



# Full wwPDB NMR Structure Validation Report ⓘ

Apr 26, 2016 – 01:58 PM BST

PDB ID : 1BYN  
Title : SOLUTION STRUCTURE OF THE CALCIUM-BOUND FIRST C2-DOMAIN OF SYNAPTOTAGMIN I  
Authors : Shao, X.; Fernandez, I.; Sudhof, T.C.; Rizo, J.  
Deposited on : 1998-10-18

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

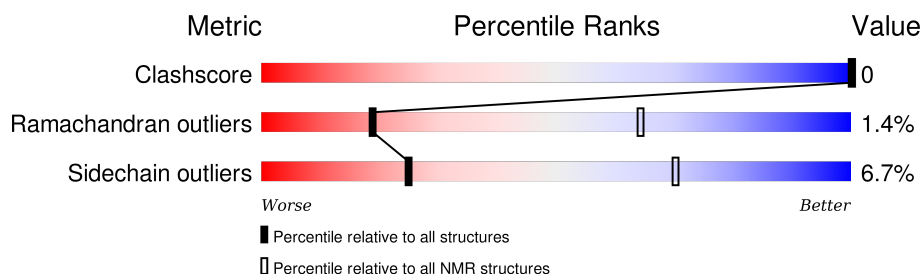
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment 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	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	128	 89% 6% • •

## 2 Ensemble composition and analysis

This entry contains 20 models. Model 6 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:142-A:264 (123)	0.25	6

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 3 single-model clusters were found.

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

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

There are 2 unique types of molecules in this entry. The entry contains 2084 atoms, of which 1038 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called PROTEIN (SYNAPTOTAGMIN I).

Mol	Chain	Residues	Atoms						Trace
1	A	128	Total	C	H	N	O	S	0
			2081	677	1038	169	194	3	

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	188	ASP	GLU	CONFLICT	UNP P21707

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

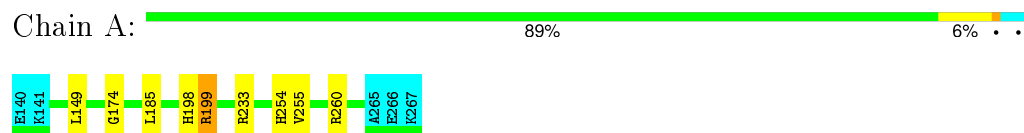
Mol	Chain	Residues	Atoms	
2	A	3	Total	Ca
			3	3

## 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: PROTEIN (SYNAPTOTAGMIN I)

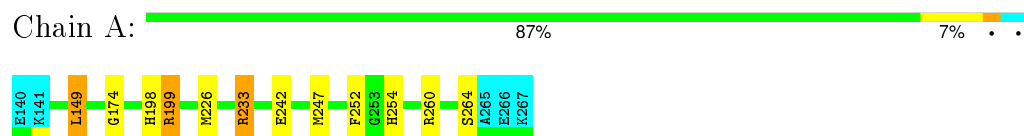


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

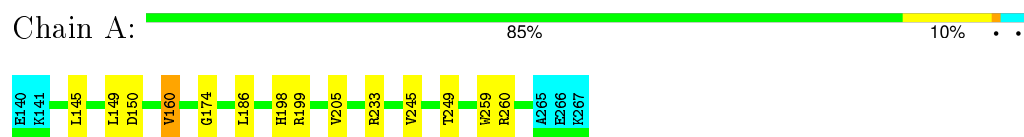
#### 4.2.1 Score per residue for model 1

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



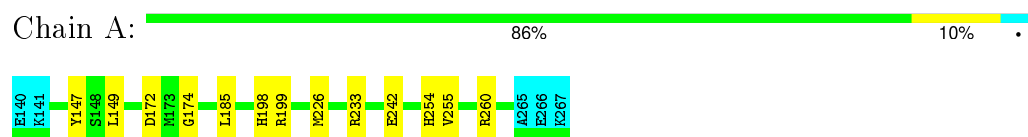
#### 4.2.2 Score per residue for model 2

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



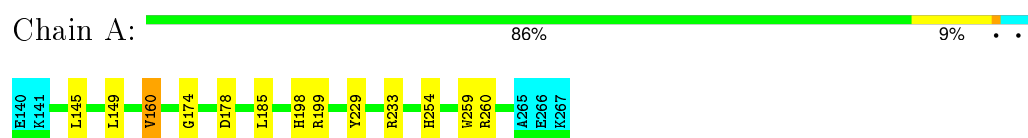
### 4.2.3 Score per residue for model 3

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



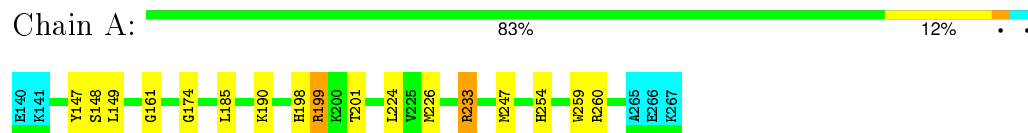
### 4.2.4 Score per residue for model 4

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



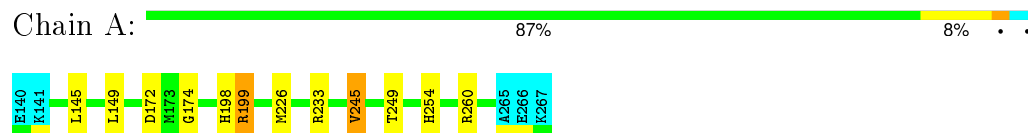
### 4.2.5 Score per residue for model 5

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



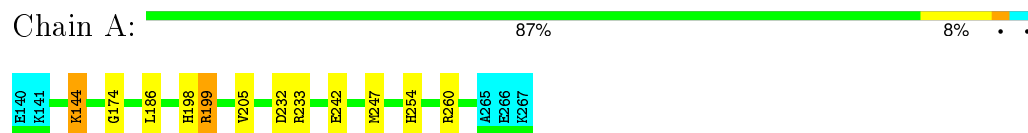
### 4.2.6 Score per residue for model 6 (medoid)

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



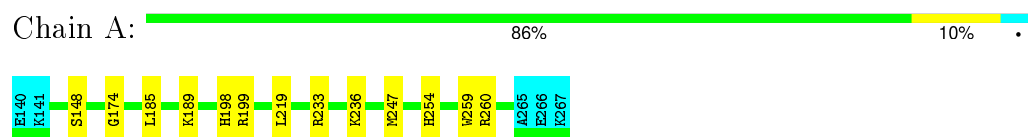
### 4.2.7 Score per residue for model 7

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



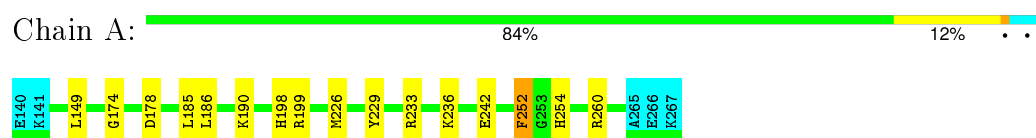
#### 4.2.8 Score per residue for model 8

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



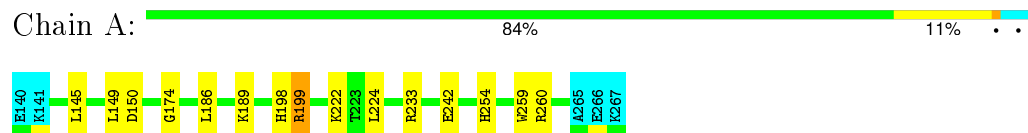
#### 4.2.9 Score per residue for model 9

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



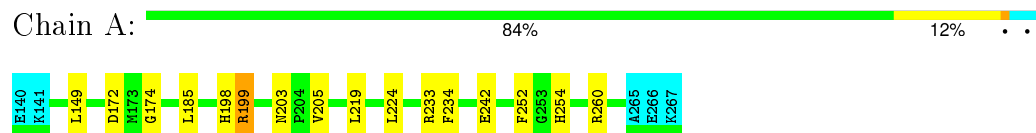
#### 4.2.10 Score per residue for model 10

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



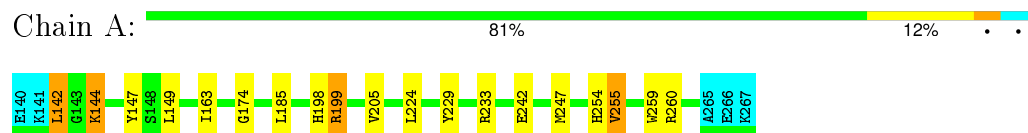
#### 4.2.11 Score per residue for model 11

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



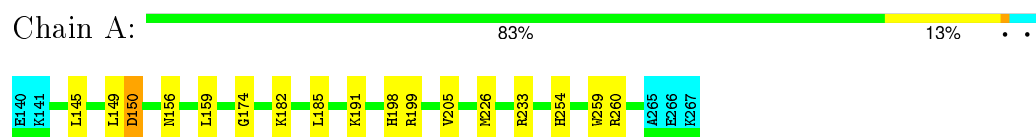
#### 4.2.12 Score per residue for model 12

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



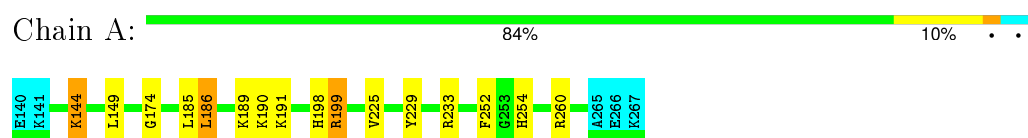
#### 4.2.13 Score per residue for model 13

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



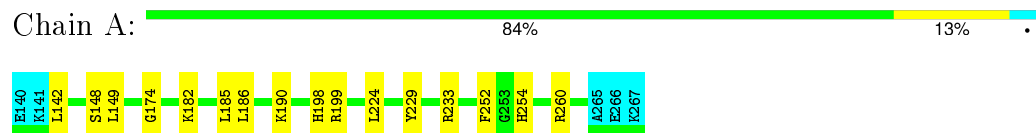
#### 4.2.14 Score per residue for model 14

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



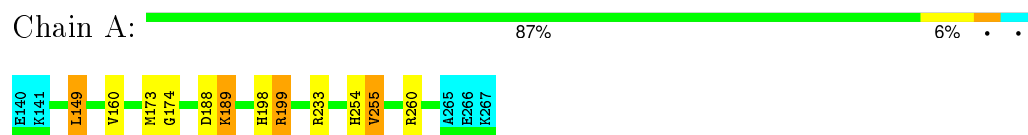
#### 4.2.15 Score per residue for model 15

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



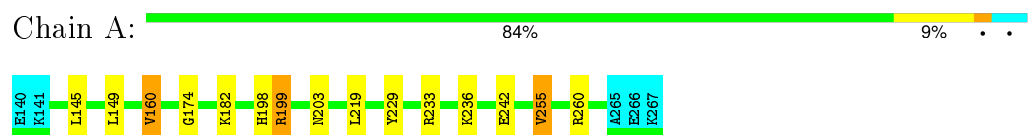
#### 4.2.16 Score per residue for model 16

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



#### 4.2.17 Score per residue for model 17

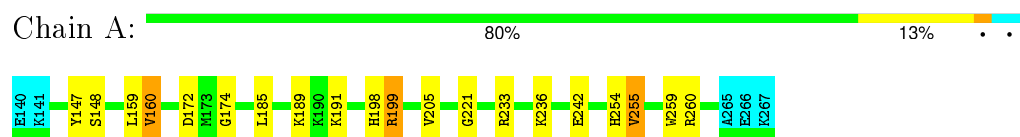
- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)





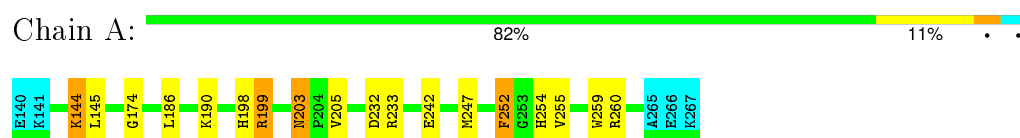
#### 4.2.18 Score per residue for model 18

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



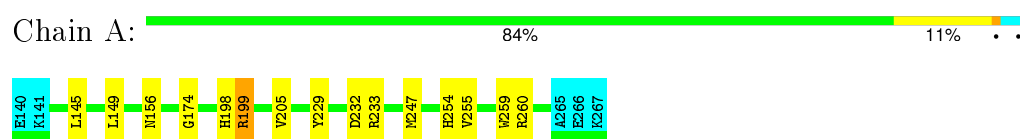
#### 4.2.19 Score per residue for model 19

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



#### 4.2.20 Score per residue for model 20

- Molecule 1: PROTEIN (SYNAPTOTAGMIN I)



## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *SIMULATED ANNEALING*.

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *FEWEST RESTRAINTS VIOLATIONS*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
DISCOVER	refinement	
DISCOVER	structure solution	
NMRCHITECT	structure solution	

No chemical shift data was provided. 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

Bond lengths and bond angles in the following residue types are not validated in this section: CA

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	0.97±0.01	0±0/1028 (0.0±0.0%)	1.25±0.03	8±1/1390 (0.6±0.1%)
All	All	0.97	0/20560 (0.0%)	1.25	157/27800 (0.6%)

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±0.4
All	All	0	3

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	260	ARG	NE-CZ-NH1	8.33	124.46	120.30	3	20
1	A	233	ARG	NE-CZ-NH1	8.11	124.36	120.30	20	20
1	A	199	ARG	NE-CZ-NH1	7.87	124.23	120.30	14	20
1	A	178	ASP	N-CA-CB	-6.66	98.61	110.60	4	2
1	A	160	VAL	CA-CB-CG1	6.60	120.80	110.90	16	1
1	A	252	PHE	N-CA-CB	-6.55	98.81	110.60	14	6
1	A	245	VAL	CA-CB-CG1	6.31	120.36	110.90	2	1
1	A	142	LEU	CB-CG-CD2	6.29	121.69	111.00	12	1
1	A	245	VAL	CA-CB-CG2	6.28	120.32	110.90	6	2
1	A	236	LYS	N-CA-CB	-6.18	99.48	110.60	8	1
1	A	201	THR	CA-CB-CG2	6.04	120.86	112.40	5	1
1	A	254	HIS	CB-CA-C	6.03	122.47	110.40	6	4

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	149	LEU	N-CA-CB	-5.94	98.52	110.40	6	5
1	A	159	LEU	N-CA-CB	-5.84	98.72	110.40	18	1
1	A	144	LYS	CA-CB-CG	5.76	126.08	113.40	14	4
1	A	186	LEU	N-CA-CB	-5.69	99.03	110.40	10	7
1	A	150	ASP	CA-CB-CG	5.55	125.61	113.40	13	1
1	A	190	LYS	N-CA-CB	-5.50	100.70	110.60	9	1
1	A	219	LEU	CA-CB-CG	5.41	127.75	115.30	17	2
1	A	147	TYR	CB-CG-CD2	-5.39	117.76	121.00	3	1
1	A	198	HIS	CG-ND1-CE1	-5.34	98.75	105.70	19	20
1	A	186	LEU	CA-CB-CG	5.28	127.45	115.30	15	1
1	A	160	VAL	CA-CB-CG2	5.25	118.78	110.90	4	4
1	A	190	LYS	CB-CA-C	5.24	120.88	110.40	9	2
1	A	259	TRP	CD1-NE1-CE2	-5.17	104.35	109.00	20	10
1	A	260	ARG	NE-CZ-NH2	-5.15	117.72	120.30	3	1
1	A	203	ASN	N-CA-CB	-5.11	101.40	110.60	19	1
1	A	254	HIS	CG-ND1-CE1	-5.11	99.06	105.70	16	16
1	A	159	LEU	CB-CA-C	5.05	119.79	110.20	18	1

There are no chirality outliers.

All unique planar outliers are listed below.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	147	TYR	Sidechain	3

## 6.2 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	1002	994	994	0±0
All	All	20100	19880	19880	1

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:186:LEU:HD13	1:A:225:VAL:HG21	0.47	1.85	14	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	123/128 (96%)	109±2 (89±1%)	12±2 (10±2%)	2±1 (1±1%)	19	64
All	All	2460/2560 (96%)	2184 (89%)	242 (10%)	34 (1%)	19	64

All 8 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	174	GLY	20
1	A	255	VAL	6
1	A	156	ASN	2
1	A	233	ARG	2
1	A	188	ASP	1
1	A	189	LYS	1
1	A	161	GLY	1
1	A	221	GLY	1

### 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	111/115 (97%)	104±2 (93±2%)	7±2 (7±2%)	25	70
All	All	2220/2300 (97%)	2071 (93%)	149 (7%)	25	70

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	149	LEU	13
1	A	199	ARG	13
1	A	185	LEU	11
1	A	242	GLU	10
1	A	145	LEU	8
1	A	205	VAL	8
1	A	247	MET	7
1	A	229	TYR	7
1	A	226	MET	6
1	A	224	LEU	5
1	A	189	LYS	5
1	A	255	VAL	5
1	A	148	SER	4
1	A	160	VAL	4
1	A	172	ASP	4
1	A	144	LYS	4
1	A	232	ASP	3
1	A	203	ASN	3
1	A	150	ASP	3
1	A	190	LYS	3
1	A	236	LYS	3
1	A	182	LYS	3
1	A	191	LYS	3
1	A	249	THR	2
1	A	142	LEU	2
1	A	252	PHE	2
1	A	264	SER	1
1	A	222	LYS	1
1	A	245	VAL	1
1	A	159	LEU	1
1	A	219	LEU	1
1	A	234	PHE	1
1	A	173	MET	1
1	A	163	ILE	1

### 6.3.3 RNA ⓘ

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](#)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

## 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

No chemical shift data were provided