



# wwPDB NMR Structure Validation Summary Report ⓘ

May 28, 2020 – 10:00 pm BST

PDB ID : 2IXQ  
Title : The solution structure of the invasive tip complex from Afa-Dr fibrils  
Authors : Cota, E.; Jones, C.; Simpson, P.; Altroff, H.; Anderson, K.L.; du Merle, L.;  
Guignot, J.; Servin, A.; Le Bouguenec, C.; Mardon, H.; Matthews, S.  
Deposited on : 2006-07-10

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 : 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 : 2.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

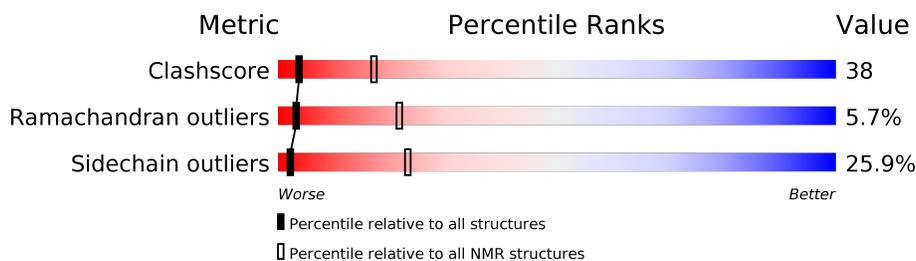
# 1 Overall quality at a glance

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	A	142	
2	B	143	

## 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 2 unique types of molecules in this entry. The entry contains 4179 atoms, of which 2016 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Protein AfaD.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	142	2061	653	986	205	213	4	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	SER	ALA	conflict	UNP Q47038
A	123	ASP	-	expression tag	UNP Q47038
A	124	ASN	-	expression tag	UNP Q47038
A	125	LYS	-	expression tag	UNP Q47038
A	126	GLN	-	expression tag	UNP Q47038
A	127	GLY	-	expression tag	UNP Q47038
A	128	PHE	-	expression tag	UNP Q47038
A	129	THR	-	expression tag	UNP Q47038
A	130	PRO	-	expression tag	UNP Q47038
A	131	SER	-	expression tag	UNP Q47038
A	132	GLY	-	expression tag	UNP Q47038
A	133	THR	-	expression tag	UNP Q47038
A	134	THR	-	expression tag	UNP Q47038
A	135	GLY	-	expression tag	UNP Q47038
A	136	THR	-	expression tag	UNP Q47038
A	137	THR	-	expression tag	UNP Q47038
A	138	LYS	-	expression tag	UNP Q47038
A	139	LEU	-	expression tag	UNP Q47038
A	140	THR	-	expression tag	UNP Q47038
A	141	VAL	-	expression tag	UNP Q47038
A	142	THR	-	expression tag	UNP Q47038

- Molecule 2 is a protein called Afimbrial adhesin AFA-III.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	143	2118	677	1030	188	220	3	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	124	ASP	-	expression tag	UNP Q57254
B	125	ASN	-	expression tag	UNP Q57254
B	126	LYS	-	expression tag	UNP Q57254
B	127	GLN	-	expression tag	UNP Q57254
B	128	GLY	-	expression tag	UNP Q57254
B	129	PHE	-	expression tag	UNP Q57254
B	130	THR	-	expression tag	UNP Q57254
B	131	PRO	-	expression tag	UNP Q57254
B	132	SER	-	expression tag	UNP Q57254
B	133	GLY	-	expression tag	UNP Q57254
B	134	THR	-	expression tag	UNP Q57254
B	135	THR	-	expression tag	UNP Q57254
B	136	GLY	-	expression tag	UNP Q57254
B	137	THR	-	expression tag	UNP Q57254
B	138	THR	-	expression tag	UNP Q57254
B	139	LYS	-	expression tag	UNP Q57254
B	140	LEU	-	expression tag	UNP Q57254
B	141	THR	-	expression tag	UNP Q57254
B	142	VAL	-	expression tag	UNP Q57254
B	143	THR	-	expression tag	UNP Q57254

## 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.

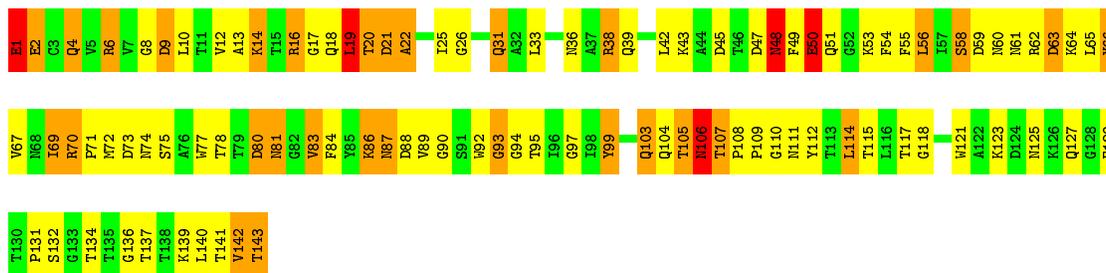
### • Molecule 1: Protein AfaD

Chain A: 



### • Molecule 2: Afimbrial adhesin AFA-III

Chain B: 



## 5 Refinement protocol and experimental data overview [i](#)

Of the ? 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
CNS	refinement	

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

COVALENT-GEOMETRY INFOmissingINFO

### 5.1 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	A	1075	986	989	97
2	B	1088	1030	1023	68
All	All	2163	2016	2012	160

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 38.

5 of 160 clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)
2:B:19:LEU:CB	2:B:19:LEU:CD1	1.58	1.80
2:B:4:GLN:NE2	2:B:4:GLN:CG	1.48	1.69
1:A:141:VAL:CA	1:A:141:VAL:CG1	1.41	1.97
1:A:138:LYS:NZ	1:A:138:LYS:CD	1.39	1.82
1:A:141:VAL:CG2	1:A:141:VAL:CA	1.38	1.98

### 5.2 Torsion angles [i](#)

#### 5.2.1 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	
1	A	140/142 (99%)	98 (70%)	34 (24%)	8 (6%)	3	22
2	B	141/143 (99%)	112 (79%)	21 (15%)	8 (6%)	3	22
All	All	281/285 (99%)	210 (75%)	55 (20%)	16 (6%)	3	22

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

Mol	Chain	Res	Type
2	B	103	GLN
1	A	124	ASN
1	A	120	PRO
2	B	19	LEU
1	A	102	GLN

### 5.2.2 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	
1	A	111/111 (100%)	80 (72%)	31 (28%)	2	19
2	B	117/117 (100%)	89 (76%)	28 (24%)	2	27
All	All	228/228 (100%)	169 (74%)	59 (26%)	2	23

5 of 59 residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
2	B	137	THR
1	A	142	THR
2	B	115	THR
1	A	16	LEU
2	B	58	SER

### 5.2.3 RNA [i](#)

There are no RNA molecules in this entry.

### 5.3 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 5.4 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 5.5 Ligand geometry [i](#)

There are no ligands in this entry.

### 5.6 Other polymers [i](#)

There are no such molecules in this entry.

### 5.7 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	31
2	B	13

The worst 5 of 44 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	97:MET	C	98:ALA	N	1.20
1	A	24:THR	C	25:GLY	N	1.19
1	A	83:VAL	C	84:SER	N	1.19
1	A	88:GLN	C	89:GLU	N	1.19
1	A	134:THR	C	135:GLY	N	1.19

## 6 Chemical shift validation

No chemical shift data were provided