



# Full wwPDB NMR Structure Validation Report ⓘ

May 29, 2020 – 12:27 am BST

PDB ID : 2MR3  
Title : A subunit of 26S proteasome lid complex  
Authors : Wu, Y.; Hu, Y.; Jin, C.  
Deposited on : 2014-06-30

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

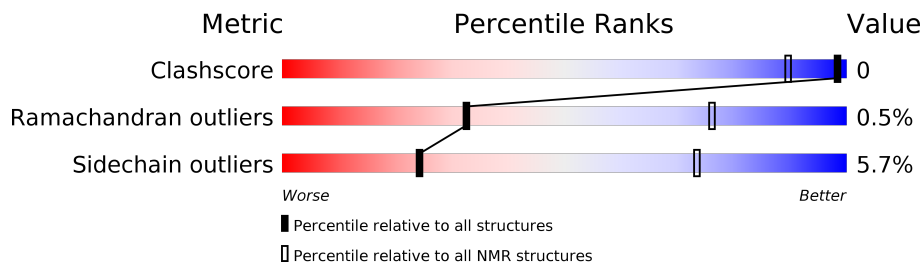
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 78%.

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	393	

## 2 Ensemble composition and analysis

This entry contains 20 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:6-A:110, A:124-A:356 (338)	1.58	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 4 clusters and 5 single-model clusters were found.

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 5868 atoms, of which 2940 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called 26S proteasome regulatory subunit RPN9.

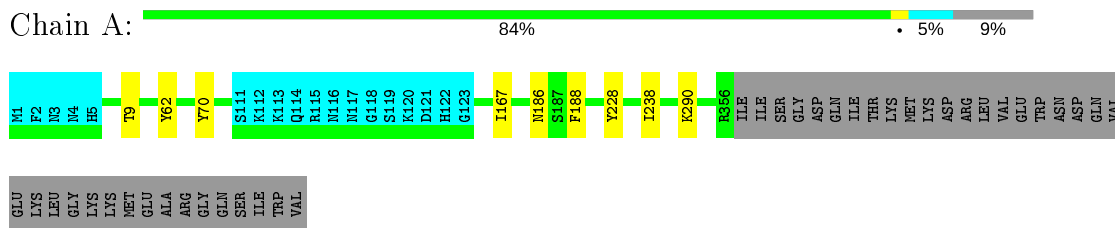
Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	356	5868	1889	2940	474	558	7	0

## 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: 26S proteasome regulatory subunit RPN9

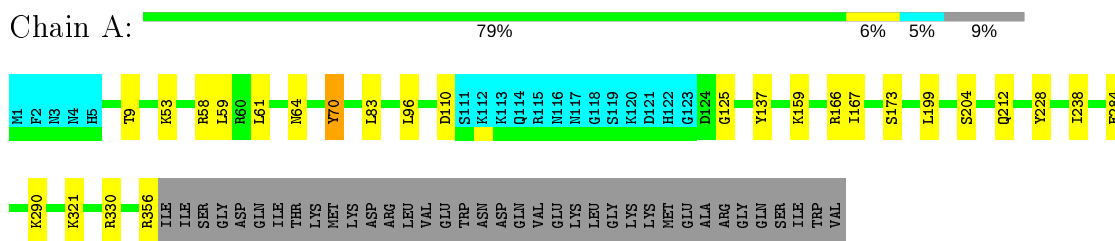


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

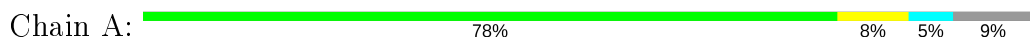
#### 4.2.1 Score per residue for model 1

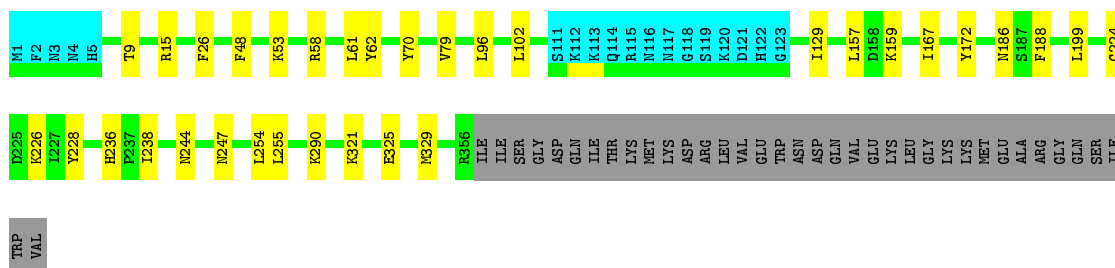
- Molecule 1: 26S proteasome regulatory subunit RPN9



#### 4.2.2 Score per residue for model 2

- Molecule 1: 26S proteasome regulatory subunit RPN9

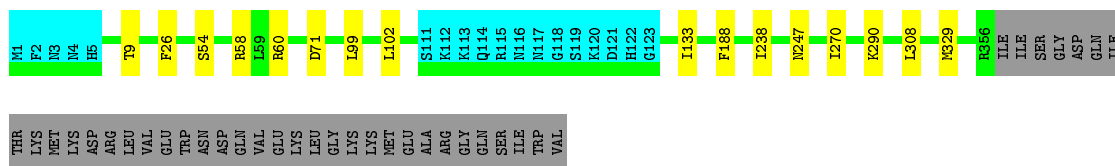




### 4.2.3 Score per residue for model 3

- Molecule 1: 26S proteasome regulatory subunit RPN9

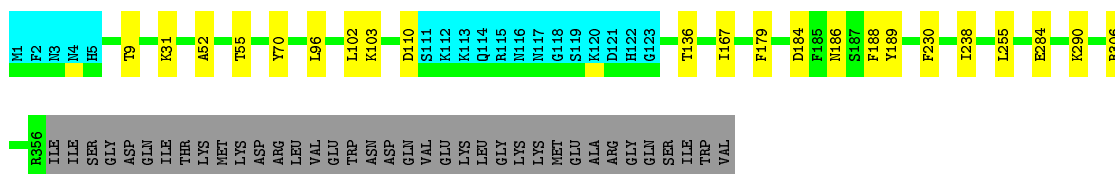
Chain A: 82% 5% 9%



### 4.2.4 Score per residue for model 4

- Molecule 1: 26S proteasome regulatory subunit RPN9

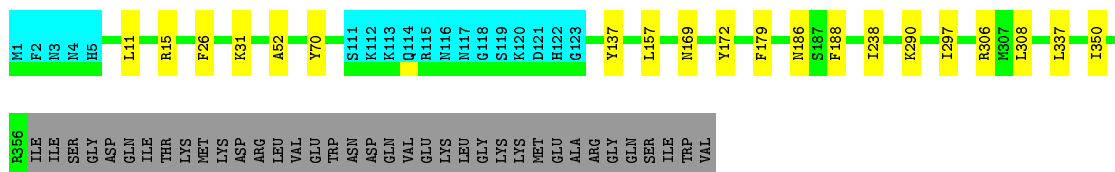
Chain A: 80% 6% 5% 9%



### 4.2.5 Score per residue for model 5

- Molecule 1: 26S proteasome regulatory subunit RPN9

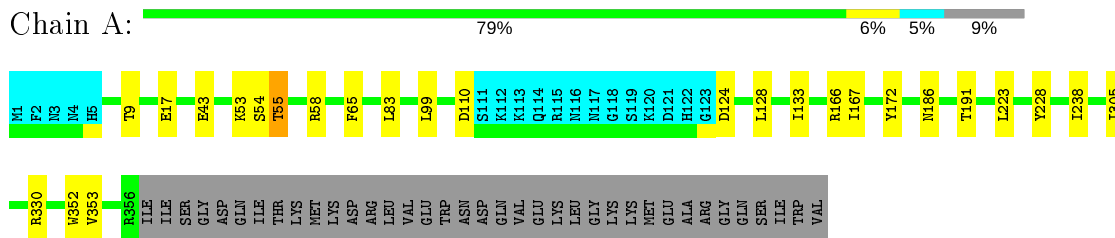
Chain A: 81% 5% 5% 9%





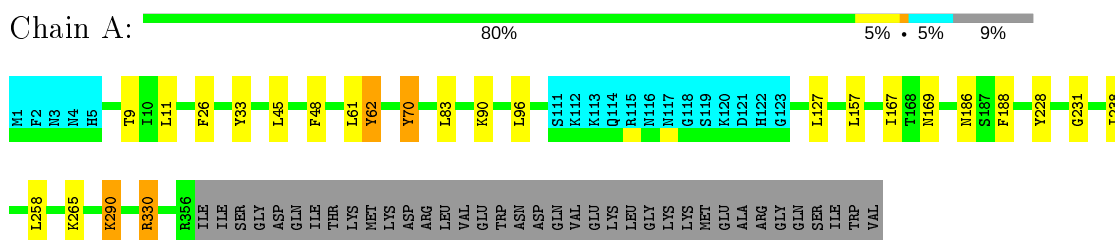
#### 4.2.10 Score per residue for model 10

- Molecule 1: 26S proteasome regulatory subunit RPN9



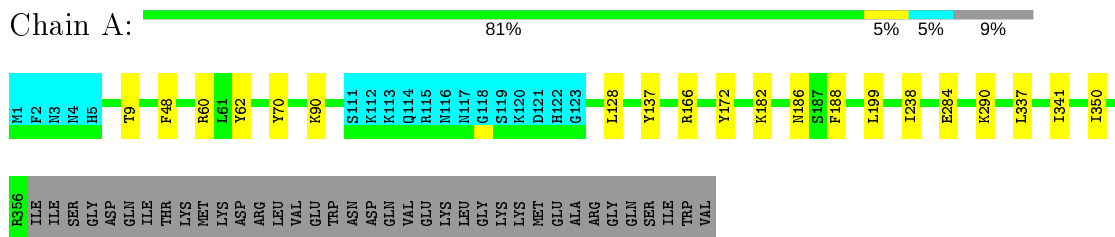
#### 4.2.11 Score per residue for model 11

- Molecule 1: 26S proteasome regulatory subunit RPN9



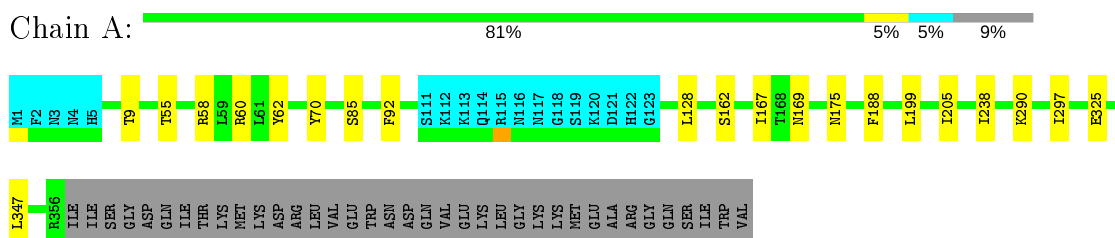
#### 4.2.12 Score per residue for model 12

- Molecule 1: 26S proteasome regulatory subunit RPN9



#### 4.2.13 Score per residue for model 13

- Molecule 1: 26S proteasome regulatory subunit RPN9

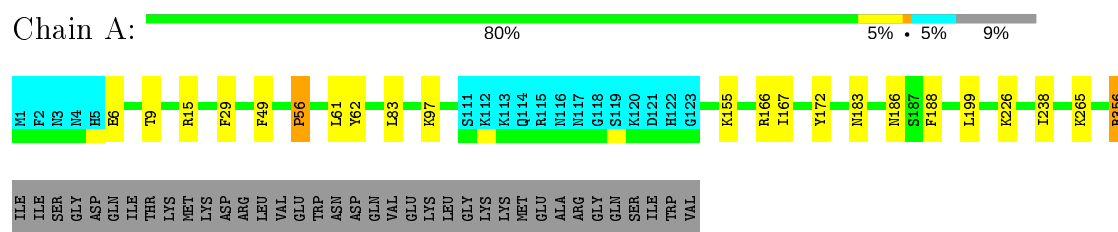






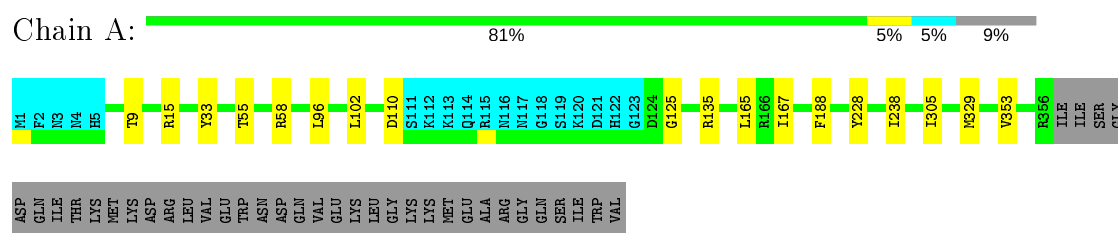
### 4.2.18 Score per residue for model 18

- Molecule 1: 26S proteasome regulatory subunit RPN9



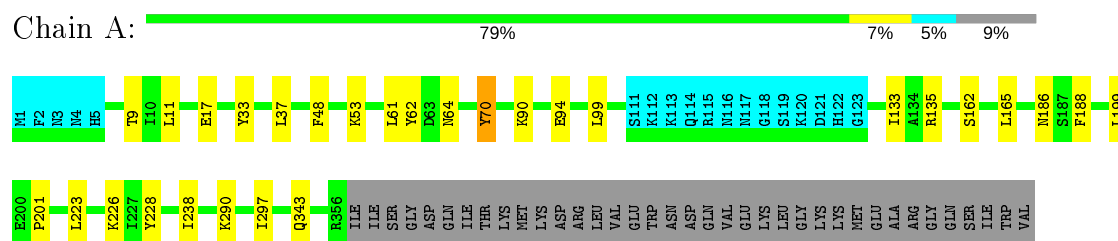
### 4.2.19 Score per residue for model 19

- Molecule 1: 26S proteasome regulatory subunit RPN9



### 4.2.20 Score per residue for model 20

- Molecule 1: 26S proteasome regulatory subunit RPN9



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

The models were refined using the following method: *simulated annealing*.

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
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	4165
Number of shifts mapped to atoms	4165
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	78%

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	2782	2802	2802	1±1
All	All	55640	56040	56040	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 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:172:TYR:CD1	1:A:191:THR:HG23	0.49	2.43	10	1
1:A:99:LEU:HB2	1:A:133:ILE:HD11	0.49	1.85	3	4
1:A:99:LEU:CB	1:A:133:ILE:HD11	0.46	2.40	20	3
1:A:228:TYR:CE1	1:A:290:LYS:HE3	0.44	2.48	11	1
1:A:337:LEU:CD1	1:A:350:ILE:HD12	0.43	2.44	12	2
1:A:96:LEU:HD21	1:A:136:THR:HB	0.41	1.93	4	1
1:A:49:PHE:CD2	1:A:84:ALA:HB2	0.41	2.50	6	1
1:A:228:TYR:CZ	1:A:290:LYS:HE3	0.41	2.50	20	1
1:A:79:VAL:HG21	1:A:129:ILE:HD13	0.41	1.91	2	1
1:A:228:TYR:CE2	1:A:290:LYS:HE3	0.41	2.51	15	1
1:A:26:PHE:CE1	1:A:29:PHE:CD2	0.41	3.09	6	1
1:A:169:ASN:HA	1:A:172:TYR:CE2	0.40	2.51	17	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	337/393 (86%)	316±3 (94±1%)	19±3 (6±1%)	2±1 (1±0%)	32	76
All	All	6740/7860 (86%)	6328 (94%)	376 (6%)	36 (1%)	32	76

All 22 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	52	ALA	4
1	A	55	THR	3
1	A	110	ASP	3
1	A	162	SER	3
1	A	228	TYR	2
1	A	56	PRO	2
1	A	125	GLY	2
1	A	352	TRP	2
1	A	71	ASP	2
1	A	201	PRO	1
1	A	53	LYS	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	237	PRO	1
1	A	231	GLY	1
1	A	158	ASP	1
1	A	159	LYS	1
1	A	54	SER	1
1	A	230	PHE	1
1	A	224	GLY	1
1	A	205	ILE	1
1	A	179	PHE	1
1	A	6	GLU	1
1	A	92	PHE	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	319/368 (87%)	301±4 (94±1%)	18±4 (6±1%)	24	73
All	All	6380/7360 (87%)	6014 (94%)	366 (6%)	24	73

All 115 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	238	ILE	20
1	A	9	THR	18
1	A	188	PHE	13
1	A	167	ILE	13
1	A	186	ASN	13
1	A	290	LYS	11
1	A	70	TYR	9
1	A	53	LYS	8
1	A	31	LYS	7
1	A	169	ASN	7
1	A	96	LEU	7
1	A	199	LEU	7
1	A	258	LEU	6
1	A	61	LEU	6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	48	PHE	6
1	A	58	ARG	5
1	A	128	LEU	5
1	A	62	TYR	5
1	A	102	LEU	5
1	A	90	LYS	5
1	A	11	LEU	5
1	A	157	LEU	5
1	A	83	LEU	4
1	A	247	ASN	4
1	A	329	MET	4
1	A	305	ILE	4
1	A	55	THR	4
1	A	49	PHE	4
1	A	255	LEU	4
1	A	297	ILE	4
1	A	103	LYS	4
1	A	26	PHE	4
1	A	59	LEU	3
1	A	284	GLU	3
1	A	226	LYS	3
1	A	265	LYS	3
1	A	137	TYR	3
1	A	330	ARG	3
1	A	64	ASN	3
1	A	353	VAL	3
1	A	325	GLU	3
1	A	323	ASN	3
1	A	147	ARG	3
1	A	166	ARG	3
1	A	33	TYR	3
1	A	244	ASN	3
1	A	127	LEU	3
1	A	156	THR	2
1	A	37	LEU	2
1	A	194	LEU	2
1	A	155	LYS	2
1	A	135	ARG	2
1	A	184	ASP	2
1	A	321	LYS	2
1	A	57	LEU	2
1	A	228	TYR	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	43	GLU	2
1	A	54	SER	2
1	A	314	SER	2
1	A	306	ARG	2
1	A	165	LEU	2
1	A	131	SER	2
1	A	86	LEU	2
1	A	356	ARG	2
1	A	124	ASP	2
1	A	343	GLN	2
1	A	326	HIS	2
1	A	270	ILE	2
1	A	10	ILE	2
1	A	347	LEU	2
1	A	85	SER	2
1	A	60	ARG	2
1	A	308	LEU	2
1	A	223	LEU	2
1	A	133	ILE	1
1	A	126	ILE	1
1	A	150	LEU	1
1	A	274	ILE	1
1	A	187	SER	1
1	A	319	LEU	1
1	A	97	LYS	1
1	A	15	ARG	1
1	A	179	PHE	1
1	A	78	VAL	1
1	A	149	LEU	1
1	A	110	ASP	1
1	A	29	PHE	1
1	A	183	ASN	1
1	A	163	ILE	1
1	A	159	LYS	1
1	A	94	GLU	1
1	A	341	ILE	1
1	A	182	LYS	1
1	A	22	LEU	1
1	A	205	ILE	1
1	A	38	TRP	1
1	A	254	LEU	1
1	A	212	GLN	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	41	LEU	1
1	A	153	LEU	1
1	A	171	PHE	1
1	A	172	TYR	1
1	A	140	LYS	1
1	A	173	SER	1
1	A	28	GLN	1
1	A	25	LEU	1
1	A	204	SER	1
1	A	168	THR	1
1	A	203	THR	1
1	A	175	ASN	1
1	A	45	LEU	1
1	A	236	HIS	1
1	A	289	GLN	1
1	A	92	PHE	1
1	A	65	PHE	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 78% for the well-defined parts and 77% 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	4165
Number of shifts mapped to atoms	4165
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	5

#### 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
$^{13}\text{C}_\alpha$	383	$-0.64 \pm 0.07$	Should be applied
$^{13}\text{C}_\beta$	371	$0.41 \pm 0.05$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}'$	344	$-0.32 \pm 0.13$	None needed ( $< 0.5$ ppm)
$^{15}\text{N}$	370	$0.15 \pm 0.13$	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 78%, i.e. 3408 atoms were assigned a chemical shift out of a possible 4356. 0 out of 69 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	1653/1672 (99%)	662/667 (99%)	663/676 (98%)	328/329 (100%)
Sidechain	1677/2296 (73%)	970/1338 (72%)	687/870 (79%)	20/88 (23%)

*Continued on next page...*

Continued from previous page...

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	78/388 (20%)	43/208 (21%)	32/171 (19%)	3/9 (33%)
Overall	3408/4356 (78%)	1675/2213 (76%)	1382/1717 (80%)	351/426 (82%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 77%, i.e. 3534 atoms were assigned a chemical shift out of a possible 4591. 0 out of 69 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Backbone	1723/1762 (98%)	691/703 (98%)	691/712 (97%)	341/347 (98%)
Sidechain	1731/2418 (72%)	1003/1412 (71%)	708/907 (78%)	20/99 (20%)
Aromatic	80/411 (19%)	45/221 (20%)	32/179 (18%)	3/11 (27%)
Overall	3534/4591 (77%)	1739/2336 (74%)	1431/1798 (80%)	364/457 (80%)

#### 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	A	40	GLN	HB3	0.43	3.37 – 0.67	-5.9
1	A	20	PRO	HA	2.53	6.05 – 2.75	-5.7
1	A	346	GLU	HA	2.03	6.30 – 2.20	-5.4
1	A	36	LYS	HA	1.94	6.46 – 2.06	-5.3
1	A	238	ILE	H	11.83	11.73 – 4.83	5.1

#### 6.1.5 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:

