



# Full wwPDB NMR Structure Validation Report ⓘ

May 28, 2020 – 11:17 pm BST

PDB ID : 2L50  
Title : Solution structure of apo S100A16  
Authors : Babini, E.; Bertini, I.; Borsi, V.; Calderone, V.; Hu, X.; Luchinat, C.; Parigi, G.  
Deposited on : 2010-10-22

This is a Full wwPDB NMR Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 following versions of software and data (see [references ⓘ](#)) 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

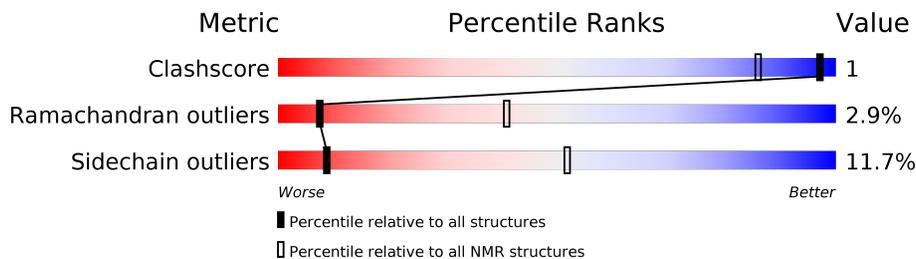
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 73%.

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  $\geq 3$ , 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	102	 76% 8% 16%
1	B	102	 75% 10% 15%

## 2 Ensemble composition and analysis i

This entry contains 30 models. Model 2 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	Residue range (total)	Backbone RMSD (Å)	Medoid model
1	A:5-A:23, A:33-A:99, B:108-B:126, B:136-B:203 (173)	0.62	2

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 3 clusters and 3 single-model clusters were found.

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 3278 atoms, of which 1636 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Protein S100-A16.

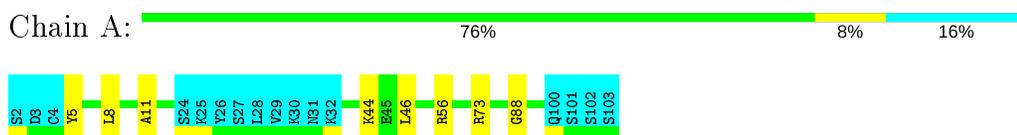
Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	102	1639	517	818	139	162	3	0
1	B	102	1639	517	818	139	162	3	0

## 4 Residue-property plots [i](#)

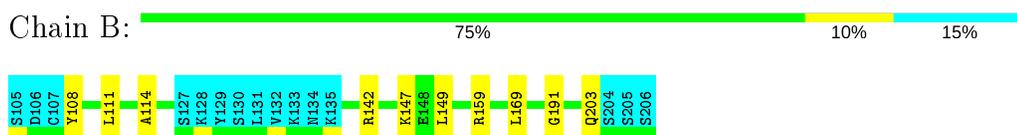
### 4.1 Average score per residue in the NMR ensemble

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: Protein S100-A16



- Molecule 1: Protein S100-A16



### 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: Protein S100-A16

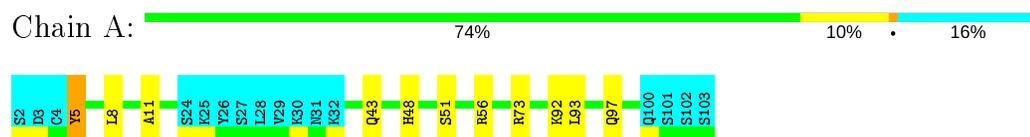


- Molecule 1: Protein S100-A16



### 4.2.2 Score per residue for model 2 (medoid)

- Molecule 1: Protein S100-A16

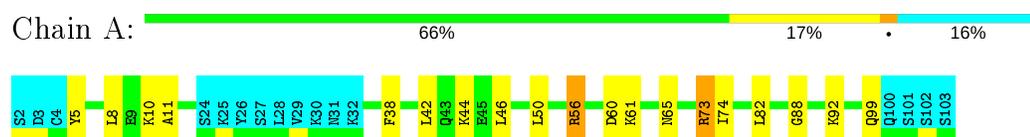


- Molecule 1: Protein S100-A16

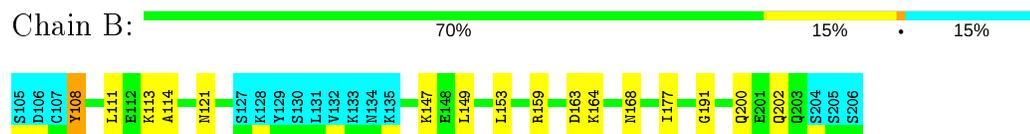


### 4.2.3 Score per residue for model 3

- Molecule 1: Protein S100-A16



- Molecule 1: Protein S100-A16



### 4.2.4 Score per residue for model 4

- Molecule 1: Protein S100-A16



- Molecule 1: Protein S100-A16



### 4.2.5 Score per residue for model 5

- Molecule 1: Protein S100-A16

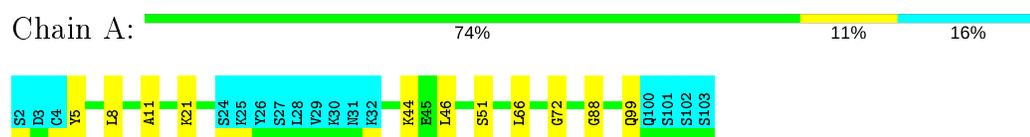


- Molecule 1: Protein S100-A16

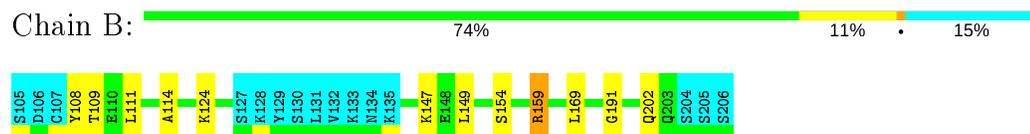


### 4.2.6 Score per residue for model 6

- Molecule 1: Protein S100-A16

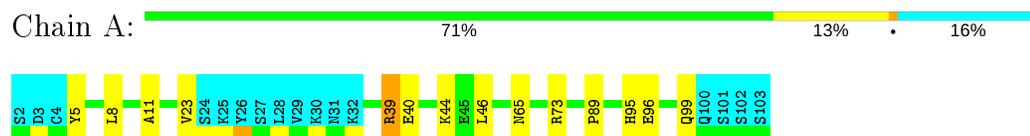


- Molecule 1: Protein S100-A16

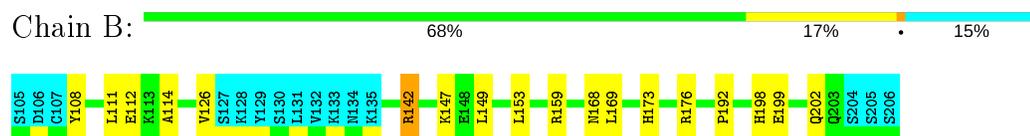


### 4.2.7 Score per residue for model 7

- Molecule 1: Protein S100-A16



- Molecule 1: Protein S100-A16



### 4.2.8 Score per residue for model 8

- Molecule 1: Protein S100-A16

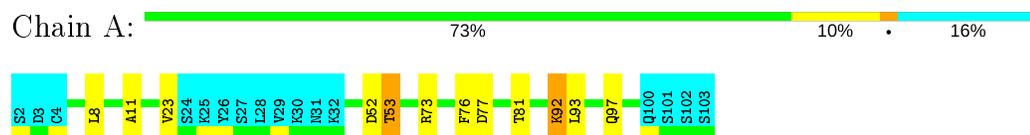


- Molecule 1: Protein S100-A16

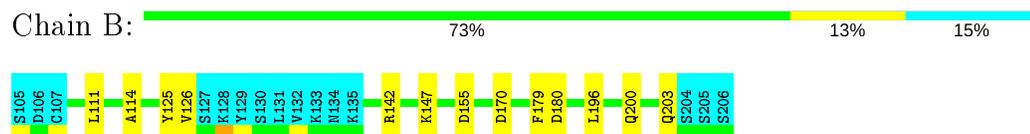


### 4.2.9 Score per residue for model 9

- Molecule 1: Protein S100-A16



- Molecule 1: Protein S100-A16

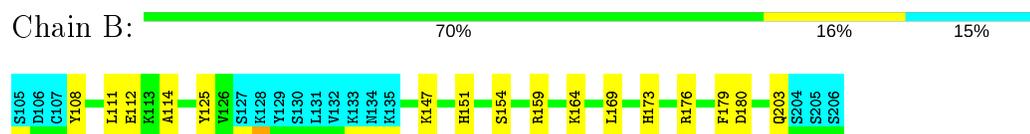


### 4.2.10 Score per residue for model 10

- Molecule 1: Protein S100-A16

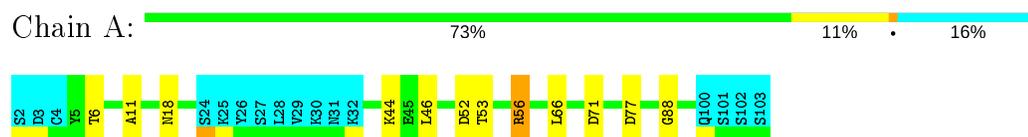


- Molecule 1: Protein S100-A16

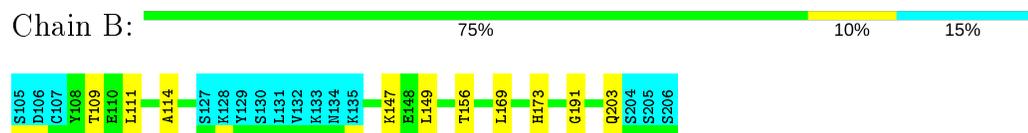


### 4.2.11 Score per residue for model 11

- Molecule 1: Protein S100-A16

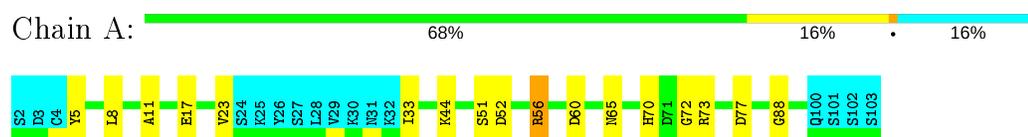


- Molecule 1: Protein S100-A16

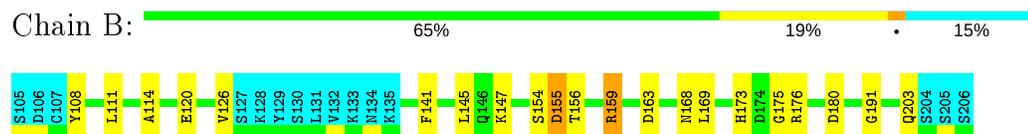


### 4.2.12 Score per residue for model 12

- Molecule 1: Protein S100-A16



- Molecule 1: Protein S100-A16



### 4.2.13 Score per residue for model 13

- Molecule 1: Protein S100-A16

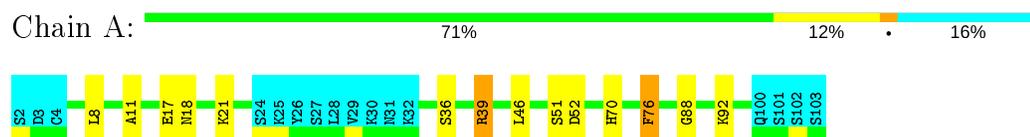


- Molecule 1: Protein S100-A16



#### 4.2.14 Score per residue for model 14

- Molecule 1: Protein S100-A16

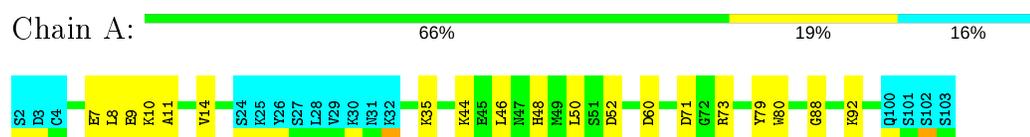


- Molecule 1: Protein S100-A16

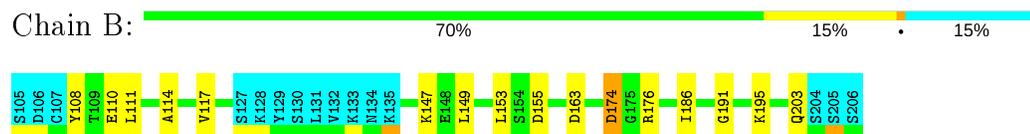


#### 4.2.15 Score per residue for model 15

- Molecule 1: Protein S100-A16

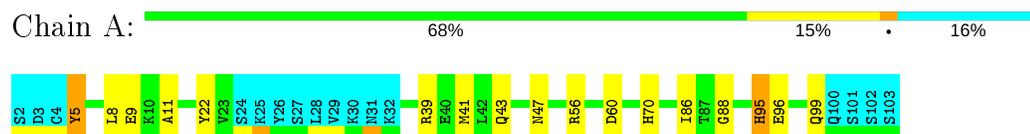


- Molecule 1: Protein S100-A16



#### 4.2.16 Score per residue for model 16

- Molecule 1: Protein S100-A16



- Molecule 1: Protein S100-A16

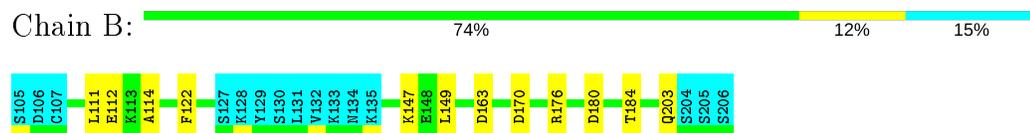


### 4.2.17 Score per residue for model 17

- Molecule 1: Protein S100-A16

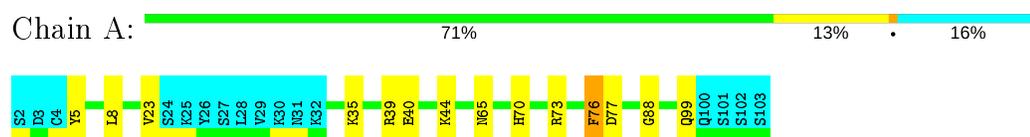


- Molecule 1: Protein S100-A16



### 4.2.18 Score per residue for model 18

- Molecule 1: Protein S100-A16

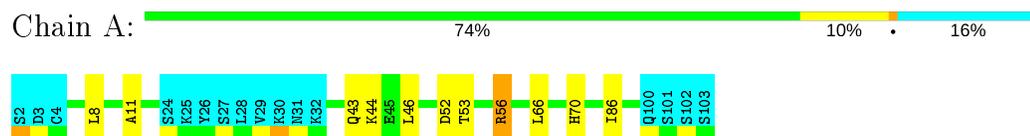


- Molecule 1: Protein S100-A16

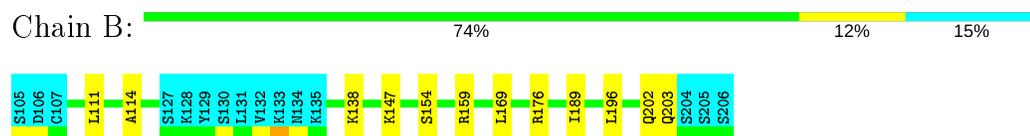


### 4.2.19 Score per residue for model 19

- Molecule 1: Protein S100-A16



- Molecule 1: Protein S100-A16

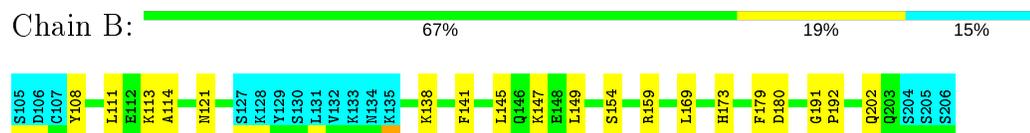


#### 4.2.20 Score per residue for model 20

- Molecule 1: Protein S100-A16

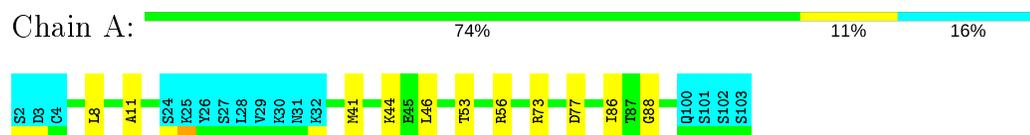


- Molecule 1: Protein S100-A16

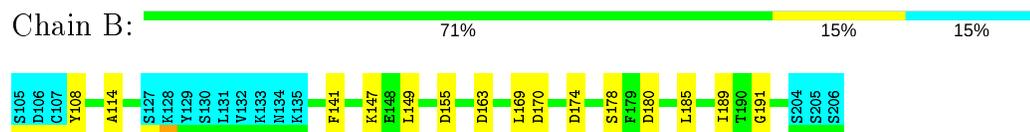


#### 4.2.21 Score per residue for model 21

- Molecule 1: Protein S100-A16

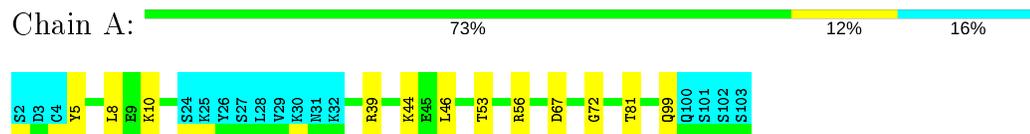


- Molecule 1: Protein S100-A16

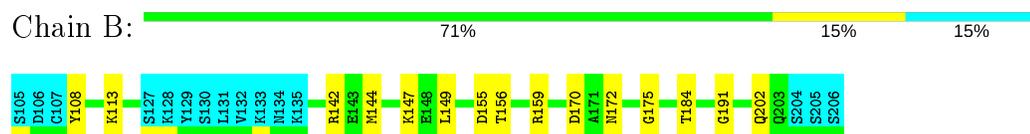


#### 4.2.22 Score per residue for model 22

- Molecule 1: Protein S100-A16

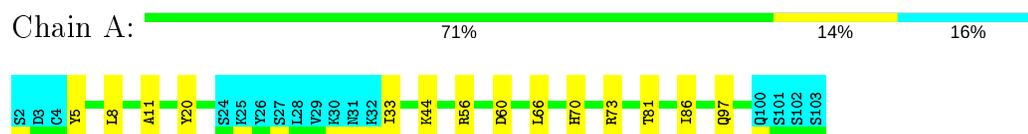


- Molecule 1: Protein S100-A16

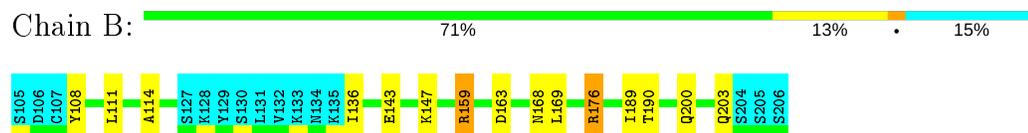


### 4.2.23 Score per residue for model 23

- Molecule 1: Protein S100-A16

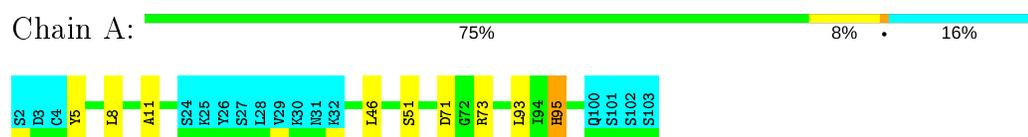


- Molecule 1: Protein S100-A16



### 4.2.24 Score per residue for model 24

- Molecule 1: Protein S100-A16

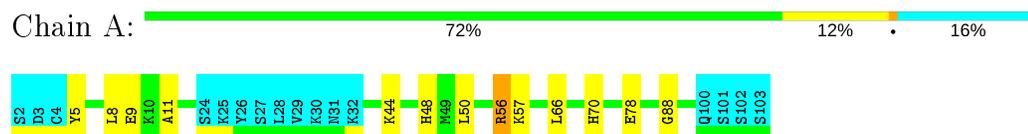


- Molecule 1: Protein S100-A16

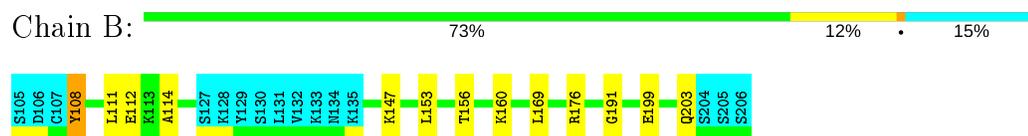


### 4.2.25 Score per residue for model 25

- Molecule 1: Protein S100-A16

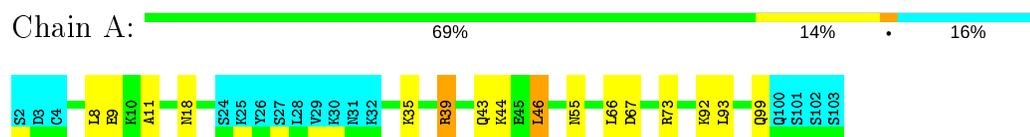


- Molecule 1: Protein S100-A16



#### 4.2.26 Score per residue for model 26

- Molecule 1: Protein S100-A16



- Molecule 1: Protein S100-A16

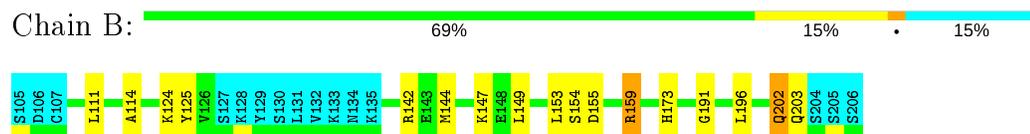


#### 4.2.27 Score per residue for model 27

- Molecule 1: Protein S100-A16

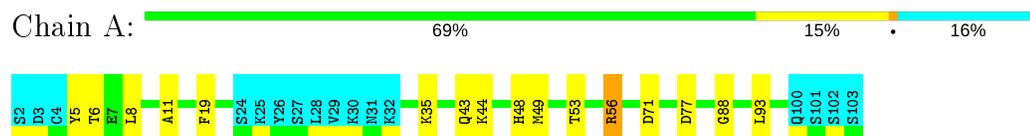


- Molecule 1: Protein S100-A16

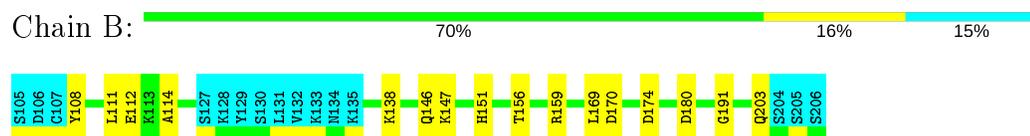


#### 4.2.28 Score per residue for model 28

- Molecule 1: Protein S100-A16



- Molecule 1: Protein S100-A16

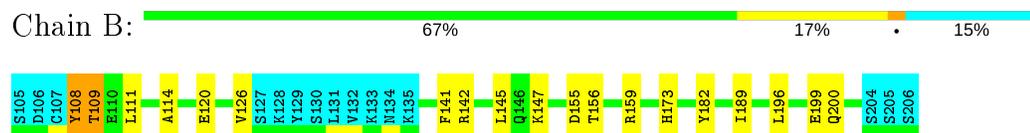


### 4.2.29 Score per residue for model 29

- Molecule 1: Protein S100-A16



- Molecule 1: Protein S100-A16

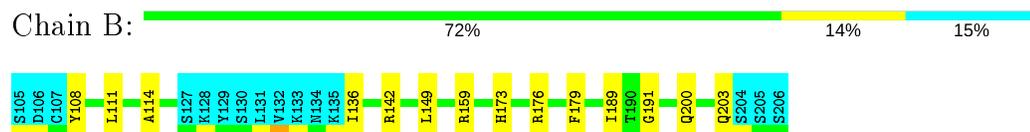


### 4.2.30 Score per residue for model 30

- Molecule 1: Protein S100-A16



- Molecule 1: Protein S100-A16



## 5 Refinement protocol and experimental data overview (i)

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

Of the 350 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	structure solution	10.0
CYANA	structure solution	2.1
ProcheckNMR	structure solution	
TALOS	structure solution	
WHAT IF	structure solution	
WHAT IF	refinement	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 6 of this report.

Chemical shift file(s)	input_cs.cif
Number of chemical shift lists	1
Total number of shifts	1904
Number of shifts mapped to atoms	1904
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	73%

No validations of the models with respect to experimental NMR restraints is performed at this time.

COVALENT-GEOMETRY INFOmissingINFO

### 5.1 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	699	695	695	2±1
1	B	708	703	703	2±1
All	All	42210	41940	41940	54

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:11:ALA:HB1	1:B:114:ALA:HB1	0.72	1.61	12	27
1:A:11:ALA:HB1	1:B:114:ALA:CB	0.50	2.36	15	3
1:A:11:ALA:CB	1:B:114:ALA:HB1	0.47	2.38	12	3
1:A:21:LYS:HE3	1:A:22:TYR:CE1	0.47	2.44	29	1
1:B:155:ASP:CG	1:B:156:THR:H	0.47	2.14	29	1
1:A:52:ASP:CG	1:A:53:THR:H	0.45	2.15	29	1
1:B:151:HIS:CD2	1:B:151:HIS:H	0.45	2.28	5	1
1:A:14:VAL:HG21	1:B:110:GLU:HG2	0.45	1.88	15	1
1:A:7:GLU:HG2	1:B:117:VAL:HG21	0.43	1.90	15	1
1:A:38:PHE:CZ	1:A:42:LEU:HD11	0.43	2.48	3	2
1:B:141:PHE:CE2	1:B:145:LEU:HD11	0.43	2.49	20	2
1:B:141:PHE:CZ	1:B:185:LEU:HB3	0.43	2.48	21	1
1:A:35:LYS:HE2	1:A:78:GLU:OE2	0.42	2.15	5	1
1:B:151:HIS:H	1:B:151:HIS:CD2	0.42	2.32	4	1
1:B:125:TYR:CE2	1:B:144:MET:HG3	0.42	2.49	5	1
1:A:38:PHE:CE2	1:A:42:LEU:HD11	0.42	2.49	29	1
1:B:141:PHE:CE1	1:B:145:LEU:HD11	0.42	2.49	12	1
1:A:48:HIS:CD2	1:B:108:TYR:HA	0.41	2.50	15	1
1:A:80:TRP:CZ2	1:B:186:ILE:HG21	0.41	2.50	15	1
1:A:48:HIS:CD2	1:A:48:HIS:H	0.41	2.33	17	1
1:A:38:PHE:CZ	1:A:82:LEU:HB3	0.40	2.51	3	1
1:A:52:ASP:CG	1:A:53:THR:N	0.40	2.75	19	1

## 5.2 Torsion angles [i](#)

### 5.2.1 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	86/102 (84%)	76±2 (88±3%)	8±2 (9±3%)	2±1 (3±1%)	8	42
1	B	87/102 (85%)	75±3 (87±3%)	9±3 (10±3%)	3±1 (3±1%)	7	40
All	All	5190/6120 (85%)	4530 (87%)	509 (10%)	151 (3%)	7	41

All 33 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	88	GLY	20
1	B	191	GLY	19
1	B	108	TYR	15
1	A	5	TYR	14
1	B	203	GLN	9
1	B	155	ASP	7
1	A	52	ASP	6
1	B	154	SER	6
1	A	71	ASP	5
1	A	23	VAL	5
1	B	126	VAL	5
1	A	51	SER	4
1	A	89	PRO	4
1	A	47	ASN	3
1	B	192	PRO	3
1	B	174	ASP	3
1	A	53	THR	3
1	B	156	THR	2
1	A	50	LEU	2
1	B	136	ILE	2
1	B	178	SER	2
1	A	70	HIS	1
1	B	173	HIS	1
1	B	172	ASN	1
1	A	35	LYS	1
1	A	49	MET	1
1	A	75	SER	1
1	B	180	ASP	1
1	A	6	THR	1
1	B	149	LEU	1
1	B	153	LEU	1
1	B	109	THR	1
1	A	33	ILE	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	76/92 (83%)	67±2 (88±3%)	9±2 (12±3%)	8	51
1	B	77/92 (84%)	68±2 (89±3%)	9±2 (11±3%)	9	52
All	All	4590/5520 (83%)	4053 (88%)	537 (12%)	9	52

All 93 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	8	LEU	28
1	B	111	LEU	27
1	A	44	LYS	25
1	B	147	LYS	25
1	A	46	LEU	18
1	B	149	LEU	16
1	B	169	LEU	16
1	A	77	ASP	13
1	A	66	LEU	12
1	A	70	HIS	12
1	A	43	GLN	12
1	B	202	GLN	11
1	B	180	ASP	11
1	A	92	LYS	11
1	B	112	GLU	10
1	B	146	GLN	9
1	A	9	GLU	9
1	B	173	HIS	9
1	A	99	GLN	9
1	B	151	HIS	8
1	A	60	ASP	8
1	A	97	GLN	7
1	A	65	ASN	7
1	B	170	ASP	7
1	B	163	ASP	7
1	A	93	LEU	7
1	B	200	GLN	7
1	A	48	HIS	7
1	B	203	GLN	7
1	B	195	LYS	7
1	B	196	LEU	7
1	B	168	ASN	6
1	B	189	ILE	6
1	A	86	ILE	6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	10	LYS	5
1	A	81	THR	5
1	A	18	ASN	5
1	A	67	ASP	5
1	B	153	LEU	5
1	A	71	ASP	4
1	B	138	LYS	4
1	A	56	ARG	4
1	A	5	TYR	4
1	B	124	LYS	4
1	B	121	ASN	4
1	B	113	LYS	4
1	A	35	LYS	4
1	A	6	THR	4
1	A	41	MET	3
1	A	40	GLU	3
1	A	53	THR	3
1	B	120	GLU	3
1	A	50	LEU	3
1	A	51	SER	3
1	A	21	LYS	3
1	A	95	HIS	3
1	B	109	THR	3
1	B	144	MET	3
1	B	174	ASP	3
1	B	156	THR	3
1	B	108	TYR	3
1	A	17	GLU	2
1	B	164	LYS	2
1	B	159	ARG	2
1	A	75	SER	2
1	A	73	ARG	2
1	A	74	ILE	2
1	A	96	GLU	2
1	B	199	GLU	2
1	B	155	ASP	2
1	A	61	LYS	2
1	A	57	LYS	2
1	B	143	GLU	2
1	B	160	LYS	2
1	B	184	THR	2
1	B	148	GLU	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	B	176	ARG	2
1	B	154	SER	2
1	B	178	SER	2
1	B	158	ASN	1
1	A	55	ASN	1
1	A	69	ASN	1
1	B	190	THR	1
1	B	198	HIS	1
1	B	182	TYR	1
1	A	47	ASN	1
1	B	110	GLU	1
1	B	177	ILE	1
1	A	52	ASP	1
1	A	78	GLU	1
1	A	36	SER	1
1	B	142	ARG	1
1	A	39	ARG	1

### 5.2.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

### 5.4 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 5.5 Ligand geometry [i](#)

There are no ligands in this entry.

### 5.6 Other polymers [i](#)

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 [i](#)

The completeness of assignment taking into account all chemical shift lists is 73% for the well-defined parts and 72% for the entire structure.

### 6.1 Chemical shift list 1

File name: input\_cs.cif

Chemical shift list name: *assigned\_chem\_shift\_list\_1*

#### 6.1.1 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	1904
Number of shifts mapped to atoms	1904
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	2

#### 6.1.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	200	$2.00 \pm 0.03$	Should be applied
$^{13}\text{C}_\beta$	190	$2.92 \pm 0.07$	Should be applied
$^{13}\text{C}'$	194	$2.23 \pm 0.07$	Should be applied
$^{15}\text{N}$	194	$0.23 \pm 0.24$	None needed ( $< 0.5$ ppm)

#### 6.1.3 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 73%, i.e. 1609 atoms were assigned a chemical shift out of a possible 2199. 26 out of 26 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	845/861 (98%)	338/344 (98%)	338/346 (98%)	169/171 (99%)
Sidechain	762/1154 (66%)	390/672 (58%)	370/431 (86%)	2/51 (4%)

*Continued on next page...*

Continued from previous page...

	<b>Total</b>	<b><sup>1</sup>H</b>	<b><sup>13</sup>C</b>	<b><sup>15</sup>N</b>
Aromatic	2/184 (1%)	2/98 (2%)	0/78 (0%)	0/8 (0%)
Overall	1609/2199 (73%)	730/1114 (66%)	708/855 (83%)	171/230 (74%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 72%, i.e. 1848 atoms were assigned a chemical shift out of a possible 2558. 30 out of 30 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	<b>Total</b>	<b><sup>1</sup>H</b>	<b><sup>13</sup>C</b>	<b><sup>15</sup>N</b>
Backbone	980/1016 (96%)	392/406 (97%)	394/408 (97%)	194/202 (96%)
Sidechain	866/1342 (65%)	442/786 (56%)	422/496 (85%)	2/60 (3%)
Aromatic	2/200 (1%)	2/106 (2%)	0/86 (0%)	0/8 (0%)
Overall	1848/2558 (72%)	836/1298 (64%)	816/990 (82%)	196/270 (73%)

#### 6.1.4 Statistically unusual chemical shifts [i](#)

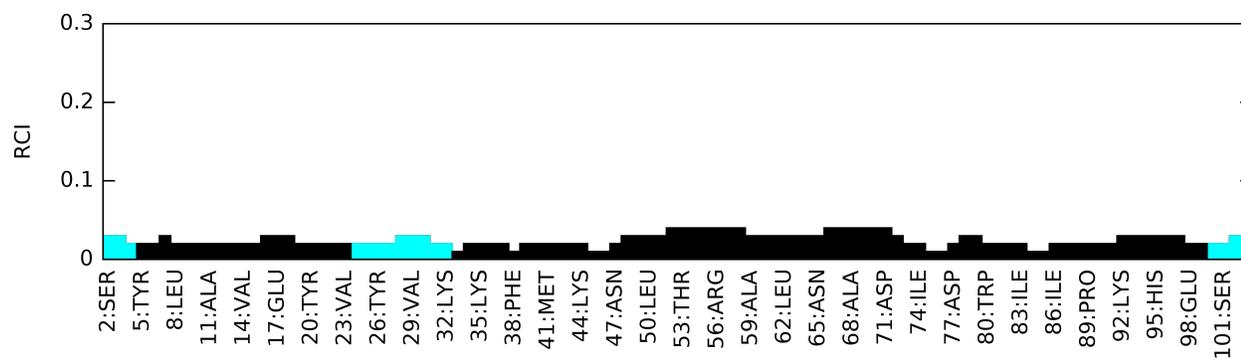
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	B	187	GLY	HA3	1.68	5.80 – 2.00	-5.8
1	A	84	GLY	HA3	1.68	5.80 – 2.00	-5.8

#### 6.1.5 Random Coil Index (RCI) plots [i](#)

The images below report *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:



Random coil index (RCI) for chain B:

