wwPDB NMR Structure Validation Summary Report (i)

Jun 6, 2023 – 01:16 AM EDT

PDB ID	:	2MCA
BMRB ID	:	19433
Title	:	NMR structure of the protein YP_002937094.1 from Eubacterium rectale
Authors	:	Proudfoot, A.; Serrano, P.; Geralt, M.; Dutta, S.; Wuthrich, K.; Joint Center for Structural Genomics (JCSG)
Deposited on	:	2013-08-19

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 (1)) were used in the production of this report:

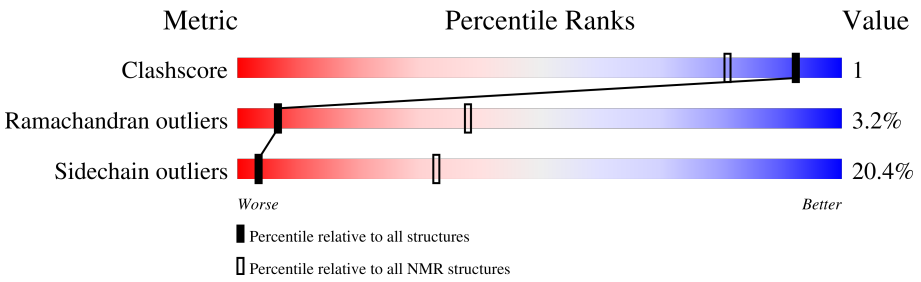
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.33

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment is 87%.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and a grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain
1	A	103	74% 15% • 11%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 15 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *closest to the average*.

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

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:8-A:99 (92)	0.56	15

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

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

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

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1454 atoms, of which 655 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Uncharacterized protein.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	103	1454	497	655	134	167	1	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP C4ZHE8

4 Residue-property plots (i)

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: Uncharacterized protein

Chain A: 74% 15% • 11%

G1 A2 Q3 D4 G5 K6 E7 T8 T9 R12 Y18 K22 N23 I24 K25 N29 K32 E33 Q34 N37 T38 L39 S49 Y55 T60 T81 G100 E101 I102 D103

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

The representative model is number 15. Colouring as in section 4.1 above.

- Molecule 1: Uncharacterized protein

Chain A: 71% 15% • 11%

G1 A2 Q3 D4 G5 K6 E7 N15 Q16 T17 Y18 K22 N23 N29 D30 G31 K32 E33 Q34 N37 T38 L39 H42 S49 D58 V59 T60 T81 L87 D88 E89 N90 G100 E101 I102 D103

5 Refinement protocol and experimental data overview (i)

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

Of the 80 calculated structures, 20 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
CYANA	refinement	
UNIO	structure solution	
OPALp	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	1144
Number of shifts mapped to atoms	1037
Number of unparsed shifts	0
Number of shifts with mapping errors	107
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	87%

6 Model quality (i)

6.1 Standard geometry (i)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the (average) root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	0.60±0.01	0±0/735 (0.0± 0.0%)	1.06±0.03	1±1/996 (0.1± 0.1%)
All	All	0.60	0/14700 (0.0%)	1.06	24/19920 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	Chirality	Planarity
1	A	0.0±0.0	0.8±0.7
All	All	0	15

There are no bond-length outliers.

5 of 12 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	A	95	VAL	CG1-CB-CG2	8.51	124.52	110.90	17	3
1	A	93	VAL	CA-CB-CG1	8.04	122.95	110.90	16	1
1	A	18	TYR	CB-CG-CD2	-7.42	116.55	121.00	14	8
1	A	58	ASP	CB-CG-OD1	6.73	124.36	118.30	15	2
1	A	18	TYR	CB-CG-CD1	6.39	124.83	121.00	6	2

There are no chirality outliers.

5 of 7 unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	12	ARG	Sidechain	5
1	A	82	ARG	Sidechain	5
1	A	49	SER	Peptide	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	59	VAL	Peptide	1
1	A	87	LEU	Peptide	1

6.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	722	602	700	2±2
All	All	14440	12040	14000	41

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 1.

5 of 25 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:71:GLY:HA2	1:A:95:VAL:HG11	0.69	1.65	10	1
1:A:15:ASN:HB2	1:A:39:LEU:HD12	0.63	1.71	12	5
1:A:71:GLY:HA3	1:A:95:VAL:HG11	0.62	1.71	13	2
1:A:71:GLY:CA	1:A:95:VAL:HG11	0.61	2.26	10	3
1:A:17:THR:HG22	1:A:87:LEU:HD12	0.61	1.71	19	3

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	92/103 (89%)	79±2 (85±2%)	10±2 (11±2%)	3±1 (3±1%)	7	38
All	All	1840/2060 (89%)	1572 (85%)	209 (11%)	59 (3%)	7	38

5 of 8 unique Ramachandran outliers are listed below. They are sorted by the frequency of

occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	23	ASN	20
1	A	37	ASN	15
1	A	89	GLU	8
1	A	31	GLY	6
1	A	72	HIS	6

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	82/89 (92%)	65±3 (80±3%)	17±3 (20±3%)	3	33
All	All	1640/1780 (92%)	1305 (80%)	335 (20%)	3	33

5 of 53 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	23	ASN	20
1	A	22	LYS	19
1	A	18	TYR	18
1	A	29	ASN	18
1	A	55	TYR	14

6.3.3 RNA (i)

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates (i)

There are no monosaccharides in this entry.

6.6 Ligand geometry (i)

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 (i)

The completeness of assignment taking into account all chemical shift lists is 87% for the well-defined parts and 85% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

7.1.1 Bookkeeping (i)

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	1144
Number of shifts mapped to atoms	1037
Number of unparsed shifts	0
Number of shifts with mapping errors	107
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	5

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

- No matching atom found in the structure. First 5 (of 107) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	3	GLN	HB2	1.921	0.025	2
1	A	3	GLN	HG2	2.291	0.025	2
1	A	4	ASP	HB2	2.63	0.025	2
1	A	6	LYS	HB2	1.78	0.025	2
1	A	7	GLU	HB2	1.849	0.025	2
1	A	7	GLU	HG2	2.15	0.025	2
1	A	11	ILE	HG12	0.01	0.025	2
1	A	12	ARG	HB2	1.278	0.025	2
1	A	12	ARG	HG2	1.284	0.025	2
1	A	12	ARG	HD2	2.671	0.025	2
1	A	13	LEU	HB2	1.478	0.025	2
1	A	14	ILE	HG12	0.621	0.025	2
1	A	15	ASN	HB2	2.115	0.025	2
1	A	16	GLN	HB2	2.341	0.025	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	16	GLN	HG2	2.336	0.025	2
1	A	18	TYR	HB2	2.901	0.025	2
1	A	19	PHE	HB2	2.986	0.025	2
1	A	20	ASN	HB2	2.532	0.025	2
1	A	22	LYS	HB2	1.666	0.025	2
1	A	22	LYS	HG2	1.025	0.025	2
1	A	22	LYS	HD2	1.511	0.025	2
1	A	22	LYS	HE2	2.783	0.025	2
1	A	23	ASN	HB2	2.604	0.025	2
1	A	24	ILE	HG12	0.824	0.025	2
1	A	25	LYS	HB2	1.514	0.025	2
1	A	25	LYS	HG2	1.135	0.025	2
1	A	25	LYS	HD2	1.449	0.025	2
1	A	25	LYS	HE2	2.703	0.025	2
1	A	28	TRP	HB2	3.353	0.025	2
1	A	29	ASN	HB2	2.943	0.025	2
1	A	30	ASP	HB2	2.682	0.025	2
1	A	32	LYS	HG2	1.054	0.025	2
1	A	32	LYS	HD2	1.441	0.025	2
1	A	32	LYS	HE2	2.967	0.025	2
1	A	33	GLU	HB2	1.705	0.025	2
1	A	33	GLU	HG2	1.925	0.025	2
1	A	34	GLN	HB2	1.454	0.025	2
1	A	34	GLN	HG2	1.856	0.025	2
1	A	37	ASN	HB2	2.825	0.025	2
1	A	39	LEU	HB2	1.635	0.025	2
1	A	41	SER	HB2	3.306	0.025	2
1	A	42	HIS	HB2	3.342	0.025	2
1	A	43	ASP	HB2	2.629	0.025	2
1	A	44	SER	HB2	3.521	0.025	2
1	A	45	ILE	HG12	1.562	0.025	2
1	A	46	ASP	HB2	2.336	0.025	2
1	A	47	PHE	HB2	2.927	0.025	2
1	A	48	SER	HB2	3.826	0.025	2
1	A	49	SER	HB2	3.507	0.025	2
1	A	50	ASP	HB2	2.722	0.025	2
1	A	53	SER	HB2	3.806	0.025	2
1	A	55	TYR	HB2	3.288	0.025	2
1	A	56	LYS	HB2	1.842	0.025	2
1	A	56	LYS	HG2	1.426	0.025	2
1	A	56	LYS	HD2	1.734	0.025	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	56	LYS	HE2	3.02	0.025	2
1	A	57	MET	HB2	1.915	0.025	2
1	A	57	MET	HG2	2.379	0.025	2
1	A	58	ASP	HB2	2.653	0.025	2
1	A	64	GLN	HB2	2.358	0.025	2
1	A	64	GLN	HG2	2.365	0.025	2
1	A	65	SER	HB2	3.676	0.025	2
1	A	67	GLU	HB2	2.002	0.025	2
1	A	67	GLU	HG2	2.175	0.025	2
1	A	68	LYS	HB2	1.738	0.025	2
1	A	68	LYS	HG2	1.273	0.025	2
1	A	68	LYS	HD2	1.543	0.025	2
1	A	68	LYS	HE2	2.836	0.025	2
1	A	69	PHE	HB2	2.829	0.025	2
1	A	72	HIS	HB2	3.005	0.025	2
1	A	73	PHE	HB2	3.072	0.025	2
1	A	74	LYS	HB2	1.845	0.025	2
1	A	74	LYS	HG2	1.334	0.025	2
1	A	74	LYS	HD2	1.657	0.025	2
1	A	74	LYS	HE2	2.949	0.025	2
1	A	76	LEU	HB2	1.35	0.025	2
1	A	79	LYS	HB2	1.784	0.025	2
1	A	79	LYS	HG2	1.333	0.025	2
1	A	79	LYS	HD2	1.584	0.025	2
1	A	79	LYS	HE2	2.903	0.025	2
1	A	80	ASP	HB2	2.554	0.025	2
1	A	82	ARG	HB2	1.553	0.025	2
1	A	82	ARG	HG2	1.422	0.025	2
1	A	82	ARG	HD2	3.061	0.025	2
1	A	84	PHE	HB2	2.625	0.025	2
1	A	85	ILE	HG12	0.613	0.025	2
1	A	86	GLU	HB2	1.906	0.025	2
1	A	86	GLU	HG2	2.152	0.025	2
1	A	87	LEU	HB2	1.622	0.025	2
1	A	88	ASP	HB2	3.299	0.025	2
1	A	89	GLU	HB2	1.993	0.025	2
1	A	89	GLU	HG2	2.182	0.025	2
1	A	90	ASN	HB2	2.78	0.025	2
1	A	92	ASP	HB2	2.513	0.025	2
1	A	94	GLN	HB2	1.58	0.025	2
1	A	94	GLN	HG2	2.329	0.025	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	96	PHE	HB2	2.889	0.025	2
1	A	97	ILE	HG12	0.578	0.025	2
1	A	98	PRO	HB2	2.158	0.025	2
1	A	98	PRO	HG2	1.848	0.025	2
1	A	98	PRO	HD2	4.107	0.025	2
1	A	99	GLN	HB2	1.968	0.025	2
1	A	99	GLN	HG2	2.327	0.025	2
1	A	101	GLU	HB2	2.004	0.025	2
1	A	101	GLU	HG2	2.21	0.025	2
1	A	102	ILE	HG12	1.369	0.025	2
1	A	103	ASP	HB2	2.57	0.025	2

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	100	2.68 ± 0.13	Should be applied
$^{13}C_{\beta}$	92	2.58 ± 0.11	Should be applied
$^{13}C'$	0	—	None (insufficient data)
^{15}N	100	0.60 ± 0.38	None needed (imprecise)

7.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 87%, i.e. 1061 atoms were assigned a chemical shift out of a possible 1219. 0 out of 13 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	373/466 (80%)	191/191 (100%)	91/184 (49%)	91/91 (100%)
Sidechain	600/649 (92%)	407/418 (97%)	182/207 (88%)	11/24 (46%)
Aromatic	88/104 (85%)	44/52 (85%)	43/49 (88%)	1/3 (33%)
Overall	1061/1219 (87%)	642/661 (97%)	316/440 (72%)	103/118 (87%)

7.1.4 Statistically unusual chemical shifts (i)

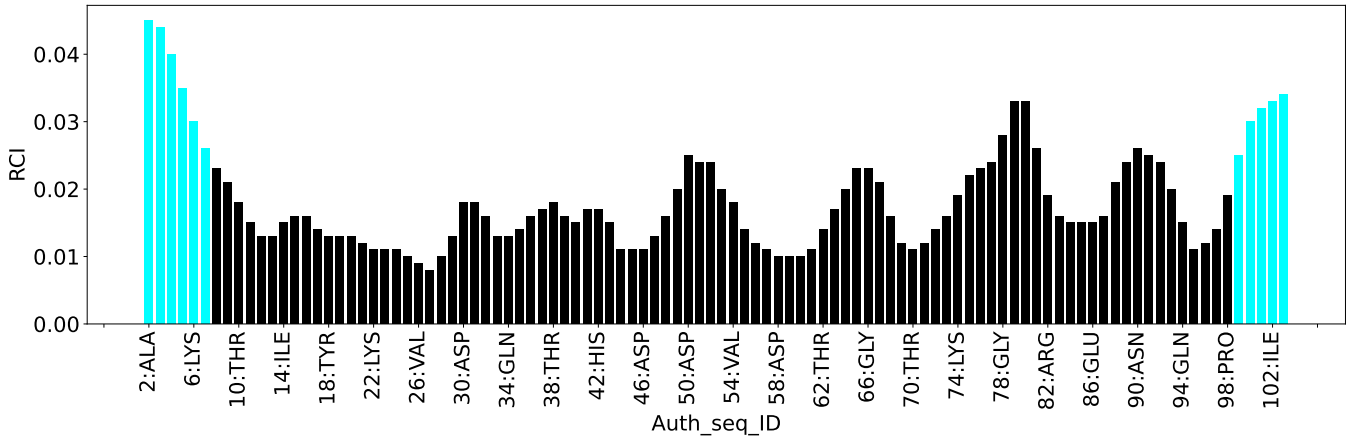
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	32	LYS	HB3	0.33	0.46 – 3.04	-5.5
1	A	11	ILE	HD11	-0.76	-0.72 – 2.09	-5.1
1	A	11	ILE	HD12	-0.76	-0.72 – 2.09	-5.1
1	A	11	ILE	HD13	-0.76	-0.72 – 2.09	-5.1
1	A	11	ILE	HB	0.33	0.35 – 3.22	-5.1

7.1.5 Random Coil Index (RCI) plots (i)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:



8 NMR restraints analysis (i)

8.1 Conformationally restricting restraints (i)

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	2136
Intra-residue (|i-j|=0)	408
Sequential (|i-j|=1)	614
Medium range (|i-j|>1 and |i-j|<5)	227
Long range (|i-j|≥5)	887
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	569
Number of restraints per residue	20.7
Number of long range restraints per residue[1]	8.6

[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 (i)

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 (i)

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)	30.7	0.2
0.2-0.5 (Medium)	59.0	0.5
>0.5 (Large)	104.8	3.14

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)

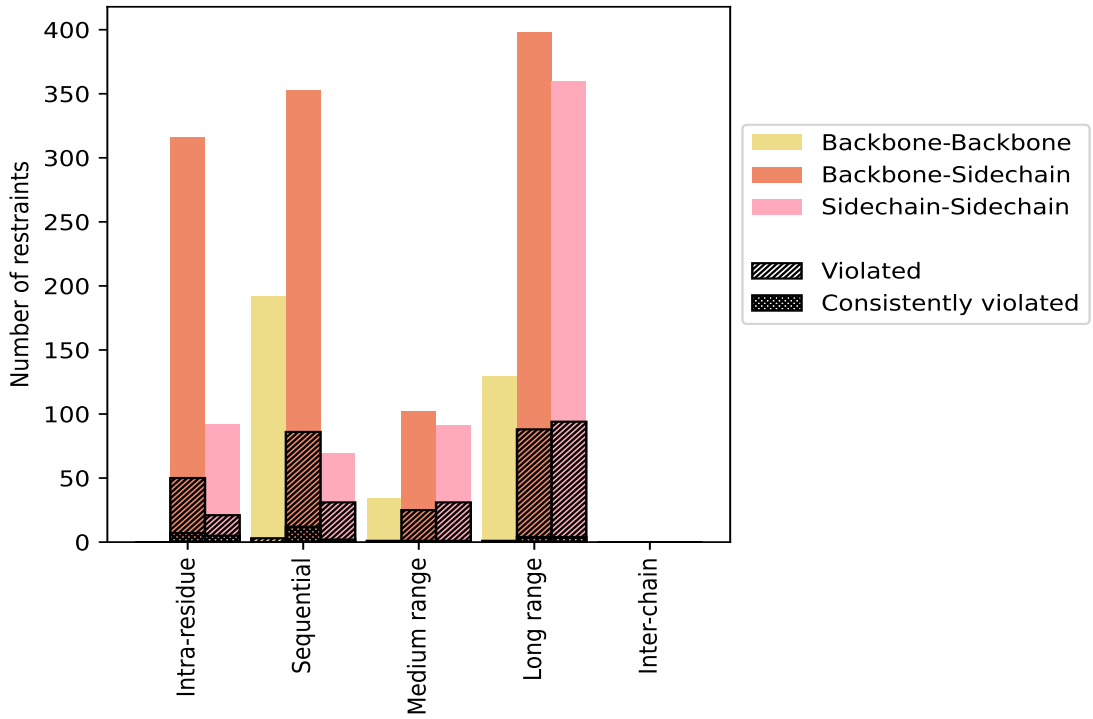
9.1 Summary of distance violations (i)

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)	408	19.1	71	17.4	3.3	12	2.9	0.6
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	316	14.8	50	15.8	2.3	7	2.2	0.3
Sidechain-Sidechain	92	4.3	21	22.8	1.0	5	5.4	0.2
Sequential (|i-j|=1)	614	28.7	120	19.5	5.6	14	2.3	0.7
Backbone-Backbone	192	9.0	3	1.6	0.1	0	0.0	0.0
Backbone-Sidechain	353	16.5	86	24.4	4.0	12	3.4	0.6
Sidechain-Sidechain	69	3.2	31	44.9	1.5	2	2.9	0.1
Medium range (|i-j|>1 & |i-j|<5)	227	10.6	57	25.1	2.7	2	0.9	0.1
Backbone-Backbone	34	1.6	1	2.9	0.0	0	0.0	0.0
Backbone-Sidechain	102	4.8	25	24.5	1.2	1	1.0	0.0
Sidechain-Sidechain	91	4.3	31	34.1	1.5	1	1.1	0.0
Long range (|i-j|≥5)	887	41.5	183	20.6	8.6	8	0.9	0.4
Backbone-Backbone	129	6.0	1	0.8	0.0	0	0.0	0.0
Backbone-Sidechain	398	18.6	88	22.1	4.1	4	1.0	0.2
Sidechain-Sidechain	360	16.9	94	26.1	4.4	4	1.1	0.2
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	2136	100.0	431	20.2	20.2	36	1.7	1.7
Backbone-Backbone	355	16.6	5	1.4	0.2	0	0.0	0.0
Backbone-Sidechain	1169	54.7	249	21.3	11.7	24	2.1	1.1
Sidechain-Sidechain	612	28.7	177	28.9	8.3	12	2.0	0.6

[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 (i)



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 (i)

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	44	67	25	67	0	203	0.63	2.24	0.46	0.49
2	39	63	22	70	0	194	0.67	2.59	0.49	0.53
3	37	54	15	60	0	166	0.63	1.98	0.46	0.5
4	39	57	21	72	0	189	0.66	2.61	0.48	0.53
5	38	56	25	83	0	202	0.67	2.12	0.47	0.56
6	42	63	20	59	0	184	0.64	2.15	0.44	0.53
7	40	72	18	79	0	209	0.71	2.45	0.52	0.56
8	39	54	21	71	0	185	0.62	2.24	0.44	0.54
9	36	61	25	74	0	196	0.67	2.34	0.51	0.55
10	42	65	27	68	0	202	0.69	2.41	0.46	0.57
11	39	61	17	70	0	187	0.67	2.64	0.48	0.54

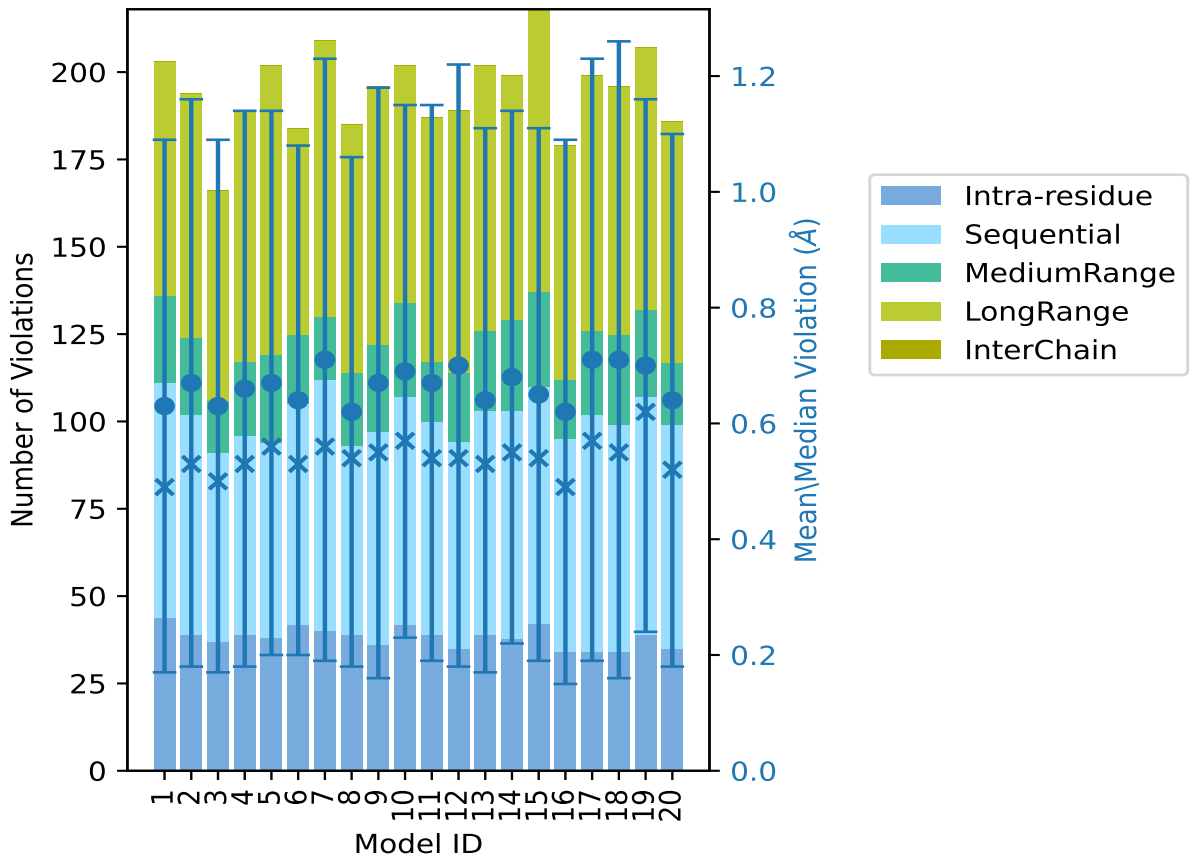
Continued on next page...

Continued from previous page...

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				
12	35	59	20	75	0	189	0.7	2.96	0.52	0.54
13	39	64	23	76	0	202	0.64	2.5	0.47	0.53
14	38	65	26	70	0	199	0.68	2.38	0.46	0.55
15	42	68	27	81	0	218	0.65	2.36	0.46	0.54
16	34	61	17	67	0	179	0.62	2.72	0.47	0.49
17	34	68	24	73	0	199	0.71	3.14	0.52	0.57
18	34	65	26	71	0	196	0.71	3.02	0.55	0.55
19	39	68	25	75	0	207	0.7	2.56	0.46	0.62
20	35	64	18	69	0	186	0.64	2.3	0.46	0.52

[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

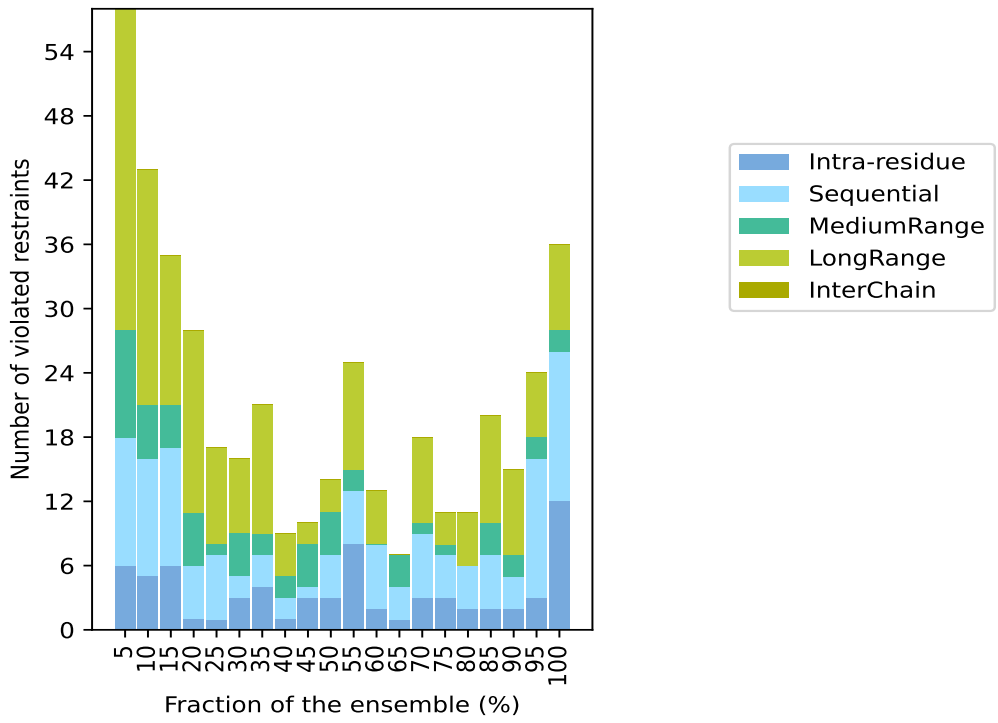
9.3 Distance violation statistics for the ensemble (i)

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, 1705(IR:337, SQ:494, MR:170, LR:704, 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]	%
6	12	10	30	0	58	1	5.0
5	11	5	22	0	43	2	10.0
6	11	4	14	0	35	3	15.0
1	5	5	17	0	28	4	20.0
1	6	1	9	0	17	5	25.0
3	2	4	7	0	16	6	30.0
4	3	2	12	0	21	7	35.0
1	2	2	4	0	9	8	40.0
3	1	4	2	0	10	9	45.0
3	4	4	3	0	14	10	50.0
8	5	2	10	0	25	11	55.0
2	6	0	5	0	13	12	60.0
1	3	3	0	0	7	13	65.0
3	6	1	8	0	18	14	70.0
3	4	1	3	0	11	15	75.0
2	4	0	5	0	11	16	80.0
2	5	3	10	0	20	17	85.0
2	3	2	8	0	15	18	90.0
3	13	2	6	0	24	19	95.0
12	14	2	8	0	36	20	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

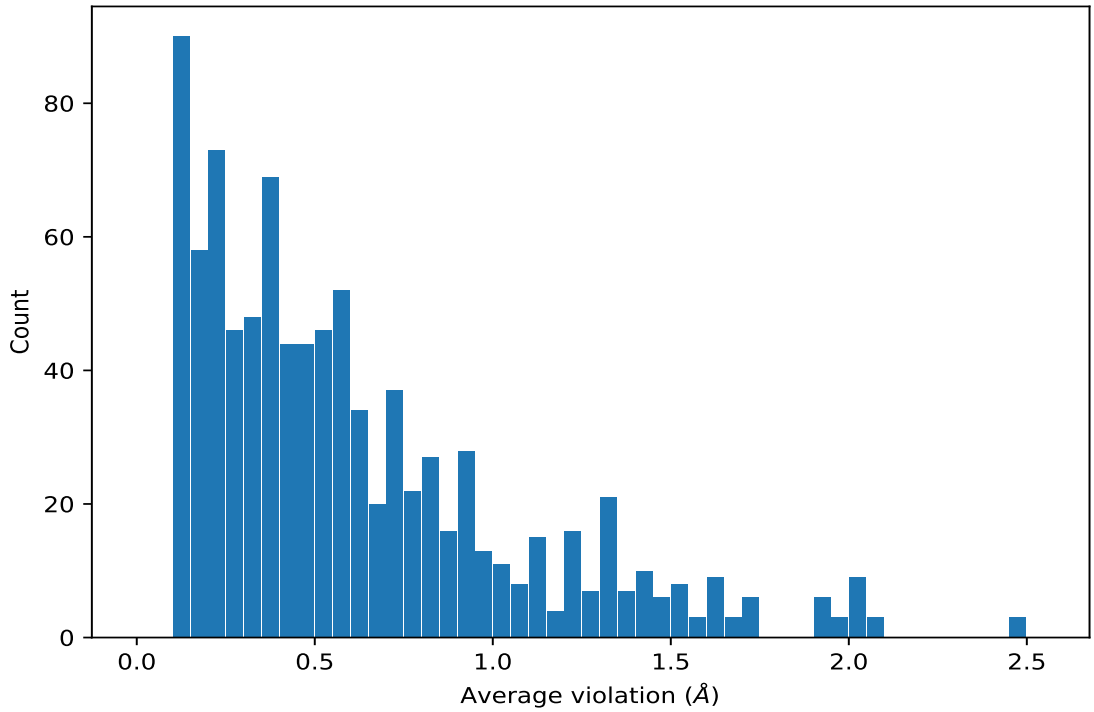
9.3.1 Bar graph : Distance violation statistics for the ensemble (i)



9.4 Most violated distance restraints in the ensemble (i)

9.4.1 Histogram : Distribution of mean distance violations (i)

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 (i)

The following table provides the mean and the standard deviation of the violations for the 10 worst performing restraints, sorted by number of violated models and the mean 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	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,1373)	1:A:22:LYS:HG3	1:A:62:THR:HB	20	1.43	0.12	1.45
(1,998)	1:A:98:PRO:HA	1:A:99:GLN:HG3	20	1.32	0.26	1.4
(1,1982)	1:A:69:PHE:HD1	1:A:95:VAL:HG11	20	1.21	0.49	1.32
(1,1982)	1:A:69:PHE:HD1	1:A:95:VAL:HG12	20	1.21	0.49	1.32
(1,1982)	1:A:69:PHE:HD1	1:A:95:VAL:HG13	20	1.21	0.49	1.32
(1,1982)	1:A:69:PHE:HD2	1:A:95:VAL:HG11	20	1.21	0.49	1.32
(1,1982)	1:A:69:PHE:HD2	1:A:95:VAL:HG12	20	1.21	0.49	1.32
(1,1982)	1:A:69:PHE:HD2	1:A:95:VAL:HG13	20	1.21	0.49	1.32
(1,2097)	1:A:28:TRP:HD1	1:A:29:ASN:HB3	20	1.17	0.3	1.23
(1,909)	1:A:23:ASN:HB3	1:A:24:ILE:HA	20	1.14	0.05	1.15
(1,644)	1:A:75:GLY:H	1:A:76:LEU:HB3	20	1.13	0.19	1.18
(1,2054)	1:A:13:LEU:HD11	1:A:47:PHE:HE1	20	1.07	0.2	1.0
(1,2054)	1:A:13:LEU:HD11	1:A:47:PHE:HE2	20	1.07	0.2	1.0
(1,2054)	1:A:13:LEU:HD12	1:A:47:PHE:HE1	20	1.07	0.2	1.0
(1,2054)	1:A:13:LEU:HD12	1:A:47:PHE:HE2	20	1.07	0.2	1.0
(1,2054)	1:A:13:LEU:HD13	1:A:47:PHE:HE1	20	1.07	0.2	1.0

Continued on next page...

Continued from previous page...

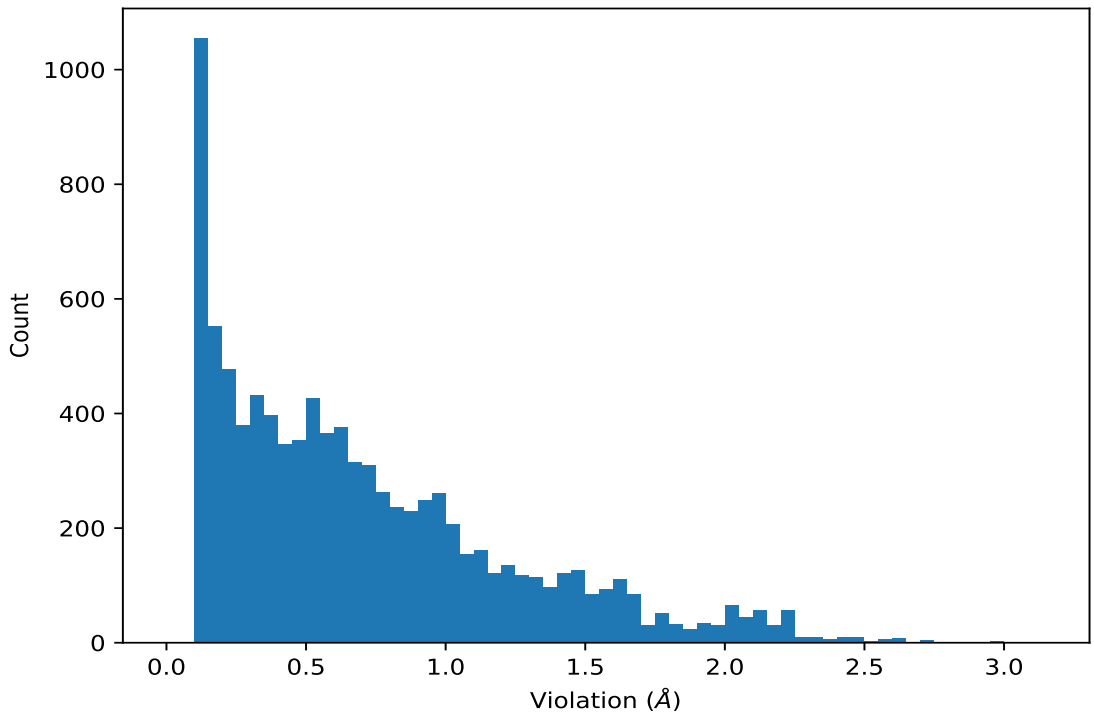
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,2054)	1:A:13:LEU:HD13	1:A:47:PHE:HE2	20	1.07	0.2	1.0
(1,1078)	1:A:98:PRO:HA	1:A:99:GLN:HB3	20	1.01	0.09	1.02
(1,913)	1:A:22:LYS:HB3	1:A:23:ASN:HA	20	1.0	0.06	0.99
(1,1097)	1:A:16:GLN:HB3	1:A:85:ILE:HG21	20	0.94	0.17	0.99

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

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)
(1,1313)	1:A:25:LYS:HG3	1:A:33:GLU:HG3	17	3.14

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,1313)	1:A:25:LYS:HG3	1:A:33:GLU:HG3	18	3.02
(1,507)	1:A:36:VAL:HG11	1:A:37:ASN:HD22	12	2.96
(1,507)	1:A:36:VAL:HG12	1:A:37:ASN:HD22	12	2.96
(1,507)	1:A:36:VAL:HG13	1:A:37:ASN:HD22	12	2.96
(1,1313)	1:A:25:LYS:HG3	1:A:33:GLU:HG3	16	2.72
(1,1498)	1:A:54:VAL:HG21	1:A:56:LYS:HA	18	2.7
(1,1498)	1:A:54:VAL:HG22	1:A:56:LYS:HA	18	2.7
(1,1498)	1:A:54:VAL:HG23	1:A:56:LYS:HA	18	2.7
(1,1313)	1:A:25:LYS:HG3	1:A:33:GLU:HG3	11	2.64
(1,440)	1:A:9:THR:H	1:A:54:VAL:HG11	18	2.63
(1,440)	1:A:9:THR:H	1:A:54:VAL:HG12	18	2.63
(1,440)	1:A:9:THR:H	1:A:54:VAL:HG13	18	2.63
(1,1417)	1:A:67:GLU:HG3	1:A:93:VAL:HG11	4	2.61
(1,1417)	1:A:67:GLU:HG3	1:A:93:VAL:HG12	4	2.61
(1,1417)	1:A:67:GLU:HG3	1:A:93:VAL:HG13	4	2.61
(1,1498)	1:A:54:VAL:HG21	1:A:56:LYS:HA	2	2.59
(1,1498)	1:A:54:VAL:HG22	1:A:56:LYS:HA	2	2.59
(1,1498)	1:A:54:VAL:HG23	1:A:56:LYS:HA	2	2.59
(1,1498)	1:A:54:VAL:HG21	1:A:56:LYS:HA	19	2.56

10 Dihedral-angle violation analysis (i)

No dihedral-angle restraints found