



# Full wwPDB NMR Structure Validation Report ⓘ

May 29, 2020 – 12:20 am BST

PDB ID : 2N6D  
Title : NMR structure of the 140-315 fragment of the N-acetylglucosamine-1-phosphate transferase, alpha and beta subunits  
Authors : Serrano, P.; Geralt, M.; Wuthrich, K.; Joint Center for Structural Genomics (JCSG)  
Deposited on : 2015-08-19

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

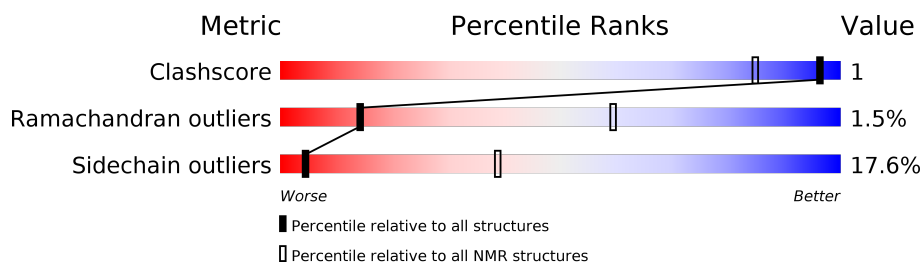
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*


The overall completeness of chemical shifts assignment is 85%.

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	172	 90% 7% •

## 2 Ensemble composition and analysis

This entry contains 19 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:1-A:147, A:154-A:172 (166)	0.88	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 1 single-model cluster was found.

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 2736 atoms, of which 1389 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called N-acetylglucosamine-1-phosphotransferase subunits alpha/beta.

Mol	Chain	Residues	Atoms						Trace
1	A	172	Total	C	H	N	O	S	0
			2736	868	1389	219	256	4	

There is a discrepancy between the modelled and reference sequences:

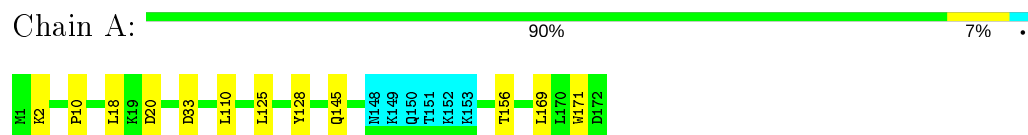
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	INITIATING METHIONINE	UNP Q3T906

## 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: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta

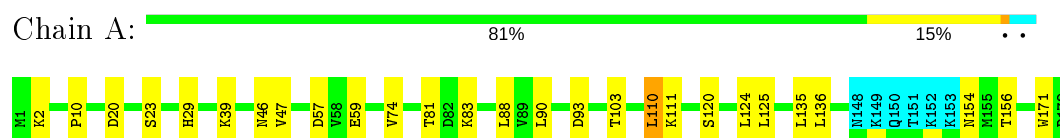


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

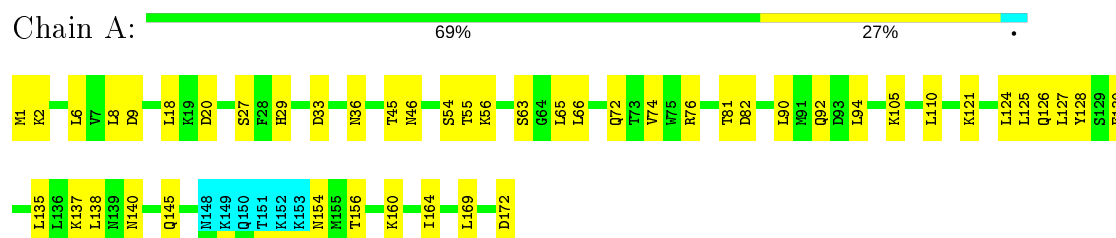
#### 4.2.1 Score per residue for model 1

- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta



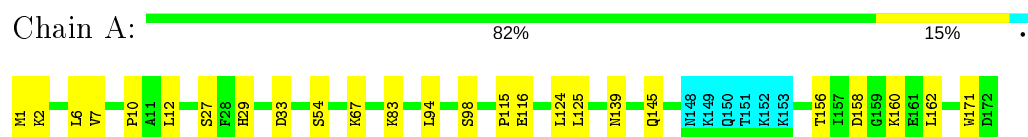
#### 4.2.2 Score per residue for model 2 (medoid)

- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta



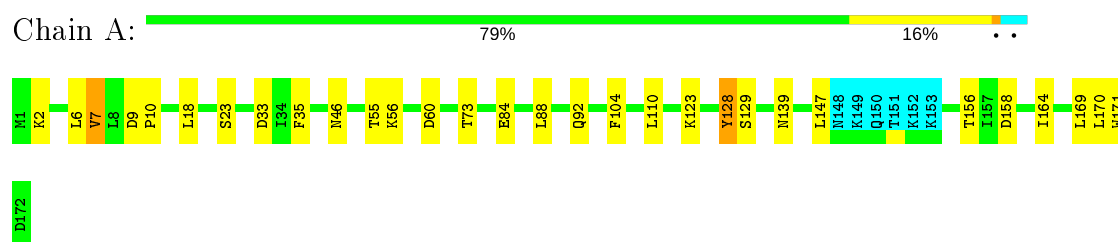
### 4.2.3 Score per residue for model 3

- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta



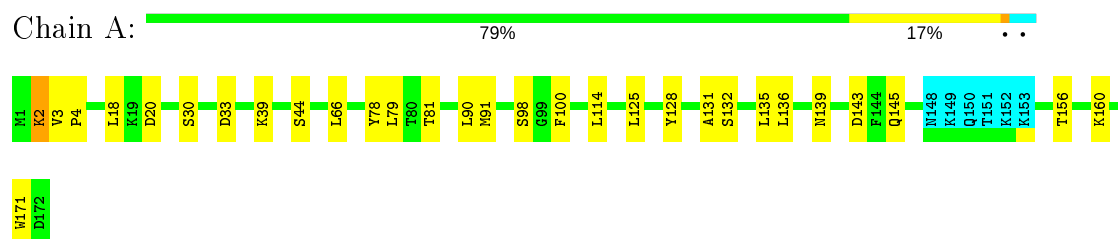
### 4.2.4 Score per residue for model 4

- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta



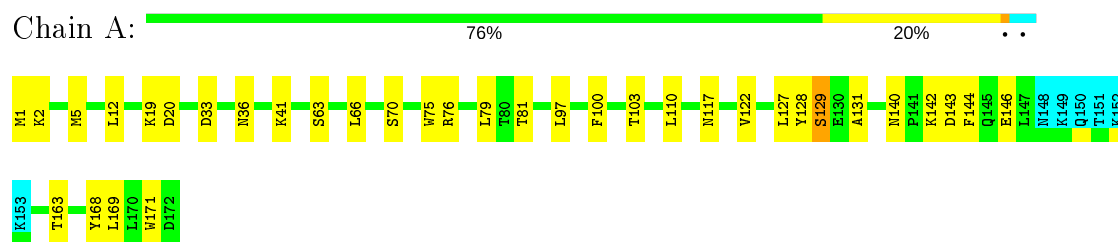
### 4.2.5 Score per residue for model 5

- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta



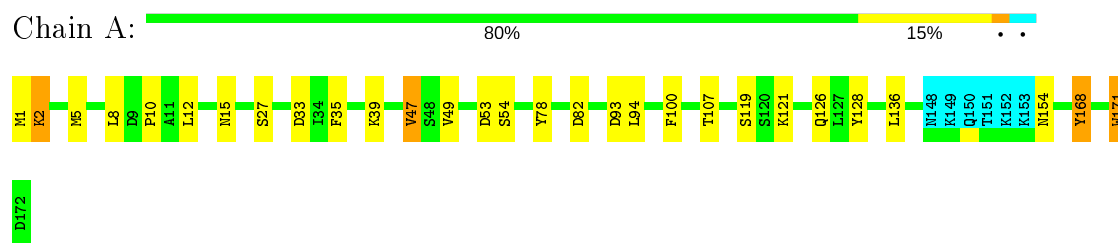
### 4.2.6 Score per residue for model 6

- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta



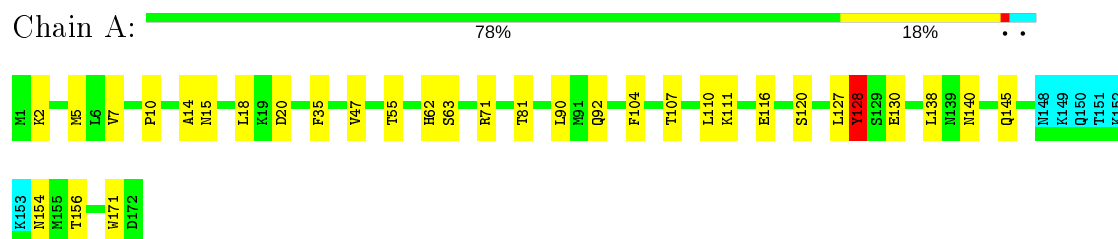
### 4.2.7 Score per residue for model 7

- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta



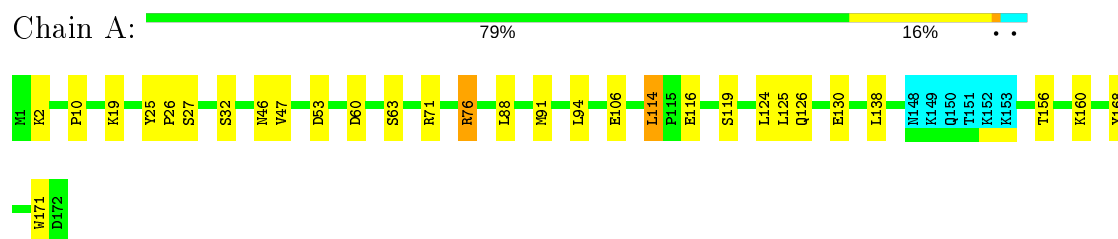
### 4.2.8 Score per residue for model 8

- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta



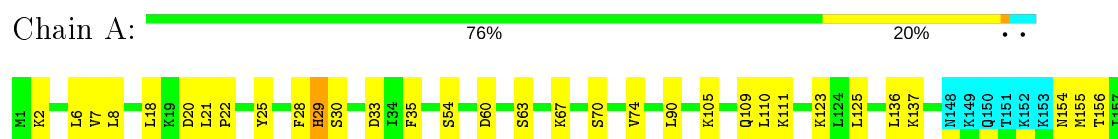
### 4.2.9 Score per residue for model 9

- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta



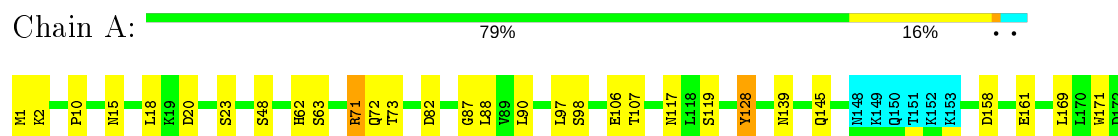
### 4.2.10 Score per residue for model 10

- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta



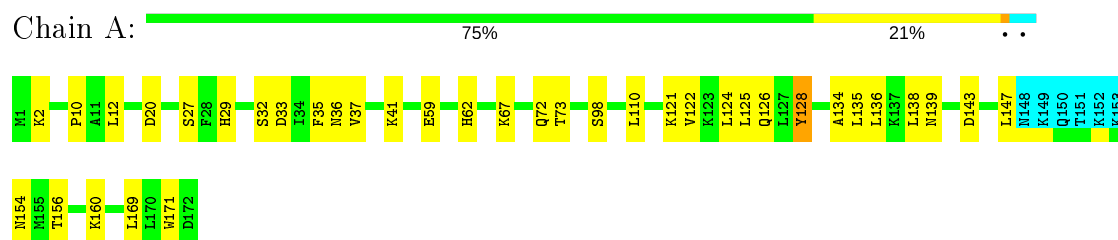
#### 4.2.11 Score per residue for model 11

- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta



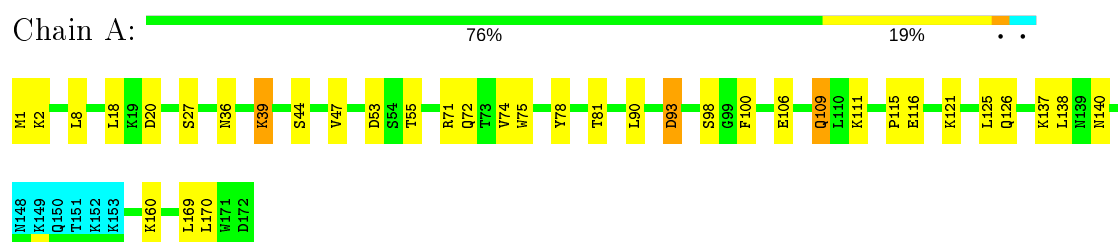
#### 4.2.12 Score per residue for model 12

- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta



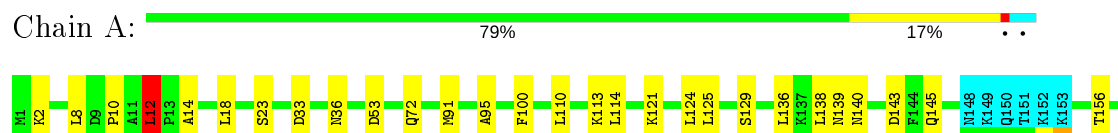
#### 4.2.13 Score per residue for model 13

- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta



#### 4.2.14 Score per residue for model 14

- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta

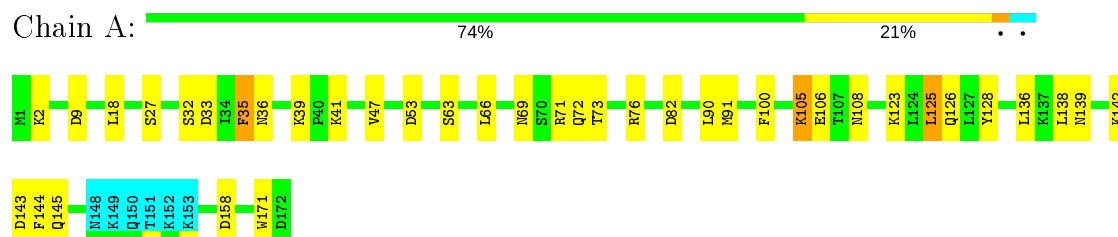






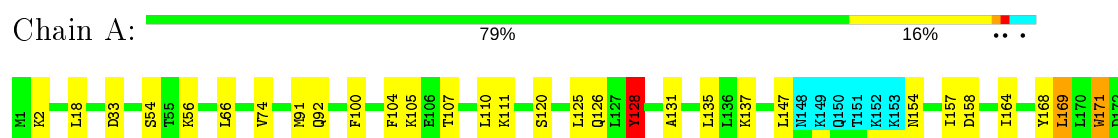
#### 4.2.15 Score per residue for model 15

- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta



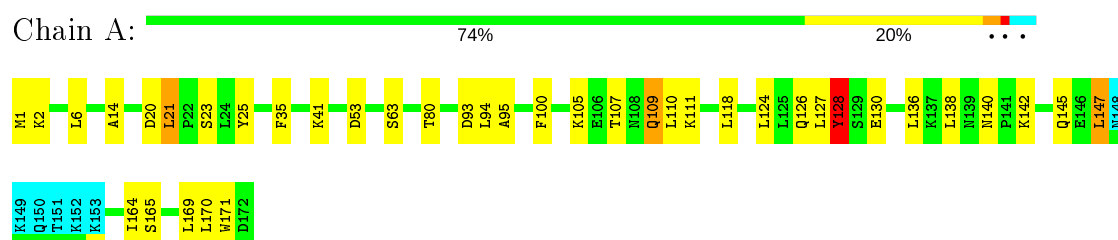
#### 4.2.16 Score per residue for model 16

- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta



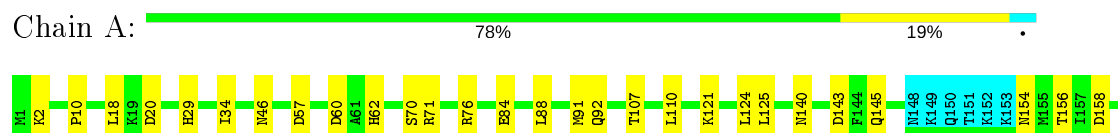
#### 4.2.17 Score per residue for model 17

- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta



#### 4.2.18 Score per residue for model 18

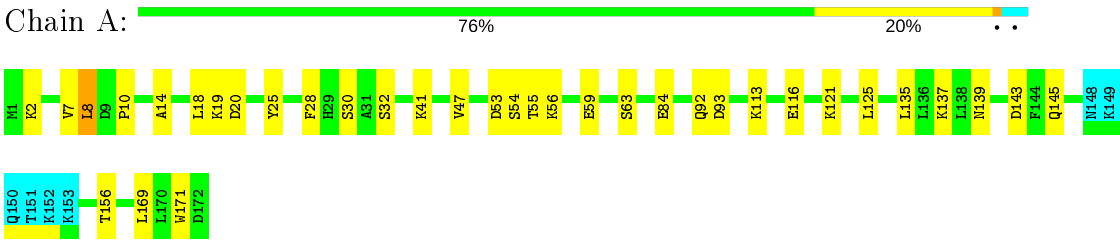
- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta





4.2.19 Score per residue for model 19

- Molecule 1: N-acetylglucosamine-1-phosphotransferase subunits alpha/beta



## 5 Refinement protocol and experimental data overview

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

Of the 80 calculated structures, 19 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	

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	2041
Number of shifts mapped to atoms	2041
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	85%

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	1296	1329	1331	2±2
All	All	24624	25251	25289	38

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:110:LEU:HD11	1:A:124:LEU:HD22	0.65	1.69	1	1
1:A:114:LEU:HD12	1:A:114:LEU:C	0.57	2.20	9	1
1:A:114:LEU:HD12	1:A:116:GLU:H	0.53	1.62	9	1
1:A:3:VAL:HG12	1:A:79:LEU:H	0.53	1.64	5	1
1:A:12:LEU:HD22	1:A:12:LEU:C	0.52	2.25	14	1
1:A:118:LEU:HD13	1:A:147:LEU:HD13	0.51	1.82	17	1
1:A:94:LEU:HD21	1:A:135:LEU:HD12	0.50	1.82	2	1
1:A:6:LEU:HD12	1:A:28:PHE:CE2	0.49	2.42	10	1
1:A:88:LEU:HD12	1:A:168:TYR:CG	0.48	2.43	9	1
1:A:39:LYS:HE2	1:A:47:VAL:CG1	0.48	2.38	15	3
1:A:122:VAL:HG12	1:A:138:LEU:HA	0.46	1.87	12	1
1:A:138:LEU:HD22	1:A:144:PHE:CD2	0.46	2.46	15	1
1:A:122:VAL:HG11	1:A:144:PHE:CE1	0.46	2.46	6	1
1:A:5:MET:SD	1:A:79:LEU:HD23	0.45	2.51	6	1
1:A:104:PHE:CZ	1:A:157:ILE:HG21	0.45	2.47	16	1
1:A:2:LYS:C	1:A:2:LYS:HE2	0.44	2.33	7	1
1:A:127:LEU:HD23	1:A:128:TYR:N	0.43	2.28	17	2
1:A:80:THR:HG22	1:A:170:LEU:HD23	0.43	1.90	17	1
1:A:39:LYS:HE3	1:A:75:TRP:CH2	0.42	2.50	13	1
1:A:21:LEU:HD23	1:A:22:PRO:N	0.42	2.28	10	1
1:A:128:TYR:N	1:A:128:TYR:CD1	0.42	2.87	16	1
1:A:114:LEU:HD22	1:A:119:SER:CB	0.41	2.45	9	1
1:A:114:LEU:CD1	1:A:116:GLU:H	0.41	2.27	9	1
1:A:124:LEU:HD21	1:A:134:ALA:HB1	0.41	1.92	12	1
1:A:88:LEU:C	1:A:88:LEU:HD12	0.41	2.36	1	1
1:A:95:ALA:HB1	1:A:164:ILE:HD11	0.41	1.92	17	1
1:A:35:PHE:CD1	1:A:125:LEU:HD13	0.41	2.51	15	1
1:A:100:PHE:CZ	1:A:128:TYR:CZ	0.41	3.09	16	1
1:A:21:LEU:C	1:A:21:LEU:HD12	0.41	2.37	17	1
1:A:21:LEU:O	1:A:21:LEU:HD12	0.40	2.16	17	1
1:A:35:PHE:CZ	1:A:49:VAL:HG11	0.40	2.51	7	1
1:A:169:LEU:HD11	1:A:171:TRP:CE3	0.40	2.51	16	1
1:A:8:LEU:HD11	1:A:28:PHE:CE2	0.40	2.51	19	1
1:A:2:LYS:HE3	1:A:78:TYR:CD1	0.40	2.51	5	1
1:A:21:LEU:HD23	1:A:21:LEU:C	0.40	2.37	10	1

## 5.2 Torsion angles ⓘ

### 5.2.1 Protein backbone ⓘ

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	164/172 (95%)	139±4 (85±3%)	22±4 (13±3%)	3±1 (2±1%)	14	59
All	All	3116/3268 (95%)	2649 (85%)	419 (13%)	48 (2%)	14	59

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	47	VAL	5
1	A	154	ASN	4
1	A	14	ALA	4
1	A	130	GLU	4
1	A	93	ASP	4
1	A	109	GLN	3
1	A	116	GLU	3
1	A	131	ALA	3
1	A	129	SER	2
1	A	115	PRO	2
1	A	103	THR	2
1	A	143	ASP	2
1	A	84	GLU	1
1	A	26	PRO	1
1	A	142	LYS	1
1	A	29	HIS	1
1	A	4	PRO	1
1	A	91	MET	1
1	A	88	LEU	1
1	A	105	LYS	1
1	A	7	VAL	1
1	A	23	SER	1

### 5.2.2 Protein sidechains ⓘ

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	151/157 (96%)	124±5 (82±3%)	27±5 (18±3%)	4	39

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2869/2983 (96%)	2363 (82%)	506 (18%)	4 39

All 109 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	2	LYS	19
1	A	171	TRP	17
1	A	125	LEU	13
1	A	156	THR	12
1	A	18	LEU	12
1	A	20	ASP	12
1	A	110	LEU	11
1	A	33	ASP	11
1	A	169	LEU	10
1	A	145	GLN	10
1	A	63	SER	9
1	A	136	LEU	8
1	A	126	GLN	8
1	A	139	ASN	8
1	A	90	LEU	8
1	A	100	PHE	7
1	A	27	SER	7
1	A	158	ASP	7
1	A	53	ASP	7
1	A	140	ASN	7
1	A	1	MET	7
1	A	8	LEU	6
1	A	81	THR	6
1	A	138	LEU	6
1	A	35	PHE	6
1	A	72	GLN	6
1	A	107	THR	6
1	A	121	LYS	6
1	A	29	HIS	6
1	A	36	ASN	6
1	A	54	SER	6
1	A	111	LYS	6
1	A	160	LYS	6
1	A	92	GLN	6
1	A	124	LEU	6
1	A	55	THR	5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	12	LEU	5
1	A	105	LYS	5
1	A	66	LEU	5
1	A	74	VAL	5
1	A	7	VAL	5
1	A	91	MET	5
1	A	128	TYR	5
1	A	135	LEU	5
1	A	137	LYS	5
1	A	46	ASN	5
1	A	143	ASP	5
1	A	98	SER	5
1	A	154	ASN	4
1	A	6	LEU	4
1	A	82	ASP	4
1	A	23	SER	4
1	A	71	ARG	4
1	A	32	SER	4
1	A	73	THR	4
1	A	94	LEU	4
1	A	106	GLU	4
1	A	56	LYS	4
1	A	62	HIS	4
1	A	41	LYS	4
1	A	9	ASP	3
1	A	170	LEU	3
1	A	120	SER	3
1	A	60	ASP	3
1	A	59	GLU	3
1	A	70	SER	3
1	A	19	LYS	3
1	A	123	LYS	3
1	A	30	SER	3
1	A	164	ILE	3
1	A	147	LEU	3
1	A	172	ASP	3
1	A	163	THR	3
1	A	15	ASN	3
1	A	67	LYS	3
1	A	76	ARG	3
1	A	113	LYS	2
1	A	109	GLN	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	104	PHE	2
1	A	57	ASP	2
1	A	117	ASN	2
1	A	25	TYR	2
1	A	168	TYR	2
1	A	97	LEU	2
1	A	83	LYS	2
1	A	39	LYS	2
1	A	5	MET	2
1	A	129	SER	2
1	A	44	SER	2
1	A	119	SER	2
1	A	84	GLU	2
1	A	142	LYS	2
1	A	127	LEU	2
1	A	88	LEU	2
1	A	69	ASN	1
1	A	48	SER	1
1	A	114	LEU	1
1	A	21	LEU	1
1	A	108	ASN	1
1	A	34	ILE	1
1	A	37	VAL	1
1	A	146	GLU	1
1	A	161	GLU	1
1	A	93	ASP	1
1	A	65	LEU	1
1	A	116	GLU	1
1	A	132	SER	1
1	A	155	MET	1
1	A	45	THR	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

The completeness of assignment taking into account all chemical shift lists is 85% for the well-defined parts and 85% 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

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	2041
Number of shifts mapped to atoms	2041
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

#### 6.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	166	$2.58 \pm 0.11$	Should be applied
$^{13}\text{C}_\beta$	160	$2.45 \pm 0.08$	Should be applied
$^{13}\text{C}'$	87	$2.86 \pm 0.22$	Should be applied
$^{15}\text{N}$	156	$1.21 \pm 0.26$	Should be applied

#### 6.1.3 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 85%, i.e. 1725 atoms were assigned a chemical shift out of a possible 2030. 35 out of 39 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	705/804 (88%)	310/319 (97%)	245/332 (74%)	150/153 (98%)
Sidechain	897/1093 (82%)	540/640 (84%)	340/417 (82%)	17/36 (47%)

*Continued on next page...*

Continued from previous page...

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	123/133 (92%)	65/71 (92%)	54/58 (93%)	4/4 (100%)
Overall	1725/2030 (85%)	915/1030 (89%)	639/807 (79%)	171/193 (89%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 85%, i.e. 1795 atoms were assigned a chemical shift out of a possible 2120. 35 out of 39 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Backbone	731/834 (88%)	322/331 (97%)	253/344 (74%)	156/159 (98%)
Sidechain	941/1153 (82%)	569/676 (84%)	353/436 (81%)	19/41 (46%)
Aromatic	123/133 (92%)	65/71 (92%)	54/58 (93%)	4/4 (100%)
Overall	1795/2120 (85%)	956/1078 (89%)	660/838 (79%)	179/204 (88%)

#### 6.1.4 Statistically unusual chemical shifts ⓘ

There are no statistically unusual chemical shifts.

#### 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:

