Full wwPDB NMR Structure Validation Report (i)

Feb 20, 2022 – 02:49 AM EST

PDB ID	:	1U3M
Title	:	NMR structure of the chicken prion protein fragment 128-242
Authors	:	Lysek, D.A.; Calzolai, L.; Guntert, P.; Wuthrich, K.
Deposited on	:	2004-07-22

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:

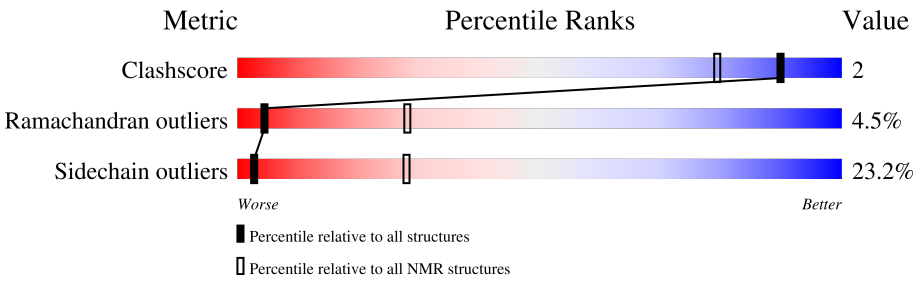
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.26
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.26

1 Overall quality at a glance (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	A	117	59% 21% 20%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 8 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *fewest violations*.

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:135-A:199, A:212-A:240 (94)	0.52	8

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 4 clusters and 5 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1819 atoms, of which 884 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called prion-like protein.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	117	1819	585	884	163	180	7	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	126	GLY	-	cloning artifact	UNP P27177
A	127	SER	-	cloning artifact	UNP P27177

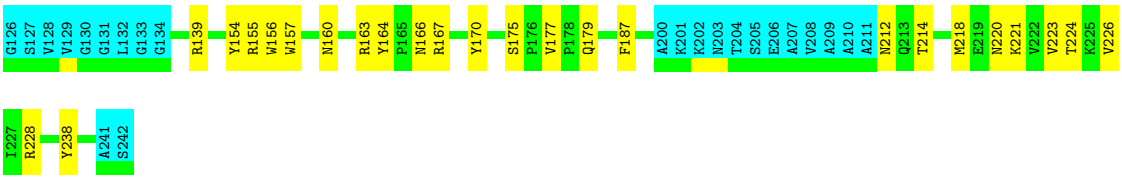
4 Residue-property plots (i)

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: prion-like protein

Chain A: 59% 21% 20%



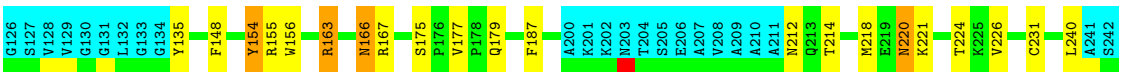
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: prion-like protein

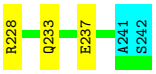
Chain A: 62% 15% • 20%



4.2.2 Score per residue for model 2

• Molecule 1: prion-like protein

Chain A: 58% 21% • 20%



4.2.3 Score per residue for model 3

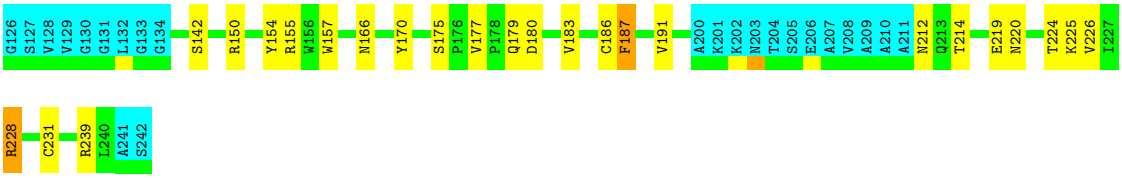
- Molecule 1: prion-like protein

Chain A:

4.2.4 Score per residue for model 4

- Molecule 1: prion-like protein

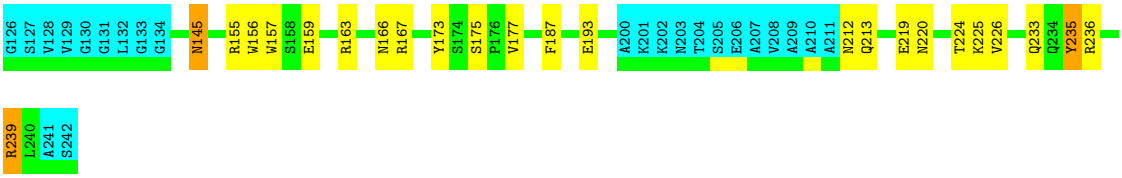
Chain A:



4.2.5 Score per residue for model 5

- Molecule 1: prion-like protein

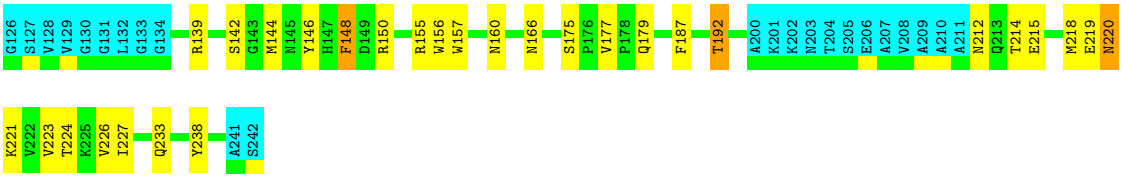
Chain A:



4.2.6 Score per residue for model 6

- Molecule 1: prion-like protein

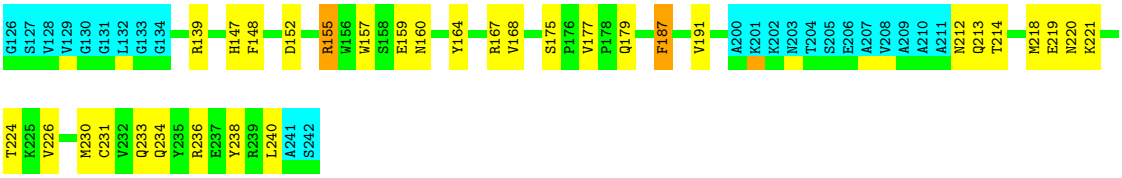
Chain A:



4.2.7 Score per residue for model 7

- Molecule 1: prion-like protein

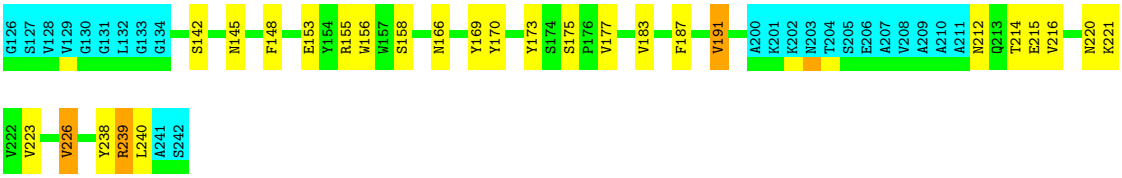
Chain A:



4.2.8 Score per residue for model 8 (medoid)

- Molecule 1: prion-like protein

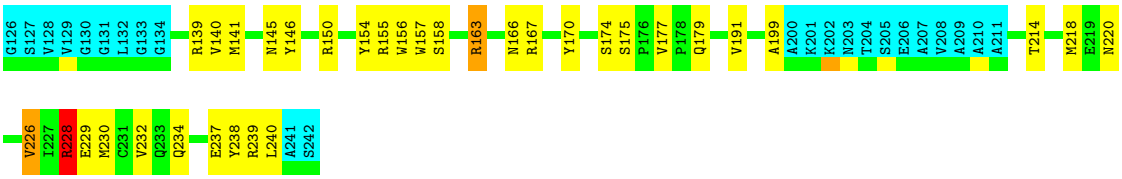
Chain A:



4.2.9 Score per residue for model 9

- Molecule 1: prion-like protein

Chain A:



4.2.10 Score per residue for model 10

- Molecule 1: prion-like protein

Chain A: 56% 22% • 20%

G126 S127 V128 V129 G130 G131 L132 G133 G134 Y135 R139 N145 Y146 H147 F148 D152 E153 Y154 R155 W156 W157 N160 R163 R167 S175 P176 V177 P178 Q179 F187 N188 A200 K201 K202 N203 T204 S205 E206 A207 V208 A209 A210 A211 N212 E215 E219 N220 V223

I227 M230 R236 E237 Y238 A241 S242

4.2.11 Score per residue for model 11

- Molecule 1: prion-like protein

Chain A: 60% 19% • 20%

G126 S127 V128 V129 G130 G131 L132 G133 G134 N145 Y154 R155 W156 N160 R163 N166 Y170 S175 P176 V177 P178 Q179 D180 V183 F187 A200 K201 K202 N203 T204 S205 E206 A207 V208 A209 A210 A211 T214 M218 E219 N220 V223 T224 R228 V232

Q233 Q234 Y238 A241 S242

4.2.12 Score per residue for model 12

- Molecule 1: prion-like protein

Chain A: 57% 20% • 20%

G126 S127 V128 V129 G130 G131 L132 G133 G134 V140 F148 D152 R155 W156 W157 N160 R163 R167 Y170 R171 S175 P176 V177 P178 Q179 F187 E193 Y194 A200 K201 K202 N203 T204 S205 E206 A207 V208 A209 A210 A211 N212 Q213 T214 M218 E219 N220 K221

T224 R228 C231 Y238 R239 L240 A241 S242

4.2.13 Score per residue for model 13

- Molecule 1: prion-like protein

Chain A: 50% 26% • • 20%

G126 S127 V128 V129 G130 G131 L132 G133 G134 R139 M144 N145 Y146 H147 F148 Y154 R155 W156 N160 R163 R167 R171 S175 P176 V177 P178 Q179 D180 V183 I189 T190 V191 T192 A200 K201 K202 N203 T204 S205 E206 A207 V208 A209 A210 A211 N212 Q213 T214 E215

M218 E219 N220 K221 V222 V223 V226 I227 R228 Q233 Q234 Y238 R239 L240 A241 S242

4.2.14 Score per residue for model 14

- Molecule 1: prion-like protein

Chain A: 59% 18% • 20%

G126 S127 V128 V129 G130 G131 L132 G133 G134 Y146 Y154 R155 W156 W157 S158 Y164 P165 N166 R167 Y170 R171 D172 Y173 S174 S175 P176 V177 P178 Q179 F187 T192 A200 K201 K202 N203 T204 S205 E206 A207 V208 A209 A210 A211 N212 Q213 T214 M218 E219 N220 T224 K225

V226 E237 Y238 R239 L240 A241 S242

4.2.15 Score per residue for model 15

- Molecule 1: prion-like protein

Chain A: 56% 21% • • 20%

G126 S127 V128 V129 G130 G131 L132 G133 G134 R139 D152 R155 W156 W157 S158 E159 N160 R163 Y164 P165 N166 R167 V168 D172 Y173 S174 S175 P176 V177 P178 Q179 D180 F187 T192 E193 A200 K201 K202 N203 T204 S205 E206 A207 V208 A209 A210 A211 N212 M218 E219 N220

V223 V226 I227 R228 E229 M230 Y238 A241 S242

4.2.16 Score per residue for model 16

- Molecule 1: prion-like protein

Chain A: 56% 21% • 20%

G126 S127 V128 V129 G130 G131 L132 G133 G134 R139 S142 G143 M144 N145 Y146 H147 R155 W156 Y164 P165 N166 R167 D172 S175 P176 V177 P178 Q179 V191 Y194 A200 K201 K202 N203 T204 S205 E206 A207 V208 A209 A210 A211 N212 Q213 T214 M218 E219 N220 K221 V222

V223 V226 I227 R228 E229 M230 C231 V232 Y238 A241 S242

4.2.17 Score per residue for model 17

- Molecule 1: prion-like protein

Chain A: 56% 22% • 20%

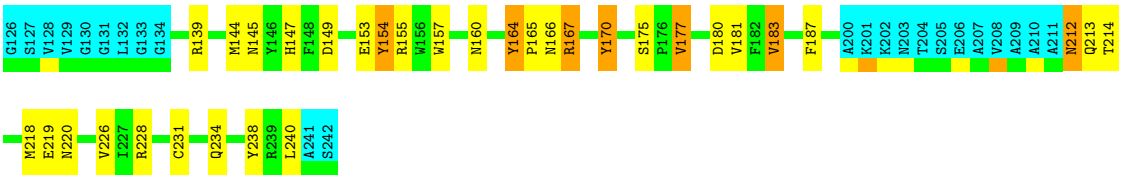
G126 S127 V128 V129 G130 G131 L132 G133 G134 S142 Y146 D149 Y154 R155 W156 W157 N160 Y164 R167 V168 Y169 Y170 R171 D172 S175 P176 V177 P178 Q179 F187 N188 A200 K201 K202 N203 T204 S205 E206 A207 V208 A209 A210 A211 N212 Q213 T214 N220 K221 V222

V223 T224 K225 V226 I227 M230 Y238 R239 L240 A241 S242

4.2.18 Score per residue for model 18

- Molecule 1: prion-like protein

Chain A: 52% 22% 6% 20%



4.2.19 Score per residue for model 19

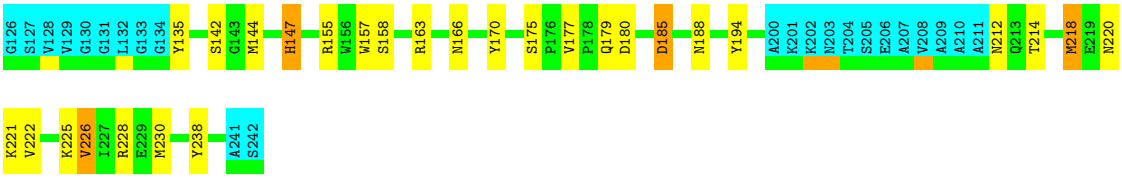
- Molecule 1: prion-like protein

Chain A: 52% 23% 5% 20%

4.2.20 Score per residue for model 20

- Molecule 1: prion-like protein

Chain A: 56% 21% • 20%



5 Refinement protocol and experimental data overview (i)

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

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *structures with the least restraint violations*.

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

Software name	Classification	Version
DYANA	structure solution	6.0
DYANA	refinement	6.0

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	0.71±0.01	0±0/819 (0.0± 0.0%)	1.16±0.05	2±2/1113 (0.2± 0.1%)
All	All	0.71	0/16380 (0.0%)	1.16	34/22260 (0.2%)

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	2.1±0.9
All	All	0	43

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	223	VAL	CA-CB-CG2	8.61	123.81	110.90	19	1
1	A	163	ARG	NE-CZ-NH2	-7.46	116.57	120.30	15	1
1	A	226	VAL	CA-CB-CG2	6.68	120.92	110.90	16	1
1	A	228	ARG	CD-NE-CZ	6.66	132.93	123.60	9	1
1	A	154	TYR	CB-CG-CD2	-6.51	117.09	121.00	18	3
1	A	168	VAL	CG1-CB-CG2	-6.22	100.95	110.90	7	1
1	A	191	VAL	CA-CB-CG2	5.96	119.84	110.90	4	4
1	A	164	TYR	CB-CG-CD2	-5.93	117.44	121.00	15	1
1	A	228	ARG	NE-CZ-NH1	5.77	123.18	120.30	9	1
1	A	185	ASP	CB-CG-OD2	-5.76	113.11	118.30	20	1
1	A	146	TYR	CB-CG-CD2	-5.65	117.61	121.00	13	1
1	A	192	THR	CA-CB-CG2	5.61	120.26	112.40	6	1
1	A	239	ARG	NE-CZ-NH2	-5.57	117.52	120.30	3	3
1	A	177	VAL	CB-CA-C	5.49	121.83	111.40	13	1
1	A	177	VAL	CG1-CB-CG2	5.49	119.68	110.90	13	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	187	PHE	CB-CG-CD2	-5.46	116.98	120.80	19	1
1	A	139	ARG	NE-CZ-NH2	-5.20	117.70	120.30	10	1
1	A	238	TYR	CB-CG-CD2	-5.18	117.89	121.00	15	1
1	A	191	VAL	CA-CB-CG1	5.18	118.66	110.90	8	1
1	A	150	ARG	CD-NE-CZ	5.15	130.81	123.60	6	1
1	A	163	ARG	NE-CZ-NH1	5.13	122.87	120.30	20	1
1	A	167	ARG	NE-CZ-NH1	5.13	122.87	120.30	3	1
1	A	226	VAL	CA-CB-CG1	5.12	118.58	110.90	9	3
1	A	139	ARG	NE-CZ-NH1	5.08	122.84	120.30	10	2

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	155	ARG	Sidechain	4
1	A	228	ARG	Sidechain	4
1	A	163	ARG	Sidechain	3
1	A	171	ARG	Sidechain	3
1	A	173	TYR	Sidechain	3
1	A	236	ARG	Sidechain	3
1	A	239	ARG	Sidechain	3
1	A	194	TYR	Sidechain	3
1	A	154	TYR	Sidechain	2
1	A	167	ARG	Sidechain	2
1	A	169	TYR	Sidechain	2
1	A	146	TYR	Sidechain	2
1	A	150	ARG	Sidechain	2
1	A	170	TYR	Sidechain	2
1	A	135	TYR	Sidechain	1
1	A	235	TYR	Sidechain	1
1	A	238	TYR	Sidechain	1
1	A	164	TYR	Sidechain	1
1	A	142	SER	Peptide	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	796	741	741	3±1
All	All	15920	14820	14820	51

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 2.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:187:PHE:CE1	1:A:224:THR:HG23	0.61	2.30	4	3
1:A:187:PHE:CZ	1:A:224:THR:HG22	0.60	2.31	19	1
1:A:226:VAL:HG12	1:A:230:MET:CE	0.56	2.31	17	2
1:A:220:ASN:O	1:A:224:THR:HG23	0.56	2.00	17	9
1:A:177:VAL:HG22	1:A:181:VAL:HG21	0.54	1.78	3	1
1:A:148:PHE:CE2	1:A:156:TRP:CE3	0.53	2.96	2	1
1:A:191:VAL:HG23	1:A:199:ALA:CB	0.52	2.35	9	1
1:A:228:ARG:O	1:A:232:VAL:HG23	0.50	2.06	9	3
1:A:183:VAL:HG23	1:A:231:CYS:HB2	0.49	1.84	18	1
1:A:187:PHE:CE1	1:A:191:VAL:HG11	0.48	2.43	19	1
1:A:226:VAL:HG12	1:A:230:MET:HE2	0.48	1.86	16	1
1:A:187:PHE:CE1	1:A:224:THR:HG22	0.47	2.44	19	1
1:A:164:TYR:CE1	1:A:226:VAL:HG21	0.46	2.45	19	1
1:A:220:ASN:HA	1:A:223:VAL:HG22	0.46	1.86	19	1
1:A:212:ASN:O	1:A:216:VAL:HG13	0.46	2.11	8	1
1:A:177:VAL:HG22	1:A:181:VAL:HG11	0.45	1.87	18	1
1:A:218:MET:HE2	1:A:222:VAL:HG23	0.44	1.90	16	1
1:A:168:VAL:HG21	1:A:230:MET:CB	0.44	2.43	15	1
1:A:187:PHE:HA	1:A:227:ILE:CG2	0.43	2.43	17	3
1:A:140:VAL:HG13	1:A:166:ASN:ND2	0.43	2.28	9	1
1:A:218:MET:HA	1:A:218:MET:CE	0.43	2.43	14	2
1:A:187:PHE:CZ	1:A:224:THR:HG23	0.43	2.48	7	3
1:A:218:MET:O	1:A:222:VAL:HG12	0.43	2.13	20	1
1:A:191:VAL:HG12	1:A:196:ILE:HD12	0.43	1.90	19	1
1:A:218:MET:HE2	1:A:218:MET:O	0.43	2.14	16	1
1:A:173:TYR:CG	1:A:177:VAL:HG11	0.42	2.49	14	1
1:A:183:VAL:HG13	1:A:231:CYS:HB2	0.42	1.90	4	1
1:A:141:MET:SD	1:A:230:MET:HG2	0.41	2.55	9	1
1:A:146:TYR:CD1	1:A:148:PHE:CZ	0.41	3.09	6	1
1:A:183:VAL:HG21	1:A:232:VAL:HG22	0.41	1.92	11	1
1:A:148:PHE:CZ	1:A:156:TRP:CE3	0.41	3.09	10	1
1:A:223:VAL:HA	1:A:226:VAL:HG12	0.41	1.93	13	1
1:A:223:VAL:O	1:A:227:ILE:HD12	0.40	2.16	17	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	94/117 (80%)	81±2 (86±2%)	9±2 (9±2%)	4±1 (5±1%)	4 28
All	All	1880/2340 (80%)	1623 (86%)	172 (9%)	85 (5%)	4 28

All 16 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	175	SER	20
1	A	166	ASN	13
1	A	212	ASN	9
1	A	148	PHE	7
1	A	146	TYR	6
1	A	164	TYR	5
1	A	147	HIS	4
1	A	145	ASN	4
1	A	144	MET	4
1	A	163	ARG	4
1	A	135	TYR	2
1	A	172	ASP	2
1	A	149	ASP	2
1	A	176	PRO	1
1	A	177	VAL	1
1	A	165	PRO	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	87/99 (88%)	67±3 (77±4%)	20±3 (23±4%)	3 28

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1740/1980 (88%)	1337 (77%)	403 (23%)	3 28

All 60 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	177	VAL	20
1	A	220	ASN	20
1	A	155	ARG	19
1	A	179	GLN	17
1	A	214	THR	16
1	A	238	TYR	15
1	A	218	MET	14
1	A	226	VAL	14
1	A	157	TRP	14
1	A	167	ARG	13
1	A	156	TRP	12
1	A	221	LYS	11
1	A	160	ASN	11
1	A	187	PHE	10
1	A	212	ASN	10
1	A	228	ARG	10
1	A	139	ARG	10
1	A	240	LEU	9
1	A	223	VAL	9
1	A	170	TYR	9
1	A	219	GLU	9
1	A	154	TYR	8
1	A	180	ASP	8
1	A	213	GLN	7
1	A	163	ARG	6
1	A	142	SER	6
1	A	233	GLN	6
1	A	158	SER	6
1	A	192	THR	6
1	A	234	GLN	6
1	A	230	MET	5
1	A	237	GLU	4
1	A	153	GLU	4
1	A	145	ASN	4
1	A	215	GLU	4
1	A	152	ASP	4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	164	TYR	4
1	A	231	CYS	3
1	A	150	ARG	3
1	A	225	LYS	3
1	A	193	GLU	3
1	A	147	HIS	3
1	A	183	VAL	3
1	A	188	ASN	3
1	A	166	ASN	2
1	A	149	ASP	2
1	A	239	ARG	2
1	A	159	GLU	2
1	A	191	VAL	2
1	A	172	ASP	2
1	A	186	CYS	1
1	A	235	TYR	1
1	A	174	SER	1
1	A	229	GLU	1
1	A	140	VAL	1
1	A	144	MET	1
1	A	175	SER	1
1	A	189	ILE	1
1	A	196	ILE	1
1	A	185	ASP	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