



# Full wwPDB NMR Structure Validation Report ⓘ

Apr 27, 2016 – 12:38 AM BST

PDB ID : 2KUD  
Title : NMR structure of the PASTA domain 1 and 2 of Mycobacterium tuberculosis of PknB  
Authors : Barthe, P.; Mukamolova, G.; Roumestand, C.; Cohen-Gonsaud, M.  
Deposited on : 2010-02-17

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  
Mogul : unknown  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
RCI : v\_1n\_11\_5\_13\_A (Berjanski et al., 2005)  
PANAV : Wang et al. (2010)  
ShiftChecker : rb-20027457  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027457

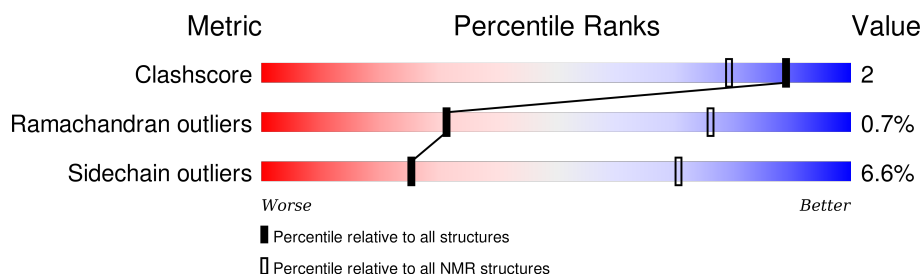
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 81%.

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	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

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	140	 91% . . .

## 2 Ensemble composition and analysis

This entry contains 30 models. Model 10 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:359-A:491 (133)	0.82	10

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 2009 atoms, of which 1010 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Serine/threonine-protein kinase pknB.

Mol	Chain	Residues	Atoms					Trace
1	A	137	Total	C	H	N	O	0
			2009	617	1010	177	205	

There are 3 discrepancies between the modelled and reference sequences:

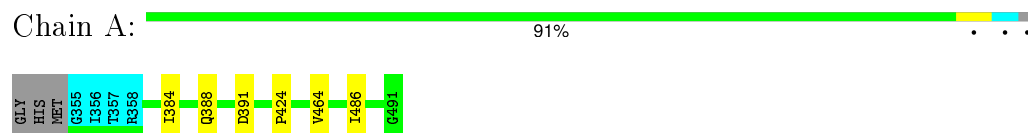
Chain	Residue	Modelled	Actual	Comment	Reference
A	352	GLY	-	EXPRESSION TAG	UNP P0A5S4
A	353	HIS	-	EXPRESSION TAG	UNP P0A5S4
A	354	MET	-	EXPRESSION TAG	UNP P0A5S4

## 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: Serine/threonine-protein kinase pknB

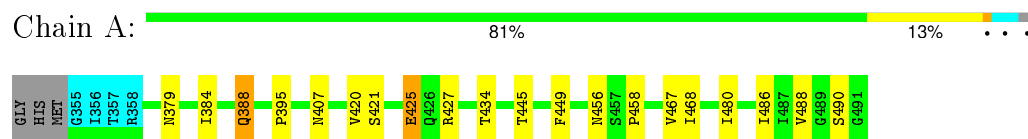


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

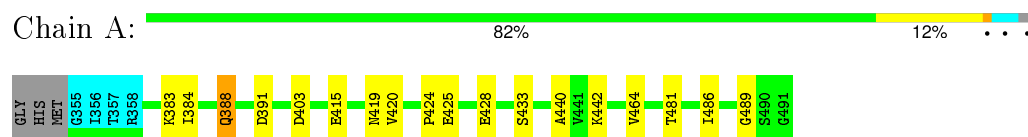
#### 4.2.1 Score per residue for model 1

- Molecule 1: Serine/threonine-protein kinase pknB



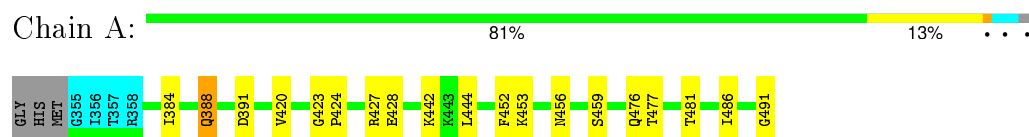
#### 4.2.2 Score per residue for model 2

- Molecule 1: Serine/threonine-protein kinase pknB



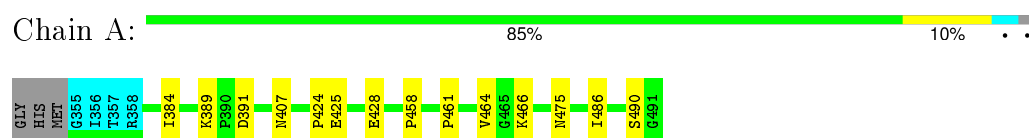
### 4.2.3 Score per residue for model 3

- Molecule 1: Serine/threonine-protein kinase pknB



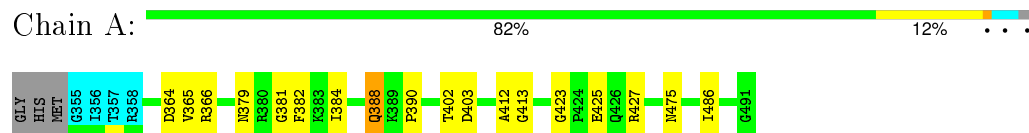
### 4.2.4 Score per residue for model 4

- Molecule 1: Serine/threonine-protein kinase pknB



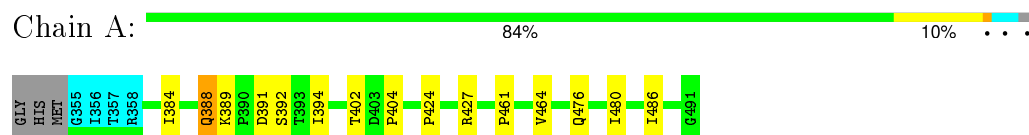
### 4.2.5 Score per residue for model 5

- Molecule 1: Serine/threonine-protein kinase pknB



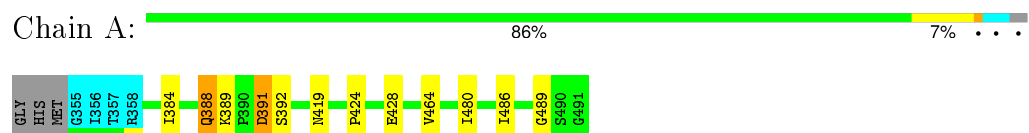
### 4.2.6 Score per residue for model 6

- Molecule 1: Serine/threonine-protein kinase pknB



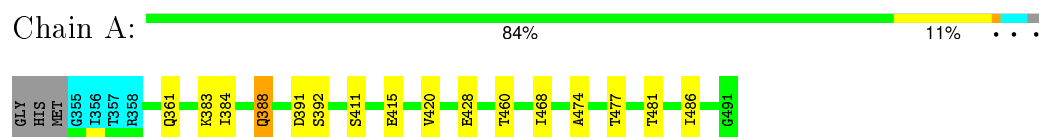
### 4.2.7 Score per residue for model 7

- Molecule 1: Serine/threonine-protein kinase pknB



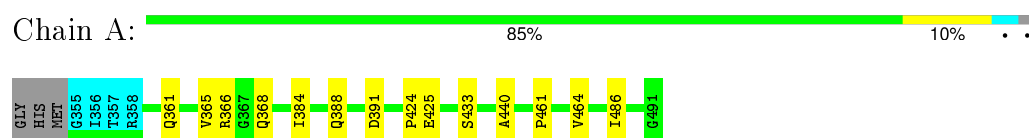
### 4.2.8 Score per residue for model 8

- Molecule 1: Serine/threonine-protein kinase pknB



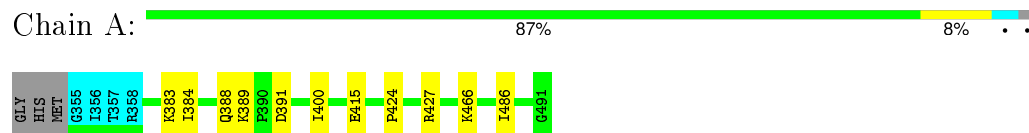
### 4.2.9 Score per residue for model 9

- Molecule 1: Serine/threonine-protein kinase pknB



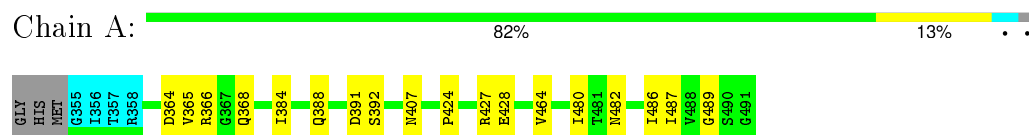
### 4.2.10 Score per residue for model 10 (medoid)

- Molecule 1: Serine/threonine-protein kinase pknB



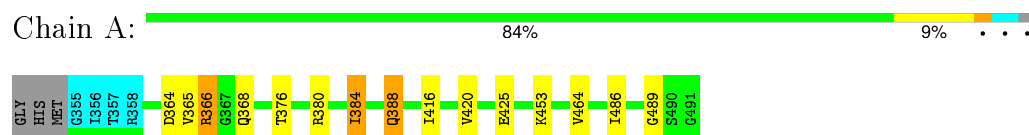
### 4.2.11 Score per residue for model 11

- Molecule 1: Serine/threonine-protein kinase pknB



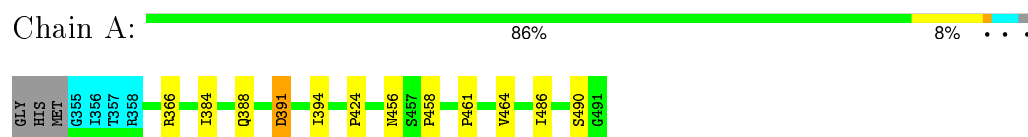
### 4.2.12 Score per residue for model 12

- Molecule 1: Serine/threonine-protein kinase pknB



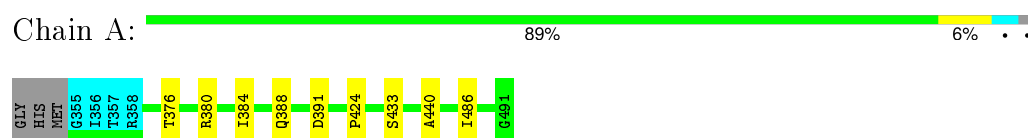
### 4.2.13 Score per residue for model 13

- Molecule 1: Serine/threonine-protein kinase pknB



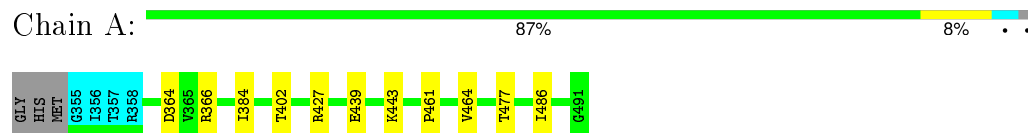
### 4.2.14 Score per residue for model 14

- Molecule 1: Serine/threonine-protein kinase pknB



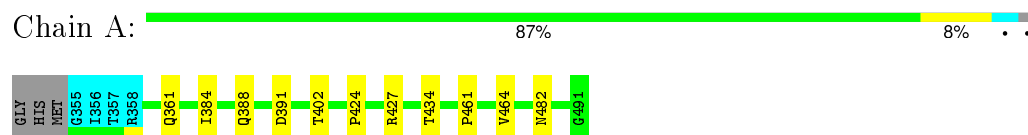
### 4.2.15 Score per residue for model 15

- Molecule 1: Serine/threonine-protein kinase pknB



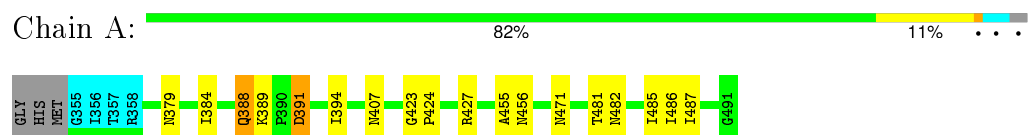
### 4.2.16 Score per residue for model 16

- Molecule 1: Serine/threonine-protein kinase pknB



### 4.2.17 Score per residue for model 17

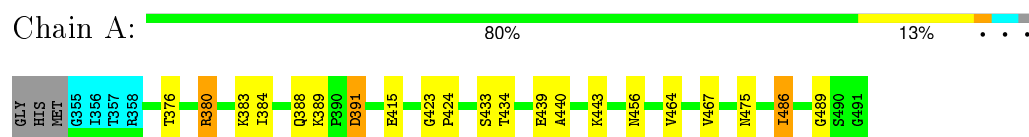
- Molecule 1: Serine/threonine-protein kinase pknB





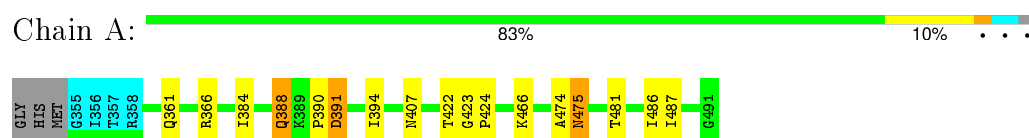
## 4.2.18 Score per residue for model 18

- Molecule 1: Serine/threonine-protein kinase pknB



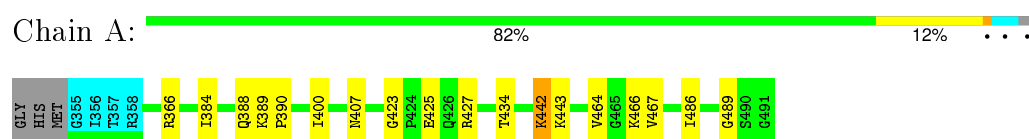
## 4.2.19 Score per residue for model 19

- Molecule 1: Serine/threonine-protein kinase pknB



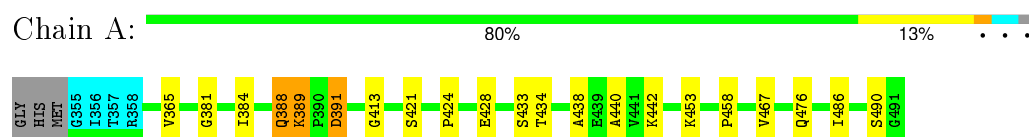
## 4.2.20 Score per residue for model 20

- Molecule 1: Serine/threonine-protein kinase pknB



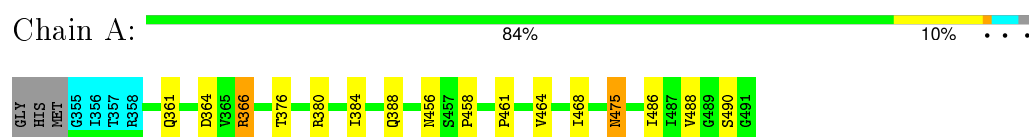
## 4.2.21 Score per residue for model 21

- Molecule 1: Serine/threonine-protein kinase pknB



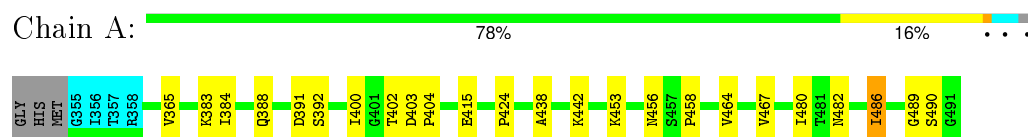
## 4.2.22 Score per residue for model 22

- Molecule 1: Serine/threonine-protein kinase pknB



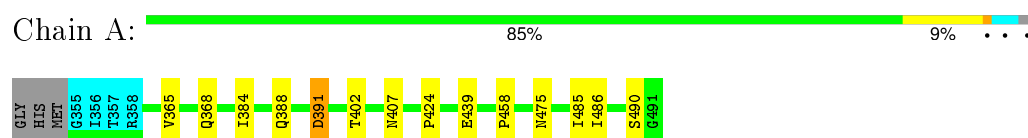
#### 4.2.23 Score per residue for model 23

- Molecule 1: Serine/threonine-protein kinase pknB



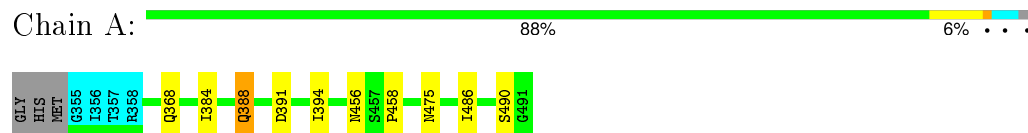
#### 4.2.24 Score per residue for model 24

- Molecule 1: Serine/threonine-protein kinase pknB



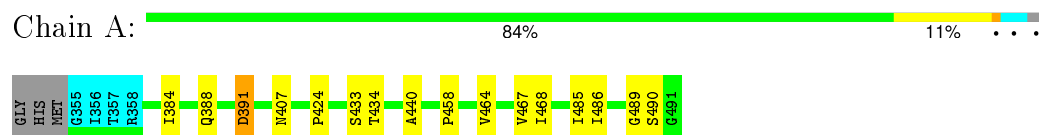
#### 4.2.25 Score per residue for model 25

- Molecule 1: Serine/threonine-protein kinase pknB



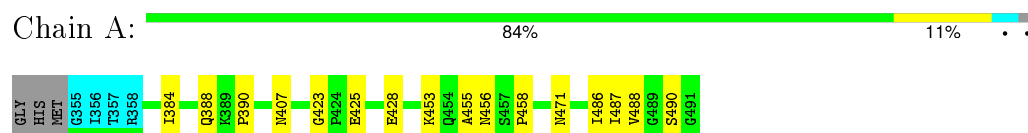
#### 4.2.26 Score per residue for model 26

- Molecule 1: Serine/threonine-protein kinase pknB



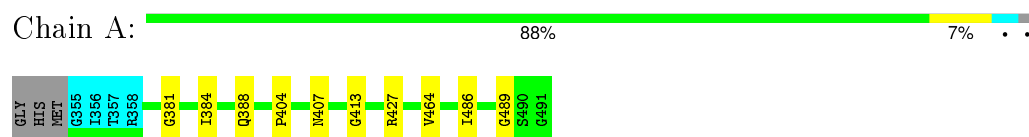
#### 4.2.27 Score per residue for model 27

- Molecule 1: Serine/threonine-protein kinase pknB



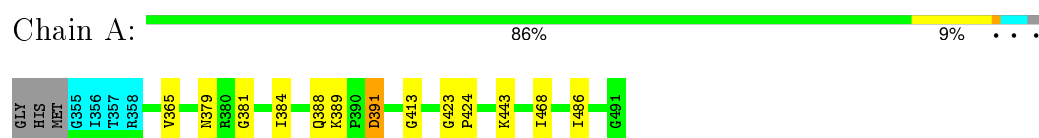
#### 4.2.28 Score per residue for model 28

- Molecule 1: Serine/threonine-protein kinase pknB



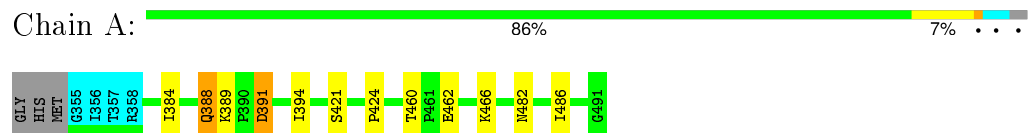
#### 4.2.29 Score per residue for model 29

- Molecule 1: Serine/threonine-protein kinase pknB



#### 4.2.30 Score per residue for model 30

- Molecule 1: Serine/threonine-protein kinase pknB



## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *restrained molecular dynamics in a hydrated environment*.

Of the 30 calculated structures, 30 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
CNS	refinement	1.2

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)	BMRB entry 16732
Number of chemical shift lists	1
Total number of shifts	1446
Number of shifts mapped to atoms	1446
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	81%

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](#)

There are no covalent bond-length or bond-angle outliers.

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	0.1±0.2
All	All	0	2

There are no bond-length outliers.

There are no bond-angle outliers.

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	366	ARG	Sidechain	1
1	A	380	ARG	Sidechain	1

### 6.2 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	969	976	976	4±2
All	All	29070	29280	29280	134

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 2.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:458:PRO:HA	1:A:490:SER:O	0.60	1.95	1	10
1:A:464:VAL:HA	1:A:489:GLY:O	0.59	1.98	7	9
1:A:392:SER:OG	1:A:480:ILE:HB	0.57	1.98	6	3
1:A:391:ASP:O	1:A:424:PRO:HA	0.56	1.99	16	20
1:A:456:ASN:OD1	1:A:488:VAL:HB	0.55	2.01	1	3
1:A:433:SER:HA	1:A:440:ALA:HB2	0.53	1.80	14	6
1:A:388:GLN:HA	1:A:420:VAL:O	0.53	2.04	1	5
1:A:439:GLU:O	1:A:443:LYS:HG2	0.53	2.03	18	2
1:A:384:ILE:HA	1:A:416:ILE:O	0.52	2.05	12	1
1:A:438:ALA:O	1:A:442:LYS:HG2	0.51	2.05	21	2
1:A:390:PRO:HA	1:A:423:GLY:O	0.51	2.05	20	3
1:A:381:GLY:O	1:A:413:GLY:HA2	0.51	2.06	21	4
1:A:376:THR:O	1:A:380:ARG:HG2	0.51	2.05	22	2
1:A:383:LYS:O	1:A:415:GLU:HA	0.51	2.06	23	5
1:A:364:ASP:OD1	1:A:366:ARG:HG3	0.50	2.06	12	2
1:A:392:SER:O	1:A:481:THR:HB	0.50	2.07	8	1
1:A:461:PRO:O	1:A:464:VAL:HG12	0.49	2.06	9	6
1:A:391:ASP:HB3	1:A:394:ILE:HG23	0.49	1.84	19	5
1:A:376:THR:O	1:A:380:ARG:HD3	0.48	2.09	18	2
1:A:395:PRO:HD2	1:A:421:SER:OG	0.47	2.09	1	1
1:A:459:SER:O	1:A:491:GLY:HA3	0.46	2.10	3	1
1:A:388:GLN:HE21	1:A:388:GLN:N	0.46	2.08	17	4
1:A:382:PHE:CZ	1:A:412:ALA:HA	0.46	2.45	5	1
1:A:434:THR:HA	1:A:467:VAL:O	0.46	2.10	18	5
1:A:427:ARG:NE	1:A:427:ARG:HA	0.46	2.25	6	1
1:A:388:GLN:N	1:A:388:GLN:HE21	0.46	2.09	5	3
1:A:394:ILE:HD12	1:A:421:SER:HB2	0.45	1.89	30	1
1:A:388:GLN:HB2	1:A:422:THR:OG1	0.45	2.12	19	1
1:A:467:VAL:HG22	1:A:486:ILE:HD11	0.44	1.88	23	2
1:A:461:PRO:O	1:A:464:VAL:HG22	0.44	2.12	13	1
1:A:364:ASP:OD2	1:A:366:ARG:HD2	0.44	2.12	5	1
1:A:455:ALA:O	1:A:487:ILE:HA	0.43	2.13	27	2
1:A:365:VAL:O	1:A:368:GLN:HG2	0.43	2.14	24	1
1:A:364:ASP:HB3	1:A:366:ARG:NH1	0.43	2.29	15	1
1:A:475:ASN:N	1:A:475:ASN:HD22	0.43	2.12	22	2
1:A:425:GLU:O	1:A:480:ILE:HG12	0.42	2.13	1	1
1:A:392:SER:CB	1:A:480:ILE:HB	0.42	2.44	7	1
1:A:460:THR:HB	1:A:462:GLU:OE1	0.42	2.13	30	1
1:A:471:ASN:CB	1:A:485:ILE:HB	0.42	2.45	17	1
1:A:365:VAL:HB	1:A:368:GLN:NE2	0.42	2.29	12	1
1:A:453:LYS:HE3	1:A:453:LYS:HA	0.41	1.92	21	1
1:A:442:LYS:HA	1:A:442:LYS:HE3	0.41	1.91	20	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:444:LEU:HD23	1:A:452:PHE:CE2	0.41	2.50	3	1
1:A:402:THR:HG22	1:A:403:ASP:H	0.41	1.76	5	2
1:A:445:THR:HA	1:A:449:PHE:O	0.41	2.15	1	1
1:A:442:LYS:HA	1:A:442:LYS:HE2	0.41	1.91	3	1
1:A:389:LYS:O	1:A:421:SER:HA	0.41	2.16	21	1
1:A:390:PRO:O	1:A:391:ASP:HB2	0.41	2.16	19	1
1:A:364:ASP:OD2	1:A:366:ARG:HD3	0.40	2.17	11	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	132/140 (94%)	127±2 (96±2%)	4±2 (3±2%)	1±1 (1±1%)	31	76
All	All	3960/4200 (94%)	3800 (96%)	133 (3%)	27 (1%)	31	76

All 5 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	391	ASP	11
1	A	365	VAL	6
1	A	423	GLY	5
1	A	404	PRO	3
1	A	474	ALA	2

### 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	109/114 (96%)	102±2 (93±2%)	7±2 (7±2%)	25	71
All	All	3270/3420 (96%)	3055 (93%)	215 (7%)	25	71

All 34 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	384	ILE	30
1	A	486	ILE	29
1	A	388	GLN	28
1	A	389	LYS	10
1	A	427	ARG	10
1	A	407	ASN	10
1	A	428	GLU	8
1	A	425	GLU	8
1	A	475	ASN	7
1	A	456	ASN	6
1	A	466	LYS	5
1	A	468	ILE	5
1	A	361	GLN	5
1	A	482	ASN	5
1	A	366	ARG	5
1	A	402	THR	4
1	A	379	ASN	4
1	A	453	LYS	4
1	A	481	THR	4
1	A	368	GLN	3
1	A	476	GLN	3
1	A	400	ILE	3
1	A	477	THR	3
1	A	487	ILE	2
1	A	442	LYS	2
1	A	485	ILE	2
1	A	443	LYS	2
1	A	419	ASN	2
1	A	411	SER	1
1	A	460	THR	1
1	A	439	GLU	1
1	A	471	ASN	1
1	A	434	THR	1
1	A	403	ASP	1



### 6.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 6.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 6.7 Other polymers [i](#)

There are no such molecules in this entry.

## 6.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 7 Chemical shift validation [i](#)

The completeness of assignment taking into account all chemical shift lists is 81% for the well-defined parts and 80% for the entire structure.

### 7.1 Chemical shift list 1

File name: BMRB entry 16732

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

#### 7.1.1 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	1446
Number of shifts mapped to atoms	1446
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

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

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	135	$-0.27 \pm 0.15$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}_\beta$	123	$-0.04 \pm 0.06$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}'$	134	$-0.06 \pm 0.08$	None needed ( $< 0.5$ ppm)
$^{15}\text{N}$	126	$-0.58 \pm 0.20$	Should be applied

#### 7.1.3 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 81%, i.e. 1210 atoms were assigned a chemical shift out of a possible 1496. 0 out of 19 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	636/643 (99%)	253/255 (99%)	261/266 (98%)	122/122 (100%)
Sidechain	553/811 (68%)	414/470 (88%)	120/306 (39%)	19/35 (54%)

*Continued on next page...*

Continued from previous page...

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	21/42 (50%)	21/23 (91%)	0/18 (0%)	0/1 (0%)
Overall	1210/1496 (81%)	688/748 (92%)	381/590 (65%)	141/158 (89%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 80%, i.e. 1242 atoms were assigned a chemical shift out of a possible 1547. 0 out of 19 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Backbone	651/663 (98%)	259/263 (98%)	267/274 (97%)	125/126 (99%)
Sidechain	570/842 (68%)	427/488 (88%)	123/316 (39%)	20/38 (53%)
Aromatic	21/42 (50%)	21/23 (91%)	0/18 (0%)	0/1 (0%)
Overall	1242/1547 (80%)	707/774 (91%)	390/608 (64%)	145/165 (88%)

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

There are no statistically unusual chemical shifts.

#### 7.1.5 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:

