WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB NMR Structure Validation Summary Report (i)

Aug 10, 2021 – 12:06 PM EDT

PDB ID : 2LP8
Title : SOLUTION STRUCTURE OF AN APOPTOSIS ACTIVATING PHOTO-SWITCHABLE BAK PEPTIDE BOUND to BCL-XL
Authors : Wysoczanski, P.; Mart, R.J.; Loveridge, J.E.; Williams, C.; Whittaker, S.B.-M.; Crump, M.P.; Allemann, R.K.
Deposited on : 2012-02-03

The following versions of software and data (see references (i)) were used in the production of this report:

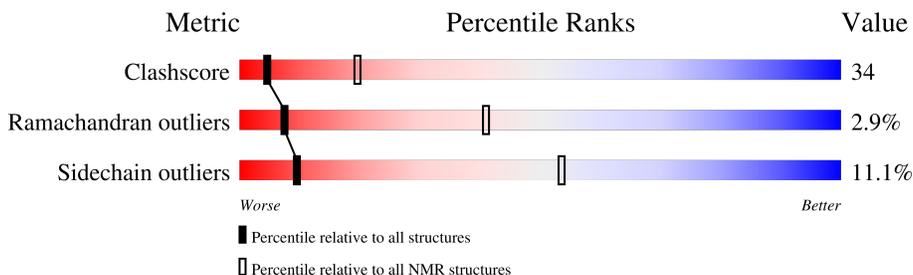|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.23.1
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.1

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 78%.

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	Percentile Ranks	Value
Clashscore		34
Ramachandran outliers		2.9%
Sidechain outliers		11.1%

Worse　　　*Better*

▮ Percentile relative to all structures

▯ Percentile relative to all NMR structures

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	185	34% · 44% · 21%
2	B	18	22% · 33% · 44%

2　Ensemble composition and analysis ⓘ

This entry contains 21 models. Model 21 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:3-A:32,　　　A:41-A:157, B:1005-B:1014 (157)	0.36	21

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 2 clusters and 3 single-model clusters were found.

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

3 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 3198 atoms, of which 1536 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Bcl-2-like protein 1.

Mol	Chain	Residues	Atoms						Trace
1	A	185	Total	C	H	N	O	S	0
			2928	949	1411	274	290	4	

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	170	LEU	-	expression tag	UNP Q07817
A	171	GLU	-	expression tag	UNP Q07817
A	172	HIS	-	expression tag	UNP Q07817
A	173	HIS	-	expression tag	UNP Q07817
A	174	HIS	-	expression tag	UNP Q07817
A	175	HIS	-	expression tag	UNP Q07817
A	176	HIS	-	expression tag	UNP Q07817
A	177	HIS	-	expression tag	UNP Q07817
A	178	LEU	-	expression tag	UNP Q07817
A	179	GLU	-	expression tag	UNP Q07817
A	180	HIS	-	expression tag	UNP Q07817
A	181	HIS	-	expression tag	UNP Q07817
A	182	HIS	-	expression tag	UNP Q07817
A	183	HIS	-	expression tag	UNP Q07817
A	184	HIS	-	expression tag	UNP Q07817
A	185	HIS	-	expression tag	UNP Q07817

- Molecule 2 is a protein called Bcl-2 homologous antagonist/killer.

Mol	Chain	Residues	Atoms						Trace
2	B	18	Total	C	H	N	O	S	1
			228	69	113	24	20	2	

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1001	ACE	-	acetylation	UNP Q16611
B	1003	CYS	GLN	engineered mutation	UNP Q16611
B	1007	ALA	GLN	engineered mutation	UNP Q16611
B	1010	ALA	ILE	engineered mutation	UNP Q16611

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1011	PHE	ILE	engineered mutation	UNP Q16611
B	1014	CYS	ASP	engineered mutation	UNP Q16611
B	1018	NH2	-	amidation	UNP Q16611

- Molecule 3 is 3,3'-(E)-diazene-1,2-diylbis{6-[(chloroacetyl)amino]benzenesulfonic acid} (three-letter code: 33B) (formula: $C_{16}H_{14}Cl_2N_4O_8S_2$).
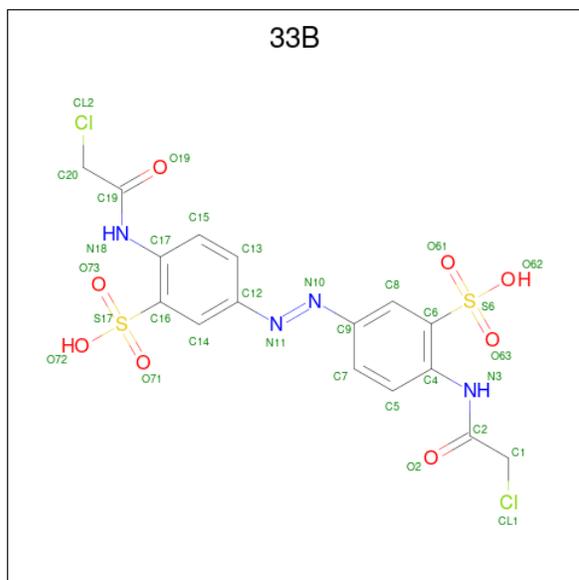

| Mol	Chain	Residues	Atoms					
			Total	C	H	N	O	S
3	B	1	42	16	12	4	8	2

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: Bcl-2-like protein 1
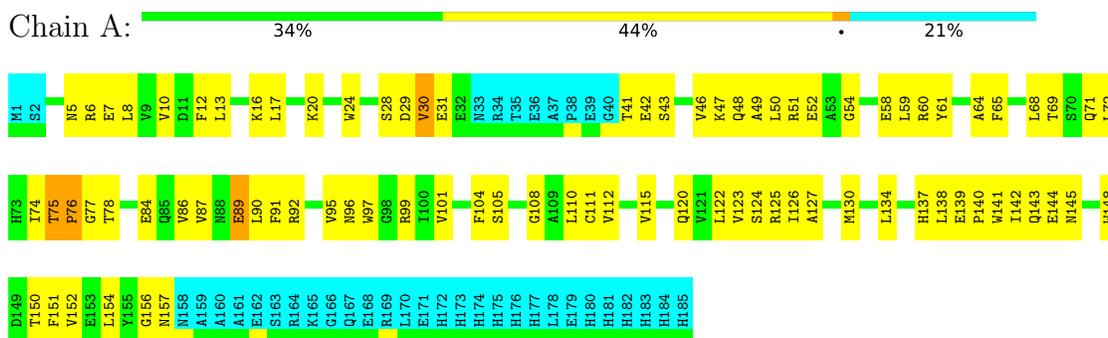
● Molecule 2: Bcl-2 homologous antagonist/killer
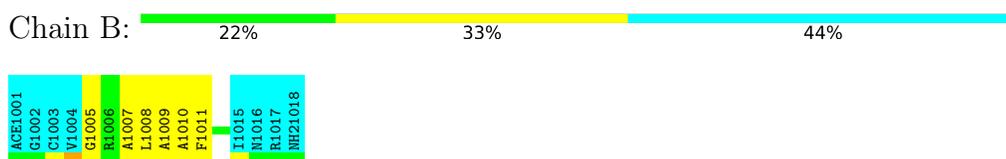
4.2 Residue scores for the representative (medoid) model from the NMR ensemble
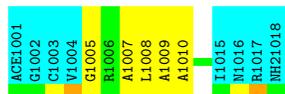
The representative model is number 21. Colouring as in section 4.1 above.

● Molecule 1: Bcl-2-like protein 1

• Molecule 2: Bcl-2 homologous antagonist/killer

Chain B: 28%   28%   44%

5 Refinement protocol and experimental data overview ⓘ

The models were refined using the following method: *simulated annealing.*

Of the 500 calculated structures, 21 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
CNS	refinement	
ARIA	structure solution	2.3
ARIA	refinement	2.3

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	2
Total number of shifts	1950
Number of shifts mapped to atoms	1950
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	78%

6 Model quality ⓘ

6.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: NH2, ACE, 33B

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	0.39±0.04	0±0/1227 (0.0± 0.0%)	0.51±0.04	0±0/1664 (0.0± 0.0%)
2	B	0.40±0.13	0±0/68 (0.0± 0.0%)	0.48±0.07	0±0/91 (0.0± 0.0%)
All	All	0.40	0/27195 (0.0%)	0.51	1/36855 (0.0%)

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	0.2±0.4
All	All	0	4

There are no bond-length outliers.

All unique angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	99	ARG	NE-CZ-NH1	6.17	123.39	120.30	18	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	155	TYR	Sidechain	2
1	A	61	TYR	Sidechain	1
1	A	99	ARG	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	1197	1124	1122	73±12
2	B	67	62	62	13±3
3	B	30	12	10	12±3
All	All	27174	25158	25074	1798

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 34.

5 of 508 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Models Total
1:A:89:GLU:HB2	2:B:1009:ALA:HB2	1.02	1.28	18	16
1:A:89:GLU:HB3	2:B:1009:ALA:HB2	0.99	1.30	2	3
1:A:86:VAL:HG13	2:B:1008:LEU:HG	0.92	1.41	2	3
1:A:13:LEU:HG	1:A:108:GLY:HA2	0.91	1.39	11	21
1:A:89:GLU:HG2	2:B:1005:GLY:HA2	0.91	1.39	20	7

6.3 Torsion angles ⓘ

6.3.1 Protein backbone ⓘ

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	147/185 (79%)	132±2 (90±1%)	11±2 (8±1%)	4±1 (3±1%)	8	43
2	B	10/18 (56%)	8±1 (82±10%)	1±1 (11±9%)	1±0 (6±5%)	3	19
All	All	3297/4263 (77%)	2939 (89%)	263 (8%)	95 (3%)	7	41

5 of 12 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	77	GLY	21
1	A	30	VAL	20
1	A	76	PRO	20
1	A	29	ASP	11
2	B	1005	GLY	11

6.3.2   Protein sidechains ⓘ

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	127/159 (80%)	113±2 (89±2%)	14±2 (11±2%)	9	52
2	B	5/10 (50%)	5±0 (96±8%)	0±0 (4±8%)	36	84
All	All	2772/3549 (78%)	2465 (89%)	307 (11%)	9	53

5 of 61 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	THR	21
1	A	110	LEU	21
1	A	138	LEU	14
1	A	92	ARG	14
1	A	17	LEU	12

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 ⓘ

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
3	33B	B	1100	2	29,31,33	1.54±0.05	4±0 (14±1%)

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
3	33B	B	1100	2	40,46,48	1.26±0.03	4±1 (11±3%)

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
3	33B	B	1100	2	-	0±0,25,25,29	0±0,2,2,2

All unique bond 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(Å)	Ideal(Å)	Models	
								Worst	Total
3	B	1100	33B	C17-N18	5.08	1.32	1.41	7	21
3	B	1100	33B	C4-N3	4.39	1.33	1.41	2	21
3	B	1100	33B	C9-N10	2.87	1.31	1.44	7	21
3	B	1100	33B	C12-N11	2.80	1.32	1.44	7	21

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
3	B	1100	33B	C1-C2	2.27	1.55	1.50	21	2

5 of 10 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
3	B	1100	33B	C4-N3-C2	3.98	117.75	127.49	2	21
3	B	1100	33B	C17-N18-C19	3.93	117.87	127.49	6	21
3	B	1100	33B	O61-S6-C6	2.67	111.03	106.51	4	5
3	B	1100	33B	C16-C17-N18	2.53	117.36	120.37	7	2
3	B	1100	33B	O71-S17-C16	2.42	110.60	106.51	9	13

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 33B B 1100

Bond lengths

Bond angles

Torsions

Rings
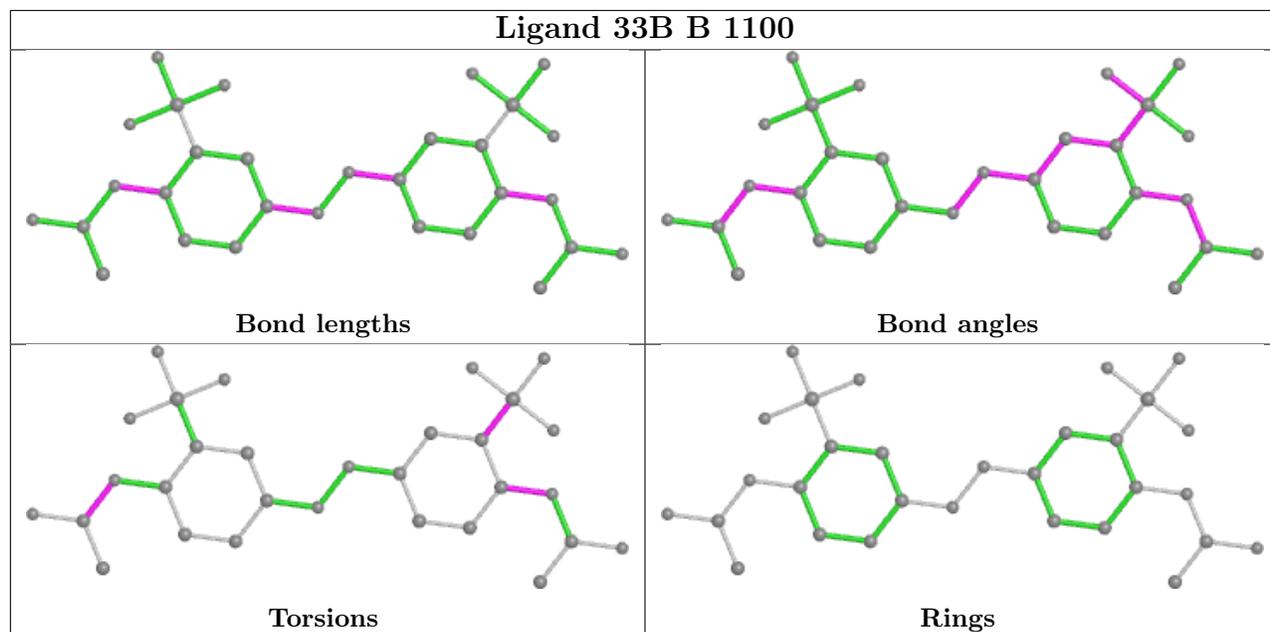
6.7 Other polymers ⓘ

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 78% for the well-defined parts and 69% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *BCL-XL*

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	1867
Number of shifts mapped to atoms	1867
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	5

7.1.2 Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_\alpha$	164	-0.79 ± 0.15	Should be applied
$^{13}C_\beta$	151	0.09 ± 0.04	None needed (< 0.5 ppm)
$^{13}C'$	152	-0.34 ± 0.09	None needed (< 0.5 ppm)
^{15}N	160	0.54 ± 0.25	Should be applied

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 76%, i.e. 1460 atoms were assigned a chemical shift out of a possible 1932. 1 out of 27 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	704/781 (90%)	283/312 (91%)	281/314 (89%)	140/155 (90%)
Sidechain	737/928 (79%)	452/541 (84%)	270/341 (79%)	15/46 (33%)

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	19/223 (9%)	9/117 (8%)	5/97 (5%)	5/9 (56%)
Overall	1460/1932 (76%)	744/970 (77%)	556/752 (74%)	160/210 (76%)

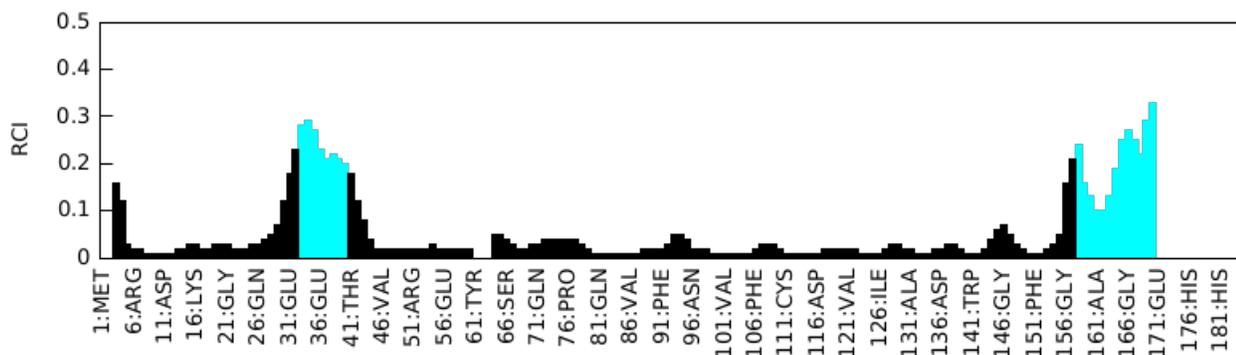### 7.1.4   Statistically unusual chemical shifts   ⓘ

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.

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	145	ASN	HB3	0.88	4.41 − 1.11	-5.7
1	A	95	VAL	HB	0.17	3.59 − 0.39	-5.7
1	A	95	VAL	HG11	-0.55	2.13 − -0.47	-5.3
1	A	95	VAL	HG12	-0.55	2.13 − -0.47	-5.3
1	A	95	VAL	HG13	-0.55	2.13 − -0.47	-5.3

7.1.5   Random Coil Index (RCI) plots   ⓘ

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:

7.2   Chemical shift list 2

File name: working_cs.cif

Chemical shift list name: *PHOTOSWITCHABLEBAK*

7.2.1   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	83
Number of shifts mapped to atoms	83
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

7.2.2   Chemical shift referencing ⓘ

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

7.2.3   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 2%, i.e. 38 atoms were assigned a chemical shift out of a possible 1932. 0 out of 27 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	17/781 (2%)	17/312 (5%)	0/314 (0%)	0/155 (0%)
Sidechain	16/928 (2%)	16/541 (3%)	0/341 (0%)	0/46 (0%)
Aromatic	5/223 (2%)	5/117 (4%)	0/97 (0%)	0/9 (0%)
Overall	38/1932 (2%)	38/970 (4%)	0/752 (0%)	0/210 (0%)

7.2.4   Statistically unusual chemical shifts ⓘ

There are no statistically unusual chemical shifts.

7.2.5   Random Coil Index (RCI) plots ⓘ

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 B: