wwPDB NMR Structure Validation Summary Report (i)

Apr 26, 2016 – 02:17 PM BST

PDB ID	:	1BPV
Title	:	TITIN MODULE A71 FROM HUMAN CARDIAC MUSCLE, NMR, 50 STRUCTURES
Authors	:	Muhle-Goll, C.; Pastore, A.; Nilges, M.
Deposited on	:	1998-08-11

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
http://wwpdb.org/validation/2016/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:

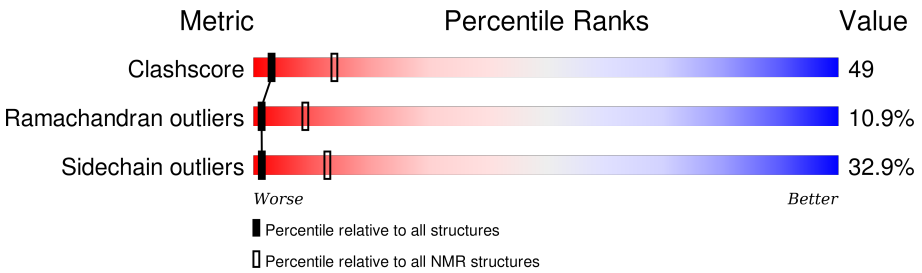
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027457

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	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

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	112	17% 56% 15% • 7%

2 Ensemble composition and analysis (i)

This entry contains 50 models. Model 49 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative.

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:2-A:100 (99)	0.37	49

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 6 clusters and 15 single-model clusters were found.

Cluster number	Models
1	2, 3, 5, 6, 7, 8, 12, 15, 16, 18, 21, 31, 33, 34, 36, 37, 38, 42, 44, 45, 49
2	9, 13, 25, 27, 29, 43
3	1, 48
4	47, 50
5	11, 39
6	23, 46
Single-model clusters	4; 10; 14; 17; 19; 20; 22; 24; 26; 28; 30; 32; 35; 40; 41

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1603 atoms, of which 796 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called TITIN.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	104	1603	513	796	136	157	1	0

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

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: TITIN

Chain A: 17% 56% 15% • 7%

MET HIS HIS HIS HIS HIS HIS SER S0 P1 I2 D3 P4 P5 G6 K7 P8 V9 P10 L11 N12 I13 T14 R15 H16 T17 V18 T19 L20 K21 W22 A23 K24 P25 T28 G29 G30 F31 K32 I33 T34 S35 Y36 I37 V38 E39 K40 R41 D42 L43 G46 R47 W48 L49 K50 A51 N52 F53

S54 N55 I56 L57 E58 N59 E60 F61 T62 V63 S64 G65 L66 T67 E68 D69 A70 A71 Y72 E73 F74 R75 V76 I77 A78 K79 N80 A81 A82 G83 A84 I85 S86 P87 P88 S89 E90 P91 S92 D93 A94 I95 T96 C97 R98 D99 D100 V101 E102 A103

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

The representative model is number 49. Colouring as in section 4.1 above.

- Molecule 1: TITIN

Chain A: 29% 41% 15% • • 7%

MET HIS HIS HIS HIS HIS HIS SER S0 P1 I2 D3 P4 P5 P8 V9 P10 L11 N12 I13 T14 R15 V18 T19 L20 K21 W22 A23 K24 P25 E26 Y27 T28 G29 G30 F31 K32 I33 Y36 I37 V38 E39 K40 R41 D42 L43 R47 W48 L49 K50 A51 N52 F53 S54 N55 I56

N59 E60 F61 T62 V63 S64 G65 L66 T67 E68 D69 A70 A71 Y72 E73 F74 R75 V76 I77 A78 K79 N80 I85 S86 P87 P88 S89 E90 P91 S92 D93 A94 I95 T96 C97 R98 D99 D100 V101 E102 A103

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *SIMULATED ANNEALING*.

Of the 1 calculated structures, 50 were deposited, based on the following criterion: *LEAST RESTRAINT VIOLATION*.

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

Software name	Classification	Version
XPLOR 3.851 (EXTENDED VERSION)	refinement	VERSION)
XPLOR (EXTENDED VERSION)	structure solution	VERSION)

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)

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	772	764	764	75±10
All	All	38600	38200	38195	3739

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 49.

5 of 902 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:14:THR:CG2	1:A:100:ASP:OD1	1.44	1.65	38	2
1:A:21:LYS:CG	1:A:60:GLU:OE1	1.44	1.66	31	7
1:A:42:ASP:OD1	1:A:47:ARG:CG	1.37	1.73	48	5
1:A:21:LYS:HB3	1:A:60:GLU:OE1	1.36	1.14	2	3
1:A:3:ASP:OD2	1:A:27:TYR:CB	1.34	1.74	27	6

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	99/112 (88%)	69±3 (69±3%)	20±3 (20±3%)	11±2 (11±2%)	1 9
All	All	4950/5600 (88%)	3432 (69%)	976 (20%)	542 (11%)	1 9

5 of 52 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	33	ILE	43
1	A	29	GLY	39
1	A	25	PRO	37
1	A	87	PRO	33
1	A	90	GLU	32

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	84/96 (88%)	56±4 (67±4%)	28±4 (33±4%)	1 12
All	All	4200/4800 (88%)	2817 (67%)	1383 (33%)	1 12

5 of 65 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	20	LEU	50
1	A	14	THR	50
1	A	47	ARG	49
1	A	42	ASP	49
1	A	24	LYS	47

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 (i)

There are no chain breaks in this entry.

7 Chemical shift validation (i)

No chemical shift data were provided