wwPDB NMR Structure Validation Summary Report (i)

Jun 12, 2024 – 12:56 PM EDT

PDB ID	:	2VIL
Title	:	REFINED STRUCTURE OF THE ACTIN-SEVERING DOMAIN VILLIN 14T, DETERMINED BY SOLUTION NMR, 11 STRUCTURES
Authors	:	Markus, M.A.; Matsudaira, P.; Wagner, G.
Deposited on	:	1997-01-16

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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

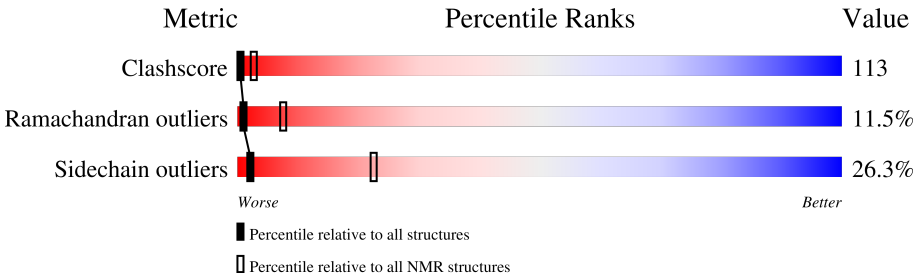
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

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

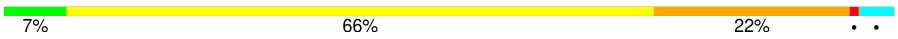
The overall completeness of chemical shifts assignment was not calculated.

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	126	

2 Ensemble composition and analysis (i)

This entry contains 11 models. Model 2 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:1-A:121 (121)	0.84	2

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 1 clusters and 4 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1968 atoms, of which 971 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called VILLIN 14T.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	126	1968	631	971	168	193	5	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: VILLIN 14T

Chain A: 7% 66% 22%

V1 E2 L3 S4 K5 K6 V7 T8 G9 K10 L11 D12 K13 T14 T15 I18 Q19 I20 W21 R22 I23 E24 N25 M26 E27 M28 V29 P30 V31 P32 T33 K34 S35 Y36 G37 N38 F39 Y40 E41 G42 D43 C44 Y45 V46 L47 L48 S49 T50 R51 K52 T53 F57 S58 Y59 N60 I61 H62 Y63

W64 L65 G66 K67 N68 S69 S70 Q71 D72 E73 Q74 G75 A76 A77 A78 I79 Y80 T81 T82 Q83 M84 D85 E86 Y87 L88 G89 S90 V91 A92 V93 Q94 H95 R96 E97 V98 Q99 G100 H101 E102 S103 E104 T105 F106 R107 A108 Y109 F110 K111 Q112 G113 L114 I115 Y116 K117 Q118 G119 G120 V121 A122 S123

G124 M125 K126

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

The representative model is number 2. Colouring as in section 4.1 above.

- Molecule 1: VILLIN 14T

Chain A: 13% 59% 22%

V1 E2 L3 S4 K5 K6 V7 T8 G9 K10 L11 D12 K13 T14 G17 I18 Q19 I20 W21 R22 I23 E24 N25 M26 E27 M28 V29 P30 V31 P32 T33 K34 S35 Y36 G37 N38 F39 Y40 E41 G42 D43 C44 Y45 V46 L47 L48 S49 T50 R51 K52 T53 F57 S58 Y59 N60 I61 H62 Y63

W64 L65 G66 K67 N68 S69 S70 Q71 D72 E73 Q74 G75 I79 Y80 T81 M84 D85 E86 Y87 L88 G89 S90 V91 A92 V93 Q94 H95 R96 E97 V98 Q99 G100 H101 E102 S103 E104 T105 F106 R107 A108 Y109 F110 K111 Q112 G113 L114 I115 Y116 K117 Q118 G119 G120 V121 A122 S123 G124 M125 K126

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *distance geometry*.

Of the 20 calculated structures, 11 were deposited, based on the following criterion: *NOE VIOLATIONS < 0.5 A DIHEDRAL VIOL < 5 DEGREES*.

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

Software name	Classification	Version
X-PLOR	refinement	3.1
DGII	structure solution	
IN INSIGHT II	structure solution	II

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	4.0±0.0
All	All	0	44

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

All unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	22	ARG	Sidechain	11
1	A	51	ARG	Sidechain	11
1	A	96	ARG	Sidechain	11
1	A	107	ARG	Sidechain	11

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	964	936	936	214±14
All	All	10604	10296	10296	2353

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 113.

5 of 1060 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:11:LEU:HD12	1:A:88:LEU:HD21	1.14	1.17	7	4
1:A:18:ILE:HD12	1:A:47:LEU:HD11	1.09	1.16	8	5
1:A:11:LEU:HD11	1:A:88:LEU:HD11	1.09	1.14	6	1
1:A:3:LEU:HD13	1:A:4:SER:N	1.01	1.71	11	3
1:A:3:LEU:HD21	1:A:80:TYR:CG	1.00	1.91	1	1

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	120/126 (95%)	83±3 (69±2%)	24±3 (20±2%)	14±3 (12±2%)	1 7
All	All	1320/1386 (95%)	908 (69%)	260 (20%)	152 (12%)	1 7

5 of 39 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	66	GLY	11
1	A	67	LYS	11
1	A	68	ASN	10
1	A	69	SER	10
1	A	103	SER	9

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	103/106 (97%)	76±3 (74±3%)	27±3 (26±3%)	2 22
All	All	1133/1166 (97%)	835 (74%)	298 (26%)	2 22

5 of 71 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	33	THR	11
1	A	106	PHE	11
1	A	29	VAL	10
1	A	39	PHE	10
1	A	65	LEU	10

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