



Full wwPDB NMR Structure Validation Report ⓘ

Apr 26, 2016 – 03:40 PM BST

PDB ID : 1M39
Title : Solution structure of the C-terminal fragment (F86-I165) of the human centrin 2 in calcium saturated form
Authors : Matei, E.; Miron, S.; Blouquit, Y.; Duchambon, P.; Durussel, P.; Cox, J.A.; Craescu, C.T.
Deposited on : 2002-06-27

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
<http://wwpdb.org/validation/2016/NMRValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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

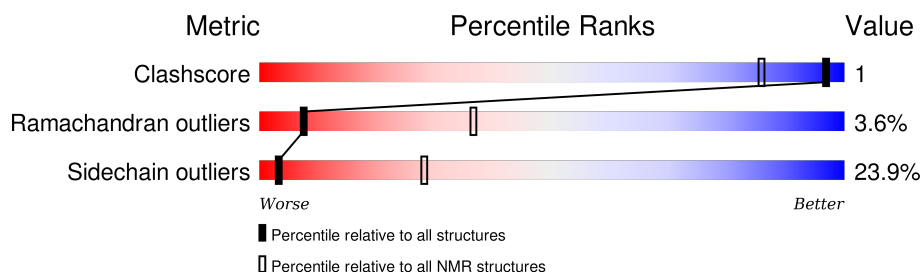
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 45%.

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	89	<div> <div>56%</div> <div>18%</div> <div>•</div> <div>12%</div> <div>10%</div> </div>

2 Ensemble composition and analysis

This entry contains 25 models. Model 2 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *closest to the average*.

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:86-A:91, A:103-A:165 (69)	0.65	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 6 clusters and 3 single-model clusters were found.

Cluster number	Models
1	2, 3, 5, 7, 10, 11, 20, 25
2	6, 9, 12, 16
3	13, 18, 22
4	4, 8, 23
5	15, 19
6	21, 24
Single-model clusters	1; 14; 17

3 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 1279 atoms, of which 630 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Caltractin, isoform 1.

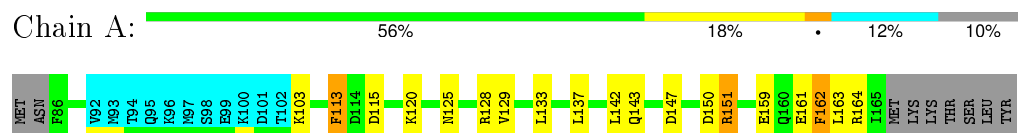
Mol	Chain	Residues	Atoms						Trace
1	A	80	Total	C	H	N	O	S	0
			1279	404	630	103	139	3	

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: Caltractin, isoform 1

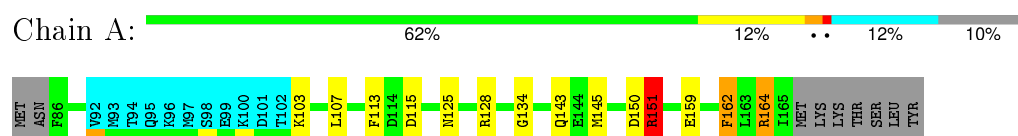


4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

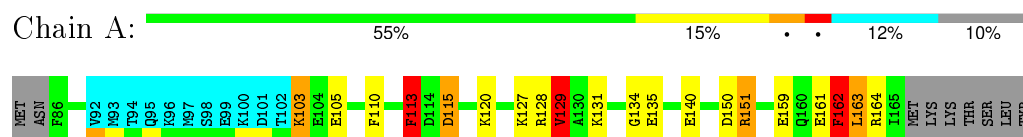
4.2.1 Score per residue for model 1

- Molecule 1: Caltractin, isoform 1



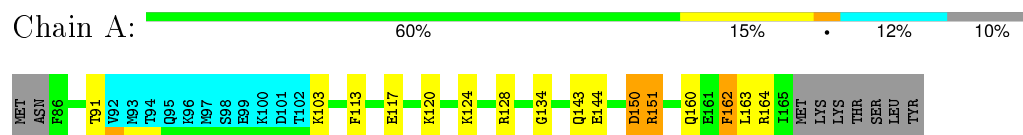
4.2.2 Score per residue for model 2 (medoid)

- Molecule 1: Caltractin, isoform 1



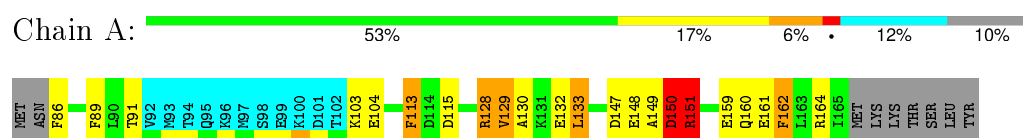
4.2.3 Score per residue for model 3

- Molecule 1: Caltractin, isoform 1



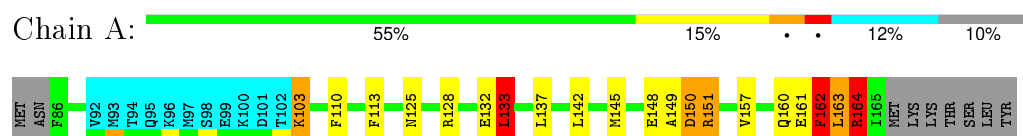
4.2.4 Score per residue for model 4

- Molecule 1: Caltractin, isoform 1



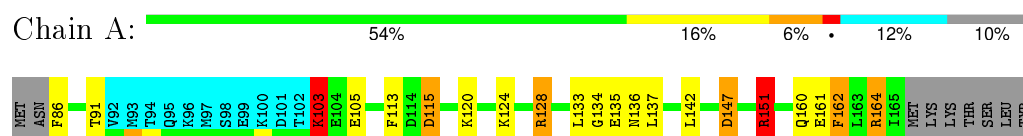
4.2.5 Score per residue for model 5

- Molecule 1: Caltractin, isoform 1



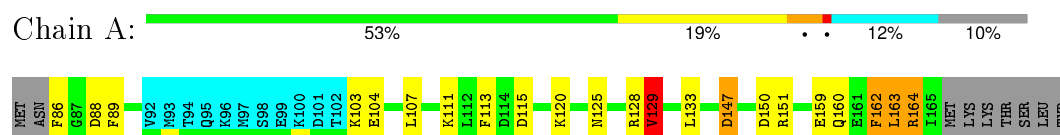
4.2.6 Score per residue for model 6

- Molecule 1: Caltractin, isoform 1



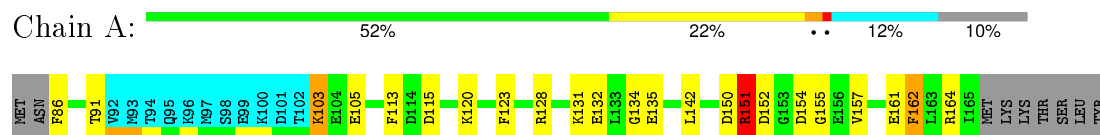
4.2.7 Score per residue for model 7

- Molecule 1: Caltractin, isoform 1



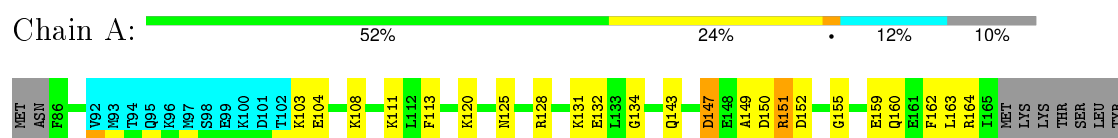
4.2.8 Score per residue for model 8

- Molecule 1: Caltractin, isoform 1



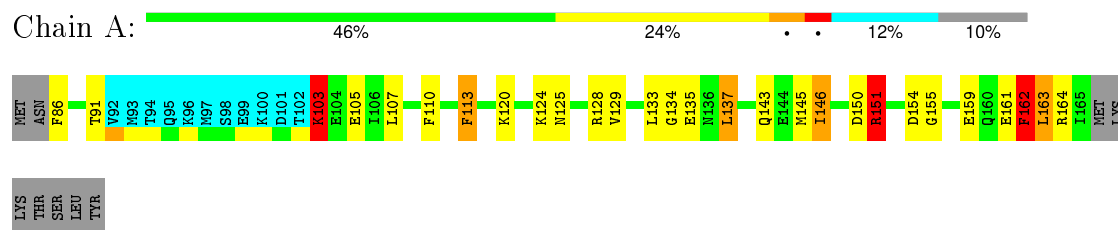
4.2.9 Score per residue for model 9

- Molecule 1: Caltractin, isoform 1



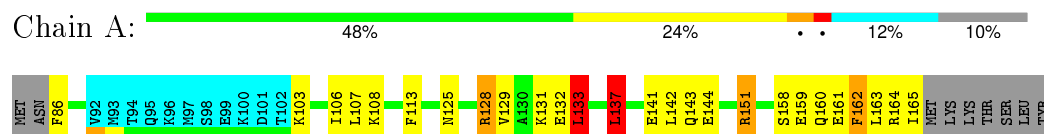
4.2.10 Score per residue for model 10

- Molecule 1: Caltractin, isoform 1



4.2.11 Score per residue for model 11

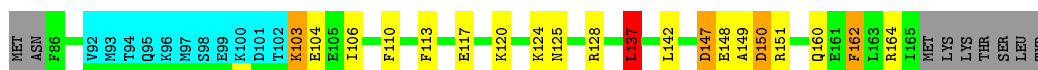
- Molecule 1: Caltractin, isoform 1



4.2.12 Score per residue for model 12

- Molecule 1: Caltractin, isoform 1





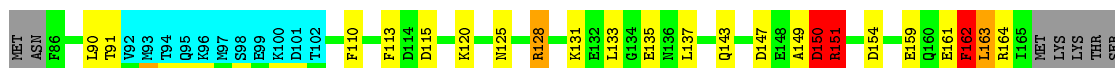
4.2.13 Score per residue for model 13

- Molecule 1: Caltractin, isoform 1



4.2.14 Score per residue for model 14

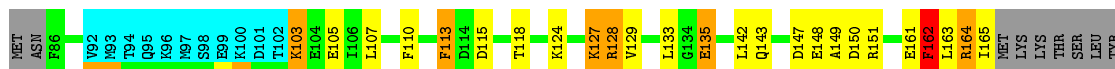
- Molecule 1: Caltractin, isoform 1



LEU
TYR

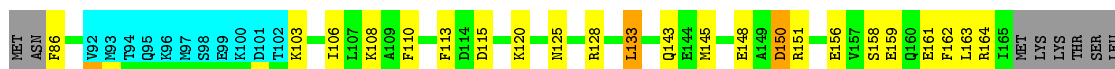
4.2.15 Score per residue for model 15

- Molecule 1: Caltractin, isoform 1



4.2.16 Score per residue for model 16

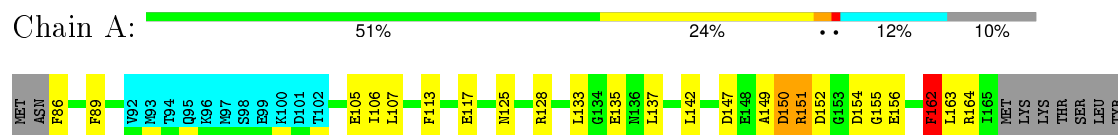
- Molecule 1: Caltractin, isoform 1



TYR

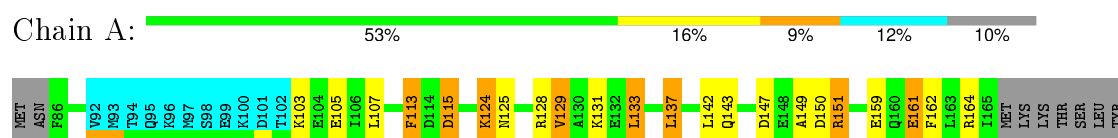
4.2.17 Score per residue for model 17

- Molecule 1: Caltractin, isoform 1



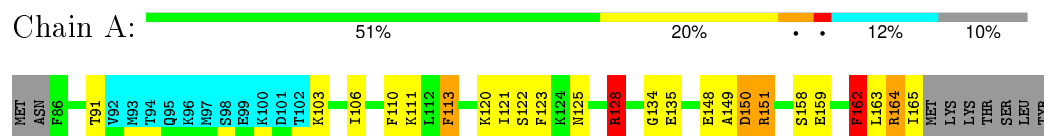
4.2.18 Score per residue for model 18

- Molecule 1: Caltractin, isoform 1



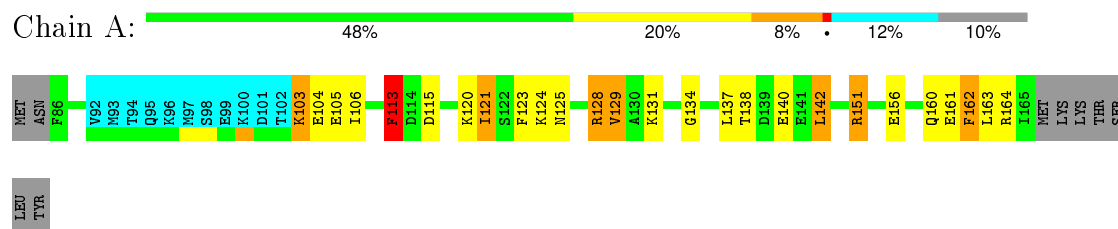
4.2.19 Score per residue for model 19

- Molecule 1: Caltractin, isoform 1



4.2.20 Score per residue for model 20

- Molecule 1: Caltractin, isoform 1



4.2.21 Score per residue for model 21

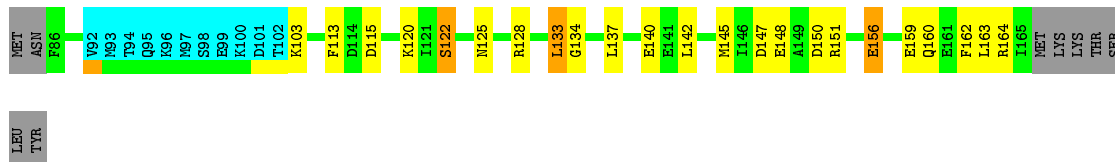
- Molecule 1: Caltractin, isoform 1





4.2.22 Score per residue for model 22

- Molecule 1: Caltractin, isoform 1



4.2.23 Score per residue for model 23

- Molecule 1: Caltractin, isoform 1



4.2.24 Score per residue for model 24

- Molecule 1: Caltractin, isoform 1



4.2.25 Score per residue for model 25

- Molecule 1: Caltractin, isoform 1



5 Refinement protocol and experimental data overview

The models were refined using the following method: *distance geometry, restraint energy minimization*.

Of the 76 calculated structures, 25 were deposited, based on the following criterion: *structures with the least restraint violations, structures with the lowest energy*.

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

Software name	Classification	Version
DGII	structure solution	I2000
DISCOVER	refinement	I2000

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	BMRB entry 5503
Number of chemical shift lists	1
Total number of shifts	540
Number of shifts mapped to atoms	540
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	45%

No validations of the models with respect to experimental NMR restraints is performed at this time.

6 Model quality

6.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the (average) root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	1.26±0.02	0±0/567 (0.0±0.0%)	1.67±0.07	9±3/755 (1.2±0.4%)
All	All	1.26	0/14175 (0.0%)	1.67	221/18875 (1.2%)

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	2.2±1.1
All	All	0	55

There are no bond-length outliers.

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	142	LEU	CB-CG-CD1	12.42	132.12	111.00	13	7
1	A	151	ARG	NE-CZ-NH1	11.75	126.18	120.30	23	25
1	A	128	ARG	NE-CZ-NH1	10.32	125.46	120.30	24	25
1	A	133	LEU	N-CA-CB	-10.23	89.95	110.40	15	8
1	A	113	PHE	CB-CG-CD2	-9.35	114.25	120.80	2	5
1	A	133	LEU	CB-CG-CD1	9.05	126.38	111.00	14	7
1	A	164	ARG	NE-CZ-NH1	8.64	124.62	120.30	21	25
1	A	129	VAL	CA-CB-CG2	8.48	123.62	110.90	7	4
1	A	151	ARG	NE-CZ-NH2	-7.36	116.62	120.30	18	4
1	A	113	PHE	CB-CG-CD1	7.22	125.86	120.80	21	2
1	A	150	ASP	N-CA-CB	-6.93	98.13	110.60	12	20
1	A	149	ALA	N-CA-CB	-6.92	100.41	110.10	19	10
1	A	86	PHE	CB-CG-CD2	-6.71	116.10	120.80	24	3
1	A	128	ARG	NE-CZ-NH2	-6.70	116.95	120.30	19	6
1	A	105	GLU	N-CA-CB	-6.56	98.79	110.60	24	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	89	PHE	CB-CG-CD2	-6.45	116.28	120.80	21	3
1	A	128	ARG	CA-CB-CG	6.40	127.49	113.40	24	2
1	A	163	LEU	CB-CG-CD2	6.31	121.72	111.00	14	4
1	A	151	ARG	N-CA-CB	-6.29	99.28	110.60	18	3
1	A	147	ASP	CA-CB-CG	6.26	127.17	113.40	12	5
1	A	150	ASP	CA-CB-CG	6.24	127.14	113.40	16	1
1	A	164	ARG	NE-CZ-NH2	-6.07	117.27	120.30	17	7
1	A	162	PHE	CB-CG-CD1	-5.94	116.64	120.80	10	2
1	A	86	PHE	CB-CG-CD1	5.86	124.90	120.80	24	1
1	A	137	LEU	CA-CB-CG	5.82	128.69	115.30	25	2
1	A	142	LEU	CA-CB-CG	5.79	128.63	115.30	17	4
1	A	128	ARG	N-CA-CB	5.77	120.98	110.60	20	1
1	A	137	LEU	N-CA-CB	-5.74	98.91	110.40	11	1
1	A	162	PHE	CA-CB-CG	5.71	127.59	113.90	19	3
1	A	162	PHE	N-CA-CB	-5.69	100.36	110.60	17	1
1	A	103	LYS	CA-CB-CG	5.67	125.87	113.40	6	3
1	A	123	PHE	CB-CG-CD1	-5.60	116.88	120.80	19	1
1	A	133	LEU	CB-CG-CD2	5.56	120.45	111.00	11	2
1	A	123	PHE	CB-CG-CD2	-5.55	116.92	120.80	20	2
1	A	121	ILE	CA-CB-CG1	5.54	121.52	111.00	20	1
1	A	129	VAL	CA-CB-CG1	5.53	119.20	110.90	25	1
1	A	161	GLU	N-CA-CB	-5.50	100.70	110.60	24	1
1	A	153	GLY	N-CA-C	5.42	126.66	113.10	25	1
1	A	161	GLU	CB-CA-C	5.41	121.22	110.40	6	2
1	A	124	LYS	CA-CB-CG	5.41	125.30	113.40	18	1
1	A	86	PHE	N-CA-C	-5.41	96.40	111.00	24	1
1	A	130	ALA	CB-CA-C	5.31	118.07	110.10	4	1
1	A	127	LYS	CB-CA-C	5.27	120.95	110.40	2	2
1	A	89	PHE	CB-CG-CD1	5.27	124.49	120.80	21	1
1	A	161	GLU	CA-CB-CG	5.25	124.96	113.40	8	2
1	A	157	VAL	CA-CB-CG2	5.20	118.70	110.90	8	1
1	A	121	ILE	N-CA-CB	5.19	122.74	110.80	21	1
1	A	132	GLU	N-CA-CB	-5.09	101.44	110.60	24	1
1	A	137	LEU	CB-CA-C	5.07	119.83	110.20	10	1
1	A	151	ARG	N-CA-C	5.06	124.67	111.00	4	1
1	A	103	LYS	CB-CG-CD	5.03	124.69	111.60	10	1
1	A	118	THR	CA-CB-CG2	5.02	119.42	112.40	15	1

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	162	PHE	Sidechain	25
1	A	113	PHE	Sidechain	11
1	A	151	ARG	Sidechain	10
1	A	128	ARG	Sidechain	4
1	A	86	PHE	Sidechain	3
1	A	123	PHE	Sidechain	1
1	A	89	PHE	Sidechain	1

6.2 Too-close contacts

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	562	540	539	1±1
All	All	14050	13500	13475	27

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 1.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:129:VAL:O	1:A:133:LEU:HD22	0.60	1.96	11	1
1:A:138:THR:O	1:A:142:LEU:HD13	0.54	2.02	21	2
1:A:113:PHE:CE1	1:A:129:VAL:HG21	0.48	2.44	21	2
1:A:149:ALA:HB1	1:A:157:VAL:HG22	0.47	1.85	13	1
1:A:137:LEU:HD12	1:A:142:LEU:HD11	0.47	1.86	11	1
1:A:110:PHE:HB3	1:A:162:PHE:CE1	0.47	2.45	5	7
1:A:113:PHE:CE1	1:A:129:VAL:HG11	0.47	2.45	2	3
1:A:89:PHE:CZ	1:A:129:VAL:HG22	0.45	2.46	7	1
1:A:163:LEU:HD12	1:A:164:ARG:H	0.43	1.74	5	1
1:A:143:GLN:NE2	1:A:146:ILE:HG21	0.42	2.29	10	2
1:A:122:SER:OG	1:A:156:GLU:OE1	0.42	2.34	22	1
1:A:143:GLN:HE21	1:A:146:ILE:HG21	0.42	1.75	10	1
1:A:129:VAL:HG23	1:A:133:LEU:HD12	0.41	1.92	7	1
1:A:107:LEU:HA	1:A:162:PHE:CE1	0.41	2.51	1	2
1:A:149:ALA:HB1	1:A:157:VAL:CG2	0.40	2.46	5	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	67/89 (75%)	58±2 (87±3%)	6±2 (9±3%)	2±2 (4±2%)	7	36
All	All	1675/2225 (75%)	1460 (87%)	155 (9%)	60 (4%)	7	36

All 10 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	134	GLY	12
1	A	103	LYS	11
1	A	115	ASP	10
1	A	150	ASP	8
1	A	137	LEU	7
1	A	155	GLY	5
1	A	154	ASP	3
1	A	117	GLU	2
1	A	135	GLU	1
1	A	136	ASN	1

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	61/81 (75%)	46±3 (76±6%)	15±3 (24±6%)	3	28
All	All	1525/2025 (75%)	1161 (76%)	364 (24%)	3	28

All 54 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	103	LYS	23
1	A	113	PHE	20
1	A	162	PHE	18
1	A	151	ARG	18
1	A	163	LEU	18
1	A	125	ASN	17
1	A	120	LYS	16
1	A	159	GLU	14
1	A	147	ASP	13
1	A	160	GLN	11
1	A	143	GLN	11
1	A	161	GLU	10
1	A	148	GLU	9
1	A	131	LYS	9
1	A	129	VAL	9
1	A	105	GLU	8
1	A	115	ASP	8
1	A	150	ASP	8
1	A	135	GLU	7
1	A	164	ARG	7
1	A	106	ILE	7
1	A	137	LEU	7
1	A	124	LYS	7
1	A	107	LEU	6
1	A	156	GLU	5
1	A	132	GLU	5
1	A	104	GLU	5
1	A	128	ARG	5
1	A	145	MET	5
1	A	133	LEU	5
1	A	86	PHE	4
1	A	165	ILE	4
1	A	140	GLU	4
1	A	158	SER	4
1	A	121	ILE	4
1	A	108	LYS	4
1	A	152	ASP	3
1	A	111	LYS	3
1	A	122	SER	3
1	A	90	LEU	2
1	A	89	PHE	2
1	A	136	ASN	2
1	A	144	GLU	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	110	PHE	2
1	A	127	LYS	1
1	A	139	ASP	1
1	A	123	PHE	1
1	A	91	THR	1
1	A	117	GLU	1
1	A	146	ILE	1
1	A	88	ASP	1
1	A	141	GLU	1
1	A	154	ASP	1
1	A	142	LEU	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 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

The completeness of assignment taking into account all chemical shift lists is 45% for the well-defined parts and 42% for the entire structure.

7.1 Chemical shift list 1

File name: BMRB entry 5503

Chemical shift list name: *assigned_chem_shift_list_1*

7.1.1 Bookkeeping

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	540
Number of shifts mapped to atoms	540
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

7.1.2 Chemical shift referencing

No chemical shift referencing corrections were calculated (not enough data).

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 45%, i.e. 393 atoms were assigned a chemical shift out of a possible 868. 0 out of 10 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹ H	¹³ C	¹⁵ N
Backbone	138/345 (40%)	138/138 (100%)	0/138 (0%)	0/69 (0%)
Sidechain	226/469 (48%)	226/271 (83%)	0/178 (0%)	0/20 (0%)
Aromatic	29/54 (54%)	29/30 (97%)	0/24 (0%)	0/0 (—%)
Overall	393/868 (45%)	393/439 (90%)	0/340 (0%)	0/89 (0%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 42%, i.e. 419 atoms were assigned a chemical shift out of a possible 1003. 0 out of 11 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	148/400 (37%)	148/160 (92%)	0/160 (0%)	0/80 (0%)
Sidechain	242/549 (44%)	242/318 (76%)	0/208 (0%)	0/23 (0%)
Aromatic	29/54 (54%)	29/30 (97%)	0/24 (0%)	0/0 (—%)
Overall	419/1003 (42%)	419/508 (82%)	0/392 (0%)	0/103 (0%)

7.1.4 Statistically unusual chemical shifts [i](#)

There are no statistically unusual chemical shifts.

7.1.5 Random Coil Index (RCI) plots [i](#)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.

Random coil index (RCI) for chain A:

