wwPDB NMR Structure Validation Summary Report

Apr 11, 2024 – 05:05 PM EDT

PDB ID	:	2MME
EMDB ID	:	EMD-5352
BMRB ID	:	18651
Title	:	Hybrid structure of the Shigella flexneri MxiH Type three secretion system needle
Authors	:	Demers, J.P.; Habenstein, B.; Loquet, A.; Vasa, S.K.; Becker, S.; Baker, D.; Lange, A.; Sgourakis, N.G.
Deposited on	:	2014-03-14

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:

EMDB validation analysis	:	NOT EXECUTED
MolProbity	:	4.02b-467
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)
Validation Pipeline (wwPDB-VP)	:	2.36.1

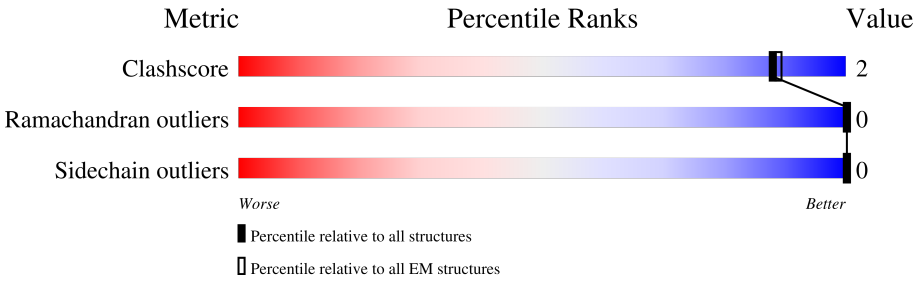
1 Overall quality at a glance

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

The reported resolution of this entry is 7.70 Å.

The overall completeness of chemical shifts assignment is 1%.

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	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	85	85% 13% .
1	B	85	95% . .
1	C	85	95% . .
1	D	85	95% . .
1	E	85	95% . .
1	F	85	95% . .
1	G	85	95% . .
1	H	85	95% . .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	I	85	95%
1	J	85	95%
1	K	85	95%
1	L	85	95%
1	M	85	95%
1	N	85	95%
1	O	85	95%
1	P	85	95%
1	Q	85	95%
1	R	85	95%
1	S	85	95%
1	T	85	95%
1	U	85	95%
1	V	85	95%
1	W	85	95%
1	X	85	95%
1	Y	85	95%
1	Z	85	95%
1	a	85	98%
1	b	85	98%
1	c	85	98%

2 Ensemble composition and analysis

This entry contains 10 models. Model 4 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:12-A:83, B:1-B:83, C:1-C:83, D:1-D:83, E:1-E:83, F:1-F:83, G:1-G:83, H:1-H:83, I:1-I:83, J:1-J:83, K:1-K:83, L:1-L:83, M:1-M:83, N:1-N:83, O:1-O:83, P:1-P:83, Q:1-Q:83, R:1-R:83, S:1-S:83, T:1-T:83, U:1-U:83, V:1-V:83, W:1-W:83, X:1-X:83, Y:1-Y:83, Z:1-Z:83, a:1-a:83, b:1-b:83, c:1-c:83 (2396)	0.23	4

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

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

NmrClust was unable to cluster the ensemble.

Error message: No clusters in NmrClust output

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 37120 atoms, of which 18357 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 MxiH.

Mol	Chain	Residues	Atoms						AltConf	Trace
1	A	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	B	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	C	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	D	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	E	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	F	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	G	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	H	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	I	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	J	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	K	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	L	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	M	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	N	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	O	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	P	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	Q	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms						AltConf	Trace
1	R	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	S	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	T	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	U	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	V	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	W	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	X	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	Y	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	Z	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	a	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	b	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		
1	c	83	Total	C	H	N	O	S	0	
			1280	402	633	107	137	1		

There are 58 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP Q6XVY0
A	0	HIS	-	expression tag	UNP Q6XVY0
B	-1	GLY	-	expression tag	UNP Q6XVY0
B	0	HIS	-	expression tag	UNP Q6XVY0
C	-1	GLY	-	expression tag	UNP Q6XVY0
C	0	HIS	-	expression tag	UNP Q6XVY0
D	-1	GLY	-	expression tag	UNP Q6XVY0
D	0	HIS	-	expression tag	UNP Q6XVY0
E	-1	GLY	-	expression tag	UNP Q6XVY0
E	0	HIS	-	expression tag	UNP Q6XVY0
F	-1	GLY	-	expression tag	UNP Q6XVY0
F	0	HIS	-	expression tag	UNP Q6XVY0
G	-1	GLY	-	expression tag	UNP Q6XVY0
G	0	HIS	-	expression tag	UNP Q6XVY0
H	-1	GLY	-	expression tag	UNP Q6XVY0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	0	HIS	-	expression tag	UNP Q6XVY0
I	-1	GLY	-	expression tag	UNP Q6XVY0
I	0	HIS	-	expression tag	UNP Q6XVY0
J	-1	GLY	-	expression tag	UNP Q6XVY0
J	0	HIS	-	expression tag	UNP Q6XVY0
K	-1	GLY	-	expression tag	UNP Q6XVY0
K	0	HIS	-	expression tag	UNP Q6XVY0
L	-1	GLY	-	expression tag	UNP Q6XVY0
L	0	HIS	-	expression tag	UNP Q6XVY0
M	-1	GLY	-	expression tag	UNP Q6XVY0
M	0	HIS	-	expression tag	UNP Q6XVY0
N	-1	GLY	-	expression tag	UNP Q6XVY0
N	0	HIS	-	expression tag	UNP Q6XVY0
O	-1	GLY	-	expression tag	UNP Q6XVY0
O	0	HIS	-	expression tag	UNP Q6XVY0
P	-1	GLY	-	expression tag	UNP Q6XVY0
P	0	HIS	-	expression tag	UNP Q6XVY0
Q	-1	GLY	-	expression tag	UNP Q6XVY0
Q	0	HIS	-	expression tag	UNP Q6XVY0
R	-1	GLY	-	expression tag	UNP Q6XVY0
R	0	HIS	-	expression tag	UNP Q6XVY0
S	-1	GLY	-	expression tag	UNP Q6XVY0
S	0	HIS	-	expression tag	UNP Q6XVY0
T	-1	GLY	-	expression tag	UNP Q6XVY0
T	0	HIS	-	expression tag	UNP Q6XVY0
U	-1	GLY	-	expression tag	UNP Q6XVY0
U	0	HIS	-	expression tag	UNP Q6XVY0
V	-1	GLY	-	expression tag	UNP Q6XVY0
V	0	HIS	-	expression tag	UNP Q6XVY0
W	-1	GLY	-	expression tag	UNP Q6XVY0
W	0	HIS	-	expression tag	UNP Q6XVY0
X	-1	GLY	-	expression tag	UNP Q6XVY0
X	0	HIS	-	expression tag	UNP Q6XVY0
Y	-1	GLY	-	expression tag	UNP Q6XVY0
Y	0	HIS	-	expression tag	UNP Q6XVY0
Z	-1	GLY	-	expression tag	UNP Q6XVY0
Z	0	HIS	-	expression tag	UNP Q6XVY0
a	-1	GLY	-	expression tag	UNP Q6XVY0
a	0	HIS	-	expression tag	UNP Q6XVY0
b	-1	GLY	-	expression tag	UNP Q6XVY0
b	0	HIS	-	expression tag	UNP Q6XVY0
c	-1	GLY	-	expression tag	UNP Q6XVY0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
c	0	HIS	-	expression tag	UNP Q6XVY0

4 Residue-property plots

4.1 Average score per residue in the NMR ensemble

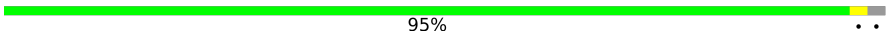
These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for 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.

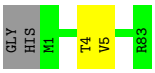
• Molecule 1: MxiH

Chain A:

• Molecule 1: MxiH

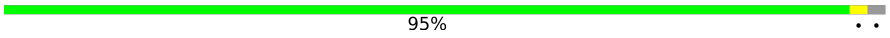
Chain B:

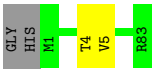




• Molecule 1: MxiH

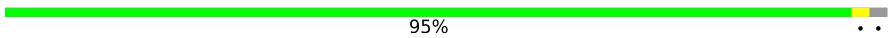
Chain C:

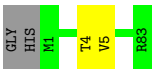




• Molecule 1: MxiH

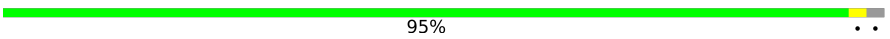
Chain D:

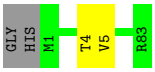




• Molecule 1: MxiH

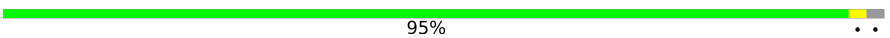
Chain E:

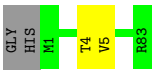




• Molecule 1: MxiH

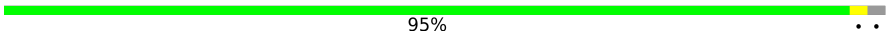
Chain F:

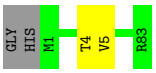




• Molecule 1: MxiH

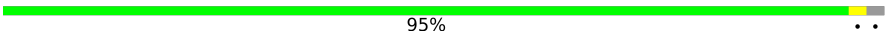
Chain G:

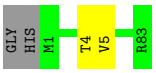




• Molecule 1: MxiH

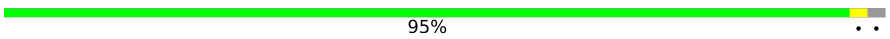
Chain H:

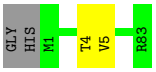




• Molecule 1: MxiH

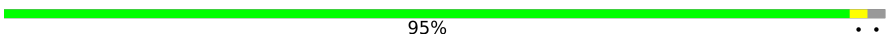
Chain I:

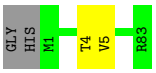




• Molecule 1: MxiH

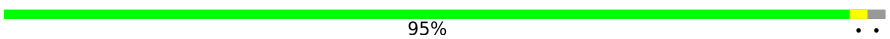
Chain J:

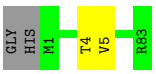




• Molecule 1: MxiH

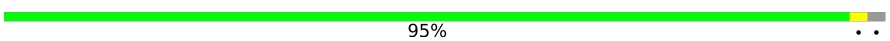
Chain K:

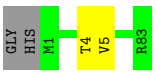




• Molecule 1: MxiH

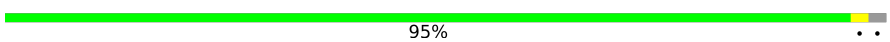
Chain L:

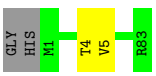




• Molecule 1: MxiH

Chain M:



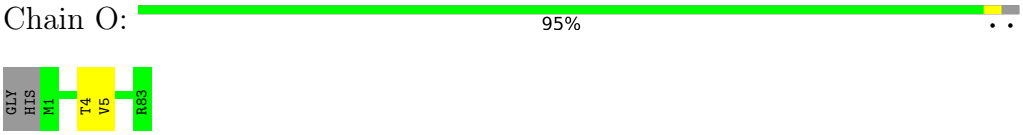


• Molecule 1: MxiH

Chain N:

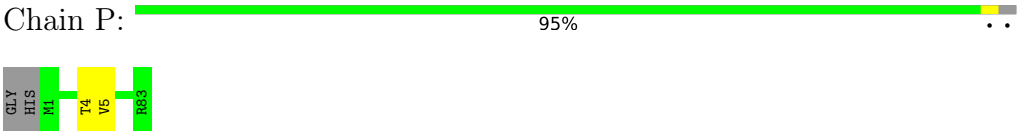
• Molecule 1: MxiH

Chain O:



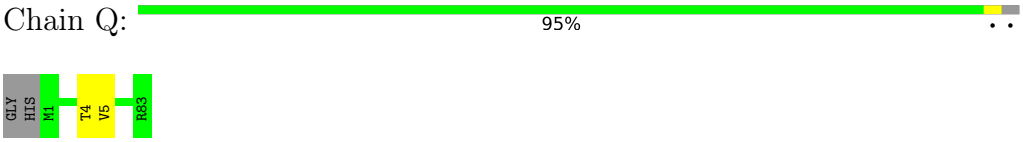
• Molecule 1: MxiH

Chain P:



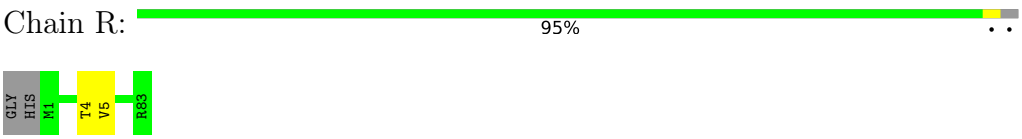
• Molecule 1: MxiH

Chain Q:



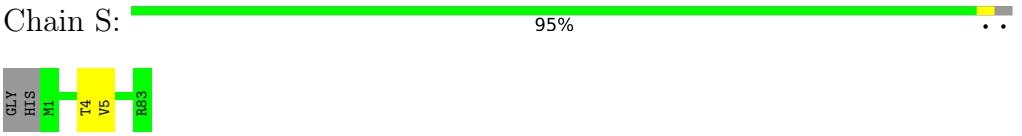
• Molecule 1: MxiH

Chain R:



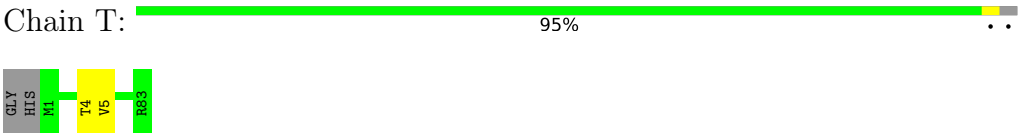
• Molecule 1: MxiH

Chain S:



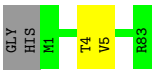
• Molecule 1: MxiH

Chain T:

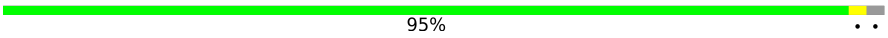


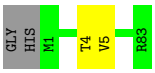
• Molecule 1: MxiH

Chain U:

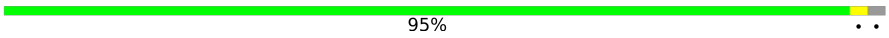


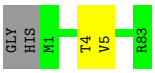
• Molecule 1: MxiH

Chain V: 

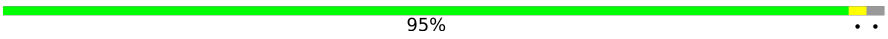


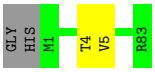
• Molecule 1: MxiH

Chain W: 

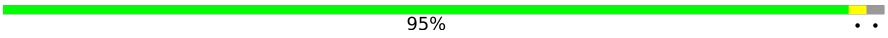


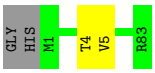
• Molecule 1: MxiH

Chain X: 

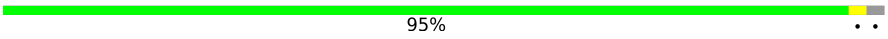


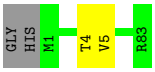
• Molecule 1: MxiH

Chain Y: 

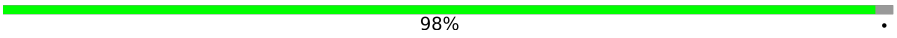


• Molecule 1: MxiH

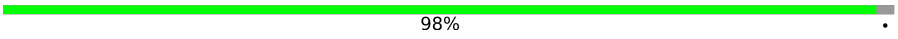
Chain Z: 



• Molecule 1: MxiH

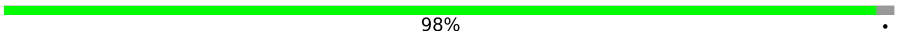
Chain a: 

• Molecule 1: MxiH

Chain b: 

• Molecule 1: MxiH

Chain c:

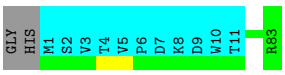


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 4. Colouring as in section 4.1 above.

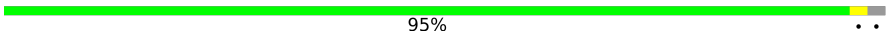
- Molecule 1: MxiH

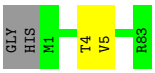
Chain A:



- Molecule 1: MxiH

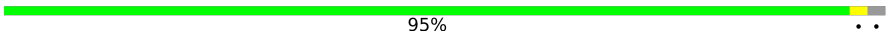
Chain B:

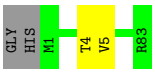




- Molecule 1: MxiH

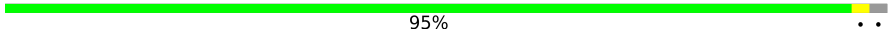
Chain C:

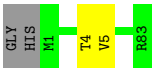




- Molecule 1: MxiH

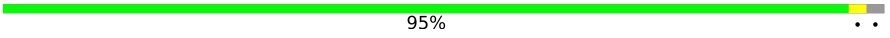
Chain D:

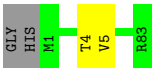




- Molecule 1: MxiH

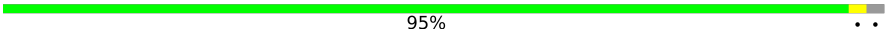
Chain E:

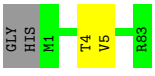




- Molecule 1: MxiH

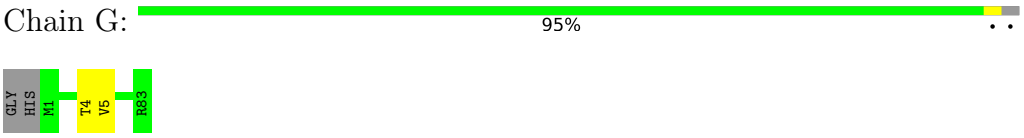
Chain F:





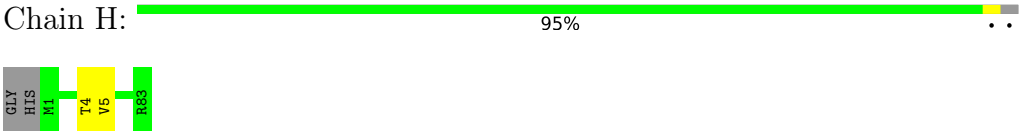
• Molecule 1: MxiH

Chain G:



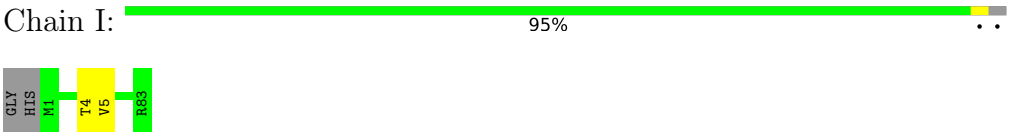
• Molecule 1: MxiH

Chain H:



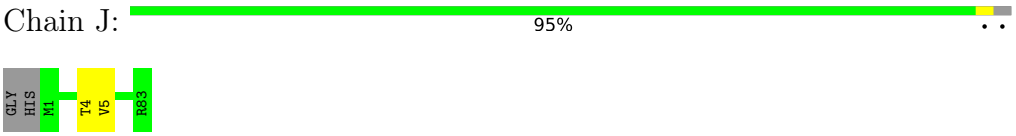
• Molecule 1: MxiH

Chain I:



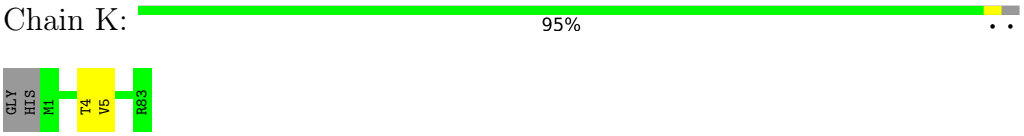
• Molecule 1: MxiH

Chain J:



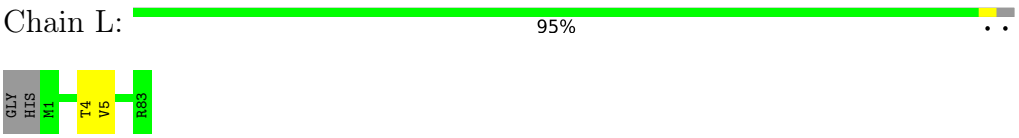
• Molecule 1: MxiH

Chain K:



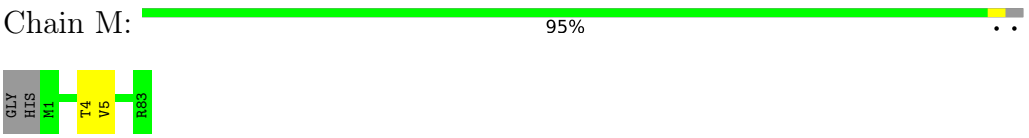
• Molecule 1: MxiH

Chain L:



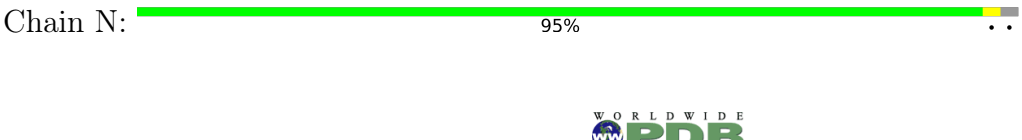
• Molecule 1: MxiH

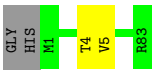
Chain M:



• Molecule 1: MxiH

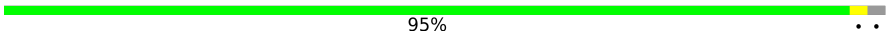
Chain N:

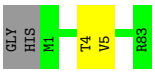




• Molecule 1: MxiH

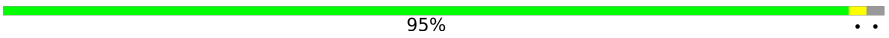
Chain O:

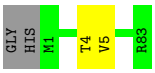




• Molecule 1: MxiH

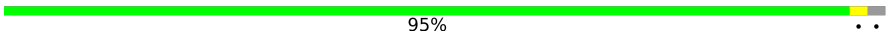
Chain P:

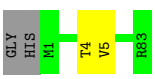




• Molecule 1: MxiH

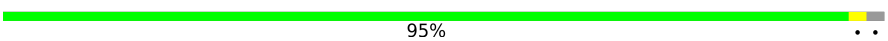
Chain Q:

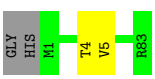




• Molecule 1: MxiH

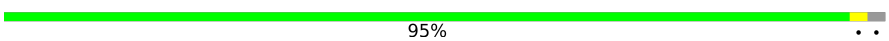
Chain R:

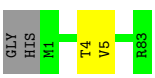




• Molecule 1: MxiH

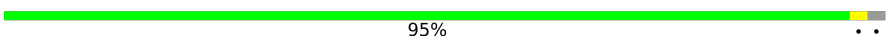
Chain S:

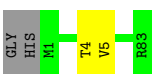




• Molecule 1: MxiH

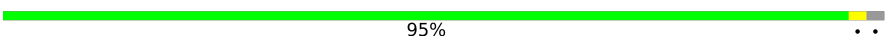
Chain T:

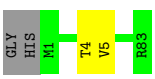




• Molecule 1: MxiH

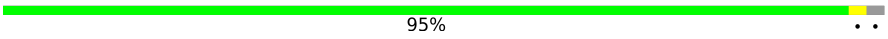
Chain U:

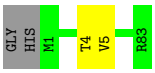




• Molecule 1: MxiH

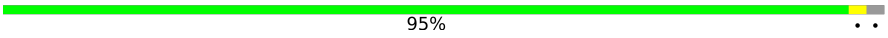
Chain V:

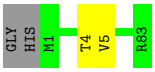




• Molecule 1: MxiH

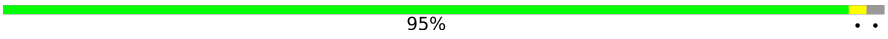
Chain W:

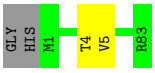




• Molecule 1: MxiH

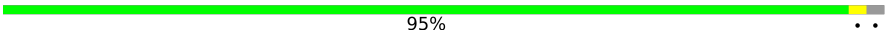
Chain X:

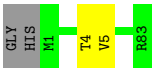




• Molecule 1: MxiH

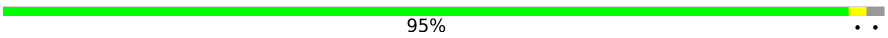
Chain Y:

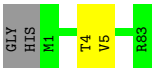




• Molecule 1: MxiH

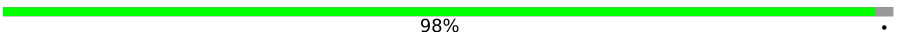
Chain Z:





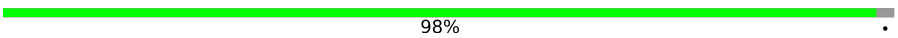
• Molecule 1: MxiH

Chain a:



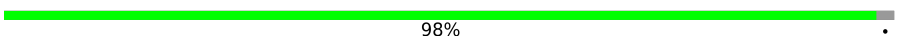
• Molecule 1: MxiH

Chain b:



• Molecule 1: MxiH

Chain c:



5 Refinement protocol and experimental data overview

The models were refined using the following method: *Rosetta fold-and-dock, Rosetta symmetric relax.*

Of the 5000 calculated structures, 10 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
Rosetta	refinement	

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	487
Number of shifts mapped to atoms	485
Number of unparsed shifts	0
Number of shifts with mapping errors	2
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	1%

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

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

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

6.2 Too-close contacts

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	559	549	549	0±0
1	B	647	633	635	3±1
1	C	647	633	635	3±1
1	D	647	633	635	3±1
1	E	647	633	635	3±0
1	F	647	633	635	3±1
1	G	647	633	635	3±1
1	H	647	633	635	3±1
1	I	647	633	635	3±0
1	J	647	633	635	3±1
1	K	647	633	635	3±1
1	L	647	633	635	3±1
1	M	647	633	635	3±1
1	N	647	633	635	3±0
1	O	647	633	635	3±1
1	P	647	633	635	3±1
1	Q	647	633	635	3±1
1	R	647	633	635	3±1
1	S	647	633	635	3±1
1	T	647	633	635	3±1
1	U	647	633	635	3±1
1	V	647	633	635	3±1
1	W	647	633	635	3±1
1	X	647	633	635	3±1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	Y	647	633	635	3±1
1	Z	647	633	635	3±1
All	All	186750	182730	183290	702

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 2.

5 of 111 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:N:5:VAL:O	1:N:5:VAL:HG13	0.61	1.96	1	9
1:J:5:VAL:O	1:J:5:VAL:HG13	0.61	1.96	1	9
1:K:5:VAL:HG13	1:K:5:VAL:O	0.61	1.96	1	9
1:S:5:VAL:HG13	1:S:5:VAL:O	0.61	1.96	1	9
1:E:5:VAL:O	1:E:5:VAL:HG13	0.61	1.96	1	9

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	71/85 (84%)	70±0 (99±0%)	1±0 (1±0%)	0±0 (0±0%)	100	100
1	B	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	C	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	D	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	E	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	F	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	G	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	H	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	I	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	J	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	K	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	L	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	M	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	N	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	O	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	P	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	Q	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	R	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	S	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	T	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	U	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	V	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	W	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	X	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	Y	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	Z	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	a	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	b	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
1	c	81/85 (95%)	78±0 (97±1%)	2±0 (3±1%)	0±0 (0±0%)	100	100
All	All	23390/24650 (95%)	22680 (97%)	710 (3%)	0 (0%)	100	100

There are no Ramachandran outliers.

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	63/75 (84%)	63±0 (100±0%)	0±0 (0±0%)	100	100
1	B	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	C	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	D	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	F	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	G	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	H	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	I	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	J	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	K	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	L	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	M	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	N	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	O	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	P	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	Q	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	R	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	S	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	T	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	U	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	V	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	W	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	X	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	Y	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	Z	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	a	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	b	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
1	c	74/75 (99%)	74±0 (100±0%)	0±0 (0±0%)	100	100
All	All	21350/21750 (98%)	21350 (100%)	0 (0%)	100	100

There are no protein residues with a non-rotameric sidechain to report.

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)

There are no ligands in this entry.

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

The completeness of assignment taking into account all chemical shift lists is 1% for the well-defined parts and 1% 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

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	487
Number of shifts mapped to atoms	485
Number of unparsed shifts	0
Number of shifts with mapping errors	2
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	1

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

- No matching atom found in the structure. All 2 occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	0	HIS	CD2	119.938	0.051	1
1	A	0	HIS	CE1	136.657	0.048	1

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	82	-0.82 ± 0.15	Should be checked
$^{13}C_{\beta}$	80	0.41 ± 0.21	None needed (< 0.5 ppm)
$^{13}C'$	81	-0.55 ± 0.21	Should be applied
^{15}N	82	0.03 ± 0.38	None needed (< 0.5 ppm)

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 1%, i.e. 415 atoms were assigned a chemical shift out of a possible 31728. 0 out of 433 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	213/11866 (2%)	0/4764 (0%)	143/4792 (3%)	70/2310 (3%)
Sidechain	187/18163 (1%)	0/11781 (0%)	171/5716 (3%)	16/666 (2%)
Aromatic	15/1699 (1%)	0/806 (0%)	15/865 (2%)	0/28 (0%)
Overall	415/31728 (1%)	0/17351 (0%)	329/11373 (3%)	86/3004 (3%)

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

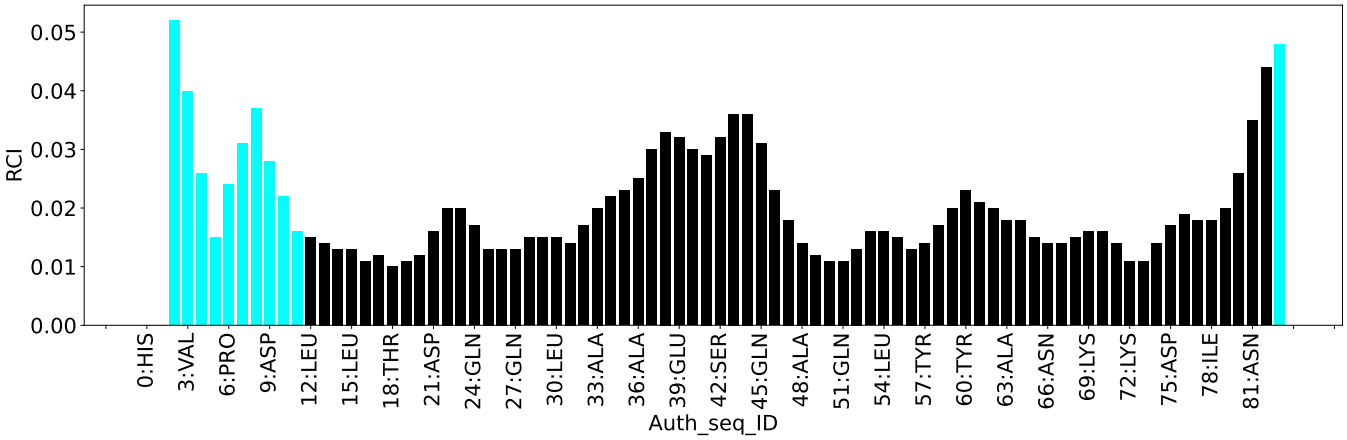
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	64	GLN	NE2	122.66	103.38 – 120.35	6.4

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:


RCI
0.05
0.04
0.03
0.02
0.01
0.00
0:HIS
3:VAL
6:PRO
9:ASP
12:LEU
15:LEU
18:THR
21:ASP
24:GLN
27:GLN
30:LEU
33:ALA
36:ALA
39:GLU
42:SER
45:GLN
48:ALA
51:GLN
54:LEU
57:TYR
60:TYR
63:ALA
66:ASN
69:LYS
72:LYS
75:ASP
78:ILE
81:ASN
Auth_seq_ID

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	988
Intra-residue (|i-j|=0)	0
Sequential (|i-j|=1)	0
Medium range (|i-j|>1 and |i-j|<5)	387
Long range (|i-j|≥5)	601
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	988
Number of restraints per residue	0.4
Number of long range restraints per residue[1]	0.2

[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)	0.6	0.47
>0.5 (Large)	46.9	39.47

8.2.2 Average number of dihedral-angle violations per model (i)

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

9 Distance violation analysis (i)

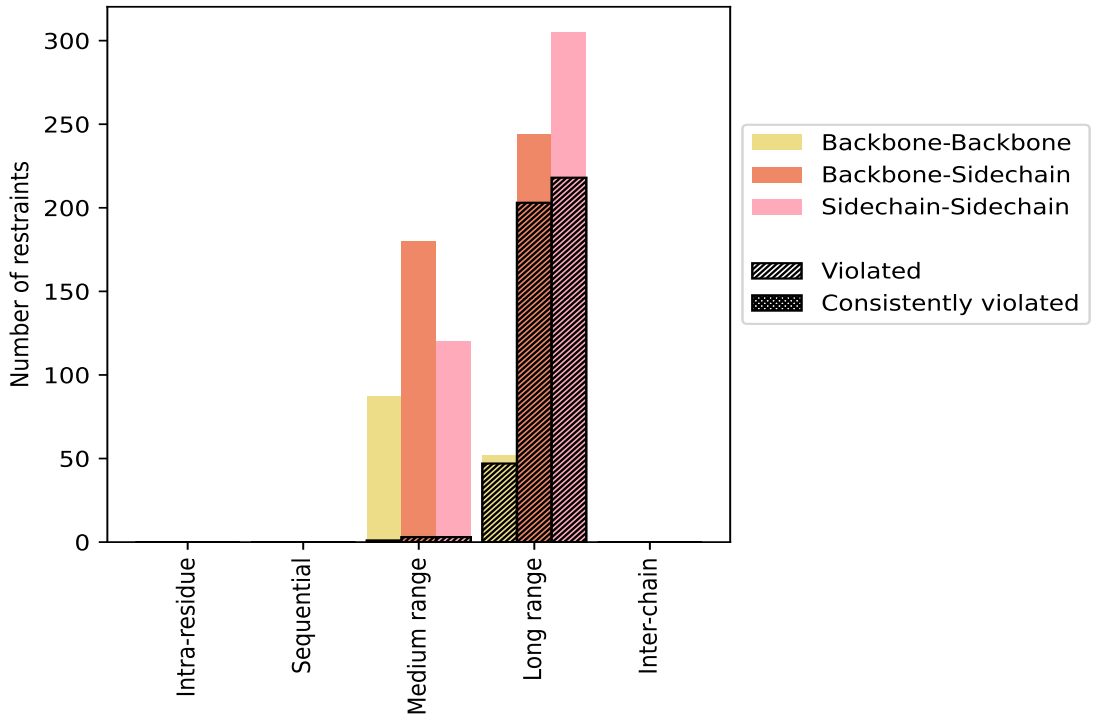
9.1 Summary of distance violations (i)

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)	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
Sequential (|i-j|=1)	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
Medium range (|i-j|>1 & |i-j|<5)	387	39.2	7	1.8	0.7	0	0.0	0.0
Backbone-Backbone	87	8.8	1	1.1	0.1	0	0.0	0.0
Backbone-Sidechain	180	18.2	3	1.7	0.3	0	0.0	0.0
Sidechain-Sidechain	120	12.1	3	2.5	0.3	0	0.0	0.0
Long range (|i-j|≥5)	601	60.8	468	77.9	47.4	0	0.0	0.0
Backbone-Backbone	52	5.3	47	90.4	4.8	0	0.0	0.0
Backbone-Sidechain	244	24.7	203	83.2	20.5	0	0.0	0.0
Sidechain-Sidechain	305	30.9	218	71.5	22.1	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	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	988	100.0	475	48.1	48.1	0	0.0	0.0
Backbone-Backbone	139	14.1	48	34.5	4.9	0	0.0	0.0
Backbone-Sidechain	424	42.9	206	48.6	20.9	0	0.0	0.0
Sidechain-Sidechain	425	43.0	221	52.0	22.4	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

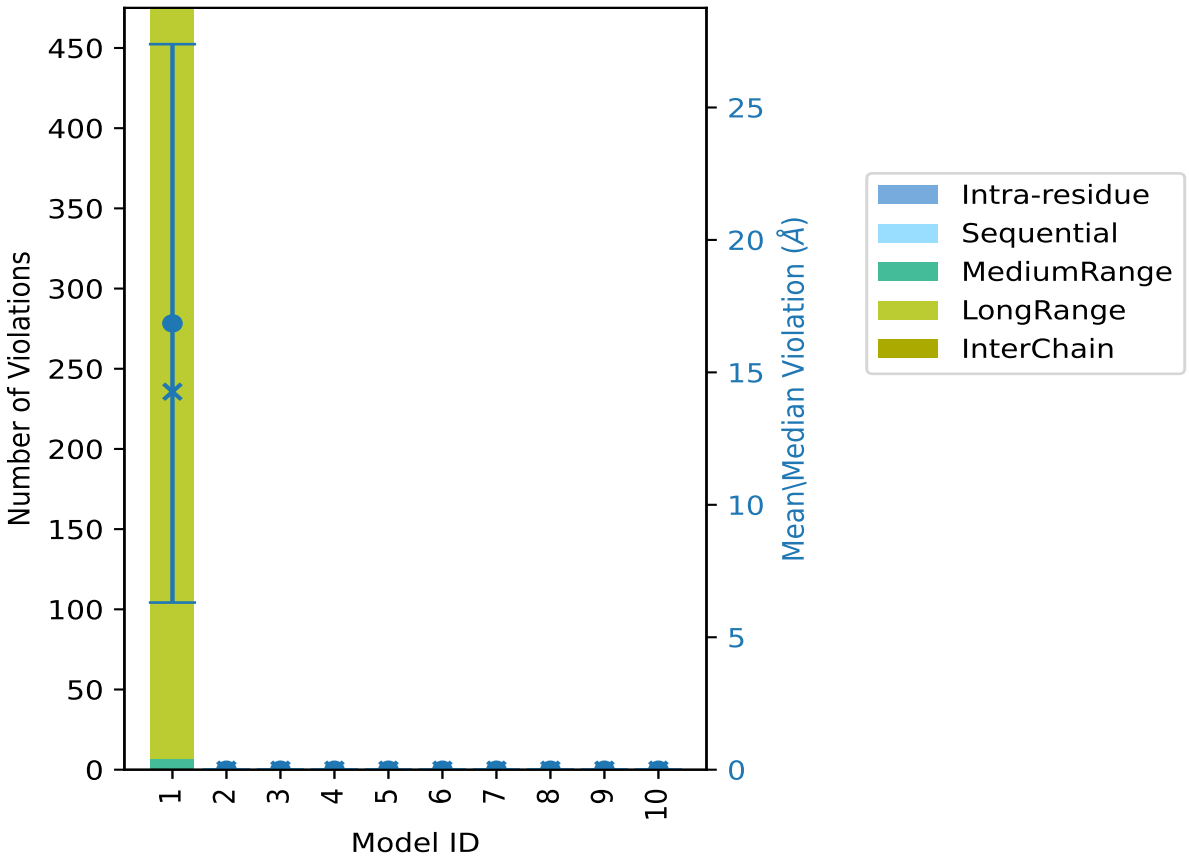
The following table provides the distance violation statistics for each model in the ensemble. Violations less than 0.1 Å are not included in the statistics.

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	7	468	0	475	16.85	39.47	10.54	14.27
2	0	0	0	0	0	0	0.0	0.0	0.0	0.0
3	0	0	0	0	0	0	0.0	0.0	0.0	0.0
4	0	0	0	0	0	0	0.0	0.0	0.0	0.0
5	0	0	0	0	0	0	0.0	0.0	0.0	0.0
6	0	0	0	0	0	0	0.0	0.0	0.0	0.0
7	0	0	0	0	0	0	0.0	0.0	0.0	0.0
8	0	0	0	0	0	0	0.0	0.0	0.0	0.0
9	0	0	0	0	0	0	0.0	0.0	0.0	0.0
10	0	0	0	0	0	0	0.0	0.0	0.0	0.0

[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, 513(IR:0, SQ:0, MR:380, LR:133, 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	0	7	468	0	475	1	10.0
0	0	0	0	0	0	2	20.0
0	0	0	0	0	0	3	30.0
0	0	0	0	0	0	4	40.0

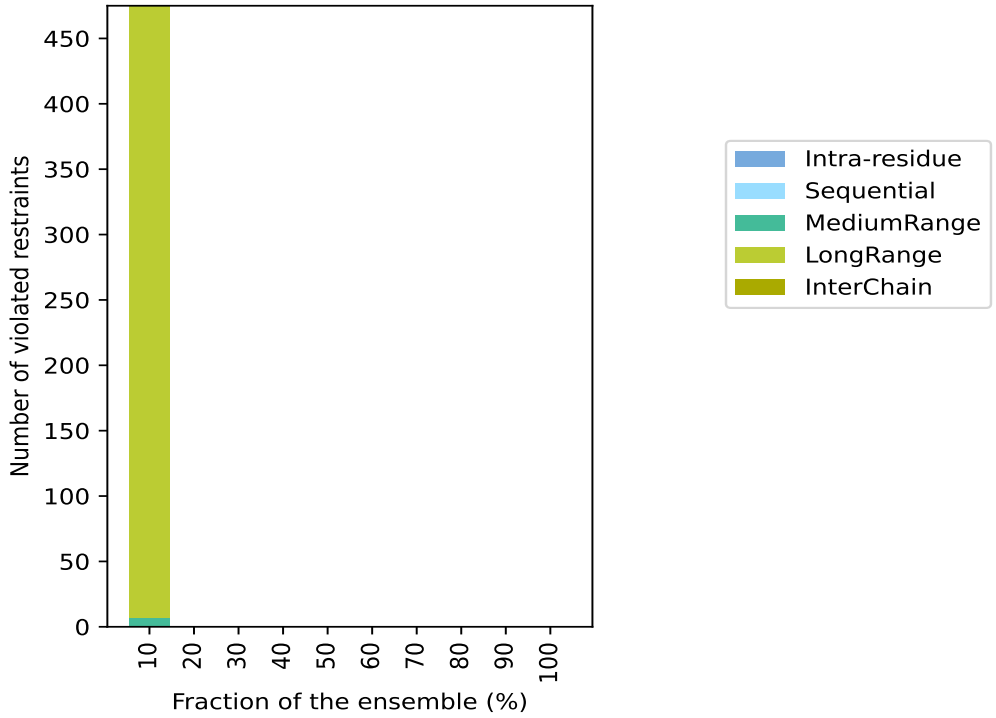
Continued on next page...

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	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



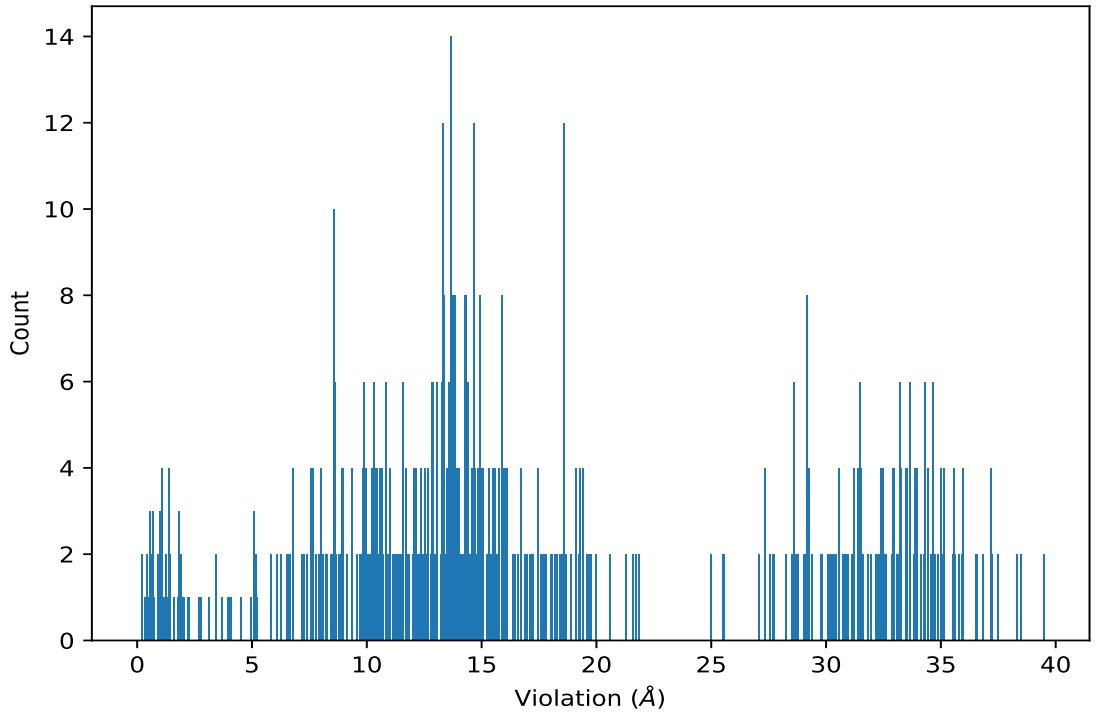
9.4 Most violated distance restraints in the ensemble

No violations found

9.5 All violated distance restraints (i)

9.5.1 Histogram : Distribution of distance violations (i)

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 (i)

The following table provides the 10 worst performing restraints, sorted by the violation 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)
(1,18)	1:48:A:ALA:CB	1:83:A:ARG:C	1	39.47
(1,18)	1:48:A:ALA:CB	1:83:A:ARG:C	1	39.47
(1,534)	1:44:A:PRO:CD	1:78:A:ILE:CB	1	38.48
(1,534)	1:44:A:PRO:CD	1:78:A:ILE:CB	1	38.48
(1,524)	1:44:A:PRO:CB	1:78:A:ILE:CG2	1	38.32
(1,524)	1:44:A:PRO:CB	1:78:A:ILE:CG2	1	38.32
(1,515)	1:44:A:PRO:CA	1:78:A:ILE:CG2	1	37.48
(1,515)	1:44:A:PRO:CA	1:78:A:ILE:CG2	1	37.48
(1,522)	1:44:A:PRO:CB	1:78:A:ILE:CB	1	37.23
(1,522)	1:44:A:PRO:CB	1:78:A:ILE:CB	1	37.23
(1,535)	1:44:A:PRO:CD	1:78:A:ILE:CD1	1	37.19

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,535)	1:44:A:PRO:CD	1:78:A:ILE:CD1	1	37.19
(1,543)	1:44:A:PRO:CG	1:78:A:ILE:CD1	1	37.15
(1,543)	1:44:A:PRO:CG	1:78:A:ILE:CD1	1	37.15
(1,347)	1:47:A:LEU:CD1	1:79:A:ILE:CA	1	36.82
(1,347)	1:47:A:LEU:CD1	1:79:A:ILE:CA	1	36.82
(1,424)	1:43:A:ASN:CA	1:78:A:ILE:CD1	1	36.59
(1,424)	1:43:A:ASN:CA	1:78:A:ILE:CD1	1	36.59
(1,346)	1:47:A:LEU:CD1	1:78:A:ILE:CG2	1	36.53

10 Dihedral-angle violation analysis (i)

No dihedral-angle restraints found