



# Full wwPDB NMR Structure Validation Report ⓘ

May 28, 2020 – 11:55 pm BST

PDB ID : 2MOQ  
Title : Solution Structure and Molecular determinants of Hemoglobin Binding of the first NEAT Domain of IsdB in Staphylococcus aureus  
Authors : Fonner, B.A.; Tripet, B.P.; Eilers, B.J.; Stanisich, J.; Sullivan-Springhetti, R.K.; Moore, R.; Lui, M.; Lei, B.; Copie, V.  
Deposited on : 2014-04-29

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

We welcome your comments at [validation@mail.wwpdb.org](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

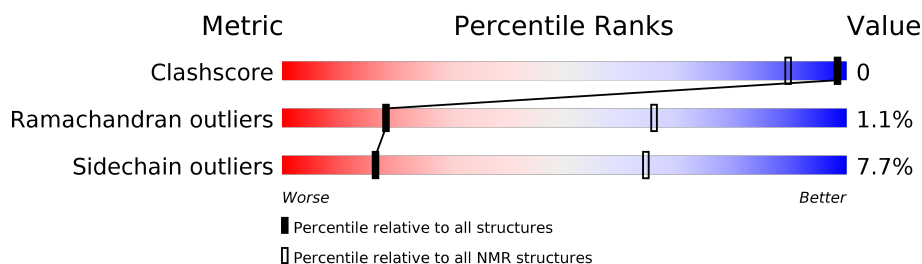
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 60%.

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	163	

## 2 Ensemble composition and analysis ⓘ

This entry contains 20 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:127-A:262 (136)	0.81	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 2 clusters. No single-model clusters were found.

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 2414 atoms, of which 1195 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Iron-regulated surface determinant protein B.

Mol	Chain	Residues	Atoms						Trace
1	A	148	Total	C	H	N	O	S	0
			2414	779	1195	202	236	2	

There are 15 discrepancies between the modelled and reference sequences:

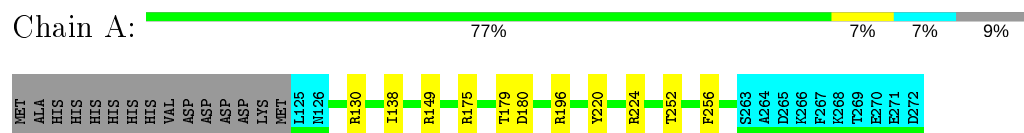
Chain	Residue	Modelled	Actual	Comment	Reference
A	110	MET	-	EXPRESSION TAG	UNP Q8NX66
A	111	ALA	-	EXPRESSION TAG	UNP Q8NX66
A	112	HIS	-	EXPRESSION TAG	UNP Q8NX66
A	113	HIS	-	EXPRESSION TAG	UNP Q8NX66
A	114	HIS	-	EXPRESSION TAG	UNP Q8NX66
A	115	HIS	-	EXPRESSION TAG	UNP Q8NX66
A	116	HIS	-	EXPRESSION TAG	UNP Q8NX66
A	117	HIS	-	EXPRESSION TAG	UNP Q8NX66
A	118	VAL	-	EXPRESSION TAG	UNP Q8NX66
A	119	ASP	-	EXPRESSION TAG	UNP Q8NX66
A	120	ASP	-	EXPRESSION TAG	UNP Q8NX66
A	121	ASP	-	EXPRESSION TAG	UNP Q8NX66
A	122	ASP	-	EXPRESSION TAG	UNP Q8NX66
A	123	LYS	-	EXPRESSION TAG	UNP Q8NX66
A	124	MET	-	EXPRESSION TAG	UNP Q8NX66

## 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: Iron-regulated surface determinant protein B

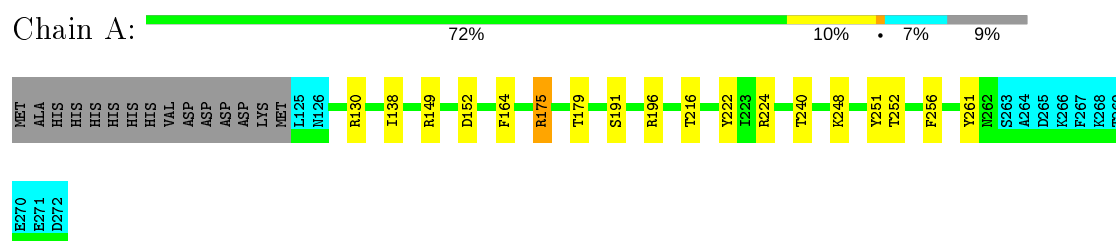


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

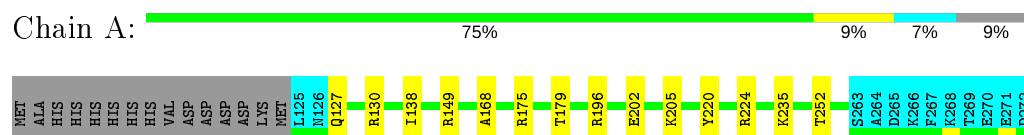
#### 4.2.1 Score per residue for model 1

- Molecule 1: Iron-regulated surface determinant protein B



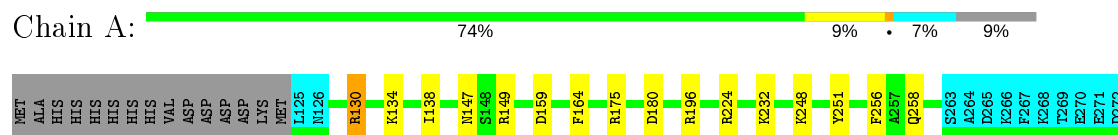
#### 4.2.2 Score per residue for model 2 (medoid)

- Molecule 1: Iron-regulated surface determinant protein B



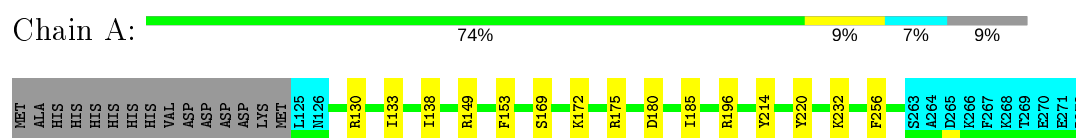
### 4.2.3 Score per residue for model 3

- Molecule 1: Iron-regulated surface determinant protein B



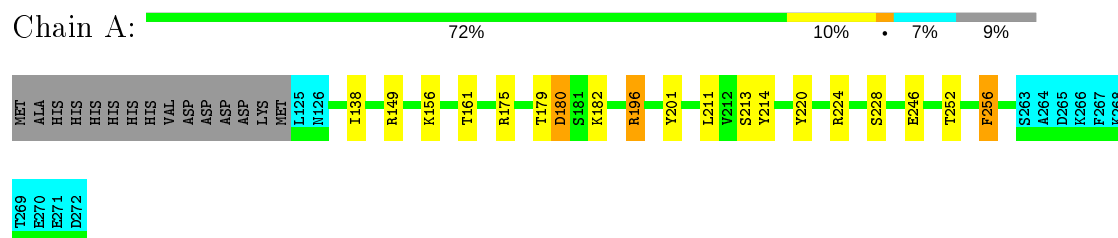
### 4.2.4 Score per residue for model 4

- Molecule 1: Iron-regulated surface determinant protein B



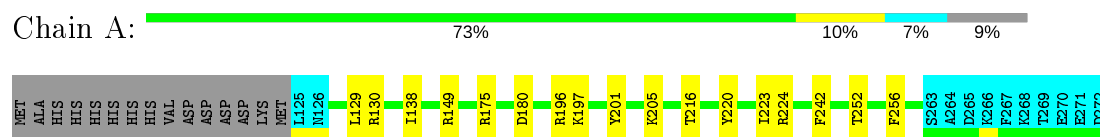
### 4.2.5 Score per residue for model 5

- Molecule 1: Iron-regulated surface determinant protein B



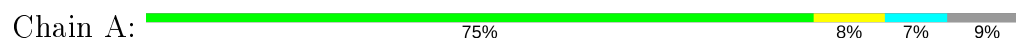
### 4.2.6 Score per residue for model 6

- Molecule 1: Iron-regulated surface determinant protein B



### 4.2.7 Score per residue for model 7

- Molecule 1: Iron-regulated surface determinant protein B





#### 4.2.8 Score per residue for model 8

- Molecule 1: Iron-regulated surface determinant protein B

Chain A: 76% 7% 7% 9%



#### 4.2.9 Score per residue for model 9

- Molecule 1: Iron-regulated surface determinant protein B

Chain A: 75% 8% 7% 9%



#### 4.2.10 Score per residue for model 10

- Molecule 1: Iron-regulated surface determinant protein B

Chain A: 73% 10% 7% 9%

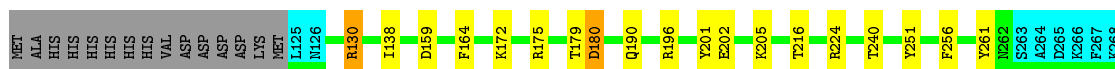


D272

#### 4.2.11 Score per residue for model 11

- Molecule 1: Iron-regulated surface determinant protein B

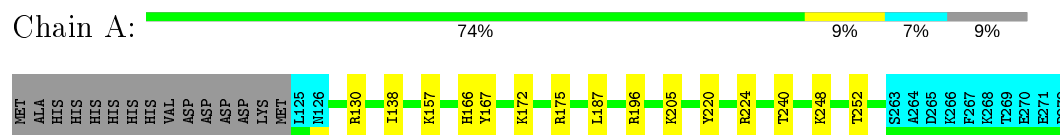
Chain A: 72% 10% 7% 9%



T269  
E270  
E271  
D272

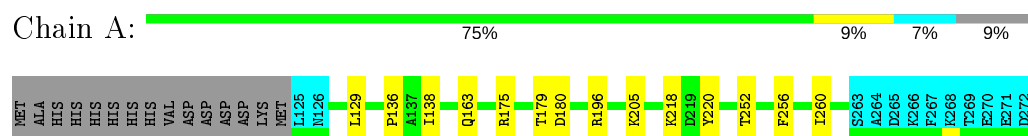
### 4.2.12 Score per residue for model 12

- Molecule 1: Iron-regulated surface determinant protein B



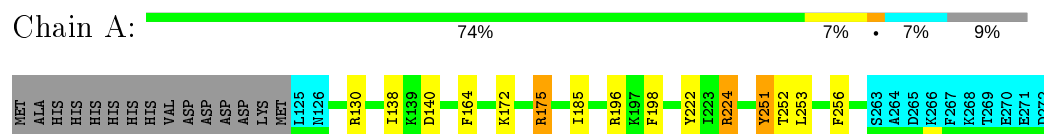
### 4.2.13 Score per residue for model 13

- Molecule 1: Iron-regulated surface determinant protein B



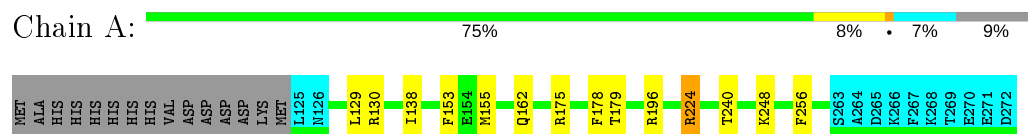
### 4.2.14 Score per residue for model 14

- Molecule 1: Iron-regulated surface determinant protein B



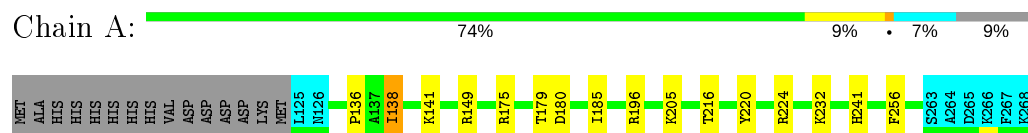
### 4.2.15 Score per residue for model 15

- Molecule 1: Iron-regulated surface determinant protein B



### 4.2.16 Score per residue for model 16

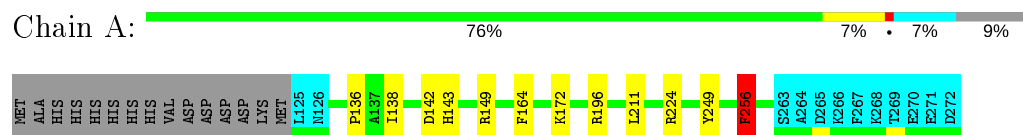
- Molecule 1: Iron-regulated surface determinant protein B





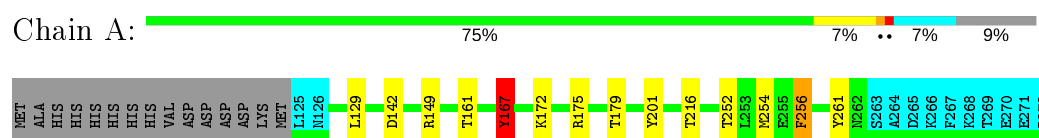
### 4.2.17 Score per residue for model 17

- Molecule 1: Iron-regulated surface determinant protein B



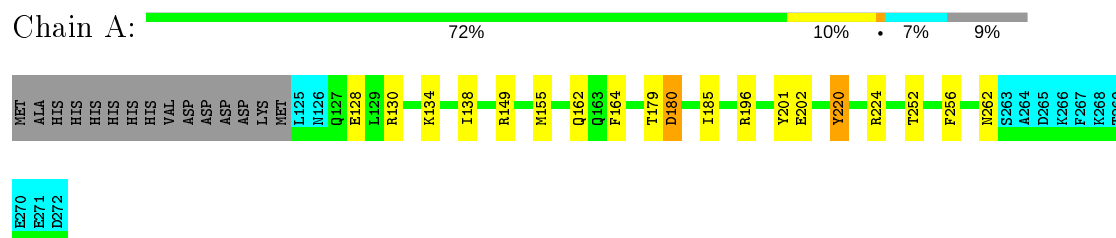
### 4.2.18 Score per residue for model 18

- Molecule 1: Iron-regulated surface determinant protein B



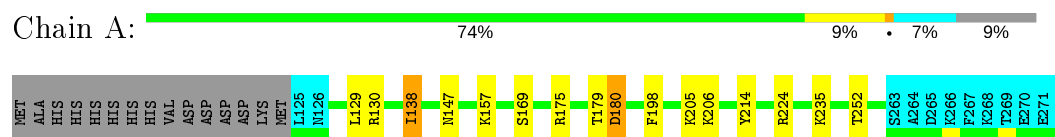
### 4.2.19 Score per residue for model 19

- Molecule 1: Iron-regulated surface determinant protein B



### 4.2.20 Score per residue for model 20

- Molecule 1: Iron-regulated surface determinant protein B



## 5 Refinement protocol and experimental data overview

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

Of the 200 calculated structures, 20 were deposited, based on the following criterion: *structures with the least restraint violations*.

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

Software name	Classification	Version
CYANA	refinement	
AMBER	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	1280
Number of shifts mapped to atoms	1280
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	60%

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	1121	1106	1106	0±1
All	All	22420	22120	22120	10

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 0.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:133:ILE:HG22	1:A:214:TYR:CE1	0.59	2.33	4	1
1:A:138:ILE:H	1:A:138:ILE:HD12	0.56	1.61	20	1
1:A:179:THR:HG22	1:A:180:ASP:H	0.52	1.62	19	6
1:A:138:ILE:HD13	1:A:138:ILE:H	0.46	1.71	16	1
1:A:246:GLU:H	1:A:246:GLU:CD	0.45	2.15	5	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	136/163 (83%)	124±2 (91±2%)	10±2 (7±2%)	2±1 (1±1%)	18	66
All	All	2720/3260 (83%)	2488 (91%)	201 (7%)	31 (1%)	18	66

All 9 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	180	ASP	11
1	A	164	PHE	7
1	A	251	TYR	4
1	A	136	PRO	3
1	A	147	ASN	2
1	A	140	ASP	1
1	A	127	GLN	1
1	A	167	TYR	1
1	A	196	ARG	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	123/148 (83%)	114±2 (92±2%)	9±2 (8±2%)	16	64
All	All	2460/2960 (83%)	2271 (92%)	189 (8%)	16	64

All 61 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	138	ILE	19
1	A	256	PHE	15
1	A	252	THR	13
1	A	220	TYR	10
1	A	205	LYS	8
1	A	172	LYS	7
1	A	129	LEU	7
1	A	201	TYR	6
1	A	248	LYS	5
1	A	185	ILE	5
1	A	216	THR	5
1	A	198	PHE	4
1	A	261	TYR	4
1	A	202	GLU	4
1	A	179	THR	4
1	A	232	LYS	4
1	A	240	THR	4
1	A	134	LYS	3
1	A	157	LYS	3
1	A	162	GLN	3
1	A	130	ARG	3
1	A	182	LYS	2
1	A	258	GLN	2
1	A	155	MET	2
1	A	222	TYR	2
1	A	159	ASP	2
1	A	224	ARG	2
1	A	214	TYR	2
1	A	251	TYR	2
1	A	161	THR	2
1	A	153	PHE	2
1	A	142	ASP	2
1	A	169	SER	2
1	A	235	LYS	2
1	A	166	HIS	1
1	A	167	TYR	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	254	MET	1
1	A	213	SER	1
1	A	152	ASP	1
1	A	175	ARG	1
1	A	141	LYS	1
1	A	156	LYS	1
1	A	144	SER	1
1	A	206	LYS	1
1	A	187	LEU	1
1	A	191	SER	1
1	A	196	ARG	1
1	A	228	SER	1
1	A	244	ASN	1
1	A	190	GLN	1
1	A	260	ILE	1
1	A	164	PHE	1
1	A	128	GLU	1
1	A	241	HIS	1
1	A	247	GLU	1
1	A	253	LEU	1
1	A	197	LYS	1
1	A	163	GLN	1
1	A	218	LYS	1
1	A	262	ASN	1
1	A	223	ILE	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 [i](#)

The completeness of assignment taking into account all chemical shift lists is 60% for the well-defined parts and 61% 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	1280
Number of shifts mapped to atoms	1280
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	1

#### 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
$^{13}\text{C}_\alpha$	143	$-0.41 \pm 0.13$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}_\beta$	138	$-0.37 \pm 0.13$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}'$	0	—	None (insufficient data)
$^{15}\text{N}$	131	$-0.04 \pm 0.45$	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 60%, i.e. 1060 atoms were assigned a chemical shift out of a possible 1760. 7 out of 14 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	495/666 (74%)	245/265 (92%)	131/272 (48%)	119/129 (92%)
Sidechain	548/917 (60%)	299/544 (55%)	249/329 (76%)	0/44 (0%)

*Continued on next page...*

Continued from previous page...

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	17/177 (10%)	16/94 (17%)	0/79 (0%)	1/4 (25%)
Overall	1060/1760 (60%)	560/903 (62%)	380/680 (56%)	120/177 (68%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 61%, i.e. 1161 atoms were assigned a chemical shift out of a possible 1905. 7 out of 15 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Backbone	542/726 (75%)	268/289 (93%)	143/296 (48%)	131/141 (93%)
Sidechain	601/993 (61%)	330/588 (56%)	271/358 (76%)	0/47 (0%)
Aromatic	18/186 (10%)	17/99 (17%)	0/83 (0%)	1/4 (25%)
Overall	1161/1905 (61%)	615/976 (63%)	414/737 (56%)	132/192 (69%)

#### 6.1.4 Statistically unusual chemical shifts ⓘ

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	218	LYS	HB2	0.50	3.03 – 0.53	-5.1

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

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:

