wwPDB NMR Structure Validation Summary Report

Jun 4, 2023 – 09:01 AM EDT

PDB ID	:	2LCV
BMRB ID	:	17634
Title	:	Structure of the Cytidine Repressor DNA-Binding Domain; an alternate calculation
Authors	:	Moody, C.L.; Tretyachenko-Ladokhina, V.; Senear, D.F.; Cocco, M.J.
Deposited on	:	2011-05-10

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:

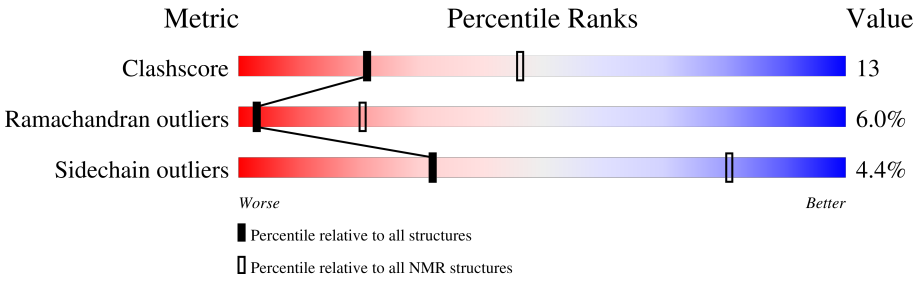
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal 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 70%.

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	67	48% 15% • • 30%

2 Ensemble composition and analysis

This entry contains 11 models. Model 11 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *minimized 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:10-A:53 (44)	0.66	11

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 2 single-model clusters were found.

Cluster number	Models
1	1, 2, 5, 11
2	3, 4, 9
3	6, 10
Single-model clusters	7; 8

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 727 atoms, of which 376 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called HTH-type transcriptional repressor CytR.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	47	727	216	376	67	66	2	0

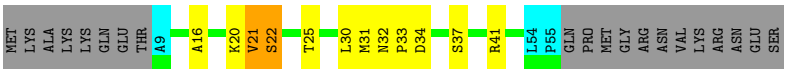
4 Residue-property plots (i)

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: HTH-type transcriptional repressor CytR

Chain A:

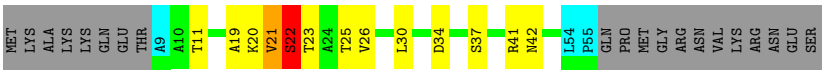


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 11. Colouring as in section 4.1 above.

- Molecule 1: HTH-type transcriptional repressor CytR

Chain A:



5 Refinement protocol and experimental data overview

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

Of the 100 calculated structures, 11 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
X-PLOR NIH	refinement	
VNMR	structure solution	
NMRPipe	structure solution	
ANALYSIS - CCPN	structure solution	

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	607
Number of shifts mapped to atoms	458
Number of unparsed shifts	0
Number of shifts with mapping errors	149
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	70%

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	331	353	353	9±4
All	All	3641	3883	3883	99

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 13.

5 of 67 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:25:THR:HG23	1:A:26:VAL:N	0.67	2.03	10	1
1:A:12:MET:SD	1:A:53:TYR:CD2	0.60	2.95	7	1
1:A:17:LEU:HD23	1:A:17:LEU:C	0.60	2.17	9	1
1:A:51:VAL:HG13	1:A:53:TYR:H	0.59	1.58	8	1
1:A:17:LEU:O	1:A:17:LEU:HD13	0.58	1.97	3	1

6.3 Torsion angles

6.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the backbone conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	44/67 (66%)	39±2 (88±4%)	3±1 (6±3%)	3±1 (6±2%)	3 20
All	All	484/737 (66%)	426 (88%)	29 (6%)	29 (6%)	3 20

5 of 7 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	22	SER	11
1	A	21	VAL	10
1	A	33	PRO	3
1	A	34	ASP	2
1	A	35	LYS	1

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	35/55 (64%)	33±2 (96±5%)	2±2 (4±5%)	32 81
All	All	385/605 (64%)	368 (96%)	17 (4%)	32 81

5 of 9 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	22	SER	5
1	A	32	ASN	3
1	A	17	LEU	2
1	A	42	ASN	2
1	A	11	THR	1

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

The completeness of assignment taking into account all chemical shift lists is 70% for the well-defined parts and 72% 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	607
Number of shifts mapped to atoms	458
Number of unparsed shifts	0
Number of shifts with mapping errors	149
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

The following errors were found when reading this chemical shift list.

- Chemical shift has been reported more than once. All 4 occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	17	LEU	HD12	1.0083	.	.
1	A	17	LEU	HD13	1.0083	.	.
1	A	54	LEU	HD12	0.9388	.	.
1	A	54	LEU	HD13	0.9388	.	.

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 149) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	2	LYS	CA	56.4808	.	1
1	A	2	LYS	CB	33.2612	.	1
1	A	3	ALA	CA	52.346	.	1
1	A	3	ALA	CB	19.6097	.	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	3	ALA	H	8.3759	.	1
1	A	3	ALA	HA	4.378	.	1
1	A	3	ALA	HB1	1.4162	.	.
1	A	3	ALA	HB2	1.4162	.	.
1	A	3	ALA	HB3	1.4162	.	.
1	A	3	ALA	N	126.1333	.	1
1	A	4	LYS	CA	56.2606	.	1
1	A	4	LYS	CB	33.3463	.	1
1	A	4	LYS	H	8.2816	.	1
1	A	4	LYS	N	121.32	.	1
1	A	5	LYS	CA	56.4255	.	1
1	A	5	LYS	CB	33.2387	.	1
1	A	5	LYS	H	8.3559	.	1
1	A	5	LYS	N	123.2627	.	1
1	A	6	GLN	CA	56.0064	.	1
1	A	6	GLN	CB	29.7147	.	1
1	A	6	GLN	CG	33.7464	.	1
1	A	6	GLN	H	8.4244	.	1
1	A	6	GLN	HA	4.3464	.	1
1	A	6	GLN	HB2	2.0634	.	.
1	A	6	GLN	HB3	2.1344	.	.
1	A	6	GLN	HG2	2.423	.	.
1	A	6	GLN	HG3	2.423	.	.
1	A	6	GLN	N	122.0811	.	1
1	A	7	GLU	CA	56.8527	.	1
1	A	7	GLU	CB	30.4222	.	1
1	A	7	GLU	CG	36.3866	.	1
1	A	7	GLU	H	8.5469	.	1
1	A	7	GLU	HA	4.4033	.	1
1	A	7	GLU	HB2	2.0076	.	.
1	A	7	GLU	HB3	2.1444	.	.
1	A	7	GLU	HG2	2.327	.	.
1	A	7	GLU	HG3	2.327	.	.
1	A	7	GLU	N	122.5524	.	1
1	A	8	THR	CA	61.3682	.	1
1	A	8	THR	CB	70.2225	.	1
1	A	8	THR	CG2	21.7074	.	1
1	A	8	THR	H	8.09	.	1
1	A	8	THR	HA	4.4082	.	1
1	A	8	THR	HB	4.2943	.	1
1	A	8	THR	HG21	1.2598	.	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	8	THR	HG22	1.2598	.	.
1	A	8	THR	HG23	1.2598	.	.
1	A	8	THR	N	114.864	.	1
1	A	56	GLN	CA	53.6697	.	1
1	A	56	GLN	CB	29.1753	.	1
1	A	56	GLN	CG	33.2954	.	1
1	A	56	GLN	H	8.5063	.	1
1	A	56	GLN	HA	4.6181	.	1
1	A	56	GLN	HB2	1.9717	.	.
1	A	56	GLN	HB3	2.1396	.	.
1	A	56	GLN	HE21	6.8594	.	.
1	A	56	GLN	HG2	2.4437	.	.
1	A	56	GLN	HG3	2.4437	.	.
1	A	56	GLN	N	121.523	.	1
1	A	57	PRO	CA	63.4954	.	1
1	A	57	PRO	CB	31.962	.	1
1	A	57	PRO	CD	50.5656	.	1
1	A	57	PRO	CG	28.0885	.	1
1	A	57	PRO	HA	4.468	.	1
1	A	57	PRO	HB2	1.9608	.	.
1	A	57	PRO	HB3	2.3392	.	.
1	A	57	PRO	HD2	3.6999	.	.
1	A	57	PRO	HD3	3.7281	.	.
1	A	57	PRO	HG2	2.0595	.	.
1	A	57	PRO	HG3	2.0829	.	.
1	A	58	MET	CA	55.7476	.	1
1	A	58	MET	CB	33.2079	.	1
1	A	58	MET	CG	32.0467	.	1
1	A	58	MET	H	8.4291	.	1
1	A	58	MET	HA	4.5254	.	1
1	A	58	MET	HB2	2.0995	.	.
1	A	58	MET	HB3	2.1668	.	.
1	A	58	MET	HE1	0.9553	.	.
1	A	58	MET	HE2	0.9553	.	.
1	A	58	MET	HE3	0.9553	.	.
1	A	58	MET	HG2	2.6377	.	.
1	A	58	MET	HG3	2.6966	.	.
1	A	58	MET	N	120.2457	.	1
1	A	59	GLY	CA	45.5484	.	1
1	A	59	GLY	H	8.3868	.	1
1	A	59	GLY	HA2	4.0029	.	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	59	GLY	HA3	4.0029	.	.
1	A	59	GLY	N	109.8508	.	1
1	A	60	ARG	CA	56.3717	.	1
1	A	60	ARG	CB	30.8592	.	1
1	A	60	ARG	H	8.1936	.	1
1	A	60	ARG	HA	4.3597	.	1
1	A	60	ARG	HB2	1.7982	.	.
1	A	60	ARG	HB3	1.8819	.	.
1	A	60	ARG	HD2	3.2585	.	.
1	A	60	ARG	HD3	3.2585	.	.
1	A	60	ARG	HG2	1.6575	.	.
1	A	60	ARG	HG3	1.6865	.	.
1	A	60	ARG	N	120.2487	.	1
1	A	61	ASN	CA	53.4535	.	1
1	A	61	ASN	CB	38.973	.	1
1	A	61	ASN	H	8.5094	.	1
1	A	61	ASN	HA	4.7591	.	1
1	A	61	ASN	HB2	2.8013	.	.
1	A	61	ASN	HB3	2.8851	.	.
1	A	61	ASN	N	119.5324	.	1
1	A	62	VAL	CA	62.3503	.	1
1	A	62	VAL	CB	32.8521	.	1
1	A	62	VAL	CG1	21.5386	.	.
1	A	62	VAL	CG2	21.5386	.	.
1	A	62	VAL	H	7.9053	.	1
1	A	62	VAL	HA	4.1399	.	1
1	A	62	VAL	HB	2.1196	.	1
1	A	62	VAL	HG11	0.932	.	.
1	A	62	VAL	HG12	0.932	.	.
1	A	62	VAL	HG13	0.932	.	.
1	A	62	VAL	N	120.0893	.	1
1	A	63	LYS	CA	56.3005	.	1
1	A	63	LYS	CB	32.9944	.	1
1	A	63	LYS	H	8.3311	.	1
1	A	63	LYS	HA	4.3714	.	1
1	A	63	LYS	HB2	1.7884	.	.
1	A	63	LYS	HB3	1.8695	.	.
1	A	63	LYS	HD2	1.7508	.	.
1	A	63	LYS	HD3	1.7508	.	.
1	A	63	LYS	HG2	1.4693	.	.
1	A	63	LYS	HG3	1.5101	.	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	63	LYS	N	125.1308	.	1
1	A	64	ARG	CA	56.1167	.	1
1	A	64	ARG	CB	31.0832	.	1
1	A	64	ARG	H	8.3473	.	1
1	A	64	ARG	HA	4.3759	.	1
1	A	64	ARG	N	122.7721	.	1
1	A	65	ASN	CA	53.4787	.	1
1	A	65	ASN	CB	39.0579	.	1
1	A	65	ASN	H	8.4899	.	1
1	A	65	ASN	N	120.1232	.	1
1	A	66	GLU	CA	56.6788	.	1
1	A	66	GLU	CB	30.5799	.	1
1	A	66	GLU	H	8.3461	.	1
1	A	66	GLU	HA	4.4157	.	1
1	A	66	GLU	N	121.679	.	1
1	A	67	SER	CA	60.1037	.	1
1	A	67	SER	CB	65.0112	.	1
1	A	67	SER	H	7.9331	.	1
1	A	67	SER	HA	4.3055	.	1
1	A	67	SER	HB2	3.8635	.	.
1	A	67	SER	HB3	3.8842	.	.
1	A	67	SER	N	122.1271	.	1

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	66	-0.49 ± 0.37	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	64	0.02 ± 0.19	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	62	0.83 ± 0.91	None needed (imprecise)

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 70%, i.e. 418 atoms were assigned a chemical shift out of a possible 594. 0 out of 8 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	174/219 (79%)	87/88 (99%)	44/88 (50%)	43/43 (100%)
Sidechain	242/366 (66%)	174/238 (73%)	68/108 (63%)	0/20 (0%)
Aromatic	2/9 (22%)	2/4 (50%)	0/5 (0%)	0/0 (—%)
Overall	418/594 (70%)	263/330 (80%)	112/201 (56%)	43/63 (68%)

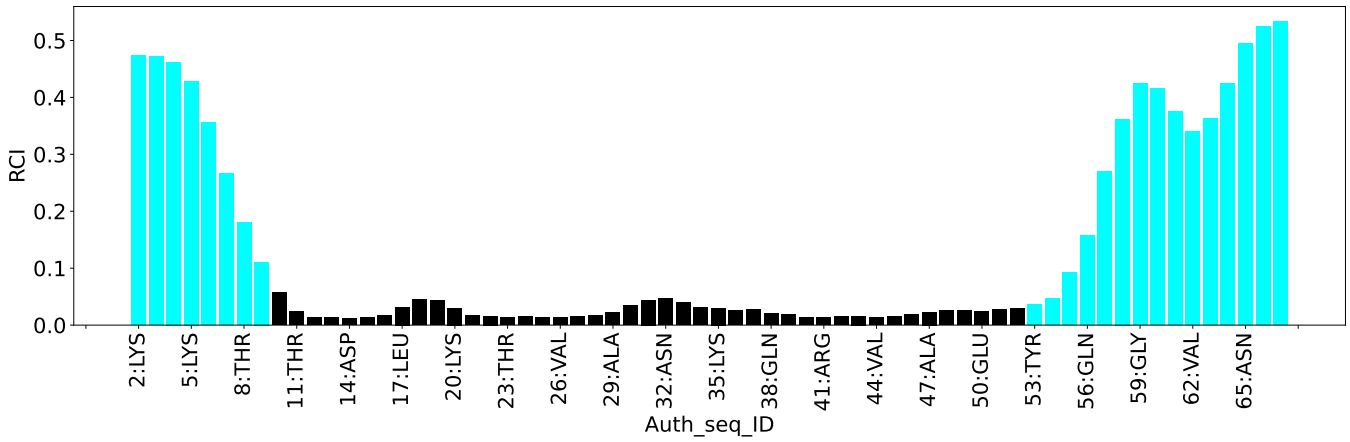
7.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

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

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	555
Intra-residue (|i-j|=0)	232
Sequential (|i-j|=1)	190
Medium range (|i-j|>1 and |i-j|<5)	94
Long range (|i-j|≥5)	39
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	0
Number of restraints per residue	8.3
Number of long range restraints per residue[1]	0.6

[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

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)	31.6	0.2
0.2-0.5 (Medium)	3.0	0.36
>0.5 (Large)	0.1	0.64

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)

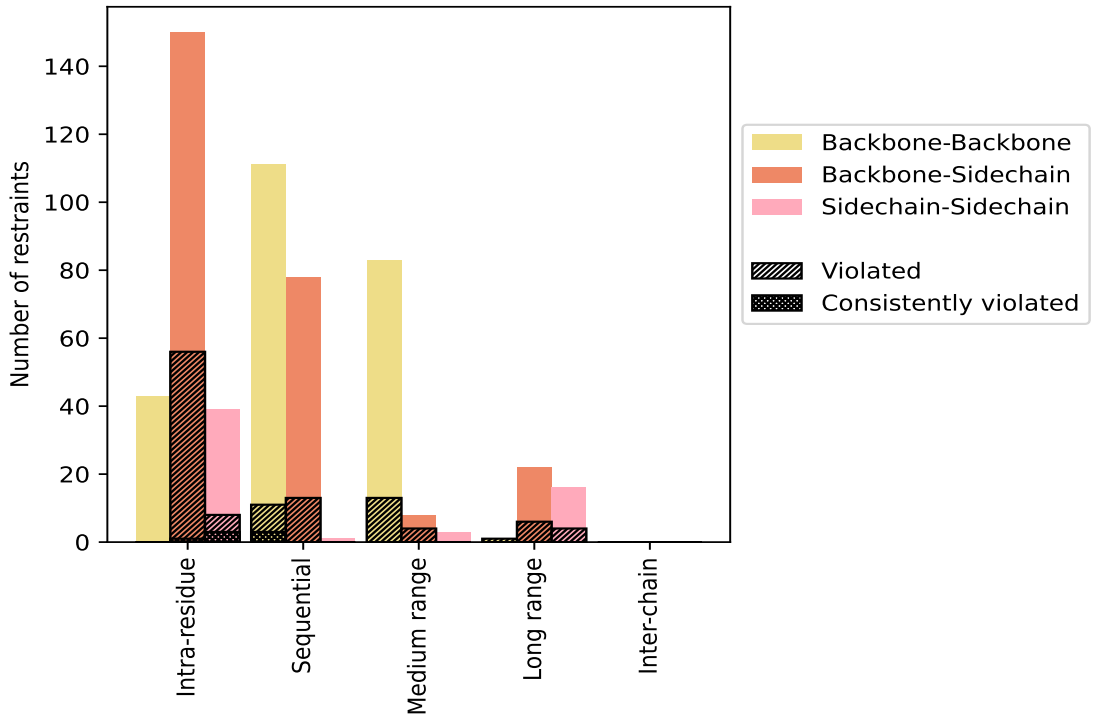
9.1 Summary of distance violations (i)

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)	232	41.8	64	27.6	11.5	4	1.7	0.7
Backbone-Backbone	43	7.7	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	150	27.0	56	37.3	10.1	1	0.7	0.2
Sidechain-Sidechain	39	7.0	8	20.5	1.4	3	7.7	0.5
Sequential (|i-j|=1)	190	34.2	24	12.6	4.3	3	1.6	0.5
Backbone-Backbone	111	20.0	11	9.9	2.0	3	2.7	0.5
Backbone-Sidechain	78	14.1	13	16.7	2.3	0	0.0	0.0
Sidechain-Sidechain	1	0.2	0	0.0	0.0	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	94	16.9	17	18.1	3.1	0	0.0	0.0
Backbone-Backbone	83	15.0	13	15.7	2.3	0	0.0	0.0
Backbone-Sidechain	8	1.4	4	50.0	0.7	0	0.0	0.0
Sidechain-Sidechain	3	0.5	0	0.0	0.0	0	0.0	0.0
Long range (|i-j|≥5)	39	7.0	11	28.2	2.0	0	0.0	0.0
Backbone-Backbone	1	0.2	1	100.0	0.2	0	0.0	0.0
Backbone-Sidechain	22	4.0	6	27.3	1.1	0	0.0	0.0
Sidechain-Sidechain	16	2.9	4	25.0	0.7	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	555	100.0	116	20.9	20.9	7	1.3	1.3
Backbone-Backbone	238	42.9	25	10.5	4.5	3	1.3	0.5
Backbone-Sidechain	258	46.5	79	30.6	14.2	1	0.4	0.2
Sidechain-Sidechain	59	10.6	12	20.3	2.2	3	5.1	0.5

[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

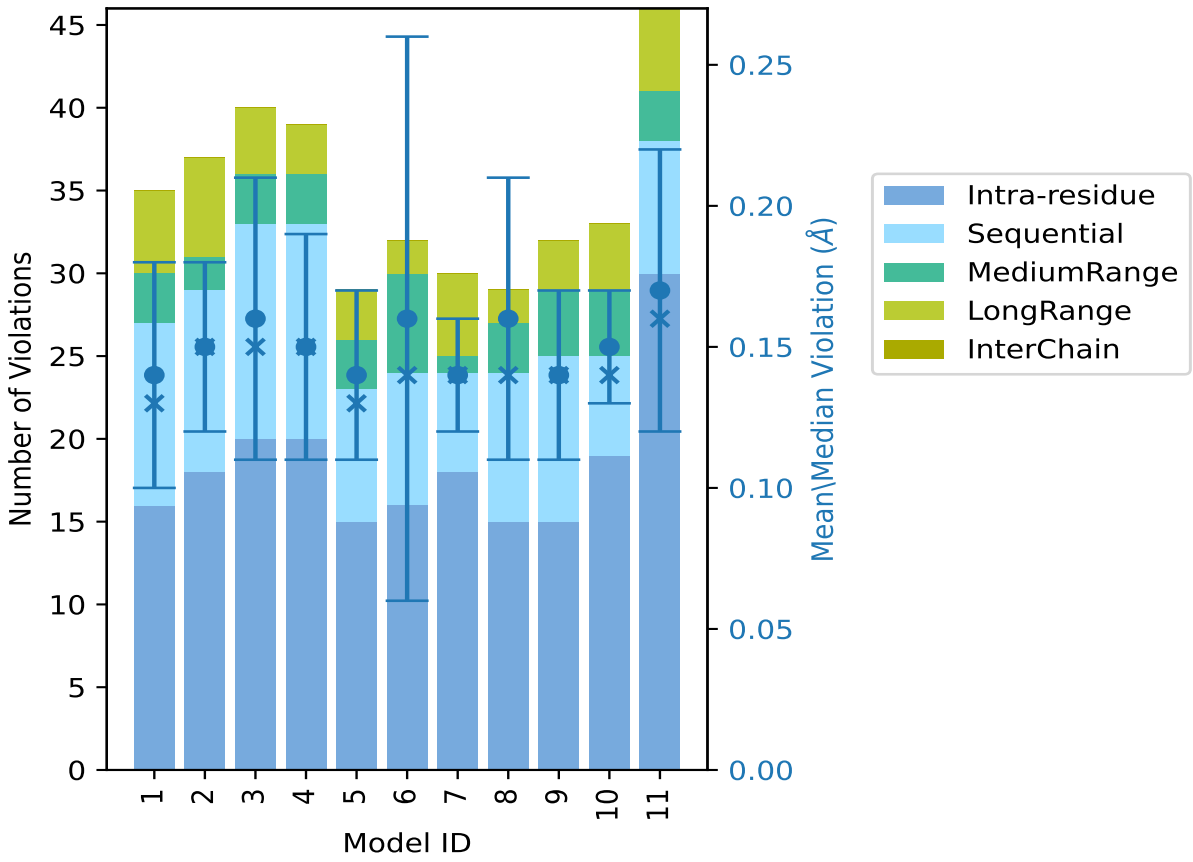
9.2 Distance violation statistics for each model

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	16	11	3	5	0	35	0.14	0.29	0.04	0.13
2	18	11	2	6	0	37	0.15	0.2	0.03	0.15
3	20	13	3	4	0	40	0.16	0.36	0.05	0.15
4	20	13	3	3	0	39	0.15	0.27	0.04	0.15
5	15	8	3	3	0	29	0.14	0.23	0.03	0.13
6	16	8	6	2	0	32	0.16	0.64	0.1	0.14
7	18	6	1	5	0	30	0.14	0.23	0.02	0.14
8	15	9	3	2	0	29	0.16	0.34	0.05	0.14
9	15	10	4	3	0	32	0.14	0.24	0.03	0.14
10	19	6	4	4	0	33	0.15	0.2	0.02	0.14
11	30	8	3	5	0	46	0.17	0.33	0.05	0.16

[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, 439(IR:168, SQ:166, MR:77, LR:28, 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]	%
25	8	11	3	0	47	1	9.1
9	1	2	0	0	12	2	18.2
10	4	2	3	0	19	3	27.3

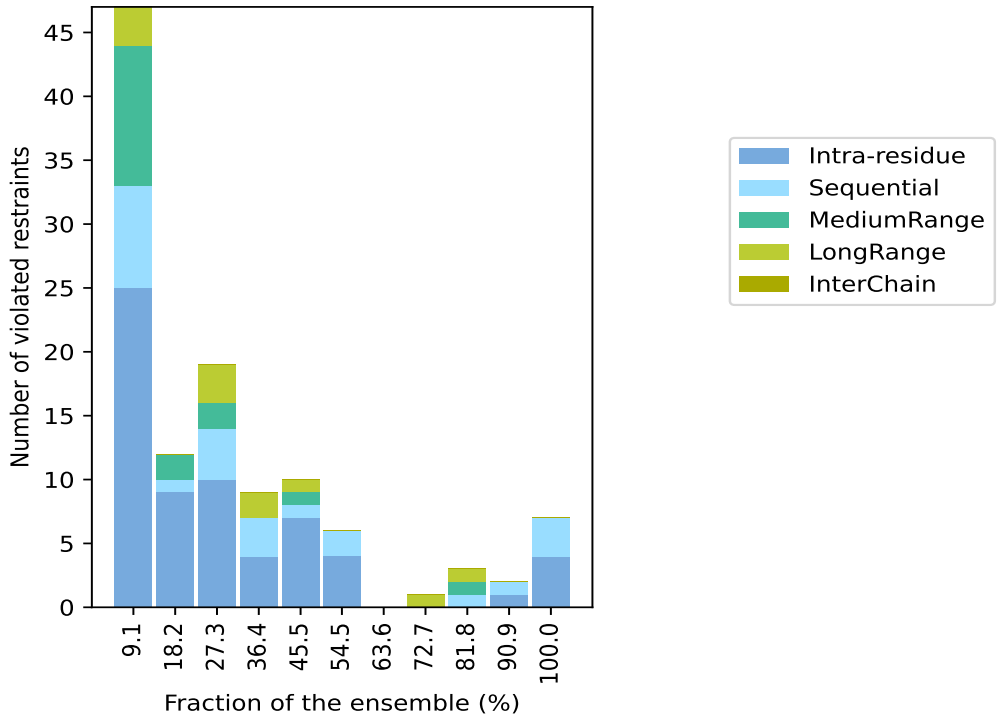
Continued on next page...

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]	%
4	3	0	2	0	9	4	36.4
7	1	1	1	0	10	5	45.5
4	2	0	0	0	6	6	54.5
0	0	0	0	0	0	7	63.6
0	0	0	1	0	1	8	72.7
0	1	1	1	0	3	9	81.8
1	1	0	0	0	2	10	90.9
4	3	0	0	0	7	11	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

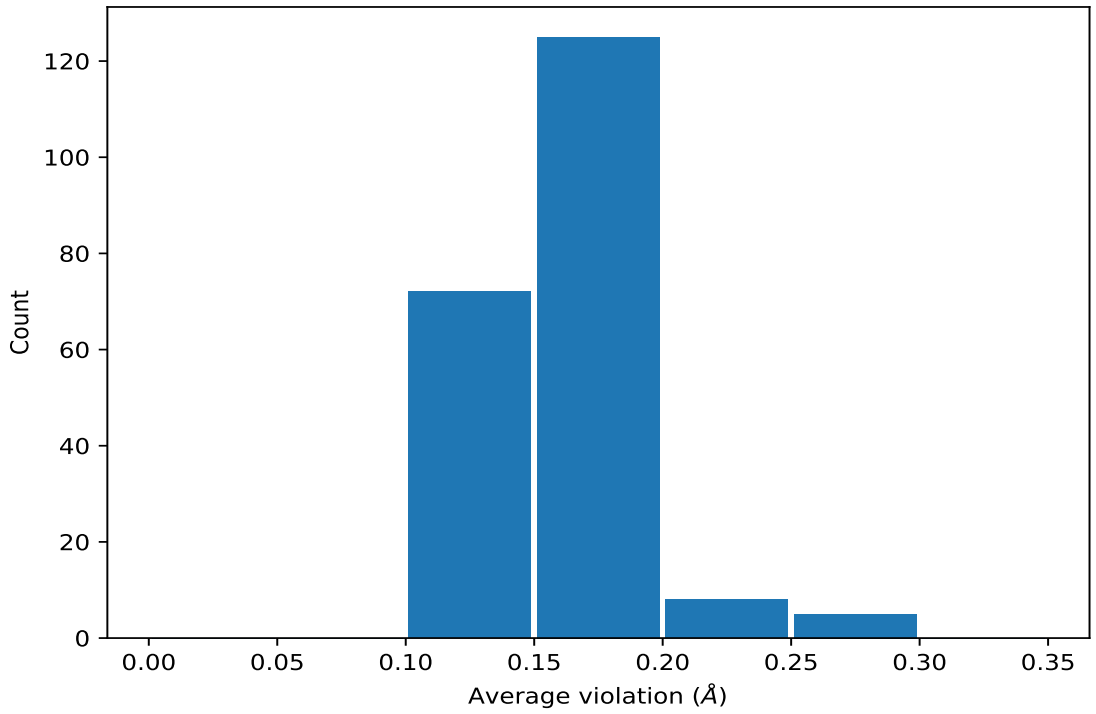


9.4 Most violated distance restraints in the ensemble

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 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)
(1,386)	1:A:20:LYS:HB2	1:A:20:LYS:HG2	11	0.19	0.0	0.19
(1,386)	1:A:20:LYS:HB2	1:A:20:LYS:HG3	11	0.19	0.0	0.19
(1,386)	1:A:20:LYS:HB3	1:A:20:LYS:HG2	11	0.19	0.0	0.19
(1,386)	1:A:20:LYS:HB3	1:A:20:LYS:HG3	11	0.19	0.0	0.19
(1,363)	1:A:17:LEU:HG	1:A:17:LEU:HD21	11	0.15	0.01	0.15
(1,363)	1:A:17:LEU:HG	1:A:17:LEU:HD22	11	0.15	0.01	0.15
(1,363)	1:A:17:LEU:HG	1:A:17:LEU:HD23	11	0.15	0.01	0.15
(1,191)	1:A:13:LYS:H	1:A:14:ASP:HA	11	0.15	0.02	0.15
(1,467)	1:A:39:ALA:HB1	1:A:39:ALA:HA	11	0.15	0.0	0.15
(1,467)	1:A:39:ALA:HB2	1:A:39:ALA:HA	11	0.15	0.0	0.15
(1,467)	1:A:39:ALA:HB3	1:A:39:ALA:HA	11	0.15	0.0	0.15
(1,553)	1:A:55:PRO:HB2	1:A:55:PRO:HG2	11	0.14	0.01	0.14
(1,553)	1:A:55:PRO:HB2	1:A:55:PRO:HG3	11	0.14	0.01	0.14
(1,553)	1:A:55:PRO:HB3	1:A:55:PRO:HG2	11	0.14	0.01	0.14
(1,553)	1:A:55:PRO:HB3	1:A:55:PRO:HG3	11	0.14	0.01	0.14

Continued on next page...

Continued from previous page...

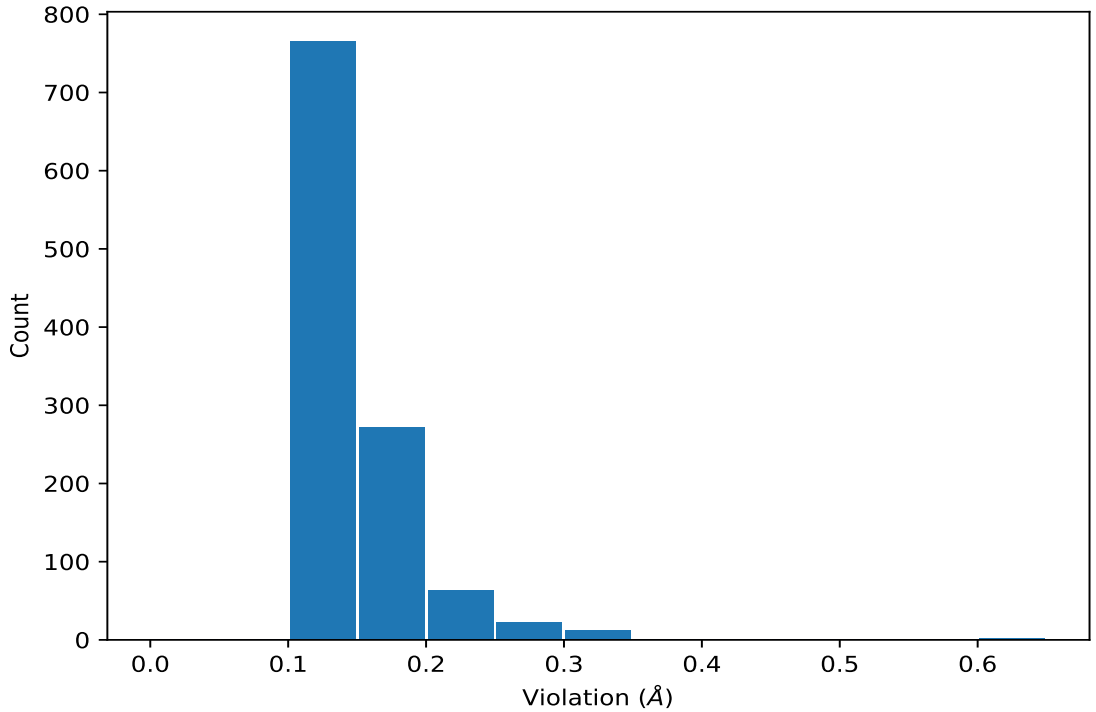
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,222)	1:A:22:SER:H	1:A:21:VAL:HA	11	0.14	0.01	0.14
(1,107)	1:A:37:SER:H	1:A:38:GLN:H	11	0.13	0.01	0.13
(1,518)	1:A:50:GLU:H	1:A:50:GLU:HG2	10	0.17	0.05	0.15
(1,518)	1:A:50:GLU:H	1:A:50:GLU:HG3	10	0.17	0.05	0.15
(1,195)	1:A:14:ASP:H	1:A:15:VAL:HG11	10	0.13	0.02	0.12
(1,195)	1:A:14:ASP:H	1:A:15:VAL:HG12	10	0.13	0.02	0.12
(1,195)	1:A:14:ASP:H	1:A:15:VAL:HG13	10	0.13	0.02	0.12
(1,9)	1:A:14:ASP:HA	1:A:51:VAL:HG11	9	0.16	0.07	0.13

[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)
(1,338)	1:A:13:LYS:HA	1:A:13:LYS:HE2	6	0.64
(1,338)	1:A:13:LYS:HA	1:A:13:LYS:HE3	6	0.64
(1,103)	1:A:35:LYS:HA	1:A:37:SER:H	3	0.36
(1,9)	1:A:14:ASP:HA	1:A:51:VAL:HG11	8	0.34
(1,9)	1:A:14:ASP:HA	1:A:51:VAL:HG12	8	0.34
(1,9)	1:A:14:ASP:HA	1:A:51:VAL:HG13	8	0.34
(1,9)	1:A:14:ASP:HA	1:A:51:VAL:HG21	8	0.34
(1,9)	1:A:14:ASP:HA	1:A:51:VAL:HG22	8	0.34
(1,9)	1:A:14:ASP:HA	1:A:51:VAL:HG23	8	0.34
(1,492)	1:A:45:GLU:H	1:A:45:GLU:HB2	11	0.33
(1,492)	1:A:45:GLU:H	1:A:45:GLU:HB3	11	0.33
(1,321)	1:A:53:TYR:H	1:A:54:LEU:HG	6	0.32
(1,358)	1:A:17:LEU:H	1:A:17:LEU:HD21	3	0.31
(1,358)	1:A:17:LEU:H	1:A:17:LEU:HD22	3	0.31
(1,358)	1:A:17:LEU:H	1:A:17:LEU:HD23	3	0.31
(1,97)	1:A:34:ASP:H	1:A:35:LYS:H	1	0.29
(1,247)	1:A:29:ALA:H	1:A:28:ARG:HD2	8	0.28
(1,247)	1:A:29:ALA:H	1:A:28:ARG:HD3	8	0.28
(1,220)	1:A:21:VAL:H	1:A:20:LYS:HD2	4	0.27
(1,220)	1:A:21:VAL:H	1:A:20:LYS:HD3	4	0.27
(1,98)	1:A:34:ASP:HA	1:A:35:LYS:H	3	0.26

10 Dihedral-angle violation analysis (i)

No dihedral-angle restraints found