WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB NMR Structure Validation Report ⓘ

May 7, 2024 – 01:23 pm BST

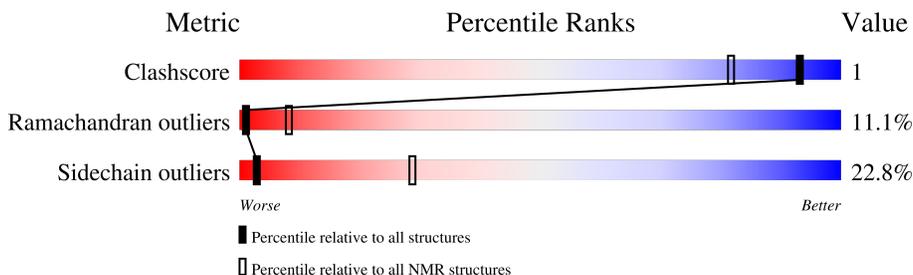|		
PDB ID	:	1HHN
Title	:	Calreticulin P-domain
Authors	:	Ellgaard, L.; Riek, R.; Herrmann, T.; Guntert, P.; Braun, D.; Helenius, A.; Wuthrich, K.
Deposited on	:	2000-12-22

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	Percentile Ranks	Value
Clashscore		1
Ramachandran outliers		11.1%
Sidechain outliers		22.8%

Worse — *Better*

▮ Percentile relative to all structures
▯ Percentile relative to all NMR structures

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	A	101	59% 19% • 20%

2 Ensemble composition and analysis ⓘ

This entry contains 20 models. Model 18 is the overall representative, medoid model (most similar to other models). The authors have identified model 4 as representative.

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:188-A:208,　A:262-A:288 (48)	1.43	18
2	A:221-A:253 (33)	0.50	15

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 5 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1595 atoms, of which 754 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called CALRETICULIN.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	101	1595	528	754	135	177	1	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	188	SER	PRO	cloning artifact	UNP P18418

4 Residue-property plots ⓘ

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: CALRETICULIN

Chain A:
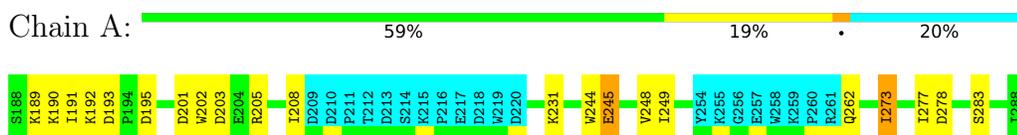
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: CALRETICULIN

Chain A:
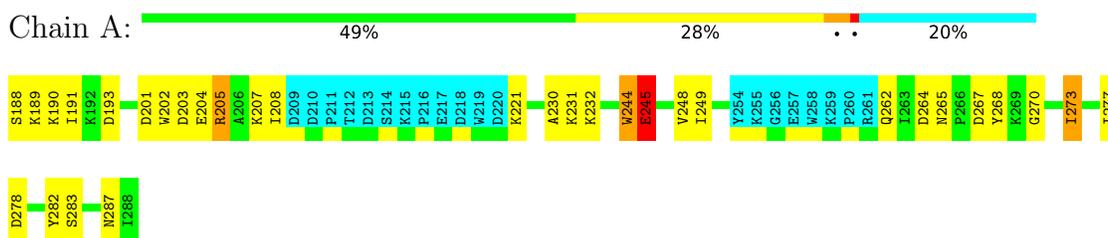
4.2.2 Score per residue for model 2

• Molecule 1: CALRETICULIN

Chain A:
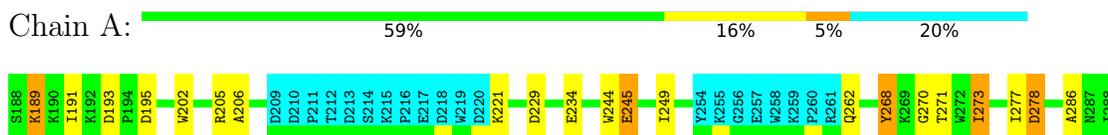
4.2.3 Score per residue for model 3

• Molecule 1: CALRETICULIN

Chain A:

4.2.4 Score per residue for model 4

• Molecule 1: CALRETICULIN

Chain A:

4.2.5 Score per residue for model 5

• Molecule 1: CALRETICULIN

Chain A:

4.2.6 Score per residue for model 6

• Molecule 1: CALRETICULIN

Chain A:

4.2.7 Score per residue for model 7

• Molecule 1: CALRETICULIN

Chain A:

S188 K189 K190 I191 K192 D193 P194 D195 E200 D201 W202 I208 D209 D210 P211 T212 D213 S214 K215 P216 E217 D218 W219 D220 D229 A230 K231 W244 E245 I249 Y254 K255 G256 E257 W258 K259 P260 R261 Q262 I263 D264 N265 P266 K269 I273 I277 D278 N279 S283

I288
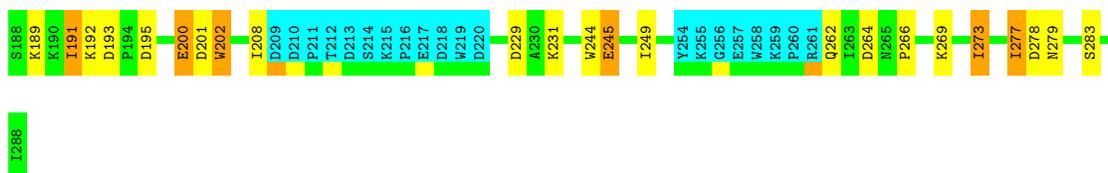
4.2.8　Score per residue for model 8

• Molecule 1: CALRETICULIN

Chain A:　　　55%　　　17%　　8%　　20%

S188 K189 K190 I191 K192 A196 E200 D201 W202 D203 E204 I208 D209 D210 P211 T212 D213 S214 K215 P216 E217 D218 W219 D220 D229 E239 W244 E245 V248 I249 Y254 K255 G256 E257 W258 K259 P260 R261 Q262 W272 I273 I277 D278 Y282 S283 N287 I288

4.2.9　Score per residue for model 9

• Molecule 1: CALRETICULIN

Chain A:　　　51%　　　24%　　5%　　20%

S188 K189 K190 I191 K192 D193 P194 D195 K198 P199 E200 D201 W202 D203 K207 I208 D209 D210 P211 T212 D213 S214 K215 P216 E217 D218 W219 D220 D229 A230 K231 K232 E243 W244 E245 V248 I249 Y254 K255 G256 E257 W258 K259 P260 R261 D264 D267 G270 T271 W272 I273

E276 I277 D278 D285 A286 N287 I288
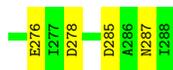
4.2.10　Score per residue for model 10

• Molecule 1: CALRETICULIN

Chain A:　　　53%　　　22%　　• •　　20%

S188 K189 K190 I191 K192 D193 P194 D195 D201 W202 D203 E204 R205 A206 K207 I208 D209 D210 P211 T212 D213 S214 K215 P216 E217 D218 W219 D220 E223 D227 W244 E245 Y254 K255 G256 E257 W258 K259 P260 R261 Q262 I263 D264 D267 Y268 K269 G270 I273 I277 S283 P284

D285 A286 N287 I288

4.2.11　Score per residue for model 11

• Molecule 1: CALRETICULIN

Chain A:　　　56%　　　19%　　5%　　20%
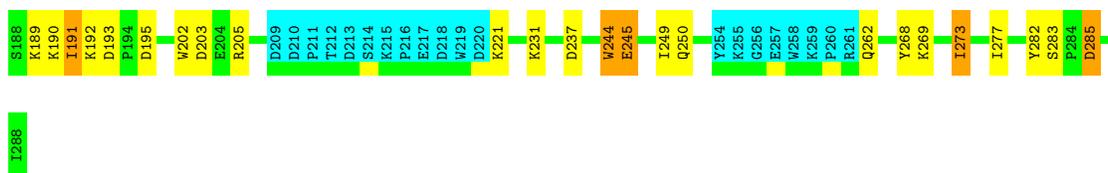
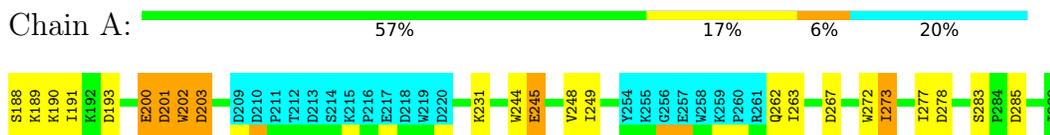### 4.2.12 Score per residue for model 12

• Molecule 1: CALRETICULIN

Chain A:

4.2.13 Score per residue for model 13
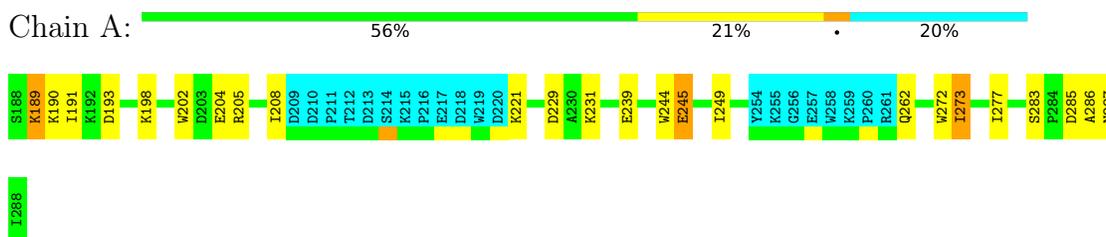
• Molecule 1: CALRETICULIN

Chain A:

4.2.14 Score per residue for model 14
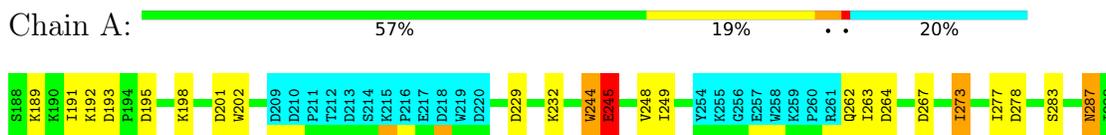
• Molecule 1: CALRETICULIN

Chain A:

4.2.15 Score per residue for model 15

• Molecule 1: CALRETICULIN

Chain A:

4.2.16 Score per residue for model 16

• Molecule 1: CALRETICULIN

Chain A:
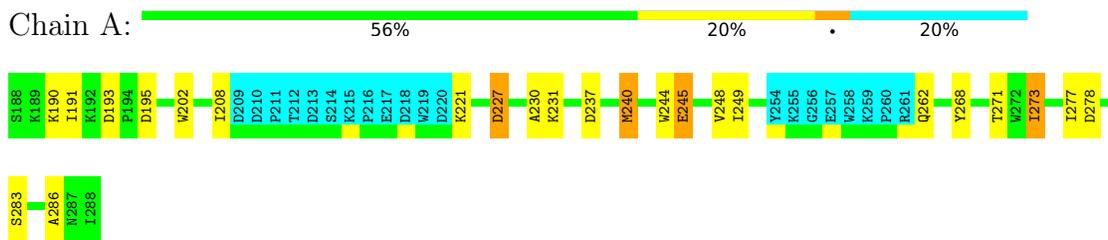
4.2.17 Score per residue for model 17

• Molecule 1: CALRETICULIN

Chain A:
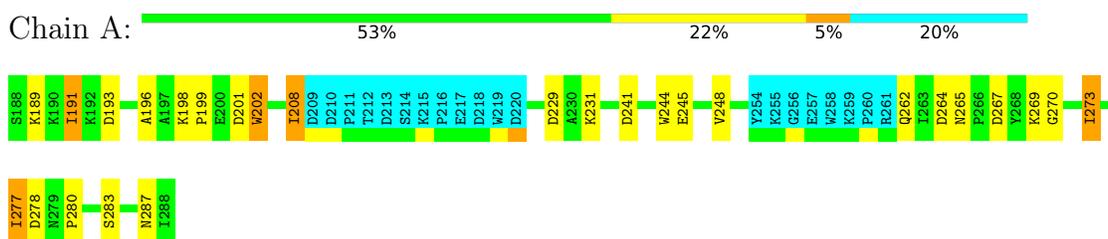
4.2.18 Score per residue for model 18 (medoid)

• Molecule 1: CALRETICULIN

Chain A:
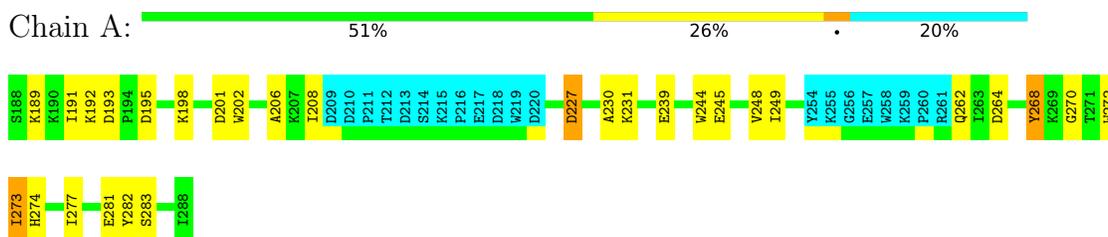
4.2.19 Score per residue for model 19

• Molecule 1: CALRETICULIN

Chain A:
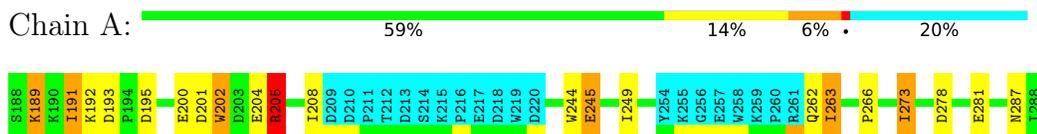
4.2.20 Score per residue for model 20

• Molecule 1: CALRETICULIN

Chain A:
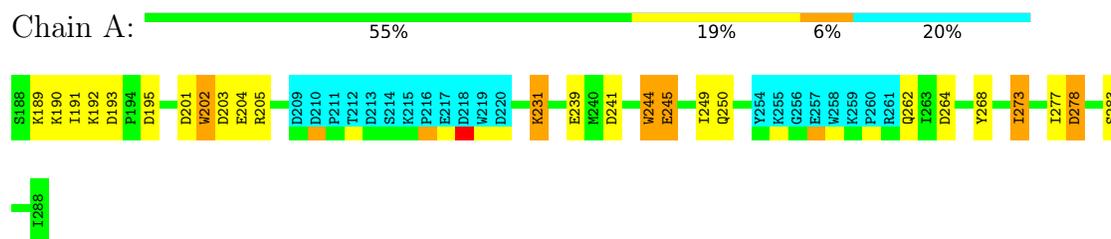
5 Refinement protocol and experimental data overview ⓘ

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

Of the ? calculated structures, 20 were deposited, based on the following criterion: *?*.

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

Software name	Classification	Version
OPALp	refinement	
DYANA	structure solution	

No chemical shift data was provided.

6 Model quality ⓘ

6.1 Standard geometry ⓘ

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	0.62±0.01	0±0/691 (0.0± 0.0%)	1.18±0.03	0±1/946 (0.0± 0.1%)
All	All	0.62	0/13820 (0.0%)	1.18	8/18920 (0.0%)

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	0.3±0.6
All	All	0	7

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(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	240	MET	CA-CB-CG	6.91	125.04	113.30	16	1
1	A	205	ARG	NE-CZ-NH1	5.87	123.23	120.30	19	2
1	A	205	ARG	NE-CZ-NH2	-5.54	117.53	120.30	15	2
1	A	264	ASP	CB-CG-OD1	5.50	123.25	118.30	6	1
1	A	282	TYR	CB-CG-CD2	-5.13	117.92	121.00	8	1
1	A	189	LYS	CB-CA-C	5.08	120.55	110.40	8	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	268	TYR	Sidechain	2
1	A	200	GLU	Peptide	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	266	PRO	Peptide	1
1	A	188	SER	Peptide	1
1	A	273	ILE	Peptide	1
1	A	205	ARG	Sidechain	1

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	667	605	604	2±1
All	All	13340	12100	12080	34

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	Models Total
1:A:227:ASP:HB2	1:A:230:ALA:HB2	0.57	1.77	16	2
1:A:191:ILE:HD13	1:A:192:LYS:H	0.54	1.62	11	4
1:A:208:ILE:HD12	1:A:265:ASN:HB2	0.54	1.78	17	1
1:A:189:LYS:HE2	1:A:282:TYR:CD2	0.53	2.37	1	1
1:A:231:LYS:HE3	1:A:231:LYS:H	0.48	1.68	20	2
1:A:244:TRP:CG	1:A:245:GLU:N	0.48	2.82	6	6
1:A:189:LYS:HE3	1:A:282:TYR:CD2	0.47	2.44	11	1
1:A:206:ALA:HB2	1:A:268:TYR:CE2	0.47	2.45	2	1
1:A:206:ALA:HA	1:A:268:TYR:CG	0.46	2.45	10	2
1:A:231:LYS:N	1:A:231:LYS:HE3	0.45	2.25	4	1
1:A:232:LYS:HB3	1:A:244:TRP:CD2	0.45	2.46	9	1
1:A:277:ILE:HD11	1:A:280:PRO:HD3	0.44	1.90	4	2
1:A:202:TRP:CG	1:A:202:TRP:O	0.43	2.71	4	1
1:A:198:LYS:HE3	1:A:199:PRO:O	0.43	2.14	12	1
1:A:244:TRP:CD1	1:A:245:GLU:HG3	0.42	2.49	2	1
1:A:283:SER:H	1:A:288:ILE:HD12	0.42	1.73	10	1
1:A:231:LYS:N	1:A:231:LYS:HE2	0.42	2.29	18	1
1:A:200:GLU:O	1:A:201:ASP:CB	0.41	2.69	8	1
1:A:234:GLU:CD	1:A:234:GLU:H	0.41	2.18	12	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:277:ILE:HD12	1:A:278:ASP:C	0.41	2.36	7	1
1:A:202:TRP:CD1	1:A:204:GLU:HB2	0.41	2.51	8	1
1:A:231:LYS:H	1:A:231:LYS:CE	0.40	2.29	20	1

6.3   Torsion angles ⓘ

6.3.1   Protein backbone ⓘ

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	79/101 (78%)	56±2 (71±3%)	14±3 (17±4%)	9±2 (11±2%)	1	8
All	All	1580/2020 (78%)	1129 (71%)	275 (17%)	176 (11%)	1	8

All 29 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	245	GLU	20
1	A	273	ILE	20
1	A	244	TRP	17
1	A	283	SER	16
1	A	201	ASP	11
1	A	190	LYS	10
1	A	272	TRP	8
1	A	270	GLY	7
1	A	189	LYS	7
1	A	205	ARG	6
1	A	278	ASP	6
1	A	203	ASP	6
1	A	202	TRP	5
1	A	287	ASN	5
1	A	285	ASP	5
1	A	200	GLU	4
1	A	191	ILE	4
1	A	267	ASP	3
1	A	286	ALA	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	196	ALA	2
1	A	276	GLU	2
1	A	204	GLU	2
1	A	230	ALA	1
1	A	271	THR	1
1	A	279	ASN	1
1	A	284	PRO	1
1	A	199	PRO	1
1	A	263	ILE	1
1	A	266	PRO	1

6.3.2 Protein sidechains ⓘ

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	74/93 (80%)	57±2 (77±3%)	17±2 (23±3%)	3	29
All	All	1480/1860 (80%)	1142 (77%)	338 (23%)	3	29

All 50 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	202	TRP	20
1	A	273	ILE	20
1	A	193	ASP	19
1	A	262	GLN	18
1	A	277	ILE	18
1	A	191	ILE	17
1	A	249	ILE	17
1	A	245	GLU	14
1	A	208	ILE	12
1	A	231	LYS	12
1	A	248	VAL	11
1	A	195	ASP	11
1	A	189	LYS	10
1	A	278	ASP	10
1	A	229	ASP	9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	203	ASP	8
1	A	264	ASP	8
1	A	192	LYS	8
1	A	221	LYS	7
1	A	198	LYS	6
1	A	268	TYR	5
1	A	227	ASP	5
1	A	239	GLU	5
1	A	204	GLU	4
1	A	205	ARG	4
1	A	232	LYS	4
1	A	287	ASN	4
1	A	201	ASP	4
1	A	263	ILE	4
1	A	285	ASP	4
1	A	271	THR	4
1	A	200	GLU	3
1	A	267	ASP	3
1	A	269	LYS	3
1	A	281	GLU	3
1	A	188	SER	2
1	A	207	LYS	2
1	A	234	GLU	2
1	A	223	GLU	2
1	A	237	ASP	2
1	A	250	GLN	2
1	A	288	ILE	2
1	A	282	TYR	2
1	A	241	ASP	2
1	A	265	ASN	1
1	A	190	LYS	1
1	A	283	SER	1
1	A	243	GLU	1
1	A	240	MET	1
1	A	274	HIS	1

6.3.3　RNA ⓘ

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.6 Ligand geometry ⓘ

There are no ligands in this entry.

6.7 Other polymers ⓘ

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 (i)

No chemical shift data were provided