Full wwPDB NMR Structure Validation Report (i)

Apr 21, 2024 – 02:21 AM EDT

PDB ID	:	2LYA
BMRB ID	:	18715
Title	:	Structure of HIV-1 myr(-) matrix protein in complex with 1,2-dioctanoyl-sn-phosphatidylcholine
Authors	:	Vlach, J.; Saad, J.
Deposited on	:	2012-09-14

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:

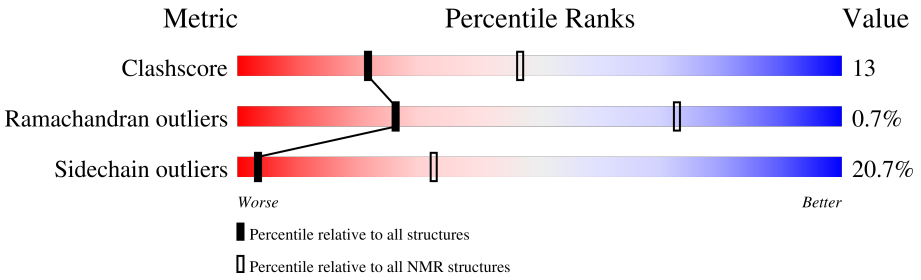
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (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	137	48% 23% • 23% •

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 13 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:7-A:105 (99)	0.47	13

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 5 single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2164 atoms, of which 1096 are hydrogens and 0 are deuteriums.

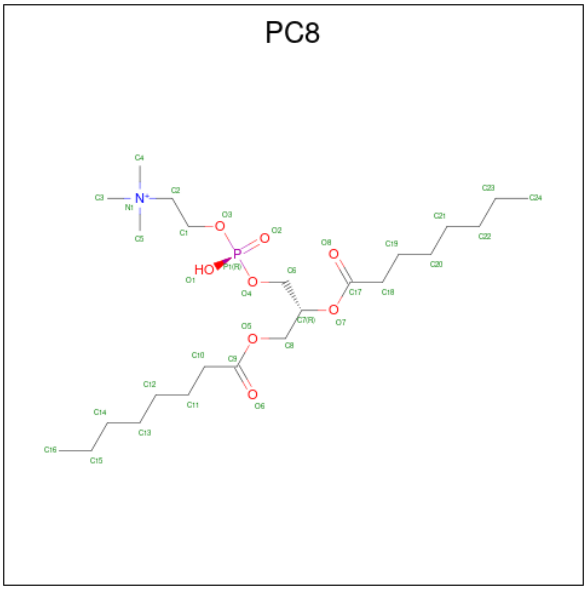
- Molecule 1 is a protein called Matrix protein p17.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	131	2082	640	1048	191	201	2	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	133	HIS	-	expression tag	UNP P12497
A	134	HIS	-	expression tag	UNP P12497
A	135	HIS	-	expression tag	UNP P12497
A	136	HIS	-	expression tag	UNP P12497
A	137	HIS	-	expression tag	UNP P12497
A	138	HIS	-	expression tag	UNP P12497

- Molecule 2 is 1,2-DIOCTANOYL-SN-GLYCERO-3-PHOSPHOCHOLINE (three-letter code: PC8) (formula: $C_{24}H_{49}NO_8P$).



Mol	Chain	Residues	Atoms					
			Total	C	H	N	O	P
2	A	1	82	24	48	1	8	1

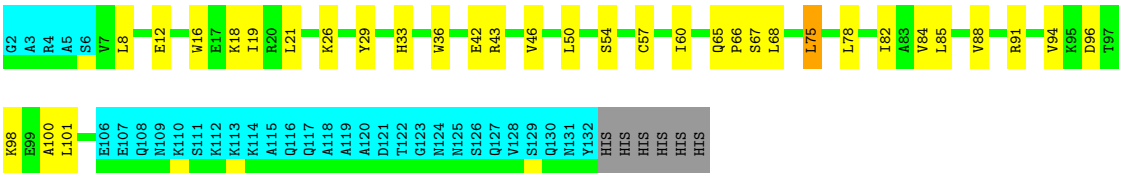
4 Residue-property plots (i)

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: Matrix protein p17

Chain A: 48% 23% • 23% •



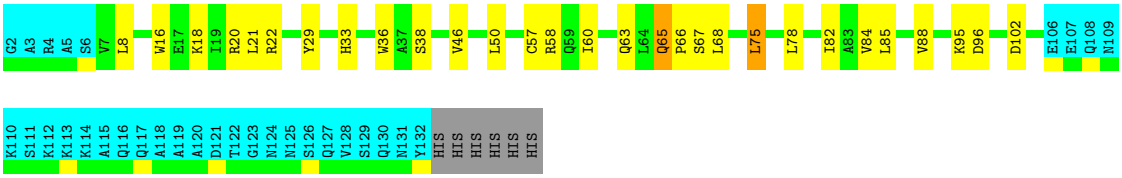
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Matrix protein p17

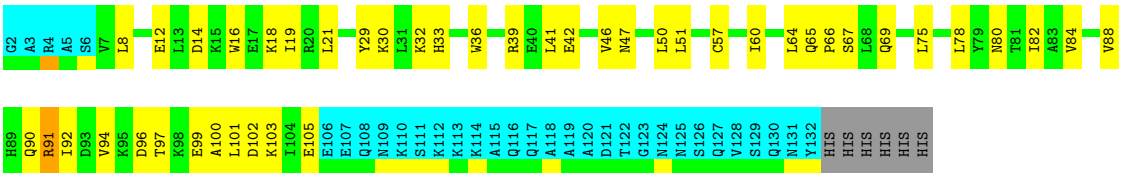
Chain A: 51% 20% • 23% •



4.2.2 Score per residue for model 2

• Molecule 1: Matrix protein p17

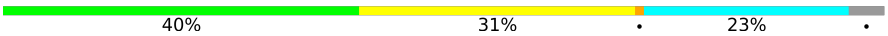
Chain A: 40% 31% • 23% •

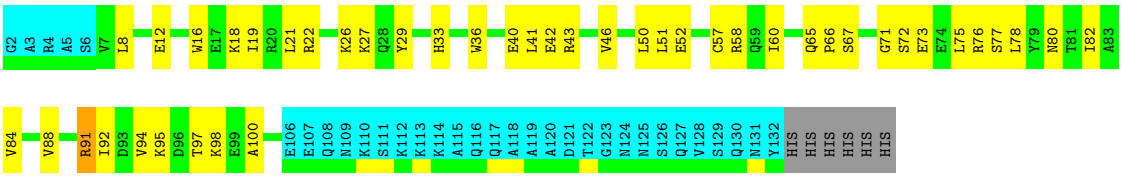


4.2.3 Score per residue for model 3

- Molecule 1: Matrix protein p17

Chain A:

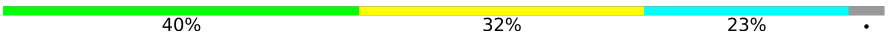


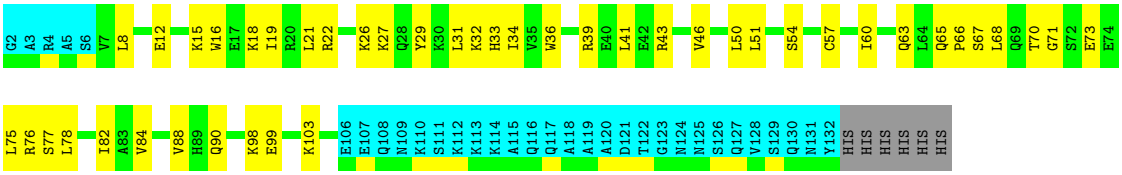


4.2.4 Score per residue for model 4

- Molecule 1: Matrix protein p17

Chain A:

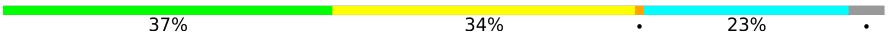


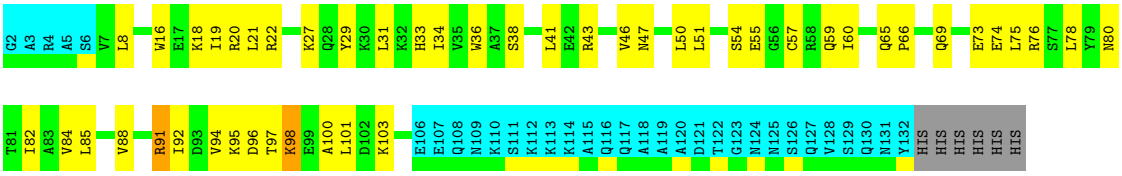


4.2.5 Score per residue for model 5

- Molecule 1: Matrix protein p17

Chain A:





4.2.6 Score per residue for model 6

- Molecule 1: Matrix protein p17

Chain A: 46% 26% 23% •

G2 A3 R4 A5 S6 L13 K18 L21 Y29 K30 L31 K32 H33 W36 R39 E42 R43 V46 N47 L50 S54 C57 I60 Q63 L64 Q65 P66 S67 L68 E73 E74 L75 R76 S77 L78 I82 A83 V84 V88 R91 V94 K95

D96 A100 L101 E106 E107 Q108 N109 K110 S111 K112 K113 K114 A115 Q116 Q117 A118 A119 A120 D121 T122 G123 N124 N125 S126 Q127 V128 S129 Q130 N131 Y132 HIS HIS HIS HIS HIS HIS

4.2.7 Score per residue for model 7

- Molecule 1: Matrix protein p17

Chain A: 42% 29% • 23% •

G2 A3 R4 A5 S6 L13 D14 K18 I19 K26 K27 Q28 Y29 K30 S38 R39 E40 L41 E42 V46 L50 L51 E52 T53 S54 E55 G56 C57 R58 Q59 I60 Q63 L64 Q65 P66 S67 L68 E74 L75 L78 Y79 I82 A83 V84 V88 H89 Q90 D93

V94 K95 D96 T97 K98 E99 A100 L101 D102 K103 I104 E105 E106 E107 Q108 N109 K110 S111 K112 K113 K114 A115 Q116 Q117 A118 A119 A120 D121 T122 G123 N124 N125 S126 Q127 V128 S129 Q130 N131 Y132 HIS HIS HIS HIS HIS

4.2.8 Score per residue for model 8

- Molecule 1: Matrix protein p17

Chain A: 43% 26% • 23% •

G2 A3 R4 A5 S6 V7 L8 E12 W16 I19 R20 L21 K26 K27 Q28 Y29 K30 L31 K32 H33 W36 R39 E40 L41 E42 R43 V46 L50 L51 C57 R58 Q59 I60 Q65 P66 S67 L68 G71 L75 L78 I82 A83 V84 V88

R91 I92 D93 V94 K95 D96 A100 L101 D102 K103 E106 E107 Q108 N109 K110 S111 K112 K113 K114 A115 Q116 Q117 A118 A119 A120 D121 T122 G123 N124 N125 S126 Q127 V128 S129 Q130 N131 Y132 HIS HIS HIS HIS HIS

4.2.9 Score per residue for model 9

- Molecule 1: Matrix protein p17

Chain A: 45% 26% • 23% •

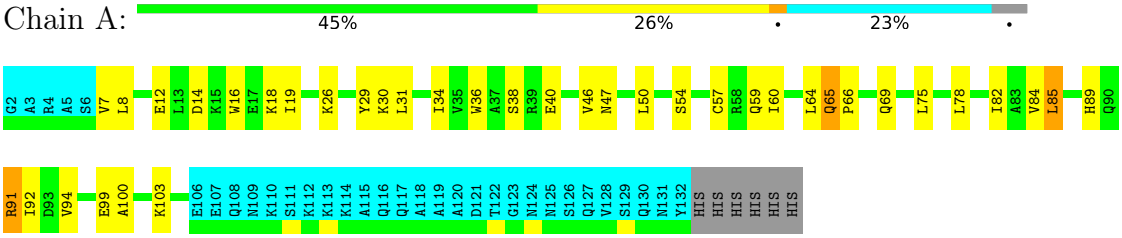
G2 A3 R4 A5 S6 V7 L8 E12 K15 W16 I19 R20 L21 K27 Q28 Y29 H33 W36 E40 L41 E42 V46 L50 L51 S54 C57 R58 Q59 I60 L64 Q65 P66 S67 L68 L75 L78 I82 A83 V84 L85 V88 R91 I92

D93 V94 K95 D96 T97 K98 E99 A100 L101 D102 K103 E106 E107 Q108 N109 K110 S111 K112 K113 K114 A115 Q116 Q117 A118 A119 A120 D121 T122 G123 N124 N125 S126 Q127 V128 S129 Q130 N131 Y132 HIS HIS HIS HIS HIS

4.2.10 Score per residue for model 10

- Molecule 1: Matrix protein p17

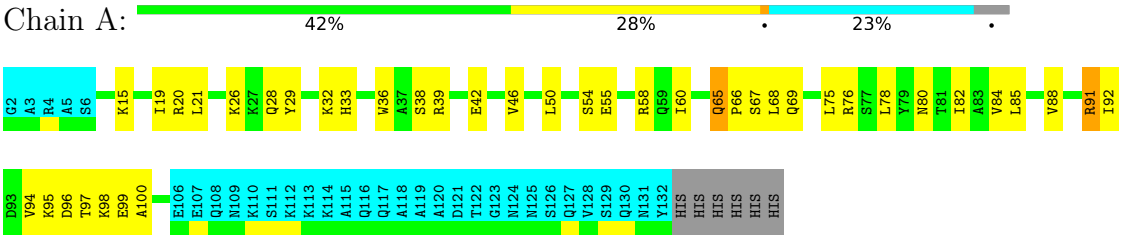
Chain A:



4.2.11 Score per residue for model 11

- Molecule 1: Matrix protein p17

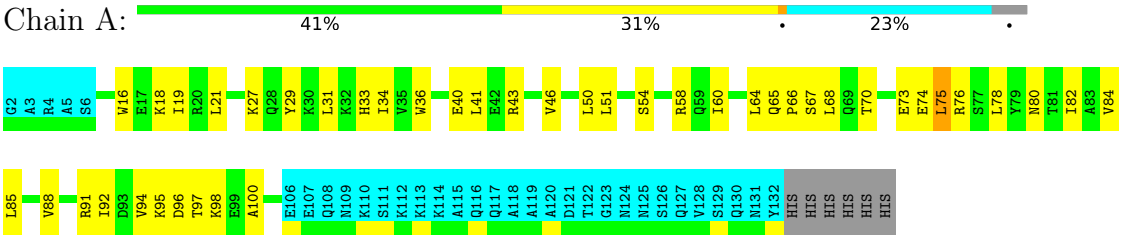
Chain A:



4.2.12 Score per residue for model 12

- Molecule 1: Matrix protein p17

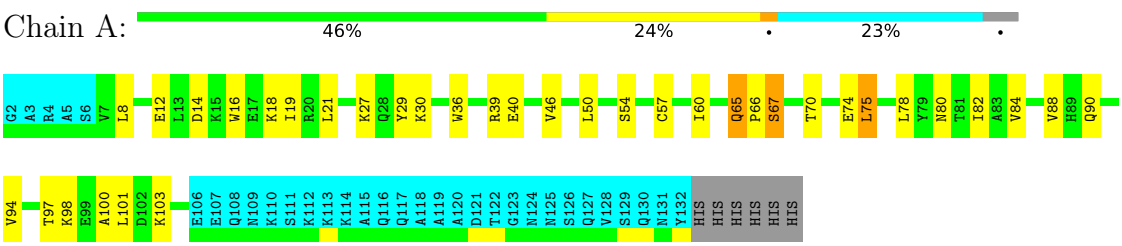
Chain A:



4.2.13 Score per residue for model 13 (medoid)

- Molecule 1: Matrix protein p17

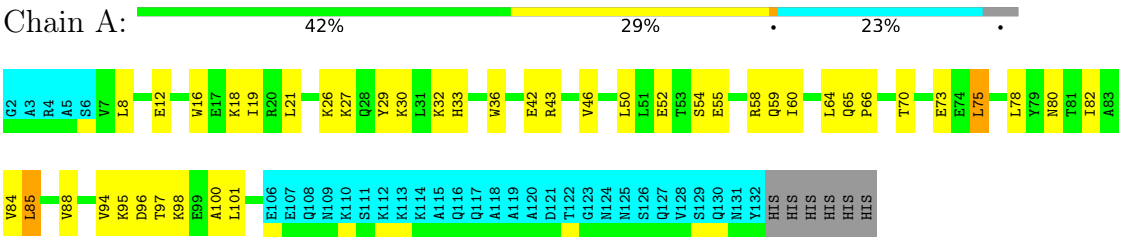
Chain A:



4.2.14 Score per residue for model 14

- Molecule 1: Matrix protein p17

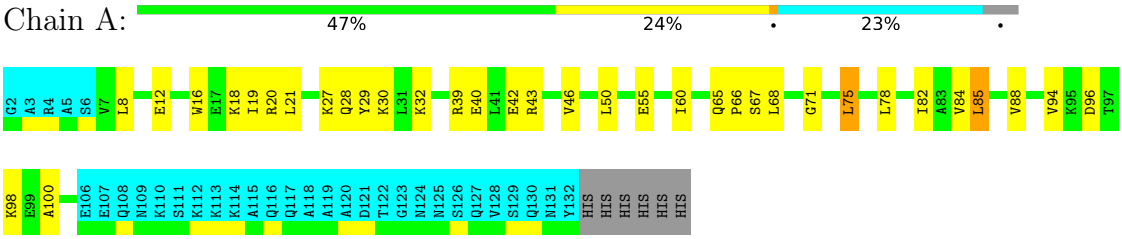
Chain A:



4.2.15 Score per residue for model 15

- Molecule 1: Matrix protein p17

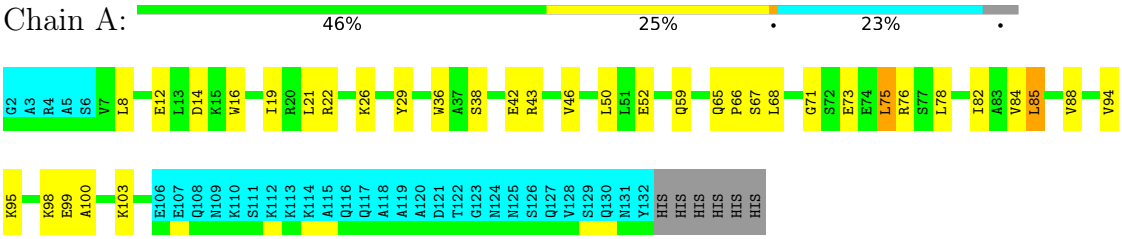
Chain A:



4.2.16 Score per residue for model 16

- Molecule 1: Matrix protein p17

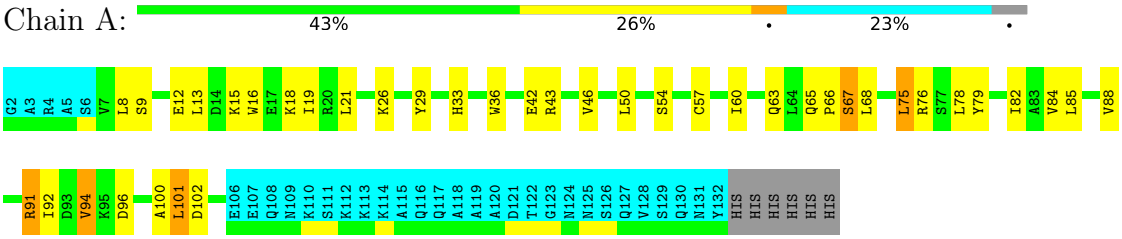
Chain A:



4.2.17 Score per residue for model 17

- Molecule 1: Matrix protein p17

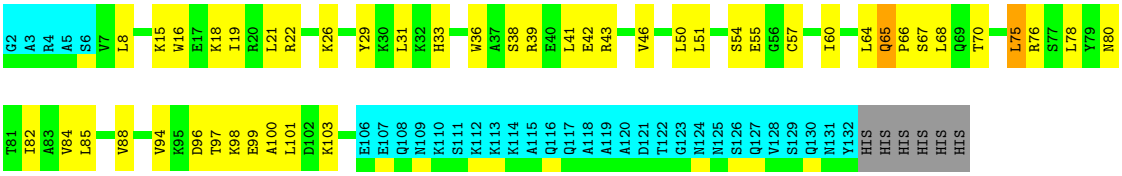
Chain A:



4.2.18 Score per residue for model 18

- Molecule 1: Matrix protein p17

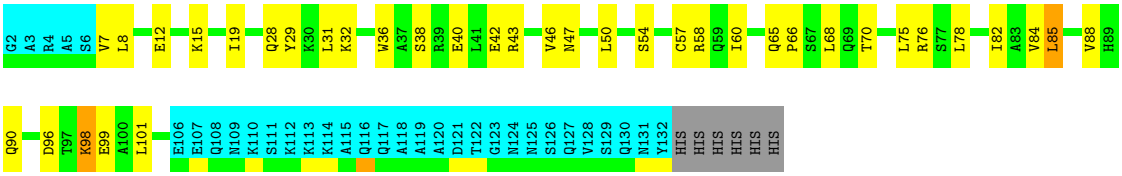
Chain A: 39% 32% • 23% •



4.2.19 Score per residue for model 19

- Molecule 1: Matrix protein p17

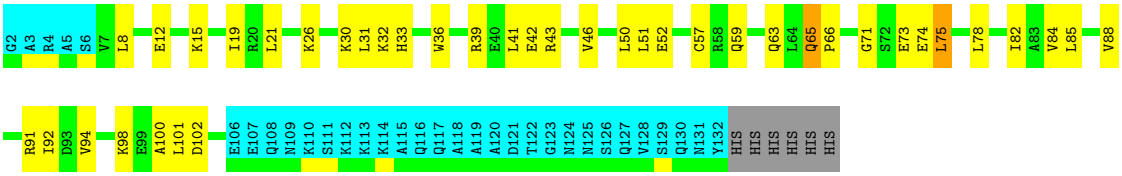
Chain A: 45% 26% • 23% •



4.2.20 Score per residue for model 20

- Molecule 1: Matrix protein p17

Chain A: 43% 28% • 23% •



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *simulated annealing, torsion angle dynamics.*

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy.*

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

Software name	**Classification**	**Version**
CYANA	structure solution	2.1
CYANA	refinement	2.1

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	1445
Number of shifts mapped to atoms	1445
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: PC8

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	796	822	821	22±3
2	A	34	48	48	0±1
All	All	16600	17400	17380	443

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 13.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:78:LEU:HD12	1:A:82:ILE:HD11	0.77	1.56	12	19
1:A:97:THR:HG22	1:A:101:LEU:HD11	0.76	1.58	7	1
1:A:94:VAL:HG21	1:A:100:ALA:HB2	0.72	1.62	9	14
1:A:29:TYR:CE2	1:A:84:VAL:HG21	0.68	2.24	8	18
1:A:46:VAL:HG13	1:A:50:LEU:HD12	0.67	1.63	18	20
1:A:60:ILE:HG21	1:A:82:ILE:HD13	0.67	1.65	7	5
1:A:41:LEU:HD12	1:A:51:LEU:HD11	0.66	1.66	8	7
1:A:78:LEU:HD23	2:A:300:PC8:H48	0.65	1.66	7	1
1:A:21:LEU:HD21	1:A:33:HIS:CE1	0.65	2.26	20	14
1:A:16:TRP:CZ3	1:A:85:LEU:HD12	0.64	2.27	10	3
1:A:84:VAL:O	1:A:88:VAL:HG23	0.62	1.95	6	18

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:50:LEU:HB2	1:A:60:ILE:HD11	0.62	1.69	9	4
1:A:19:ILE:HG21	1:A:84:VAL:CG1	0.62	2.24	20	16
1:A:46:VAL:CG1	1:A:50:LEU:HD12	0.62	2.25	18	20
1:A:41:LEU:HD11	1:A:51:LEU:HD21	0.61	1.71	18	1
1:A:60:ILE:CG2	1:A:82:ILE:HD13	0.61	2.25	7	17
1:A:98:LYS:HA	1:A:101:LEU:HD12	0.61	1.72	7	3
1:A:92:ILE:HG22	1:A:94:VAL:HG13	0.61	1.73	3	10
1:A:94:VAL:HG11	1:A:100:ALA:HA	0.59	1.75	20	8
1:A:21:LEU:HD23	1:A:29:TYR:CZ	0.58	2.34	15	13
1:A:8:LEU:HD22	1:A:12:GLU:HB3	0.58	1.74	20	12
1:A:80:ASN:OD1	1:A:97:THR:HG21	0.57	1.99	3	2
1:A:16:TRP:HA	1:A:19:ILE:HD12	0.57	1.76	8	10
1:A:80:ASN:CG	1:A:97:THR:HG21	0.57	2.19	14	3
1:A:80:ASN:HB3	1:A:97:THR:HG21	0.57	1.77	11	2
1:A:71:GLY:HA2	1:A:75:LEU:HD13	0.56	1.77	3	1
1:A:68:LEU:HD11	1:A:79:TYR:CD2	0.56	2.35	7	1
1:A:64:LEU:HD11	2:A:300:PC8:C24	0.56	2.31	7	1
1:A:64:LEU:HD22	1:A:75:LEU:HG	0.56	1.77	2	6
1:A:36:TRP:CZ3	1:A:78:LEU:HD22	0.55	2.36	13	18
1:A:19:ILE:HD13	1:A:84:VAL:CG1	0.55	2.31	20	8
1:A:71:GLY:HA3	1:A:75:LEU:HD22	0.55	1.77	8	5
1:A:67:SER:HB3	1:A:75:LEU:HD21	0.55	1.78	8	3
1:A:13:LEU:O	1:A:13:LEU:HD23	0.55	2.01	6	2
1:A:97:THR:HG22	1:A:101:LEU:CD1	0.55	2.29	7	1
1:A:7:VAL:HG11	1:A:85:LEU:HD21	0.55	1.79	19	1
1:A:19:ILE:HG21	1:A:84:VAL:HG11	0.54	1.78	10	6
1:A:21:LEU:HD23	1:A:29:TYR:CE1	0.52	2.39	3	13
1:A:92:ILE:CG2	1:A:94:VAL:HG13	0.51	2.35	8	5
1:A:64:LEU:HD11	2:A:300:PC8:H48	0.51	1.80	9	2
1:A:41:LEU:HD12	1:A:51:LEU:CD1	0.51	2.35	8	3
1:A:8:LEU:HD21	1:A:85:LEU:CD1	0.50	2.36	1	2
1:A:60:ILE:HG22	1:A:82:ILE:HD13	0.50	1.83	6	9
1:A:65:GLN:HA	1:A:68:LEU:HD12	0.50	1.83	18	12
1:A:67:SER:CB	1:A:75:LEU:HD21	0.50	2.36	16	2
1:A:50:LEU:CB	1:A:60:ILE:HD11	0.49	2.37	19	3
1:A:8:LEU:HD13	1:A:16:TRP:HB2	0.49	1.84	2	12
1:A:67:SER:HB3	1:A:75:LEU:HD11	0.49	1.84	4	2
1:A:68:LEU:HD11	1:A:79:TYR:CE2	0.49	2.42	7	1
1:A:19:ILE:HD13	1:A:84:VAL:HG13	0.48	1.85	18	4
1:A:65:GLN:CB	1:A:66:PRO:CD	0.48	2.91	16	20
1:A:94:VAL:HG11	1:A:100:ALA:CA	0.48	2.38	20	4

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:31:LEU:HD13	1:A:34:ILE:HD12	0.47	1.85	5	2
1:A:79:TYR:CD2	1:A:101:LEU:HD21	0.47	2.44	17	1
1:A:51:LEU:HD22	1:A:85:LEU:HD22	0.47	1.87	9	1
1:A:19:ILE:HG21	1:A:84:VAL:HG13	0.47	1.84	18	10
1:A:8:LEU:HD23	1:A:89:HIS:CE1	0.47	2.44	10	1
1:A:8:LEU:HD11	1:A:85:LEU:CD1	0.47	2.40	20	1
1:A:8:LEU:HD11	1:A:85:LEU:HD12	0.46	1.86	17	1
1:A:80:ASN:CB	1:A:97:THR:HG21	0.46	2.41	5	1
1:A:7:VAL:HG21	1:A:34:ILE:CG2	0.45	2.41	10	1
1:A:16:TRP:CH2	1:A:85:LEU:HD12	0.45	2.47	10	2
1:A:64:LEU:HD11	2:A:300:PC8:H46	0.45	1.88	7	1
1:A:79:TYR:CE2	1:A:101:LEU:HD22	0.44	2.47	7	1
1:A:8:LEU:HD22	1:A:12:GLU:CB	0.44	2.42	9	1
1:A:12:GLU:HB3	1:A:88:VAL:HG11	0.43	1.90	17	4
1:A:67:SER:HB2	1:A:75:LEU:HD21	0.43	1.90	16	1
1:A:64:LEU:HD21	2:A:300:PC8:H41	0.43	1.90	14	1
1:A:21:LEU:HD21	1:A:33:HIS:NE2	0.43	2.29	17	1
1:A:71:GLY:CA	1:A:75:LEU:HD13	0.43	2.43	3	1
1:A:57:CYS:HA	1:A:60:ILE:HD12	0.42	1.91	18	1
1:A:31:LEU:HD12	1:A:34:ILE:HD12	0.42	1.91	12	1
1:A:19:ILE:HG23	1:A:96:ASP:N	0.41	2.31	8	1
1:A:13:LEU:HD23	1:A:13:LEU:O	0.41	2.16	17	1
1:A:65:GLN:N	1:A:66:PRO:HD2	0.41	2.31	8	17
1:A:79:TYR:HD2	1:A:101:LEU:HD21	0.41	1.74	17	1
1:A:8:LEU:HD11	1:A:16:TRP:CE3	0.40	2.51	15	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	99/137 (72%)	94±2 (95±2%)	5±2 (5±2%)	1±1 (1±1%)	26	73
All	All	1980/2740 (72%)	1875 (95%)	92 (5%)	13 (1%)	26	73

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	91	ARG	10
1	A	94	VAL	3

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	87/117 (74%)	69±3 (79±4%)	18±3 (21±4%)	3 32
All	All	1740/2340 (74%)	1379 (79%)	361 (21%)	3 32

All 47 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	75	LEU	16
1	A	18	LYS	14
1	A	57	CYS	14
1	A	96	ASP	14
1	A	42	GLU	14
1	A	98	LYS	14
1	A	43	ARG	13
1	A	54	SER	13
1	A	26	LYS	11
1	A	103	LYS	10
1	A	76	ARG	10
1	A	58	ARG	9
1	A	95	LYS	9
1	A	30	LYS	9
1	A	39	ARG	9
1	A	67	SER	9
1	A	99	GLU	9
1	A	101	LEU	9
1	A	38	SER	8
1	A	32	LYS	8
1	A	91	ARG	8
1	A	27	LYS	8
1	A	40	GLU	8
1	A	73	GLU	8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	85	LEU	8
1	A	65	GLN	7
1	A	15	LYS	7
1	A	22	ARG	6
1	A	63	GLN	6
1	A	70	THR	6
1	A	55	GLU	6
1	A	59	GLN	6
1	A	31	LEU	6
1	A	14	ASP	5
1	A	47	ASN	5
1	A	90	GLN	5
1	A	52	GLU	5
1	A	74	GLU	5
1	A	28	GLN	5
1	A	20	ARG	4
1	A	102	ASP	4
1	A	69	GLN	4
1	A	105	GLU	2
1	A	77	SER	2
1	A	72	SER	1
1	A	93	ASP	1
1	A	9	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)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
2	PC8	A	300	-	33,33,33	0.97±0.00	1±0 (3±0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of angles that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
2	PC8	A	300	-	39,41,41	1.13±0.00	3±0 (7±0%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	PC8	A	300	-	-	0±0,37,37,37	-

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
2	A	300	PC8	P1-O2	2.26	1.58	1.50	11	20

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
2	A	300	PC8	C7-O7-C17	4.26	128.28	117.79	11	20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
2	A	300	PC8	O7-C17-C18	2.98	117.93	111.50	13	20
2	A	300	PC8	C8-O5-C9	2.56	126.61	117.12	15	20

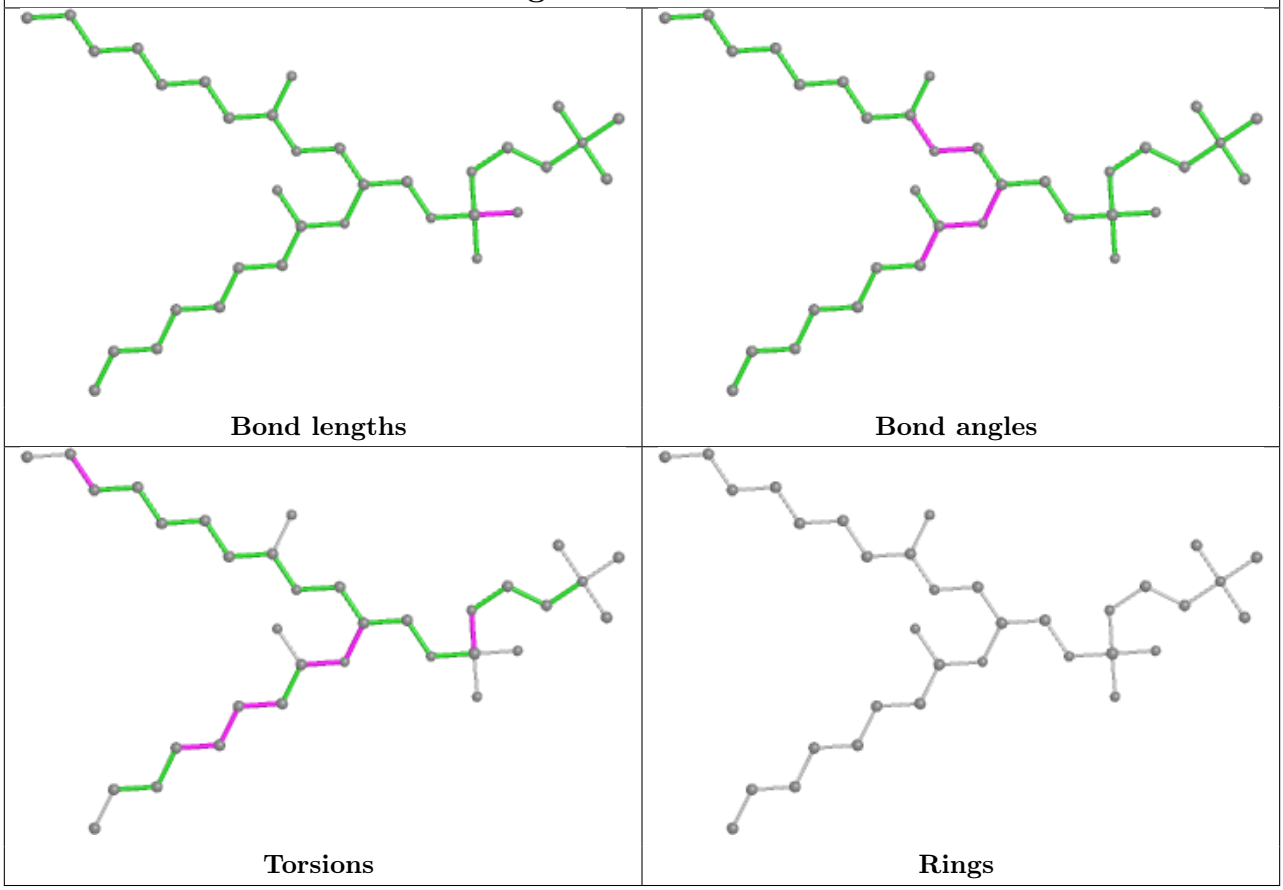
There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand PC8 A 300



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)

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 (i)

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	1445
Number of shifts mapped to atoms	1445
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 (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	131	-0.46 ± 0.12	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	119	0.08 ± 0.09	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	128	0.45 ± 0.26	None needed (< 0.5 ppm)

7.1.3 Completeness of resonance assignments (i)

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. 1159 atoms were assigned a chemical shift out of a possible 1411. 0 out of 20 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	398/497 (80%)	203/203 (100%)	99/198 (50%)	96/96 (100%)
Sidechain	702/838 (84%)	476/541 (88%)	225/259 (87%)	1/38 (3%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	59/76 (78%)	30/37 (81%)	27/34 (79%)	2/5 (40%)
Overall	1159/1411 (82%)	709/781 (91%)	351/491 (71%)	99/139 (71%)

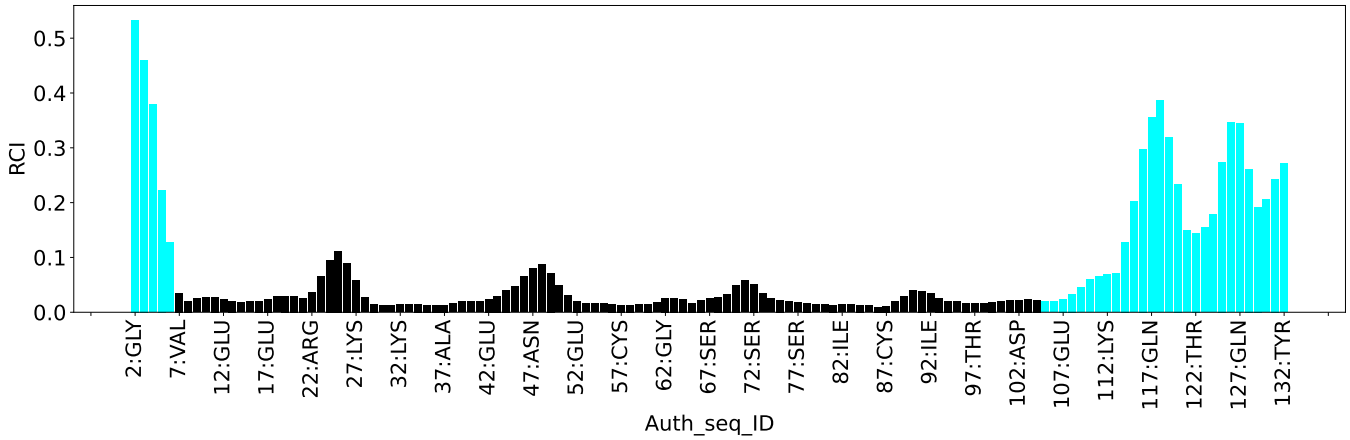
The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 80%, i.e. 1445 atoms were assigned a chemical shift out of a possible 1803. 0 out of 21 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	526/659 (80%)	267/269 (99%)	131/262 (50%)	128/128 (100%)
Sidechain	856/1059 (81%)	581/679 (86%)	274/326 (84%)	1/54 (2%)
Aromatic	63/85 (74%)	32/41 (78%)	29/39 (74%)	2/5 (40%)
Overall	1445/1803 (80%)	880/989 (89%)	434/627 (69%)	131/187 (70%)

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.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: