Full wwPDB NMR Structure Validation Report (i)

Feb 27, 2022 – 01:49 PM EST

PDB ID	:	2D7O
Title	:	Solution structure of the 17th Filamin domain from human Filamin C
Authors	:	Tomizawa, T.; Kigawa, T.; Koshiba, S.; Inoue, M.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2005-11-24

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:

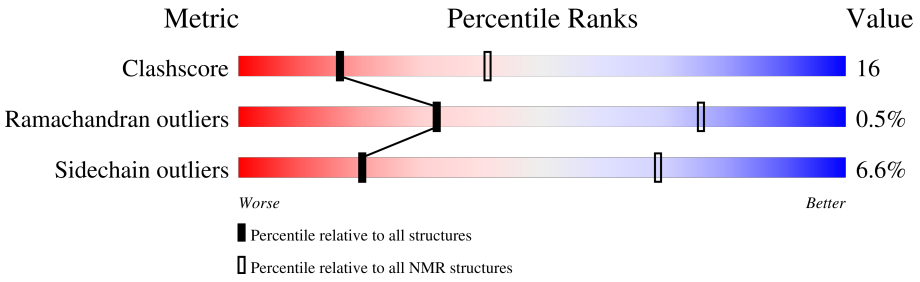
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

1 Overall quality at a glance (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	111	54% 16% • 29%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 12 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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:15-A:37, A:44-A:99 (79)	0.17	12

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	1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 17, 18, 19
2	16, 20
Single-model clusters	2

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1553 atoms, of which 766 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Filamin-C.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	111	1553	483	766	136	164	4	0

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	cloning artifact	UNP Q14315
A	2	SER	-	cloning artifact	UNP Q14315
A	3	SER	-	cloning artifact	UNP Q14315
A	4	GLY	-	cloning artifact	UNP Q14315
A	5	SER	-	cloning artifact	UNP Q14315
A	6	SER	-	cloning artifact	UNP Q14315
A	7	GLY	-	cloning artifact	UNP Q14315
A	106	SER	-	cloning artifact	UNP Q14315
A	107	GLY	-	cloning artifact	UNP Q14315
A	108	PRO	-	cloning artifact	UNP Q14315
A	109	SER	-	cloning artifact	UNP Q14315
A	110	SER	-	cloning artifact	UNP Q14315
A	111	GLY	-	cloning artifact	UNP Q14315

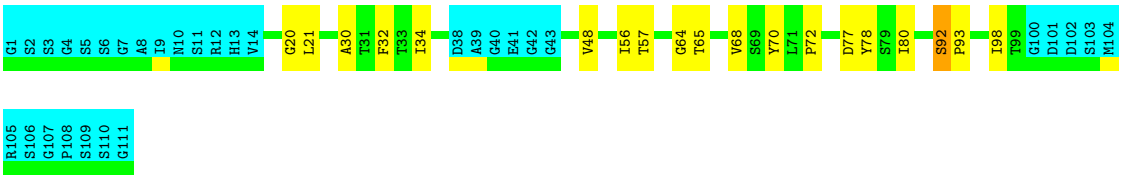
4 Residue-property plots (i)

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: Filamin-C

Chain A: 54% 16% • 29%



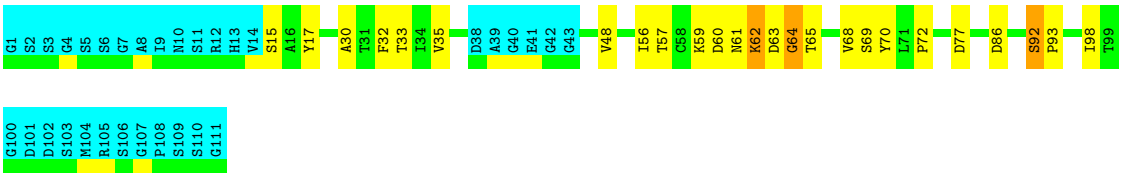
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Filamin-C

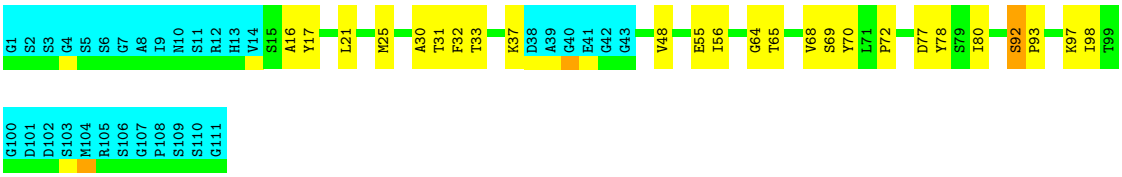
Chain A: 49% 20% • 29%



4.2.2 Score per residue for model 2

• Molecule 1: Filamin-C

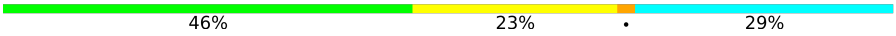
Chain A: 49% 22% • 29%

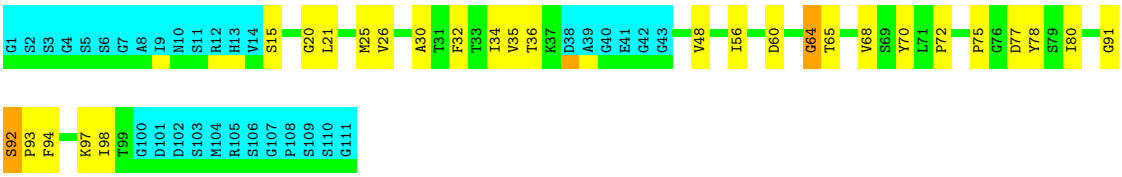


4.2.3 Score per residue for model 3

- Molecule 1: Filamin-C

Chain A:





4.2.4 Score per residue for model 4

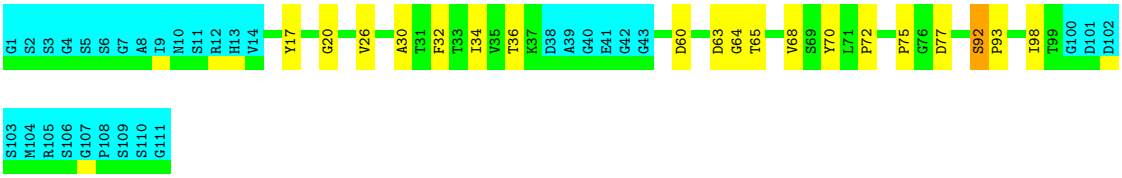
- Molecule 1: Filamin-C

Chain A:

4.2.5 Score per residue for model 5

- Molecule 1: Filamin-C

Chain A:



4.2.6 Score per residue for model 6

- Molecule 1: Filamin-C

Chain A: 49% 22% • 29%

G1 S2 S3 G4 S5 S6 G7 A8 I9 N10 S11 R12 H13 V14 G20 V26 A30 T31 F32 T33 I34 V35 T36 K37 D38 A39 G40 E41 G42 G43 L46 A47 V48 I56 T57 D60 N61 K62 D63 V68 S69 Y70 G76 D77 V82 R83 F84 S92 P93

K97 I98 T99 G100 D101 D102 S103 M104 R105 S106 G107 P108 S109 S110 G111

4.2.7 Score per residue for model 7

- Molecule 1: Filamin-C

Chain A: 48% 22% •• 29%

G1 S2 S3 G4 S5 S6 G7 A8 I9 N10 S11 R12 H13 V14 Y17 G20 L21 G24 A30 T31 F32 D38 A39 G40 E41 G42 G43 L46 I56 D60 N61 K62 D63 G64 T65 V68 S69 Y70 L71 P72 D77 Y78 S79 I80 I81 V82 R83 F84 S92

P93 K97 I98 T99 G100 D101 D102 S103 M104 R105 S106 G107 P108 S109 S110 G111

4.2.8 Score per residue for model 8

- Molecule 1: Filamin-C

Chain A: 57% 14% • 29%

G1 S2 S3 G4 S5 S6 G7 A8 I9 N10 S11 R12 H13 V14 G20 L21 K28 P29 A30 T31 F32 V35 D38 A39 G40 E41 G42 G43 V48 I56 G64 T65 V68 S69 Y70 I80 S92 P93 I98 T99 G100 D101 D102 S103 M104 R105 S106 G107 P108

S109 S110 G111

4.2.9 Score per residue for model 9

- Molecule 1: Filamin-C

Chain A: 56% 14% • 29%

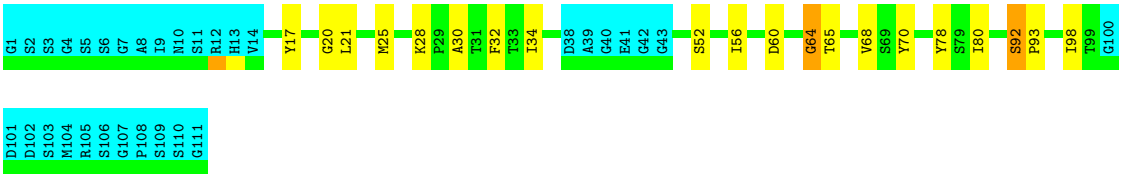
G1 S2 S3 G4 S5 S6 G7 A8 I9 N10 S11 R12 H13 V14 L21 M25 A30 T31 F32 T33 I34 D38 A39 G40 E41 G42 G43 V48 I56 T57 V68 S69 Y70 L71 P72 D77 Y78 S79 I80 S92 P93 I98 T99 G100 D101 D102 S103 M104 R105 S106

G107 P108 S109 S110 G111

4.2.10 Score per residue for model 10

- Molecule 1: Filamin-C

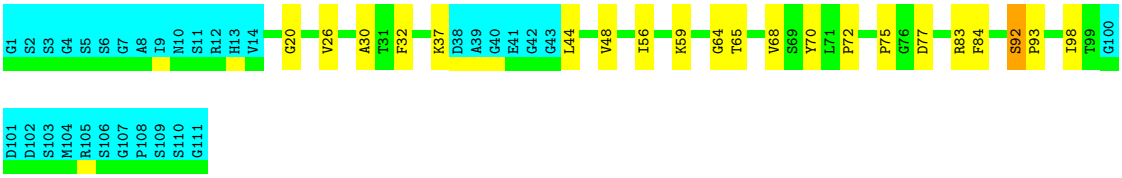
Chain A: 53% 16% • 29%



4.2.11 Score per residue for model 11

- Molecule 1: Filamin-C

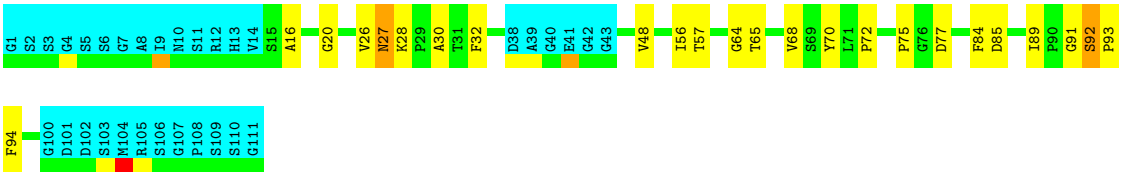
Chain A: 52% 18% • 29%



4.2.12 Score per residue for model 12 (medoid)

- Molecule 1: Filamin-C

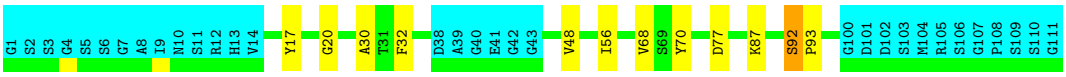
Chain A: 50% 20% • 29%



4.2.13 Score per residue for model 13

- Molecule 1: Filamin-C

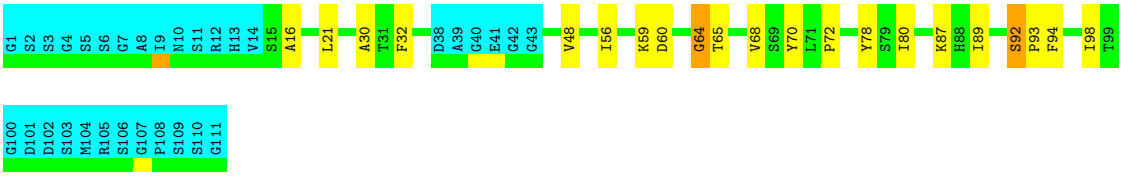
Chain A: 60% 10% • 29%



4.2.14 Score per residue for model 14

- Molecule 1: Filamin-C

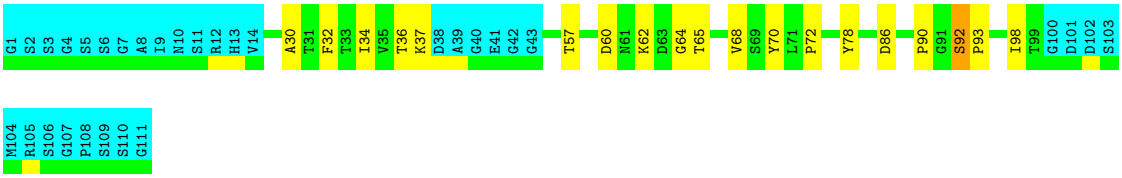
Chain A: 52% 17% • 29%



4.2.15 Score per residue for model 15

- Molecule 1: Filamin-C

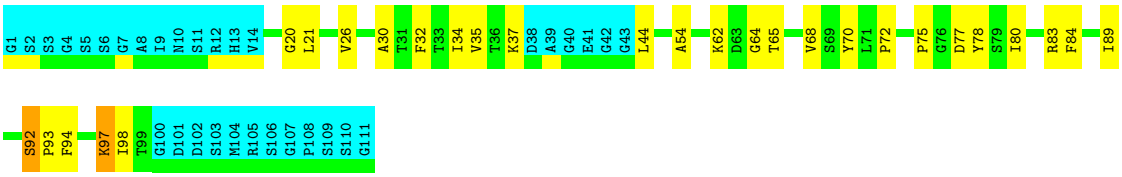
Chain A: 54% 16% • 29%



4.2.16 Score per residue for model 16

- Molecule 1: Filamin-C

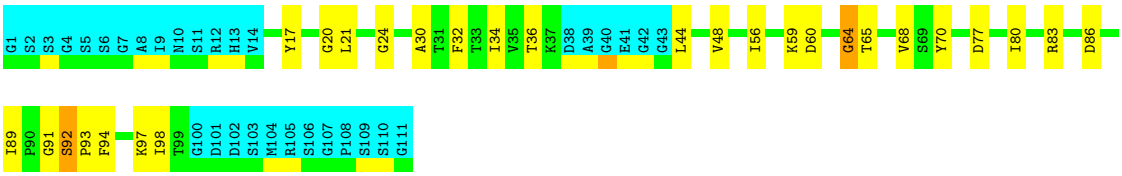
Chain A: 46% 23% • 29%



4.2.17 Score per residue for model 17

- Molecule 1: Filamin-C

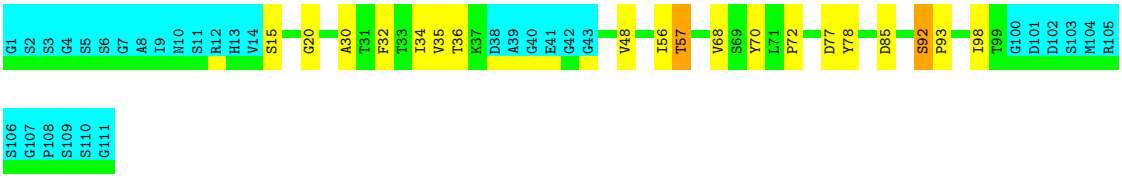
Chain A: 46% 23% • 29%



4.2.18 Score per residue for model 18

- Molecule 1: Filamin-C

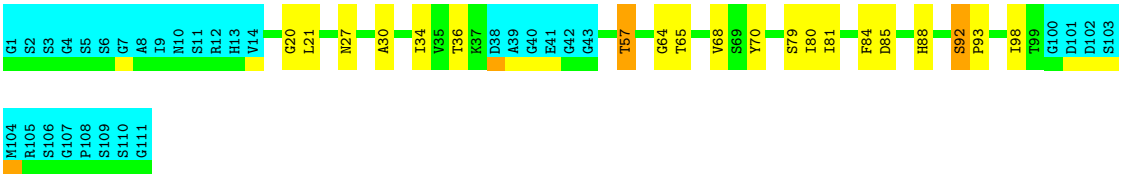
Chain A: 54% 15% • 29%



4.2.19 Score per residue for model 19

- Molecule 1: Filamin-C

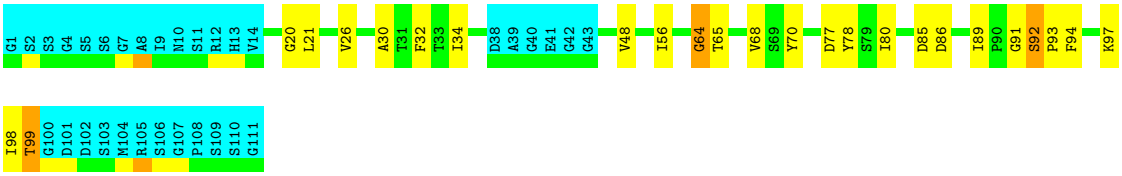
Chain A: 53% 16% • 29%



4.2.20 Score per residue for model 20

- Molecule 1: Filamin-C

Chain A: 49% 20% • 29%



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *torsion angle dynamics, restrainted molecular dynamics.*

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *target function,structures with the lowest energy,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	2.0.17
CYANA	refinement	2.0.17

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.

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	584	590	588	18±4
All	All	11680	11800	11760	365

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 16.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:70:TYR:OH	1:A:98:ILE:HD11	0.82	1.75	9	17
1:A:20:GLY:HA3	1:A:30:ALA:HB1	0.79	1.52	18	15
1:A:30:ALA:HB3	1:A:70:TYR:CE2	0.78	2.14	20	19
1:A:26:VAL:HG23	1:A:99:THR:O	0.75	1.82	20	2
1:A:48:VAL:HG23	1:A:56:ILE:HD11	0.73	1.58	4	14
1:A:20:GLY:CA	1:A:30:ALA:HB1	0.72	2.14	8	14
1:A:32:PHE:CE2	1:A:34:ILE:HD11	0.69	2.23	20	6
1:A:48:VAL:HG23	1:A:56:ILE:CD1	0.64	2.22	4	6
1:A:21:LEU:HD23	1:A:80:ILE:HD13	0.62	1.72	20	10
1:A:57:THR:O	1:A:68:VAL:HG13	0.60	1.96	18	7
1:A:30:ALA:CB	1:A:70:TYR:CE2	0.60	2.84	5	11
1:A:34:ILE:HG22	1:A:36:THR:HG23	0.58	1.74	17	7
1:A:16:ALA:HB1	1:A:21:LEU:HD11	0.57	1.74	2	2
1:A:91:GLY:O	1:A:94:PHE:CD1	0.57	2.57	20	5

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:70:TYR:O	1:A:70:TYR:CD1	0.57	2.58	13	6
1:A:89:ILE:HD11	1:A:94:PHE:HE1	0.56	1.60	14	2
1:A:46:LEU:HD23	1:A:82:VAL:HG22	0.56	1.77	7	2
1:A:56:ILE:CG2	1:A:68:VAL:CG1	0.54	2.85	20	13
1:A:21:LEU:CD2	1:A:80:ILE:HD13	0.54	2.33	10	7
1:A:17:TYR:CD1	1:A:17:TYR:O	0.54	2.61	17	4
1:A:70:TYR:CD1	1:A:70:TYR:O	0.54	2.61	19	12
1:A:91:GLY:O	1:A:94:PHE:CE1	0.53	2.61	20	2
1:A:70:TYR:C	1:A:70:TYR:CD1	0.52	2.82	14	3
1:A:81:ILE:HG23	1:A:88:HIS:CE1	0.52	2.39	19	1
1:A:44:LEU:HD12	1:A:83:ARG:O	0.52	2.04	17	3
1:A:30:ALA:HB3	1:A:70:TYR:CD2	0.51	2.39	5	4
1:A:78:TYR:CD1	1:A:98:ILE:HG13	0.51	2.40	18	8
1:A:70:TYR:CD1	1:A:70:TYR:C	0.51	2.84	9	10
1:A:61:ASN:O	1:A:63:ASP:N	0.50	2.45	1	2
1:A:21:LEU:HD11	1:A:32:PHE:CD2	0.50	2.41	8	2
1:A:17:TYR:CE1	1:A:33:THR:CG2	0.50	2.94	1	2
1:A:34:ILE:HD12	1:A:68:VAL:CG2	0.50	2.37	20	1
1:A:31:THR:HG22	1:A:69:SER:OG	0.49	2.07	2	2
1:A:78:TYR:CD2	1:A:98:ILE:HG13	0.49	2.42	2	1
1:A:89:ILE:HD11	1:A:94:PHE:HE2	0.49	1.67	17	3
1:A:15:SER:O	1:A:35:VAL:HG23	0.48	2.08	18	3
1:A:32:PHE:CZ	1:A:68:VAL:HB	0.48	2.44	9	19
1:A:26:VAL:HG21	1:A:75:PRO:HD3	0.48	1.85	3	6
1:A:30:ALA:HB3	1:A:70:TYR:CZ	0.47	2.44	10	4
1:A:64:GLY:O	1:A:65:THR:CG2	0.47	2.63	2	15
1:A:70:TYR:CE1	1:A:72:PRO:HD3	0.46	2.45	1	3
1:A:54:ALA:HB2	1:A:72:PRO:HA	0.46	1.87	4	2
1:A:92:SER:HA	1:A:93:PRO:C	0.46	2.31	6	20
1:A:89:ILE:HD11	1:A:94:PHE:CE2	0.46	2.45	17	3
1:A:70:TYR:CD1	1:A:72:PRO:HD3	0.46	2.46	3	10
1:A:30:ALA:CB	1:A:70:TYR:CZ	0.46	2.99	6	4
1:A:84:PHE:CE2	1:A:85:ASP:OD2	0.46	2.69	19	1
1:A:16:ALA:HB1	1:A:21:LEU:CD1	0.46	2.40	14	2
1:A:72:PRO:HB3	1:A:78:TYR:CZ	0.46	2.45	15	6
1:A:17:TYR:O	1:A:17:TYR:CD1	0.46	2.69	5	1
1:A:24:GLY:N	1:A:97:LYS:O	0.46	2.49	17	2
1:A:15:SER:O	1:A:35:VAL:CG2	0.46	2.64	18	1
1:A:89:ILE:HD11	1:A:94:PHE:CE1	0.45	2.45	14	1
1:A:20:GLY:HA2	1:A:30:ALA:HB1	0.44	1.87	8	1
1:A:17:TYR:CD1	1:A:17:TYR:C	0.44	2.91	10	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:64:GLY:O	1:A:65:THR:HG23	0.44	2.12	7	10
1:A:61:ASN:O	1:A:62:LYS:C	0.43	2.54	7	2
1:A:78:TYR:CE1	1:A:98:ILE:HG13	0.43	2.48	10	2
1:A:78:TYR:CE2	1:A:98:ILE:HG13	0.43	2.48	2	1
1:A:78:TYR:CD1	1:A:98:ILE:CG1	0.43	3.02	14	3
1:A:62:LYS:N	1:A:62:LYS:HD2	0.43	2.29	7	1
1:A:85:ASP:C	1:A:86:ASP:OD1	0.43	2.57	20	1
1:A:37:LYS:N	1:A:64:GLY:O	0.43	2.52	4	1
1:A:17:TYR:O	1:A:17:TYR:CG	0.43	2.71	7	2
1:A:61:ASN:CB	1:A:63:ASP:OD1	0.42	2.68	6	1
1:A:21:LEU:HD23	1:A:80:ILE:CD1	0.42	2.43	14	2
1:A:16:ALA:HB3	1:A:94:PHE:CE2	0.42	2.49	12	1
1:A:61:ASN:C	1:A:63:ASP:N	0.42	2.72	1	1
1:A:26:VAL:C	1:A:27:ASN:OD1	0.41	2.59	12	1
1:A:48:VAL:CG2	1:A:56:ILE:HD11	0.41	2.40	4	1
1:A:97:LYS:HE3	1:A:98:ILE:N	0.41	2.31	16	1
1:A:56:ILE:HG22	1:A:68:VAL:CG1	0.41	2.45	8	1
1:A:17:TYR:CZ	1:A:33:THR:HG21	0.41	2.51	2	1
1:A:62:LYS:N	1:A:62:LYS:CD	0.40	2.83	7	1
1:A:76:GLY:N	1:A:98:ILE:O	0.40	2.54	6	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	79/111 (71%)	71±2 (90±3%)	8±2 (10±2%)	0±1 (1±1%)	32	76
All	All	1580/2220 (71%)	1418 (90%)	154 (10%)	8 (1%)	32	76

All 2 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	64	GLY	6
1	A	62	LYS	2

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	66/87 (76%)	62±1 (93±2%)	4±1 (7±2%)	20	69
All	All	1320/1740 (76%)	1233 (93%)	87 (7%)	20	69

All 20 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	92	SER	20
1	A	77	ASP	15
1	A	60	ASP	9
1	A	97	LYS	5
1	A	59	LYS	4
1	A	25	MET	4
1	A	37	LYS	4
1	A	62	LYS	4
1	A	57	THR	4
1	A	86	ASP	3
1	A	28	LYS	3
1	A	63	ASP	2
1	A	27	ASN	2
1	A	87	LYS	2
1	A	69	SER	1
1	A	55	GLU	1
1	A	52	SER	1
1	A	85	ASP	1
1	A	79	SER	1
1	A	99	THR	1

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