Full wwPDB NMR Structure Validation Report (i)

May 28, 2020 – 10:45 pm BST

PDB ID	:	2KS4
Title	:	NMR structure of the sea anemone actinoporin Sticholysin
Authors	:	Castrillo, I.; Santoro, J.; Bruix, M.
Deposited on	:	2009-12-29

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 following versions of software and data (see references (1)) were used in the production of this report:

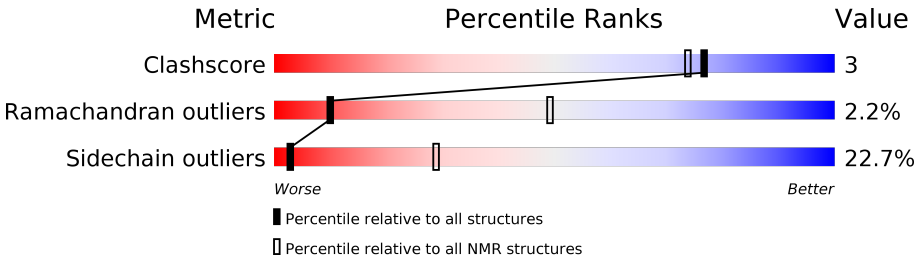
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance

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

The overall completeness of chemical shifts assignment was not calculated.

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	176	73% 18% • 9%

2 Ensemble composition and analysis

This entry contains 20 models. Model 8 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:3-A:76, A:83-A:105, A:113-A:176 (161)	0.39	8

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 5 clusters and 2 single-model clusters were found.

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

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 2705 atoms, of which 1338 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Sticholysin-1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	176	2705	873	1338	229	258	7	0

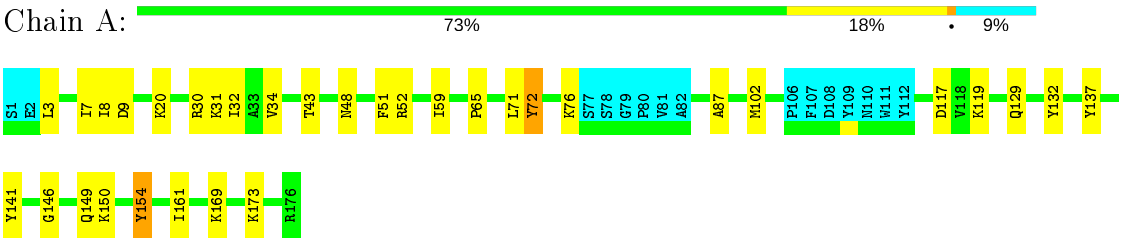
4 Residue-property plots

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA 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: Sticholysin-1

Chain A:



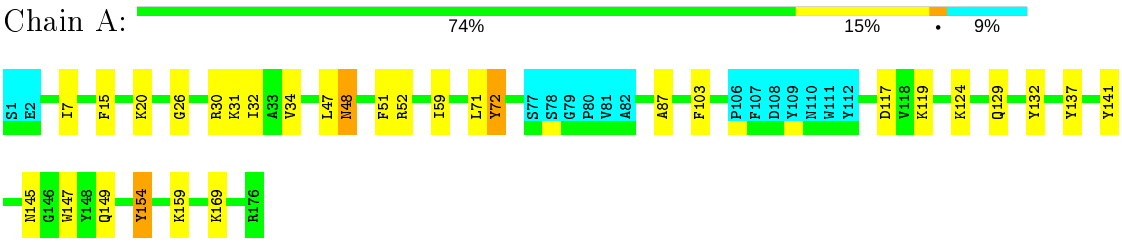
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

- Molecule 1: Sticholysin-1

Chain A:



4.2.2 Score per residue for model 2

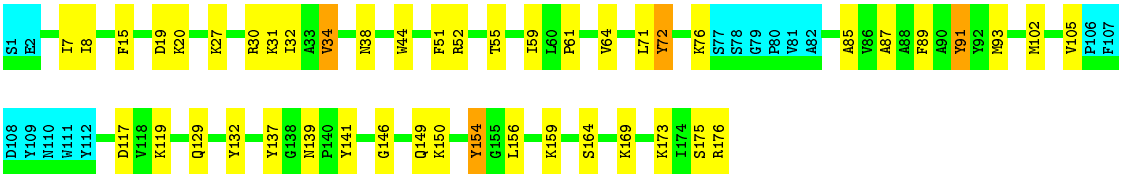
- Molecule 1: Sticholysin-1

Chain A:

4.2.3 Score per residue for model 3

- Molecule 1: Sticholysin-1

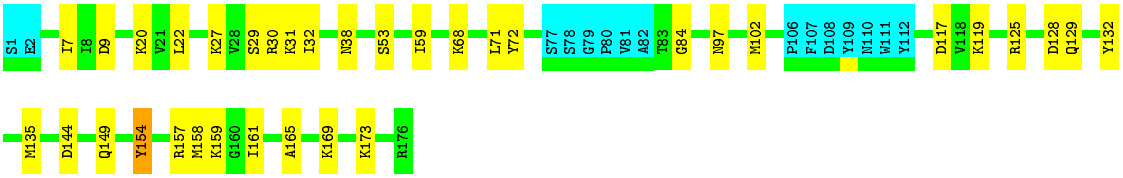
Chain A:



4.2.4 Score per residue for model 4

- Molecule 1: Sticholysin-1

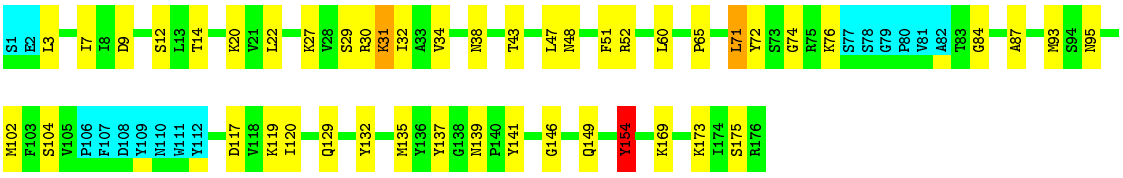
Chain A:



4.2.5 Score per residue for model 5

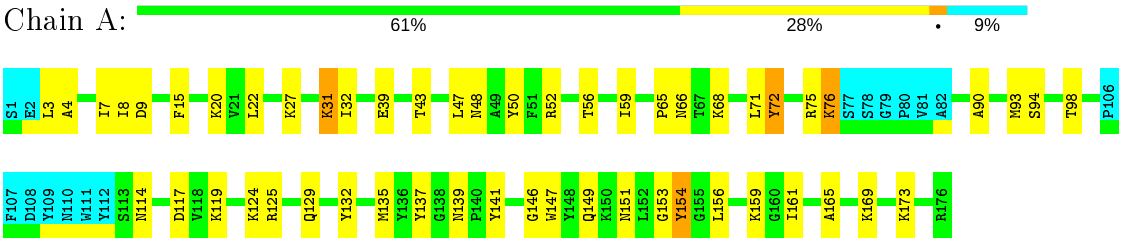
- Molecule 1: Sticholysin-1

Chain A:



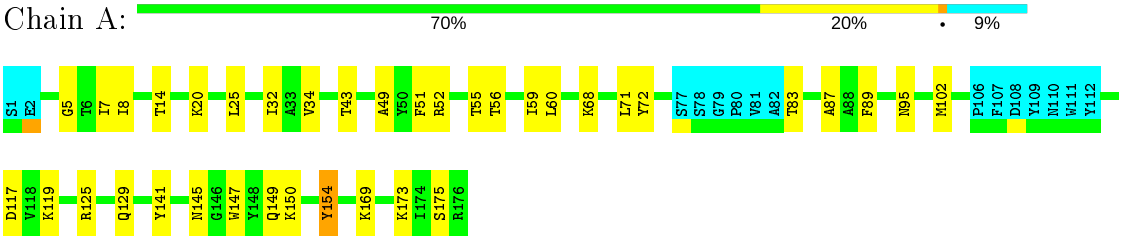
4.2.6 Score per residue for model 6

- Molecule 1: Sticholysin-1



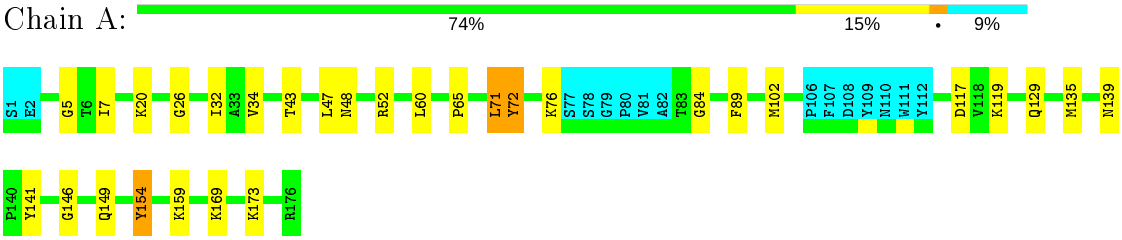
4.2.7 Score per residue for model 7

- Molecule 1: Sticholysin-1



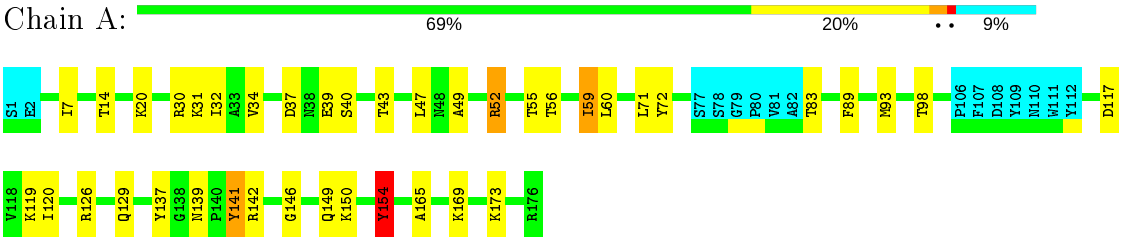
4.2.8 Score per residue for model 8 (medoid)

- Molecule 1: Sticholysin-1



4.2.9 Score per residue for model 9

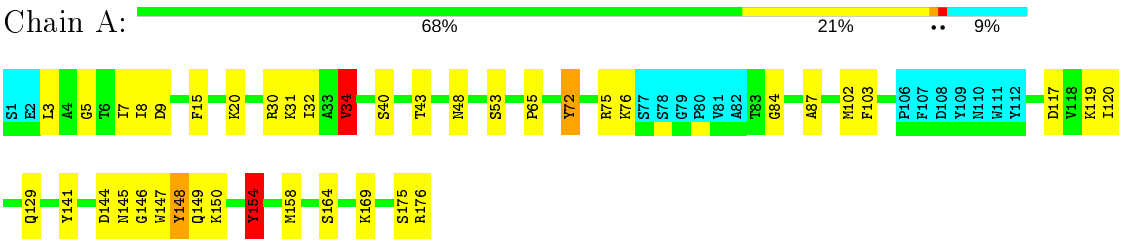
- Molecule 1: Sticholysin-1



4.2.10 Score per residue for model 10

- Molecule 1: Sticholysin-1

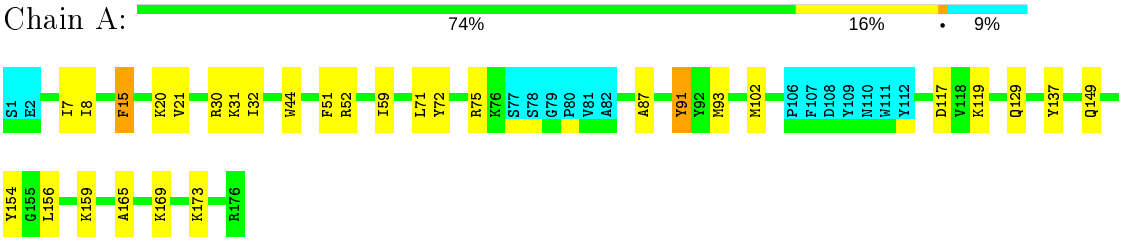
Chain A:



4.2.11 Score per residue for model 11

- Molecule 1: Sticholysin-1

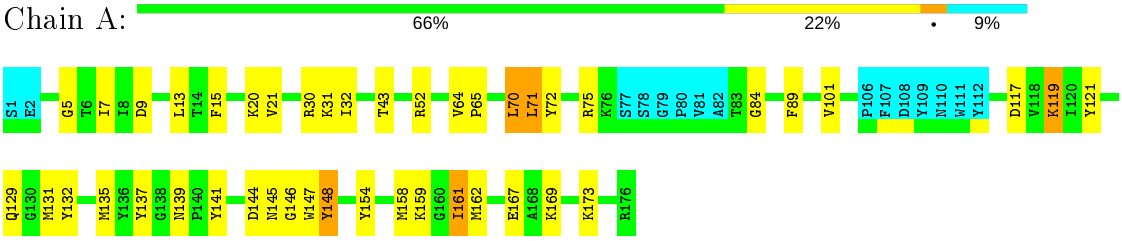
Chain A:



4.2.12 Score per residue for model 12

- Molecule 1: Sticholysin-1

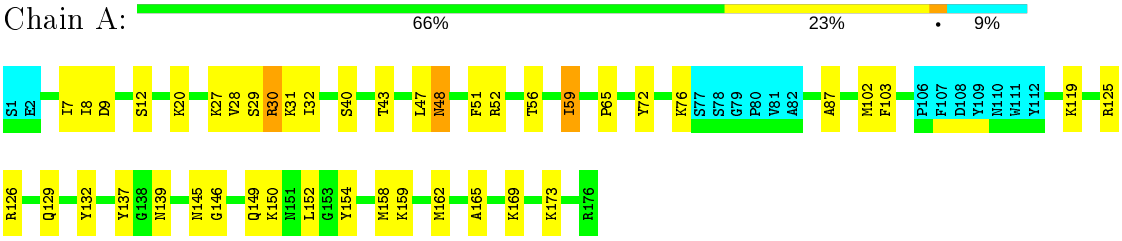
Chain A:



4.2.13 Score per residue for model 13

- Molecule 1: Sticholysin-1

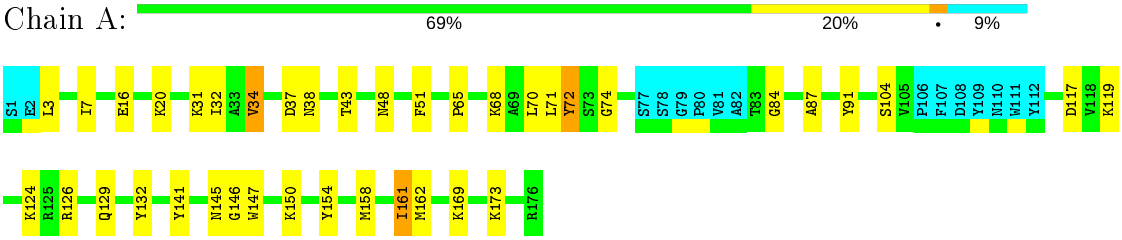
Chain A:



4.2.14 Score per residue for model 14

- Molecule 1: Sticholysin-1

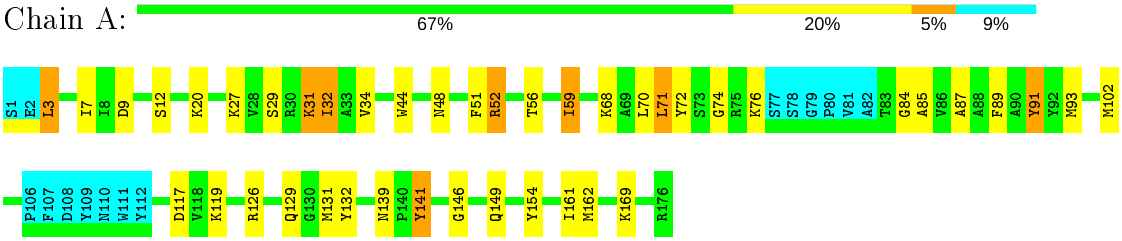
Chain A:



4.2.15 Score per residue for model 15

- Molecule 1: Sticholysin-1

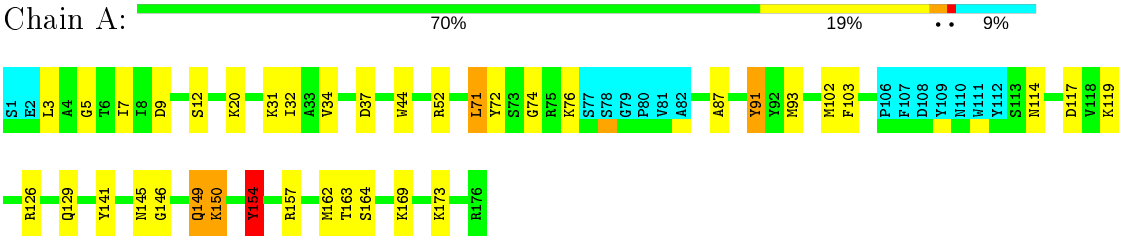
Chain A:



4.2.16 Score per residue for model 16

- Molecule 1: Sticholysin-1

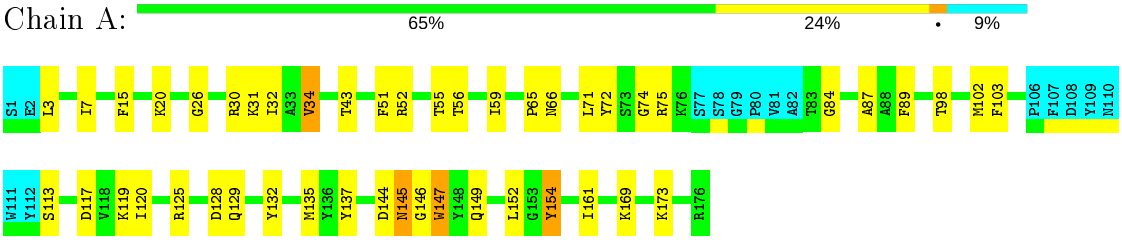
Chain A:



4.2.17 Score per residue for model 17

- Molecule 1: Sticholysin-1

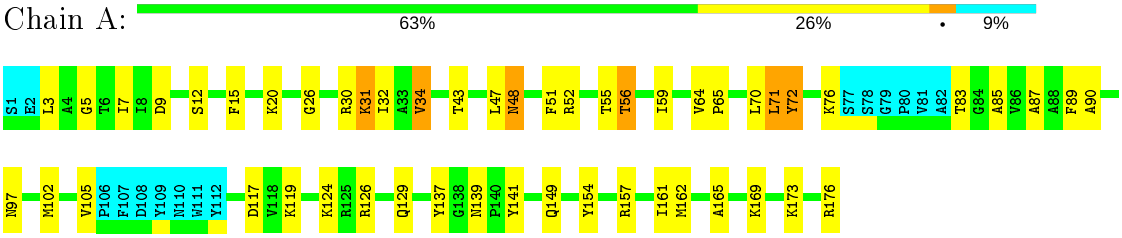
Chain A:



4.2.18 Score per residue for model 18

- Molecule 1: Sticholysin-1

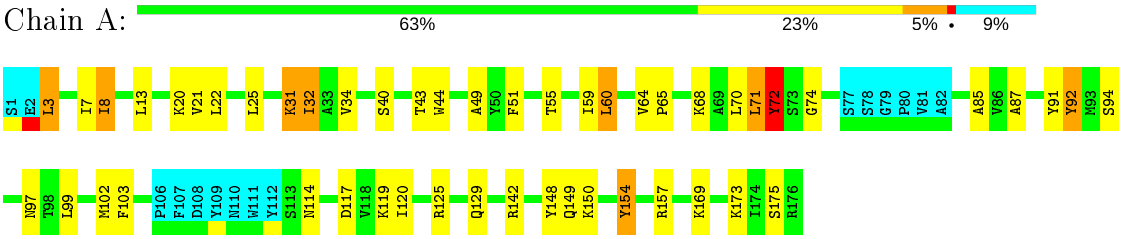
Chain A:



4.2.19 Score per residue for model 19

- Molecule 1: Sticholysin-1

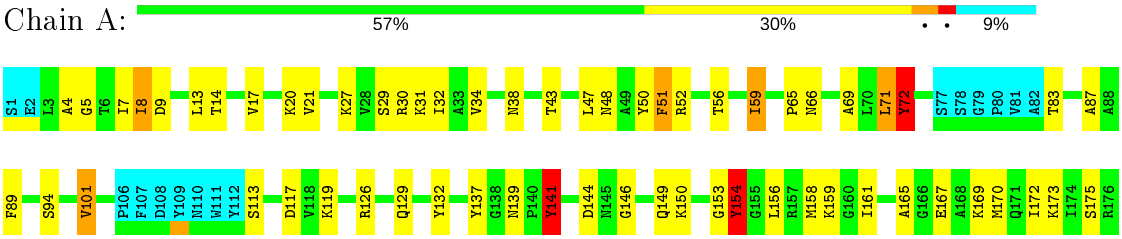
Chain A:



4.2.20 Score per residue for model 20

- Molecule 1: Sticholysin-1

Chain A:



5 Refinement protocol and experimental data overview

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

Of the 20 calculated structures, 20 were deposited, based on the following criterion: *all calculated structures submitted*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
CYANA	structure solution	
AMBER	refinement	

No chemical shift data was provided. No validations of the models with respect to experimental NMR restraints is performed at this time.

COVALENT-GEOMETRY INFOmissingINFO

5.1 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	1245	1237	1237	8±4
All	All	24900	24740	24740	150

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 3.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:34:VAL:HG22	1:A:72:TYR:HB2	0.68	1.65	18	14
1:A:87:ALA:HB3	1:A:103:PHE:CZ	0.64	2.28	13	6
1:A:3:LEU:HD12	1:A:74:GLY:H	0.63	1.54	19	6
1:A:43:THR:HG22	1:A:65:PRO:HA	0.58	1.75	14	11
1:A:72:TYR:CE2	1:A:89:PHE:CE2	0.56	2.94	20	4
1:A:147:TRP:H	1:A:161:ILE:HD13	0.56	1.60	17	2
1:A:32:ILE:HG13	1:A:85:ALA:HB1	0.56	1.78	15	2
1:A:152:LEU:HD21	1:A:158:MET:HB2	0.56	1.78	13	1
1:A:3:LEU:HD12	1:A:74:GLY:N	0.53	2.18	19	3
1:A:34:VAL:HG13	1:A:72:TYR:CB	0.53	2.34	10	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:47:LEU:HD11	1:A:90:ALA:HB3	0.52	1.81	18	1
1:A:51:PHE:CE1	1:A:87:ALA:HB1	0.52	2.40	11	4
1:A:47:LEU:HD12	1:A:48:ASN:N	0.51	2.20	20	6
1:A:72:TYR:CE2	1:A:89:PHE:CD2	0.51	2.99	17	7
1:A:51:PHE:CD1	1:A:87:ALA:HB1	0.50	2.41	20	11
1:A:3:LEU:HD11	1:A:31:LYS:HB2	0.50	1.84	18	1
1:A:3:LEU:HD13	1:A:31:LYS:HB2	0.49	1.84	5	3
1:A:119:LYS:HE3	1:A:121:TYR:CE1	0.49	2.42	12	1
1:A:120:ILE:HD11	1:A:152:LEU:HD13	0.49	1.84	17	1
1:A:89:PHE:CE2	1:A:101:VAL:HG11	0.49	2.43	20	1
1:A:147:TRP:H	1:A:161:ILE:HD12	0.49	1.68	14	1
1:A:89:PHE:CZ	1:A:101:VAL:HG11	0.48	2.44	12	2
1:A:49:ALA:HB2	1:A:60:LEU:HD22	0.48	1.85	19	1
1:A:56:THR:HG21	1:A:59:ILE:HG23	0.48	1.86	20	6
1:A:119:LYS:CE	1:A:121:TYR:CE1	0.47	2.97	12	1
1:A:34:VAL:O	1:A:71:LEU:HD13	0.47	2.09	8	6
1:A:64:VAL:HG21	1:A:70:LEU:HD23	0.47	1.86	12	1
1:A:64:VAL:HG21	1:A:70:LEU:HD13	0.47	1.86	19	2
1:A:98:THR:HG21	1:A:125:ARG:O	0.46	2.10	17	1
1:A:22:LEU:HD23	1:A:31:LYS:HD2	0.46	1.87	5	1
1:A:44:TRP:CD1	1:A:91:TYR:CZ	0.46	3.04	16	3
1:A:49:ALA:HB2	1:A:60:LEU:HD12	0.46	1.86	7	2
1:A:13:LEU:HD11	1:A:71:LEU:CD2	0.46	2.41	12	2
1:A:13:LEU:HD23	1:A:69:ALA:CB	0.46	2.41	20	1
1:A:4:ALA:HB1	1:A:71:LEU:HB3	0.45	1.87	20	1
1:A:44:TRP:HA	1:A:91:TYR:CE1	0.45	2.47	3	4
1:A:44:TRP:CZ2	1:A:99:LEU:CD2	0.43	3.01	19	1
1:A:90:ALA:HB1	1:A:98:THR:CG2	0.43	2.44	6	1
1:A:152:LEU:HD21	1:A:158:MET:CB	0.43	2.42	13	1
1:A:149:GLN:NE2	1:A:150:LYS:H	0.43	2.11	16	1
1:A:61:PRO:HG2	1:A:64:VAL:HG23	0.43	1.89	3	1
1:A:158:MET:SD	1:A:172:ILE:HG21	0.43	2.53	20	1
1:A:161:ILE:HD13	1:A:162:MET:N	0.42	2.29	12	2
1:A:85:ALA:HB3	1:A:105:VAL:HG22	0.42	1.90	3	2
1:A:3:LEU:HD11	1:A:31:LYS:HG3	0.42	1.91	6	1
1:A:28:VAL:HG23	1:A:30:ARG:O	0.42	2.14	13	1
1:A:72:TYR:CZ	1:A:89:PHE:CZ	0.42	3.08	8	1
1:A:51:PHE:CZ	1:A:59:ILE:HG21	0.42	2.50	15	1
1:A:120:ILE:HB	1:A:154:TYR:CE2	0.42	2.49	10	2
1:A:47:LEU:HD21	1:A:98:THR:HG23	0.42	1.92	9	1
1:A:56:THR:CG2	1:A:59:ILE:HD12	0.42	2.45	15	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:87:ALA:HB3	1:A:103:PHE:CE2	0.42	2.49	16	1
1:A:59:ILE:HB	1:A:72:TYR:CE2	0.42	2.49	20	2
1:A:170:MET:SD	1:A:172:ILE:CD1	0.41	3.08	20	1
1:A:154:TYR:N	1:A:154:TYR:CD1	0.41	2.87	16	1
1:A:34:VAL:HG13	1:A:72:TYR:HB3	0.41	1.91	10	1
1:A:15:PHE:CD1	1:A:15:PHE:C	0.41	2.93	11	1
1:A:76:LYS:HE2	1:A:83:THR:O	0.41	2.15	18	1
1:A:120:ILE:HB	1:A:154:TYR:CZ	0.41	2.51	19	1
1:A:17:VAL:O	1:A:21:VAL:HG23	0.41	2.17	20	1
1:A:120:ILE:HB	1:A:154:TYR:CE1	0.40	2.51	5	1

5.2 Torsion angles

5.2.1 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	160/176 (91%)	135±3 (85±2%)	21±2 (13±1%)	4±1 (2±1%)	10	49
All	All	3200/3520 (91%)	2709 (85%)	421 (13%)	70 (2%)	10	49

All 15 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	146	GLY	13
1	A	84	GLY	9
1	A	5	GLY	7
1	A	165	ALA	7
1	A	55	THR	7
1	A	147	TRP	5
1	A	145	ASN	5
1	A	26	GLY	4
1	A	154	TYR	3
1	A	153	GLY	3
1	A	8	ILE	3
1	A	4	ALA	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	117	ASP	1
1	A	164	SER	1
1	A	27	LYS	1

5.2.2 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	127/140 (91%)	98±4 (77±4%)	29±4 (23±4%)	3	29
All	All	2540/2800 (91%)	1964 (77%)	576 (23%)	3	29

All 84 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	169	LYS	20
1	A	129	GLN	20
1	A	32	ILE	20
1	A	7	ILE	20
1	A	119	LYS	20
1	A	154	TYR	20
1	A	20	LYS	20
1	A	31	LYS	18
1	A	149	GLN	18
1	A	71	LEU	18
1	A	117	ASP	17
1	A	173	LYS	17
1	A	52	ARG	16
1	A	141	TYR	15
1	A	102	MET	13
1	A	137	TYR	12
1	A	30	ARG	11
1	A	9	ASP	11
1	A	59	ILE	11
1	A	132	TYR	11
1	A	139	ASN	10
1	A	150	LYS	9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	159	LYS	9
1	A	48	ASN	8
1	A	76	LYS	8
1	A	15	PHE	8
1	A	93	MET	7
1	A	8	ILE	7
1	A	161	ILE	7
1	A	126	ARG	6
1	A	68	LYS	6
1	A	125	ARG	6
1	A	27	LYS	6
1	A	175	SER	6
1	A	38	ASN	6
1	A	135	MET	6
1	A	162	MET	5
1	A	156	LEU	5
1	A	12	SER	5
1	A	144	ASP	5
1	A	29	SER	5
1	A	91	TYR	5
1	A	40	SER	4
1	A	124	LYS	4
1	A	145	ASN	4
1	A	72	TYR	4
1	A	158	MET	4
1	A	14	THR	4
1	A	75	ARG	4
1	A	94	SER	3
1	A	39	GLU	3
1	A	114	ASN	3
1	A	66	ASN	3
1	A	60	LEU	3
1	A	70	LEU	3
1	A	21	VAL	3
1	A	97	ASN	3
1	A	83	THR	3
1	A	22	LEU	3
1	A	37	ASP	3
1	A	3	LEU	3
1	A	95	ASN	2
1	A	47	LEU	2
1	A	104	SER	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	113	SER	2
1	A	164	SER	2
1	A	131	MET	2
1	A	167	GLU	2
1	A	92	TYR	2
1	A	148	TYR	2
1	A	157	ARG	2
1	A	34	VAL	2
1	A	56	THR	2
1	A	142	ARG	2
1	A	53	SER	2
1	A	128	ASP	2
1	A	43	THR	2
1	A	176	ARG	1
1	A	147	TRP	1
1	A	151	ASN	1
1	A	163	THR	1
1	A	16	GLU	1
1	A	19	ASP	1
1	A	101	VAL	1

5.2.3 RNA

There are no RNA molecules in this entry.

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

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

5.4 Carbohydrates

There are no carbohydrates in this entry.

5.5 Ligand geometry

There are no ligands in this entry.

5.6 Other polymers

There are no such molecules in this entry.

5.7 Polymer linkage issues

There are no chain breaks in this entry.

6 Chemical shift validation

No chemical shift data were provided