Full wwPDB NMR Structure Validation Report (i)

Apr 4, 2024 – 02:03 AM EDT

PDB ID	:	2KBT
BMRB ID	:	16053
Title	:	Attachment of an NMR-invisible solubility enhancement tag (INSET) using a sortase-mediated protein ligation method
Authors	:	Kumeta, H.; Kobashigawa, Y.; Ogura, K.; Inagaki, F.
Deposited on	:	2008-12-07

This is a Full wwPDB NMR Structure Validation 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:

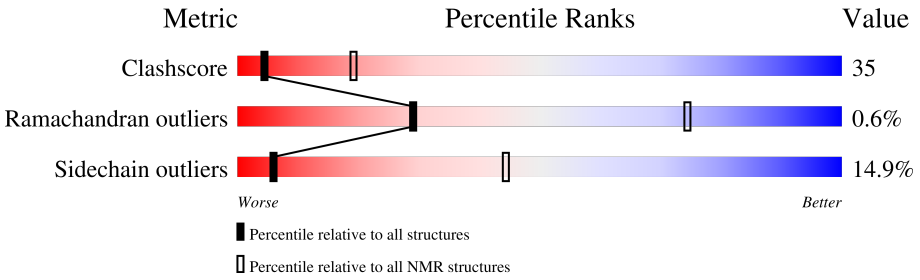
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.1

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 91%.

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	142	18% 17% • 11% 51%

2 Ensemble composition and analysis

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: *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:785-A:814, A:819-A:841 (53)	0.41	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 1 single-model cluster was found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1102 atoms, of which 536 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Proto-oncogene vav,Immunoglobulin G-binding protein G.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	69	1102	362	536	96	107	1	0

There are 26 discrepancies between the modelled and reference sequences:

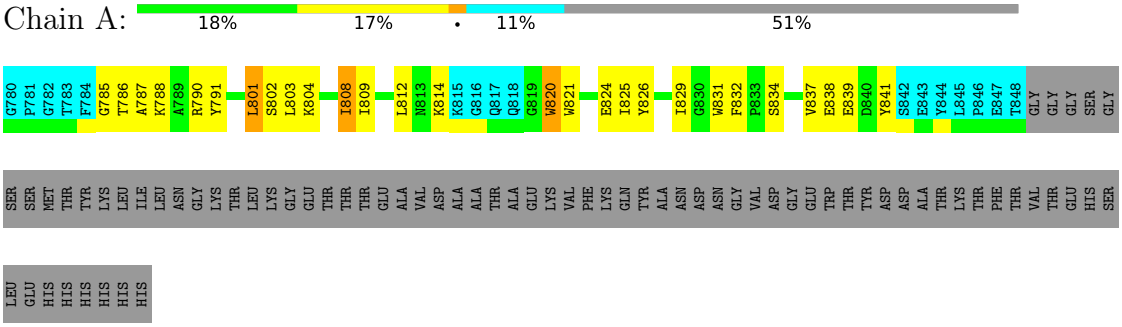
Chain	Residue	Modelled	Actual	Comment	Reference
A	780	GLY	-	expression tag	UNP P27870
A	781	PRO	-	expression tag	UNP P27870
A	782	GLY	-	expression tag	UNP P27870
A	783	THR	-	expression tag	UNP P27870
A	845	LEU	-	linker	UNP P27870
A	846	PRO	-	linker	UNP P27870
A	847	GLU	-	linker	UNP P27870
A	848	THR	-	linker	UNP P27870
A	849	GLY	-	linker	UNP P27870
A	850	GLY	-	linker	UNP P27870
A	851	GLY	-	linker	UNP P27870
A	852	SER	-	linker	UNP P27870
A	853	GLY	-	linker	UNP P27870
A	854	SER	-	linker	UNP P27870
A	855	SER	-	linker	UNP P27870
A	856	MET	-	linker	UNP P27870
A	912	HIS	-	expression tag	UNP P06654
A	913	SER	-	expression tag	UNP P06654
A	914	LEU	-	expression tag	UNP P06654
A	915	GLU	-	expression tag	UNP P06654
A	916	HIS	-	expression tag	UNP P06654
A	917	HIS	-	expression tag	UNP P06654
A	918	HIS	-	expression tag	UNP P06654
A	919	HIS	-	expression tag	UNP P06654
A	920	HIS	-	expression tag	UNP P06654
A	921	HIS	-	expression tag	UNP P06654

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: Proto-oncogene vav,Immunoglobulin G-binding protein G

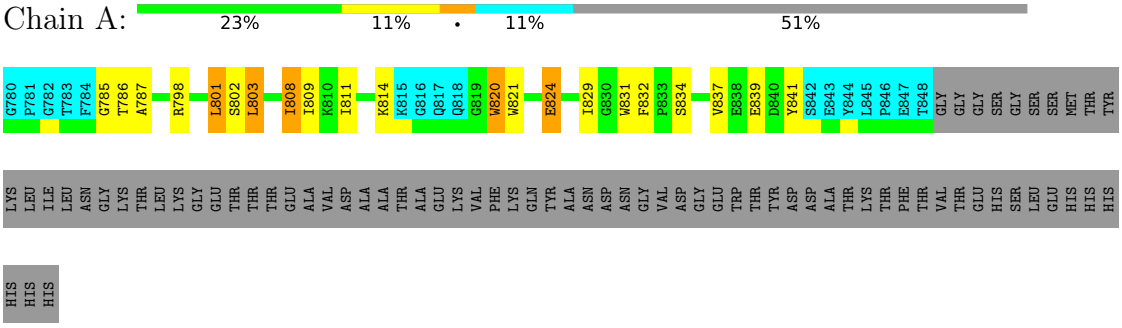


4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

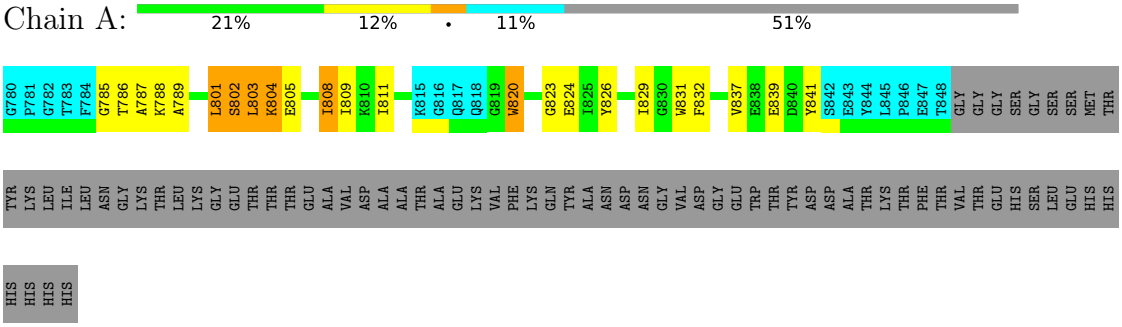
• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G



4.2.2 Score per residue for model 2

• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G

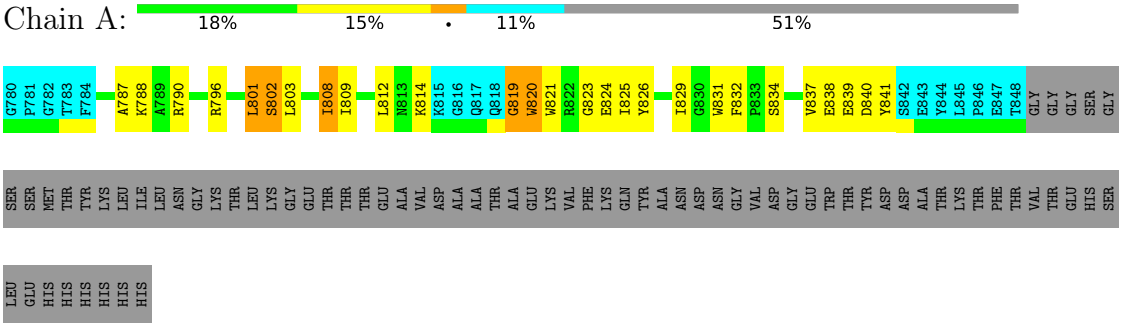
Chain A:



4.2.3 Score per residue for model 3

• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G

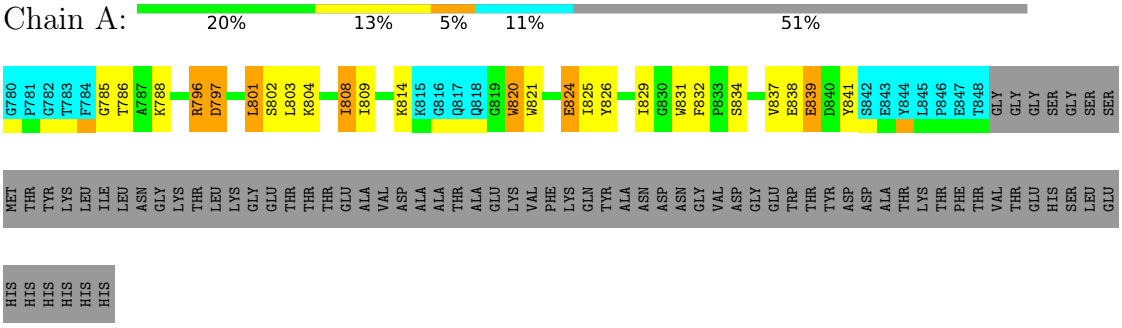
Chain A:



4.2.4 Score per residue for model 4

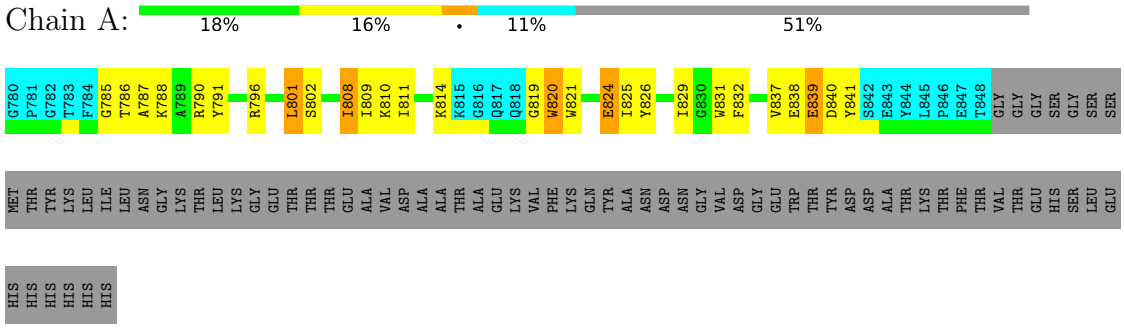
• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G

Chain A:



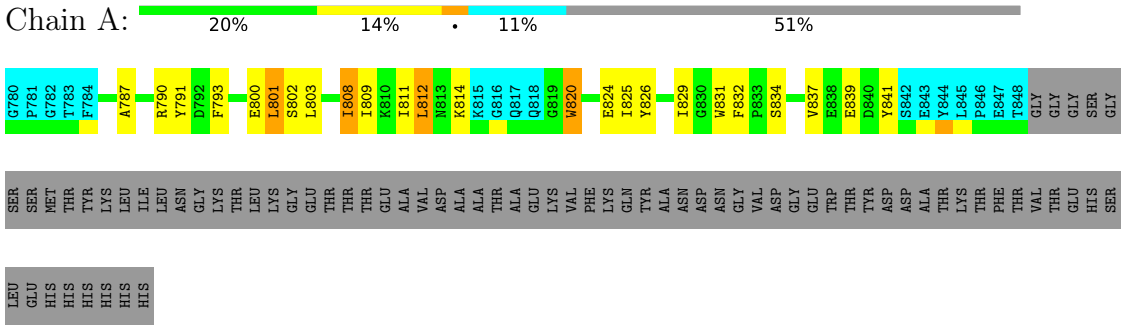
4.2.5 Score per residue for model 5

• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G



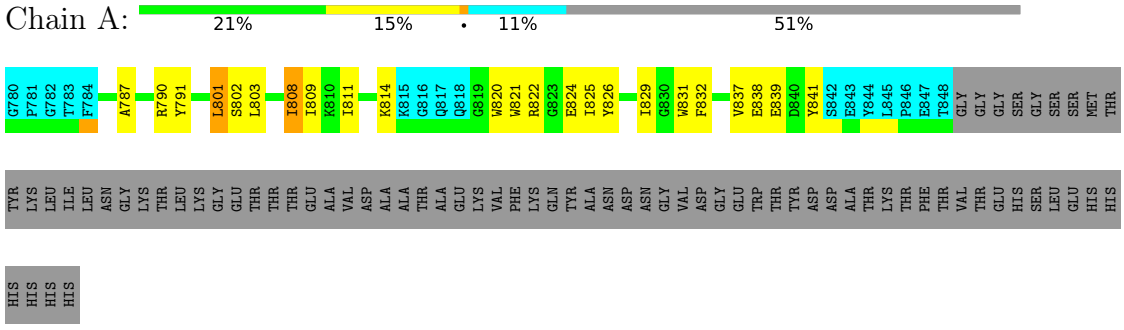
4.2.6 Score per residue for model 6

• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G



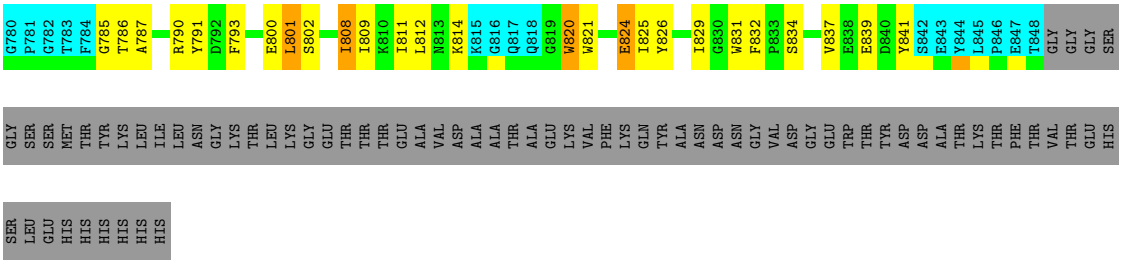
4.2.7 Score per residue for model 7

• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G



4.2.8 Score per residue for model 8

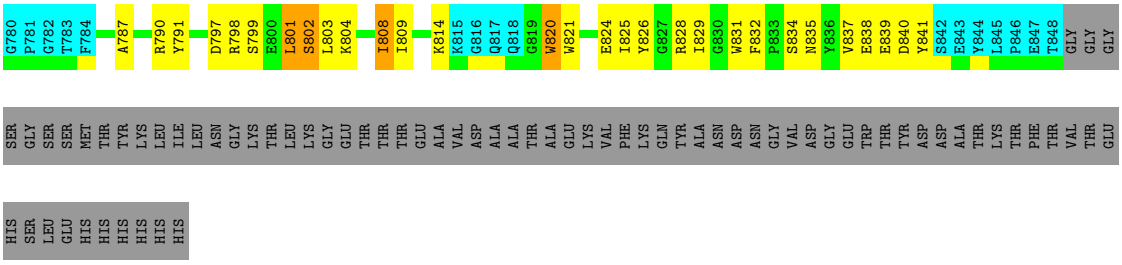
• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G



4.2.9 Score per residue for model 9

• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G

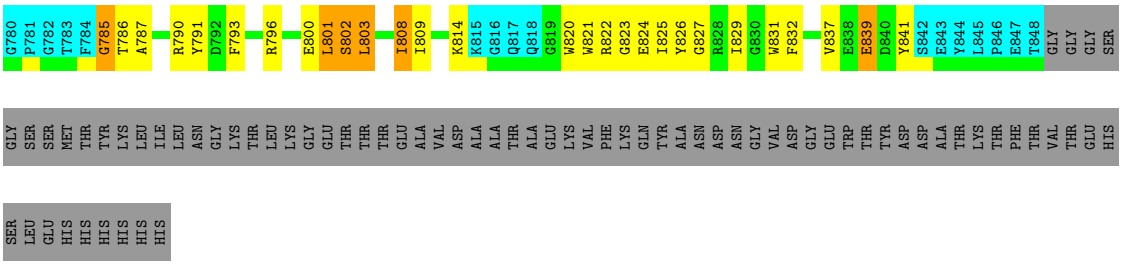
Chain A: 17% 18% • 11% 51%



4.2.10 Score per residue for model 10

• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G

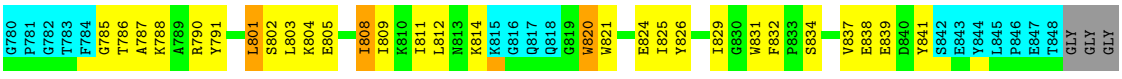
Chain A: 18% 15% • 11% 51%

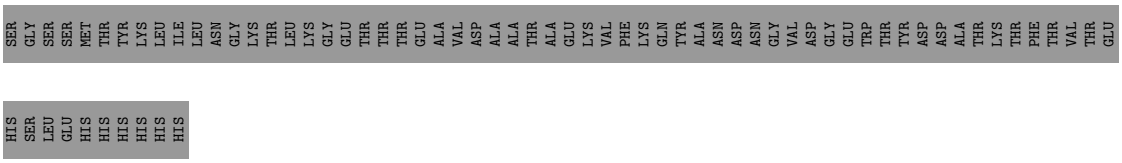


4.2.11 Score per residue for model 11

• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G

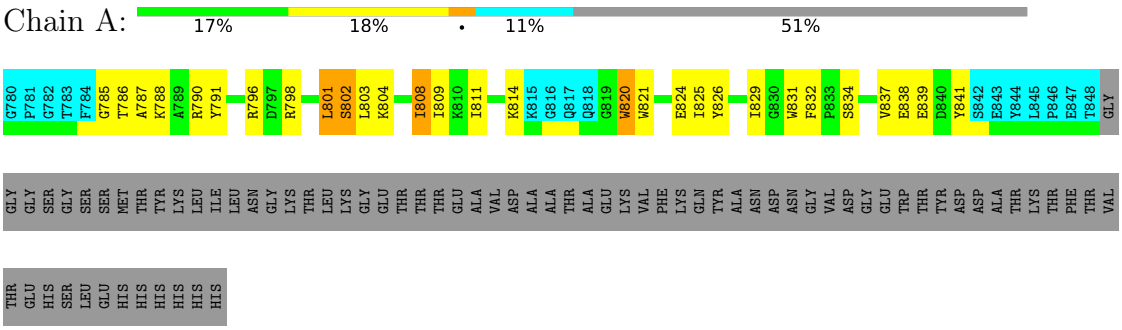
Chain A: 17% 18% • 11% 51%





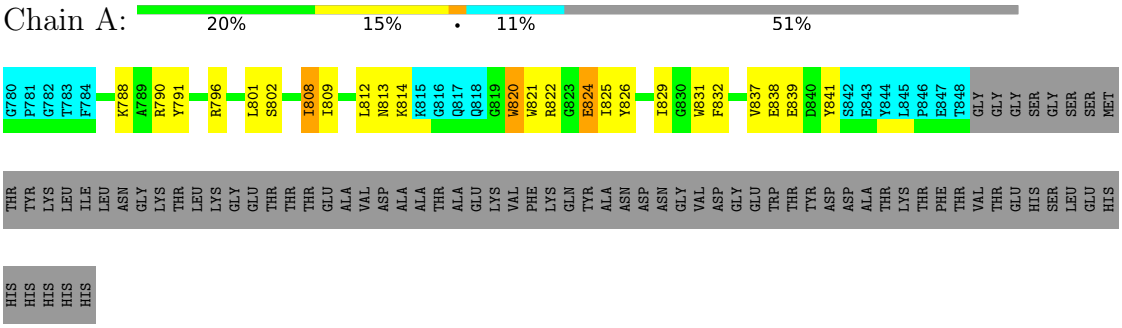
4.2.12 Score per residue for model 12

• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G



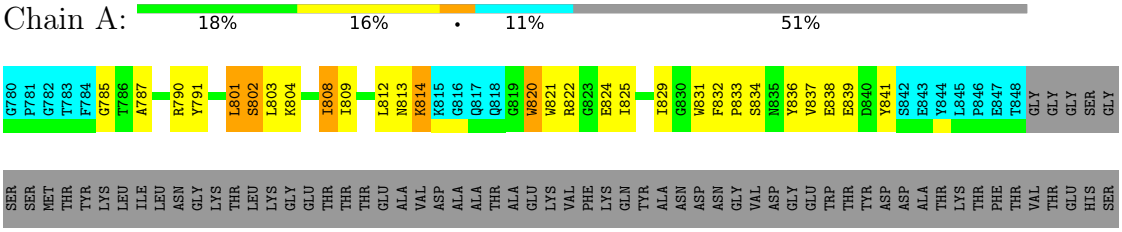
4.2.13 Score per residue for model 13

• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G



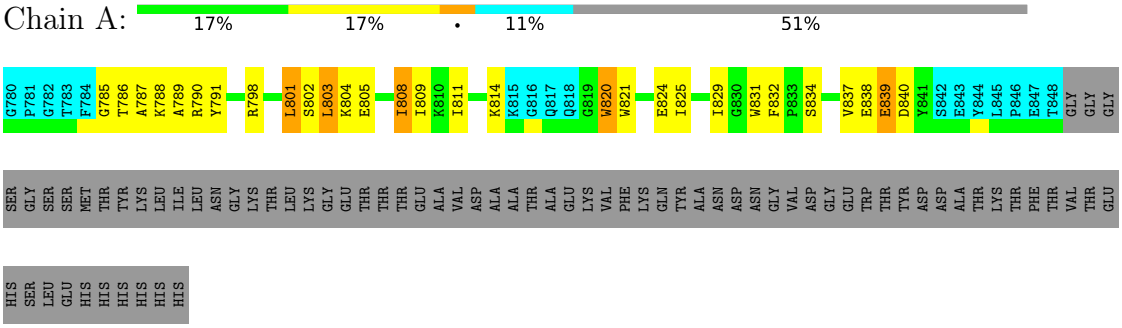
4.2.14 Score per residue for model 14

• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G



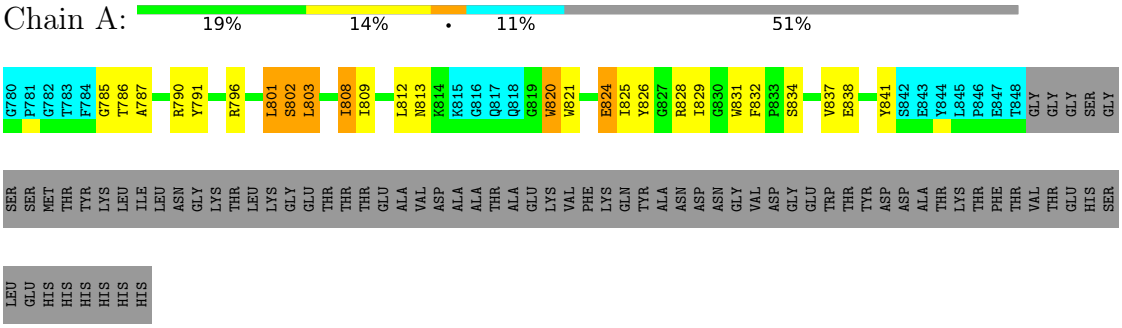
4.2.15 Score per residue for model 15 (medoid)

• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G



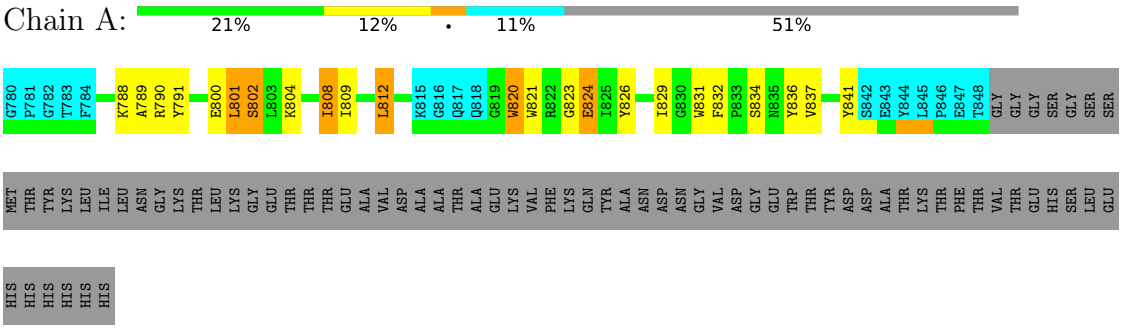
4.2.16 Score per residue for model 16

• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G



4.2.17 Score per residue for model 17

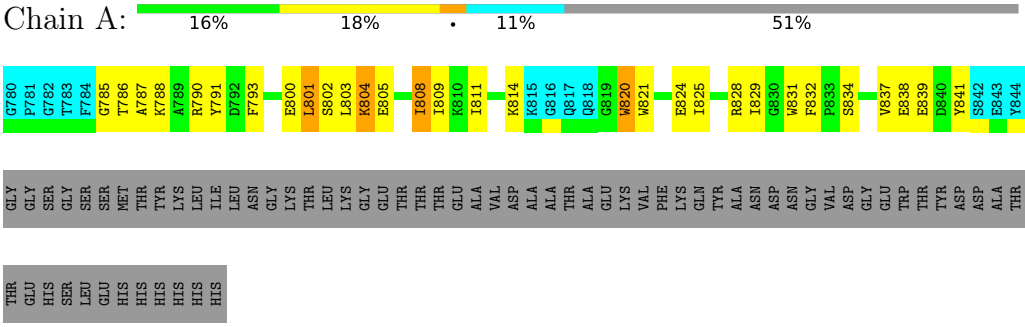
• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G



4.2.18 Score per residue for model 18

• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G

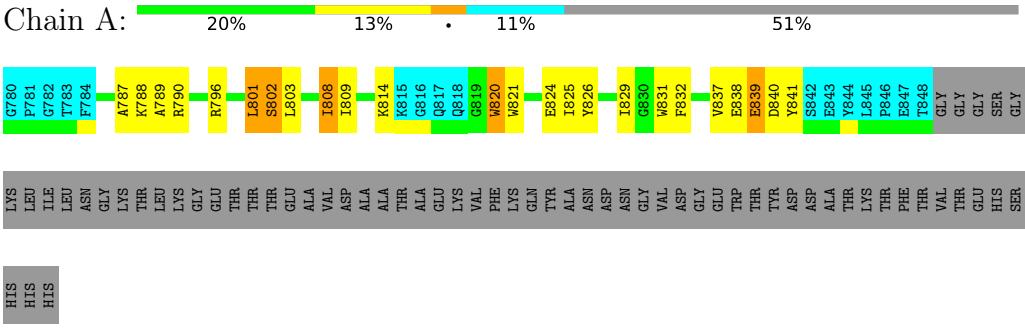
Chain A:



4.2.19 Score per residue for model 19

• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G

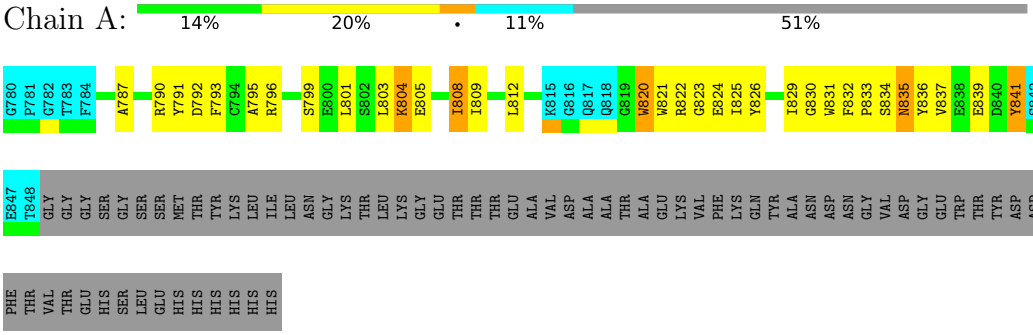
Chain A:



4.2.20 Score per residue for model 20

• Molecule 1: Proto-oncogene vav,Immunoglobulin G-binding protein G

Chain A:



5 Refinement protocol and experimental data overview

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

Of the 100 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
Sparky	refinement	3.110
CYANA	structure solution	2.1

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	862
Number of shifts mapped to atoms	862
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	91%

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	444	423	423	30±6
All	All	8880	8460	8460	606

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 35.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:811:ILE:HG21	1:A:814:LYS:NZ	0.90	1.80	7	1
1:A:809:ILE:HG21	1:A:832:PHE:CZ	0.87	2.04	3	20
1:A:824:GLU:HB2	1:A:829:ILE:HD13	0.76	1.58	12	17
1:A:820:TRP:CE2	1:A:831:TRP:CE3	0.74	2.75	16	17
1:A:824:GLU:HB3	1:A:829:ILE:HD13	0.71	1.62	17	3
1:A:808:ILE:HD13	1:A:808:ILE:N	0.71	2.01	10	20
1:A:803:LEU:HD11	1:A:809:ILE:HG12	0.68	1.64	3	7
1:A:811:ILE:HG21	1:A:814:LYS:HZ1	0.68	1.49	7	1
1:A:812:LEU:CB	1:A:822:ARG:HH21	0.67	2.01	14	1
1:A:787:ALA:HB2	1:A:811:ILE:HD11	0.67	1.67	12	9
1:A:809:ILE:CG2	1:A:832:PHE:CZ	0.66	2.78	12	20
1:A:832:PHE:CE2	1:A:837:VAL:HG21	0.63	2.29	10	17
1:A:803:LEU:HD21	1:A:837:VAL:CG1	0.63	2.24	16	2
1:A:785:GLY:O	1:A:786:THR:HG23	0.63	1.93	10	11

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:832:PHE:HE2	1:A:837:VAL:HG21	0.62	1.55	12	18
1:A:809:ILE:HG21	1:A:832:PHE:CE1	0.61	2.30	15	18
1:A:808:ILE:O	1:A:825:ILE:HD11	0.61	1.96	9	9
1:A:787:ALA:CB	1:A:811:ILE:HD11	0.60	2.25	18	5
1:A:820:TRP:NE1	1:A:831:TRP:CE3	0.60	2.69	15	14
1:A:832:PHE:CD2	1:A:837:VAL:HG11	0.60	2.32	15	5
1:A:821:TRP:O	1:A:832:PHE:CD1	0.60	2.55	13	9
1:A:790:ARG:C	1:A:791:TYR:CD1	0.60	2.75	14	9
1:A:821:TRP:O	1:A:832:PHE:CE1	0.59	2.55	14	11
1:A:820:TRP:NE1	1:A:831:TRP:CZ3	0.59	2.70	4	13
1:A:824:GLU:CB	1:A:829:ILE:HD13	0.59	2.26	17	1
1:A:821:TRP:CH2	1:A:839:GLU:OE1	0.59	2.55	3	1
1:A:801:LEU:HD23	1:A:802:SER:N	0.58	2.13	17	18
1:A:834:SER:HA	1:A:837:VAL:HG22	0.57	1.75	20	5
1:A:793:PHE:CE2	1:A:800:GLU:CD	0.56	2.79	8	4
1:A:793:PHE:CE2	1:A:800:GLU:OE1	0.55	2.59	10	1
1:A:801:LEU:HD11	1:A:824:GLU:N	0.54	2.17	19	4
1:A:826:TYR:CD2	1:A:828:ARG:NH2	0.54	2.76	9	1
1:A:801:LEU:HD11	1:A:823:GLY:C	0.53	2.24	10	4
1:A:814:LYS:CD	1:A:821:TRP:CE2	0.52	2.92	8	1
1:A:826:TYR:N	1:A:826:TYR:CD1	0.52	2.75	3	10
1:A:808:ILE:N	1:A:808:ILE:CD1	0.52	2.70	16	15
1:A:811:ILE:HG21	1:A:814:LYS:HZ3	0.52	1.63	7	1
1:A:821:TRP:CZ3	1:A:834:SER:OG	0.51	2.59	12	1
1:A:812:LEU:HB2	1:A:822:ARG:HH21	0.51	1.62	14	1
1:A:820:TRP:CZ3	1:A:831:TRP:CG	0.51	2.99	10	2
1:A:814:LYS:HA	1:A:821:TRP:CD1	0.51	2.41	5	6
1:A:825:ILE:HG22	1:A:826:TYR:CD2	0.50	2.41	10	8
1:A:814:LYS:CD	1:A:821:TRP:CD1	0.50	2.95	7	2
1:A:824:GLU:CG	1:A:825:ILE:N	0.50	2.75	11	12
1:A:821:TRP:CH2	1:A:834:SER:OG	0.50	2.57	15	1
1:A:825:ILE:HG23	1:A:826:TYR:CE1	0.50	2.42	8	7
1:A:821:TRP:CZ3	1:A:834:SER:HB2	0.50	2.42	14	9
1:A:808:ILE:O	1:A:825:ILE:CD1	0.49	2.60	7	7
1:A:790:ARG:C	1:A:791:TYR:CG	0.49	2.86	18	9
1:A:820:TRP:CE3	1:A:831:TRP:HB3	0.49	2.42	11	14
1:A:793:PHE:CZ	1:A:833:PRO:HG3	0.49	2.42	20	1
1:A:785:GLY:O	1:A:786:THR:CG2	0.49	2.60	10	4
1:A:814:LYS:CE	1:A:821:TRP:CE2	0.49	2.96	1	1
1:A:796:ARG:CZ	1:A:797:ASP:CG	0.49	2.81	4	1
1:A:812:LEU:HB3	1:A:822:ARG:HH21	0.49	1.66	14	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:809:ILE:HG21	1:A:832:PHE:CE2	0.48	2.43	13	4
1:A:820:TRP:CE3	1:A:831:TRP:CB	0.48	2.96	10	2
1:A:821:TRP:CH2	1:A:839:GLU:OE2	0.48	2.66	12	1
1:A:825:ILE:HG23	1:A:826:TYR:CD1	0.48	2.43	3	9
1:A:801:LEU:HD23	1:A:802:SER:H	0.48	1.68	5	17
1:A:837:VAL:HG23	1:A:839:GLU:OE2	0.48	2.08	14	1
1:A:825:ILE:O	1:A:828:ARG:CD	0.48	2.61	18	1
1:A:814:LYS:HG2	1:A:821:TRP:CD1	0.48	2.44	1	3
1:A:814:LYS:NZ	1:A:821:TRP:CE3	0.48	2.81	3	1
1:A:821:TRP:CZ3	1:A:839:GLU:HG3	0.47	2.43	20	2
1:A:789:ALA:HA	1:A:837:VAL:HG12	0.47	1.84	19	1
1:A:787:ALA:HB1	1:A:838:GLU:O	0.47	2.10	16	3
1:A:790:ARG:HB2	1:A:791:TYR:CE1	0.47	2.44	16	1
1:A:793:PHE:CD2	1:A:800:GLU:OE2	0.47	2.68	6	2
1:A:814:LYS:HG3	1:A:821:TRP:CD1	0.47	2.44	7	2
1:A:824:GLU:OE1	1:A:825:ILE:CA	0.47	2.62	13	1
1:A:821:TRP:CH2	1:A:834:SER:HB2	0.46	2.45	4	5
1:A:822:ARG:HG3	1:A:831:TRP:CD1	0.46	2.46	7	2
1:A:822:ARG:NE	1:A:831:TRP:NE1	0.46	2.63	7	1
1:A:790:ARG:O	1:A:791:TYR:CG	0.46	2.69	15	6
1:A:788:LYS:HB2	1:A:841:TYR:CE2	0.46	2.46	19	5
1:A:821:TRP:CE3	1:A:834:SER:HB2	0.46	2.44	15	2
1:A:821:TRP:CH2	1:A:839:GLU:HG3	0.46	2.46	9	1
1:A:826:TYR:CD1	1:A:826:TYR:N	0.45	2.80	8	3
1:A:814:LYS:HD2	1:A:821:TRP:CD1	0.45	2.46	10	2
1:A:801:LEU:CB	1:A:831:TRP:O	0.45	2.65	20	1
1:A:836:TYR:O	1:A:837:VAL:CG1	0.45	2.65	14	3
1:A:793:PHE:CE2	1:A:833:PRO:HD3	0.45	2.47	20	1
1:A:803:LEU:HD11	1:A:809:ILE:CG1	0.45	2.40	3	1
1:A:788:LYS:N	1:A:838:GLU:O	0.45	2.50	18	5
1:A:787:ALA:O	1:A:809:ILE:N	0.45	2.50	7	6
1:A:825:ILE:CG2	1:A:826:TYR:CD2	0.45	3.00	10	2
1:A:839:GLU:C	1:A:841:TYR:N	0.45	2.70	7	11
1:A:790:ARG:C	1:A:791:TYR:CD2	0.45	2.90	8	2
1:A:790:ARG:O	1:A:790:ARG:NH1	0.45	2.50	7	1
1:A:824:GLU:OE2	1:A:825:ILE:CA	0.45	2.65	16	2
1:A:814:LYS:HE2	1:A:821:TRP:CZ2	0.44	2.48	1	1
1:A:790:ARG:NH2	1:A:838:GLU:OE1	0.44	2.50	16	1
1:A:787:ALA:N	1:A:809:ILE:O	0.44	2.50	20	1
1:A:788:LYS:O	1:A:838:GLU:N	0.44	2.50	13	2
1:A:801:LEU:HD11	1:A:824:GLU:CA	0.44	2.43	12	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:801:LEU:CD1	1:A:823:GLY:C	0.44	2.86	10	2
1:A:796:ARG:CZ	1:A:797:ASP:OD2	0.43	2.65	4	1
1:A:822:ARG:HB2	1:A:831:TRP:CD2	0.43	2.48	10	1
1:A:808:ILE:CG2	1:A:841:TYR:OH	0.43	2.65	13	2
1:A:822:ARG:HD2	1:A:831:TRP:CZ2	0.43	2.48	7	1
1:A:820:TRP:CD1	1:A:831:TRP:CE3	0.43	3.07	4	2
1:A:824:GLU:OE1	1:A:825:ILE:N	0.43	2.50	13	1
1:A:814:LYS:NZ	1:A:839:GLU:CD	0.43	2.72	3	1
1:A:793:PHE:CE2	1:A:833:PRO:CD	0.43	3.01	20	1
1:A:801:LEU:HD13	1:A:830:GLY:N	0.43	2.29	20	1
1:A:787:ALA:CB	1:A:838:GLU:O	0.43	2.67	9	3
1:A:790:ARG:NH2	1:A:838:GLU:CD	0.43	2.72	3	2
1:A:824:GLU:OE2	1:A:825:ILE:N	0.43	2.52	16	2
1:A:839:GLU:O	1:A:841:TYR:N	0.43	2.51	19	8
1:A:789:ALA:HB2	1:A:809:ILE:CD1	0.42	2.44	2	3
1:A:801:LEU:HD22	1:A:832:PHE:CD2	0.42	2.49	20	1
1:A:820:TRP:CZ3	1:A:831:TRP:CD2	0.42	3.07	10	1
1:A:790:ARG:CG	1:A:838:GLU:CD	0.42	2.88	19	1
1:A:814:LYS:HD3	1:A:821:TRP:CE2	0.42	2.49	3	2
1:A:820:TRP:CD2	1:A:831:TRP:CE3	0.42	3.07	10	1
1:A:811:ILE:HG21	1:A:814:LYS:CE	0.42	2.42	7	1
1:A:814:LYS:HD2	1:A:821:TRP:CE2	0.42	2.50	7	3
1:A:833:PRO:HG2	1:A:836:TYR:CD1	0.42	2.50	14	1
1:A:814:LYS:HZ1	1:A:839:GLU:CD	0.42	2.18	3	1
1:A:812:LEU:HD12	1:A:823:GLY:HA2	0.42	1.91	17	1
1:A:801:LEU:CD2	1:A:832:PHE:CD2	0.42	3.02	20	1
1:A:803:LEU:HD11	1:A:809:ILE:CD1	0.41	2.45	14	2
1:A:836:TYR:C	1:A:837:VAL:HG13	0.41	2.35	14	3
1:A:790:ARG:HB2	1:A:791:TYR:CE2	0.41	2.50	13	1
1:A:822:ARG:HG3	1:A:831:TRP:CE2	0.41	2.51	20	2
1:A:821:TRP:CZ3	1:A:839:GLU:OE1	0.41	2.74	3	1
1:A:800:GLU:CG	1:A:801:LEU:N	0.41	2.83	6	1
1:A:791:TYR:CD1	1:A:791:TYR:N	0.41	2.88	10	1
1:A:808:ILE:HG21	1:A:841:TYR:OH	0.41	2.16	13	1
1:A:836:TYR:CD1	1:A:836:TYR:N	0.41	2.89	14	1
1:A:812:LEU:HD12	1:A:823:GLY:CA	0.41	2.46	20	2
1:A:835:ASN:OD1	1:A:835:ASN:N	0.41	2.53	20	1
1:A:821:TRP:CZ3	1:A:834:SER:CB	0.41	3.03	15	1
1:A:792:ASP:OD2	1:A:792:ASP:N	0.40	2.55	20	1
1:A:803:LEU:HD21	1:A:832:PHE:HB2	0.40	1.93	2	1
1:A:804:LYS:O	1:A:805:GLU:C	0.40	2.60	2	5

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:839:GLU:O	1:A:840:ASP:C	0.40	2.60	5	5
1:A:824:GLU:OE2	1:A:827:GLY:C	0.40	2.60	10	1
1:A:839:GLU:CD	1:A:839:GLU:C	0.40	2.80	10	1
1:A:795:ALA:O	1:A:796:ARG:C	0.40	2.60	20	1

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	53/142 (37%)	47±2 (89±3%)	5±1 (10±3%)	0±0 (1±1%)	29	74
All	All	1060/2840 (37%)	948 (89%)	106 (10%)	6 (1%)	29	74

All 3 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	819	GLY	2
1	A	785	GLY	2
1	A	841	TYR	2

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	44/116 (38%)	37±2 (85±4%)	7±2 (15±4%)	6	44
All	All	880/2320 (38%)	749 (85%)	131 (15%)	6	44

All 18 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	808	ILE	20
1	A	801	LEU	19
1	A	820	TRP	18
1	A	803	LEU	12
1	A	802	SER	10
1	A	804	LYS	8
1	A	796	ARG	8
1	A	824	GLU	7
1	A	839	GLU	7
1	A	812	LEU	7
1	A	798	ARG	4
1	A	797	ASP	2
1	A	799	SER	2
1	A	835	ASN	2
1	A	813	ASN	2
1	A	814	LYS	1
1	A	828	ARG	1
1	A	800	GLU	1

6.3.3 RNA

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains

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

6.5 Carbohydrates

There are no monosaccharides in this entry.

6.6 Ligand geometry

There are no ligands in this entry.

6.7 Other polymers

There are no such molecules in this entry.

6.8 Polymer linkage issues

There are no chain breaks in this entry.

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 91% for the well-defined parts and 91% 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	862
Number of shifts mapped to atoms	862
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	8

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	69	-0.27 ± 0.13	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	60	0.16 ± 0.31	None needed (< 0.5 ppm)
$^{13}C'$	64	-0.12 ± 0.13	None needed (< 0.5 ppm)
^{15}N	65	-0.23 ± 0.36	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 91%, i.e. 678 atoms were assigned a chemical shift out of a possible 748. 0 out of 4 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	267/269 (99%)	111/111 (100%)	104/106 (98%)	52/52 (100%)
Sidechain	335/387 (87%)	224/247 (91%)	107/119 (90%)	4/21 (19%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	76/92 (83%)	44/44 (100%)	29/45 (64%)	3/3 (100%)
Overall	678/748 (91%)	379/402 (94%)	240/270 (89%)	59/76 (78%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 91%, i.e. 862 atoms were assigned a chemical shift out of a possible 945. 0 out of 5 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	341/348 (98%)	143/144 (99%)	133/138 (96%)	65/66 (98%)
Sidechain	431/486 (89%)	287/310 (93%)	138/152 (91%)	6/24 (25%)
Aromatic	90/111 (81%)	53/53 (100%)	34/55 (62%)	3/3 (100%)
Overall	862/945 (91%)	483/507 (95%)	305/345 (88%)	74/93 (80%)

7.1.4 Statistically unusual chemical shifts (i)

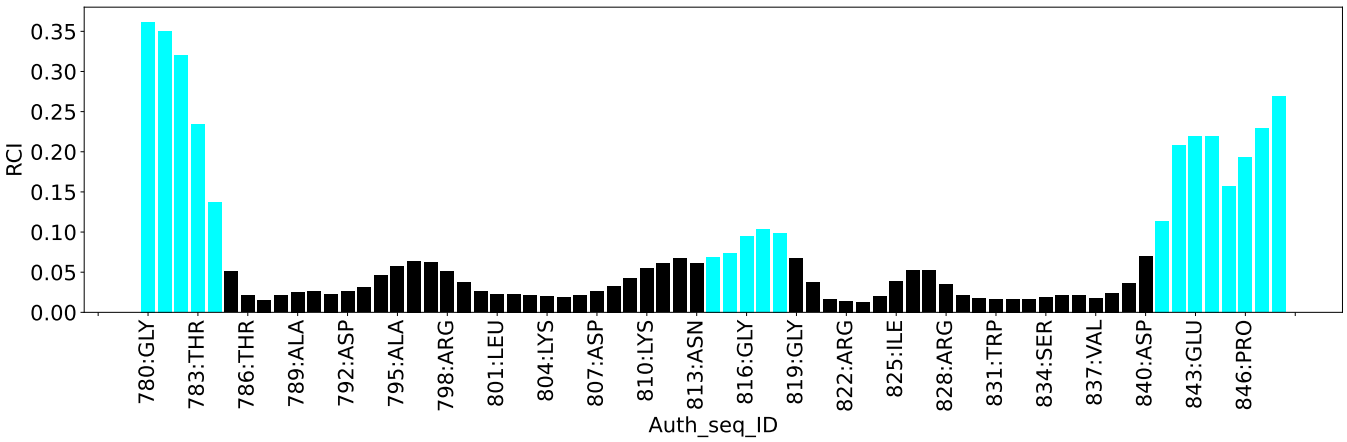
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	814	LYS	HE2	0.94	1.95 – 3.88	-10.2
1	A	822	ARG	HB2	-0.77	0.52 – 3.08	-10.0
1	A	814	LYS	HG2	-0.99	0.13 – 2.61	-9.5
1	A	834	SER	HB2	1.60	2.61 – 5.13	-9.0
1	A	834	SER	HB3	1.85	2.49 – 5.20	-7.4
1	A	833	PRO	HG2	-0.04	0.41 – 3.45	-6.5
1	A	822	ARG	HG2	0.16	0.26 – 2.87	-5.4
1	A	814	LYS	CD	34.56	23.50 – 34.42	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:


RCI
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
780:GLY
783:THR
786:THR
789:ALA
792:ASP
795:ALA
798:ARG
801:LEU
804:LYS
807:ASP
810:LYS
813:ASN
816:GLY
819:GLY
822:ARG
825:ILE
828:ARG
831:TRP
834:SER
837:VAL
840:ASP
843:GLU
846:PRO
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	1051
Intra-residue (|i-j|=0)	281
Sequential (|i-j|=1)	309
Medium range (|i-j|>1 and |i-j|<5)	75
Long range (|i-j|≥5)	386
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	7.4
Number of long range restraints per residue[1]	2.7

[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)	0.6	0.2
0.2-0.5 (Medium)	0.5	0.47
>0.5 (Large)	1.9	1.6

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)

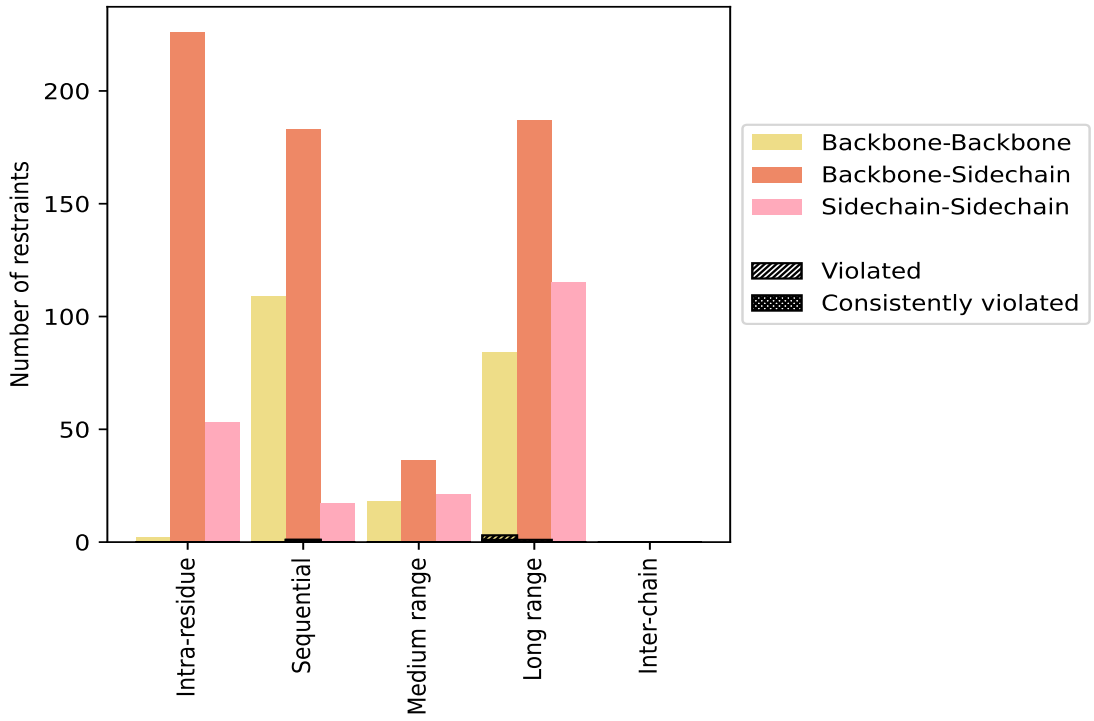
9.1 Summary of distance violations (i)

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)	281	26.7	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	2	0.2	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	226	21.5	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	53	5.0	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	309	29.4	1	0.3	0.1	1	0.3	0.1
Backbone-Backbone	109	10.4	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	183	17.4	1	0.5	0.1	1	0.5	0.1
Sidechain-Sidechain	17	1.6	0	0.0	0.0	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	75	7.1	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	18	1.7	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	36	3.4	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	21	2.0	0	0.0	0.0	0	0.0	0.0
Long range (|i-j|≥5)	386	36.7	4	1.0	0.4	1	0.3	0.1
Backbone-Backbone	84	8.0	3	3.6	0.3	1	1.2	0.1
Backbone-Sidechain	187	17.8	1	0.5	0.1	0	0.0	0.0
Sidechain-Sidechain	115	10.9	0	0.0	0.0	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	1051	100.0	5	0.5	0.5	2	0.2	0.2
Backbone-Backbone	213	20.3	3	1.4	0.3	1	0.5	0.1
Backbone-Sidechain	632	60.1	2	0.3	0.2	1	0.2	0.1
Sidechain-Sidechain	206	19.6	0	0.0	0.0	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	1	0	2	0	3	0.81	1.24	0.39	0.89
2	0	1	0	1	0	2	1.0	1.24	0.24	1.0
3	0	1	0	2	0	3	0.93	1.6	0.53	0.89
4	0	1	0	2	0	3	0.64	1.0	0.34	0.75
5	0	1	0	1	0	2	1.0	1.27	0.27	1.0
6	0	1	0	2	0	3	0.76	1.28	0.44	0.81
7	0	1	0	1	0	2	1.05	1.48	0.43	1.05
8	0	1	0	3	0	4	0.61	1.23	0.41	0.51
9	0	1	0	2	0	3	0.74	1.28	0.46	0.78
10	0	1	0	2	0	3	0.63	0.99	0.37	0.78
11	0	1	0	3	0	4	0.66	1.45	0.53	0.54

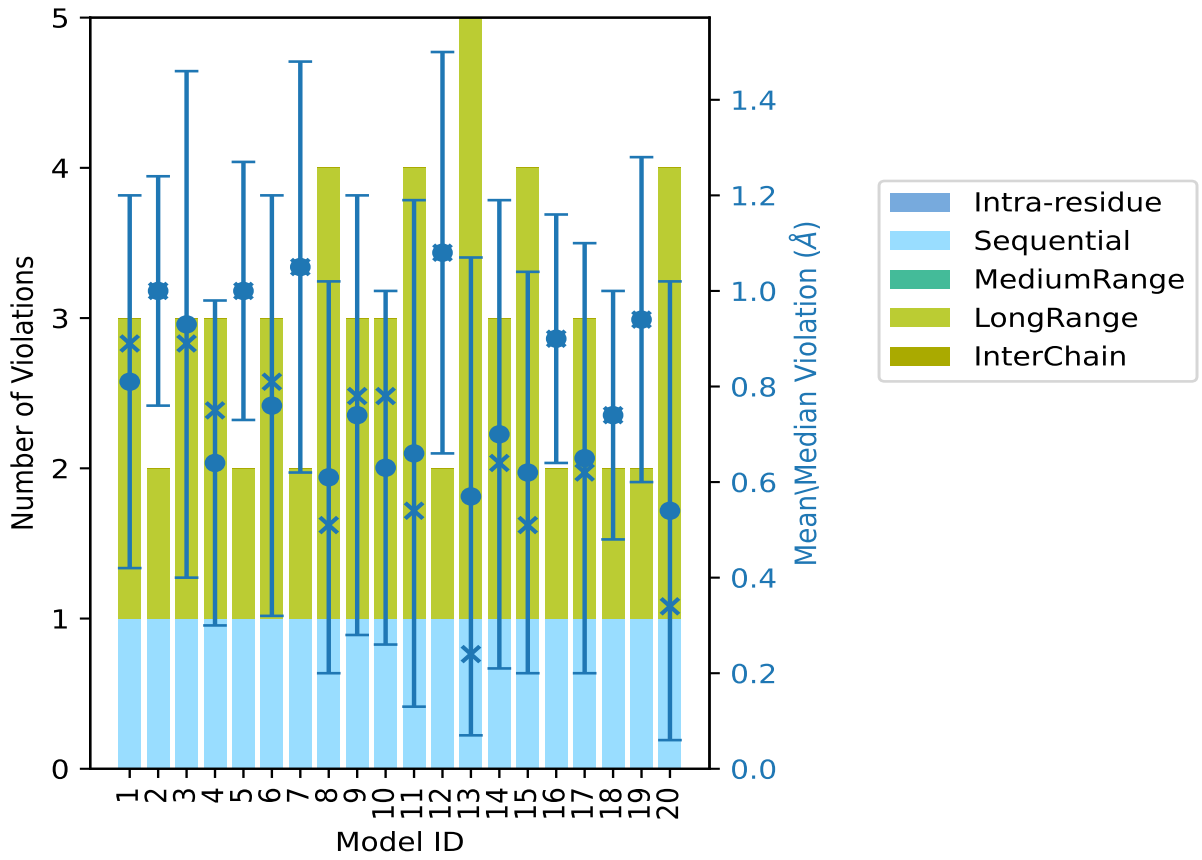
Continued on next page...

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	0	1	0	1	0	2	1.08	1.49	0.42	1.08
13	0	1	0	4	0	5	0.57	1.5	0.5	0.24
14	0	1	0	2	0	3	0.7	1.33	0.49	0.64
15	0	1	0	3	0	4	0.62	1.23	0.42	0.51
16	0	1	0	1	0	2	0.9	1.17	0.26	0.9
17	0	1	0	2	0	3	0.65	1.22	0.45	0.62
18	0	1	0	1	0	2	0.74	1.0	0.26	0.74
19	0	1	0	1	0	2	0.94	1.28	0.34	0.94
20	0	1	0	3	0	4	0.54	1.31	0.48	0.34

[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

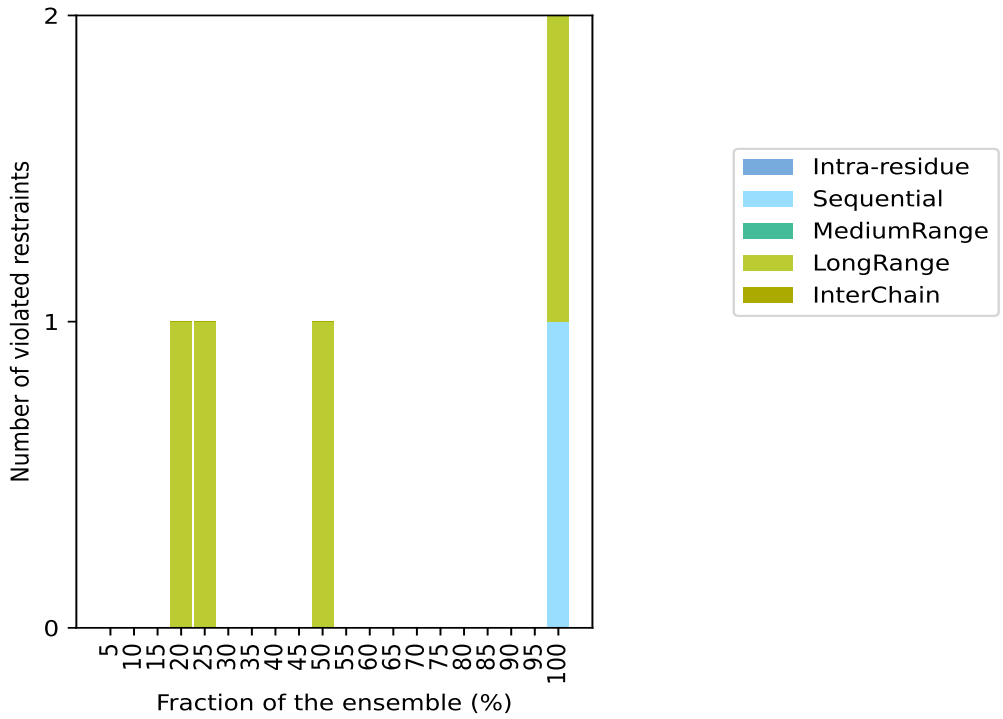
9.3 Distance violation statistics for the ensemble (i)

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, 1046(IR:281, SQ:308, MR:75, LR:382, 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	0	0	0	0	1	5.0
0	0	0	0	0	0	2	10.0
0	0	0	0	0	0	3	15.0
0	0	0	1	0	1	4	20.0
0	0	0	1	0	1	5	25.0
0	0	0	0	0	0	6	30.0
0	0	0	0	0	0	7	35.0
0	0	0	0	0	0	8	40.0
0	0	0	0	0	0	9	45.0
0	0	0	1	0	1	10	50.0
0	0	0	0	0	0	11	55.0
0	0	0	0	0	0	12	60.0
0	0	0	0	0	0	13	65.0
0	0	0	0	0	0	14	70.0
0	0	0	0	0	0	15	75.0
0	0	0	0	0	0	16	80.0
0	0	0	0	0	0	17	85.0
0	0	0	0	0	0	18	90.0
0	0	0	0	0	0	19	95.0
0	1	0	1	0	2	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

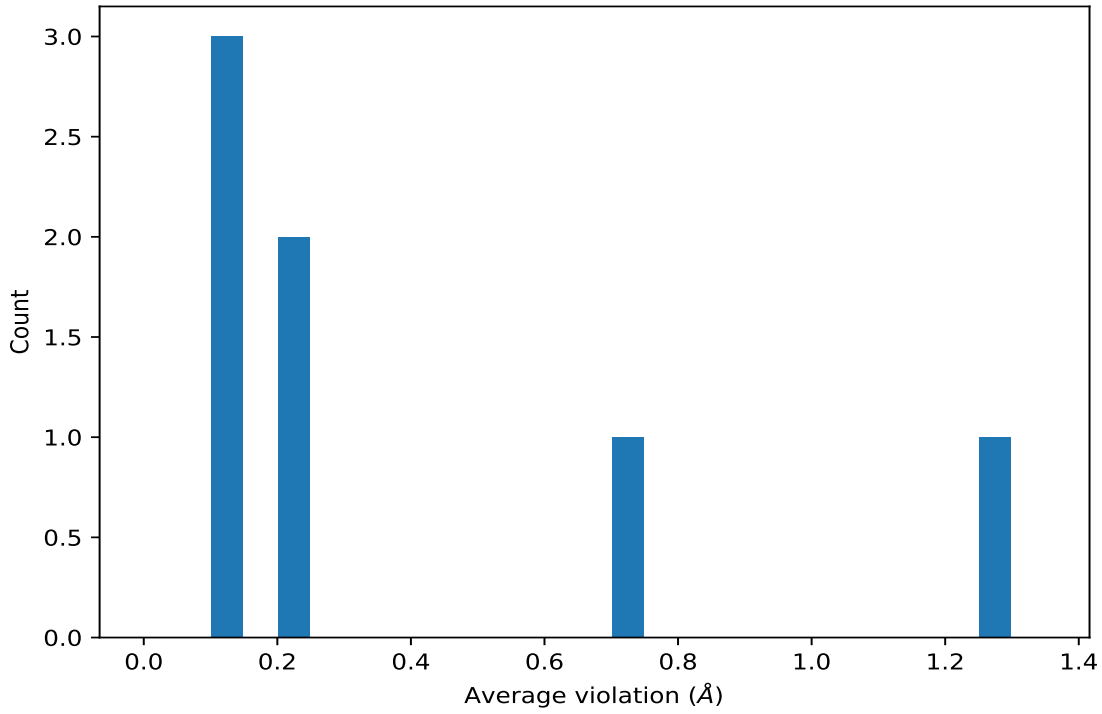
9.3.1 Bar graph : Distance violation statistics for the ensemble



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 violation for each restraint sorted by number of violated models and the mean value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint. Rows with same key represent combinatorial or ambiguous restraints and are counted as a single restraint.

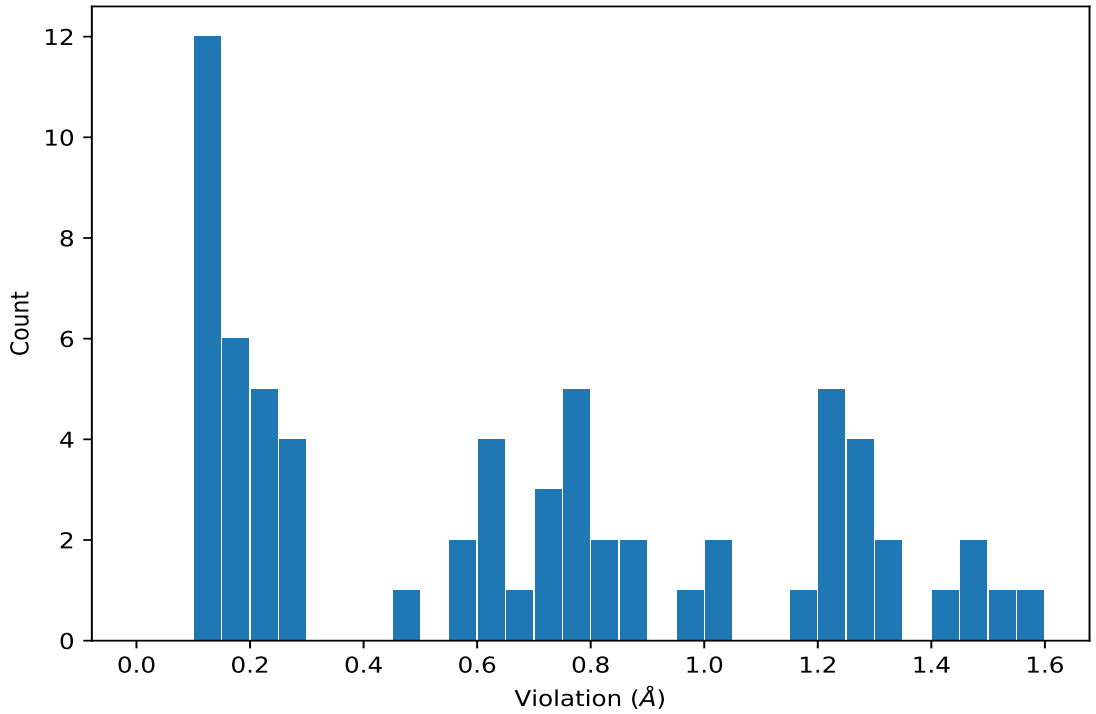
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	20	1.28	0.16	1.27
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	20	0.71	0.11	0.72
(1,433)	1:812:A:LEU:H	1:823:A:GLY:HA3	10	0.21	0.05	0.21
(1,392)	1:810:A:LYS:H	1:823:A:GLY:HA2	5	0.21	0.07	0.24
(1,553)	1:809:A:ILE:HG21	1:823:A:GLY:HA2	4	0.14	0.02	0.14
(1,553)	1:809:A:ILE:HG22	1:823:A:GLY:HA2	4	0.14	0.02	0.14
(1,553)	1:809:A:ILE:HG23	1:823:A:GLY:HA2	4	0.14	0.02	0.14

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

9.5 All violated distance restraints

9.5.1 Histogram : Distribution of distance violations

The following histogram shows the distribution of the absolute value of the violation for all violated restraints in the ensemble.



9.5.2 Table : All distance violations (i)

The following table lists the absolute value of the violation for each restraint in the ensemble sorted by its value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint. Rows with same key represent combinatorial or ambiguous restraints and are counted as a single restraint.

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	3	1.6
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	13	1.5
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	12	1.49
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	7	1.48
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	11	1.45
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	14	1.33
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	20	1.31
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	6	1.28
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	9	1.28
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	19	1.28
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	5	1.27
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	1	1.24
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	2	1.24
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	8	1.23
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	15	1.23
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	17	1.22

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	16	1.17
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	4	1.0
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	18	1.0
(1,470)	1:819:A:GLY:HA2	1:820:A:TRP:HE1	10	0.99
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	1	0.89
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	3	0.89
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	11	0.84
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	6	0.81
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	9	0.78
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	10	0.78
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	15	0.77
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	2	0.76
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	4	0.75
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	5	0.73
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	8	0.72
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	13	0.72
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	12	0.66
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	14	0.64
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	16	0.64
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	7	0.62
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	17	0.62
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	19	0.6
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	20	0.55
(1,43)	1:813:A:ASN:H	1:823:A:GLY:HA3	18	0.47
(1,433)	1:812:A:LEU:H	1:823:A:GLY:HA3	3	0.3
(1,392)	1:810:A:LYS:H	1:823:A:GLY:HA2	8	0.3
(1,433)	1:812:A:LEU:H	1:823:A:GLY:HA3	1	0.29
(1,392)	1:810:A:LYS:H	1:823:A:GLY:HA2	15	0.25
(1,433)	1:812:A:LEU:H	1:823:A:GLY:HA3	13	0.24
(1,392)	1:810:A:LYS:H	1:823:A:GLY:HA2	13	0.24
(1,433)	1:812:A:LEU:H	1:823:A:GLY:HA3	11	0.23
(1,433)	1:812:A:LEU:H	1:823:A:GLY:HA3	15	0.22
(1,433)	1:812:A:LEU:H	1:823:A:GLY:HA3	6	0.2
(1,433)	1:812:A:LEU:H	1:823:A:GLY:HA3	4	0.18
(1,433)	1:812:A:LEU:H	1:823:A:GLY:HA3	8	0.18
(1,553)	1:809:A:ILE:HG21	1:823:A:GLY:HA2	13	0.17
(1,553)	1:809:A:ILE:HG22	1:823:A:GLY:HA2	13	0.17
(1,553)	1:809:A:ILE:HG23	1:823:A:GLY:HA2	13	0.17
(1,433)	1:812:A:LEU:H	1:823:A:GLY:HA3	9	0.16
(1,553)	1:809:A:ILE:HG21	1:823:A:GLY:HA2	14	0.14
(1,553)	1:809:A:ILE:HG22	1:823:A:GLY:HA2	14	0.14
(1,553)	1:809:A:ILE:HG23	1:823:A:GLY:HA2	14	0.14

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,553)	1:809:A:ILE:HG21	1:823:A:GLY:HA2	20	0.14
(1,553)	1:809:A:ILE:HG22	1:823:A:GLY:HA2	20	0.14
(1,553)	1:809:A:ILE:HG23	1:823:A:GLY:HA2	20	0.14
(1,392)	1:810:A:LYS:H	1:823:A:GLY:HA2	20	0.14
(1,433)	1:812:A:LEU:H	1:823:A:GLY:HA3	10	0.13
(1,392)	1:810:A:LYS:H	1:823:A:GLY:HA2	11	0.12
(1,553)	1:809:A:ILE:HG21	1:823:A:GLY:HA2	17	0.11
(1,553)	1:809:A:ILE:HG22	1:823:A:GLY:HA2	17	0.11
(1,553)	1:809:A:ILE:HG23	1:823:A:GLY:HA2	17	0.11

10 Dihedral-angle violation analysis (i)

No dihedral-angle restraints found