WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB NMR Structure Validation Report (i)

Dec 25, 2024 – 08:31 PM EST

PDB ID	:	8Y2R
BMRB ID	:	36642
Title	:	NMR Solution Structure of the 2:1 Coptisine-ATF4-G4 Complex
Authors	:	Xiao, C.M.; Li, Y.P.; Liu, Y.S.; Dong, R.F.; He, X.Y.; Lin, Q.; Zang, X.; Wang, K.B.; Xia, Y.Z.; Kong, L.Y.
Deposited on	:	2024-01-27

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

MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

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 43%.

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		0

Worse   *Better*

▌ Percentile relative to all structures

▯ Percentile relative to all NMR structures

Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	210492	14027

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	X	22	32% 68%

2 Ensemble composition and analysis ⓘ

This entry contains 10 models. This entry does not contain polypeptide chains, therefore identification of well-defined residues and clustering analysis are not possible. All residues are included in the validation scores.

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 790 atoms, of which 273 are hydrogens and 0 are deuteriums.

- Molecule 1 is a DNA chain called DNA (5'-D(*TP*AP*AP*GP*GP*GP*AP*GP*GP*GP* AP*GP*GP*GP*AP*AP*GP*GP*GP*AP*GP*C*(LIG)*(LIG))-3').

Mol	Chain	Residues	Atoms						Trace
1	X	22	Total	C	H	N	O	P	0
			714	219	245	105	124	21	

- Molecule 2 is 6,7-dihydro[1,3]dioxolo[4,5-g][1,3]dioxolo[7,8]isoquino[3,2-a]isoquinolin-5-ium (three-letter code: KPT) (formula: $C_{19}H_{14}NO_4$) (labeled as "Ligand of Interest" by depositor).

KPT
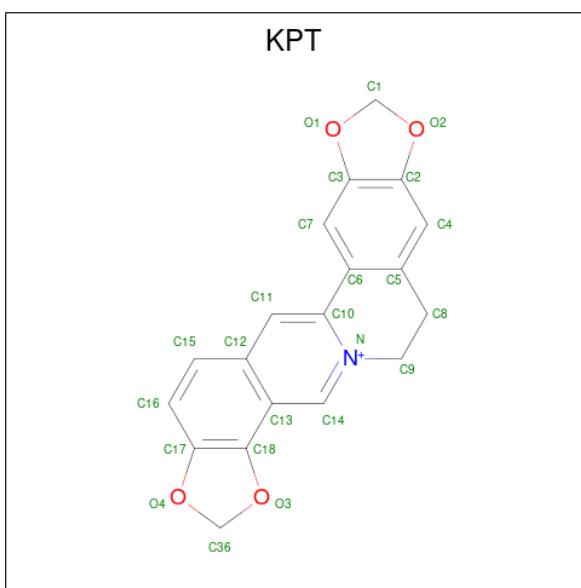
Mol	Chain	Residues	Atoms				
2	X	1	Total	C	H	N	O
			38	19	14	1	4
2	X	1	Total	C	H	N	O
			38	19	14	1	4

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: DNA (5'-D(*TP*AP*AP*GP*GP*GP*AP*GP*GP*GP*AP*GP*GP*GP*AP*A P*GP*GP*GP*AP*GP*C*(LIG)*(LIG))-3')
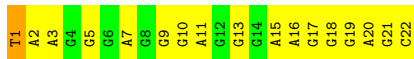
Chain X:

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: DNA (5'-D(*TP*AP*AP*GP*GP*GP*AP*GP*GP*GP*AP*GP*GP*GP*AP*A P*GP*GP*GP*AP*GP*C*(LIG)*(LIG))-3')

Chain X:

4.2.2 Score per residue for model 2

• Molecule 1: DNA (5'-D(*TP*AP*AP*GP*GP*GP*AP*GP*GP*GP*AP*GP*GP*GP*AP*A P*GP*GP*GP*AP*GP*C*(LIG)*(LIG))-3')

Chain X:

4.2.3 Score per residue for model 3

• Molecule 1: DNA (5'-D(*TP*AP*AP*GP*GP*GP*AP*GP*GP*GP*AP*GP*GP*GP*AP*A P*GP*GP*GP*AP*GP*C*(LIG)*(LIG))-3')

Chain X: 32% 68%

T1 A2 A3 G4 G5 G6 A7 G8 G9 G10 A11 G12 G13 G14 A15 A16 G17 G18 G19 A20 G21 C22

4.2.4 Score per residue for model 4

• Molecule 1: DNA (5'-D(*TP*AP*AP*GP*GP*GP*AP*GP*GP*GP*AP*GP*GP*GP*AP*A P*GP*GP*GP*AP*GP*C*(LIG)*(LIG))-3')

Chain X: 27% 73%

T1 A2 A3 G6 A7 G8 G9 G10 A11 G12 G13 G14 A15 A16 G17 G18 G19 A20 G21 C22

4.2.5 Score per residue for model 5

• Molecule 1: DNA (5'-D(*TP*AP*AP*GP*GP*GP*AP*GP*GP*GP*AP*GP*GP*GP*AP*A P*GP*GP*GP*AP*GP*C*(LIG)*(LIG))-3')

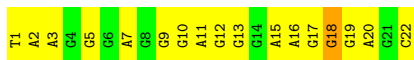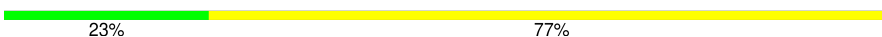Chain X: 23% 77%

T1 A2 A3 G6 A7 G8 G9 G10 A11 G12 G13 G14 A15 A16 G17 G18 G19 A20 G21 C22

4.2.6 Score per residue for model 6

• Molecule 1: DNA (5'-D(*TP*AP*AP*GP*GP*GP*AP*GP*GP*GP*AP*GP*GP*GP*AP*A P*GP*GP*GP*AP*GP*C*(LIG)*(LIG))-3')

Chain X: 32% 68%

T1 A2 A3 G4 G5 G6 A7 G8 G9 G10 A11 G12 G13 G14 A15 A16 G17 G18 G19 A20 G21 C22

4.2.7 Score per residue for model 7

• Molecule 1: DNA (5'-D(*TP*AP*AP*GP*GP*GP*AP*GP*GP*GP*AP*GP*GP*GP*AP*A P*GP*GP*GP*AP*GP*C*(LIG)*(LIG))-3')

Chain X: 32% 68%

T1 A2 A3 G4 A7 G8 G9 G10 A11 A15 A16 G17 G18 G19 A20 G21 C22

4.2.8 Score per residue for model 8

● Molecule 1: DNA (5'-D(*TP*AP*AP*GP*GP*GP*AP*GP*GP*GP*AP*GP*GP*GP*AP*A P*GP*GP*GP*AP*GP*C*(LIG)*(LIG))-3')

Chain X: 27% 73%

T1 A2 A3 G4 G5 G6 A7 G8 G9 G10 A11 G12 G13 G14 A15 A16 G17 G18 G19 A20 G21 C22

4.2.9 Score per residue for model 9

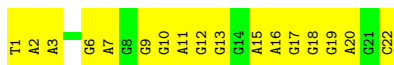● Molecule 1: DNA (5'-D(*TP*AP*AP*GP*GP*GP*AP*GP*GP*GP*AP*GP*GP*GP*AP*A P*GP*GP*GP*AP*GP*C*(LIG)*(LIG))-3')
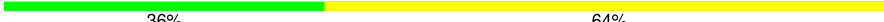
Chain X: 18% 82%

T1 A2 A3 G4 G5 G6 A7 G8 G9 G10 A11 G12 G13 G14 A15 A16 G17 G18 G19 A20 G21 C22

4.2.10 Score per residue for model 10
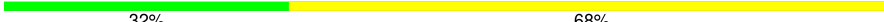
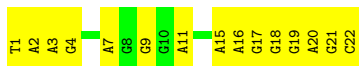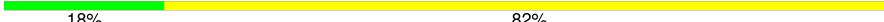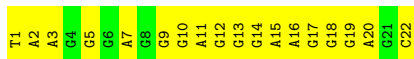● Molecule 1: DNA (5'-D(*TP*AP*AP*GP*GP*GP*AP*GP*GP*GP*AP*GP*GP*GP*AP*A P*GP*GP*GP*AP*GP*C*(LIG)*(LIG))-3')

Chain X: 36% 64%

T1 A2 A3 G6 A7 G8 G9 G10 A11 G12 G13 G14 A15 A16 G17 G18 G19 A20 G21 C22

5 Refinement protocol and experimental data overview ⓘ

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

Of the 100 calculated structures, 10 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
Amber	structure calculation	
X-PLOR NIH	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	199
Number of shifts mapped to atoms	199
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	43%

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: KPT

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	X	1.53±0.01	0±0/532 (0.0± 0.0%)	2.39±0.02	36±2/824 (4.3± 0.2%)
All	All	1.53	0/5320 (0.0%)	2.39	357/8240 (4.3%)

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	X	0.0±0.0	0.2±0.4
All	All	0	2

There are no bond-length outliers.

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	X	15	DA	N1-C6-N6	-11.58	111.65	118.60	8	10
1	X	11	DA	N1-C6-N6	-9.79	112.73	118.60	2	10
1	X	20	DA	N1-C6-N6	-9.70	112.78	118.60	9	10
1	X	16	DA	N1-C6-N6	-9.62	112.83	118.60	7	10
1	X	17	DG	O4'-C1'-N9	9.59	114.71	108.00	10	10
1	X	7	DA	N1-C6-N6	-9.57	112.86	118.60	7	10
1	X	2	DA	N1-C6-N6	-9.30	113.02	118.60	6	10
1	X	15	DA	O4'-C1'-N9	8.28	113.80	108.00	2	6
1	X	15	DA	C5-C6-N1	8.11	121.75	117.70	8	10
1	X	3	DA	N1-C6-N6	-8.10	113.74	118.60	5	10
1	X	15	DA	C4-C5-C6	-7.86	113.07	117.00	8	10
1	X	16	DA	C5-C6-N1	7.70	121.55	117.70	7	10

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models Worst	Total
1	X	11	DA	C5-C6-N1	7.66	121.53	117.70	2	10
1	X	18	DG	O4'-C1'-N9	7.64	113.35	108.00	2	6
1	X	7	DA	C5-C6-N1	7.59	121.50	117.70	3	10
1	X	2	DA	C5-C6-N1	7.58	121.49	117.70	4	10
1	X	20	DA	C5-C6-N1	7.52	121.46	117.70	6	10
1	X	20	DA	C4-C5-C6	-7.19	113.41	117.00	9	10
1	X	3	DA	C5-C6-N1	7.13	121.27	117.70	4	10
1	X	5	DG	O4'-C1'-N9	7.11	112.98	108.00	8	4
1	X	11	DA	C4-C5-C6	-7.06	113.47	117.00	2	10
1	X	1	DT	C6-C5-C7	-7.05	118.67	122.90	1	10
1	X	15	DA	P-O3'-C3'	6.96	128.05	119.70	3	1
1	X	7	DA	C4-C5-C6	-6.83	113.59	117.00	3	10
1	X	16	DA	C4-C5-C6	-6.73	113.64	117.00	3	10
1	X	2	DA	C4-C5-C6	-6.68	113.66	117.00	4	10
1	X	19	DG	O4'-C1'-N9	6.60	112.62	108.00	5	8
1	X	22	DC	N3-C2-O2	-6.50	117.35	121.90	1	10
1	X	3	DA	C4-C5-C6	-6.49	113.75	117.00	5	10
1	X	9	DG	O4'-C1'-N9	6.49	112.54	108.00	1	2
1	X	1	DT	O4'-C1'-N1	6.49	112.54	108.00	10	1
1	X	20	DA	O4'-C4'-C3'	6.14	109.69	106.00	8	10
1	X	12	DG	P-O3'-C3'	5.96	126.85	119.70	9	1
1	X	4	DG	P-O3'-C3'	5.91	126.79	119.70	7	1
1	X	5	DG	N3-C2-N2	-5.89	115.78	119.90	2	6
1	X	9	DG	N3-C2-N2	-5.88	115.78	119.90	2	10
1	X	13	DG	O4'-C1'-N9	5.87	112.11	108.00	1	6
1	X	16	DA	P-O3'-C3'	5.81	126.67	119.70	5	2
1	X	2	DA	O4'-C1'-N9	5.69	111.98	108.00	5	6
1	X	1	DT	N3-C2-O2	-5.59	118.94	122.30	10	10
1	X	21	DG	O4'-C4'-C3'	5.53	109.32	106.00	8	3
1	X	13	DG	N3-C2-N2	-5.51	116.04	119.90	5	4
1	X	14	DG	O4'-C1'-N9	5.51	111.86	108.00	9	1
1	X	19	DG	N3-C2-N2	-5.46	116.08	119.90	3	3
1	X	6	DG	N3-C2-N2	-5.36	116.15	119.90	10	3
1	X	18	DG	N3-C2-N2	-5.31	116.18	119.90	9	1
1	X	15	DA	C6-C5-N7	5.29	136.00	132.30	10	6
1	X	10	DG	O4'-C1'-N9	5.28	111.69	108.00	2	3
1	X	10	DG	N3-C2-N2	-5.28	116.21	119.90	9	2
1	X	17	DG	P-O3'-C3'	5.19	125.93	119.70	5	1
1	X	18	DG	C5-C6-N1	5.17	114.08	111.50	3	3
1	X	17	DG	N3-C2-N2	-5.13	116.31	119.90	9	2
1	X	14	DG	C5-C6-N1	5.13	114.06	111.50	4	1

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models Worst	Total
1	X	12	DG	N3-C2-N2	-5.08	116.34	119.90	5	3
1	X	19	DG	C5-C6-N1	5.06	114.03	111.50	6	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	X	1	DT	Sidechain	1
1	X	18	DG	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
All	All	5170	2730	2730	-

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 -.

There are no clashes.

6.3 Torsion angles ⓘ

6.3.1 Protein backbone ⓘ

There are no protein molecules in this entry.

6.3.2 Protein sidechains ⓘ

There are no protein molecules in this entry.

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 oligosaccharides in this entry.

6.6   Ligand geometry ⓘ

2 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	KPT	X	101	-	29,29,29	1.86±0.01	7±0 (24±1%)
2	KPT	X	102	-	29,29,29	1.86±0.01	7±0 (24±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	KPT	X	101	-	41,44,44	1.41±0.02	7±1 (17±2%)
2	KPT	X	102	-	41,44,44	1.48±0.02	7±1 (16±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	KPT	X	101	-	-	-	0±0,6,6,6
2	KPT	X	102	-	-	-	0±0,6,6,6

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
2	X	102	KPT	C14-N	6.44	1.40	1.33	4	10
2	X	101	KPT	C14-N	6.29	1.39	1.33	2	10
2	X	101	KPT	C11-C10	3.18	1.42	1.37	9	10
2	X	102	KPT	C11-C10	3.08	1.42	1.37	7	10
2	X	102	KPT	C6-C10	2.62	1.43	1.47	9	10
2	X	101	KPT	C6-C10	2.59	1.43	1.47	9	10
2	X	101	KPT	C15-C16	2.51	1.42	1.36	3	10
2	X	102	KPT	C15-C16	2.50	1.42	1.36	9	10
2	X	101	KPT	C6-C5	2.40	1.43	1.40	7	10
2	X	102	KPT	C6-C5	2.40	1.43	1.40	7	10
2	X	102	KPT	C10-N	2.33	1.41	1.37	1	10
2	X	101	KPT	C10-N	2.33	1.41	1.37	9	10
2	X	102	KPT	C7-C6	2.24	1.43	1.39	4	10
2	X	101	KPT	C7-C6	2.19	1.43	1.39	10	10
2	X	101	KPT	C9-C8	2.08	1.54	1.51	3	2

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(°)	Ideal(°)	Models Worst	Total
2	X	102	KPT	C14-N-C10	5.16	119.61	122.39	5	10
2	X	101	KPT	C14-N-C10	4.69	119.86	122.39	8	10
2	X	102	KPT	C13-C14-N	3.00	120.06	121.81	2	10
2	X	101	KPT	C9-N-C10	2.66	121.72	118.81	9	4
2	X	101	KPT	C6-C10-N	2.53	119.48	118.12	9	7
2	X	102	KPT	C14-C13-C12	2.45	120.33	117.59	2	10
2	X	101	KPT	C14-C13-C12	2.43	120.31	117.59	3	10
2	X	101	KPT	C13-C14-N	2.36	120.44	121.81	8	8
2	X	102	KPT	C11-C10-C6	2.32	122.25	124.47	1	10
2	X	101	KPT	C11-C10-C6	2.30	122.27	124.47	9	9
2	X	102	KPT	C6-C10-N	2.20	119.31	118.12	3	7
2	X	102	KPT	O3-C18-C13	2.20	128.31	123.43	4	10
2	X	101	KPT	C8-C9-N	2.18	112.73	109.75	9	1
2	X	101	KPT	C2-C4-C5	2.16	117.85	121.10	6	10
2	X	101	KPT	O3-C18-C13	2.16	128.23	123.43	8	10

Continued from previous page...

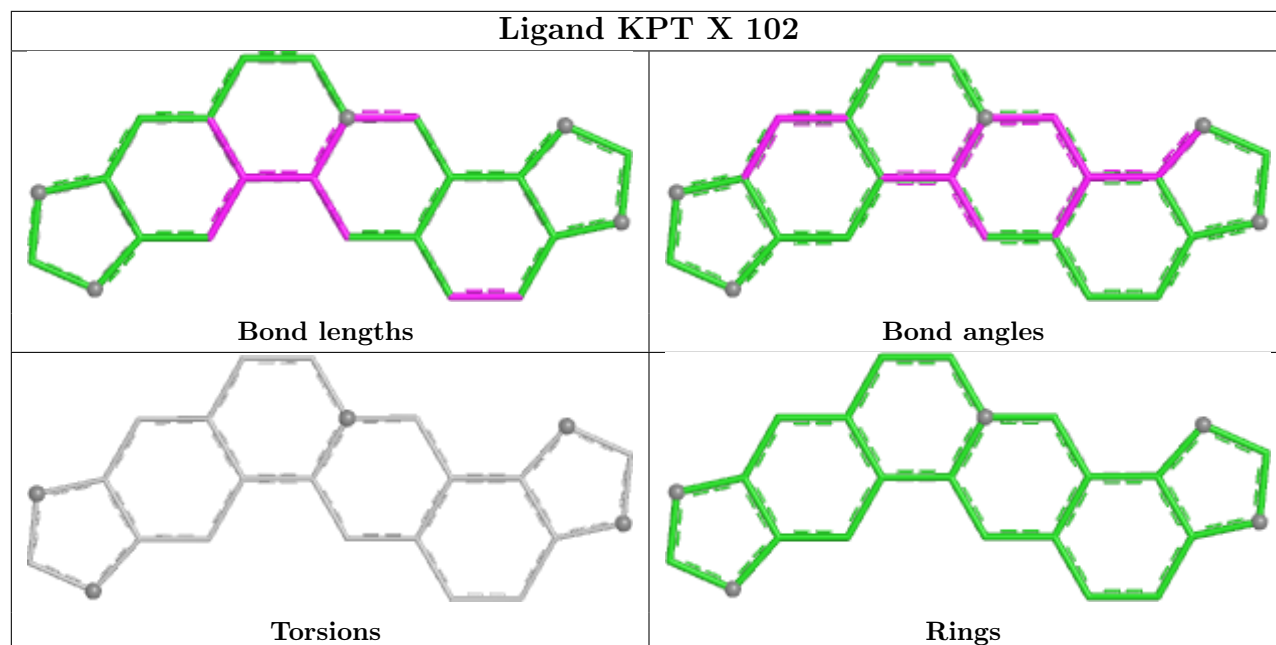| Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models Worst	Models Total
2	X	101	KPT	C8-C5-C6	2.09	121.01	118.58	2	4
2	X	102	KPT	C2-C4-C5	2.07	118.00	121.10	7	6
2	X	102	KPT	C8-C5-C6	2.03	120.94	118.58	4	4

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 KPT X 101

Bond lengths

Bond angles

Torsions

Rings

Ligand KPT X 102

Bond lengths

Bond angles

Torsions

Rings
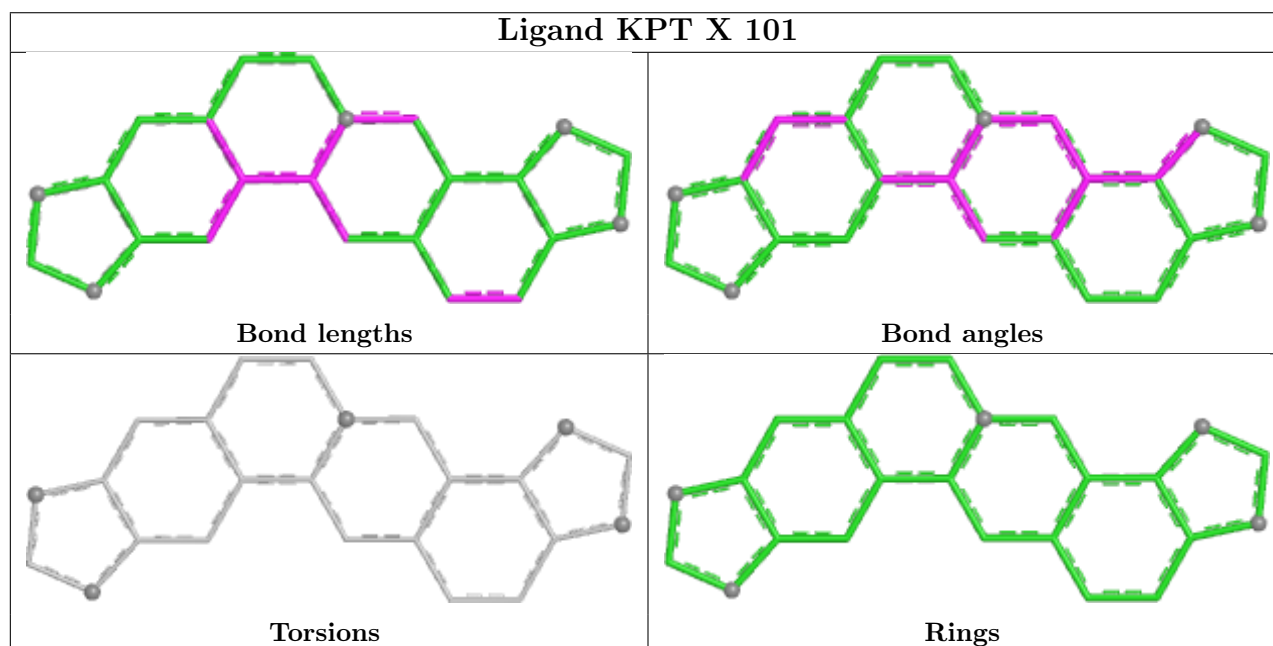
6.7 Other polymers ⓘ

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 43% for the well-defined parts and 43% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *starch_output*

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	199
Number of shifts mapped to atoms	199
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 ⓘ

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

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 43%, i.e. 196 atoms were assigned a chemical shift out of a possible 451. 0 out of 0 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Sugar	154/264 (58%)	154/154 (100%)	0/110 (0%)	0/0 (—%)
Base	42/187 (22%)	42/121 (35%)	0/31 (0%)	0/35 (0%)
Overall	196/451 (43%)	196/275 (71%)	0/141 (0%)	0/35 (0%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 43%, i.e. 196 atoms were assigned a chemical shift out of a possible 451. 0 out of 0 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Sugar	154/264 (58%)	154/154 (100%)	0/110 (0%)	0/0 (—%)
Base	42/187 (22%)	42/121 (35%)	0/31 (0%)	0/35 (0%)
Overall	196/451 (43%)	196/275 (71%)	0/141 (0%)	0/35 (0%)

7.1.4　Statistically unusual chemical shifts ⓘ

There are no statistically unusual chemical shifts.

7.1.5　Random Coil Index (RCI) plots ⓘ

No *random coil index* (RCI) plot could be generated from the current chemical shift list. RCI is only applicable to proteins

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	566
Intra-residue ($|i\text{-}j|$=0)	291
Sequential ($|i\text{-}j|$=1)	156
Medium range ($|i\text{-}j|$>1 and $|i\text{-}j|$<5)	24
Long range ($|i\text{-}j|$≥5)	47
Inter-chain	0
Hydrogen bond restraints	48
Disulfide bond restraints	0
Total dihedral-angle restraints	22
Number of unmapped restraints	39
Number of restraints per residue	24.5
Number of long range restraints per residue[1]	3.0

[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)	1.0	0.19
0.2-0.5 (Medium)	None	None
>0.5 (Large)	None	None

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
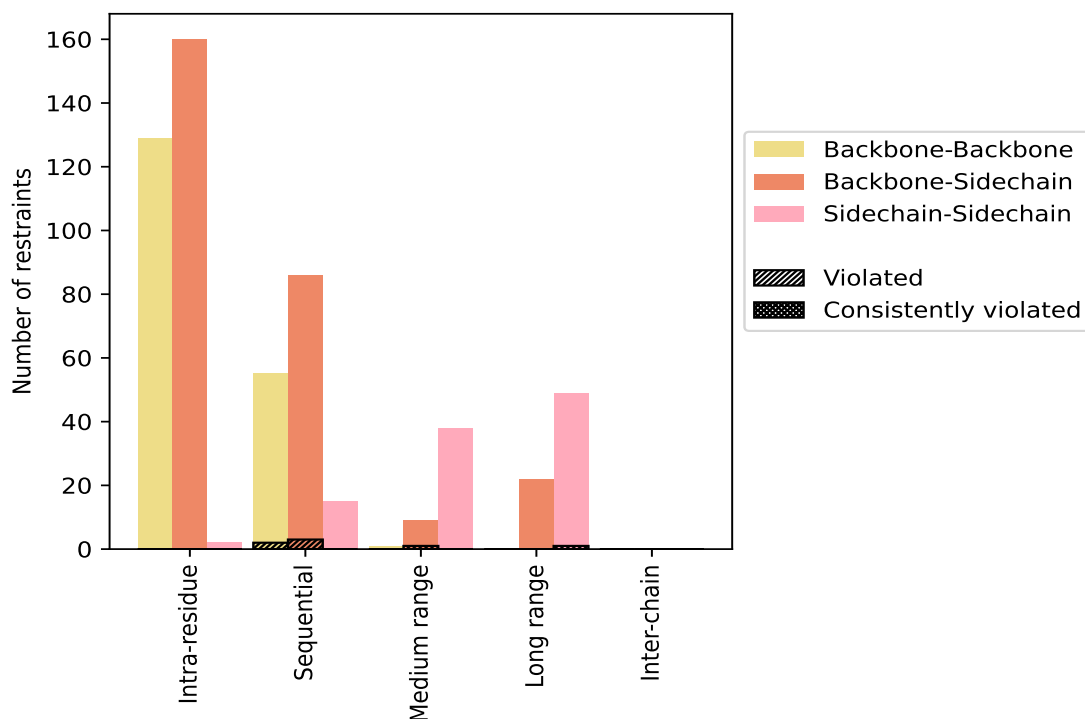
9   Distance violation analysis ⓘ

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)	291	51.4	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	129	22.8	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	160	28.3	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	2	0.4	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	156	27.6	5	3.2	0.9	0	0.0	0.0
Backbone-Backbone	55	9.7	2	3.6	0.4	0	0.0	0.0
Backbone-Sidechain	86	15.2	3	3.5	0.5	0	0.0	0.0
Sidechain-Sidechain	15	2.7	0	0.0	0.0	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	24	4.2	1	4.2	0.2	0	0.0	0.0
Backbone-Backbone	1	0.2	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	9	1.6	1	11.1	0.2	0	0.0	0.0
Sidechain-Sidechain	14	2.5	0	0.0	0.0	0	0.0	0.0
Long range (|i-j|≥5)	47	8.3	1	2.1	0.2	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	22	3.9	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	25	4.4	1	4.0	0.2	0	0.0	0.0
Inter-chain	0	0.0	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	0	0.0	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
Hydrogen bond	48	8.5	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	566	100.0	7	1.2	1.2	0	0.0	0.0
Backbone-Backbone	185	32.7	2	1.1	0.4	0	0.0	0.0
Backbone-Sidechain	277	48.9	4	1.4	0.7	0	0.0	0.0
Sidechain-Sidechain	104	18.4	1	1.0	0.2	0	0.0	0.0

[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
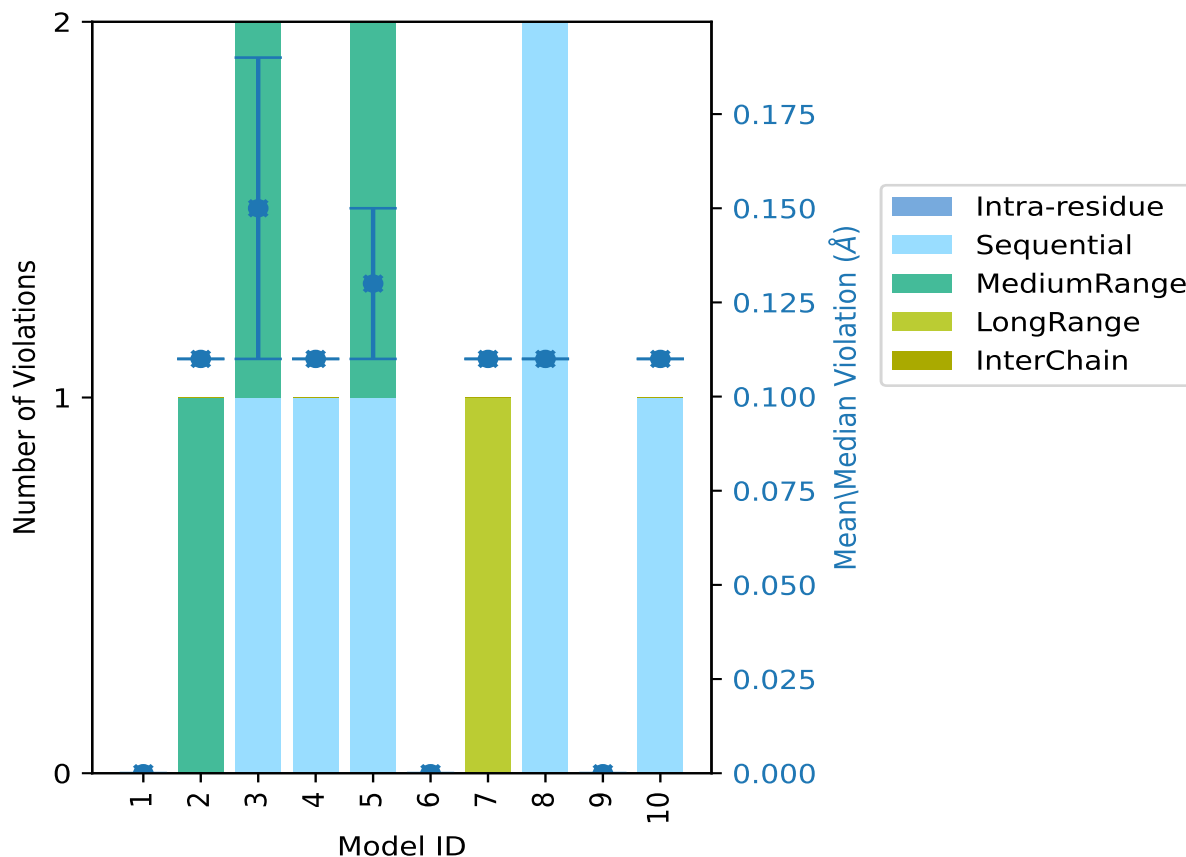
9.2 Distance violation statistics for each model ⓘ

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	0	0	0	0.0	0.0	0.0	0.0
2	0	0	1	0	0	1	0.11	0.11	0.0	0.11
3	0	1	1	0	0	2	0.15	0.19	0.04	0.15
4	0	1	0	0	0	1	0.11	0.11	0.0	0.11
5	0	1	1	0	0	2	0.13	0.15	0.02	0.13
6	0	0	0	0	0	0	0.0	0.0	0.0	0.0
7	0	0	0	1	0	1	0.11	0.11	0.0	0.11
8	0	2	0	0	0	2	0.11	0.11	0.0	0.11
9	0	0	0	0	0	0	0.0	0.0	0.0	0.0
10	0	1	0	0	0	1	0.11	0.11	0.0	0.11

[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 (i)

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 (i)

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, 511(IR:291, SQ:151, MR:23, LR:46, IC:0) 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	4	0	1	0	5	1	10.0
0	1	0	0	0	1	2	20.0
0	0	1	0	0	1	3	30.0

Continued on next page...

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

Continued from previous page...

Number of violated restraints						Fraction of the ensemble	
IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total	Count[6]	%
0	0	0	0	0	0	4	40.0
0	0	0	0	0	0	5	50.0
0	0	0	0	0	0	6	60.0
0	0	0	0	0	0	7	70.0
0	0	0	0	0	0	8	80.0
0	0	0	0	0	0	9	90.0
0	0	0	0	0	0	10	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 ⓘ
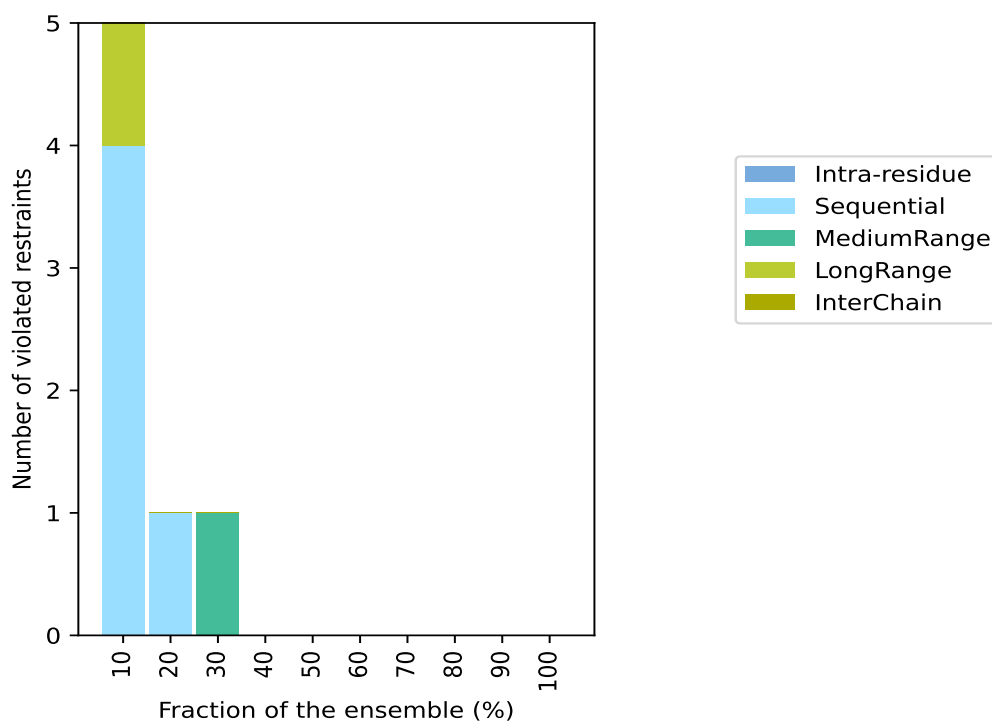
9.4   Most violated distance restraints in the ensemble ⓘ

9.4.1   Histogram : Distribution of mean distance violations ⓘ

The following histogram shows the distribution of the average value of the violation. The average is calculated for each restraint that is violated in more than one model over all the violated models

in the ensemble
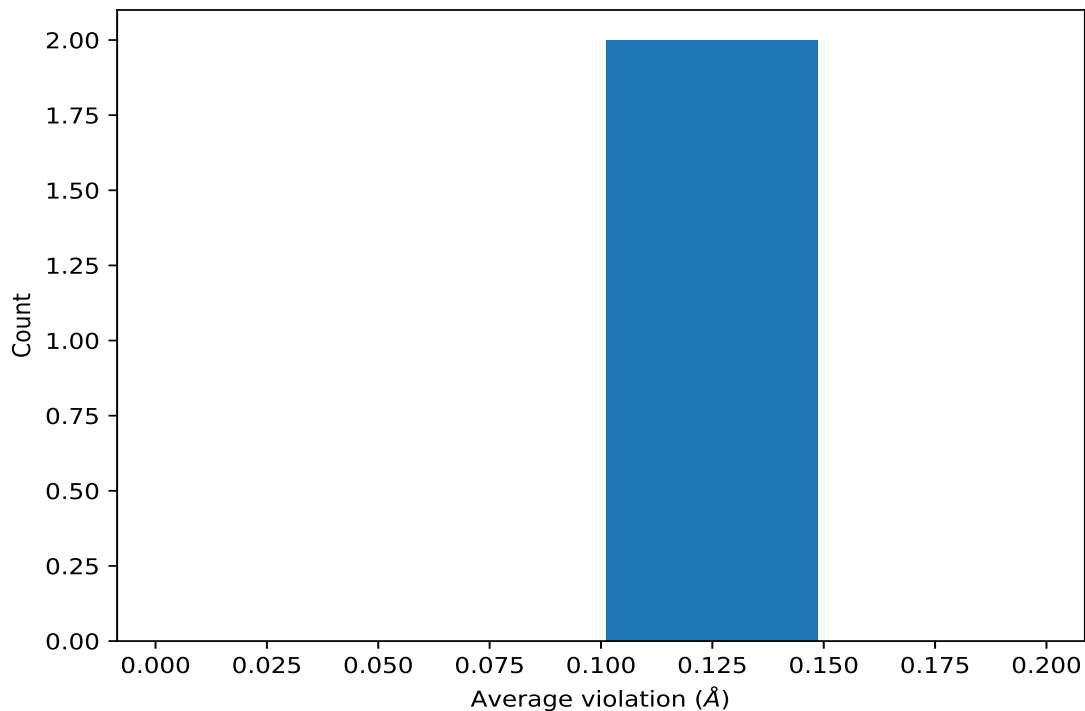
9.4.2 Table: Most violated distance restraints ⓘ

The following table provides the mean and the standard deviation of the violation for each restraint sorted by number of violated models and the mean 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	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,350)	1:17:X:DG:H2"	1:15:X:DA:H8	3	0.12	0.02	0.11
(1,320)	1:16:X:DA:H2'	1:15:X:DA:H1'	2	0.15	0.04	0.15

[1]Number of violated models, [2]Standard deviation

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.
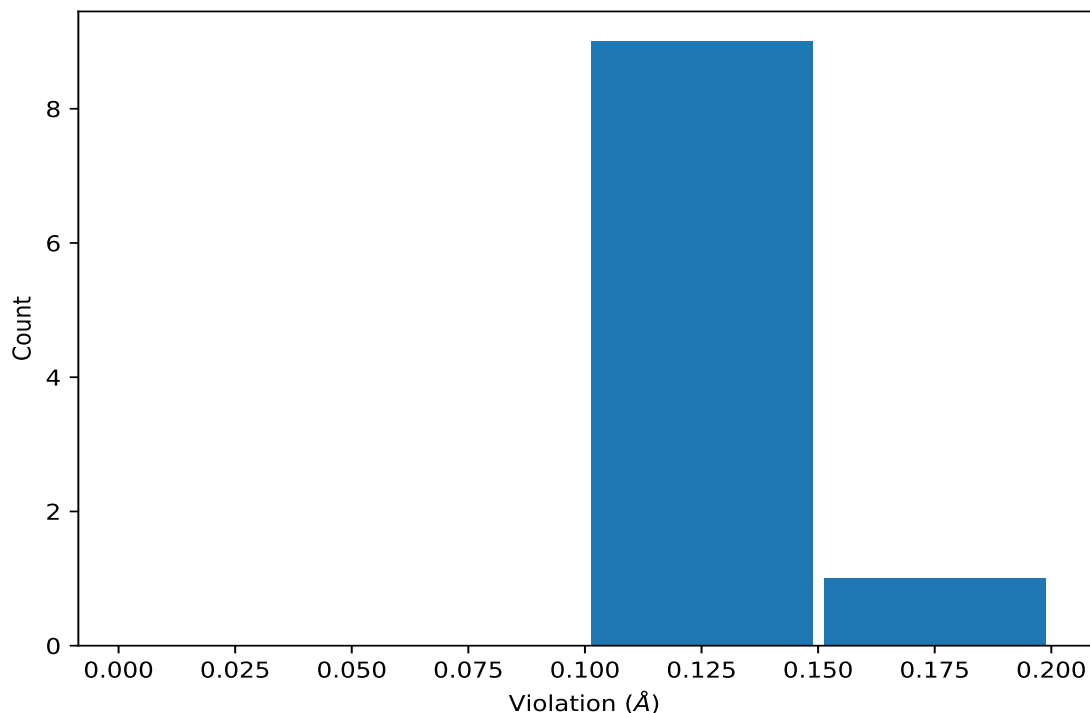
9.5.2　Table : All distance violations ⓘ

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)
(1,320)	1:16:X:DA:H2'	1:15:X:DA:H1'	3	0.19
(1,350)	1:17:X:DG:H2"	1:15:X:DA:H8	5	0.15
(1,438)	1:20:X:DA:H2	1:14:X:DG:H1	7	0.11
(1,350)	1:17:X:DG:H2"	1:15:X:DA:H8	2	0.11
(1,350)	1:17:X:DG:H2"	1:15:X:DA:H8	3	0.11
(1,320)	1:16:X:DA:H2'	1:15:X:DA:H1'	8	0.11
(1,317)	1:15:X:DA:H5"	1:16:X:DA:H8	10	0.11
(1,302)	1:15:X:DA:H2'	1:16:X:DA:H8	4	0.11
(1,274)	1:14:X:DG:H1'	1:15:X:DA:H2	5	0.11
(1,109)	1:5:X:DG:H5'	1:4:X:DG:H1'	8	0.11

10 Dihedral-angle violation analysis (i)

10.1 Summary of dihedral-angle violations (i)

The following table provides the summary of dihedral-angle violations in different dihedral-angle types. Violations less than 1° are not included in the calculation.

Angle type	Count	$\%^1$	Violated[3]			Consistently Violated[4]		
			Count	$\%^2$	$\%^1$	Count	$\%^2$	$\%^1$
CHI	22	100.0	0	0.0	0.0	0	0.0	0.0
Total	22	100.0	0	0.0	0.0	0	0.0	0.0

[1] percentage calculated with respect to total number of dihedral-angle restraints, [2] percentage calculated with respect to number of restraints in a particular dihedral-angle type, [3] violated in at least one model, [4] violated in all the models

10.1.1 Bar chart : Distribution of dihedral-angles and violations (i)
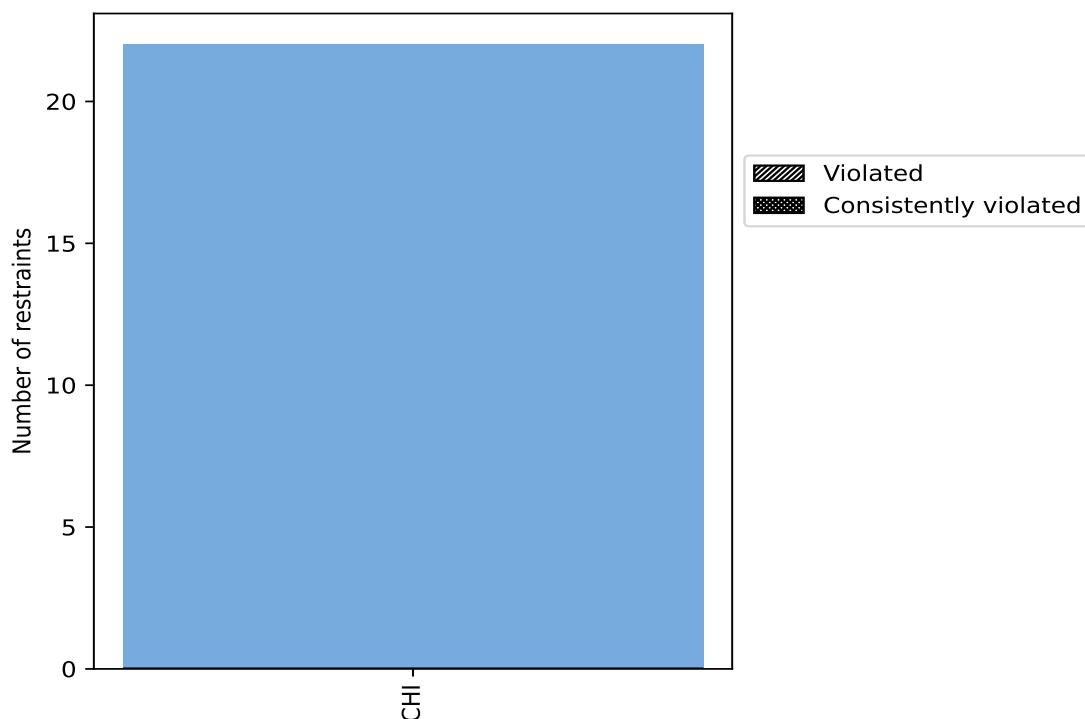
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

10.2 Dihedral-angle violation statistics for each model (i)

No violations found

10.3   Dihedral-angle violation statistics for the ensemble ⓘ

No violations found

10.4   Most violated dihedral-angle restraints in the ensemble ⓘ

No violations found

10.5   All violated dihedral-angle restraints ⓘ

No violations found