wwPDB NMR Structure Validation Summary Report (i)

Apr 21, 2024 – 03:36 AM EDT

PDB ID	:	2LQD
BMRB ID	:	17415
Title	:	Reduced and CO-bound cytochrome P450cam (CYP101A1)
Authors	:	Pochapsky, T.C.; Pochapsky, S.S.; Asciutto, E.; Madura, J.; Young, M.J.
Deposited on	:	2012-03-02

This is a wwPDB NMR Structure Validation Summary 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:

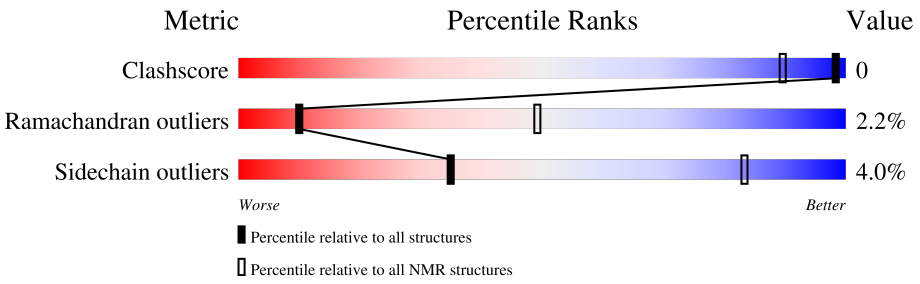
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (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 21%.

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	413	72% 23% • •

2 Ensemble composition and analysis (i)

This entry contains 1 models. Identification of well-defined residues and clustering analysis are not possible.

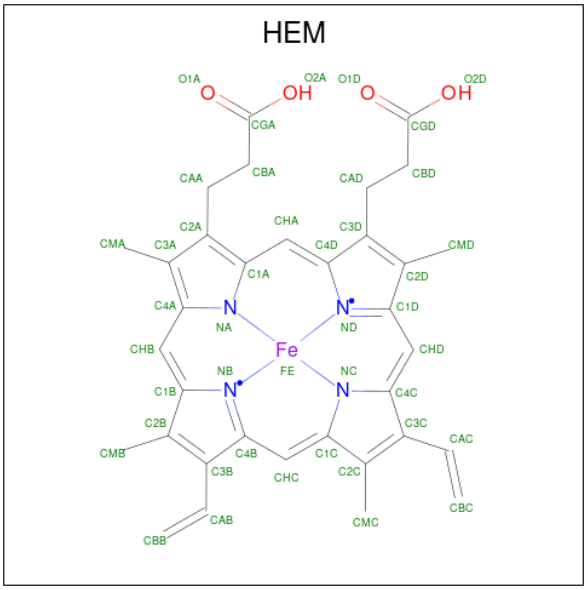
3 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6522 atoms, of which 3241 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Camphor 5-monooxygenase.

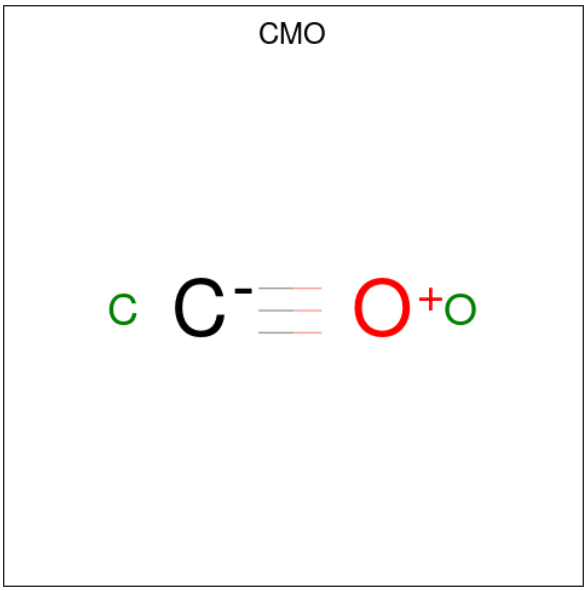
Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	405	6365	2033	3157	560	597	18	0

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



Mol	Chain	Residues	Atoms					
			Total	C	Fe	H	N	O
2	A	1	73	34	1	30	4	4

- Molecule 3 is CARBON MONOXIDE (three-letter code: CMO) (formula: CO).



Mol	Chain	Residues	Atoms		
			Total	C	O
3	A	1	2	1	1

- Molecule 4 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	
			Total	K
4	A	1	1	1

- Molecule 5 is water.

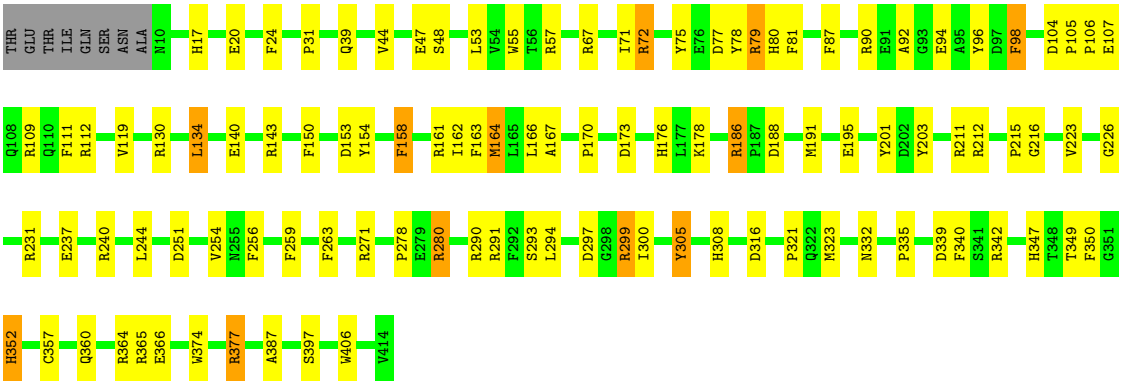
Mol	Chain	Residues	Atoms		
			Total	H	O
5	A	27	81	54	27

4 Residue-property plots (i)

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: Camphor 5-monooxygenase

Chain A: 72% 23%



5 Refinement protocol and experimental data overview (i)

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

Of the 100 calculated structures, 1 were deposited, based on the following criterion: *structures with the least restraint violations*.

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

Software name	Classification	Version
Amber	refinement	10

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	1189
Number of shifts mapped to atoms	1157
Number of unparsed shifts	0
Number of shifts with mapping errors	32
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	21%

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: HEM, K, CMO

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	1.55	14/3287 (0.4%)	1.97	106/4465 (2.4%)
All	All	1.55	14/3287 (0.4%)	1.97	106/4465 (2.4%)

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	12
All	All	0	12

5 of 14 bond outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	215	PRO	N-CD	-7.14	1.37	1.47
1	A	305	TYR	CE2-CZ	7.09	1.47	1.38
1	A	237	GLU	CD-OE1	6.22	1.32	1.25
1	A	78	TYR	CG-CD2	6.07	1.47	1.39
1	A	170	PRO	N-CD	-6.03	1.39	1.47

5 of 106 angle outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	130	ARG	NE-CZ-NH1	20.73	130.67	120.30
1	A	365	ARG	NE-CZ-NH1	12.94	126.77	120.30
1	A	67	ARG	NE-CZ-NH2	-12.35	114.13	120.30
1	A	161	ARG	NE-CZ-NH1	11.51	126.06	120.30
1	A	143	ARG	NE-CZ-NH1	11.36	125.98	120.30

There are no chirality outliers.

5 of 12 planar outliers are listed below.

Mol	Chain	Res	Type	Group
1	A	17	HIS	Mainchain
1	A	72	ARG	Sidechain
1	A	79	ARG	Sidechain
1	A	90	ARG	Sidechain
1	A	96	TYR	Sidechain

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	3208	3157	3156	1
All	All	3281	3241	3186	1

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 0.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)
1:A:158:PHE:CE1	1:A:162:ILE:HD11	0.45	2.47

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	403/413 (98%)	355 (88%)	39 (10%)	9 (2%)	10 49
All	All	403/413 (98%)	355 (88%)	39 (10%)	9 (2%)	10 49

5 of 9 Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in

the ensemble.

Mol	Chain	Res	Type
1	A	119	VAL
1	A	188	ASP
1	A	294	LEU
1	A	308	HIS
1	A	335	PRO

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	350/357 (98%)	336 (96%)	14 (4%)	35	83
All	All	350/357 (98%)	336 (96%)	14 (4%)	35	83

5 of 14 residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	A	39	GLN
1	A	44	VAL
1	A	53	LEU
1	A	164	MET
1	A	178	LYS

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)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 for Mogul analysis.

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	HEM	A	501	1,3	41,50,50	1.56	9 (21%)
3	CMO	A	502	2	0,1,1	0.00	-

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	HEM	A	501	1,3	45,82,82	1.98	15 (33%)
3	CMO	A	502	2	-	-	-

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	HEM	A	501	1,3	-	0,12,54,54	-

5 of 9 bond outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	HEM	C4D-C3D	3.75	1.51	1.45
2	A	501	HEM	C3C-C2C	3.53	1.35	1.40
2	A	501	HEM	C1D-ND	2.99	1.45	1.38
2	A	501	HEM	C3D-C2D	2.91	1.42	1.36

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	HEM	C3C-CAC	2.48	1.52	1.47

5 of 15 angle outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	501	HEM	C3B-C2B-C1B	5.08	110.25	106.49
2	A	501	HEM	CAD-CBD-CGD	4.30	122.86	113.60
2	A	501	HEM	C4B-C3B-C2B	4.29	103.70	107.11
2	A	501	HEM	C4B-CHC-C1C	4.16	128.04	122.56
2	A	501	HEM	C4C-CHD-C1D	3.55	127.24	122.56

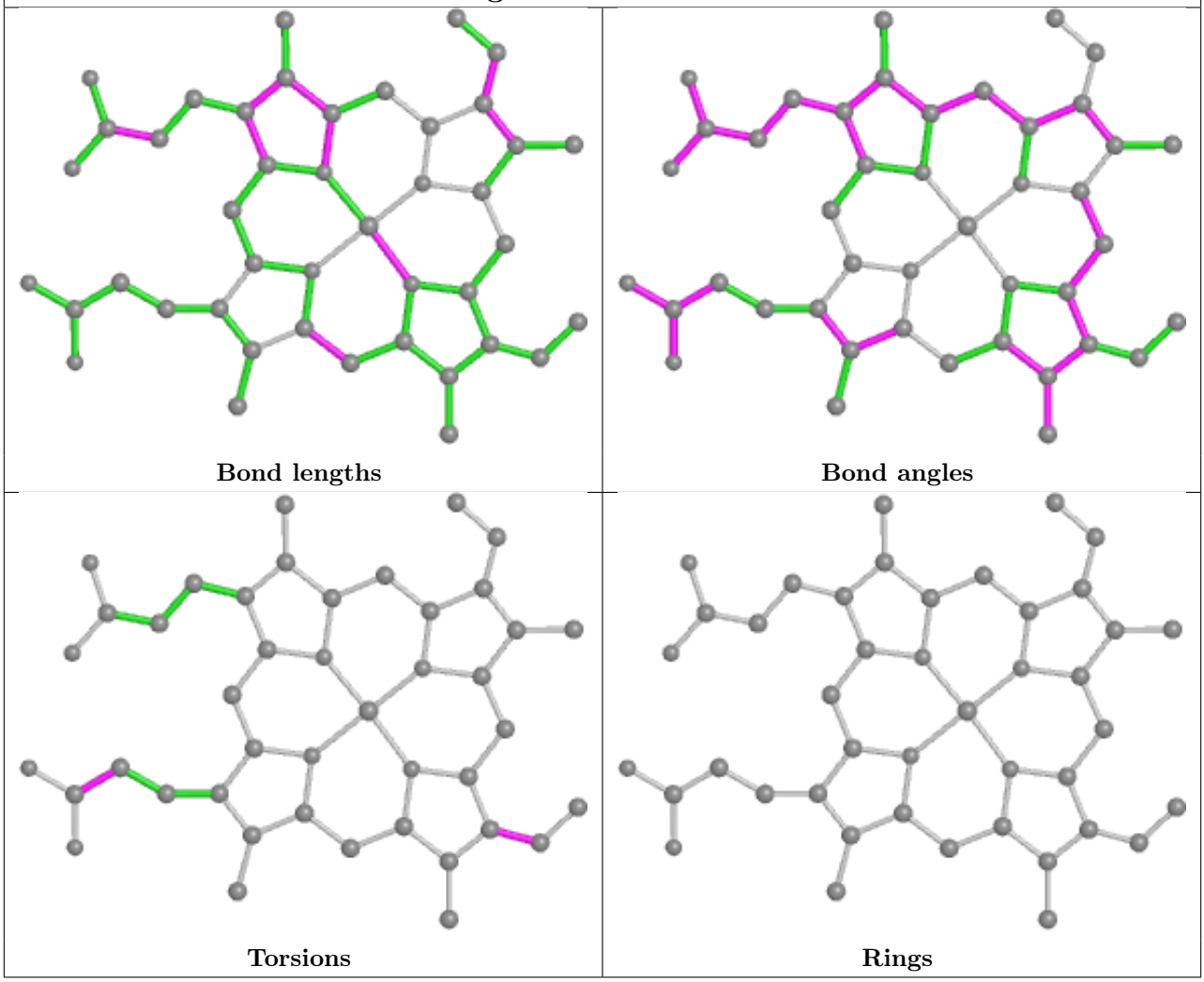
There are no chirality outliers.

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 HEM A 501



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 21% for the well-defined parts and 21% 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	1189
Number of shifts mapped to atoms	1157
Number of unparsed shifts	0
Number of shifts with mapping errors	32
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	3

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. First 5 (of 32) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	2	THR	H	8.2	0.02	1
1	A	2	THR	CA	60.7	0.1	1
1	A	2	THR	CB	68.4	0.1	1
1	A	2	THR	N	117.1	0.1	1
1	A	3	GLU	H	8.16	0.02	1
1	A	3	GLU	CA	55.2	0.1	1
1	A	3	GLU	CB	28.6	0.1	1
1	A	3	GLU	N	116.9	0.1	1
1	A	4	THR	H	8.12	0.02	1
1	A	4	THR	CA	60.7	0.1	1
1	A	4	THR	CB	68.4	0.1	1
1	A	4	THR	N	120.5	0.1	1
1	A	5	ILE	H	8.15	0.02	1
1	A	5	ILE	CA	60.2	0.1	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	5	ILE	CB	36.8	0.1	1
1	A	5	ILE	N	124.9	0.1	1
1	A	6	GLN	H	8.39	0.02	1
1	A	6	GLN	CA	54.5	0.1	1
1	A	6	GLN	CB	28.0	0.1	1
1	A	6	GLN	N	125.5	0.1	1
1	A	7	SER	H	8.29	0.02	1
1	A	7	SER	CA	57.2	0.1	1
1	A	7	SER	CB	62.7	0.1	1
1	A	7	SER	N	118.4	0.1	1
1	A	8	ASN	H	8.35	0.02	1
1	A	8	ASN	CA	51.9	0.1	1
1	A	8	ASN	CB	37.3	0.1	1
1	A	8	ASN	N	121.7	0.1	1
1	A	9	ALA	H	8.1	0.02	1
1	A	9	ALA	CA	51.1	0.1	1
1	A	9	ALA	CB	17.1	0.1	1
1	A	9	ALA	N	124.3	0.1	1

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	330	0.98 ± 0.13	Should be checked
$^{13}C_{\beta}$	266	1.99 ± 0.11	Should be checked
$^{13}C'$	0	—	None (insufficient data)
^{15}N	292	0.33 ± 0.17	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 21%, i.e. 1157 atoms were assigned a chemical shift out of a possible 5573. 0 out of 66 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

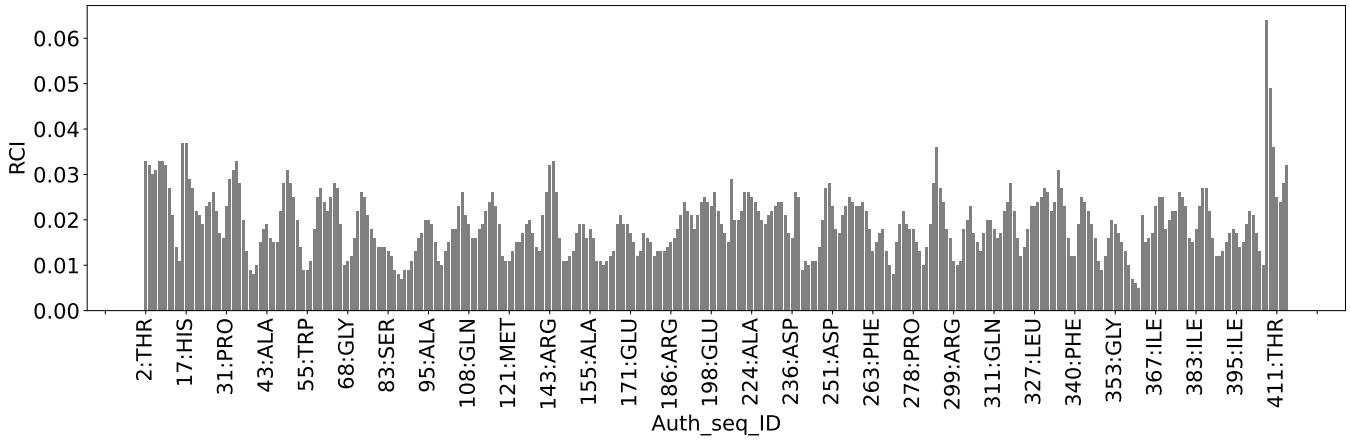
	Total	^{1}H	^{13}C	^{15}N
Backbone	893/1990 (45%)	287/805 (36%)	322/810 (40%)	284/375 (76%)
Sidechain	258/3171 (8%)	0/2065 (0%)	258/982 (26%)	0/124 (0%)
Aromatic	6/412 (1%)	3/208 (1%)	0/186 (0%)	3/18 (17%)
Overall	1157/5573 (21%)	290/3078 (9%)	580/1978 (29%)	287/517 (56%)

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	358	LEU	H	2.74	5.09 – 11.34	-8.8
1	A	357	CYS	H	4.04	5.02 – 11.74	-6.5
1	A	349	THR	H	11.82	5.19 – 11.27	5.9

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: