



# Full wwPDB NMR Structure Validation Report ⓘ

May 29, 2020 – 07:51 am BST

PDB ID : 5N6R  
Title : Solution structure of the Dbl-homology domain of Bcr-Abl  
Authors : Reckel, S.; Lohr, F.; Buchner, L.; Guntert, P.; Dotsch, V.; Hantschel, O.  
Deposited on : 2017-02-16

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

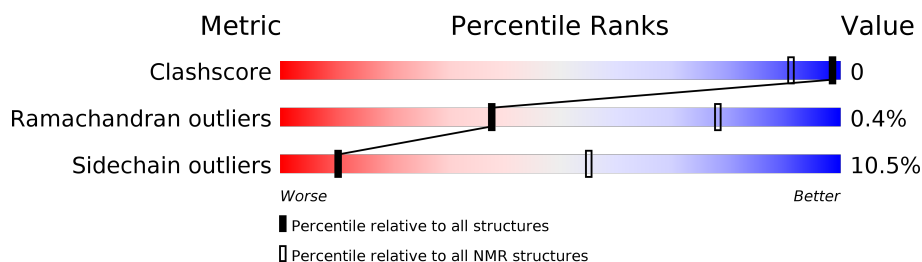
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 93%.

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	218	

## 2 Ensemble composition and analysis

This entry contains 20 models. Model 1 is the overall representative, medoid model (most similar to other models).

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:496-A:619, A:638-A:693 (180)	0.62	1

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 and 2 single-model clusters were found.

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 3500 atoms, of which 1756 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Breakpoint cluster region protein.

Mol	Chain	Residues	Atoms						Trace
1	A	218	Total	C	H	N	O	S	0
			3500	1105	1756	302	329	8	

There are 2 discrepancies between the modelled and reference sequences:

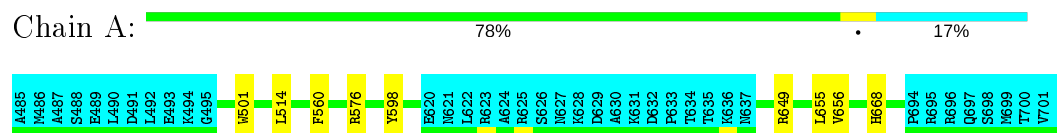
Chain	Residue	Modelled	Actual	Comment	Reference
A	485	ALA	-	expression tag	UNP P11274
A	486	MET	-	expression tag	UNP P11274

## 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: Breakpoint cluster region protein

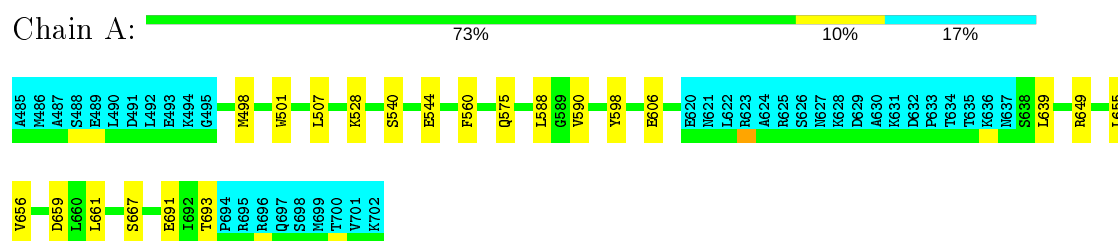


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

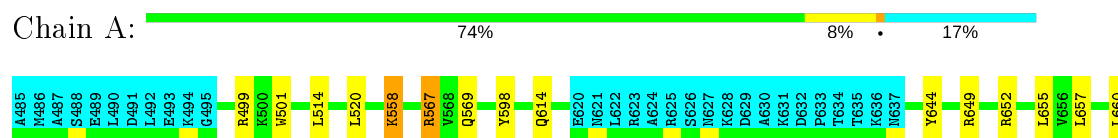
#### 4.2.1 Score per residue for model 1 (medoid)

- Molecule 1: Breakpoint cluster region protein



#### 4.2.2 Score per residue for model 2

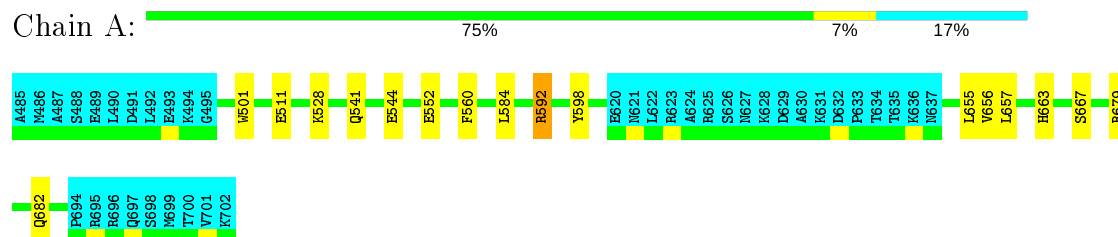
- Molecule 1: Breakpoint cluster region protein





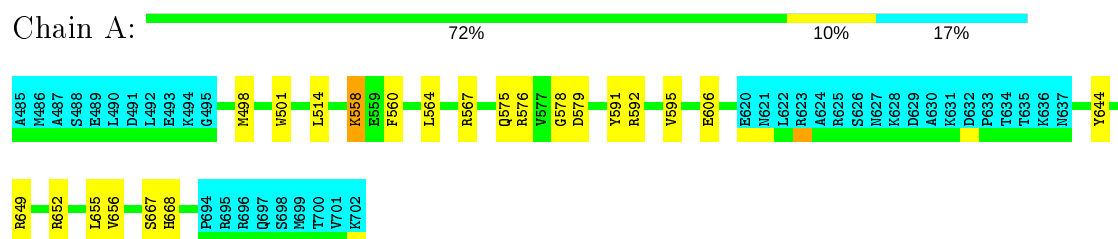
### 4.2.3 Score per residue for model 3

- Molecule 1: Breakpoint cluster region protein



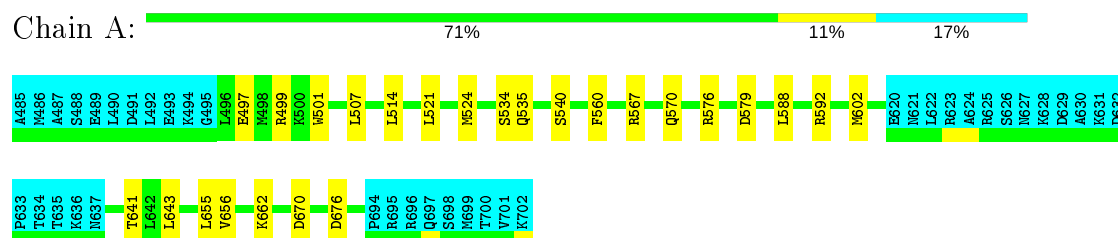
### 4.2.4 Score per residue for model 4

- Molecule 1: Breakpoint cluster region protein



### 4.2.5 Score per residue for model 5

- Molecule 1: Breakpoint cluster region protein



### 4.2.6 Score per residue for model 6

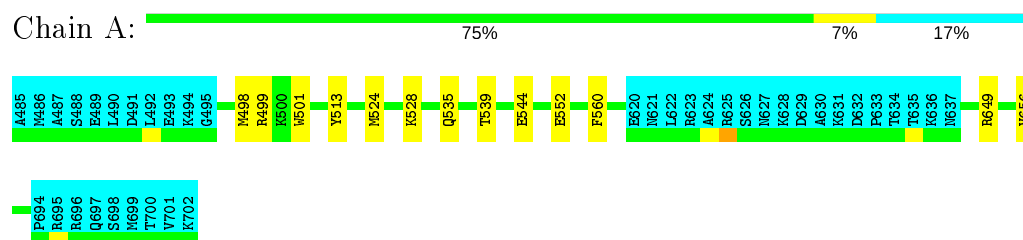
- Molecule 1: Breakpoint cluster region protein





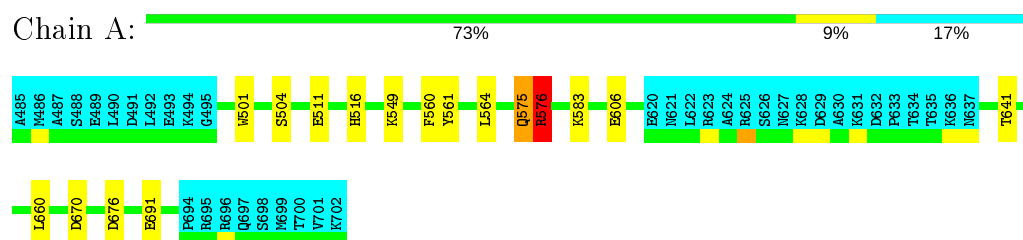
#### 4.2.7 Score per residue for model 7

- Molecule 1: Breakpoint cluster region protein



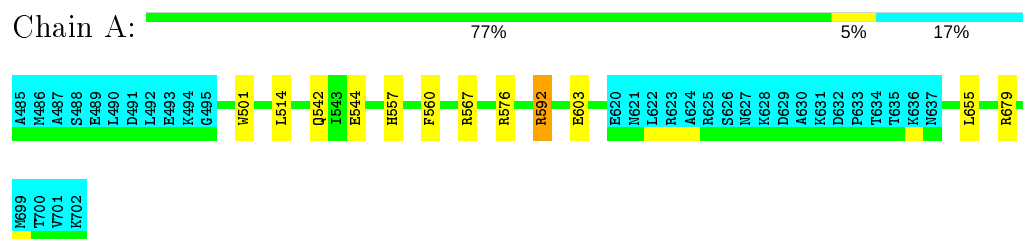
#### 4.2.8 Score per residue for model 8

- Molecule 1: Breakpoint cluster region protein



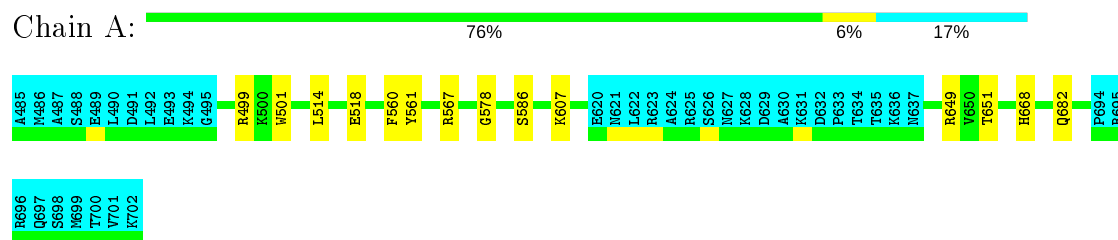
#### 4.2.9 Score per residue for model 9

- Molecule 1: Breakpoint cluster region protein



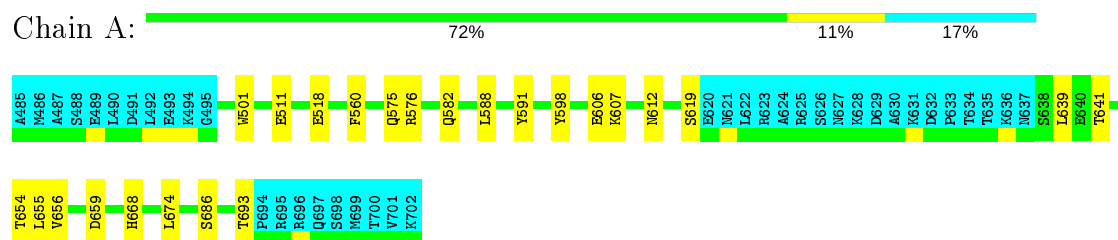
#### 4.2.10 Score per residue for model 10

- Molecule 1: Breakpoint cluster region protein



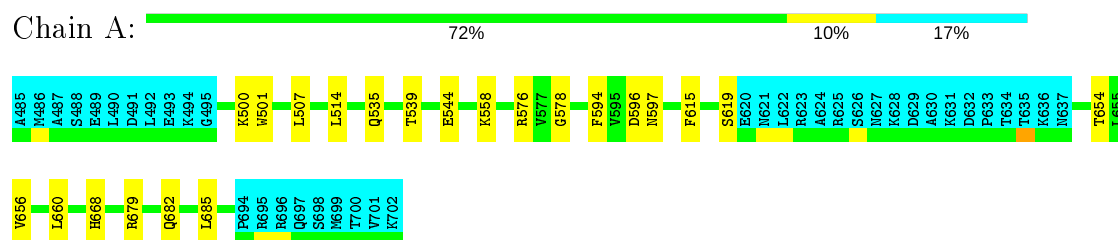
#### 4.2.11 Score per residue for model 11

- Molecule 1: Breakpoint cluster region protein



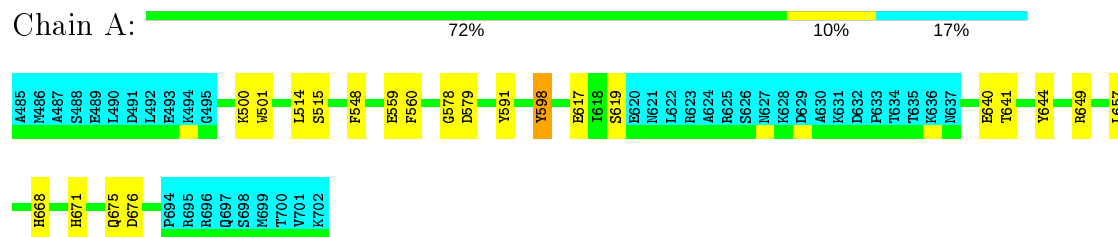
#### 4.2.12 Score per residue for model 12

- Molecule 1: Breakpoint cluster region protein



#### 4.2.13 Score per residue for model 13

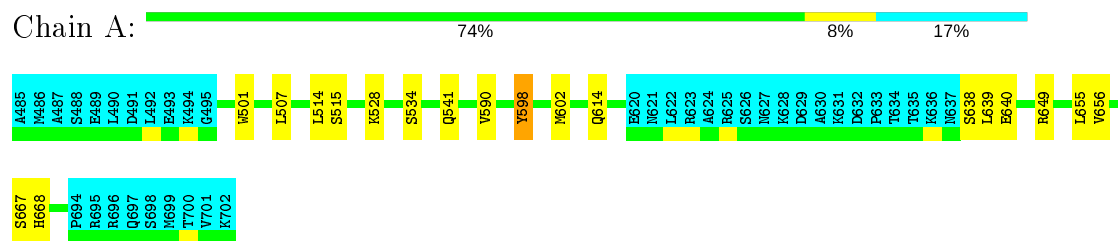
- Molecule 1: Breakpoint cluster region protein





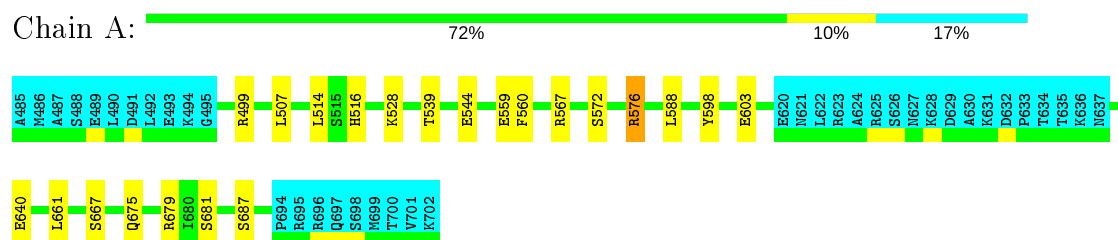
#### 4.2.14 Score per residue for model 14

- Molecule 1: Breakpoint cluster region protein



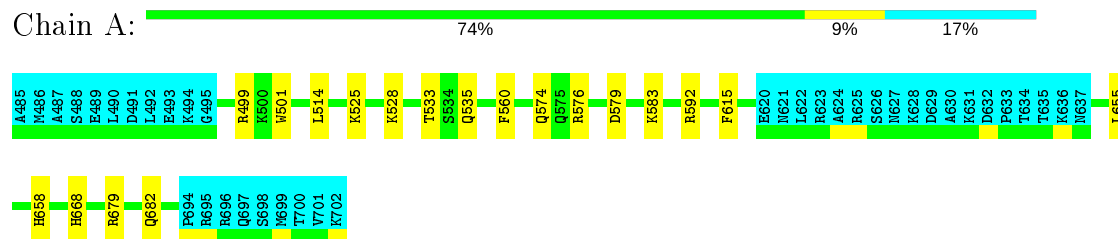
#### 4.2.15 Score per residue for model 15

- Molecule 1: Breakpoint cluster region protein



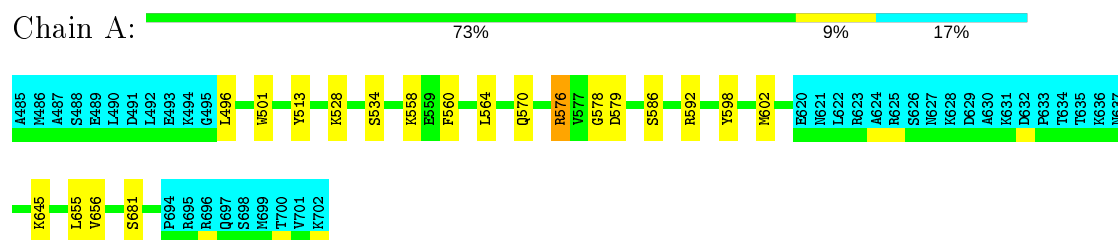
#### 4.2.16 Score per residue for model 16

- Molecule 1: Breakpoint cluster region protein



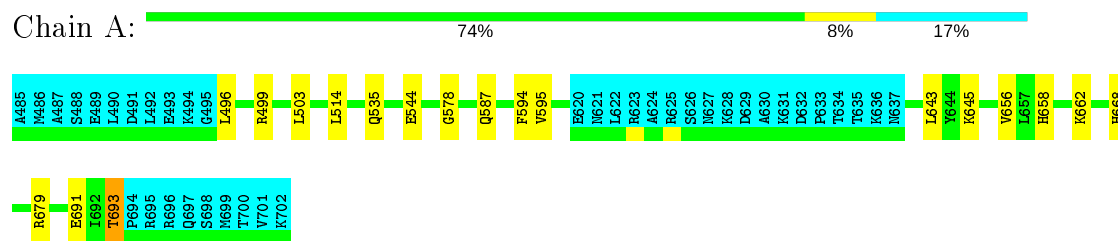
#### 4.2.17 Score per residue for model 17

- Molecule 1: Breakpoint cluster region protein



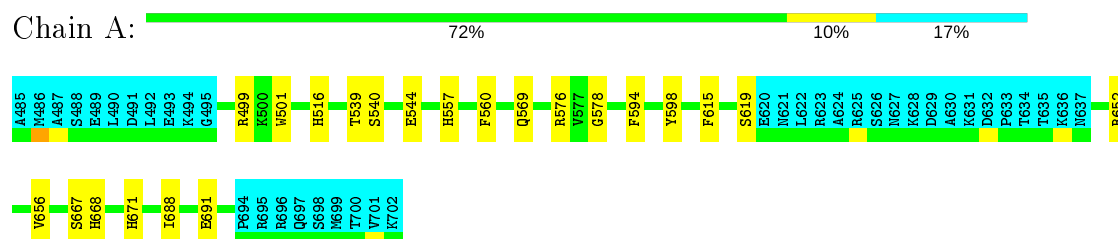
### 4.2.18 Score per residue for model 18

- Molecule 1: Breakpoint cluster region protein



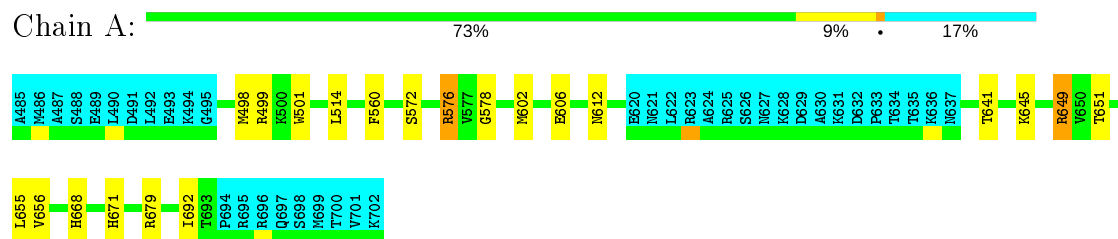
### 4.2.19 Score per residue for model 19

- Molecule 1: Breakpoint cluster region protein



### 4.2.20 Score per residue for model 20

- Molecule 1: Breakpoint cluster region protein



## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *TORSION ANGLE DYNAMICS ONFORMERS*, *NUMBER CALCULATED : NULL*.

Of the 200 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
CYANA	structure calculation	

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	2893
Number of shifts mapped to atoms	2893
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	93%

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	1447	1450	1450	0±0
All	All	28940	29000	29000	5

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:514:LEU:HD22	1:A:558:LYS:HG2	0.47	1.84	2	1
1:A:654:THR:HG21	1:A:685:LEU:HD23	0.47	1.86	12	1
1:A:514:LEU:HD21	1:A:557:HIS:CB	0.45	2.41	9	1
1:A:514:LEU:HD13	1:A:558:LYS:HG2	0.43	1.90	4	1
1:A:575:GLN:NE2	1:A:576:ARG:H	0.43	2.11	8	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	180/218 (83%)	173±2 (96±1%)	7±2 (4±1%)	1±1 (0±0%)	38	78
All	All	3600/4360 (83%)	3452 (96%)	134 (4%)	14 (0%)	38	78

All 4 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	578	GLY	8
1	A	534	SER	3
1	A	693	THR	2
1	A	688	ILE	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	161/194 (83%)	144±3 (90±2%)	17±3 (10±2%)	10	55
All	All	3220/3880 (83%)	2882 (90%)	338 (10%)	10	55

All 99 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	501	TRP	17
1	A	560	PHE	16
1	A	656	VAL	14
1	A	655	LEU	13
1	A	668	HIS	12
1	A	576	ARG	12
1	A	514	LEU	9
1	A	544	GLU	8
1	A	567	ARG	7
1	A	528	LYS	7
1	A	667	SER	7
1	A	657	LEU	6
1	A	507	LEU	5
1	A	649	ARG	5
1	A	579	ASP	5
1	A	535	GLN	5
1	A	679	ARG	5
1	A	606	GLU	5
1	A	641	THR	4
1	A	575	GLN	4
1	A	498	MET	4
1	A	516	HIS	4
1	A	682	GLN	4
1	A	588	LEU	4
1	A	602	MET	4
1	A	670	ASP	4
1	A	615	PHE	4
1	A	691	GLU	4
1	A	645	LYS	4
1	A	540	SER	4
1	A	619	SER	4
1	A	539	THR	4
1	A	558	LYS	4
1	A	660	LEU	3
1	A	499	ARG	3
1	A	524	MET	3
1	A	511	GLU	3
1	A	639	LEU	3
1	A	675	GLN	3
1	A	671	HIS	3
1	A	607	LYS	3
1	A	640	GLU	3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	598	TYR	3
1	A	564	LEU	3
1	A	676	ASP	3
1	A	661	LEU	3
1	A	595	VAL	2
1	A	603	GLU	2
1	A	500	LYS	2
1	A	594	PHE	2
1	A	552	GLU	2
1	A	515	SER	2
1	A	614	GLN	2
1	A	572	SER	2
1	A	651	THR	2
1	A	561	TYR	2
1	A	643	LEU	2
1	A	496	LEU	2
1	A	659	ASP	2
1	A	586	SER	2
1	A	533	THR	2
1	A	583	LYS	2
1	A	559	GLU	2
1	A	612	ASN	2
1	A	658	HIS	2
1	A	662	LYS	2
1	A	541	GLN	2
1	A	693	THR	2
1	A	590	VAL	2
1	A	617	GLU	2
1	A	663	HIS	2
1	A	570	GLN	2
1	A	518	GLU	2
1	A	569	GLN	2
1	A	591	TYR	1
1	A	597	ASN	1
1	A	638	SER	1
1	A	548	PHE	1
1	A	587	GLN	1
1	A	674	LEU	1
1	A	497	GLU	1
1	A	686	SER	1
1	A	692	ILE	1
1	A	654	THR	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	549	LYS	1
1	A	520	LEU	1
1	A	582	GLN	1
1	A	557	HIS	1
1	A	504	SER	1
1	A	525	LYS	1
1	A	681	SER	1
1	A	584	LEU	1
1	A	596	ASP	1
1	A	574	GLN	1
1	A	687	SER	1
1	A	542	GLN	1
1	A	538	LEU	1
1	A	503	LEU	1
1	A	521	LEU	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 [i](#)

There are no chain breaks in this entry.

## 6 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 93% for the well-defined parts and 92% for the entire structure.

### 6.1 Chemical shift list 1

File name: input\_cs.cif

Chemical shift list name: *dh\_final\_NMR-STAR.str*

#### 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	2893
Number of shifts mapped to atoms	2893
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
$^{13}\text{C}_\alpha$	218	$-0.70 \pm 0.08$	Should be applied
$^{13}\text{C}_\beta$	212	$0.18 \pm 0.04$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}'$	218	$-0.45 \pm 0.08$	None needed ( $< 0.5$ ppm)
$^{15}\text{N}$	207	$0.05 \pm 0.14$	None needed ( $< 0.5$ ppm)

#### 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 93%, i.e. 2118 atoms were assigned a chemical shift out of a possible 2272. 37 out of 37 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	882/882 (100%)	351/351 (100%)	360/360 (100%)	171/171 (100%)
Sidechain	1057/1197 (88%)	664/702 (95%)	392/448 (88%)	1/47 (2%)

*Continued on next page...*



Continued from previous page...

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Aromatic	179/193 (93%)	97/104 (93%)	80/80 (100%)	2/9 (22%)
Overall	2118/2272 (93%)	1112/1157 (96%)	832/888 (94%)	174/227 (77%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 92%, i.e. 2544 atoms were assigned a chemical shift out of a possible 2755. 41 out of 41 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	1067/1068 (100%)	424/425 (100%)	436/436 (100%)	207/207 (100%)
Sidechain	1298/1494 (87%)	821/878 (94%)	475/548 (87%)	2/68 (3%)
Aromatic	179/193 (93%)	97/104 (93%)	80/80 (100%)	2/9 (22%)
Overall	2544/2755 (92%)	1342/1407 (95%)	991/1064 (93%)	211/284 (74%)

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

There are no statistically unusual chemical shifts.

#### 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:

