W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

Full wwPDB NMR Structure Validation Report (i)

Mar 3, 2022 – 07:20 AM EST

PDB ID	:	2HD7
Title	:	Solution structure of C-teminal domain of twinfilin-1.
Authors	:	Hellman, M.H.; Paavilainen, V.O.; Annila, A.; Lappalainen, P.; Permi, P.I.
Deposited on	:	2006-06-20

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:

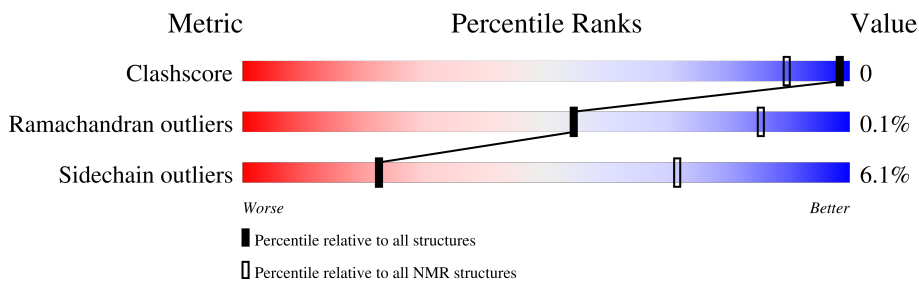|		
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (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	Percentile Ranks	Value
Clashscore		0
Ramachandran outliers		0.1%
Sidechain outliers		6.1%

Worse   *Better*

▉ Percentile relative to all structures

▯ Percentile relative to all NMR structures

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 >=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	142	88% 6% 6%

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

2   Ensemble composition and analysis ⓘ

This entry contains 15 models. Model 7 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *fewest violations*.

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:180-A:313 (134)	0.24	7

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 4 clusters and 6 single-model clusters were found.

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

3 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 2335 atoms, of which 1165 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Twinfilin-1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	142	2335	745	1165	196	224	5	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	175	ALA	-	cloning artifact	UNP Q91YR1

4 Residue-property plots ⓘ

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.
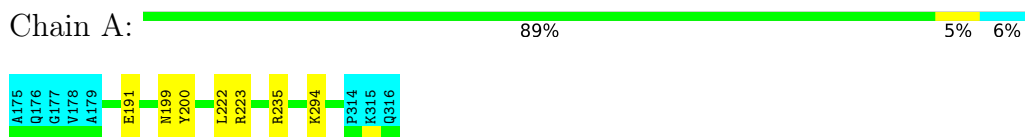
• Molecule 1: Twinfilin-1
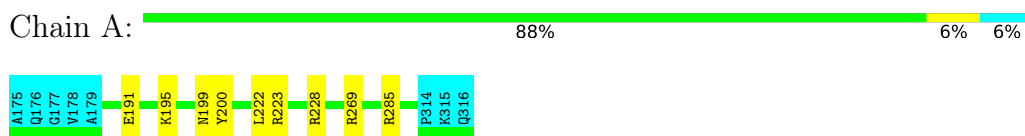
4.2 Scores per residue for each member of the ensemble
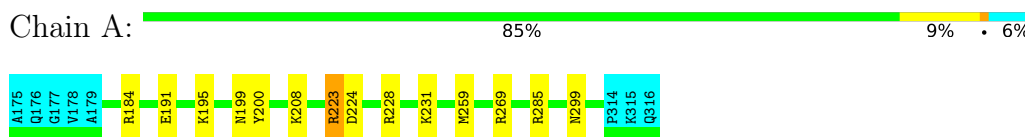
Colouring as in section 4.1 above.

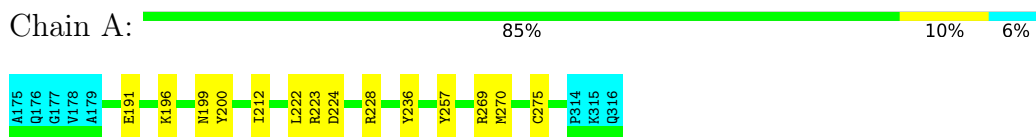### 4.2.1 Score per residue for model 1

• Molecule 1: Twinfilin-1

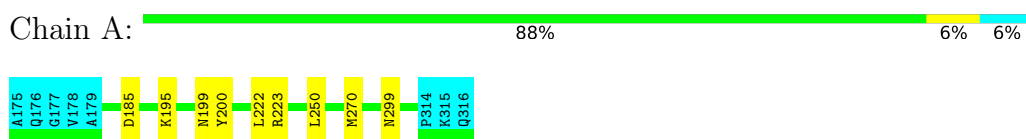
4.2.2 Score per residue for model 2

• Molecule 1: Twinfilin-1

4.2.3 Score per residue for model 3
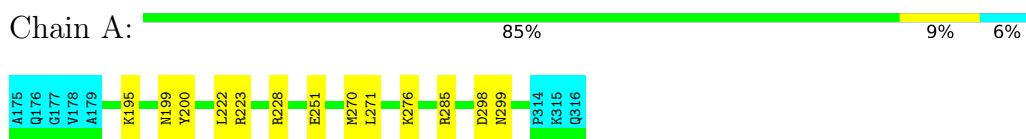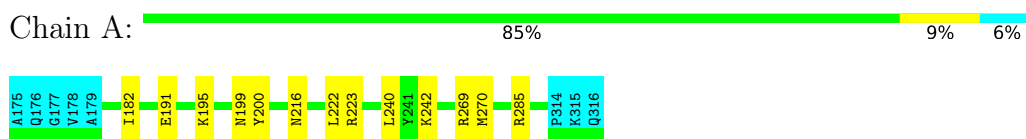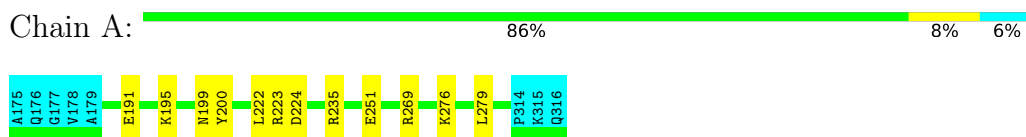
• Molecule 1: Twinfilin-1

Chain A: 85% 10% 6%

A175 Q176 G177 V178 A179 E191 K196 N199 Y200 I212 L222 R223 D224 R228 Y236 Y257 R269 M270 C275 P314 K315 Q316

4.2.4 Score per residue for model 4

• Molecule 1: Twinfilin-1

Chain A: 88% 6% 6%

A175 Q176 G177 V178 A179 D185 K195 N199 Y200 L222 R223 L250 M270 N299 P314 K315 Q316

4.2.5 Score per residue for model 5

• Molecule 1: Twinfilin-1

Chain A: 85% 9% 6%

A175 Q176 G177 V178 A179 K195 N199 Y200 L222 R223 R228 E251 M270 L271 K276 R285 D298 N299 P314 K315 Q316

4.2.6 Score per residue for model 6

• Molecule 1: Twinfilin-1

Chain A: 85% 9% 6%

A175 Q176 G177 V178 A179 I182 E191 K195 N199 Y200 N216 L222 R223 L240 Y241 K242 R269 M270 R285 P314 K315 Q316

4.2.7 Score per residue for model 7 (medoid)

• Molecule 1: Twinfilin-1

Chain A: 86% 8% 6%

A175 Q176 G177 V178 A179 E191 K195 N199 Y200 L222 R223 D224 R235 E251 R269 K276 L279 P314 K315 Q316

4.2.8 Score per residue for model 8

• Molecule 1: Twinfilin-1

Chain A:

87%   7% · 6%

A175 Q176 G177 V178 A179 K195 N199 Y200 L222 R223 R228 E251 R269 M270 R285 N299 P314 K315 Q316
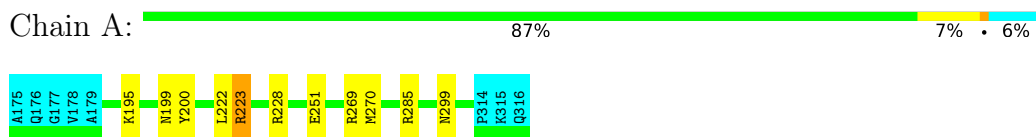
4.2.9 Score per residue for model 9

• Molecule 1: Twinfilin-1

Chain A:

85%   8% · 6%

A175 Q176 G177 V178 A179 E191 K195 N199 Y200 I212 R223 D224 R228 R235 K242 R269 M270 C275 P314 K315 Q316
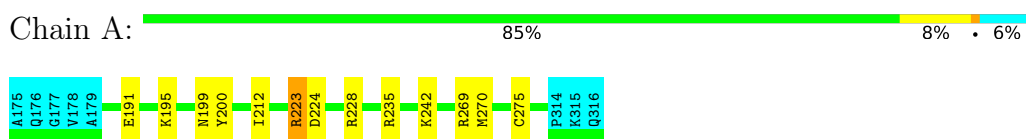
4.2.10 Score per residue for model 10

• Molecule 1: Twinfilin-1

Chain A:

87%   7%   6%

A175 Q176 G177 V178 A179 K195 N199 Y200 L222 R223 R228 R235 K242 R269 R285 P314 K315 Q316
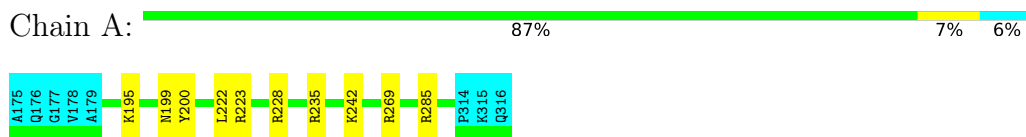
4.2.11 Score per residue for model 11

• Molecule 1: Twinfilin-1

Chain A:

87%   8%   6%

A175 Q176 G177 V178 A179 I182 E191 K195 N199 Y200 I212 L222 R223 R228 R269 R285 P314 K315 Q316
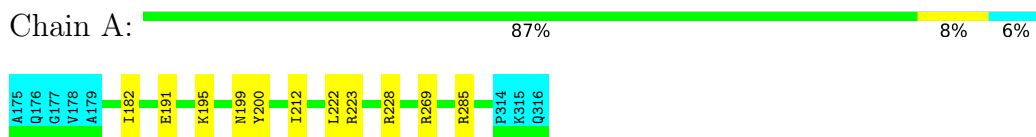
4.2.12 Score per residue for model 12

• Molecule 1: Twinfilin-1

Chain A:

89%   · · 6%

A175 Q176 G177 V178 A179 E191 K195 N199 Y200 L222 R223 D224 P314 K315 Q316
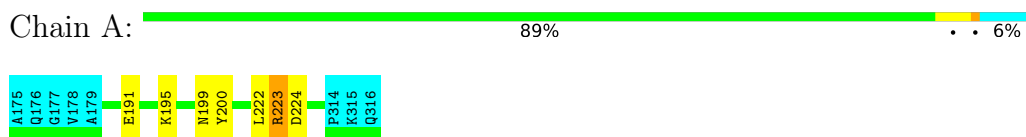
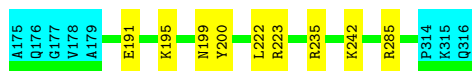#### 4.2.13　Score per residue for model 13

• Molecule 1: Twinfilin-1

Chain A:

4.2.14　Score per residue for model 14

• Molecule 1: Twinfilin-1

Chain A:

4.2.15　Score per residue for model 15

• Molecule 1: Twinfilin-1

Chain A:
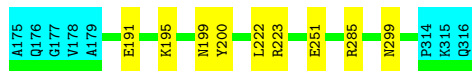
5 Refinement protocol and experimental data overview ⓘ

The models were refined using the following method: *torsion angle dynamics, simulated annealing.*

Of the 40 calculated structures, 15 were deposited, based on the following criterion: *structures with the least restraint violations.*

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

Software name	Classification	Version
CYANA	structure solution	2.1
Amber	refinement	8

No chemical shift data was provided.

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	A	0.74±0.00	0±0/1137 (0.0± 0.0%)	0.97±0.02	3±1/1536 (0.2± 0.1%)
All	All	0.74	0/17055 (0.0%)	0.97	39/23040 (0.2%)

There are no bond-length outliers.

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(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	223	ARG	NE-CZ-NH1	7.52	124.06	120.30	3	11
1	A	228	ARG	NE-CZ-NH1	7.09	123.84	120.30	9	8
1	A	235	ARG	NE-CZ-NH1	6.66	123.63	120.30	13	6
1	A	269	ARG	NE-CZ-NH1	6.40	123.50	120.30	8	5
1	A	285	ARG	NE-CZ-NH1	6.23	123.42	120.30	15	8
1	A	184	ARG	NE-CZ-NH1	5.05	122.83	120.30	2	1

There are no chirality outliers.

There are no planarity outliers.

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	1114	1105	1105	0±1
All	All	16710	16575	16575	5

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 0.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Models Total
1:A:212:ILE:HD11	1:A:275:CYS:HB3	0.54	1.80	9	2
1:A:212:ILE:HD12	1:A:212:ILE:H	0.45	1.70	11	2
1:A:236:TYR:CE1	1:A:257:TYR:CD1	0.42	3.08	3	1

6.3 Torsion angles ⓘ

6.3.1 Protein backbone ⓘ

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	134/142 (94%)	125±2 (93±1%)	9±2 (7±1%)	0±0 (0±0%)	54	85
All	All	2010/2130 (94%)	1874 (93%)	134 (7%)	2 (0%)	54	85

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	182	ILE	2

6.3.2 Protein sidechains ⓘ

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	126/131 (96%)	118±2 (94±2%)	8±2 (6±2%)	22	71
All	All	1890/1965 (96%)	1774 (94%)	116 (6%)	22	71

All 25 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)
1	A	199	ASN	15

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	200	TYR	15
1	A	222	LEU	13
1	A	195	LYS	13
1	A	191	GLU	10
1	A	223	ARG	9
1	A	270	MET	6
1	A	224	ASP	5
1	A	299	ASN	5
1	A	251	GLU	4
1	A	242	LYS	4
1	A	269	ARG	3
1	A	276	LYS	2
1	A	294	LYS	1
1	A	208	LYS	1
1	A	231	LYS	1
1	A	259	MET	1
1	A	196	LYS	1
1	A	185	ASP	1
1	A	250	LEU	1
1	A	271	LEU	1
1	A	298	ASP	1
1	A	216	ASN	1
1	A	240	LEU	1
1	A	279	LEU	1

6.3.3 RNA ⓘ

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.6 Ligand geometry ⓘ

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