wwPDB NMR Structure Validation Summary Report (i)

Feb 9, 2022 – 06:23 AM EST

PDB ID	:	1DVD
Title	:	SOLUTION NMR STRUCTURE OF HUMAN STEFIN A AT PH 5.5 AND 308K, NMR, 17 STRUCTURES
Authors	:	Martin, J.R.; Craven, C.J.; Jerala, R.; Kroon-Zitko, L.; Zerovnik, E.; Turk, V.; Waltho, J.P.
Deposited on	:	1996-02-26

This is a wwPDB NMR Structure Validation Summary 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:

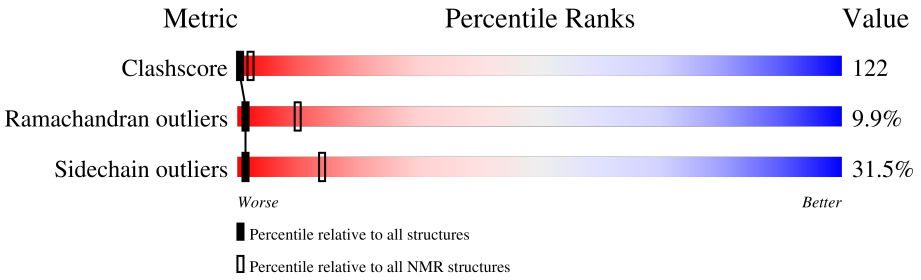
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.26
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.26

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

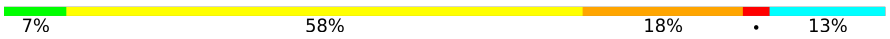
The overall completeness of chemical shifts assignment was not calculated.

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 >=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	98	

2 Ensemble composition and analysis (i)

This entry contains 17 models. Model 10 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:11-A:40, A:43-A:68, A:92-A:101, A:108-A:125 (84)	0.31	10

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 1 single-model cluster was found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1555 atoms, of which 780 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called STEFIN A.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	98	1555	492	780	126	155	2	0

4 Residue-property plots (i)

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: STEFIN A

Chain A: 7% 58% 18% • 13%

M6 I7 P8 G9 G10 L11 S12 E13 A14 K15 P16 A17 T18 P19 E20 I21 Q22 E23 I24 V25 D26 K27 V28 K29 P30 Q31 L32 E33 E34 K35 T36 N37 E38 T39 Y40 G43 K44 L45 E46 A47 V48 Q49 Y50 K51 T52 Q53 V54 V55 A56 G57 T58 N59 Y60 Y61 I62 K63 V64 R65 A66 G67

D68 N92 K93 Y94 M95 H96 L97 K98 V99 F100 K101 S102 L102A P103 G104 Q105 N105A E106 D107 L108 V109 L110 T111 G112 Y113 Q114 V115 D115A K116 N117 K118 D119 D120 E121 L122 T123 G124 F125

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

The representative model is number 10. Colouring as in section 4.1 above.

- Molecule 1: STEFIN A

Chain A: 8% 50% 23% 5% 13%

M6 I7 P8 G9 G10 L11 S12 E13 A14 K15 P16 A17 T18 P19 E20 I21 Q22 E23 I24 V25 D26 K27 V28 K29 P30 Q31 L32 E33 E34 K35 T36 N37 E38 T39 Y40 G43 K44 L45 E46 A47 V48 Q49 Y50 K51 T52 Q53 V54 V55 T58 N59 Y60 Y61 I62 K63 V64 R65 D68

Y94 M95 H96 L97 K98 V99 F100 K101 S102 L102A P103 G104 Q105 N105A E106 D107 L108 V109 L110 T111 G112 Y113 Q114 V115 D115A K116 N117 K118 D119 D120 E121 L122 T123 G124 F125

5 Refinement protocol and experimental data overview (i)

Of the ? calculated structures, 17 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

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.

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	A	0.0±0.0	0.9±0.2
All	All	0	16

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

All unique planar outliers are listed below.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	65	ARG	Sidechain	16

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	A	685	695	695	169±9
All	All	11645	11815	11815	2869

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 122.

5 of 792 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:24:ILE:HG23	1:A:99:VAL:HG21	1.04	1.29	4	17
1:A:32:LEU:HD11	1:A:97:LEU:CD1	1.01	1.83	11	12

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:32:LEU:HD11	1:A:97:LEU:HD21	0.99	1.30	8	3
1:A:32:LEU:HD11	1:A:97:LEU:HD13	0.96	1.38	7	7
1:A:32:LEU:O	1:A:36:THR:HG22	0.95	1.61	2	8

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
1	A	84/98 (86%)	61±2 (73±3%)	14±2 (17±3%)	8±2 (10±2%)	1 10
All	All	1428/1666 (86%)	1044 (73%)	243 (17%)	141 (10%)	1 10

5 of 18 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	55	VAL	17
1	A	113	TYR	17
1	A	13	GLU	15
1	A	68	ASP	12
1	A	108	LEU	12

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
1	A	75/85 (88%)	51±4 (69±5%)	24±4 (31±5%)	1 14
All	All	1275/1445 (88%)	874 (69%)	401 (31%)	1 14

5 of 56 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	18	THR	17
1	A	55	VAL	17
1	A	123	THR	17
1	A	125	PHE	16
1	A	54	VAL	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 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 (i)

No chemical shift data were provided