



# Full wwPDB NMR Structure Validation Report ⓘ

Feb 17, 2022 – 09:17 AM EST

PDB ID : 1PD7  
Title : Extended SID of Mad1 bound to the PAH2 domain of mSin3B  
Authors : Van Ingen, H.; Lasonder, E.; Jansen, J.F.; Kaan, A.M.; Spronk, C.A.; Stunnenberg, H.G.; Vuister, G.W.  
Deposited on : 2003-05-19

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:

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 : 2.26  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.26

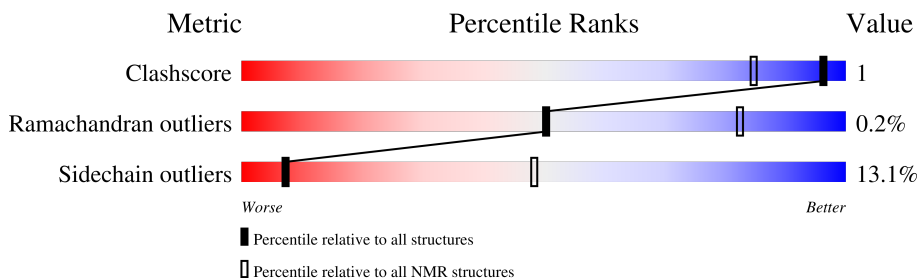
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment was not calculated.

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	85	
2	B	24	

## 2 Ensemble composition and analysis i

This entry contains 30 models. Model 24 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:5-A:39, A:54-A:78, B:201-B:215 (75)	0.29	24

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 3 clusters and 2 single-model clusters were found.

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

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

There are 2 unique types of molecules in this entry. The entry contains 1817 atoms, of which 896 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Sin3b protein.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	85	1405	454	691	122	137	1	0

- Molecule 2 is a protein called Mad1.

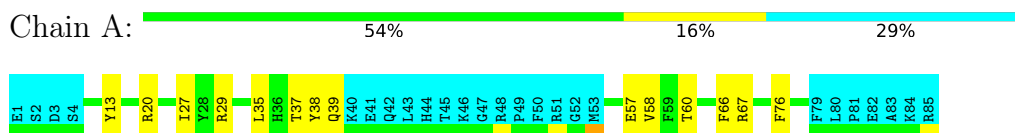
Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	24	412	125	205	40	40	2	0

## 4 Residue-property plots [i](#)

### 4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide 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: Sin3b protein



- Molecule 2: Mad1



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

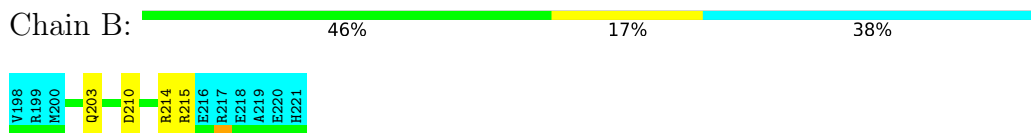
Colouring as in section 4.1 above.

#### 4.2.1 Score per residue for model 1

- Molecule 1: Sin3b protein

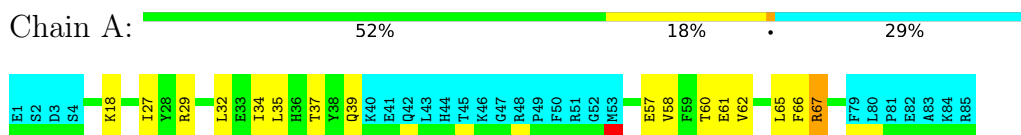


- Molecule 2: Mad1



### 4.2.2 Score per residue for model 2

- Molecule 1: Sin3b protein

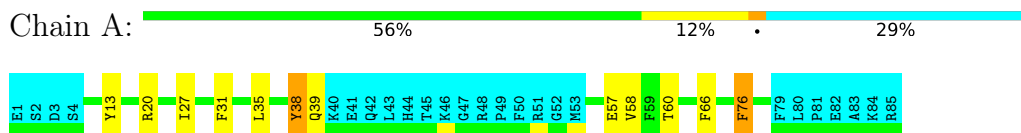


- Molecule 2: Mad1



### 4.2.3 Score per residue for model 3

- Molecule 1: Sin3b protein

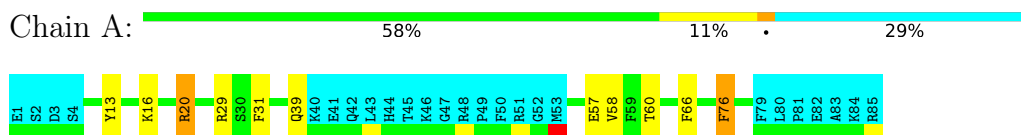


- Molecule 2: Mad1



### 4.2.4 Score per residue for model 4

- Molecule 1: Sin3b protein

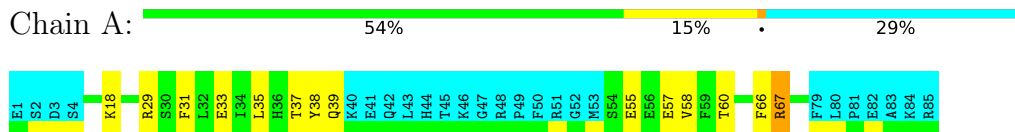


- Molecule 2: Mad1



### 4.2.5 Score per residue for model 5

- Molecule 1: Sin3b protein

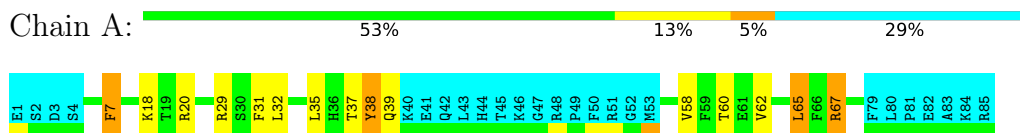


- Molecule 2: Mad1



### 4.2.6 Score per residue for model 6

- Molecule 1: Sin3b protein

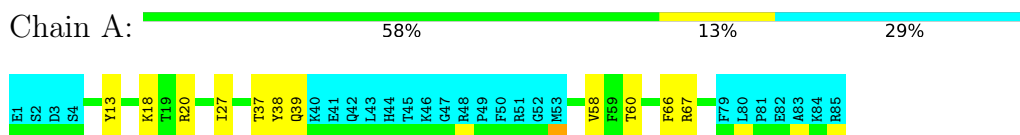


- Molecule 2: Mad1



### 4.2.7 Score per residue for model 7

- Molecule 1: Sin3b protein

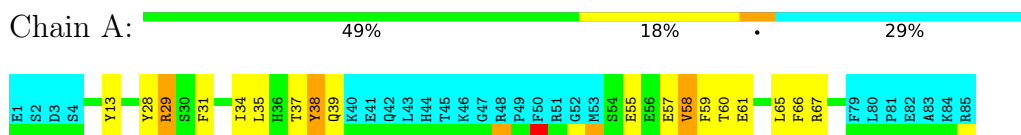


- Molecule 2: Mad1



### 4.2.8 Score per residue for model 8

- Molecule 1: Sin3b protein

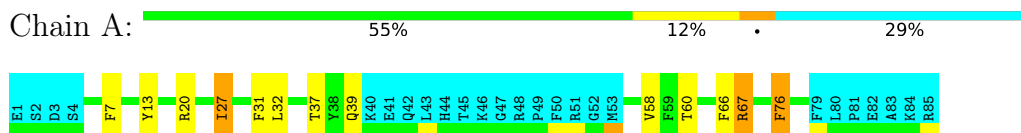


- Molecule 2: Mad1



### 4.2.9 Score per residue for model 9

- Molecule 1: Sin3b protein

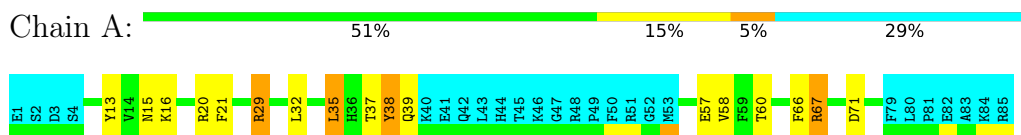


- Molecule 2: Mad1



### 4.2.10 Score per residue for model 10

- Molecule 1: Sin3b protein



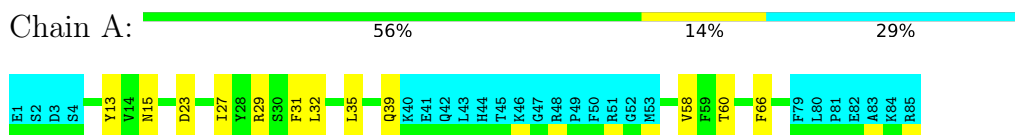
- Molecule 2: Mad1





### 4.2.11 Score per residue for model 11

- Molecule 1: Sin3b protein

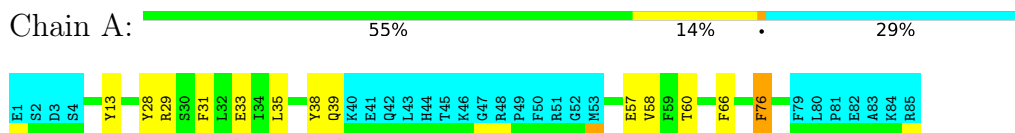


- Molecule 2: Mad1



### 4.2.12 Score per residue for model 12

- Molecule 1: Sin3b protein



- Molecule 2: Mad1



### 4.2.13 Score per residue for model 13

- Molecule 1: Sin3b protein

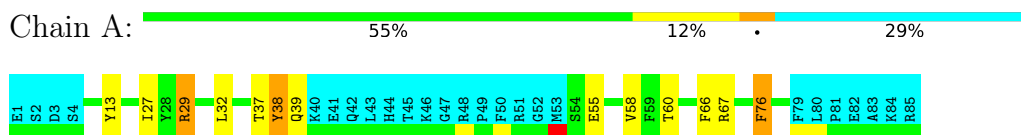


- Molecule 2: Mad1



#### 4.2.14 Score per residue for model 14

- Molecule 1: Sin3b protein

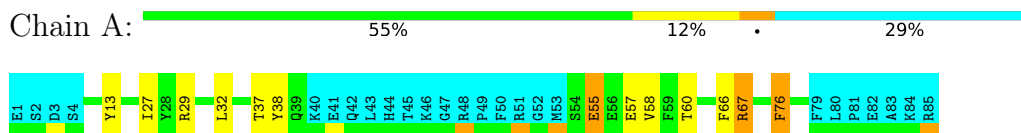


- Molecule 2: Mad1



#### 4.2.15 Score per residue for model 15

- Molecule 1: Sin3b protein

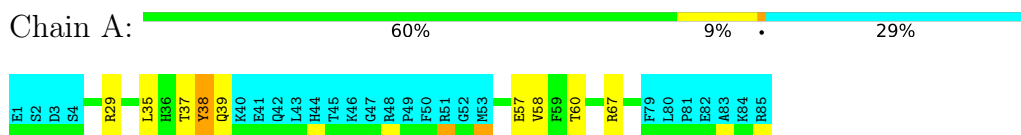


- Molecule 2: Mad1



#### 4.2.16 Score per residue for model 16

- Molecule 1: Sin3b protein

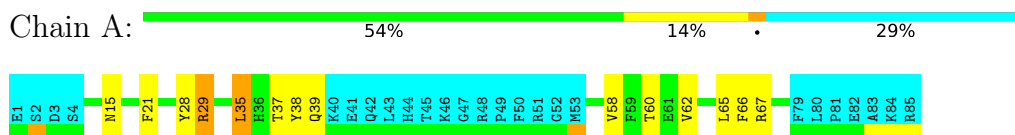


- Molecule 2: Mad1



### 4.2.17 Score per residue for model 17

- Molecule 1: Sin3b protein

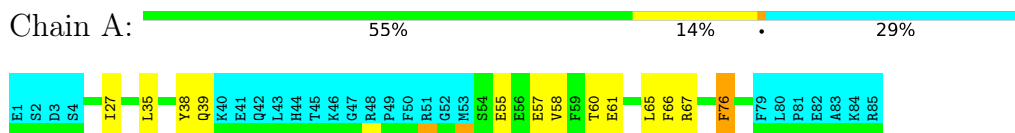


- Molecule 2: Mad1



### 4.2.18 Score per residue for model 18

- Molecule 1: Sin3b protein

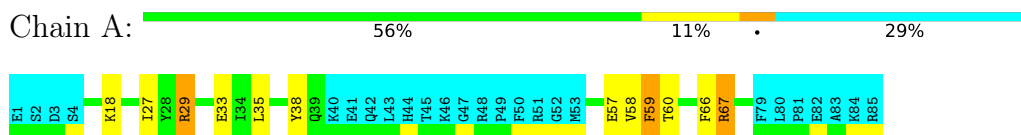


- Molecule 2: Mad1



### 4.2.19 Score per residue for model 19

- Molecule 1: Sin3b protein

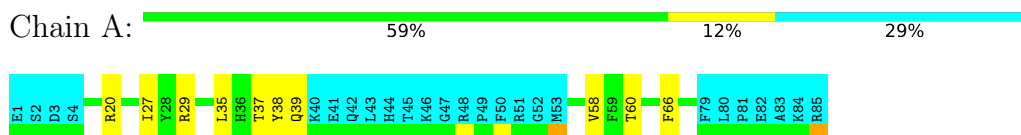


- Molecule 2: Mad1



### 4.2.20 Score per residue for model 20

- Molecule 1: Sin3b protein

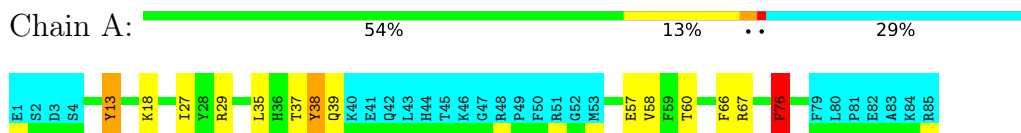


- Molecule 2: Mad1

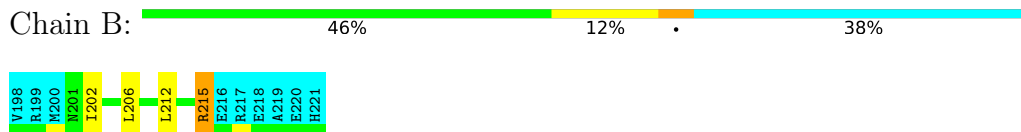


### 4.2.21 Score per residue for model 21

- Molecule 1: Sin3b protein

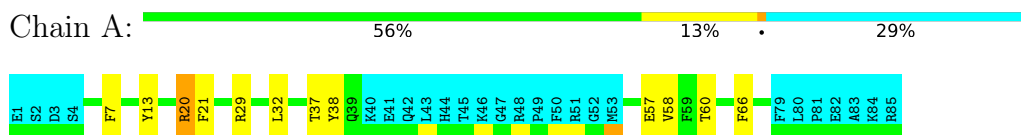


- Molecule 2: Mad1



### 4.2.22 Score per residue for model 22

- Molecule 1: Sin3b protein



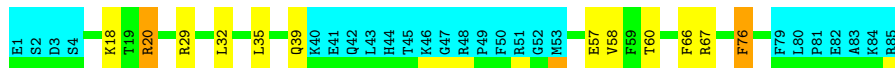
- Molecule 2: Mad1



### 4.2.23 Score per residue for model 23

- Molecule 1: Sin3b protein

Chain A:  56% 12% 29%



- Molecule 2: Mad1

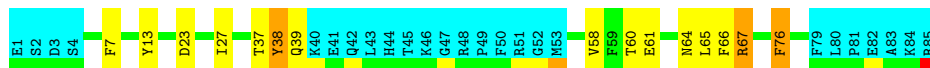
Chain B:  46% 17% 38%



### 4.2.24 Score per residue for model 24 (medoid)

- Molecule 1: Sin3b protein

Chain A:  53% 14% 29%



- Molecule 2: Mad1

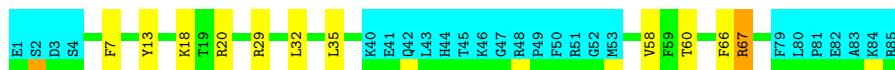
Chain B:  54% 8% 38%



### 4.2.25 Score per residue for model 25

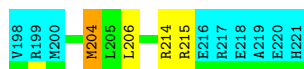
- Molecule 1: Sin3b protein

Chain A:  58% 12% 29%



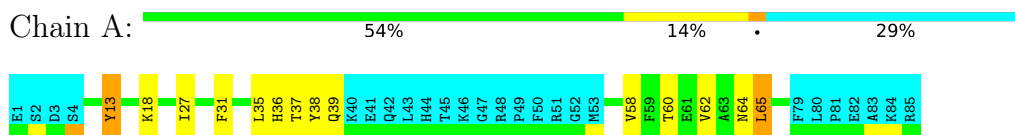
- Molecule 2: Mad1

Chain B:  46% 12% 38%



#### 4.2.26 Score per residue for model 26

- Molecule 1: Sin3b protein

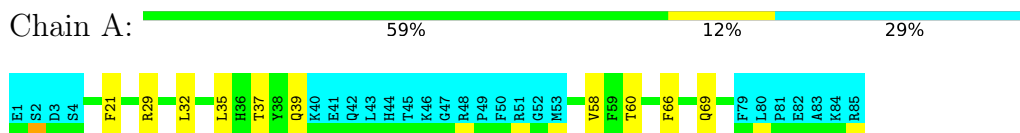


- Molecule 2: Mad1



#### 4.2.27 Score per residue for model 27

- Molecule 1: Sin3b protein

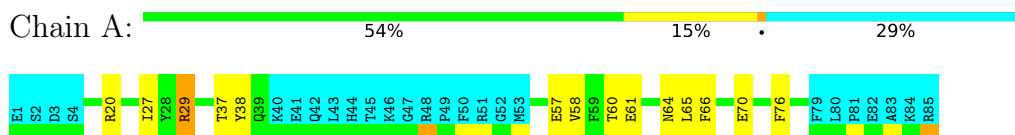


- Molecule 2: Mad1



#### 4.2.28 Score per residue for model 28

- Molecule 1: Sin3b protein

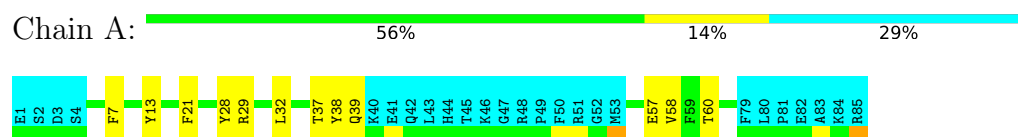


- Molecule 2: Mad1



### 4.2.29 Score per residue for model 29

- Molecule 1: Sin3b protein

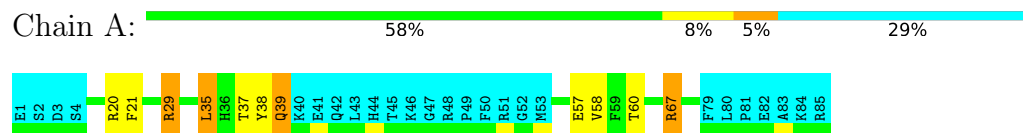


- Molecule 2: Mad1



### 4.2.30 Score per residue for model 30

- Molecule 1: Sin3b protein



- Molecule 2: Mad1



## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *simulated annealing, ITERATIVE NOE-ASSIGNMENT*.

Of the 200 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
X-PLOR	structure solution	3.581
X-PLOR	refinement	3.581

No chemical shift data was provided.



## 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	0.98±0.00	0±0/519 ( 0.0± 0.0%)	1.31±0.07	6±3/702 ( 0.9± 0.4%)
2	B	1.09±0.00	0±0/128 ( 0.0± 0.0%)	1.29±0.08	1±1/172 ( 0.8± 0.6%)
All	All	1.00	0/19410 ( 0.0%)	1.31	232/26220 ( 0.9%)

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	0.7±0.7
2	B	0.0±0.0	0.3±0.5
All	All	0	30

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	38	TYR	CB-CG-CD2	-11.57	114.06	121.00	10	12
1	A	38	TYR	CB-CG-CD1	9.77	126.86	121.00	10	8
1	A	21	PHE	CB-CG-CD2	-8.90	114.57	120.80	10	5
1	A	13	TYR	CB-CG-CD2	-8.82	115.71	121.00	10	17
1	A	76	PHE	CB-CG-CD2	-8.39	114.93	120.80	9	11
1	A	21	PHE	CB-CG-CD1	8.21	126.54	120.80	10	2
1	A	76	PHE	CB-CG-CD1	8.13	126.49	120.80	9	9
1	A	67	ARG	NE-CZ-NH1	7.94	124.27	120.30	15	14
1	A	20	ARG	NE-CZ-NH1	7.84	124.22	120.30	7	11
1	A	29	ARG	NE-CZ-NH1	7.63	124.11	120.30	30	15
1	A	66	PHE	CB-CG-CD2	-7.47	115.57	120.80	10	24
1	A	59	PHE	CB-CG-CD2	-7.21	115.76	120.80	19	2
1	A	66	PHE	CB-CG-CD1	7.20	125.84	120.80	12	25

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	59	PHE	CB-CG-CD1	7.13	125.79	120.80	19	2
2	B	214	ARG	NE-CZ-NH1	7.07	123.84	120.30	11	14
1	A	13	TYR	CB-CG-CD1	6.94	125.17	121.00	22	7
2	B	215	ARG	NE-CZ-NH1	6.89	123.75	120.30	28	15
2	B	214	ARG	NE-CZ-NH2	-6.83	116.89	120.30	4	6
1	A	29	ARG	NE-CZ-NH2	-6.48	117.06	120.30	30	2
1	A	7	PHE	CB-CG-CD2	-6.20	116.46	120.80	6	2
1	A	7	PHE	CB-CG-CD1	5.91	124.94	120.80	6	2
1	A	35	LEU	CB-CG-CD2	5.85	120.94	111.00	17	8
2	B	215	ARG	NE-CZ-NH2	-5.77	117.42	120.30	6	5
1	A	67	ARG	NE-CZ-NH2	-5.72	117.44	120.30	19	5
2	B	204	MET	CG-SD-CE	-5.64	91.18	100.20	25	2
1	A	20	ARG	NE-CZ-NH2	-5.28	117.66	120.30	6	2
2	B	206	LEU	CB-CA-C	5.24	120.15	110.20	30	1
1	A	31	PHE	CB-CG-CD2	-5.21	117.15	120.80	12	3
1	A	31	PHE	CB-CG-CD1	5.03	124.32	120.80	12	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	29	ARG	Sidechain	8
1	A	38	TYR	Sidechain	5
2	B	215	ARG	Sidechain	5
2	B	214	ARG	Sidechain	3
1	A	67	ARG	Sidechain	2
1	A	28	TYR	Sidechain	2
1	A	13	TYR	Sidechain	2
1	A	20	ARG	Sidechain	2
2	B	211	TYR	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	507	482	480	1±1
2	B	127	128	128	1±1
All	All	19020	18300	18240	31

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:39:GLN:HG2	2:B:206:LEU:HD12	0.61	1.73	30	1
1:A:35:LEU:HB2	2:B:206:LEU:HD11	0.52	1.82	10	8
1:A:34:ILE:HG22	1:A:58:VAL:HG23	0.51	1.83	8	1
1:A:35:LEU:HD13	2:B:202:ILE:HG23	0.47	1.85	16	6
1:A:62:VAL:HA	1:A:65:LEU:CD2	0.46	2.40	26	2
1:A:62:VAL:HA	1:A:65:LEU:HD23	0.46	1.87	2	2
1:A:35:LEU:HD23	1:A:58:VAL:HG22	0.46	1.88	8	1
1:A:61:GLU:O	1:A:65:LEU:HD12	0.45	2.11	8	4
1:A:35:LEU:HD12	2:B:206:LEU:HD21	0.42	1.91	13	1
1:A:76:PHE:CE1	2:B:202:ILE:HD11	0.41	2.50	21	1
1:A:34:ILE:HD13	1:A:61:GLU:HB3	0.41	1.93	2	1
1:A:62:VAL:O	1:A:65:LEU:HD23	0.41	2.16	6	1
1:A:27:ILE:HD13	1:A:27:ILE:N	0.41	2.31	9	1
1:A:55:GLU:H	1:A:55:GLU:CD	0.40	2.20	15	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	60/85 (71%)	58±1 (96±1%)	2±1 (4±1%)	0±0 (0±1%)	44	80
2	B	15/24 (62%)	14±0 (96±3%)	1±0 (4±3%)	0±0 (0±0%)	100	100
All	All	2250/3270 (69%)	2163 (96%)	82 (4%)	5 (0%)	50	82

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	67	ARG	5

### 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	56/78 (72%)	48±2 (85±3%)	8±2 (15±3%)	6	44
2	B	13/21 (62%)	12±1 (94±6%)	1±1 (6±6%)	25	74
All	All	2070/2970 (70%)	1799 (87%)	271 (13%)	7	48

All 40 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	58	VAL	30
1	A	60	THR	30
1	A	39	GLN	25
1	A	37	THR	20
1	A	57	GLU	18
1	A	27	ILE	15
1	A	32	LEU	14
1	A	76	PHE	12
1	A	38	TYR	12
1	A	18	LYS	10
1	A	67	ARG	9
1	A	31	PHE	7
1	A	29	ARG	6
1	A	55	GLU	5
1	A	7	PHE	5
2	B	202	ILE	4
2	B	212	LEU	4
1	A	28	TYR	3
2	B	203	GLN	3
2	B	210	ASP	3
1	A	35	LEU	3
1	A	20	ARG	3
1	A	33	GLU	3
1	A	15	ASN	3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	64	ASN	3
2	B	215	ARG	2
1	A	16	LYS	2
2	B	204	MET	2
1	A	65	LEU	2
1	A	23	ASP	2
2	B	213	GLU	2
2	B	214	ARG	1
1	A	71	ASP	1
2	B	201	ASN	1
1	A	59	PHE	1
1	A	13	TYR	1
1	A	36	HIS	1
1	A	21	PHE	1
1	A	69	GLN	1
1	A	70	GLU	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 monosaccharides 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

There are no chain breaks in this entry.

## 7 Chemical shift validation

No chemical shift data were provided