



# Full wwPDB NMR Structure Validation Report ⓘ

Feb 13, 2017 – 01:33 am GMT

PDB ID : 2N2E  
Title : NMR solution structure of the C-terminal domain of NisI, a lipoprotein from *Lactococcus lactis* which confers immunity against nisin  
Authors : Hacker, C.; Christ, N.A.; Korn, S.; Duchardt-Ferner, E.; Hellmich, U.A.; Duesterhus, S.; Koetter, P.; Entian, K.; Woehnert, J.  
Deposited on : 2015-05-08

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

<http://wwpdb.org/validation/2016/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	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	trunk28760
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

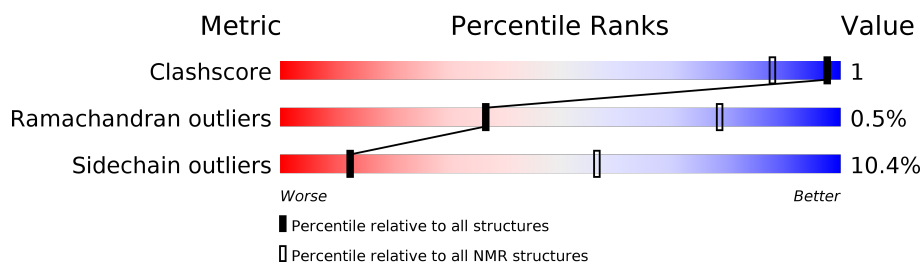
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 88%.

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	125131	11601
Ramachandran outliers	121729	10391
Sidechain outliers	121581	10367

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	129	

## 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: *minimized average structure*.

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:120-A:176, A:190-A:224 (92)	0.23	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 4 clusters and 5 single-model clusters were found.

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

### 3 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 1673 atoms, of which 814 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Nisin immunity protein.

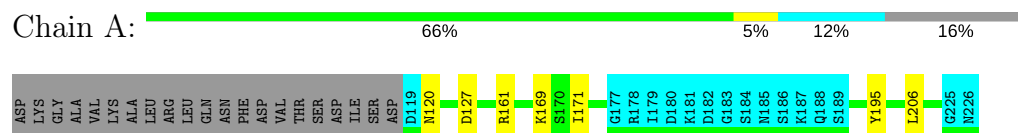
Mol	Chain	Residues	Atoms						Trace
1	A	108	Total	C	H	N	O	S	0
			1673	530	814	146	182	1	

## 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: Nisin immunity protein

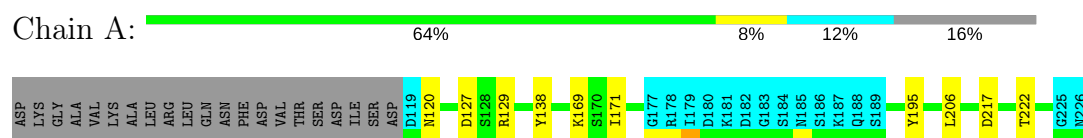


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

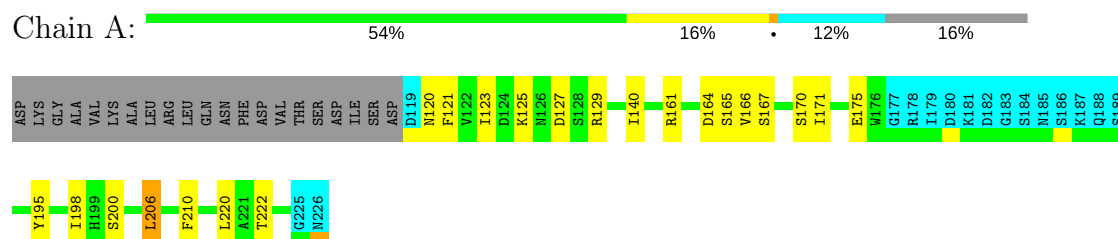
#### 4.2.1 Score per residue for model 1

- Molecule 1: Nisin immunity protein



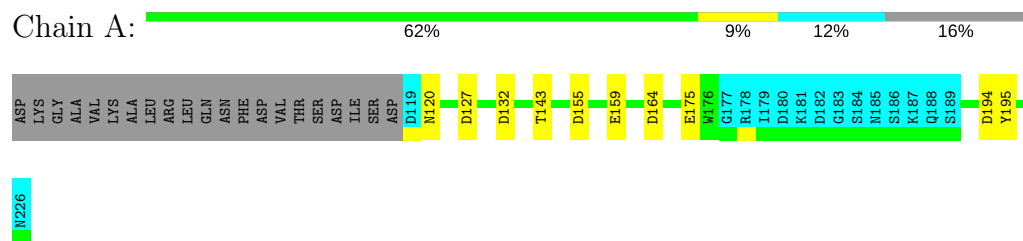
#### 4.2.2 Score per residue for model 2 (medoid)

- Molecule 1: Nisin immunity protein



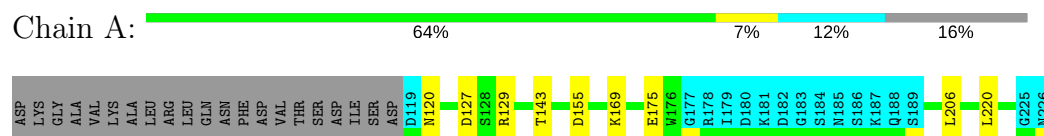
### 4.2.3 Score per residue for model 3

- Molecule 1: Nisin immunity protein



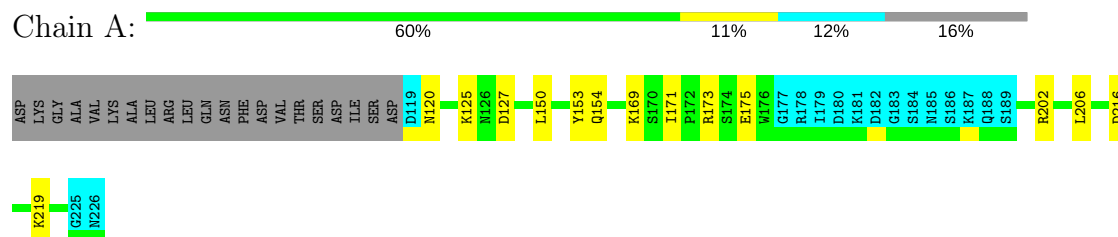
### 4.2.4 Score per residue for model 4

- Molecule 1: Nisin immunity protein



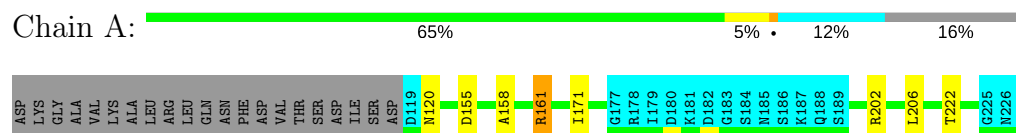
### 4.2.5 Score per residue for model 5

- Molecule 1: Nisin immunity protein



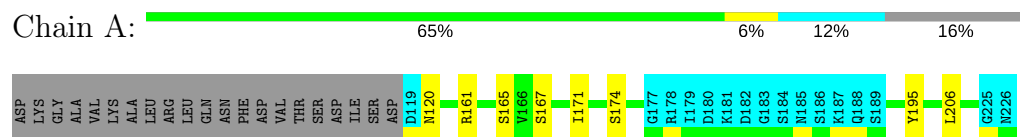
### 4.2.6 Score per residue for model 6

- Molecule 1: Nisin immunity protein



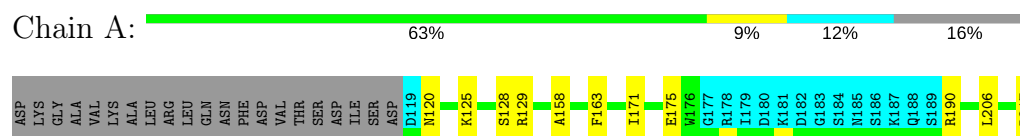
### 4.2.7 Score per residue for model 7

- Molecule 1: Nisin immunity protein



### 4.2.8 Score per residue for model 8

- Molecule 1: Nisin immunity protein



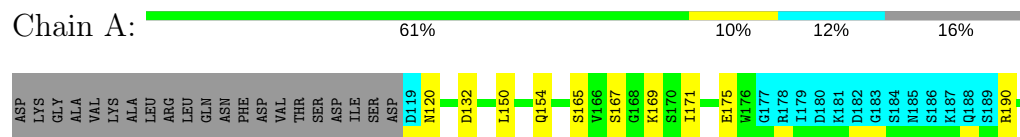
### 4.2.9 Score per residue for model 9

- Molecule 1: Nisin immunity protein



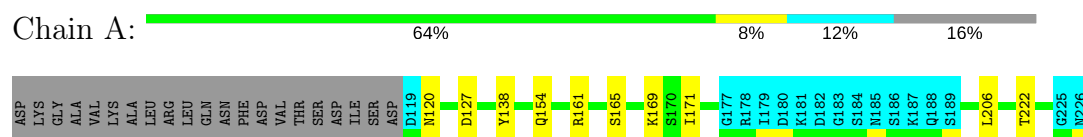
### 4.2.10 Score per residue for model 10

- Molecule 1: Nisin immunity protein



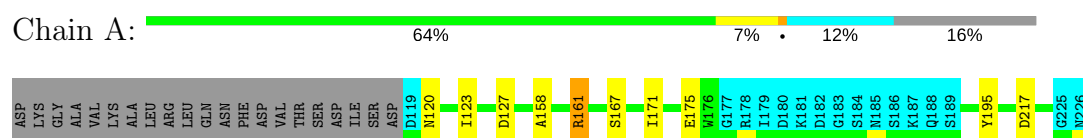
### 4.2.11 Score per residue for model 11

- Molecule 1: Nisin immunity protein



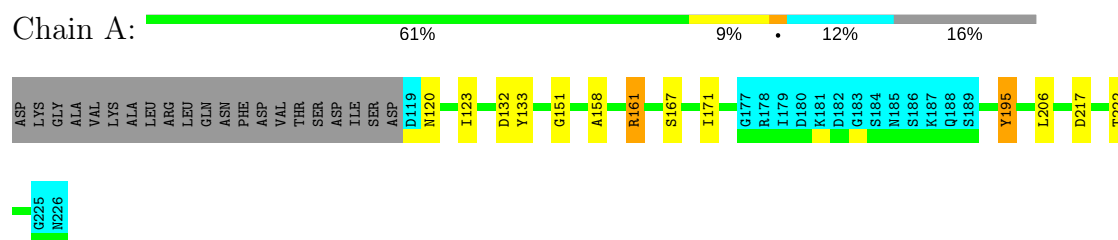
### 4.2.12 Score per residue for model 12

- Molecule 1: Nisin immunity protein



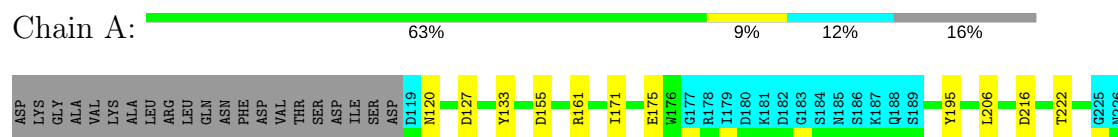
### 4.2.13 Score per residue for model 13

- Molecule 1: Nisin immunity protein



### 4.2.14 Score per residue for model 14

- Molecule 1: Nisin immunity protein



### 4.2.15 Score per residue for model 15

- Molecule 1: Nisin immunity protein



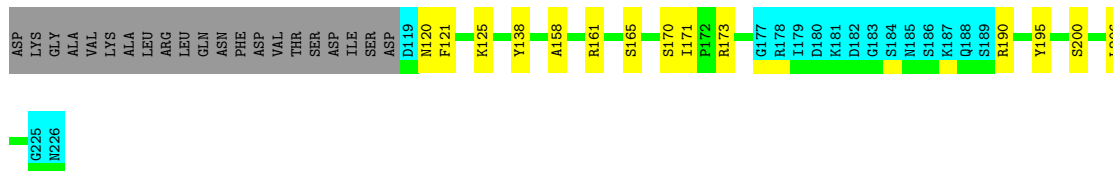




#### 4.2.16 Score per residue for model 16

- Molecule 1: Nisin immunity protein

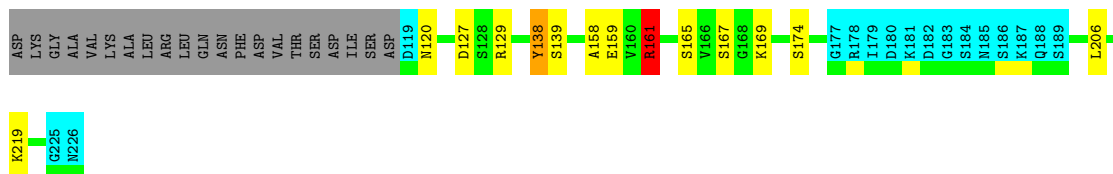
Chain A: 60% 11% 12% 16%



#### 4.2.17 Score per residue for model 17

- Molecule 1: Nisin immunity protein

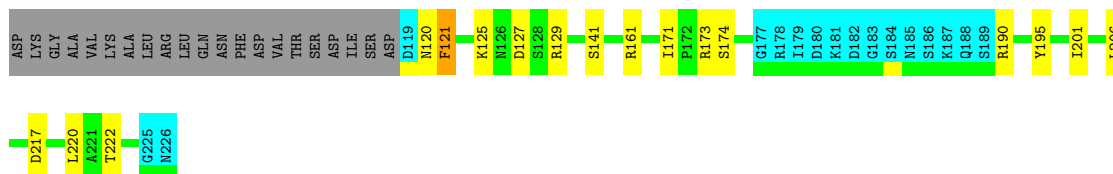
Chain A: 60% 9% 12% 16%



#### 4.2.18 Score per residue for model 18

- Molecule 1: Nisin immunity protein

Chain A: 58% 12% 12% 16%

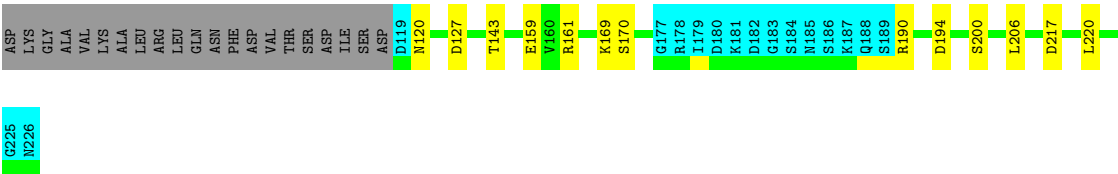


#### 4.2.19 Score per residue for model 19

- Molecule 1: Nisin immunity protein

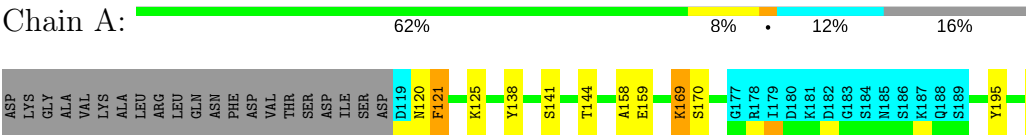
Chain A: 61% 10% 12% 16%





4.2.20 Score per residue for model 20

- Molecule 1: Nisin immunity protein



## 5 Refinement protocol and experimental data overview

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

Of the 100 calculated structures, 20 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
TOPSPIN	structure solution	
TOPSPIN	structure solution	
TOPSPIN	structure solution	
TOPSPIN	structure solution	
CYANA	refinement	
OPALp	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)	2n2e_cs.str
Number of chemical shift lists	1
Total number of shifts	1511
Number of shifts mapped to atoms	1511
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	88%

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

## 6 Model quality i

### 6.1 Standard geometry i

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	0.59±0.09	0±0/757 (0.0±0.0%)	1.06±0.17	1±1/1024 (0.1±0.1%)
All	All	0.59	0/15140 (0.0%)	1.07	19/20480 (0.1%)

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	1.3±1.1
All	All	0	26

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(°)	Ideal(°)	Models	
								Worst	Total
1	A	195	TYR	CB-CG-CD2	-6.70	116.98	121.00	7	4
1	A	173	ARG	NE-CZ-NH2	-6.40	117.10	120.30	16	2
1	A	195	TYR	C-N-CA	6.22	135.35	122.30	9	6
1	A	161	ARG	NE-CZ-NH2	-5.80	117.40	120.30	17	1
1	A	167	SER	N-CA-CB	-5.79	101.81	110.50	17	1
1	A	190	ARG	CD-NE-CZ	5.75	131.66	123.60	8	1
1	A	129	ARG	NE-CZ-NH2	-5.69	117.45	120.30	4	1
1	A	202	ARG	NE-CZ-NH2	-5.60	117.50	120.30	3	1
1	A	129	ARG	CD-NE-CZ	5.51	131.31	123.60	1	1
1	A	138	TYR	CB-CG-CD1	-5.39	117.76	121.00	17	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	161	ARG	Sidechain	5
1	A	121	PHE	Sidechain	4
1	A	138	TYR	Sidechain	4
1	A	202	ARG	Sidechain	4
1	A	190	ARG	Sidechain	4
1	A	129	ARG	Sidechain	2
1	A	133	TYR	Sidechain	2
1	A	173	ARG	Sidechain	1

## 6.2 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	743	708	708	1±2
All	All	14860	14160	14160	18

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:123:ILE:HG21	1:A:195:TYR:OH	0.62	1.94	2	1
1:A:129:ARG:O	1:A:140:ILE:HD12	0.61	1.94	2	1
1:A:121:PHE:CD2	1:A:198:ILE:HD13	0.53	2.37	2	1
1:A:210:PHE:O	1:A:220:LEU:HD22	0.53	2.02	2	1
1:A:164:ASP:CG	1:A:166:VAL:HG22	0.49	2.27	2	2
1:A:128:SER:HB2	1:A:163:PHE:CD1	0.48	2.44	8	1
1:A:123:ILE:HD13	1:A:195:TYR:CE2	0.45	2.46	13	2
1:A:201:ILE:CD1	1:A:220:LEU:HD23	0.44	2.43	18	1
1:A:143:THR:HB	1:A:220:LEU:HB3	0.43	1.91	4	3
1:A:166:VAL:HG23	1:A:167:SER:N	0.42	2.30	2	1
1:A:206:LEU:HD13	1:A:210:PHE:CG	0.42	2.50	2	1
1:A:169:LYS:H	1:A:169:LYS:HD3	0.42	1.75	20	1
1:A:121:PHE:N	1:A:121:PHE:CD1	0.40	2.90	18	1
1:A:121:PHE:CD1	1:A:121:PHE:N	0.40	2.89	20	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	92/129 (71%)	86±1 (93±1%)	6±1 (6±2%)	0±1 (0±1%)	37	78
All	All	1840/2580 (71%)	1717 (93%)	114 (6%)	9 (0%)	37	78

All 3 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	158	ALA	7
1	A	151	GLY	1
1	A	141	SER	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	83/114 (73%)	74±2 (90±2%)	9±2 (10±2%)	12	56
All	All	1660/2280 (73%)	1488 (90%)	172 (10%)	12	56

All 35 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	120	ASN	20
1	A	206	LEU	18
1	A	171	ILE	15
1	A	127	ASP	13
1	A	161	ARG	11
1	A	175	GLU	10
1	A	169	LYS	9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	222	THR	7
1	A	125	LYS	7
1	A	217	ASP	6
1	A	165	SER	6
1	A	159	GLU	5
1	A	167	SER	4
1	A	155	ASP	4
1	A	170	SER	4
1	A	174	SER	3
1	A	200	SER	3
1	A	132	ASP	3
1	A	154	GLN	3
1	A	190	ARG	2
1	A	216	ASP	2
1	A	194	ASP	2
1	A	150	LEU	2
1	A	219	LYS	2
1	A	129	ARG	1
1	A	164	ASP	1
1	A	139	SER	1
1	A	144	THR	1
1	A	215	ASN	1
1	A	191	THR	1
1	A	141	SER	1
1	A	153	TYR	1
1	A	197	GLU	1
1	A	138	TYR	1
1	A	147	ASP	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 carbohydrates in this entry.

## 6.6 Ligand geometry

There are no ligands in this entry.

## 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 88% for the well-defined parts and 88% for the entire structure.

### 7.1 Chemical shift list 1

File name: 2n2e\_cs.str

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	1511
Number of shifts mapped to atoms	1511
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	6

#### 7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	129	$-0.17 \pm 0.22$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}_\beta$	120	$-0.14 \pm 0.17$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}'$	128	$0.02 \pm 0.09$	None needed ( $< 0.5$ ppm)
$^{15}\text{N}$	127	$-1.18 \pm 0.37$	Should be applied

#### 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 88%, i.e. 993 atoms were assigned a chemical shift out of a possible 1130. 10 out of 12 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	458/458 (100%)	183/183 (100%)	184/184 (100%)	91/91 (100%)
Sidechain	450/573 (79%)	278/332 (84%)	167/215 (78%)	5/26 (19%)

*Continued on next page...*

Continued from previous page...

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	85/99 (86%)	43/52 (83%)	40/44 (91%)	2/3 (67%)
Overall	993/1130 (88%)	504/567 (89%)	391/443 (88%)	98/120 (82%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 88%, i.e. 1151 atoms were assigned a chemical shift out of a possible 1308. 10 out of 12 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Backbone	538/538 (100%)	215/215 (100%)	216/216 (100%)	107/107 (100%)
Sidechain	528/671 (79%)	329/390 (84%)	192/247 (78%)	7/34 (21%)
Aromatic	85/99 (86%)	43/52 (83%)	40/44 (91%)	2/3 (67%)
Overall	1151/1308 (88%)	587/657 (89%)	448/507 (88%)	116/144 (81%)

#### 7.1.4 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	129	ARG	NE	117.99	92.63 – 76.73	20.9
1	A	213	GLU	HG2	0.76	3.33 – 1.23	-7.2
1	A	191	THR	HB	2.04	5.82 – 2.52	-6.4
1	A	214	ILE	HG23	-0.85	2.13 – -0.57	-6.0
1	A	214	ILE	HG21	-0.85	2.13 – -0.57	-6.0
1	A	214	ILE	HG22	-0.85	2.13 – -0.57	-6.0

#### 7.1.5 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:

