



Full wwPDB NMR Structure Validation Report ⓘ

Jun 12, 2024 – 01:28 AM EDT

PDB ID : 2KCO
BMRB ID : 16089
Title : Solution NMR structure of ribosomal protein sso0164 from *Sulfolobus solfataricus*. Northeast Structural Genomics Consortium (NESG) target SsT4.
Authors : Lemak, A.; Gutmanas, A.; Yee, A.; Fares, C.; Garcia, M.; Montelione, G.T.; Arrowsmith, C.H.; Northeast Structural Genomics Consortium (NESG)
Deposited on : 2008-12-23

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 ⓘ 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 ⓘ](#)) 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
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36.2

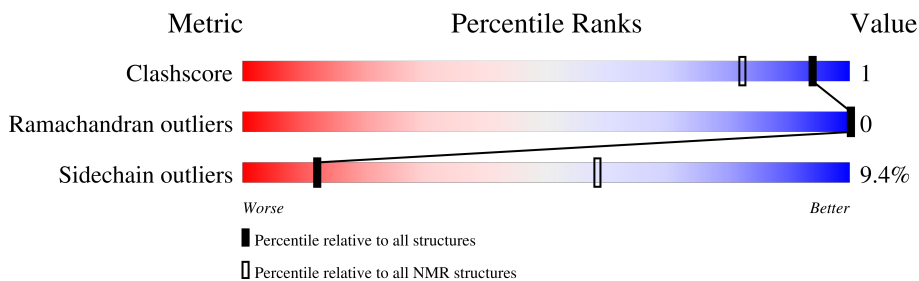
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 86%.

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	133	

2 Ensemble composition and analysis i

This entry contains 20 models. Model 19 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:39-A:47 (9)	0.23	8
2	A:61-A:126 (66)	0.56	19

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 2 clusters and 1 single-model cluster was found.

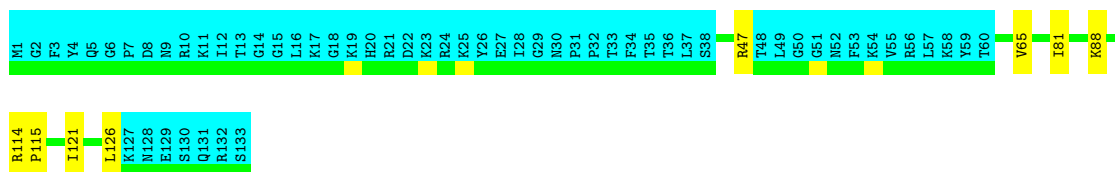
Cluster number	Models
1	1, 3, 4, 5, 6, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20
2	7, 13
Single-model clusters	2

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 2128 atoms, of which 1096 are hydrogens and 0 are deuteriums.

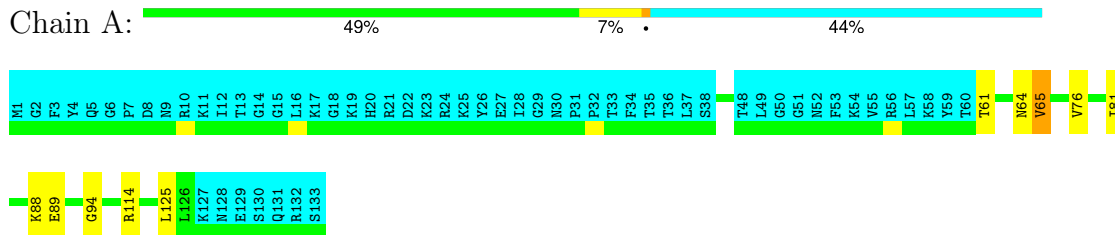
- Molecule 1 is a protein called 30S ribosomal protein S8e.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	133	2128	644	1096	196	191	1	0



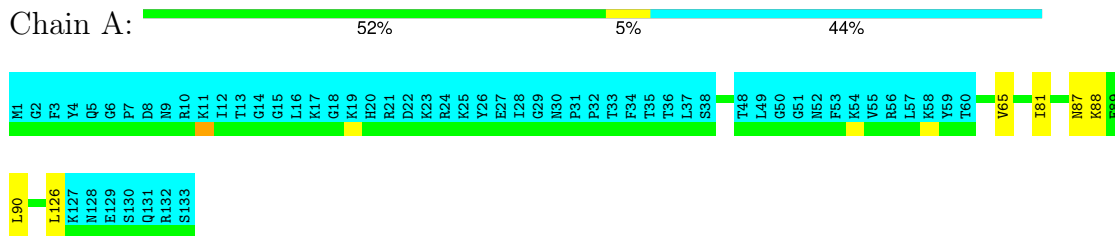
4.2.3 Score per residue for model 3

- Molecule 1: 30S ribosomal protein S8e



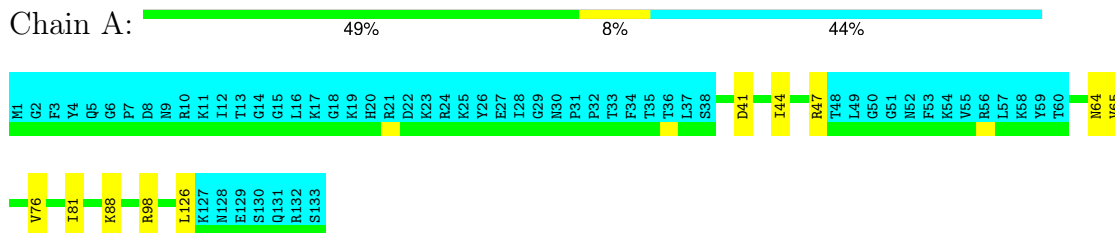
4.2.4 Score per residue for model 4

- Molecule 1: 30S ribosomal protein S8e



4.2.5 Score per residue for model 5

- Molecule 1: 30S ribosomal protein S8e



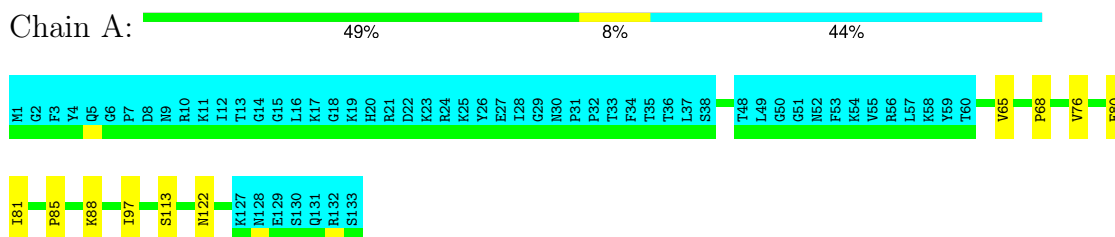
4.2.6 Score per residue for model 6

- Molecule 1: 30S ribosomal protein S8e



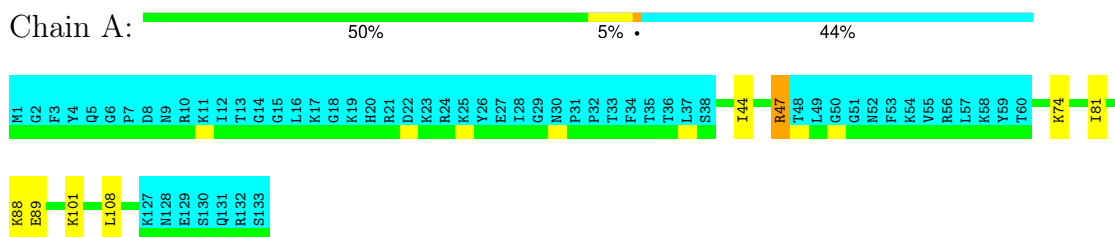
4.2.7 Score per residue for model 7

- Molecule 1: 30S ribosomal protein S8e



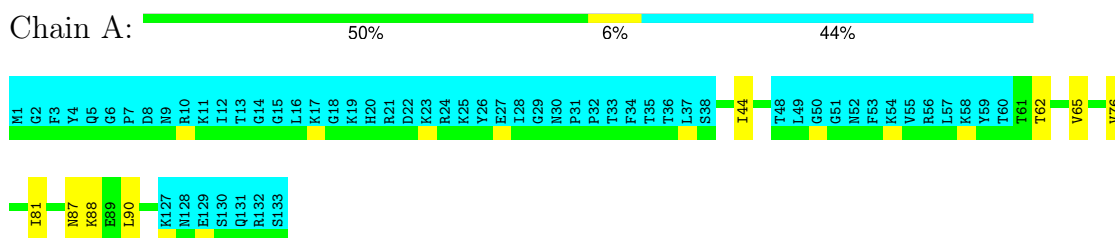
4.2.8 Score per residue for model 8

- Molecule 1: 30S ribosomal protein S8e



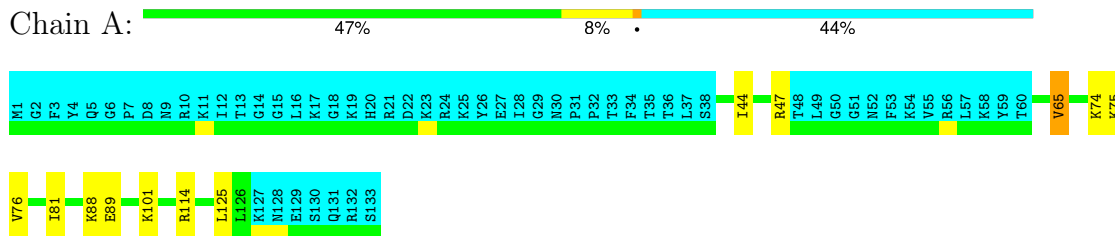
4.2.9 Score per residue for model 9

- Molecule 1: 30S ribosomal protein S8e



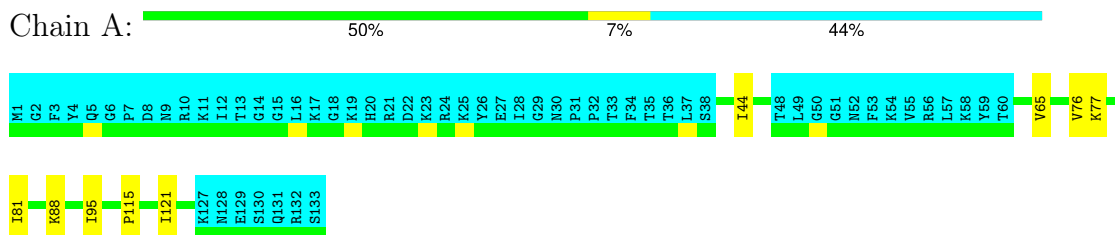
4.2.10 Score per residue for model 10

- Molecule 1: 30S ribosomal protein S8e



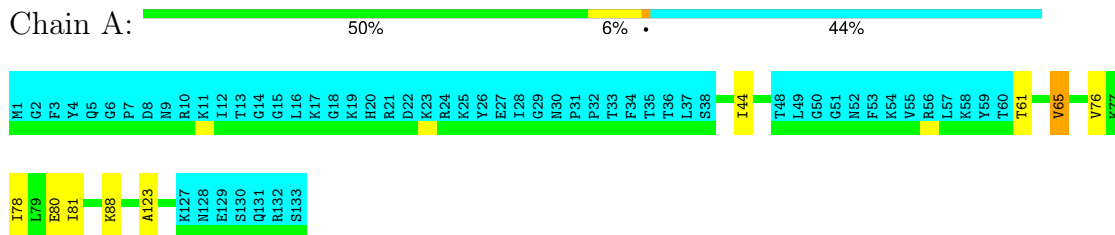
4.2.11 Score per residue for model 11

- Molecule 1: 30S ribosomal protein S8e



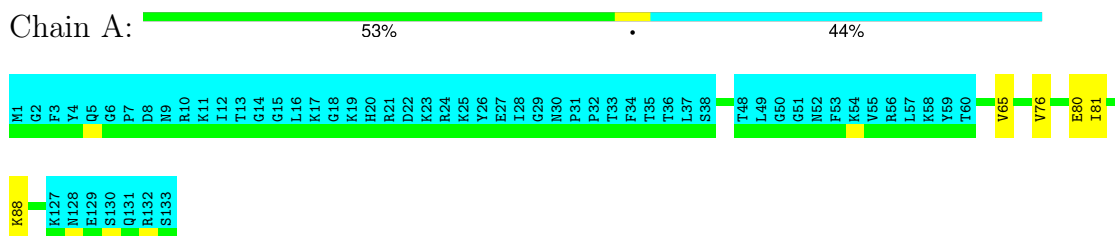
4.2.12 Score per residue for model 12

- Molecule 1: 30S ribosomal protein S8e



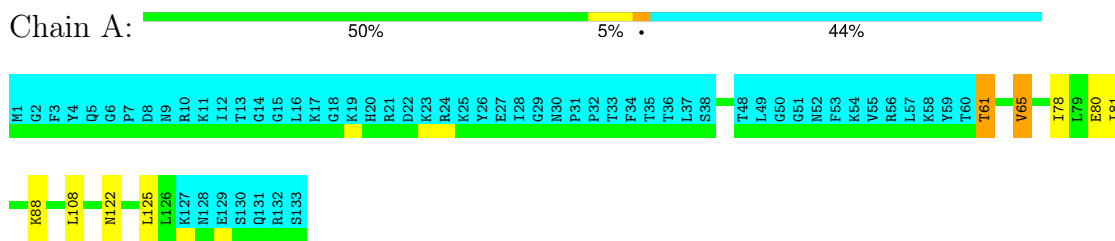
4.2.13 Score per residue for model 13

- Molecule 1: 30S ribosomal protein S8e



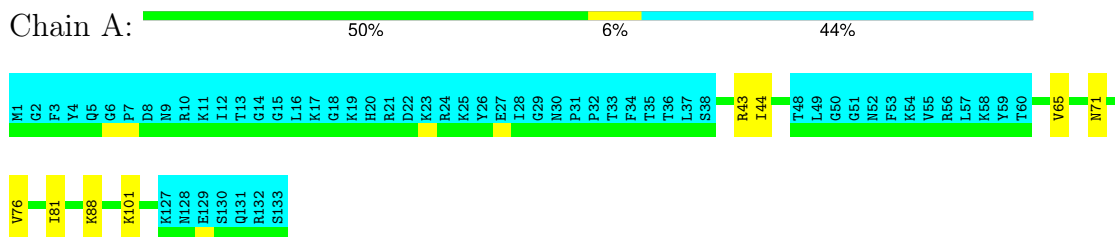
4.2.14 Score per residue for model 14

- Molecule 1: 30S ribosomal protein S8e



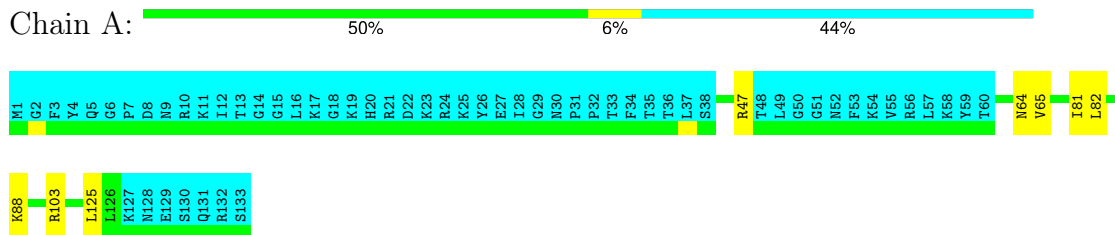
4.2.15 Score per residue for model 15

- Molecule 1: 30S ribosomal protein S8e



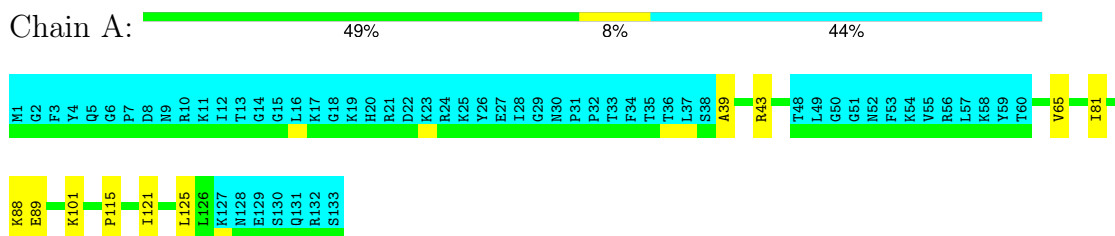
4.2.16 Score per residue for model 16

- Molecule 1: 30S ribosomal protein S8e



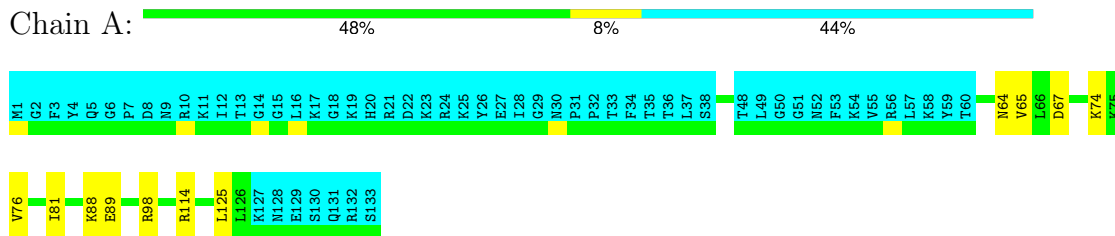
4.2.17 Score per residue for model 17

- Molecule 1: 30S ribosomal protein S8e



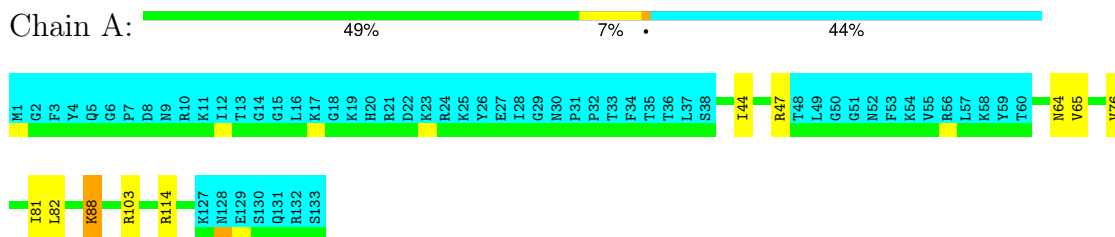
4.2.18 Score per residue for model 18

- Molecule 1: 30S ribosomal protein S8e



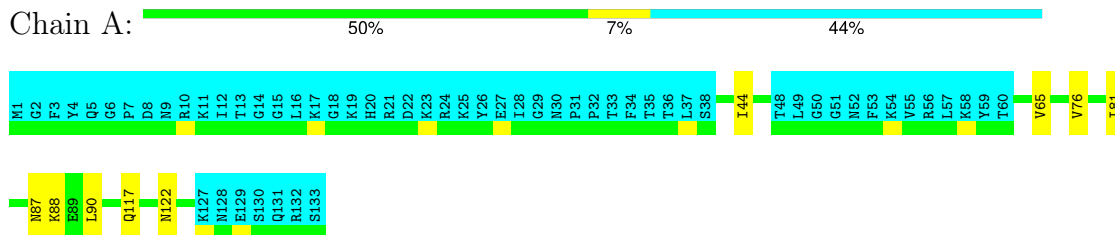
4.2.19 Score per residue for model 19 (medoid)

- Molecule 1: 30S ribosomal protein S8e



4.2.20 Score per residue for model 20

- Molecule 1: 30S ribosomal protein S8e



5 Refinement protocol and experimental data overview

The models were refined using the following method: *restrained molecular dynamics in water bath*.

Of the 100 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
CNS	refinement	

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	1407
Number of shifts mapped to atoms	1407
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	86%

6 Model quality [i](#)

6.1 Standard geometry [i](#)

There are no covalent bond-length or bond-angle outliers.

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	0.8±0.4
All	All	0	16

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

All unique planar outliers are listed below.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	65	VAL	Peptide	16

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	565	617	617	1±1
All	All	11300	12340	12340	29

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:115:PRO:HB3	1:A:121:ILE:HG13	0.53	1.80	2	1
1:A:65:VAL:HG11	1:A:125:LEU:HG	0.52	1.80	18	6

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:95:ILE:HG23	1:A:115:PRO:HG3	0.48	1.85	11	1
1:A:88:LYS:HD3	1:A:88:LYS:H	0.46	1.70	1	2
1:A:61:THR:HG22	1:A:78:ILE:HD12	0.45	1.88	14	2
1:A:113:SER:HB3	1:A:122:ASN:HB3	0.45	1.88	7	1
1:A:61:THR:HG21	1:A:94:GLY:HA3	0.45	1.88	3	1
1:A:82:LEU:HD21	1:A:103:ARG:HG3	0.44	1.89	16	2
1:A:115:PRO:HA	1:A:121:ILE:HG13	0.43	1.89	17	1
1:A:87:ASN:HD22	1:A:90:LEU:HB2	0.43	1.73	20	3
1:A:44:ILE:O	1:A:47:ARG:HB3	0.41	2.16	19	1
1:A:39:ALA:O	1:A:43:ARG:HD3	0.41	2.15	17	1
1:A:85:PRO:HB3	1:A:97:ILE:HG22	0.41	1.91	7	1
1:A:44:ILE:HD12	1:A:47:ARG:HG2	0.41	1.92	10	2
1:A:65:VAL:HA	1:A:123:ALA:O	0.40	2.15	12	1
1:A:88:LYS:HD3	1:A:88:LYS:N	0.40	2.31	1	1
1:A:121:ILE:N	1:A:121:ILE:HD12	0.40	2.31	11	1
1:A:41:ASP:O	1:A:44:ILE:HG22	0.40	2.15	5	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	75/133 (56%)	70±1 (93±2%)	5±1 (7±2%)	0±0 (0±0%)	100	100
All	All	1500/2660 (56%)	1400 (93%)	100 (7%)	0 (0%)	100	100

There are no Ramachandran outliers.

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	60/110 (55%)	54±2 (91±3%)	6±2 (9±3%)	12	58
All	All	1200/2200 (55%)	1087 (91%)	113 (9%)	12	58

All 24 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	81	ILE	20
1	A	88	LYS	20
1	A	76	VAL	13
1	A	89	GLU	6
1	A	114	ARG	6
1	A	80	GLU	5
1	A	101	LYS	5
1	A	64	ASN	5
1	A	44	ILE	5
1	A	74	LYS	4
1	A	47	ARG	4
1	A	98	ARG	3
1	A	126	LEU	3
1	A	62	THR	2
1	A	77	LYS	2
1	A	122	ASN	2
1	A	45	LYS	1
1	A	108	LEU	1
1	A	75	LYS	1
1	A	61	THR	1
1	A	43	ARG	1
1	A	71	ASN	1
1	A	67	ASP	1
1	A	117	GLN	1

6.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

6.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.6 Ligand geometry [i](#)

There are no ligands in this entry.

6.7 Other polymers [i](#)

There are no such molecules in this entry.

6.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

7 Chemical shift validation (i)

The completeness of assignment taking into account all chemical shift lists is 86% for the well-defined parts and 76% 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 (i)

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	1407
Number of shifts mapped to atoms	1407
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 (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	120	-0.27 ± 0.15	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	105	-0.08 ± 0.23	None needed (< 0.5 ppm)
$^{13}\text{C}'$	95	-0.01 ± 0.18	None needed (< 0.5 ppm)
^{15}N	102	-0.11 ± 0.29	None needed (< 0.5 ppm)

7.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 86%, i.e. 900 atoms were assigned a chemical shift out of a possible 1047. 0 out of 13 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	343/374 (92%)	142/152 (93%)	136/150 (91%)	65/72 (90%)
Sidechain	557/673 (83%)	382/439 (87%)	170/202 (84%)	5/32 (16%)

Continued on next page...

Continued from previous page...

	Total	¹ H	¹³ C	¹⁵ N
Overall	900/1047 (86%)	524/591 (89%)	306/352 (87%)	70/104 (67%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 76%, i.e. 1407 atoms were assigned a chemical shift out of a possible 1859. 0 out of 18 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹ H	¹³ C	¹⁵ N
Backbone	551/666 (83%)	234/273 (86%)	215/266 (81%)	102/127 (80%)
Sidechain	824/1128 (73%)	562/728 (77%)	252/339 (74%)	10/61 (16%)
Aromatic	32/65 (49%)	21/31 (68%)	11/32 (34%)	0/2 (0%)
Overall	1407/1859 (76%)	817/1032 (79%)	478/637 (75%)	112/190 (59%)

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:

