Full wwPDB NMR Structure Validation Report (i)

Feb 27, 2022 – 10:48 AM EST

PDB ID	:	2CUI
Title	:	Solution structure of the 31st fibronectin type III domain of the human tenascin X
Authors	:	Ohnishi, S.; Kigawa, T.; Tochio, N.; Koshiba, S.; Inoue, M.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2005-05-26

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:

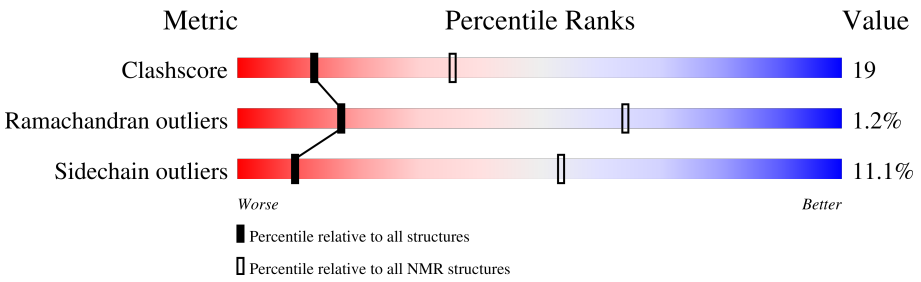
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
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	112	48% 24% • 27%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 5 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:11-A:46, A:62-A:107 (82)	0.13	5

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 11 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1686 atoms, of which 847 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Tenascin-X.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	112	1686	521	847	157	160	1	0

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	cloning artifact	UNP P22105
A	2	SER	-	cloning artifact	UNP P22105
A	3	SER	-	cloning artifact	UNP P22105
A	4	GLY	-	cloning artifact	UNP P22105
A	5	SER	-	cloning artifact	UNP P22105
A	6	SER	-	cloning artifact	UNP P22105
A	7	GLY	-	cloning artifact	UNP P22105
A	107	SER	-	cloning artifact	UNP P22105
A	108	GLY	-	cloning artifact	UNP P22105
A	109	PRO	-	cloning artifact	UNP P22105
A	110	SER	-	cloning artifact	UNP P22105
A	111	SER	-	cloning artifact	UNP P22105
A	112	GLY	-	cloning artifact	UNP P22105

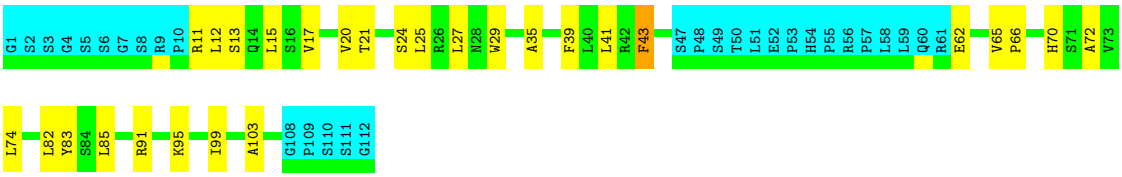
4 Residue-property plots (i)

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: Tenascin-X

Chain A: 48% 24% • 27%



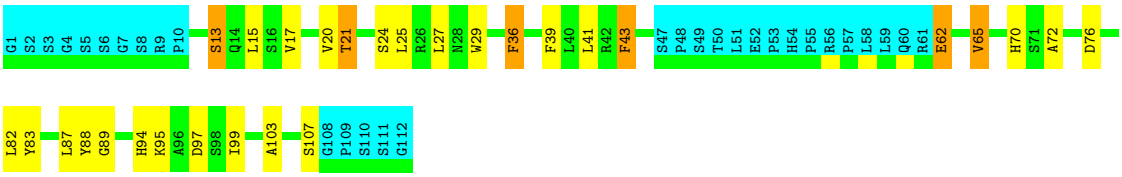
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Tenascin-X

Chain A: 47% 21% 5% 27%



4.2.2 Score per residue for model 2

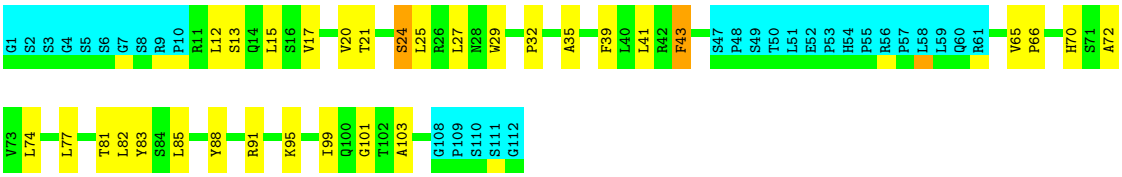
• Molecule 1: Tenascin-X

Chain A: 47% 22% • 27%

4.2.3 Score per residue for model 3

- Molecule 1: Tenascin-X

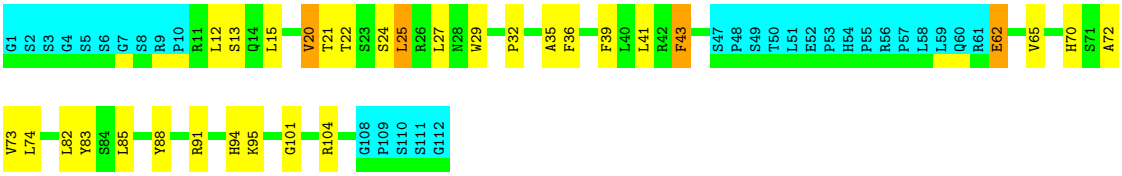
Chain A: 46% 26% • 27%



4.2.4 Score per residue for model 4

- Molecule 1: Tenascin-X

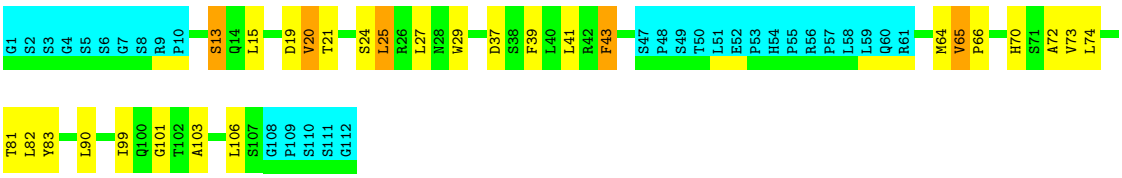
Chain A: 46% 24% • 27%



4.2.5 Score per residue for model 5 (medoid)

- Molecule 1: Tenascin-X

Chain A: 48% 21% • 27%



4.2.6 Score per residue for model 6

- Molecule 1: Tenascin-X

Chain A: 51% 17% 5% 27%

G1 S2 S3 G4 S5 S6 G7 S8 R9 P10 R11 L15 S16 V17 V20 T21 S24 L25 R26 L27 N28 W29 F39 L40 L41 R42 F43 G44 V45 P46 S47 P48 S49 T50 L51 E52 P53 H54 P55 R56 P57 L58 L59 Q60 R61 E62 L63 M64 V65 H70 L74 R75 D76 L77

R78 T81 L82 Y83 S84 L85 H94 A103 G108 P109 S110 S111 G112

4.2.7 Score per residue for model 7

- Molecule 1: Tenascin-X

Chain A: 49% 18% 6% 27%

G1 S2 S3 G4 S5 S6 G7 S8 R9 P10 R11 L12 S13 Q14 L15 S16 V17 V20 T21 S24 L25 R26 L27 N28 W29 D37 S38 F39 L40 L41 R42 F43 S47 P48 S49 T50 L51 E52 P53 H54 P55 R56 P57 L58 L59 Q60 R61 V65 P66 H70 S71 A72 V73 L74

L82 Y83 S84 L85 R91 H94 K95 A96 G101 T102 A103 G108 P109 S110 S111 G112

4.2.8 Score per residue for model 8

- Molecule 1: Tenascin-X

Chain A: 45% 25% • 27%

G1 S2 S3 G4 S5 S6 G7 S8 R9 P10 R11 L12 S13 V17 T18 D19 V20 T21 S24 L25 W29 F39 L40 L41 R42 F43 S47 P48 S49 T50 L51 E52 P53 H54 P55 R56 P57 L58 L59 Q60 R61 E62 L63 M64 V65 P66 H70 S71 A72 V73 L74 R75 D76 L77

R78 S79 G80 T81 L82 Y83 Y88 R91 K95 I99 A103 G108 P109 S110 S111 G112

4.2.9 Score per residue for model 9

- Molecule 1: Tenascin-X

Chain A: 48% 21% • 27%

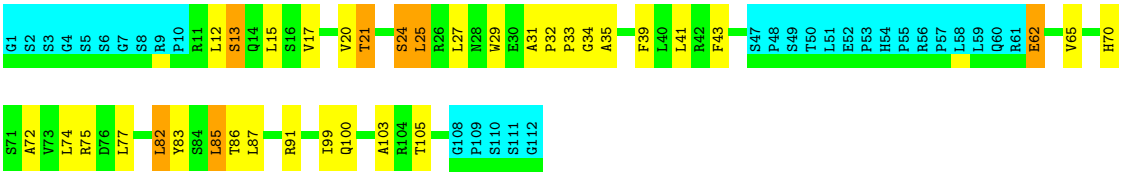
G1 S2 S3 G4 S5 S6 G7 S8 R9 P10 R11 L12 S13 V17 V20 T21 S24 L25 R26 L27 N28 W29 D37 S38 F39 F43 S47 P48 S49 T50 L51 E52 P53 H54 P55 R56 P57 L58 L59 Q60 R61 E62 V65 P66 H70 L77 R78 L82 Y83 S84

L85 T86 L87 Y88 R91 K95 A96 I99 A103 G108 P109 S110 S111 G112

4.2.10 Score per residue for model 10

- Molecule 1: Tenascin-X

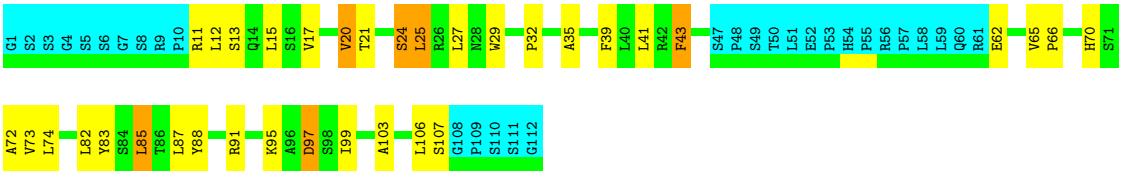
Chain A: 42% 25% 6% 27%



4.2.11 Score per residue for model 11

- Molecule 1: Tenascin-X

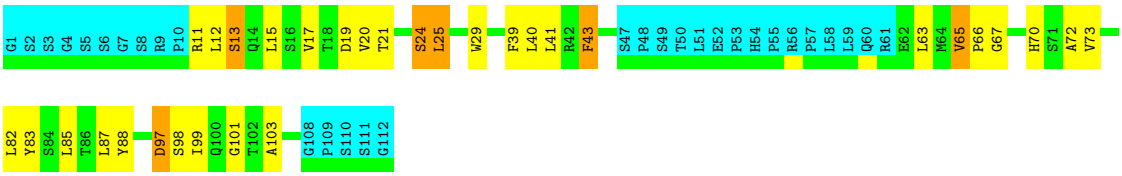
Chain A: 42% 26% 5% 27%



4.2.12 Score per residue for model 12

- Molecule 1: Tenascin-X

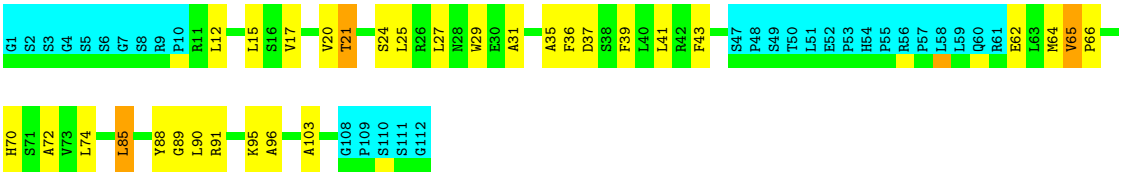
Chain A: 45% 23% 5% 27%



4.2.13 Score per residue for model 13

- Molecule 1: Tenascin-X

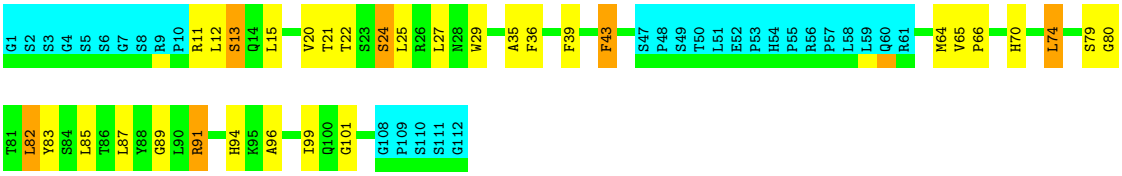
Chain A: 46% 25% • 27%



4.2.14 Score per residue for model 14

- Molecule 1: Tenascin-X

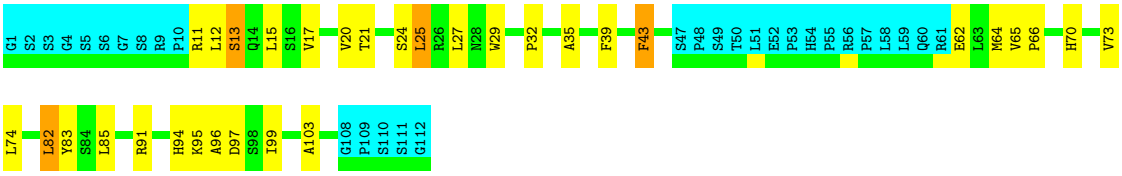
Chain A: 45% 23% 5% 27%



4.2.15 Score per residue for model 15

- Molecule 1: Tenascin-X

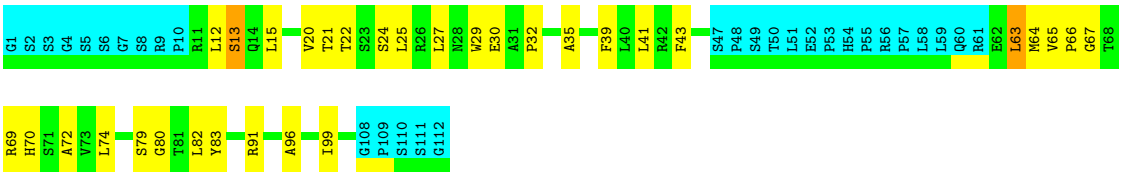
Chain A: 45% 25% • 27%



4.2.16 Score per residue for model 16

- Molecule 1: Tenascin-X

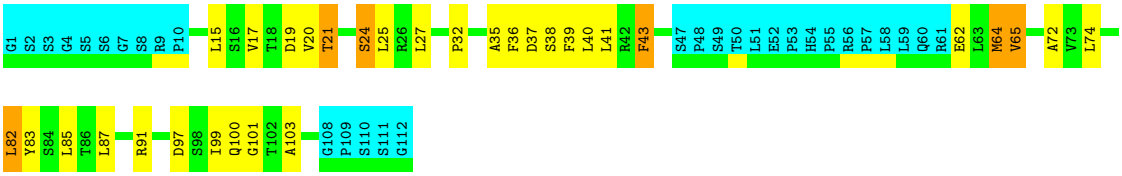
Chain A: 45% 27% • 27%



4.2.17 Score per residue for model 17

- Molecule 1: Tenascin-X

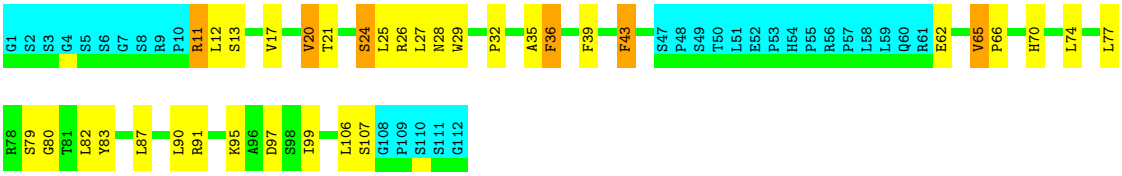
Chain A: 45% 23% 5% 27%



4.2.18 Score per residue for model 18

- Molecule 1: Tenascin-X

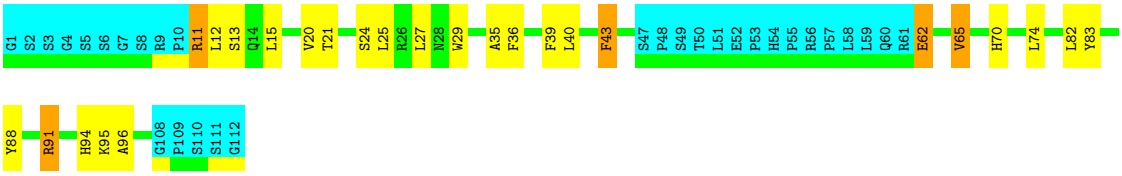
Chain A: 42% 26% 5% 27%



4.2.19 Score per residue for model 19

- Molecule 1: Tenascin-X

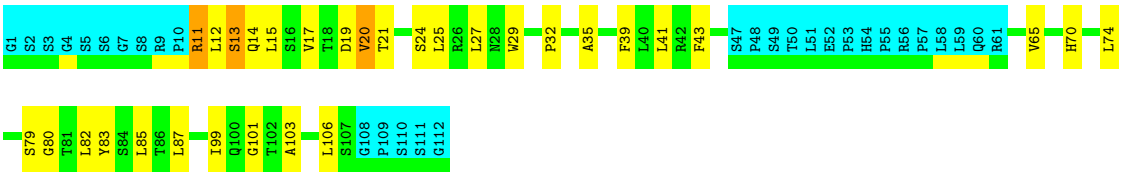
Chain A: 50% 19% • 27%



4.2.20 Score per residue for model 20

- Molecule 1: Tenascin-X

Chain A: 46% 25% • 27%



5 Refinement protocol and experimental data overview (i)

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

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *target function, 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	631	644	642	24±3
All	All	12620	12880	12840	476

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 19.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:15:LEU:HD11	1:A:27:LEU:HD22	0.91	1.38	10	16
1:A:17:VAL:HG21	1:A:103:ALA:HB2	0.88	1.43	20	14
1:A:43:PHE:CE1	1:A:74:LEU:HD12	0.83	2.09	2	4
1:A:15:LEU:CD1	1:A:27:LEU:HD22	0.78	2.09	10	14
1:A:82:LEU:HD13	1:A:83:TYR:N	0.74	1.98	3	19
1:A:40:LEU:HD11	1:A:62:GLU:HG3	0.70	1.63	17	1
1:A:17:VAL:HG22	1:A:27:LEU:CD2	0.70	2.16	7	3
1:A:91:ARG:HG2	1:A:96:ALA:HB2	0.69	1.65	2	1
1:A:17:VAL:HG21	1:A:103:ALA:CB	0.69	2.18	20	13
1:A:27:LEU:HD21	1:A:85:LEU:HD13	0.67	1.67	3	7
1:A:39:PHE:CD1	1:A:87:LEU:HD21	0.65	2.26	12	1
1:A:41:LEU:HD22	1:A:72:ALA:HB1	0.64	1.69	11	11
1:A:63:LEU:HD21	1:A:65:VAL:CG2	0.63	2.22	16	2
1:A:12:LEU:HD21	1:A:39:PHE:CE1	0.63	2.29	18	15

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:43:PHE:CE1	1:A:74:LEU:CD1	0.62	2.82	2	4
1:A:12:LEU:HD21	1:A:39:PHE:CZ	0.62	2.28	14	4
1:A:41:LEU:HD21	1:A:74:LEU:HD21	0.62	1.71	11	6
1:A:91:ARG:HB2	1:A:96:ALA:HB2	0.62	1.70	19	7
1:A:25:LEU:O	1:A:73:VAL:HG13	0.61	1.95	11	6
1:A:15:LEU:HD23	1:A:101:GLY:HA3	0.61	1.70	20	5
1:A:17:VAL:CG2	1:A:103:ALA:HB2	0.60	2.27	2	10
1:A:29:TRP:CE2	1:A:87:LEU:HD21	0.59	2.32	18	2
1:A:15:LEU:HD11	1:A:27:LEU:CD2	0.59	2.24	13	2
1:A:43:PHE:N	1:A:43:PHE:CD1	0.59	2.70	18	16
1:A:17:VAL:HG11	1:A:25:LEU:HD22	0.58	1.75	10	3
1:A:74:LEU:HD22	1:A:74:LEU:N	0.58	2.14	19	10
1:A:43:PHE:CZ	1:A:74:LEU:HD12	0.58	2.34	17	4
1:A:27:LEU:HD11	1:A:85:LEU:HD13	0.57	1.77	13	3
1:A:17:VAL:HG22	1:A:27:LEU:HD23	0.57	1.76	7	2
1:A:82:LEU:HD13	1:A:82:LEU:C	0.56	2.21	15	16
1:A:32:PRO:HG2	1:A:35:ALA:HB2	0.56	1.77	2	10
1:A:87:LEU:HB3	1:A:99:ILE:HG22	0.55	1.78	14	4
1:A:36:PHE:CE2	1:A:97:ASP:OD1	0.55	2.59	17	1
1:A:82:LEU:C	1:A:82:LEU:HD13	0.54	2.22	18	1
1:A:36:PHE:CZ	1:A:97:ASP:OD2	0.53	2.61	1	2
1:A:41:LEU:HD22	1:A:72:ALA:CB	0.53	2.33	7	9
1:A:62:GLU:N	1:A:62:GLU:CD	0.53	2.62	4	3
1:A:82:LEU:C	1:A:82:LEU:CD1	0.52	2.78	15	12
1:A:39:PHE:CD1	1:A:87:LEU:CD2	0.52	2.93	12	1
1:A:85:LEU:HD12	1:A:101:GLY:O	0.52	2.04	7	3
1:A:91:ARG:O	1:A:91:ARG:NE	0.51	2.43	7	2
1:A:35:ALA:O	1:A:91:ARG:CD	0.50	2.60	3	2
1:A:13:SER:O	1:A:99:ILE:CD1	0.50	2.60	12	12
1:A:21:THR:N	1:A:24:SER:O	0.50	2.44	12	20
1:A:13:SER:O	1:A:99:ILE:HD13	0.50	2.07	3	5
1:A:88:TYR:CD1	1:A:95:LYS:HD3	0.50	2.42	3	9
1:A:41:LEU:HD11	1:A:85:LEU:HB3	0.50	1.84	13	1
1:A:39:PHE:N	1:A:65:VAL:O	0.49	2.45	16	16
1:A:29:TRP:CZ2	1:A:70:HIS:HA	0.49	2.42	18	15
1:A:15:LEU:CG	1:A:27:LEU:HD22	0.49	2.36	17	5
1:A:37:ASP:N	1:A:90:LEU:O	0.49	2.45	5	3
1:A:29:TRP:O	1:A:70:HIS:CB	0.49	2.60	14	6
1:A:74:LEU:HB3	1:A:77:LEU:HD21	0.49	1.82	8	2
1:A:43:PHE:CD1	1:A:74:LEU:HG	0.49	2.43	6	1
1:A:36:PHE:CD2	1:A:39:PHE:CZ	0.49	3.01	13	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:35:ALA:O	1:A:36:PHE:CD1	0.49	2.66	4	2
1:A:17:VAL:HG22	1:A:27:LEU:HD21	0.49	1.83	7	1
1:A:27:LEU:HD21	1:A:85:LEU:CD1	0.49	2.36	3	2
1:A:78:ARG:NH2	1:A:81:THR:HG21	0.48	2.23	6	1
1:A:41:LEU:HB2	1:A:87:LEU:HD12	0.48	1.85	20	3
1:A:39:PHE:CE1	1:A:89:GLY:HA2	0.48	2.43	14	2
1:A:20:VAL:O	1:A:106:LEU:HD12	0.47	2.10	11	4
1:A:90:LEU:HD22	1:A:95:LYS:HG2	0.47	1.86	18	1
1:A:64:MET:SD	1:A:65:VAL:N	0.47	2.88	17	1
1:A:29:TRP:CD1	1:A:70:HIS:HB3	0.47	2.45	4	18
1:A:20:VAL:O	1:A:106:LEU:CD1	0.46	2.63	20	2
1:A:74:LEU:N	1:A:74:LEU:CD1	0.46	2.77	14	1
1:A:65:VAL:HG12	1:A:66:PRO:HD2	0.46	1.87	9	12
1:A:40:LEU:HD12	1:A:63:LEU:O	0.46	2.11	12	1
1:A:41:LEU:HD23	1:A:74:LEU:HD11	0.46	1.86	6	1
1:A:67:GLY:O	1:A:70:HIS:CE1	0.45	2.69	12	1
1:A:66:PRO:O	1:A:67:GLY:C	0.45	2.55	16	1
1:A:97:ASP:OD1	1:A:97:ASP:C	0.45	2.55	1	1
1:A:24:SER:HA	1:A:77:LEU:HD12	0.45	1.89	3	3
1:A:67:GLY:O	1:A:70:HIS:NE2	0.44	2.50	16	1
1:A:12:LEU:HD12	1:A:97:ASP:OD2	0.44	2.13	11	2
1:A:27:LEU:C	1:A:28:ASN:ND2	0.44	2.70	18	1
1:A:40:LEU:HD11	1:A:62:GLU:CD	0.44	2.32	19	1
1:A:81:THR:CG2	1:A:82:LEU:N	0.44	2.80	5	2
1:A:17:VAL:HG13	1:A:26:ARG:O	0.44	2.13	18	1
1:A:31:ALA:HB3	1:A:70:HIS:NE2	0.44	2.27	13	2
1:A:81:THR:HG22	1:A:82:LEU:N	0.44	2.28	5	2
1:A:21:THR:HG22	1:A:22:THR:N	0.44	2.27	14	1
1:A:66:PRO:O	1:A:69:ARG:N	0.43	2.50	16	1
1:A:79:SER:O	1:A:80:GLY:C	0.43	2.56	8	6
1:A:76:ASP:N	1:A:76:ASP:OD1	0.43	2.52	6	1
1:A:39:PHE:CD1	1:A:89:GLY:HA2	0.43	2.49	14	1
1:A:11:ARG:HD2	1:A:11:ARG:N	0.42	2.30	19	3
1:A:89:GLY:N	1:A:97:ASP:OD1	0.42	2.51	1	1
1:A:25:LEU:CD1	1:A:74:LEU:HD23	0.42	2.44	8	1
1:A:20:VAL:HG11	1:A:104:ARG:O	0.42	2.15	4	1
1:A:62:GLU:C	1:A:62:GLU:OE1	0.42	2.58	19	1
1:A:35:ALA:HB1	1:A:36:PHE:CE1	0.42	2.50	14	1
1:A:88:TYR:CE1	1:A:98:SER:HB2	0.42	2.50	12	1
1:A:63:LEU:HD21	1:A:65:VAL:HG21	0.42	1.91	6	1
1:A:15:LEU:HD11	1:A:27:LEU:HB3	0.42	1.91	17	2

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:64:MET:C	1:A:65:VAL:CG2	0.41	2.88	16	1
1:A:35:ALA:C	1:A:36:PHE:CG	0.41	2.93	4	1
1:A:12:LEU:CG	1:A:97:ASP:OD2	0.41	2.68	11	1
1:A:78:ARG:N	1:A:83:TYR:OH	0.41	2.53	9	1
1:A:86:THR:OG1	1:A:100:GLN:NE2	0.41	2.53	10	1
1:A:29:TRP:CE2	1:A:70:HIS:HA	0.41	2.51	6	1
1:A:45:VAL:HG13	1:A:78:ARG:NH2	0.41	2.30	6	1
1:A:78:ARG:CZ	1:A:81:THR:HG21	0.41	2.46	6	1
1:A:99:ILE:HG13	1:A:100:GLN:N	0.41	2.30	17	1
1:A:85:LEU:CG	1:A:101:GLY:O	0.40	2.69	7	1
1:A:75:ARG:O	1:A:76:ASP:CG	0.40	2.59	8	1
1:A:33:PRO:O	1:A:34:GLY:C	0.40	2.59	10	1
1:A:75:ARG:O	1:A:77:LEU:HD23	0.40	2.17	10	1
1:A:76:ASP:CG	1:A:76:ASP:O	0.40	2.60	1	1
1:A:82:LEU:HD22	1:A:103:ALA:O	0.40	2.17	5	1
1:A:85:LEU:O	1:A:100:GLN:OE1	0.40	2.40	10	1
1:A:74:LEU:N	1:A:74:LEU:HD22	0.40	2.31	13	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	82/112 (73%)	74±2 (90±2%)	7±2 (9±2%)	1±0 (1±0%)	17	64
All	All	1640/2240 (73%)	1477 (90%)	143 (9%)	20 (1%)	17	64

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	20	VAL	20

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	70/95 (74%)	62±2 (89±3%)	8±2 (11±3%)	9 53
All	All	1400/1900 (74%)	1245 (89%)	155 (11%)	9 53

All 28 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	25	LEU	20
1	A	43	PHE	14
1	A	13	SER	10
1	A	65	VAL	10
1	A	21	THR	9
1	A	62	GLU	9
1	A	91	ARG	9
1	A	24	SER	9
1	A	85	LEU	8
1	A	11	ARG	8
1	A	94	HIS	7
1	A	82	LEU	7
1	A	64	MET	6
1	A	19	ASP	5
1	A	107	SER	3
1	A	22	THR	3
1	A	37	ASP	3
1	A	97	ASP	3
1	A	36	PHE	2
1	A	14	GLN	2
1	A	78	ARG	1
1	A	76	ASP	1
1	A	105	THR	1
1	A	74	LEU	1
1	A	95	LYS	1
1	A	30	GLU	1
1	A	63	LEU	1
1	A	38	SER	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