Full wwPDB NMR Structure Validation Report (i)

Mar 1, 2022 – 10:38 AM EST

PDB ID	:	2G5M
Title	:	Spinophilin PDZ domain
Authors	:	Kelker, M.S.; Peti, W.
Deposited on	:	2006-02-23

This is a Full wwPDB NMR Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *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 (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

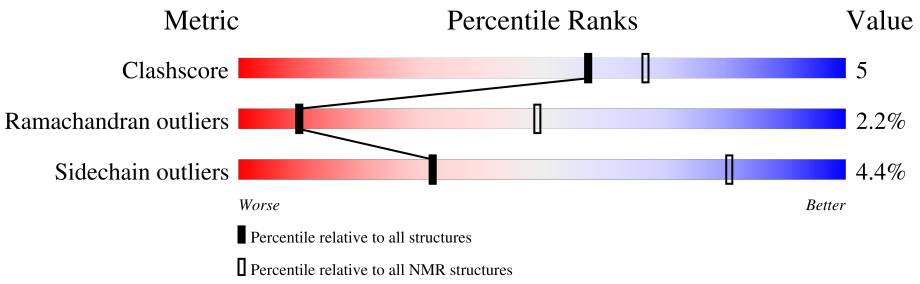
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

1 Overall quality at a glance (i)

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 >=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 <=5%

Mol	Chain	Length	Quality of chain
1	B	113	62% 10% 28%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 11 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *closest to the average*.

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	B:3-B:24, B:35-B:93 (81)	0.43	11

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 5 clusters and 2 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1706 atoms, of which 857 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Neurabin-2.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	B	113	1706	525	857	151	169	4	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	GLY	-	cloning artifact	UNP O35274
B	2	HIS	-	cloning artifact	UNP O35274
B	3	MET	-	cloning artifact	UNP O35274

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: Neurabin-2

Chain B: 62% 10% 28%

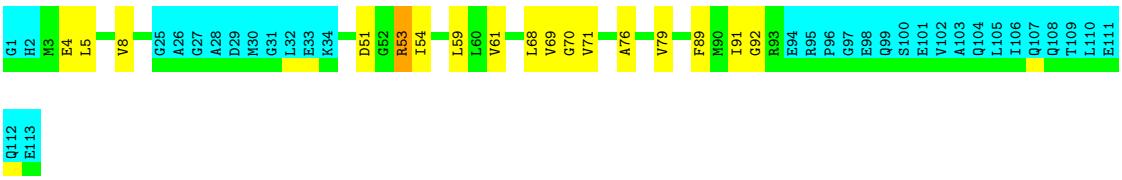
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Neurabin-2

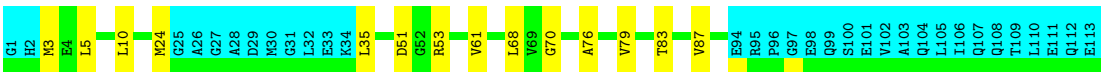
Chain B: 57% 14% • 28%



4.2.2 Score per residue for model 2

• Molecule 1: Neurabin-2

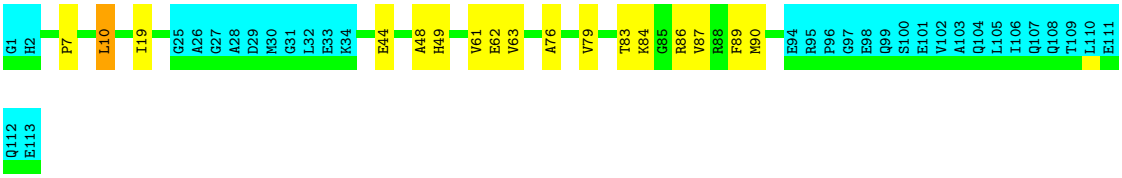
Chain B: 59% 12% 28%



4.2.3 Score per residue for model 3

- Molecule 1: Neurabin-2

Chain B: 57% 14% • 28%



4.2.4 Score per residue for model 4

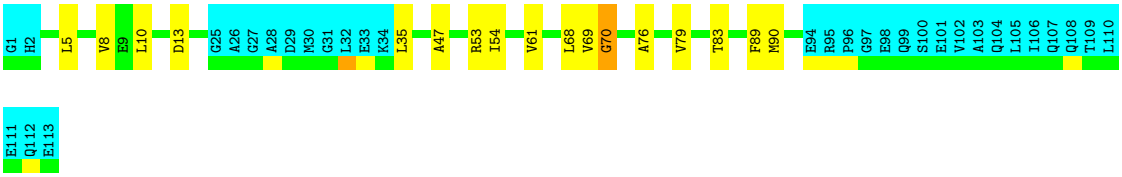
- Molecule 1: Neurabin-2

Chain B: 58% 12% • 28%

4.2.5 Score per residue for model 5

- Molecule 1: Neurabin-2

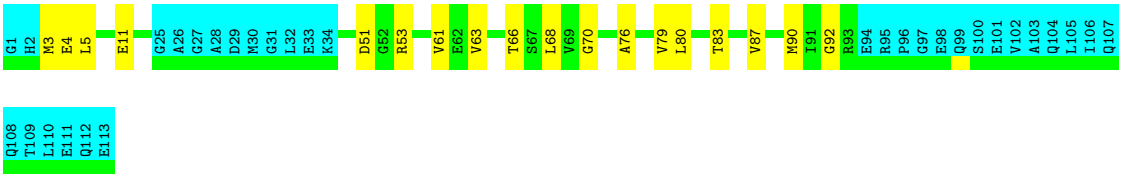
Chain B: 57% 14% • 28%



4.2.6 Score per residue for model 6

- Molecule 1: Neurabin-2

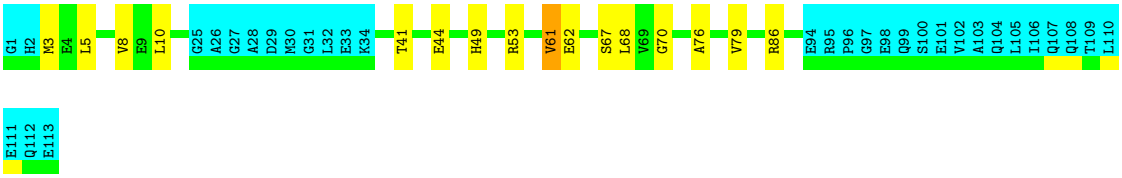
Chain B: 56% 16% 28%



4.2.7 Score per residue for model 7

- Molecule 1: Neurabin-2

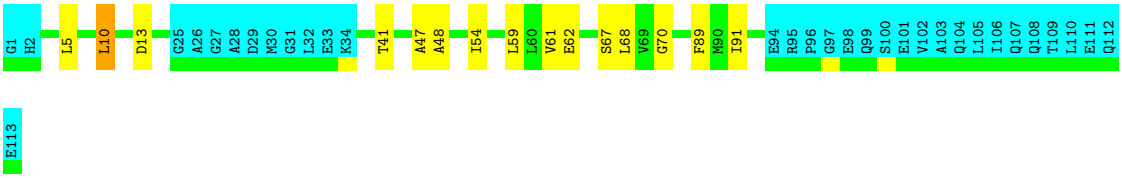
Chain B: 58% 13% • 28%



4.2.8 Score per residue for model 8

- Molecule 1: Neurabin-2

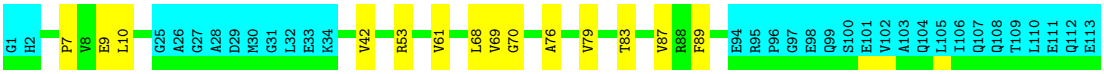
Chain B: 58% 12% • 28%



4.2.9 Score per residue for model 9

- Molecule 1: Neurabin-2

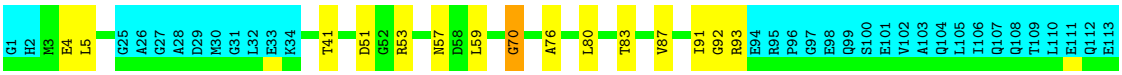
Chain B: 59% 12% 28%



4.2.10 Score per residue for model 10

- Molecule 1: Neurabin-2

Chain B: 58% 12% • 28%



4.2.11 Score per residue for model 11 (medoid)

- Molecule 1: Neurabin-2

Chain B: 54% 17% • 28%

G1 H2 M3 E4 L5 F6 P7 G25 A26 G27 A28 D29 M30 G31 L32 E33 K34 T41 R53 I54 V61 E62 V63 D64 G65 T66 S67 L68 V69 G70 A76 V79 N82 T83 V87 R88 F89 M90 R93 E94 R95 P96 G97 E98 Q99 S100 E101 V102 A103 Q104 L105 I106 Q107 Q108 T109 L110 E111 Q112 E113

4.2.12 Score per residue for model 12

- Molecule 1: Neurabin-2

Chain B: 61% 10% • 28%

G1 H2 L5 L10 E11 K12 G16 G25 A26 G27 A28 D29 M30 G31 L32 E33 K34 V42 I54 L59 G70 R86 F89 M90 I91 E94 R95 P96 G97 E98 Q99 S100 E101 V102 A103 Q104 L105 I106 Q107 Q108 T109 L110 E111 Q112 E113

4.2.13 Score per residue for model 13

- Molecule 1: Neurabin-2

Chain B: 60% 11% • 28%

G1 H2 P7 L10 G25 A26 G27 A28 D29 M30 G31 L32 E33 K34 L35 T41 A48 V61 G70 A76 V79 T83 V87 R88 F89 M90 E94 R95 P96 G97 E98 Q99 S100 E101 V102 A103 Q104 L105 I106 Q107 Q108 T109 L110 E111 Q112 E113

4.2.14 Score per residue for model 14

- Molecule 1: Neurabin-2

Chain B: 59% 12% • 28%

G1 H2 L5 L10 G18 I19 G25 A26 G27 A28 D29 M30 G31 L32 E33 K34 I54 N57 V61 L68 A76 V79 T83 V87 R88 F89 M90 E94 R95 P96 G97 E98 Q99 S100 E101 V102 A103 Q104 L105 I106 Q107 Q108 T109 L110 E111 Q112 E113

4.2.15 Score per residue for model 15

- Molecule 1: Neurabin-2

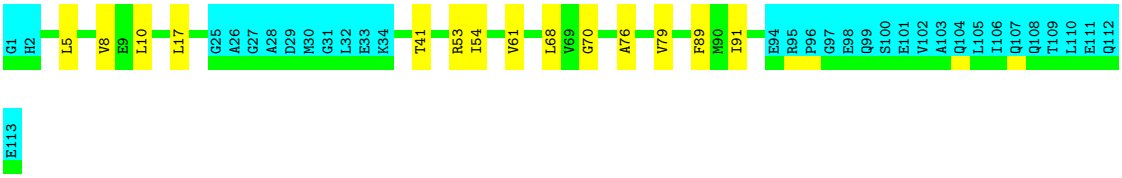
Chain B: 59% 12% • 28%

G1 H2 L5 L10 S14 I19 G25 A26 G27 A28 D29 M30 G31 L32 E33 K34 K40 R53 V61 E62 S67 L68 V69 G70 A76 V79 R86 E94 R95 P96 G97 E98 Q99 S100 E101 V102 A103 Q104 L105 I106 Q107 Q108 T109 L110 E111 Q112 E113

4.2.16 Score per residue for model 16

- Molecule 1: Neurabin-2

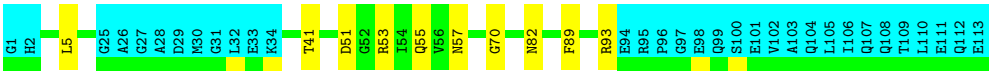
Chain B: 59% 12% 28%



4.2.17 Score per residue for model 17

- Molecule 1: Neurabin-2

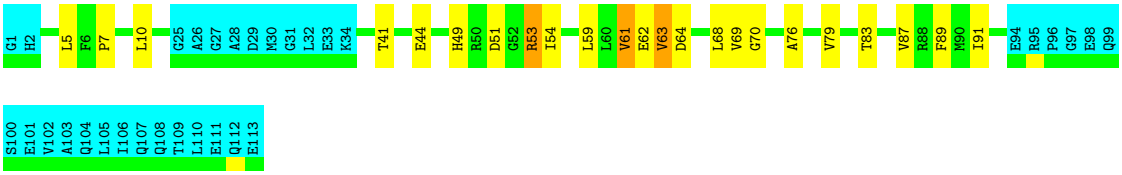
Chain B: 63% 9% 28%



4.2.18 Score per residue for model 18

- Molecule 1: Neurabin-2

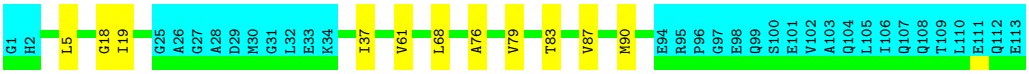
Chain B: 51% 18% • 28%



4.2.19 Score per residue for model 19

- Molecule 1: Neurabin-2

Chain B: 62% 10% 28%

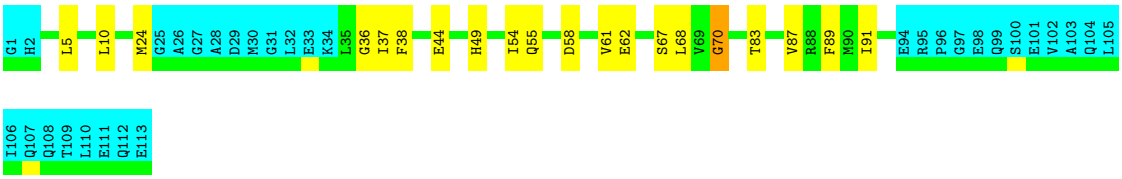


4.2.20 Score per residue for model 20

- Molecule 1: Neurabin-2

Chain B:

54%
17%
•
28%


G1
H2
L5
L10
M24
G25
A26
G27
A28
D29
M30
G31
L32
E33
K34
L35
G36
I37
F38
E44
H49
I54
Q55
D58
V61
E62
S67
L68
V69
G70
T83
V87
R88
F89
M90
I91
E94
R95
P96
G97
E98
Q99
S100
E101
V102
A103
Q104
L105
I106
Q107
Q108
T109
L110
E111
Q112
E113

5 Refinement protocol and experimental data overview (i)

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

Of the 50 calculated structures, 20 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
CYANA	structure solution	2.0
CNS	refinement	1.1
atnos/candid	structure solution	2.0

No chemical shift data was provided.

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	B	0.0±0.0	0.3±0.7
All	All	0	6

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	B	86	ARG	Sidechain	3
1	B	93	ARG	Sidechain	2
1	B	50	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	B	609	631	631	6±2
All	All	12180	12620	12620	123

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 5.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:B:55:GLN:HB2	1:B:58:ASP:OD1	0.62	1.93	20	1
1:B:61:VAL:O	1:B:68:LEU:HB2	0.62	1.94	7	15
1:B:83:THR:HB	1:B:87:VAL:HG11	0.59	1.74	10	12
1:B:76:ALA:O	1:B:79:VAL:HG22	0.58	1.98	14	2
1:B:76:ALA:O	1:B:79:VAL:HB	0.56	2.01	13	1
1:B:24:MET:HB2	1:B:38:PHE:CD1	0.55	2.35	20	1
1:B:54:ILE:HD11	1:B:89:PHE:CE2	0.53	2.38	16	9
1:B:53:ARG:HA	1:B:53:ARG:NE	0.53	2.19	10	3
1:B:51:ASP:OD2	1:B:53:ARG:HD3	0.53	2.03	6	3
1:B:8:VAL:HG13	1:B:53:ARG:HB3	0.51	1.81	7	3
1:B:59:LEU:O	1:B:91:ILE:HA	0.51	2.05	18	5
1:B:61:VAL:HB	1:B:90:MET:O	0.50	2.06	6	6
1:B:8:VAL:HG13	1:B:53:ARG:HG3	0.49	1.84	1	1
1:B:51:ASP:OD1	1:B:53:ARG:HG2	0.49	2.07	10	3
1:B:76:ALA:O	1:B:80:LEU:HG	0.49	2.07	10	2
1:B:7:PRO:HA	1:B:89:PHE:O	0.49	2.07	13	5
1:B:35:LEU:O	1:B:70:GLY:HA2	0.49	2.08	13	2
1:B:11:GLU:OE2	1:B:86:ARG:HD3	0.49	2.07	12	1
1:B:76:ALA:O	1:B:79:VAL:HG12	0.48	2.08	7	12
1:B:44:GLU:HA	1:B:49:HIS:CG	0.47	2.44	20	3
1:B:4:GLU:O	1:B:92:GLY:HA2	0.46	2.10	6	3
1:B:63:VAL:HG12	1:B:66:THR:HB	0.45	1.87	6	2
1:B:13:ASP:HB2	1:B:47:ALA:N	0.45	2.26	5	1
1:B:62:GLU:CB	1:B:67:SER:HA	0.45	2.42	20	3
1:B:62:GLU:HG2	1:B:63:VAL:O	0.44	2.11	18	2
1:B:10:LEU:HD11	1:B:48:ALA:HB2	0.44	1.88	13	3
1:B:44:GLU:H	1:B:44:GLU:CD	0.43	2.17	7	1
1:B:9:GLU:O	1:B:53:ARG:HG3	0.43	2.13	9	1
1:B:5:LEU:HA	1:B:91:ILE:O	0.42	2.14	20	2
1:B:24:MET:O	1:B:35:LEU:HG	0.42	2.15	2	1
1:B:64:ASP:HB2	1:B:87:VAL:HB	0.42	1.90	11	2
1:B:44:GLU:HA	1:B:49:HIS:HB3	0.42	1.92	3	1
1:B:13:ASP:HB2	1:B:47:ALA:HB2	0.42	1.91	8	1
1:B:68:LEU:HD12	1:B:71:VAL:CG2	0.41	2.45	1	1
1:B:79:VAL:O	1:B:83:THR:HG23	0.41	2.15	5	2
1:B:62:GLU:HB2	1:B:67:SER:HA	0.41	1.91	7	1
1:B:10:LEU:O	1:B:86:ARG:HG3	0.41	2.15	3	1
1:B:3:MET:HA	1:B:93:ARG:O	0.40	2.16	11	1
1:B:8:VAL:HB	1:B:91:ILE:HD13	0.40	1.94	16	1
1:B:4:GLU:OE2	1:B:93:ARG:HD2	0.40	2.16	10	1
1:B:10:LEU:HD21	1:B:48:ALA:HA	0.40	1.92	3	1
1:B:44:GLU:HA	1:B:49:HIS:CB	0.40	2.46	3	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	B	81/113 (72%)	69±3 (85±4%)	10±3 (12±3%)	2±1 (2±1%)	10 49
All	All	1620/2260 (72%)	1385 (85%)	199 (12%)	36 (2%)	10 49

All 12 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	B	70	GLY	16
1	B	69	VAL	4
1	B	19	ILE	4
1	B	42	VAL	2
1	B	18	GLY	2
1	B	37	ILE	2
1	B	84	LYS	1
1	B	16	GLY	1
1	B	14	SER	1
1	B	17	LEU	1
1	B	63	VAL	1
1	B	36	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	B	66/90 (73%)	63±1 (96±2%)	3±1 (4±2%)	32 81
All	All	1320/1800 (73%)	1262 (96%)	58 (4%)	32 81

All 14 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	B	5	LEU	16
1	B	10	LEU	14
1	B	41	THR	9
1	B	53	ARG	3
1	B	40	LYS	2
1	B	3	MET	2
1	B	61	VAL	2
1	B	57	ASN	2
1	B	82	ASN	2
1	B	89	PHE	2
1	B	11	GLU	1
1	B	12	LYS	1
1	B	55	GLN	1
1	B	90	MET	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 (i)

There are no chain breaks in this entry.

7 Chemical shift validation (i)

No chemical shift data were provided