wwPDB NMR Structure Validation Summary Report

Jan 31, 2024 – 12:23 am GMT

PDB ID	:	8COO
BMRB ID	:	34660
Title	:	Solution structure of Zipcode binding protein 1 (ZBP1) KH3(DD)KH4 domains in complex with N6-Methyladenosine containing RNA
Authors	:	Nicastro, G.; Abis, G.; Taylor, I.A.; Ramos, A.
Deposited on	:	2023-02-28

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:

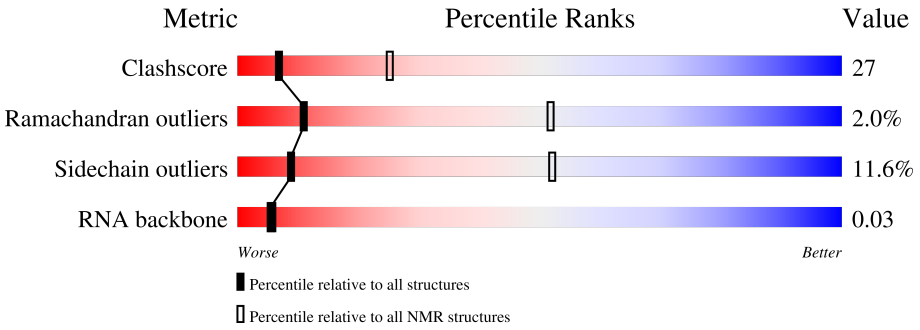
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	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)	:	2.36

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 62%.

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
RNA backbone	4643	676

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	191	37% 40% . 22%
2	B	7	14% 86%

2 Ensemble composition and analysis

This entry contains 20 models. Model 2 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:20-A:61, A:68-A:96, A:102-A:179 (149)	0.65	2

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 6 clusters and 2 single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3152 atoms, of which 1556 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Insulin-like growth factor 2 mRNA-binding protein 1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	191	2927	912	1477	261	273	4	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP O42254
A	2	ALA	-	expression tag	UNP O42254
A	3	MET	-	expression tag	UNP O42254
A	4	GLY	-	expression tag	UNP O42254
A	14	PHE	TYR	engineered mutation	UNP O42254
A	40	ASP	LYS	engineered mutation	UNP O42254
A	41	ASP	LYS	engineered mutation	UNP O42254

- Molecule 2 is a RNA chain called RNA_(5'-R(*(UP*CP*GP*GP*(6MZ)P*CP*U)-3').

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
2	B	7	225	67	79	25	48	6	0

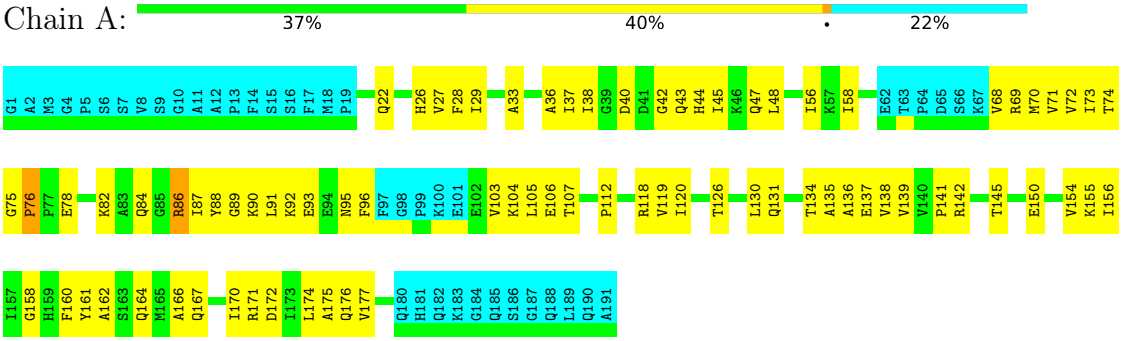
4 Residue-property plots (i)

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: Insulin-like growth factor 2 mRNA-binding protein 1

Chain A:



• Molecule 2: RNA_(5'-R(*(UP*CP*GP*GP*(6MZ)P*CP*U)-3')

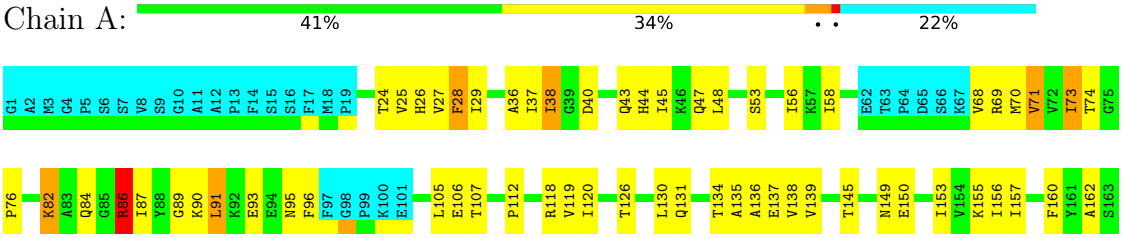
Chain B:

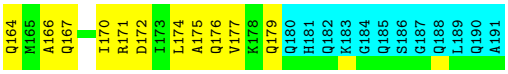
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 2. Colouring as in section 4.1 above.

• Molecule 1: Insulin-like growth factor 2 mRNA-binding protein 1

Chain A:





• Molecule 2: RNA_(5'-R(*(UP*CP*GP*GP*(6MZ)P*CP*U)-3')

Chain B:

5 Refinement protocol and experimental data overview

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

Of the 40 calculated structures, 20 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
ARIA	refinement	
ARIA	structure calculation	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	1472
Number of shifts mapped to atoms	1472
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	62%

6 Model quality

6.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 6MZ

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	0.90±0.01	0±0/1169 (0.0± 0.0%)	1.04±0.02	2±1/1581 (0.2± 0.1%)
2	B	0.30±0.02	0±0/135 (0.0± 0.0%)	0.90±0.04	0±0/206 (0.0± 0.0%)
All	All	0.85	0/26080 (0.0%)	1.03	49/35740 (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	0.5±0.5
All	All	0	9

There are no bond-length outliers.

5 of 10 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	113	ALA	N-CA-CB	-5.68	102.14	110.10	15	4
1	A	177	VAL	CA-CB-CG1	5.25	118.78	110.90	15	6
1	A	110	ARG	NE-CZ-NH1	5.20	122.90	120.30	14	6
1	A	86	ARG	NE-CZ-NH1	5.16	122.88	120.30	18	7
1	A	168	ARG	NE-CZ-NH1	5.15	122.88	120.30	10	4

There are no chirality outliers.

All unique planar outliers are listed below.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	38	ILE	Peptide	9

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	1149	1192	1188	48±7
2	B	146	79	79	24±3
All	All	25900	25420	25340	1404

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 27.

5 of 327 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:B:1:U:H6	2:B:1:U:HO5'	0.86	0.88	12	3
2:B:4:G:O2'	2:B:5:6MZ:O2P	0.74	2.06	19	20
2:B:1:U:O2'	2:B:2:C:H3'	0.74	1.82	4	7
1:A:153:ILE:HB	1:A:155:LYS:HZ1	0.73	1.44	3	3
1:A:27:VAL:HG21	1:A:87:ILE:HG21	0.72	1.60	14	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	149/191 (78%)	135±2 (91±1%)	11±2 (7±1%)	3±1 (2±1%)	11	52
All	All	2980/3820 (78%)	2697 (91%)	223 (7%)	60 (2%)	11	52

5 of 10 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	76	PRO	20
1	A	40	ASP	10

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	124	GLY	9
1	A	141	PRO	9
1	A	122	LYS	4

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	122/154 (79%)	108±3 (88±2%)	14±3 (12±2%)	9 52
All	All	2440/3080 (79%)	2157 (88%)	283 (12%)	9 52

5 of 54 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	134	THR	18
1	A	86	ARG	17
1	A	29	ILE	16
1	A	45	ILE	15
1	A	93	GLU	14

6.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers	Suiteness
2	B	6/7 (86%)	4±1 (71±13%)	2±1 (30±15%)	0.04±0.05
All	All	106/140 (76%)	85 (80%)	36 (34%)	0.05

The overall RNA backbone suiteness is 0.03.

All unique RNA backbone outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
2	B	2	C	20
2	B	6	C	20
2	B	4	G	17
2	B	3	G	15

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
2	B	7	U	13

All unique RNA pucker outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
2	B	3	G	15
2	B	2	C	13
2	B	1	U	6
2	B	6	C	2

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
2	6MZ	B	5	2	18,25,26	1.28±0.02	1±0 (5±0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of angles that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
2	6MZ	B	5	2	16,36,39	1.53±0.04	4±0 (24±2%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	6MZ	B	5	2	-	0±0,5,27,28	0±0,3,3,3

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
2	B	5	6MZ	C9-N6	4.07	1.38	1.45	18	20

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
2	B	5	6MZ	O4'-C1'-C2'	3.62	101.64	106.93	12	20
2	B	5	6MZ	C9-N6-C6	3.25	120.07	122.87	8	20
2	B	5	6MZ	O5'-C5'-C4'	2.54	117.64	108.99	16	20
2	B	5	6MZ	C2-N1-C6	2.45	118.69	116.59	18	17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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 62% for the well-defined parts and 55% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *starch_output*

7.1.1 Bookkeeping

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

Total number of shifts	1472
Number of shifts mapped to atoms	1472
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	133	-0.25 ± 0.17	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	135	0.16 ± 0.11	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	0	—	None (insufficient data)

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 62%, i.e. 1345 atoms were assigned a chemical shift out of a possible 2174. 0 out of 22 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	380/738 (51%)	256/300 (85%)	124/298 (42%)	0/140 (0%)
Sidechain	945/1224 (77%)	664/797 (83%)	281/375 (75%)	0/52 (0%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	20/100 (20%)	20/49 (41%)	0/43 (0%)	0/8 (0%)
Sugar	0/66 (0%)	0/36 (0%)	0/30 (0%)	0/0 (—%)
Base	0/46 (0%)	0/28 (0%)	0/10 (0%)	0/8 (0%)
Overall	1345/2174 (62%)	940/1210 (78%)	405/756 (54%)	0/208 (0%)

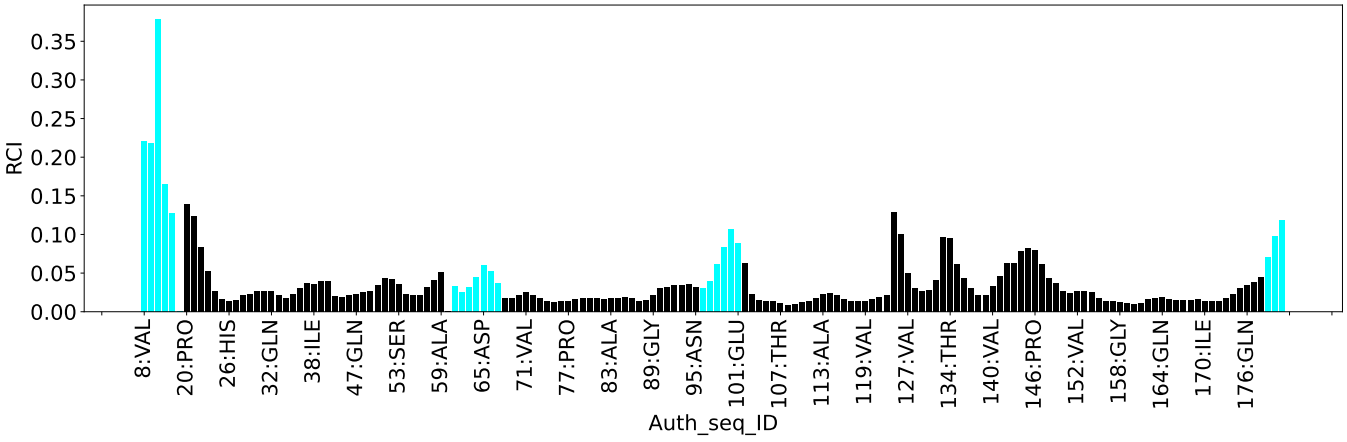
7.1.4 Statistically unusual chemical shifts (i)

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	2521
Intra-residue (|i-j|=0)	1362
Sequential (|i-j|=1)	484
Medium range (|i-j|>1 and |i-j|<5)	188
Long range (|i-j|≥5)	409
Inter-chain	67
Hydrogen bond restraints	11
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	0
Number of restraints per residue	12.7
Number of long range restraints per residue[1]	2.1

[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)	161.2	0.2
0.2-0.5 (Medium)	319.3	0.5
>0.5 (Large)	369.8	7.44

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)

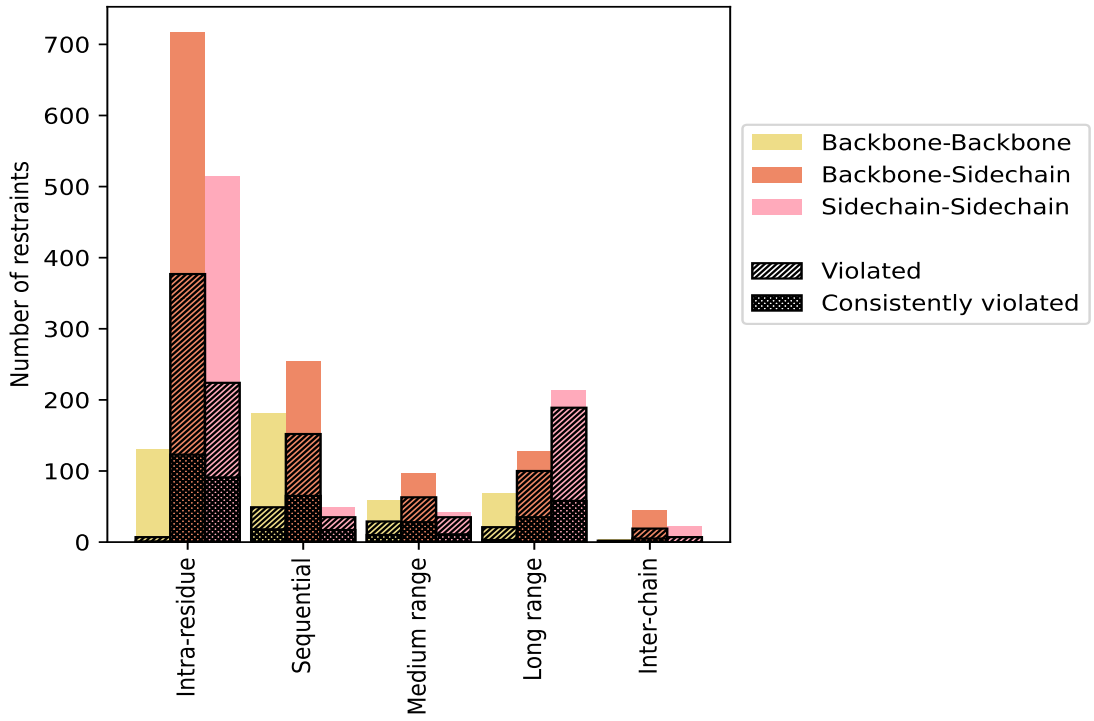
9.1 Summary of distance violations (i)

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)	1362	54.0	608	44.6	24.1	215	15.8	8.5
Backbone-Backbone	131	5.2	7	5.3	0.3	1	0.8	0.0
Backbone-Sidechain	717	28.4	377	52.6	15.0	123	17.2	4.9
Sidechain-Sidechain	514	20.4	224	43.6	8.9	91	17.7	3.6
Sequential (|i-j|=1)	484	19.2	236	48.8	9.4	100	20.7	4.0
Backbone-Backbone	181	7.2	49	27.1	1.9	18	9.9	0.7
Backbone-Sidechain	254	10.1	152	59.8	6.0	65	25.6	2.6
Sidechain-Sidechain	49	1.9	35	71.4	1.4	17	34.7	0.7
Medium range (|i-j|>1 & |i-j|<5)	188	7.5	127	67.6	5.0	49	26.1	1.9
Backbone-Backbone	58	2.3	29	50.0	1.2	10	17.2	0.4
Backbone-Sidechain	88	3.5	63	71.6	2.5	28	31.8	1.1
Sidechain-Sidechain	42	1.7	35	83.3	1.4	11	26.2	0.4
Long range (|i-j|≥5)	409	16.2	310	75.8	12.3	96	23.5	3.8
Backbone-Backbone	69	2.7	21	30.4	0.8	3	4.3	0.1
Backbone-Sidechain	127	5.0	100	78.7	4.0	35	27.6	1.4
Sidechain-Sidechain	213	8.4	189	88.7	7.5	58	27.2	2.3
Inter-chain	67	2.7	25	37.3	1.0	3	4.5	0.1
Backbone-Backbone	4	0.2	2	50.0	0.1	0	0.0	0.0
Backbone-Sidechain	42	1.7	17	40.5	0.7	3	7.1	0.1
Sidechain-Sidechain	21	0.8	6	28.6	0.2	0	0.0	0.0
Hydrogen bond	11	0.4	3	27.3	0.1	2	18.2	0.1
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	2521	100.0	1309	51.9	51.9	465	18.4	18.4
Backbone-Backbone	443	17.6	108	24.4	4.3	32	7.2	1.3
Backbone-Sidechain	1238	49.1	711	57.4	28.2	256	20.7	10.2
Sidechain-Sidechain	840	33.3	490	58.3	19.4	177	21.1	7.0

[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	394	168	84	195	12	853	0.63	5.85	0.63	0.42
2	390	171	81	198	12	852	0.63	5.48	0.63	0.42
3	393	165	80	198	12	848	0.63	5.6	0.62	0.42
4	395	168	82	204	12	861	0.62	5.56	0.62	0.41
5	385	170	85	200	12	852	0.65	5.27	0.64	0.41
6	393	170	83	203	13	862	0.65	5.61	0.67	0.42
7	395	162	85	200	12	854	0.65	6.02	0.64	0.42
8	388	175	83	208	11	865	0.65	6.04	0.65	0.43
9	396	171	81	199	16	863	0.64	5.62	0.63	0.42
10	387	164	83	198	12	844	0.65	7.05	0.66	0.42
11	393	152	89	199	11	844	0.69	7.11	0.69	0.44

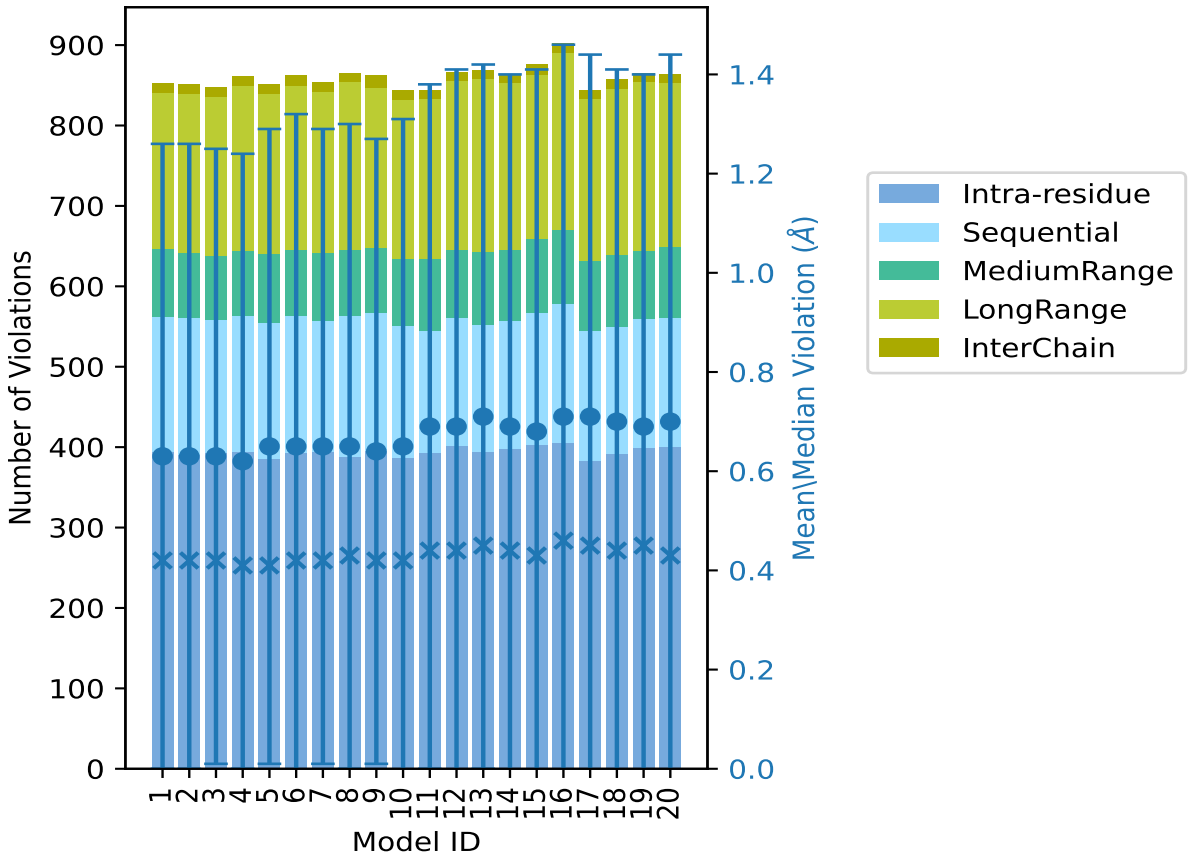
Continued on next page...

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	402	159	85	209	11	866	0.69	6.12	0.72	0.44
13	395	157	91	216	10	869	0.71	6.2	0.71	0.45
14	398	159	89	207	11	864	0.69	6.02	0.71	0.44
15	403	164	92	204	13	876	0.68	5.97	0.73	0.43
16	406	172	92	220	12	902	0.71	7.44	0.75	0.46
17	383	161	88	202	10	844	0.71	5.94	0.73	0.45
18	391	159	89	207	12	858	0.7	5.19	0.71	0.44
19	399	161	85	209	11	865	0.69	5.69	0.71	0.45
20	401	160	88	204	11	864	0.7	6.93	0.74	0.43

[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

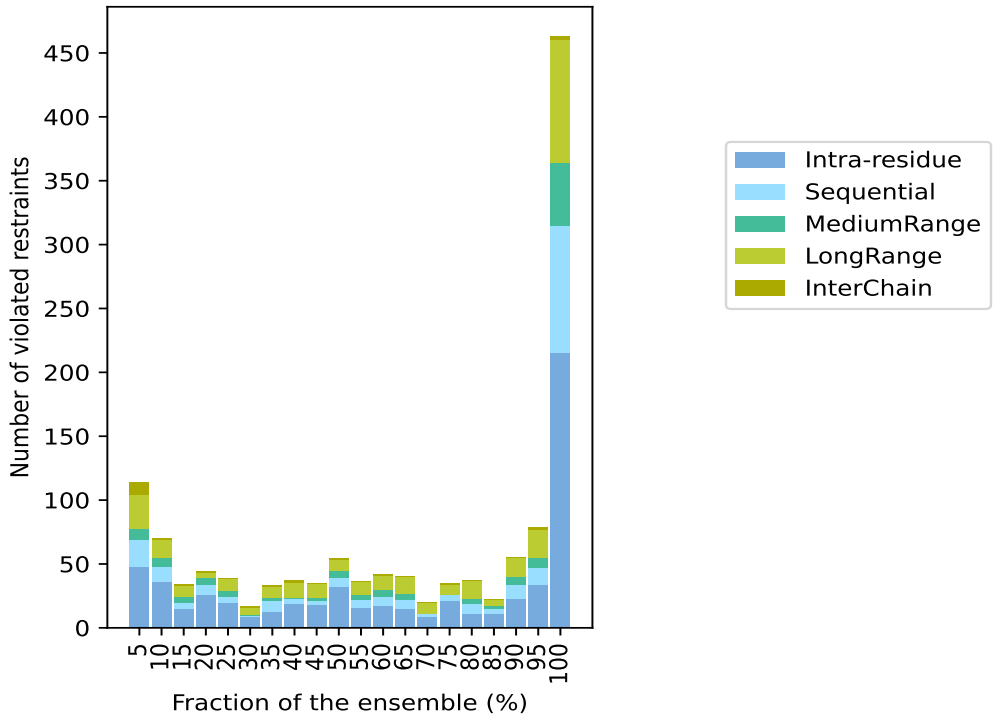
9.3 Distance violation statistics for the ensemble (i)

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, 1204(IR:754, SQ:248, MR:61, LR:99, IC:42) 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]	%
48	21	9	26	10	114	1	5.0
36	12	7	14	1	70	2	10.0
15	5	4	9	1	34	3	15.0
26	8	5	4	1	44	4	20.0
20	4	5	10	0	39	5	25.0
9	1	0	6	1	17	6	30.0
13	8	3	8	1	33	7	35.0
19	4	1	11	2	37	8	40.0
18	3	3	11	0	35	9	45.0
32	7	6	8	1	54	10	50.0
16	6	4	10	0	36	11	55.0
17	7	6	11	1	42	12	60.0
15	7	5	13	0	40	13	65.0
9	2	0	9	0	20	14	70.0
21	5	0	8	1	35	15	75.0
11	8	4	14	0	37	16	80.0
11	4	2	5	0	22	17	85.0
23	11	6	15	0	55	18	90.0
34	13	8	22	2	79	19	95.0
215	100	49	96	3	463	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

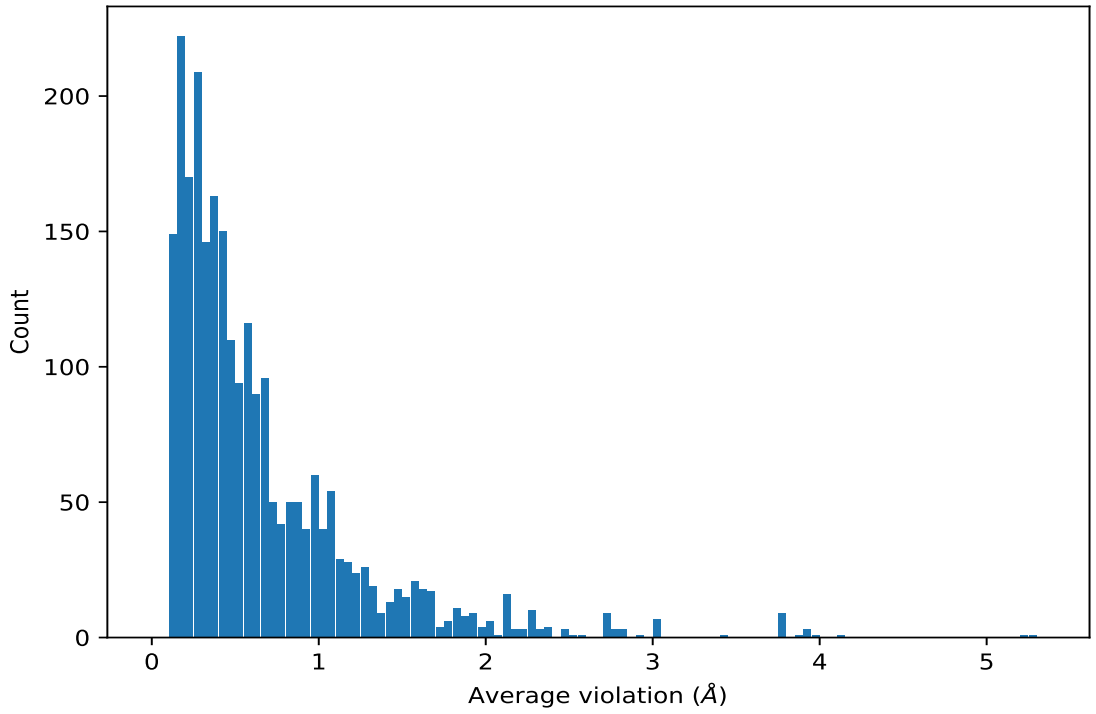
9.3.1 Bar graph : Distance violation statistics for the ensemble



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)
(1,713)	1:94:A:GLU:HG2	1:97:A:PHE:HE1	20	5.27	1.28	5.44
(1,710)	1:94:A:GLU:HA	1:90:A:LYS:HB2	20	5.21	0.48	5.11
(1,1036)	1:88:A:TYR:HD1	1:84:A:GLN:HB2	20	4.11	0.47	4.17
(1,881)	1:112:A:PRO:HB3	1:178:A:LYS:HG3	20	3.95	0.32	3.91
(1,60)	1:153:A:ILE:HA	1:152:A:VAL:HG13	20	3.9	0.03	3.9
(1,60)	1:153:A:ILE:HA	1:152:A:VAL:HG11	20	3.9	0.03	3.9
(1,60)	1:153:A:ILE:HA	1:152:A:VAL:HG12	20	3.9	0.03	3.9
(1,318)	1:112:A:PRO:HG2	1:150:A:GLU:HG2	20	3.89	0.65	3.52
(1,690)	1:38:A:ILE:HG23	1:58:A:ILE:HD11	20	3.77	0.45	3.8
(1,690)	1:38:A:ILE:HG23	1:58:A:ILE:HD12	20	3.77	0.45	3.8
(1,690)	1:38:A:ILE:HG21	1:58:A:ILE:HD12	20	3.77	0.45	3.8
(1,690)	1:38:A:ILE:HG21	1:58:A:ILE:HD13	20	3.77	0.45	3.8
(1,690)	1:38:A:ILE:HG22	1:58:A:ILE:HD13	20	3.77	0.45	3.8
(1,690)	1:38:A:ILE:HG21	1:58:A:ILE:HD11	20	3.77	0.45	3.8
(1,690)	1:38:A:ILE:HG23	1:58:A:ILE:HD13	20	3.77	0.45	3.8
(1,690)	1:38:A:ILE:HG22	1:58:A:ILE:HD12	20	3.77	0.45	3.8

Continued on next page...

Continued from previous page...

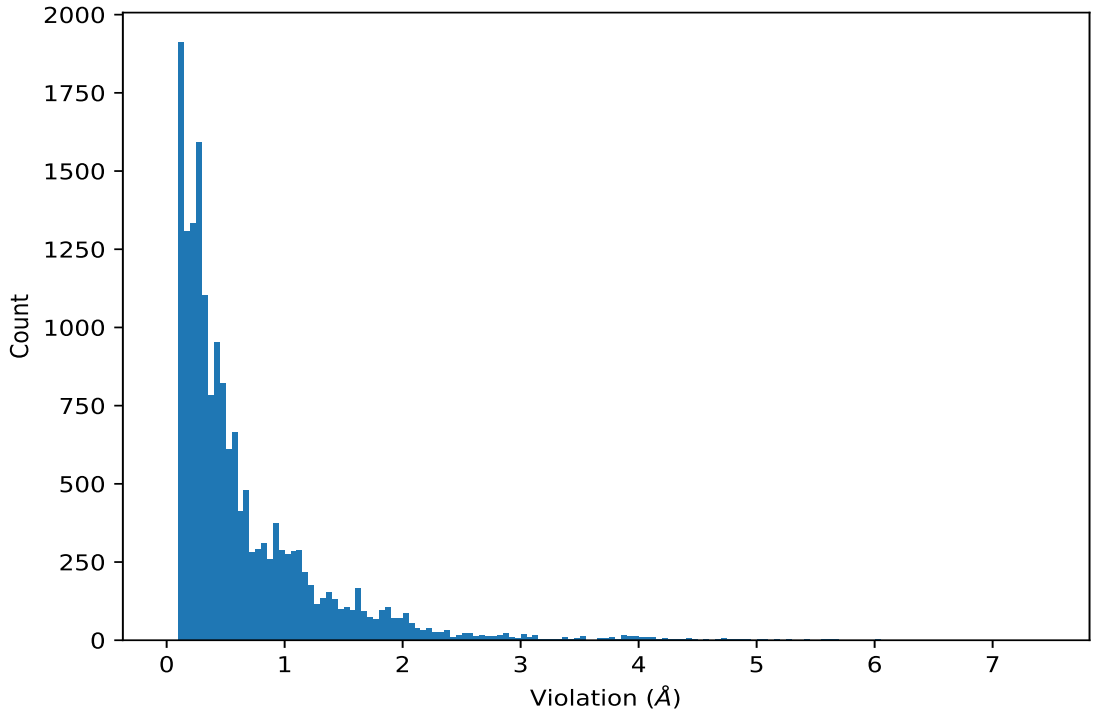
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,690)	1:38:A:ILE:HG22	1:58:A:ILE:HD11	20	3.77	0.45	3.8
(1,1051)	1:97:A:PHE:HD1	1:101:A:GLU:HG2	20	3.44	1.41	2.94
(1,239)	1:58:A:ILE:HD12	1:45:A:ILE:HG22	20	3.03	0.34	3.12
(1,239)	1:58:A:ILE:HD13	1:45:A:ILE:HG23	20	3.03	0.34	3.12
(1,239)	1:58:A:ILE:HD11	1:45:A:ILE:HG21	20	3.03	0.34	3.12
(1,239)	1:58:A:ILE:HD13	1:45:A:ILE:HG22	20	3.03	0.34	3.12
(1,239)	1:58:A:ILE:HD13	1:45:A:ILE:HG21	20	3.03	0.34	3.12
(1,239)	1:58:A:ILE:HD11	1:45:A:ILE:HG23	20	3.03	0.34	3.12
(1,239)	1:58:A:ILE:HD12	1:45:A:ILE:HG21	20	3.03	0.34	3.12
(1,2001)	1:31:A:ALA:H	1:29:A:ILE:HG13	20	2.93	0.25	2.93

[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)
(1,713)	1:94:A:GLU:HG2	1:97:A:PHE:HE1	16	7.44
(1,713)	1:94:A:GLU:HG2	1:97:A:PHE:HE1	11	7.11
(1,713)	1:94:A:GLU:HG2	1:97:A:PHE:HE1	10	7.05
(1,713)	1:94:A:GLU:HG2	1:97:A:PHE:HE1	20	6.93
(1,713)	1:94:A:GLU:HG2	1:97:A:PHE:HE1	13	6.2
(1,713)	1:94:A:GLU:HG2	1:97:A:PHE:HE1	12	6.12
(1,710)	1:94:A:GLU:HA	1:90:A:LYS:HB2	8	6.04
(1,713)	1:94:A:GLU:HG2	1:97:A:PHE:HE1	14	6.02
(1,710)	1:94:A:GLU:HA	1:90:A:LYS:HB2	7	6.02
(1,1051)	1:97:A:PHE:HD1	1:101:A:GLU:HG2	15	5.97

10 Dihedral-angle violation analysis (i)

No dihedral-angle restraints found