WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB NMR Structure Validation Summary Report (i)

Sep 7, 2023 – 08:14 pm BST

PDB ID	:	5LCB
EMDB ID	:	EMD-4033
BMRB ID	:	34012
Title	:	In situ atomic-resolution structure of the baseplate antenna complex in Chlorobaculum tepidum obtained combining solid-state NMR spectroscopy, cryo electron microscopy and polarization spectroscopy
Authors	:	Nielsen, J.T.; Kulminskaya, N.V.; Bjerring, M.; Linnanto, J.M.; Ratsep, M.; Pedersen, M.; Lambrev, P.H.; Dorogi, M.; Garab, G.; Thomsen, K.; Jegerschold, C.; Frigaard, N.U.; Lindahl, M.; Nielsen, N.C.
Deposited on	:	2016-06-20

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

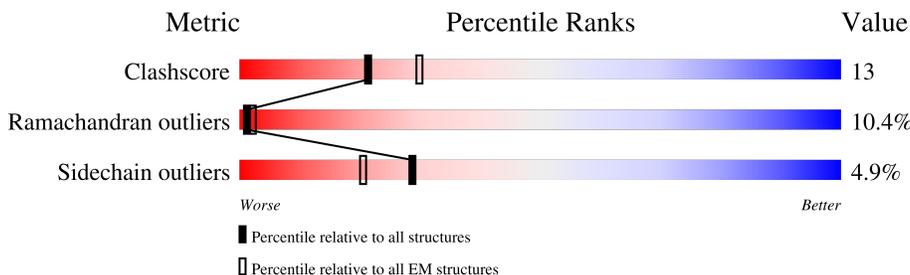|		
EMDB validation analysis	:	NOT EXECUTED
Mogul	:	1.8.4, CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	NOT EXECUTED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY, SOLID-STATE NMR

The reported resolution of this entry is 26.50 Å.

The overall completeness of chemical shifts assignment is 2%.

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		13
Ramachandran outliers		10.4%
Sidechain outliers		4.9%

Worse — *Better*

▮ Percentile relative to all structures

▯ Percentile relative to all EM structures

Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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	59	83% 15% •
1	B	59	64% 32% • •
1	C	59	69% 25% • •
1	D	59	69% 25% • •
1	E	59	64% 34% •
1	F	59	64% 29% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	59	63% 32% 5%
1	H	59	68% 27% 5%
1	I	59	76% 22% ·
1	J	59	68% 25% 7%
1	K	59	69% 27% · ·
1	L	59	66% 32% ·
1	M	59	64% 29% 7%
1	N	59	78% 20% ·

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA and RNA chains that are outliers for geometric criteria:

Mol	Chain	Compound	Res	Total models with violations	
				Chirality	Geometry
2	A	BCL	101	1	-
2	B	BCL	101	1	-
2	C	BCL	101	1	-
2	D	BCL	101	1	-
2	E	BCL	101	1	-
2	F	BCL	101	1	-
2	G	BCL	101	1	-
2	H	BCL	101	1	-
2	I	BCL	101	1	-
2	J	BCL	101	1	-
2	K	BCL	101	1	-
2	L	BCL	101	1	-
2	M	BCL	101	1	-
2	N	BCL	101	1	-

2   Ensemble composition and analysis ⓘ

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

3   Entity composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 13860 atoms, of which 6930 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Bacteriochlorophyll c-binding protein.

Mol	Chain	Residues	Atoms						AltConf	Trace
1	A	59	Total	C	H	N	O	S	0	
			850	266	421	77	81	5		
1	B	59	Total	C	H	N	O	S	0	
			850	266	421	77	81	5		
1	C	59	Total	C	H	N	O	S	0	
			850	266	421	77	81	5		
1	D	59	Total	C	H	N	O	S	0	
			850	266	421	77	81	5		
1	E	59	Total	C	H	N	O	S	0	
			850	266	421	77	81	5		
1	F	59	Total	C	H	N	O	S	0	
			850	266	421	77	81	5		
1	G	59	Total	C	H	N	O	S	0	
			850	266	421	77	81	5		
1	H	59	Total	C	H	N	O	S	0	
			850	266	421	77	81	5		
1	I	59	Total	C	H	N	O	S	0	
			850	266	421	77	81	5		
1	J	59	Total	C	H	N	O	S	0	
			850	266	421	77	81	5		
1	K	59	Total	C	H	N	O	S	0	
			850	266	421	77	81	5		
1	L	59	Total	C	H	N	O	S	0	
			850	266	421	77	81	5		
1	M	59	Total	C	H	N	O	S	0	
			850	266	421	77	81	5		
1	N	59	Total	C	H	N	O	S	0	
			850	266	421	77	81	5		

- Molecule 2 is BACTERIOCHLOROPHYLL A (three-letter code: BCL) (formula: $C_{55}H_{74}MgN_4O_6$).

BCL
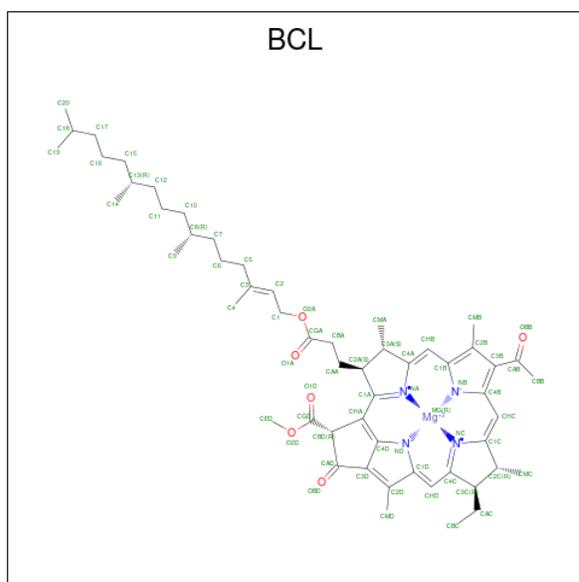
Mol	Chain	Residues	Atoms						AltConf
2	A	1	Total	C	H	Mg	N	O	
			140	55	74	1	4	6	
2	B	1	Total	C	H	Mg	N	O	
			140	55	74	1	4	6	
2	C	1	Total	C	H	Mg	N	O	
			140	55	74	1	4	6	
2	D	1	Total	C	H	Mg	N	O	
			140	55	74	1	4	6	
2	E	1	Total	C	H	Mg	N	O	
			140	55	74	1	4	6	
2	F	1	Total	C	H	Mg	N	O	
			140	55	74	1	4	6	
2	G	1	Total	C	H	Mg	N	O	
			140	55	74	1	4	6	
2	H	1	Total	C	H	Mg	N	O	
			140	55	74	1	4	6	
2	I	1	Total	C	H	Mg	N	O	
			140	55	74	1	4	6	
2	J	1	Total	C	H	Mg	N	O	
			140	55	74	1	4	6	
2	K	1	Total	C	H	Mg	N	O	
			140	55	74	1	4	6	
2	L	1	Total	C	H	Mg	N	O	
			140	55	74	1	4	6	
2	M	1	Total	C	H	Mg	N	O	
			140	55	74	1	4	6	
2	N	1	Total	C	H	Mg	N	O	
			140	55	74	1	4	6	

4 Residue-property plots ⓘ

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: Bacteriochlorophyll c-binding protein

Chain A:

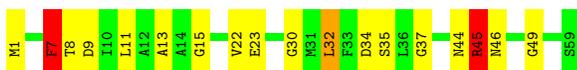• Molecule 1: Bacteriochlorophyll c-binding protein

Chain B:

• Molecule 1: Bacteriochlorophyll c-binding protein

Chain C:

• Molecule 1: Bacteriochlorophyll c-binding protein

Chain D:

• Molecule 1: Bacteriochlorophyll c-binding protein

Chain E:

• Molecule 1: Bacteriochlorophyll c-binding protein

Chain F:

M1 V6 F7 T8 D9 I10 L11 A12 A13 F18 V22 W26 E27 T28 V29 G30 L36 G37 K38 N46 A47 Y48 G49 S59

• Molecule 1: Bacteriochlorophyll c-binding protein

Chain G: 63% 32% 5%

M1 T8 D9 I10 L11 A12 A13 F18 V22 W26 E27 T28 V29 G30 L36 G37 K38 N46 A47 Y48 G49 L56 R57 G58 S59

• Molecule 1: Bacteriochlorophyll c-binding protein

Chain H: 68% 27% 5%

M1 A13 I17 F18 V22 W26 E27 T28 V29 G30 S35 L36 G37 K38 R45 N46 A47 Y48 G49 R57 G58 S59

• Molecule 1: Bacteriochlorophyll c-binding protein

Chain I: 76% 22% •

M1 A13 I17 F18 E19 V22 W26 E27 T28 L36 G37 K38 R45 G49 R57 G58 S59

• Molecule 1: Bacteriochlorophyll c-binding protein

Chain J: 68% 25% 7%

M1 G4 F7 I17 F18 E19 V20 M21 V22 H25 W26 E27 T28 V29 G30 L36 G37 K38 R45 Y48 G49 S59

• Molecule 1: Bacteriochlorophyll c-binding protein

Chain K: 69% 27% • •

M1 F7 T8 D9 F18 V22 W26 E27 T28 V29 G30 L36 G37 K38 R45 N46 A47 Y48 G49 S59

• Molecule 1: Bacteriochlorophyll c-binding protein

Chain L: 66% 32% •

M1 F7 T8 D9 I10 A13 A14 G15 R16 I17 F18 V22 W26 E27 T28 V29 G30 L36 G37 K38 R45 N46 A47 Y48 G49 S59

• Molecule 1: Bacteriochlorophyll c-binding protein

Chain M: 64% 29% 7%

M1 F7 A12 A13 A14 G15 R16 I17 F18 E19 V20 M21 V22 E23 W26 E27 T28 V29 G30 L36 G37 K38 G49 R57 G58 S59

• Molecule 1: Bacteriochlorophyll c-binding protein

Chain N:

78%　　　　　　　　　　　　　　20%　　　　　•

M1 F7 A13 A14 G15 R16 V22 E23 W26 L36 G37 K38 N46 R57 G58 S59

5   Refinement protocol and experimental data overview ⓘ

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

Of the 80 calculated structures, 1 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
Xplor-NIH	refinement	2.33
GASyCS	structure calculation	

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	207
Number of shifts mapped to atoms	207
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	2%

Note: This is a solid-state NMR structure, where hydrogen atoms are typically not assigned a chemical shift value, which may lead to lower completeness of assignment measure.

6 Model quality ⓘ

6.1 Standard geometry ⓘ

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

There are no covalent bond-length or bond-angle outliers.

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	1
1	C	0	1
1	D	0	1
1	G	0	1
1	H	0	1
1	I	0	1
1	M	0	1
1	N	0	1
All	All	0	8

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

5 of 8 planar outliers are listed below.

Mol	Chain	Res	Type	Group
1	A	57	ARG	Sidechain
1	C	45	ARG	Sidechain
1	D	45	ARG	Sidechain
1	G	57	ARG	Sidechain
1	H	57	ARG	Sidechain

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	429	421	421	10
1	B	429	421	421	15
1	C	429	421	421	10
1	D	429	421	421	14
1	E	429	421	421	15
1	F	429	421	421	16
1	G	429	421	421	11
1	H	429	421	421	13
1	I	429	421	421	11
1	J	429	421	421	13
1	K	429	421	421	12
1	L	429	421	421	10
1	M	429	421	421	17
1	N	429	421	421	6
2	A	66	74	74	4
2	B	66	74	74	7
2	C	66	74	74	4
2	D	66	74	74	7
2	E	66	74	74	5
2	F	66	74	74	3
2	G	66	74	74	3
2	H	66	74	74	5
2	I	66	74	74	2
2	J	66	74	74	4
2	K	66	74	74	2
2	L	66	74	74	4
2	M	66	74	74	5
2	N	66	74	74	4
All	All	6930	6930	6930	181

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.

5 of 181 clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)
2:H:101:BCL:HMB1	2:H:101:BCL:HBB2	0.75	1.57
2:F:101:BCL:HMB1	2:F:101:BCL:HBB2	0.74	1.58
2:G:101:BCL:HMB1	2:G:101:BCL:HBB2	0.73	1.59
2:E:101:BCL:HMB1	2:E:101:BCL:HBB2	0.71	1.61
2:C:101:BCL:HMB1	2:C:101:BCL:HBB2	0.71	1.61

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	57/59 (97%)	55 (96%)	2 (4%)	0 (0%)	100	100
1	B	57/59 (97%)	45 (79%)	6 (11%)	6 (11%)	1	9
1	C	57/59 (97%)	44 (77%)	5 (9%)	8 (14%)	1	5
1	D	57/59 (97%)	45 (79%)	5 (9%)	7 (12%)	1	6
1	E	57/59 (97%)	46 (81%)	6 (11%)	5 (9%)	1	12
1	F	57/59 (97%)	43 (75%)	5 (9%)	9 (16%)	0	4
1	G	57/59 (97%)	44 (77%)	5 (9%)	8 (14%)	1	5
1	H	57/59 (97%)	45 (79%)	5 (9%)	7 (12%)	1	6
1	I	57/59 (97%)	46 (81%)	7 (12%)	4 (7%)	2	17
1	J	57/59 (97%)	46 (81%)	5 (9%)	6 (11%)	1	9
1	K	57/59 (97%)	43 (75%)	6 (11%)	8 (14%)	1	5
1	L	57/59 (97%)	45 (79%)	6 (11%)	6 (11%)	1	9
1	M	57/59 (97%)	46 (81%)	4 (7%)	7 (12%)	1	6
1	N	57/59 (97%)	48 (84%)	7 (12%)	2 (4%)	6	35
All	All	798/826 (97%)	641 (80%)	74 (9%)	83 (10%)	1	9

5 of 83 Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	B	7	PHE
1	B	8	THR
1	B	14	ALA
1	B	32	LEU
1	B	35	SER

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	42/42 (100%)	39 (93%)	3 (7%)	18	67
1	B	42/42 (100%)	38 (90%)	4 (10%)	12	58
1	C	42/42 (100%)	37 (88%)	5 (12%)	8	51
1	D	42/42 (100%)	39 (93%)	3 (7%)	18	67
1	E	42/42 (100%)	40 (95%)	2 (5%)	29	78
1	F	42/42 (100%)	42 (100%)	0 (0%)	100	100
1	G	42/42 (100%)	42 (100%)	0 (0%)	100	100
1	H	42/42 (100%)	41 (98%)	1 (2%)	51	92
1	I	42/42 (100%)	41 (98%)	1 (2%)	51	92
1	J	42/42 (100%)	40 (95%)	2 (5%)	29	78
1	K	42/42 (100%)	40 (95%)	2 (5%)	29	78
1	L	42/42 (100%)	41 (98%)	1 (2%)	51	92
1	M	42/42 (100%)	40 (95%)	2 (5%)	29	78
1	N	42/42 (100%)	39 (93%)	3 (7%)	18	67
All	All	588/588 (100%)	559 (95%)	29 (5%)	29	78

5 of 29 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	1	MET
1	A	7	PHE
1	A	46	ASN
1	B	1	MET
1	B	7	PHE

6.3.3 RNA ⓘ

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)

14 ligands are 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	BCL	G	101	1	58,74,74	1.60	10 (17%)
2	BCL	M	101	1	58,74,74	1.64	11 (18%)
2	BCL	A	101	1	58,74,74	1.57	10 (17%)
2	BCL	L	101	1	58,74,74	1.54	11 (18%)
2	BCL	D	101	1	58,74,74	1.40	11 (18%)
2	BCL	F	101	1	58,74,74	1.69	11 (18%)
2	BCL	N	101	1	58,74,74	1.65	10 (17%)
2	BCL	K	101	1	58,74,74	1.59	11 (18%)
2	BCL	E	101	1	58,74,74	1.48	12 (20%)
2	BCL	C	101	1	58,74,74	1.43	10 (17%)
2	BCL	I	101	1	58,74,74	1.73	11 (18%)
2	BCL	H	101	1	58,74,74	1.46	12 (20%)
2	BCL	B	101	1	58,74,74	1.58	10 (17%)
2	BCL	J	101	1	58,74,74	1.72	9 (15%)

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	BCL	G	101	1	69,115,115	1.86	13 (18%)
2	BCL	M	101	1	69,115,115	1.93	16 (23%)
2	BCL	A	101	1	69,115,115	1.95	12 (17%)
2	BCL	L	101	1	69,115,115	1.87	12 (17%)
2	BCL	D	101	1	69,115,115	1.92	13 (18%)
2	BCL	F	101	1	69,115,115	1.83	13 (18%)
2	BCL	N	101	1	69,115,115	1.97	13 (18%)
2	BCL	K	101	1	69,115,115	1.87	13 (18%)
2	BCL	E	101	1	69,115,115	1.91	14 (20%)
2	BCL	C	101	1	69,115,115	1.98	14 (20%)
2	BCL	I	101	1	69,115,115	1.89	14 (20%)
2	BCL	H	101	1	69,115,115	1.91	12 (17%)
2	BCL	B	101	1	69,115,115	1.95	13 (18%)
2	BCL	J	101	1	69,115,115	1.81	13 (18%)

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	BCL	G	101	1	1,1,25,25	1,37,137,137	-
2	BCL	M	101	1	1,1,25,25	1,37,137,137	-
2	BCL	A	101	1	1,1,25,25	0,37,137,137	-
2	BCL	L	101	1	1,1,25,25	0,37,137,137	-
2	BCL	D	101	1	1,1,25,25	1,37,137,137	-
2	BCL	F	101	1	1,1,25,25	0,37,137,137	-
2	BCL	N	101	1	1,1,25,25	0,37,137,137	-
2	BCL	K	101	1	1,1,25,25	0,37,137,137	-
2	BCL	E	101	1	1,1,25,25	0,37,137,137	-
2	BCL	C	101	1	1,1,25,25	0,37,137,137	-
2	BCL	I	101	1	1,1,25,25	1,37,137,137	-
2	BCL	H	101	1	1,1,25,25	1,37,137,137	-
2	BCL	B	101	1	1,1,25,25	0,37,137,137	-
2	BCL	J	101	1	1,1,25,25	0,37,137,137	-

5 of 149 bond outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	I	101	BCL	C3C-C4C	8.63	1.62	1.51
2	F	101	BCL	C3C-C4C	7.87	1.61	1.51
2	J	101	BCL	C3C-C4C	7.82	1.61	1.51
2	N	101	BCL	C3C-C4C	7.55	1.61	1.51
2	M	101	BCL	C3C-C4C	7.24	1.60	1.51

5 of 185 angle outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	N	101	BCL	C4A-NA-C1A	9.70	111.07	106.71
2	M	101	BCL	C4A-NA-C1A	9.42	110.94	106.71
2	C	101	BCL	C4A-NA-C1A	9.30	110.89	106.71
2	A	101	BCL	C4A-NA-C1A	9.23	110.86	106.71
2	E	101	BCL	C4A-NA-C1A	8.80	110.66	106.71

5 of 14 chiral outliers are listed below.

Mol	Chain	Res	Type	Atoms
2	A	101	BCL	C8
2	B	101	BCL	C8
2	C	101	BCL	C8
2	D	101	BCL	C8
2	E	101	BCL	C8

All torsion outliers are listed below.

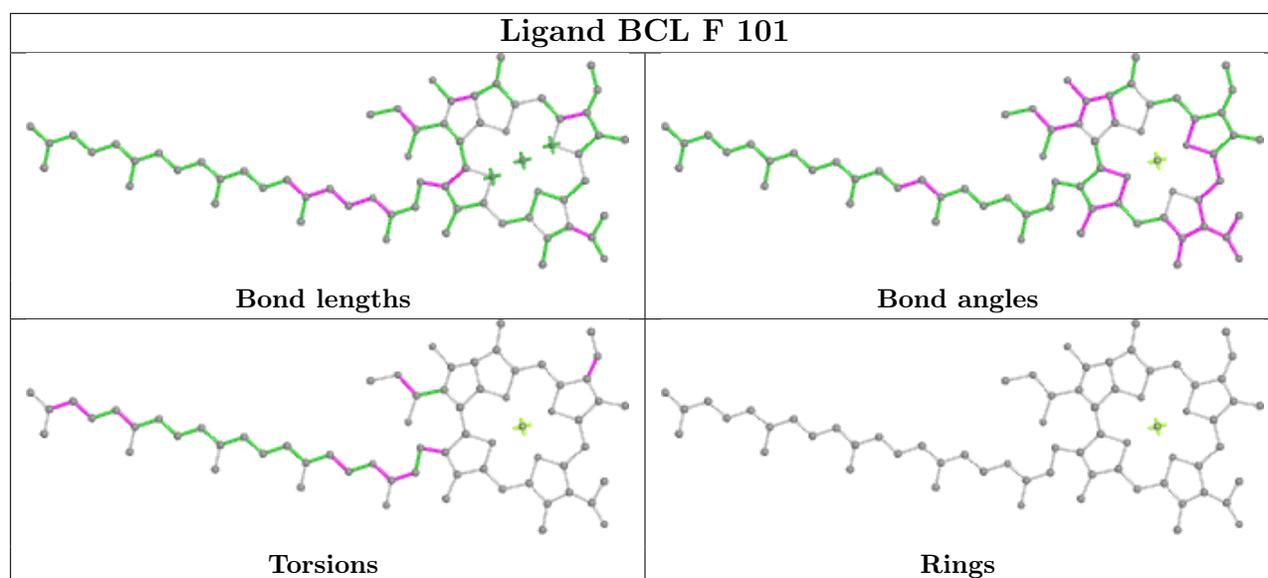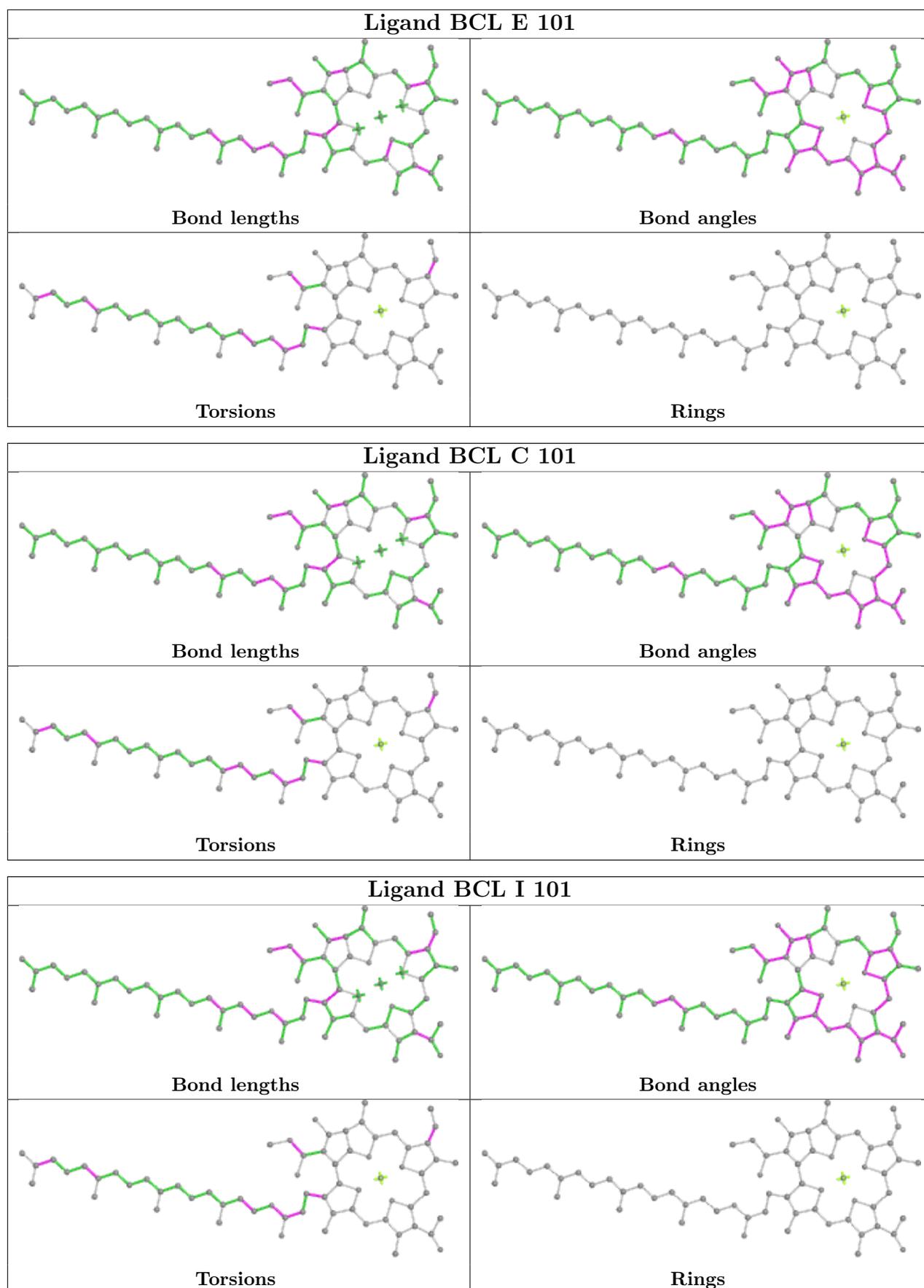| Mol	Chain	Res	Type	Atoms
2	D	101	BCL	O2A-C1-C2-C3
2	G	101	BCL	O2A-C1-C2-C3
2	H	101	BCL	O2A-C1-C2-C3
2	I	101	BCL	O2A-C1-C2-C3
2	M	101	BCL	O2A-C1-C2-C3

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
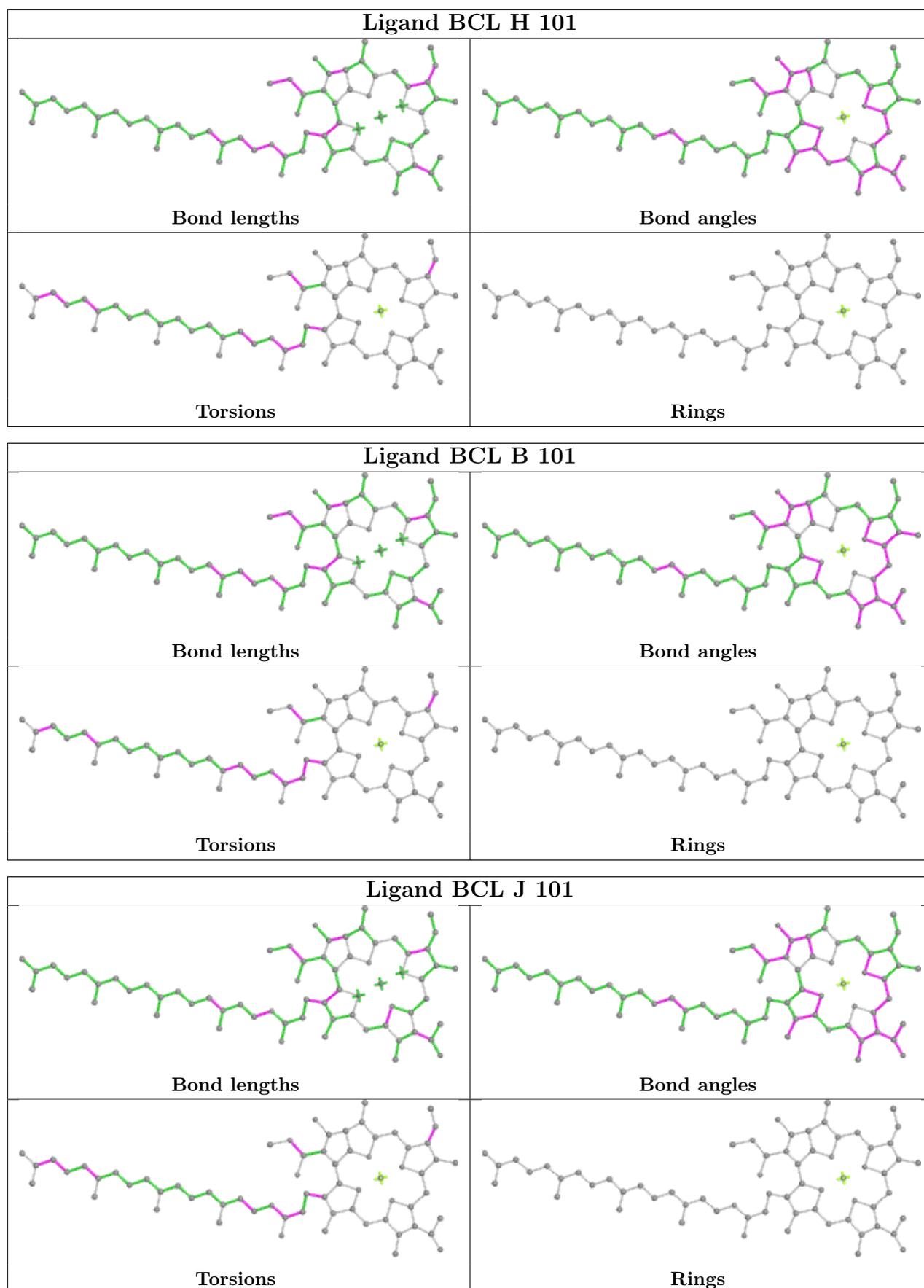
in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand BCL G 101	
\n\nBond lengths	\n\nBond angles
\n\nTorsions	\n\nRings

Ligand BCL M 101	
\n\nBond lengths	\n\nBond angles
\n\nTorsions	\n\nRings

Ligand BCL A 101

Bond lengths	Bond angles
Torsions	Rings

Ligand BCL L 101

Bond lengths	Bond angles
Torsions	Rings

Ligand BCL D 101

Bond lengths	Bond angles
Torsions	Rings

Ligand BCL F 101

Bond lengths

Bond angles

Torsions

Rings

Ligand BCL N 101

Bond lengths

Bond angles

Torsions

Rings

Ligand BCL K 101

Bond lengths

Bond angles

Torsions

Rings

Ligand BCL E 101

Bond lengths

Bond angles

Torsions

Rings

Ligand BCL C 101

Bond lengths

Bond angles

Torsions

Rings

Ligand BCL I 101

Bond lengths

Bond angles

Torsions

Rings

Ligand BCL H 101

Bond lengths

Bond angles

Torsions

Rings

Ligand BCL B 101

Bond lengths

Bond angles

Torsions

Rings

Ligand BCL J 101

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 2% for the well-defined parts and 2% for the entire structure.

7.1   Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *csmanewer31.str*

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	207
Number of shifts mapped to atoms	207
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   ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}\text{C}_\alpha$	44	-1.07 ± 0.21	Should be checked
$^{13}\text{C}_\beta$	34	0.85 ± 0.52	None needed (imprecise)
$^{13}\text{C}'$	43	-0.61 ± 0.37	None needed (imprecise)
^{15}N	44	2.02 ± 0.32	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 2%, i.e. 204 atoms were assigned a chemical shift out of a possible 10388. 0 out of 112 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	131/4312 (3%)	0/1834 (0%)	87/1652 (5%)	44/826 (5%)
Sidechain	73/5264 (1%)	0/3472 (0%)	73/1582 (5%)	0/210 (0%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	0/812 (0%)	0/406 (0%)	0/378 (0%)	0/28 (0%)
Overall	204/10388 (2%)	0/5712 (0%)	160/3612 (4%)	44/1064 (4%)

Note: This is a solid-state NMR structure, where hydrogen atoms are typically not assigned a chemical shift value, which may lead to lower completeness of assignment measure.

7.1.4 Statistically unusual chemical shifts ⓘ

There are no statistically unusual chemical shifts.

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.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:
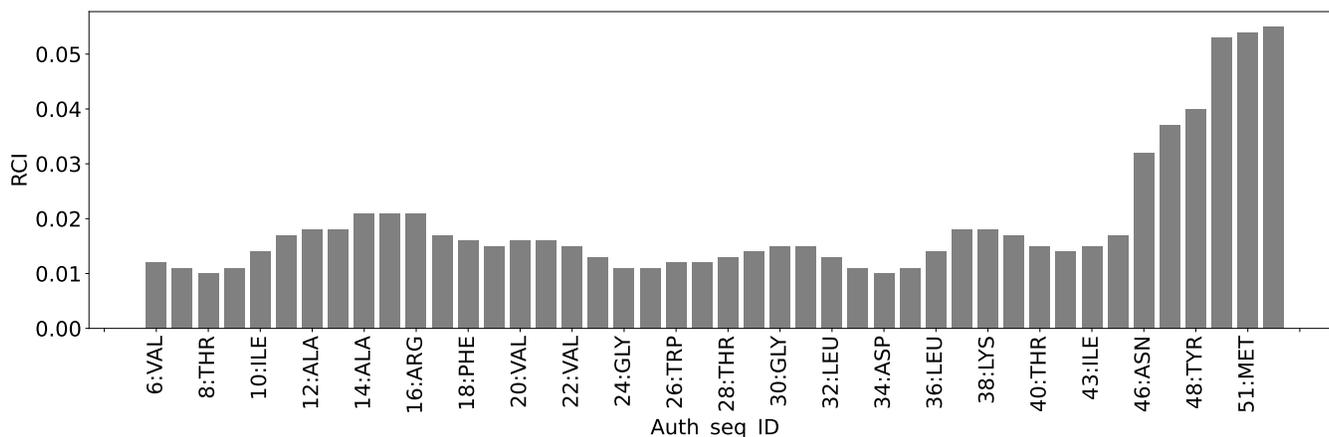
8 NMR restraints analysis ⓘ

8.1 Conformationally restricting restraints ⓘ

The following table provides the summary of experimentally observed NMR restraints in different categories. Restraints are classified into different categories based on the sequence separation of the atoms involved.

Description	Value
Total distance restraints	54
Intra-residue ($|i\text{-}j|$=0)	1
Sequential ($|i\text{-}j|$=1)	15
Medium range ($|i\text{-}j|$>1 and $|i\text{-}j|$<5)	19
Long range ($|i\text{-}j|$≥5)	9
Inter-chain	10
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	0
Number of restraints per residue	0.1
Number of long range restraints per residue[1]	0.0

[1]Long range hydrogen bonds and disulfide bonds are counted as long range restraints while calculating the number of long range restraints per residue

8.2 Residual restraint violations ⓘ

This section provides the overview of the restraint violations analysis. The violations are binned as small, medium and large violations based on its absolute value. Average number of violations per model is calculated by dividing the total number of violations in each bin by the size of the ensemble.

8.2.1 Average number of distance violations per model ⓘ

Distance violations less than 0.1 Å are not included in the calculation.

Bins (Å)	Average number of violations per model	Max (Å)
0.1-0.2 (Small)	None	None
0.2-0.5 (Medium)	None	None
>0.5 (Large)	6.0	2.39

8.2.2 Average number of dihedral-angle violations per model ⓘ

Dihedral-angle violations less than 1° are not included in the calculation.There are no dihedral-angle violations

9 Distance violation analysis ⓘ

## 9.1 Summary of distance violations ⓘ

The following table shows the summary of distance violations in different restraint categories based on the sequence separation of the atoms involved. Each category is further sub-divided into three sub-categories based on the atoms involved. Violations less than 0.1 Å are not included in the statistics.

Restraints type	Count	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
Intra-residue (|i-j|=0)	1	1.9	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	1	1.9	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	15	27.8	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	4	7.4	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	11	20.4	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	19	35.2	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	4	7.4	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	15	27.8	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Long range (|i-j|≥5)	9	16.7	4	44.4	7.4	4	44.4	7.4
Backbone-Backbone	1	1.9	1	100.0	1.9	1	100.0	1.9
Backbone-Sidechain	7	13.0	2	28.6	3.7	2	28.6	3.7
Sidechain-Sidechain	1	1.9	1	100.0	1.9	1	100.0	1.9
Inter-chain	10	18.5	2	20.0	3.7	2	20.0	3.7
Backbone-Backbone	3	5.6	1	33.3	1.9	1	33.3	1.9
Backbone-Sidechain	7	13.0	1	14.3	1.9	1	14.3	1.9
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Hydrogen bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	54	100.0	6	11.1	11.1	6	11.1	11.1
Backbone-Backbone	12	22.2	2	16.7	3.7	2	16.7	3.7
Backbone-Sidechain	41	75.9	3	7.3	5.6	3	7.3	5.6
Sidechain-Sidechain	1	1.9	1	100.0	1.9	1	100.0	1.9

[1] percentage calculated with respect to the total number of distance restraints, [2] percentage calculated with respect to the number of restraints in a particular restraint category, [3] violated in at least one model, [4] violated in all the models

9.1.1　Bar chart : Distribution of distance restraints and violations ⓘ
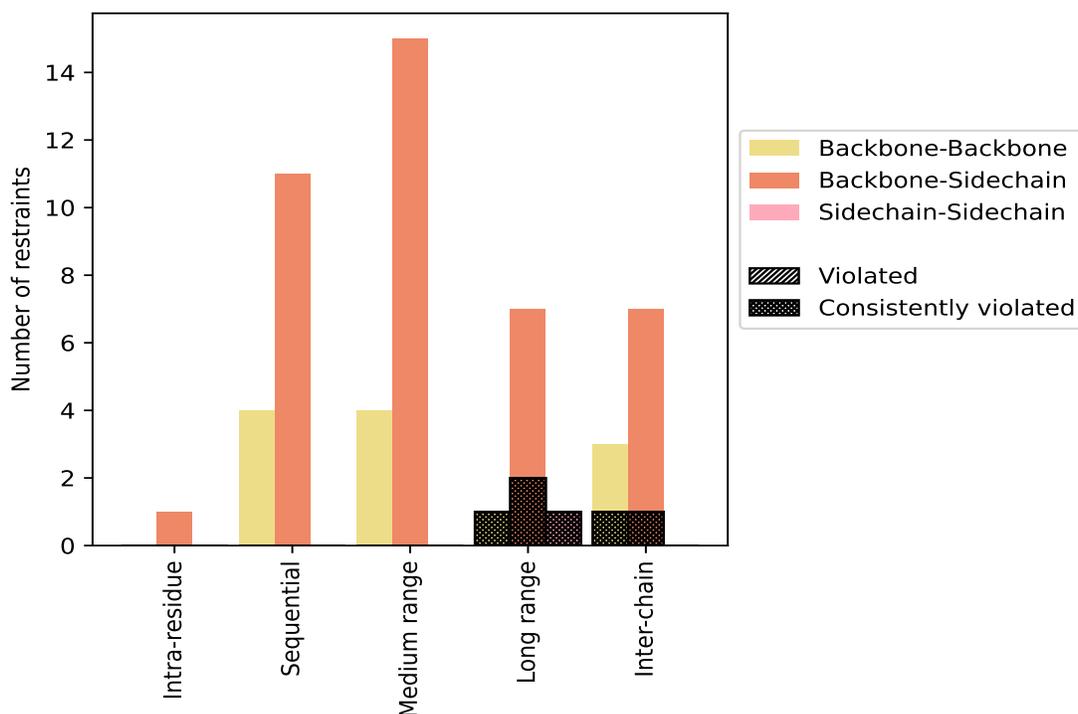
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories. The hydrogen bonds and disulfied bonds are counted in their appropriate category on the x-axis
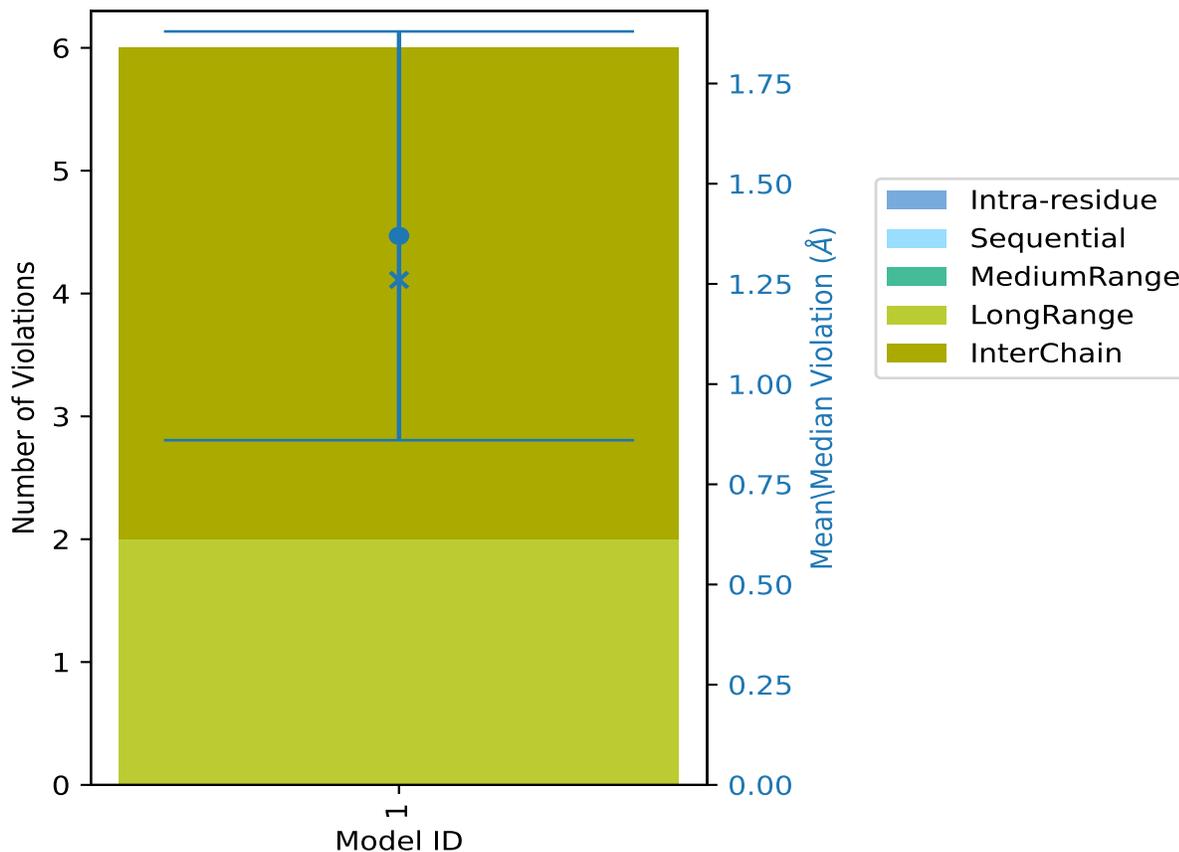
9.2　Distance violation statistics for each model ⓘ

The following table provides the distance violation statistics for each model in the ensemble. Violations less than 0.1 Å are not included in the statistics.

Model ID	Number of violations						Mean (Å)	Max (Å)	SD[6] (Å)	Median (Å)
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
1	0	0	0	2	4	6	1.37	2.39	0.51	1.26

[1]Intra-residue restraints, [2]Sequential restraints, [3]Medium range restraints, [4]Long range restraints, [5]Inter-chain restraints, [6]Standard deviation

9.2.1 Bar graph : Distance Violation statistics for each model ⓘ

The mean(dot),median(x) and the standard deviation are shown in blue with respect to the y axis on the right

9.3 Distance violation statistics for the ensemble ⓘ

Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of the ensemble. In total, 48(IR:1, SQ:15, MR:19, LR:5, IC:8) restraints are not violated in the ensemble.

Number of violated restraints						Fraction of the ensemble	
IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total	Count[6]	%
0	0	0	4	2	6	1	100.0

[1]Intra-residue restraints, [2]Sequential restraints, [3]Medium range restraints, [4]Long range restraints, [5]Inter-chain restraints, [6] Number of models with violations

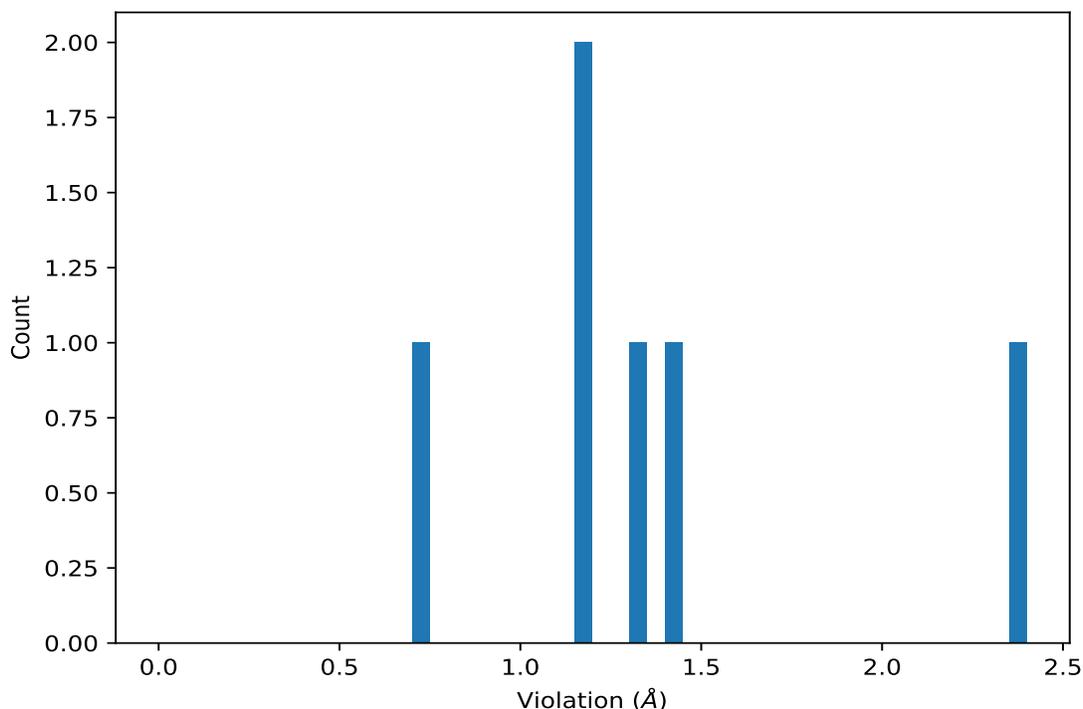### 9.3.1 Bar graph : Distance violation statistics for the ensemble ⓘ
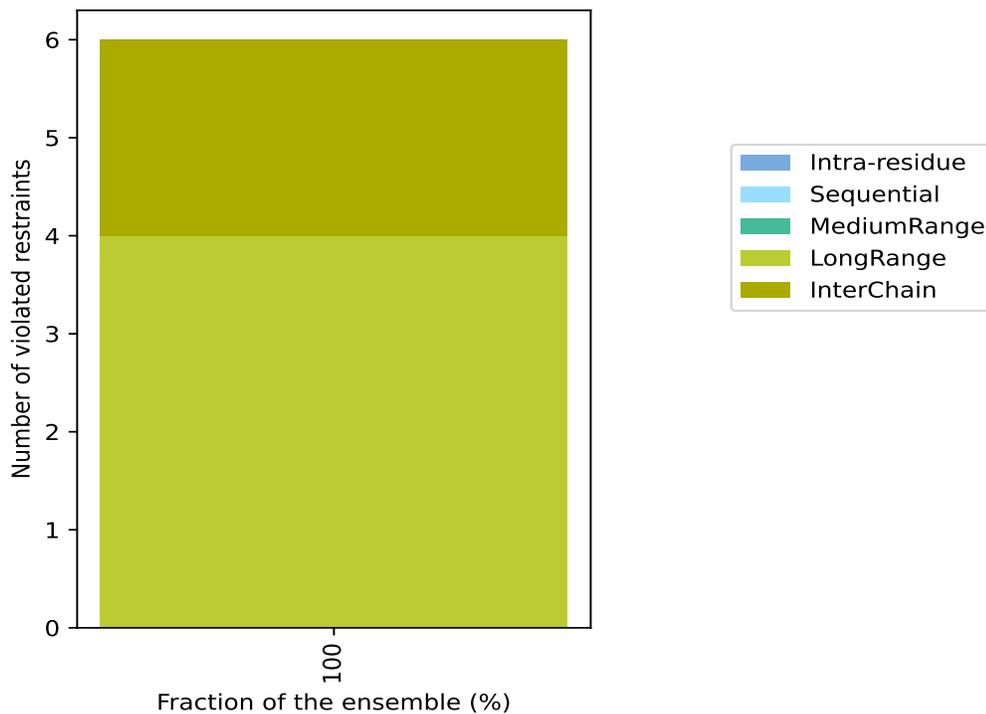
9.4 Most violated distance restraints in the ensemble ⓘ

No violations found

9.5 All violated distance restraints ⓘ

9.5.1 Histogram : Distribution of distance violations ⓘ

The following histogram shows the distribution of the absolute value of the violation for all violated restraints in the ensemble.

9.5.2 Table : All distance violations ⓘ

The following table lists the absolute value of the violation for each restraint in the ensemble sorted by its value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint. Rows with same key represent combinatorial or ambiguous restraints and are counted as a single restraint.

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,48)	1:39:B:GLY:CA	1:14:J:ALA:CA	1	2.39
(1,13)	1:11:J:LEU:CG	1:28:A:THR:CA	1	1.4
(1,7)	1:34:B:ASP:CB	1:8:J:THR:CA	1	1.33
(1,24)	1:29:E:VAL:CG2	1:24:E:GLY:CA	1	1.19
(1,2)	1:29:E:VAL:CG2	1:24:E:GLY:CA	1	1.19
(1,41)	1:47:H:ALA:CA	1:40:A:THR:C	1	0.71

10   Dihedral-angle violation analysis ⓘ

No dihedral-angle restraints found