Full wwPDB NMR Structure Validation Report (i)

Apr 21, 2024 – 07:23 AM EDT

PDB ID	:	2LRP
BMRB ID	:	18389
Title	:	Solution structure, dynamics and binding studies of CtCBM11
Authors	:	Viegas, A.; Cabrita, E.J.
Deposited on	:	2012-04-11

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 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:

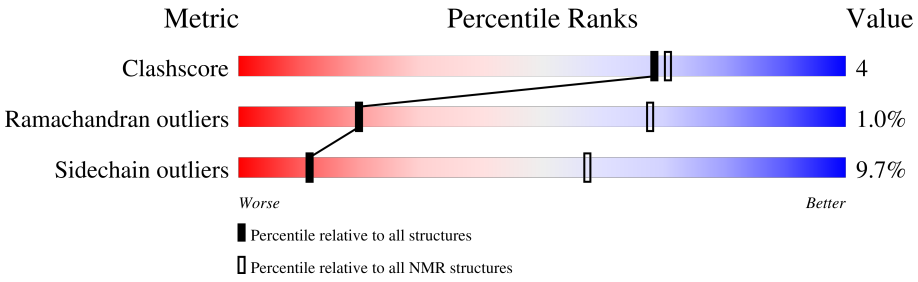
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson 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 is 82%.

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	172	83% 12% • 5%

2 Ensemble composition and analysis

This entry contains 20 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:6-A:169 (164)	1.40	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 4 clusters and 3 single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2613 atoms, of which 1276 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Endoglucanase H.

Mol	Chain	Residues	Atoms						Trace
1	A	172	Total	C	H	N	O	S	0
			2611	846	1276	216	267	6	

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP P16218
A	2	ALA	-	expression tag	UNP P16218
A	3	SER	-	expression tag	UNP P16218
A	171	LEU	-	expression tag	UNP P16218
A	172	GLU	-	expression tag	UNP P16218

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	
2	A	2	Total	Ca
			2	2

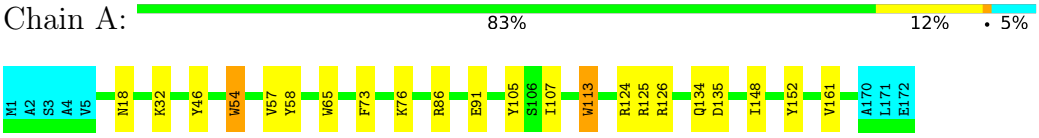
4 Residue-property plots

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: Endoglucanase H

Chain A:



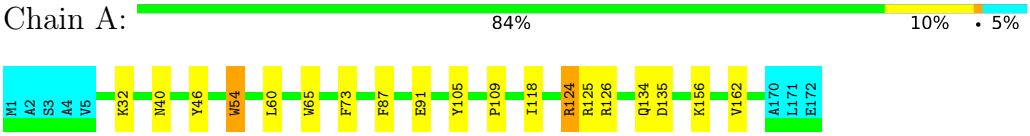
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Endoglucanase H

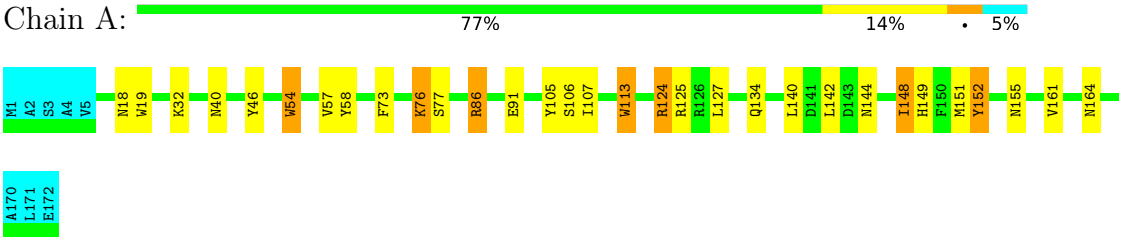
Chain A:



4.2.2 Score per residue for model 2 (medoid)

• Molecule 1: Endoglucanase H

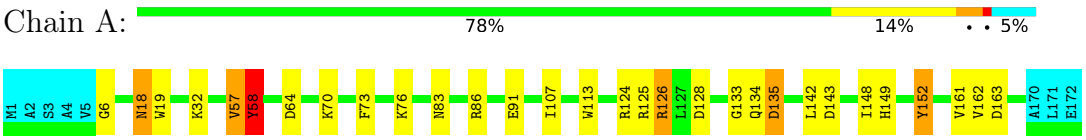
Chain A:



4.2.3 Score per residue for model 3

• Molecule 1: Endoglucanase H

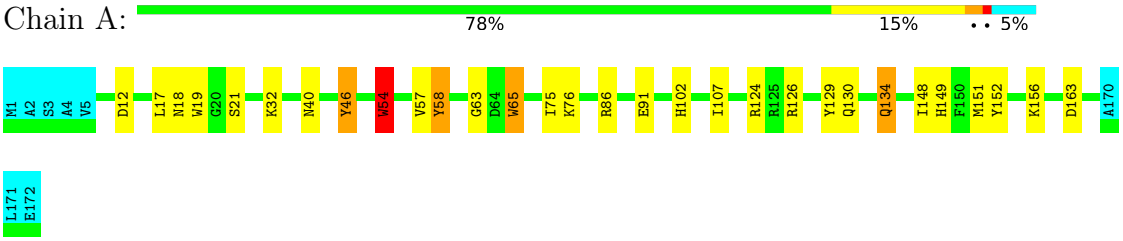
Chain A:



4.2.4 Score per residue for model 4

• Molecule 1: Endoglucanase H

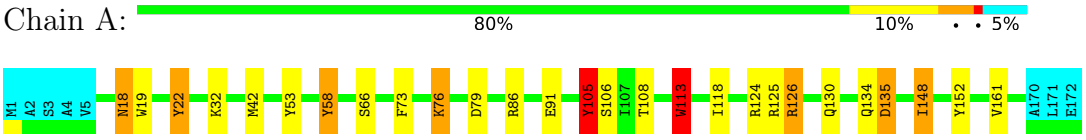
Chain A:



4.2.5 Score per residue for model 5

• Molecule 1: Endoglucanase H

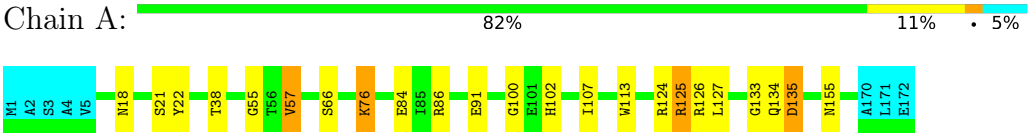
Chain A:



4.2.6 Score per residue for model 6

• Molecule 1: Endoglucanase H

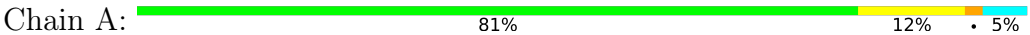
Chain A:



4.2.7 Score per residue for model 7

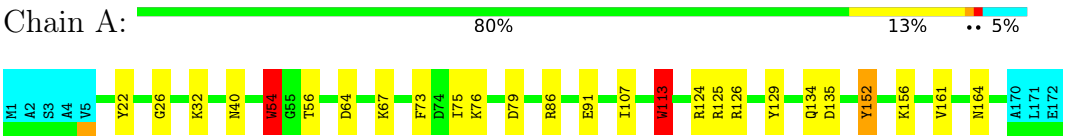
• Molecule 1: Endoglucanase H

Chain A:



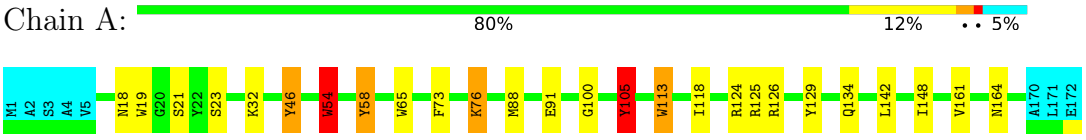
4.2.8 Score per residue for model 8

• Molecule 1: Endoglucanase H



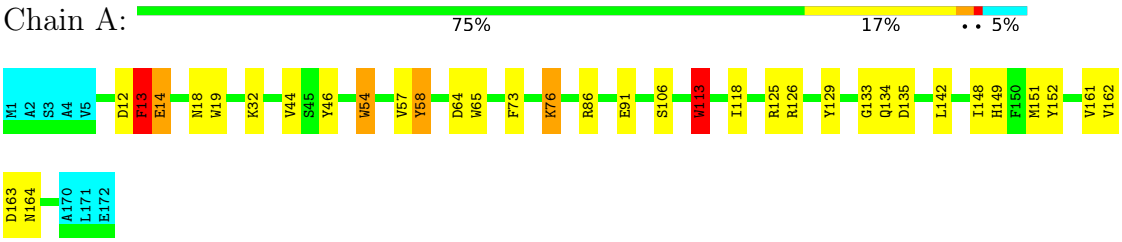
4.2.9 Score per residue for model 9

• Molecule 1: Endoglucanase H



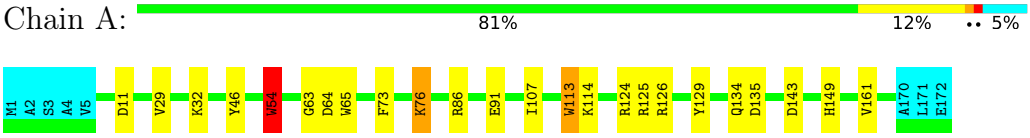
4.2.10 Score per residue for model 10

• Molecule 1: Endoglucanase H



4.2.11 Score per residue for model 11

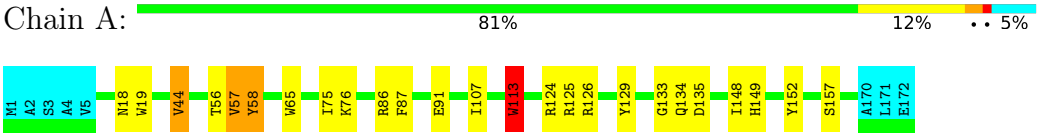
• Molecule 1: Endoglucanase H



4.2.12 Score per residue for model 12

• Molecule 1: Endoglucanase H

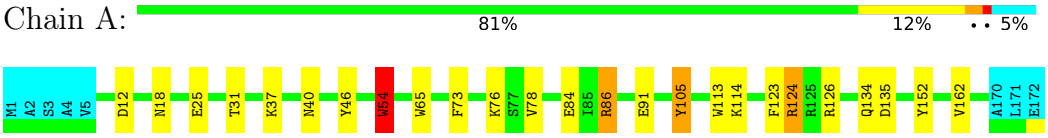
Chain A:



4.2.13 Score per residue for model 13

• Molecule 1: Endoglucanase H

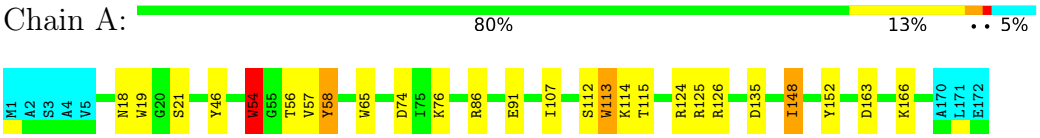
Chain A:



4.2.14 Score per residue for model 14

• Molecule 1: Endoglucanase H

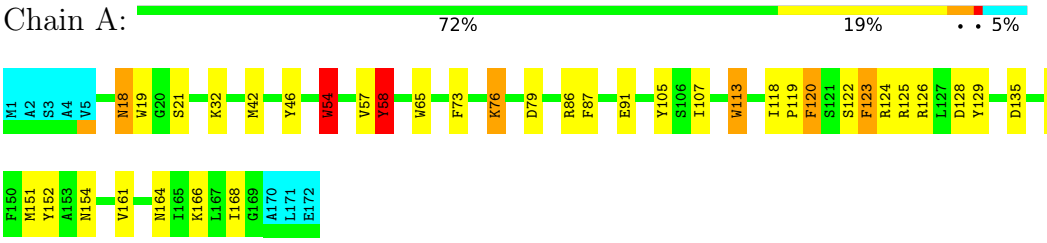
Chain A:



4.2.15 Score per residue for model 15

• Molecule 1: Endoglucanase H

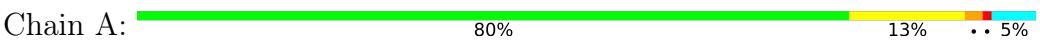
Chain A:



4.2.16 Score per residue for model 16

• Molecule 1: Endoglucanase H

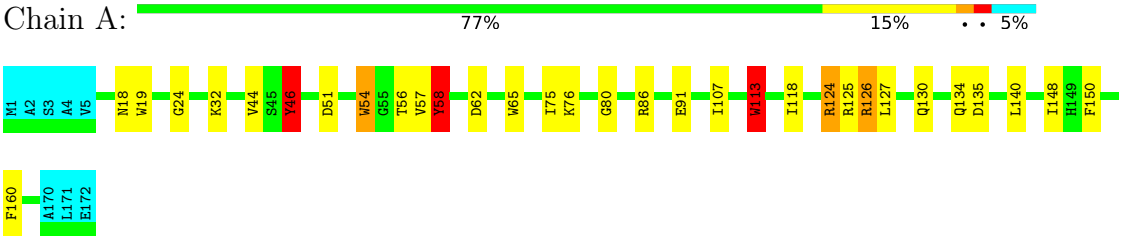
Chain A:



4.2.17 Score per residue for model 17

• Molecule 1: Endoglucanase H

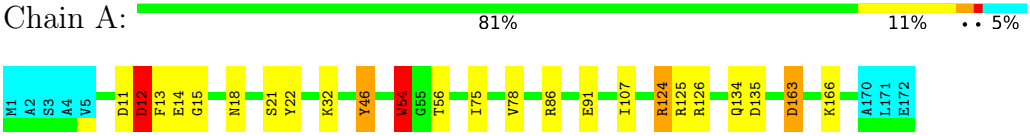
Chain A:



4.2.18 Score per residue for model 18

• Molecule 1: Endoglucanase H

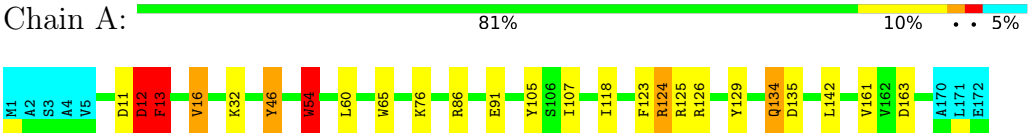
Chain A:



4.2.19 Score per residue for model 19

• Molecule 1: Endoglucanase H

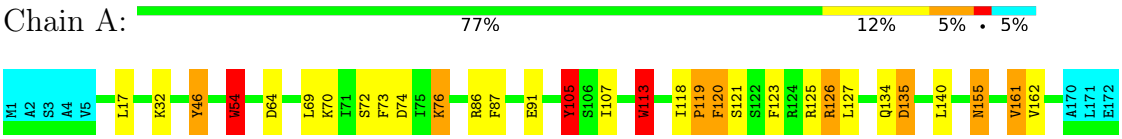
Chain A:



4.2.20 Score per residue for model 20

• Molecule 1: Endoglucanase H

Chain A:



5 Refinement protocol and experimental data overview

The models were refined using the following method: *molecular dynamics*.

Of the 20 calculated structures, 20 were deposited, based on the following criterion: *target function*.

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

Software name	Classification	Version
CYANA	structure solution	2.1
Amber	refinement	9.0

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)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	1790
Number of shifts mapped to atoms	1790
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	82%

6 Model quality (i)

6.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: CA

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	0.74±0.01	0±0/1314 (0.0± 0.0%)	1.23±0.05	11±3/1778 (0.6± 0.2%)
All	All	0.74	0/26280 (0.0%)	1.23	221/35560 (0.6%)

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	5.2±1.7
All	All	0	104

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(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	65	TRP	CD1-CG-CD2	-18.07	91.84	106.30	4	2
1	A	113	TRP	CB-CG-CD1	-12.25	111.07	127.00	20	8
1	A	65	TRP	NE1-CE2-CD2	-11.87	95.44	107.30	4	1
1	A	113	TRP	CD1-CG-CD2	-10.70	97.74	106.30	20	1
1	A	113	TRP	CG-CD1-NE1	-10.26	99.84	110.10	20	1
1	A	126	ARG	NE-CZ-NH1	10.03	125.31	120.30	10	18
1	A	54	TRP	CD1-CG-CD2	-9.80	98.46	106.30	10	15
1	A	125	ARG	NE-CZ-NH1	9.46	125.03	120.30	17	16
1	A	65	TRP	CG-CD2-CE3	-8.65	126.11	133.90	4	1
1	A	124	ARG	NE-CZ-NH1	8.64	124.62	120.30	4	15
1	A	113	TRP	CB-CG-CD2	8.52	137.67	126.60	11	9
1	A	105	TYR	CB-CG-CD2	-8.15	116.11	121.00	16	6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	86	ARG	NE-CZ-NH1	7.90	124.25	120.30	11	16
1	A	135	ASP	CB-CG-OD1	7.84	125.36	118.30	5	6
1	A	54	TRP	CE2-CD2-CG	7.82	113.56	107.30	10	3
1	A	91	GLU	OE1-CD-OE2	-7.77	113.97	123.30	3	20
1	A	113	TRP	NE1-CE2-CZ2	7.35	138.48	130.40	20	1
1	A	54	TRP	CG-CD2-CE3	-7.34	127.29	133.90	10	3
1	A	65	TRP	NE1-CE2-CZ2	-7.32	122.35	130.40	4	1
1	A	57	VAL	CA-CB-CG2	7.27	121.81	110.90	3	5
1	A	113	TRP	NE1-CE2-CD2	-7.27	100.03	107.30	20	1
1	A	54	TRP	CB-CG-CD2	7.25	136.03	126.60	8	8
1	A	152	TYR	CB-CG-CD2	-7.25	116.65	121.00	5	2
1	A	113	TRP	CB-CA-C	7.04	124.47	110.40	10	6
1	A	14	GLU	OE1-CD-OE2	-6.77	115.18	123.30	7	2
1	A	58	TYR	CB-CG-CD2	-6.69	116.98	121.00	7	4
1	A	54	TRP	NE1-CE2-CD2	-6.23	101.07	107.30	10	8
1	A	125	ARG	NE-CZ-NH2	-6.21	117.19	120.30	11	2
1	A	135	ASP	CB-CG-OD2	6.13	123.82	118.30	15	3
1	A	163	ASP	CB-CG-OD1	5.93	123.63	118.30	10	3
1	A	105	TYR	CB-CG-CD1	5.74	124.44	121.00	16	4
1	A	113	TRP	N-CA-CB	-5.73	100.29	110.60	8	4
1	A	12	ASP	C-N-CA	5.66	135.85	121.70	10	1
1	A	54	TRP	NE1-CE2-CZ2	5.61	136.57	130.40	2	3
1	A	76	LYS	CG-CD-CE	5.60	128.70	111.90	14	1
1	A	12	ASP	CB-CG-OD2	5.51	123.26	118.30	4	1
1	A	152	TYR	CA-CB-CG	5.49	123.83	113.40	3	1
1	A	54	TRP	CA-CB-CG	5.44	124.04	113.70	9	1
1	A	13	PHE	CB-CG-CD2	-5.43	117.00	120.80	10	1
1	A	54	TRP	CB-CG-CD1	-5.35	120.04	127.00	8	1
1	A	74	ASP	CB-CA-C	5.31	121.02	110.40	14	1
1	A	152	TYR	CB-CG-CD1	5.30	124.18	121.00	3	3
1	A	87	PHE	CB-CG-CD2	-5.26	117.12	120.80	15	1
1	A	65	TRP	CE2-CD2-CG	-5.26	103.09	107.30	4	1
1	A	58	TYR	CB-CG-CD1	5.26	124.16	121.00	7	1
1	A	54	TRP	CB-CA-C	5.15	120.70	110.40	9	4
1	A	163	ASP	CB-CG-OD2	5.14	122.92	118.30	18	1
1	A	16	VAL	C-N-CA	5.11	134.48	121.70	19	1
1	A	113	TRP	CA-CB-CG	5.11	123.40	113.70	20	1
1	A	135	ASP	OD1-CG-OD2	-5.03	113.74	123.30	6	1
1	A	86	ARG	NH1-CZ-NH2	-5.02	113.88	119.40	5	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	46	TYR	Peptide,Sidechain	8
1	A	105	TYR	Peptide,Sidechain	8
1	A	162	VAL	Peptide	5
1	A	75	ILE	Peptide	5
1	A	129	TYR	Sidechain,Peptide	5
1	A	133	GLY	Peptide	4
1	A	134	GLN	Peptide	4
1	A	163	ASP	Peptide	4
1	A	123	PHE	Sidechain	4
1	A	106	SER	Peptide	3
1	A	151	MET	Peptide	3
1	A	152	TYR	Sidechain	3
1	A	64	ASP	Peptide	3
1	A	44	VAL	Peptide	3
1	A	156	LYS	Peptide	2
1	A	128	ASP	Peptide	2
1	A	113	TRP	Peptide	2
1	A	125	ARG	Sidechain	2
1	A	56	THR	Peptide	2
1	A	78	VAL	Peptide	2
1	A	12	ASP	Peptide	2
1	A	22	TYR	Sidechain	1
1	A	53	TYR	Sidechain	1
1	A	100	GLY	Peptide	1
1	A	102	HIS	Peptide	1
1	A	20	GLY	Peptide	1
1	A	143	ASP	Peptide	1
1	A	149	HIS	Peptide	1
1	A	115	THR	Peptide	1
1	A	120	PHE	Peptide	1
1	A	122	SER	Peptide	1
1	A	70	LYS	Peptide	1
1	A	71	ILE	Peptide	1
1	A	87	PHE	Sidechain	1
1	A	88	MET	Peptide	1
1	A	15	GLY	Peptide	1
1	A	124	ARG	Sidechain	1
1	A	126	ARG	Peptide	1
1	A	13	PHE	Peptide	1
1	A	119	PRO	Peptide	1
1	A	155	ASN	Peptide	1

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	1281	1219	1219	9±6
All	All	25660	24380	24360	189

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 4.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:76:LYS:HE3	1:A:113:TRP:CE3	0.77	2.13	20	6
1:A:19:TRP:CZ2	1:A:58:TYR:CE2	0.76	2.73	4	10
1:A:105:TYR:CD2	1:A:107:ILE:HD12	0.75	2.17	16	2
1:A:19:TRP:CZ3	1:A:58:TYR:CG	0.64	2.85	9	10
1:A:19:TRP:CH2	1:A:58:TYR:CE2	0.62	2.88	17	10
1:A:76:LYS:HE3	1:A:113:TRP:CD2	0.62	2.30	20	2
1:A:76:LYS:CE	1:A:113:TRP:CE3	0.62	2.83	20	1
1:A:18:ASN:ND2	1:A:19:TRP:CZ2	0.61	2.69	5	10
1:A:19:TRP:CH2	1:A:58:TYR:CZ	0.61	2.89	2	7
1:A:19:TRP:CH2	1:A:58:TYR:CD2	0.59	2.91	5	7
1:A:22:TYR:CE1	1:A:57:VAL:HG23	0.58	2.32	6	1
1:A:19:TRP:CZ2	1:A:58:TYR:CD2	0.58	2.92	5	5
1:A:19:TRP:CZ3	1:A:58:TYR:CD2	0.58	2.91	5	10
1:A:11:ASP:HA	1:A:13:PHE:CE2	0.55	2.36	18	1
1:A:76:LYS:HE2	1:A:113:TRP:CE3	0.54	2.36	9	1
1:A:71:ILE:HD11	1:A:118:ILE:HD12	0.54	1.80	16	1
1:A:19:TRP:CH2	1:A:58:TYR:CE1	0.53	2.96	2	2
1:A:18:ASN:ND2	1:A:19:TRP:CE2	0.52	2.78	4	9
1:A:105:TYR:CE2	1:A:118:ILE:HD13	0.52	2.39	5	3
1:A:76:LYS:HD3	1:A:113:TRP:CE3	0.52	2.39	17	3
1:A:19:TRP:CE3	1:A:58:TYR:CD2	0.52	2.97	3	7
1:A:18:ASN:ND2	1:A:19:TRP:CH2	0.51	2.78	2	4
1:A:46:TYR:CE2	1:A:54:TRP:CD1	0.51	2.98	14	7
1:A:46:TYR:CE1	1:A:54:TRP:HB2	0.51	2.40	19	2
1:A:76:LYS:HD2	1:A:113:TRP:HA	0.50	1.82	10	2
1:A:58:TYR:CG	1:A:148:ILE:HD11	0.50	2.41	10	7
1:A:54:TRP:CZ3	1:A:152:TYR:CG	0.50	2.99	8	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:58:TYR:CG	1:A:148:ILE:HD12	0.49	2.43	5	2
1:A:46:TYR:CE2	1:A:54:TRP:NE1	0.48	2.81	11	1
1:A:63:GLY:HA2	1:A:65:TRP:CZ3	0.48	2.43	4	1
1:A:57:VAL:HG13	1:A:149:HIS:CE1	0.48	2.44	2	6
1:A:118:ILE:HG22	1:A:119:PRO:HD2	0.47	1.86	20	1
1:A:12:ASP:H	1:A:13:PHE:HB3	0.47	1.70	19	1
1:A:76:LYS:HD3	1:A:113:TRP:CD2	0.47	2.45	8	3
1:A:76:LYS:HD2	1:A:113:TRP:CZ3	0.47	2.45	17	1
1:A:58:TYR:CG	1:A:148:ILE:CD1	0.46	2.97	5	2
1:A:46:TYR:CD2	1:A:54:TRP:CD1	0.46	3.03	11	1
1:A:19:TRP:CZ3	1:A:58:TYR:CD1	0.46	3.04	2	2
1:A:54:TRP:CZ2	1:A:152:TYR:HB3	0.45	2.46	14	3
1:A:58:TYR:CD1	1:A:148:ILE:HD11	0.45	2.45	3	6
1:A:12:ASP:H	1:A:13:PHE:CB	0.45	2.25	19	1
1:A:58:TYR:CD1	1:A:148:ILE:CD1	0.45	2.99	3	2
1:A:102:HIS:CE1	1:A:129:TYR:CE1	0.45	3.05	4	1
1:A:76:LYS:HE3	1:A:113:TRP:CZ3	0.45	2.47	11	3
1:A:46:TYR:CE1	1:A:54:TRP:CB	0.44	3.00	9	2
1:A:11:ASP:HA	1:A:13:PHE:CD1	0.44	2.48	19	1
1:A:76:LYS:CE	1:A:161:VAL:HG13	0.44	2.42	7	1
1:A:70:LYS:HA	1:A:120:PHE:H	0.44	1.72	20	1
1:A:11:ASP:HA	1:A:13:PHE:CZ	0.44	2.48	18	1
1:A:76:LYS:CE	1:A:113:TRP:CD2	0.44	3.01	9	1
1:A:46:TYR:CE1	1:A:54:TRP:HB3	0.43	2.48	9	1
1:A:76:LYS:HE3	1:A:161:VAL:HG13	0.43	1.90	7	1
1:A:105:TYR:CE2	1:A:107:ILE:HD12	0.43	2.48	16	1
1:A:76:LYS:HE3	1:A:113:TRP:CE2	0.43	2.49	6	1
1:A:76:LYS:HE2	1:A:113:TRP:CE2	0.43	2.49	3	1
1:A:105:TYR:CE2	1:A:107:ILE:CD1	0.42	3.02	2	1
1:A:37:LYS:HE2	1:A:113:TRP:CE3	0.42	2.49	13	1
1:A:105:TYR:HD2	1:A:107:ILE:HD12	0.42	1.68	20	1
1:A:54:TRP:CE3	1:A:152:TYR:CD1	0.42	3.08	8	1
1:A:58:TYR:CD2	1:A:148:ILE:HD12	0.42	2.49	5	1
1:A:76:LYS:HE2	1:A:161:VAL:CG2	0.41	2.45	20	1
1:A:76:LYS:HE3	1:A:113:TRP:HA	0.40	1.93	9	1
1:A:58:TYR:CD2	1:A:148:ILE:HD11	0.40	2.51	10	1
1:A:19:TRP:CZ2	1:A:58:TYR:CZ	0.40	3.09	2	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	164/172 (95%)	149±4 (91±3%)	13±4 (8±2%)	2±1 (1±1%)	20 68
All	All	3280/3440 (95%)	2982 (91%)	266 (8%)	32 (1%)	20 68

All 22 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	164	ASN	3
1	A	66	SER	3
1	A	79	ASP	3
1	A	113	TRP	2
1	A	13	PHE	2
1	A	14	GLU	2
1	A	120	PHE	2
1	A	109	PRO	1
1	A	6	GLY	1
1	A	55	GLY	1
1	A	81	SER	1
1	A	26	GLY	1
1	A	100	GLY	1
1	A	63	GLY	1
1	A	64	ASP	1
1	A	119	PRO	1
1	A	123	PHE	1
1	A	24	GLY	1
1	A	80	GLY	1
1	A	12	ASP	1
1	A	16	VAL	1
1	A	121	SER	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	140/145 (97%)	126±3 (90±2%)	14±3 (10±2%)	12 57
All	All	2800/2900 (97%)	2529 (90%)	271 (10%)	12 57

All 76 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	32	LYS	16
1	A	134	GLN	16
1	A	54	TRP	15
1	A	76	LYS	13
1	A	73	PHE	12
1	A	135	ASP	12
1	A	107	ILE	11
1	A	65	TRP	10
1	A	161	VAL	10
1	A	58	TYR	9
1	A	124	ARG	8
1	A	142	LEU	6
1	A	18	ASN	6
1	A	21	SER	6
1	A	40	ASN	5
1	A	118	ILE	5
1	A	127	LEU	5
1	A	148	ILE	5
1	A	126	ARG	5
1	A	60	LEU	4
1	A	87	PHE	4
1	A	113	TRP	4
1	A	86	ARG	3
1	A	140	LEU	3
1	A	155	ASN	3
1	A	130	GLN	3
1	A	22	TYR	3
1	A	129	TYR	3
1	A	114	LYS	3
1	A	56	THR	3
1	A	12	ASP	3
1	A	166	LYS	3
1	A	143	ASP	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	152	TYR	2
1	A	17	LEU	2
1	A	46	TYR	2
1	A	42	MET	2
1	A	108	THR	2
1	A	84	GLU	2
1	A	67	LYS	2
1	A	164	ASN	2
1	A	13	PHE	2
1	A	77	SER	1
1	A	144	ASN	1
1	A	70	LYS	1
1	A	83	ASN	1
1	A	79	ASP	1
1	A	38	THR	1
1	A	14	GLU	1
1	A	156	LYS	1
1	A	23	SER	1
1	A	88	MET	1
1	A	64	ASP	1
1	A	125	ARG	1
1	A	151	MET	1
1	A	11	ASP	1
1	A	29	VAL	1
1	A	44	VAL	1
1	A	157	SER	1
1	A	25	GLU	1
1	A	31	THR	1
1	A	112	SER	1
1	A	105	TYR	1
1	A	123	PHE	1
1	A	154	ASN	1
1	A	168	ILE	1
1	A	28	LYS	1
1	A	30	SER	1
1	A	89	ILE	1
1	A	51	ASP	1
1	A	62	ASP	1
1	A	150	PHE	1
1	A	160	PHE	1
1	A	69	LEU	1
1	A	72	SER	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	74	ASP	1

6.3.3 RNA

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates

There are no monosaccharides in this entry.

6.6 Ligand geometry

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

6.7 Other polymers

There are no such molecules in this entry.

6.8 Polymer linkage issues

There are no chain breaks in this entry.

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 82% for the well-defined parts and 80% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

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	1790
Number of shifts mapped to atoms	1790
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	12

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	161	2.62 ± 0.10	Should be checked
$^{13}C_{\beta}$	142	2.20 ± 0.16	Should be checked
$^{13}C'$	157	2.88 ± 0.14	Should be applied
^{15}N	158	0.49 ± 0.56	None needed (< 0.5 ppm)

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 82%, i.e. 1740 atoms were assigned a chemical shift out of a possible 2130. 0 out of 18 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	791/830 (95%)	329/343 (96%)	309/328 (94%)	153/159 (96%)
Sidechain	929/1081 (86%)	626/700 (89%)	293/347 (84%)	10/34 (29%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	20/219 (9%)	15/107 (14%)	0/104 (0%)	5/8 (62%)
Overall	1740/2130 (82%)	970/1150 (84%)	602/779 (77%)	168/201 (84%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 80%, i.e. 1790 atoms were assigned a chemical shift out of a possible 2225. 0 out of 20 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	814/870 (94%)	338/359 (94%)	318/344 (92%)	158/167 (95%)
Sidechain	956/1136 (84%)	645/738 (87%)	301/364 (83%)	10/34 (29%)
Aromatic	20/219 (9%)	15/107 (14%)	0/104 (0%)	5/8 (62%)
Overall	1790/2225 (80%)	998/1204 (83%)	619/812 (76%)	173/209 (83%)

7.1.4 Statistically unusual chemical shifts (i)

The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	37	LYS	HB2	-0.36	0.58 – 2.97	-8.9
1	A	75	ILE	CG1	13.16	19.24 – 36.26	-8.6
1	A	58	TYR	HB2	0.15	1.09 – 4.72	-7.6
1	A	140	LEU	HB2	-0.88	-0.07 – 3.30	-7.4
1	A	75	ILE	CG2	27.02	10.93 – 24.12	7.2
1	A	76	LYS	HD2	0.18	0.58 – 2.64	-6.9
1	A	86	ARG	HD2	1.64	1.97 – 4.26	-6.5
1	A	101	GLU	HB3	3.28	0.95 – 3.05	6.1
1	A	86	ARG	HD3	1.64	1.81 – 4.39	-5.7
1	A	44	VAL	HB	0.30	0.43 – 3.54	-5.4
1	A	76	LYS	HB2	0.56	0.58 – 2.97	-5.1
1	A	74	ASP	HA	6.15	3.04 – 6.12	5.1

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.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A: