Full wwPDB NMR Structure Validation Report (i)

Oct 10, 2021 – 08:18 PM EDT

PDB ID	:	2K4W
Title	:	The Solution Structure of the Monomeric Copper, Zinc Superoxide Dismutase from Salmonella enterica
Authors	:	Mori, M.; Jimenez, B.; Piccioli, M.; Battistoni, A.; Sette, M.; Structural Proteomics in Europe (SPINE)
Deposited on	:	2008-06-20

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

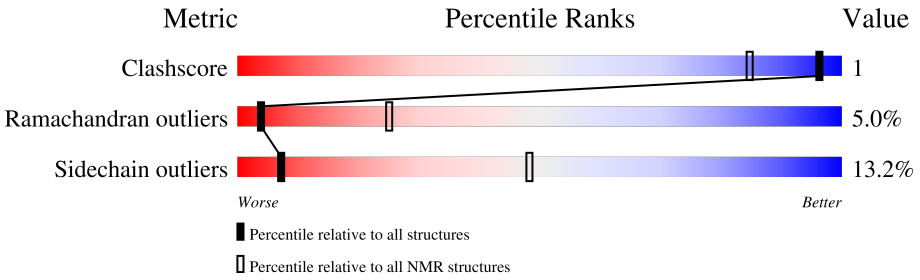
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.23.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 is 97%.

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	154	89% 11%

2 Ensemble composition and analysis

This entry contains 30 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:1-A:154 (154)	0.69	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 10 single-model clusters were found.

Cluster number	Models
1	1, 4, 6, 8, 9, 10, 14, 16, 18, 21, 26
2	19, 25, 30
3	17, 20
4	15, 28
5	23, 29
Single-model clusters	2; 3; 5; 7; 11; 12; 13; 22; 24; 27

3 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2192 atoms, of which 1088 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Superoxide dismutase [Cu-Zn].

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	154	2190	681	1088	199	217	5	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	6	GLY	GLU	engineered mutation	UNP Q704S6

- Molecule 2 is COPPER (I) ION (three-letter code: CU1) (formula: Cu).

Mol	Chain	Residues	Atoms	
			Total	Cu
2	A	1	1	1

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	
			Total	Zn
3	A	1	1	1

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: Superoxide dismutase [Cu-Zn]

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

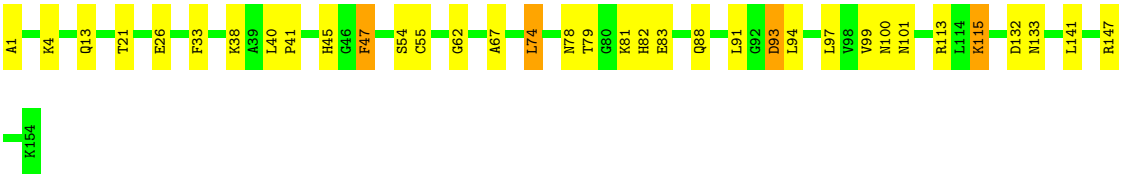
• Molecule 1: Superoxide dismutase [Cu-Zn]

Chain A: 

4.2.2 Score per residue for model 2

• Molecule 1: Superoxide dismutase [Cu-Zn]

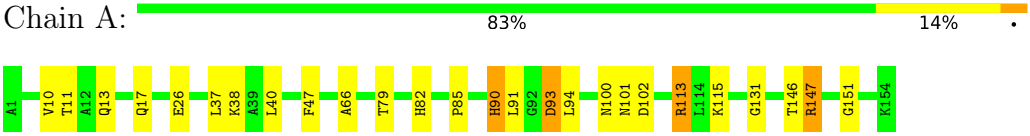
Chain A: 



4.2.3 Score per residue for model 3

• Molecule 1: Superoxide dismutase [Cu-Zn]

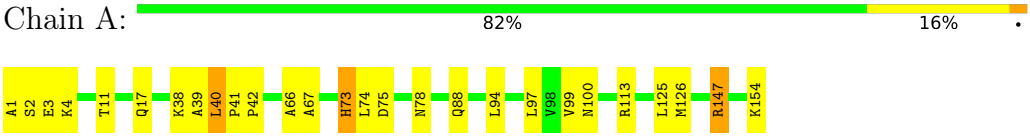
Chain A:



4.2.4 Score per residue for model 4

• Molecule 1: Superoxide dismutase [Cu-Zn]

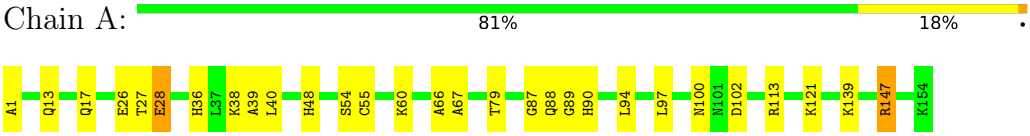
Chain A:



4.2.5 Score per residue for model 5

• Molecule 1: Superoxide dismutase [Cu-Zn]

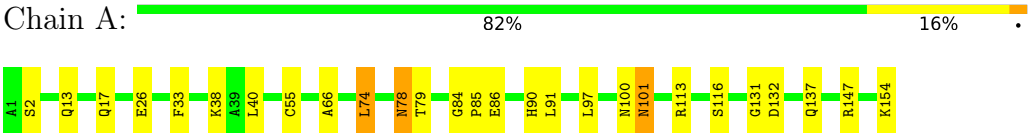
Chain A:



4.2.6 Score per residue for model 6

• Molecule 1: Superoxide dismutase [Cu-Zn]

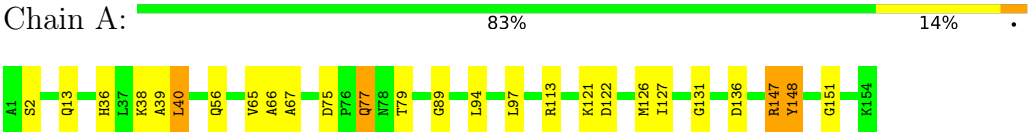
Chain A:



4.2.7 Score per residue for model 7

• Molecule 1: Superoxide dismutase [Cu-Zn]

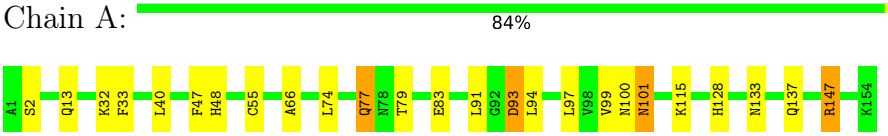
Chain A:



4.2.8 Score per residue for model 8 (medoid)

• Molecule 1: Superoxide dismutase [Cu-Zn]

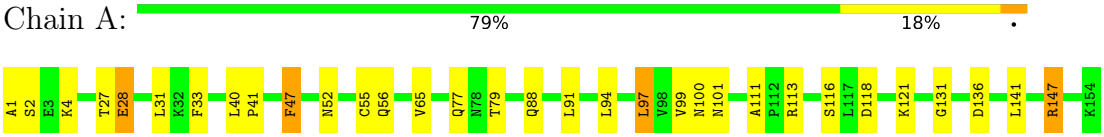
Chain A:



4.2.9 Score per residue for model 9

• Molecule 1: Superoxide dismutase [Cu-Zn]

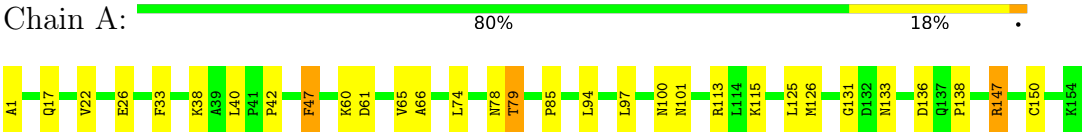
Chain A:



4.2.10 Score per residue for model 10

• Molecule 1: Superoxide dismutase [Cu-Zn]

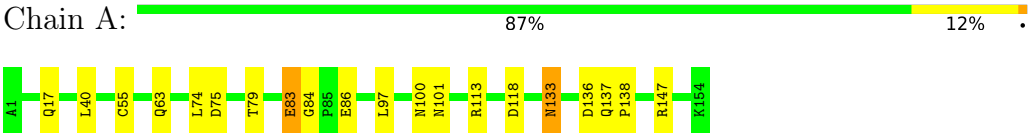
Chain A:



4.2.11 Score per residue for model 11

• Molecule 1: Superoxide dismutase [Cu-Zn]

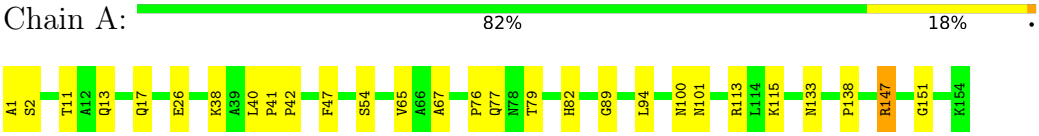
Chain A:



4.2.12 Score per residue for model 12

• Molecule 1: Superoxide dismutase [Cu-Zn]

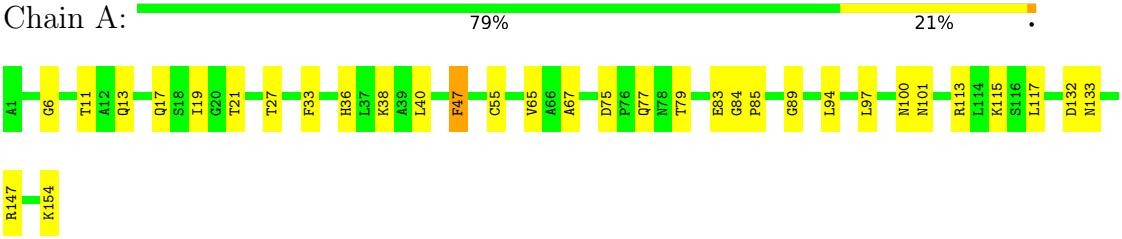
Chain A:



4.2.13 Score per residue for model 13

• Molecule 1: Superoxide dismutase [Cu-Zn]

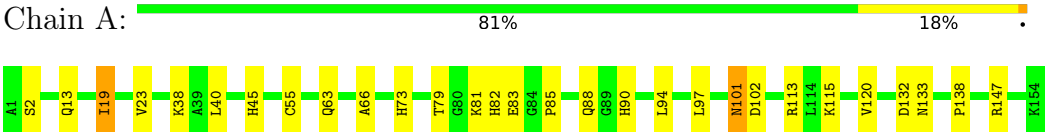
Chain A:



4.2.14 Score per residue for model 14

• Molecule 1: Superoxide dismutase [Cu-Zn]

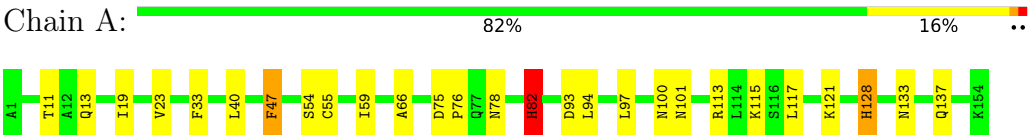
Chain A:



4.2.15 Score per residue for model 15

• Molecule 1: Superoxide dismutase [Cu-Zn]

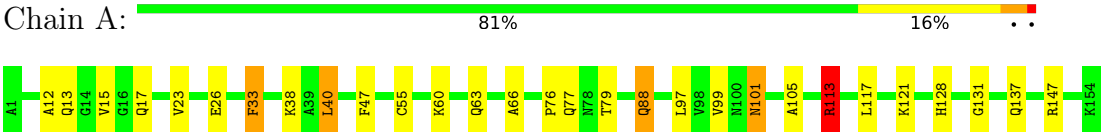
Chain A:



4.2.16 Score per residue for model 16

• Molecule 1: Superoxide dismutase [Cu-Zn]

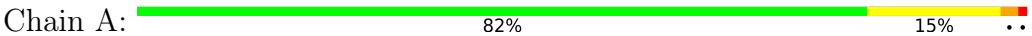
Chain A:

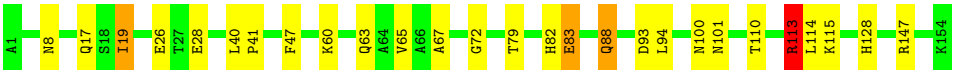


4.2.17 Score per residue for model 17

• Molecule 1: Superoxide dismutase [Cu-Zn]

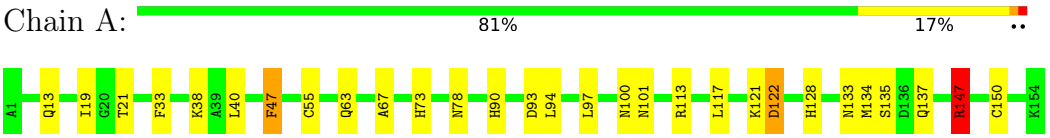
Chain A:





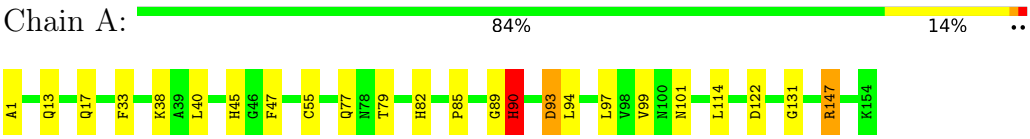
4.2.18 Score per residue for model 18

• Molecule 1: Superoxide dismutase [Cu-Zn]



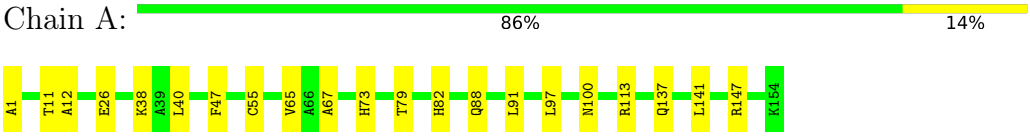
4.2.19 Score per residue for model 19

• Molecule 1: Superoxide dismutase [Cu-Zn]



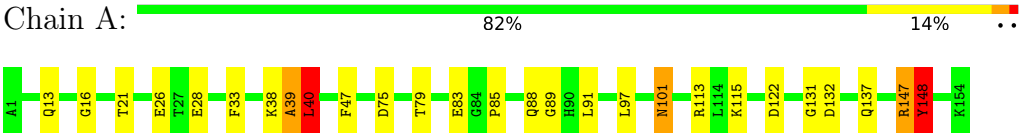
4.2.20 Score per residue for model 20

• Molecule 1: Superoxide dismutase [Cu-Zn]



4.2.21 Score per residue for model 21

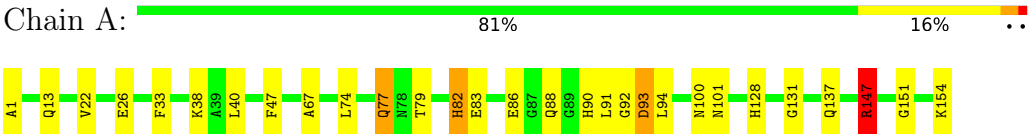
• Molecule 1: Superoxide dismutase [Cu-Zn]



4.2.22 Score per residue for model 22

• Molecule 1: Superoxide dismutase [Cu-Zn]

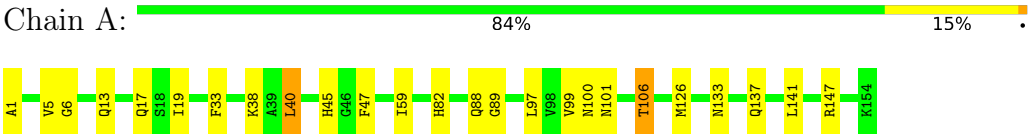
Chain A:



4.2.23 Score per residue for model 23

• Molecule 1: Superoxide dismutase [Cu-Zn]

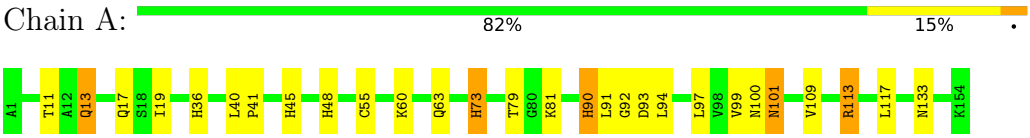
Chain A:



4.2.24 Score per residue for model 24

• Molecule 1: Superoxide dismutase [Cu-Zn]

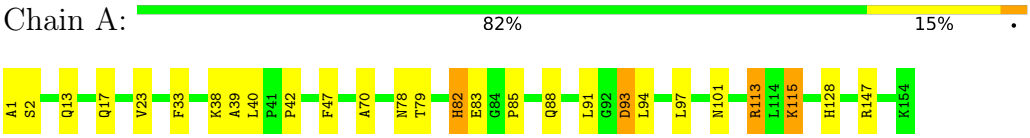
Chain A:



4.2.25 Score per residue for model 25

• Molecule 1: Superoxide dismutase [Cu-Zn]

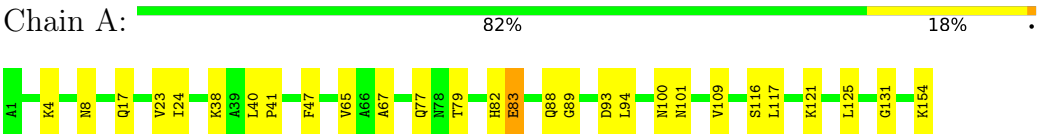
Chain A:



4.2.26 Score per residue for model 26

• Molecule 1: Superoxide dismutase [Cu-Zn]

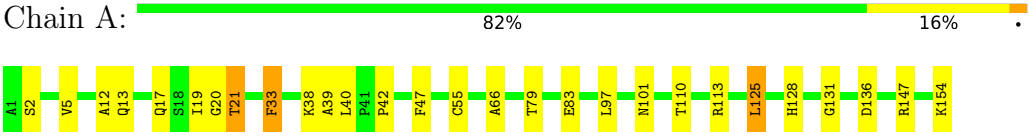
Chain A:



4.2.27 Score per residue for model 27

• Molecule 1: Superoxide dismutase [Cu-Zn]

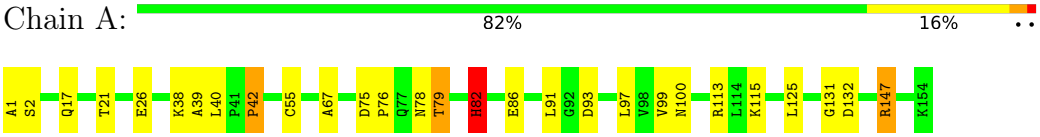
Chain A:



4.2.28 Score per residue for model 28

• Molecule 1: Superoxide dismutase [Cu-Zn]

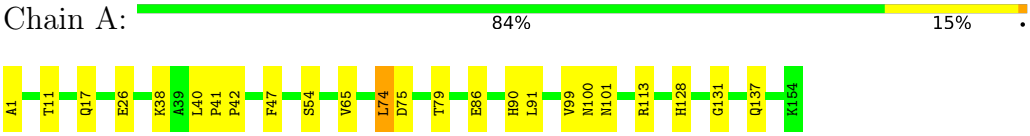
Chain A:



4.2.29 Score per residue for model 29

• Molecule 1: Superoxide dismutase [Cu-Zn]

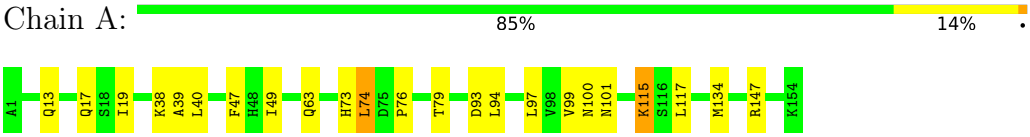
Chain A:



4.2.30 Score per residue for model 30

• Molecule 1: Superoxide dismutase [Cu-Zn]

Chain A:



5 Refinement protocol and experimental data overview

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

Of the 200 calculated structures, 30 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
Amber	refinement	8.0

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	1891
Number of shifts mapped to atoms	1891
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	97%

6 Model quality

6.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CU1, ZN

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	0.62±0.01	0±0/1123 (0.0± 0.0%)	1.08±0.03	2±1/1522 (0.2± 0.1%)
All	All	0.62	0/33690 (0.0%)	1.08	69/45660 (0.2%)

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	3.1±1.3
All	All	0	94

There are no bond-length outliers.

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
1	A	113	ARG	NE-CZ-NH1	9.66	125.13	120.30	28	24
1	A	147	ARG	NE-CZ-NH1	9.43	125.02	120.30	3	20
1	A	148	TYR	CB-CG-CD1	-7.31	116.61	121.00	21	1
1	A	26	GLU	N-CA-CB	-6.88	98.21	110.60	17	1
1	A	113	ARG	NE-CZ-NH2	-5.88	117.36	120.30	17	3
1	A	40	LEU	CA-CB-CG	5.76	128.54	115.30	21	1
1	A	117	LEU	CB-CG-CD2	5.59	120.50	111.00	18	1
1	A	28	GLU	N-CA-CB	-5.59	100.55	110.60	5	2
1	A	23	VAL	CA-CB-CG1	5.47	119.11	110.90	26	5
1	A	113	ARG	CD-NE-CZ	5.36	131.10	123.60	13	1
1	A	13	GLN	CA-CB-CG	5.34	125.15	113.40	24	1
1	A	73	HIS	CA-CB-CG	5.31	122.63	113.60	20	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	47	PHE	CB-CG-CD2	-5.25	117.12	120.80	10	3
1	A	11	THR	C-N-CA	5.17	134.64	121.70	20	1
1	A	90	HIS	CA-CB-CG	5.15	122.35	113.60	19	1
1	A	40	LEU	CB-CG-CD2	5.14	119.74	111.00	4	1
1	A	40	LEU	CB-CG-CD1	5.13	119.72	111.00	21	1
1	A	66	ALA	C-N-CA	5.01	134.23	121.70	7	1

There are no chirality outliers.

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

Mol	Chain	Res	Type	Group	Models (Total)
1	A	1	ALA	Peptide	13
1	A	131	GLY	Peptide	11
1	A	93	ASP	Peptide	11
1	A	41	PRO	Peptide	9
1	A	21	THR	Peptide	5
1	A	45	HIS	Sidechain	5
1	A	74	LEU	Peptide	5
1	A	113	ARG	Sidechain,Peptide	4
1	A	147	ARG	Sidechain	4
1	A	138	PRO	Peptide	3
1	A	115	LYS	Peptide	2
1	A	27	THR	Peptide	2
1	A	77	GLN	Peptide	2
1	A	19	ILE	Peptide	2
1	A	135	SER	Peptide	1
1	A	10	VAL	Peptide	1
1	A	87	GLY	Peptide	1
1	A	52	ASN	Peptide	1
1	A	133	ASN	Peptide	1
1	A	59	ILE	Peptide	1
1	A	128	HIS	Sidechain	1
1	A	72	GLY	Peptide	1
1	A	148	TYR	Sidechain	1
1	A	5	VAL	Peptide	1
1	A	82	HIS	Peptide	1
1	A	90	HIS	Sidechain	1
1	A	42	PRO	Peptide	1
1	A	20	GLY	Peptide	1
1	A	79	THR	Peptide	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	134	MET	Peptide	1

6.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	1102	1088	1087	2±1
All	All	33120	32640	32610	51

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:40:LEU:HD12	1:A:99:VAL:HG11	0.59	1.73	16	2
1:A:47:PHE:CZ	1:A:94:LEU:HD22	0.55	2.37	9	2
1:A:47:PHE:CE1	1:A:125:LEU:HD12	0.54	2.38	27	1
1:A:99:VAL:HG12	1:A:105:ALA:HB2	0.52	1.81	16	1
1:A:40:LEU:HD12	1:A:99:VAL:CG1	0.52	2.34	16	2
1:A:24:ILE:HD11	1:A:125:LEU:HD21	0.50	1.84	1	2
1:A:40:LEU:HA	1:A:148:TYR:CE2	0.50	2.41	21	1
1:A:73:HIS:H	1:A:73:HIS:CD2	0.49	2.24	4	1
1:A:90:HIS:CD2	1:A:92:GLY:H	0.49	2.26	24	1
1:A:83:GLU:CD	1:A:84:GLY:H	0.48	2.11	11	2
1:A:33:PHE:CZ	1:A:47:PHE:CE1	0.48	3.02	1	10
1:A:48:HIS:HA	1:A:94:LEU:HD12	0.47	1.85	5	1
1:A:94:LEU:HD23	1:A:111:ALA:CB	0.47	2.39	9	1
1:A:82:HIS:CE1	1:A:93:ASP:CG	0.46	2.89	2	2
1:A:38:LYS:HE3	1:A:38:LYS:H	0.45	1.71	1	1
1:A:33:PHE:CD2	1:A:47:PHE:CE2	0.45	3.04	9	1
1:A:82:HIS:CE1	1:A:93:ASP:OD1	0.45	2.69	15	3
1:A:22:VAL:HG11	1:A:33:PHE:CZ	0.45	2.47	10	2
1:A:48:HIS:CA	1:A:94:LEU:HD12	0.44	2.41	5	1
1:A:38:LYS:HE3	1:A:38:LYS:N	0.44	2.27	1	1
1:A:90:HIS:CE1	1:A:93:ASP:OD1	0.44	2.68	19	2
1:A:39:ALA:C	1:A:40:LEU:HD22	0.44	2.33	21	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:42:PRO:HA	1:A:99:VAL:HG23	0.44	1.90	4	2
1:A:33:PHE:CG	1:A:47:PHE:CZ	0.43	3.06	27	1
1:A:33:PHE:CE2	1:A:47:PHE:CE1	0.43	3.07	27	2
1:A:6:GLY:HA2	1:A:21:THR:HG22	0.42	1.92	13	1
1:A:49:ILE:CG2	1:A:74:LEU:HD12	0.41	2.45	30	1
1:A:22:VAL:HG11	1:A:33:PHE:CE1	0.41	2.51	22	1
1:A:40:LEU:HB3	1:A:148:TYR:CE2	0.41	2.51	7	1
1:A:40:LEU:CD1	1:A:99:VAL:HG11	0.40	2.43	16	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	152/154 (99%)	122±4 (80±3%)	23±4 (15±2%)	8±2 (5±1%)	4	25
All	All	4560/4620 (99%)	3650 (80%)	683 (15%)	227 (5%)	4	25

All 51 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	79	THR	25
1	A	55	CYS	16
1	A	67	ALA	13
1	A	66	ALA	10
1	A	82	HIS	10
1	A	101	ASN	9
1	A	65	VAL	9
1	A	78	ASN	8
1	A	85	PRO	8
1	A	39	ALA	8
1	A	89	GLY	8
1	A	133	ASN	8
1	A	11	THR	7
1	A	77	GLN	7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	76	PRO	6
1	A	121	LYS	6
1	A	90	HIS	5
1	A	75	ASP	5
1	A	54	SER	4
1	A	151	GLY	4
1	A	19	ILE	4
1	A	81	LYS	3
1	A	88	GLN	3
1	A	136	ASP	3
1	A	83	GLU	3
1	A	12	ALA	3
1	A	131	GLY	3
1	A	28	GLU	2
1	A	86	GLU	2
1	A	122	ASP	2
1	A	137	GLN	2
1	A	73	HIS	2
1	A	62	GLY	1
1	A	74	LEU	1
1	A	13	GLN	1
1	A	84	GLY	1
1	A	31	LEU	1
1	A	97	LEU	1
1	A	116	SER	1
1	A	138	PRO	1
1	A	15	VAL	1
1	A	113	ARG	1
1	A	134	MET	1
1	A	16	GLY	1
1	A	92	GLY	1
1	A	6	GLY	1
1	A	59	ILE	1
1	A	106	THR	1
1	A	70	ALA	1
1	A	115	LYS	1
1	A	91	LEU	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	115/115 (100%)	100±3 (87±2%)	15±3 (13±2%)	7 48
All	All	3450/3450 (100%)	2995 (87%)	455 (13%)	7 48

All 73 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	40	LEU	30
1	A	38	LYS	24
1	A	97	LEU	24
1	A	100	ASN	23
1	A	101	ASN	22
1	A	13	GLN	20
1	A	17	GLN	19
1	A	147	ARG	18
1	A	94	LEU	17
1	A	115	LYS	14
1	A	26	GLU	12
1	A	88	GLN	12
1	A	2	SER	11
1	A	47	PHE	11
1	A	91	LEU	11
1	A	128	HIS	10
1	A	83	GLU	9
1	A	137	GLN	9
1	A	99	VAL	7
1	A	63	GLN	7
1	A	74	LEU	6
1	A	132	ASP	6
1	A	154	LYS	6
1	A	117	LEU	6
1	A	4	LYS	5
1	A	73	HIS	5
1	A	90	HIS	5
1	A	141	LEU	5
1	A	60	LYS	5
1	A	42	PRO	5
1	A	19	ILE	5
1	A	33	PHE	4
1	A	113	ARG	4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	125	LEU	4
1	A	126	MET	4
1	A	36	HIS	4
1	A	133	ASN	3
1	A	102	ASP	3
1	A	75	ASP	3
1	A	77	GLN	3
1	A	86	GLU	3
1	A	82	HIS	3
1	A	122	ASP	3
1	A	37	LEU	2
1	A	78	ASN	2
1	A	116	SER	2
1	A	56	GLN	2
1	A	136	ASP	2
1	A	48	HIS	2
1	A	118	ASP	2
1	A	79	THR	2
1	A	150	CYS	2
1	A	8	ASN	2
1	A	28	GLU	2
1	A	110	THR	2
1	A	114	LEU	2
1	A	21	THR	2
1	A	109	VAL	2
1	A	146	THR	1
1	A	3	GLU	1
1	A	139	LYS	1
1	A	127	ILE	1
1	A	148	TYR	1
1	A	32	LYS	1
1	A	93	ASP	1
1	A	121	LYS	1
1	A	61	ASP	1
1	A	54	SER	1
1	A	27	THR	1
1	A	120	VAL	1
1	A	135	SER	1
1	A	106	THR	1
1	A	5	VAL	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

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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 97% for the well-defined parts and 97% 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	1891
Number of shifts mapped to atoms	1891
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	3

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	154	0.54 ± 0.22	Should be applied
$^{13}C_{\beta}$	129	-0.03 ± 0.20	None needed (< 0.5 ppm)
$^{13}C'$	154	0.41 ± 0.20	None needed (< 0.5 ppm)
^{15}N	154	-0.60 ± 0.33	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 97%, i.e. 1648 atoms were assigned a chemical shift out of a possible 1694. 2 out of 25 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	748/748 (100%)	297/297 (100%)	308/308 (100%)	143/143 (100%)
Sidechain	831/863 (96%)	491/503 (98%)	321/331 (97%)	19/29 (66%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	69/83 (83%)	35/46 (76%)	28/28 (100%)	6/9 (67%)
Overall	1648/1694 (97%)	823/846 (97%)	657/667 (99%)	168/181 (93%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 97%, i.e. 1648 atoms were assigned a chemical shift out of a possible 1694. 2 out of 25 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	748/748 (100%)	297/297 (100%)	308/308 (100%)	143/143 (100%)
Sidechain	831/863 (96%)	491/503 (98%)	321/331 (97%)	19/29 (66%)
Aromatic	69/83 (83%)	35/46 (76%)	28/28 (100%)	6/9 (67%)
Overall	1648/1694 (97%)	823/846 (97%)	657/667 (99%)	168/181 (93%)

7.1.4 Statistically unusual chemical shifts (i)

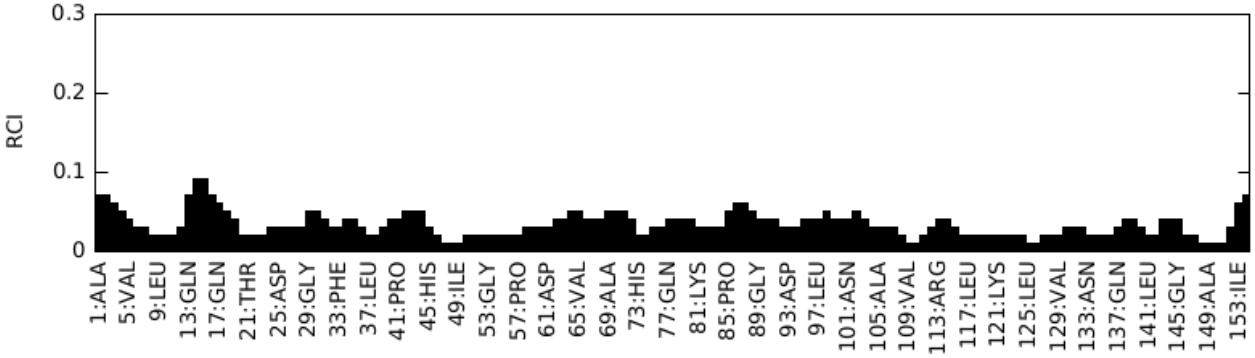
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	94	LEU	HD21	-1.12	2.14 – -0.66	-6.6
1	A	94	LEU	HD22	-1.12	2.14 – -0.66	-6.6
1	A	94	LEU	HD23	-1.12	2.14 – -0.66	-6.6

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.

Random coil index (RCI) for chain A:


RCI
0.3
0.2
0.1
0
1:ALA
5:VAL
9:LEU
13:GLN
17:GLN
21:THR
25:ASP
29:GLY
33:PHE
37:LEU
41:PRO
45:HIS
49:ILE
53:GLY
57:PRO
61:ASP
65:VAL
69:ALA
73:HIS
77:GLN
81:LYS
85:PRO
89:GLY
93:ASP
97:LEU
101:ASN
105:ALA
109:VAL
113:ARG
117:LEU
121:LYS
125:LEU
129:VAL
133:ASN
137:GLN
141:LEU
145:GLY
149:ALA
153:ILE