY-MINIMIZED..*

Of the 1 calculated structures, 1 were deposited, based on the following criterion: ?.

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

Software name	Classification	Version
X-PLOR	refinement	3.1
X-PLOR	structure solution	3.1
AMBER	structure solution	4.1

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 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	B	1.21	0/246 (0.0%)	1.89	5/377 (1.3%)
2	C	1.22	0/252 (0.0%)	1.90	7/388 (1.8%)
3	A	0.68	0/526 (0.0%)	1.19	4/706 (0.6%)
All	All	0.98	0/1024 (0.0%)	1.60	16/1471 (1.1%)

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	B	0	1
2	C	0	4
3	A	0	3
All	All	0	8

There are no bond-length outliers.

5 of 16 angle outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	147	ARG	NE-CZ-NH1	8.16	124.38	120.30
3	A	150	ARG	NE-CZ-NH2	-6.72	116.94	120.30
2	C	16	DT	O4'-C4'-C3'	6.67	110.00	106.00
3	A	150	ARG	NE-CZ-NH1	6.62	123.61	120.30
1	B	6	DC	O4'-C4'-C3'	6.22	109.73	106.00

There are no chirality outliers.

5 of 8 planar outliers are listed below.

Mol	Chain	Res	Type	Group
3	A	170	ARG	Sidechain
1	B	7	DC	Sidechain

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
2	C	22	DC	Sidechain
3	A	186	TYR	Sidechain
3	A	147	ARG	Sidechain

## 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	220	123	124	1
2	C	225	124	125	1
All	All	958	761	760	1

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 1.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)
1:B:4:DA:C2	2:C:24:DA:C2	0.46	3.04

## 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	61/63 (97%)	57 (93%)	4 (7%)	0 (0%)	100	100
All	All	61/63 (97%)	57 (93%)	4 (7%)	0 (0%)	100	100

There are no Ramachandran outliers.



### 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	47/47 (100%)	45 (96%)	2 (4%)	36	82
All	All	47/47 (100%)	45 (96%)	2 (4%)	36	82

All 2 residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
3	A	191	PHE
3	A	152	ARG

### 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 ⓘ

There are no chain breaks in this entry.

## 7 Chemical shift validation

No chemical shift data were provided