Full wwPDB NMR Structure Validation Report (i)

Nov 11, 2024 – 07:28 AM EST

PDB ID	:	2L1E
Title	:	Mouse prion protein (121-231) containing the substitution F175A
Authors	:	Christen, B.; Damberger, F.F.; Perez, D.R.; Hornemann, S.; Wuthrich, K.
Deposited on	:	2010-07-28

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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

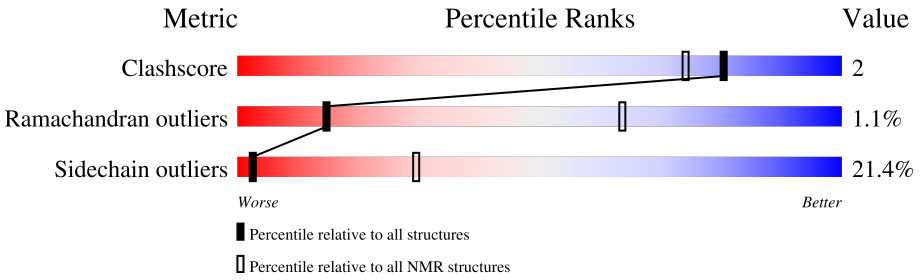
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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	210492	14027
Ramachandran outliers	207382	12486
Sidechain outliers	206894	12463

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	114	68% 18% • 12%

2 Ensemble composition and analysis

This entry contains 20 models. Model 9 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:128-A:227 (100)	0.58	9

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 3 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1799 atoms, of which 868 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Major prion protein.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	114	1799	574	868	165	183	9	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	119	GLY	-	expression tag	UNP Q4FJQ7
A	120	SER	-	expression tag	UNP Q4FJQ7
A	175	ALA	PHE	engineered mutation	UNP Q4FJQ7

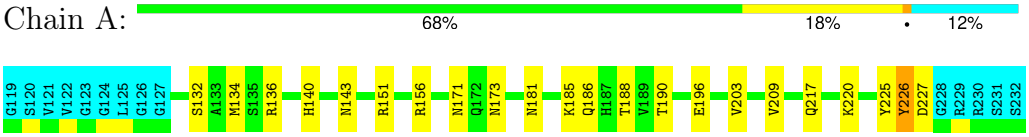
4 Residue-property plots (i)

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: Major prion protein

Chain A:



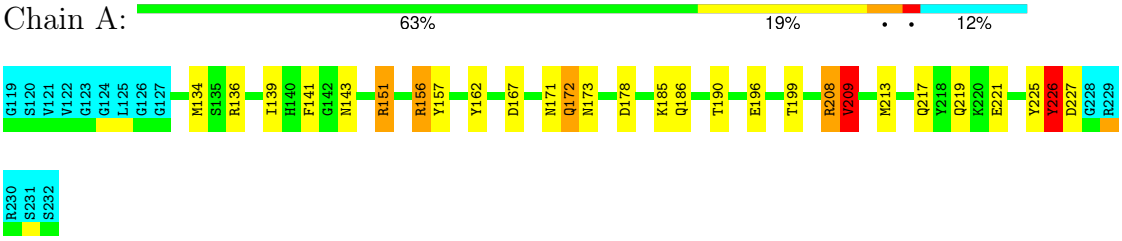
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Major prion protein

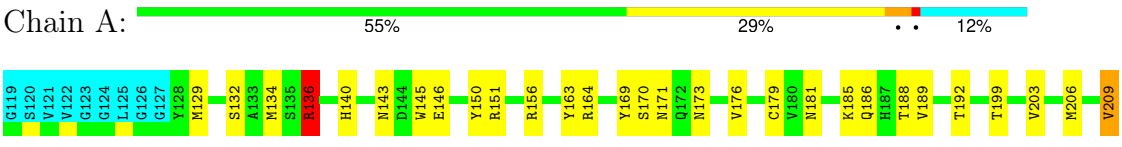
Chain A:



4.2.2 Score per residue for model 2

• Molecule 1: Major prion protein

Chain A:



4.2.3 Score per residue for model 3

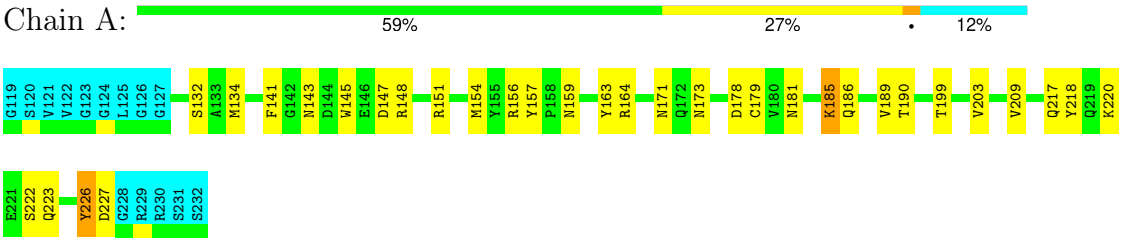
• Molecule 1: Major prion protein

Chain A:

4.2.4 Score per residue for model 4

• Molecule 1: Major prion protein

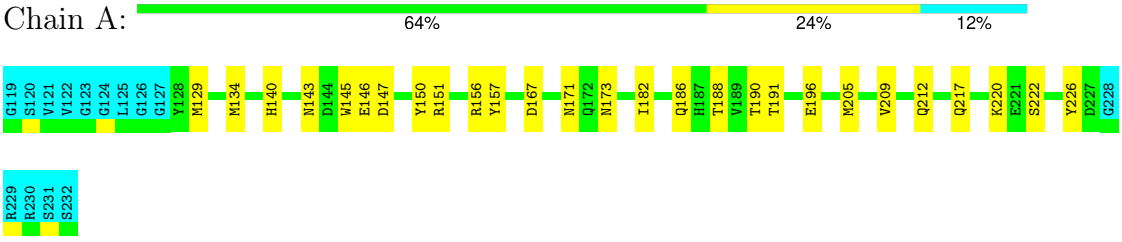
Chain A:



4.2.5 Score per residue for model 5

• Molecule 1: Major prion protein

Chain A:



4.2.6 Score per residue for model 6

• Molecule 1: Major prion protein

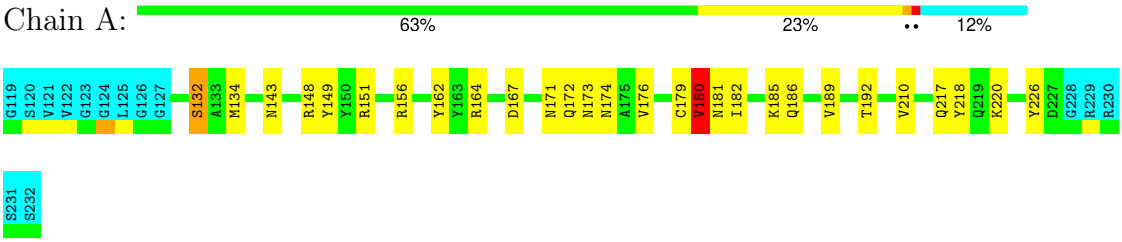
Chain A:

4.2.7 Score per residue for model 7

• Molecule 1: Major prion protein

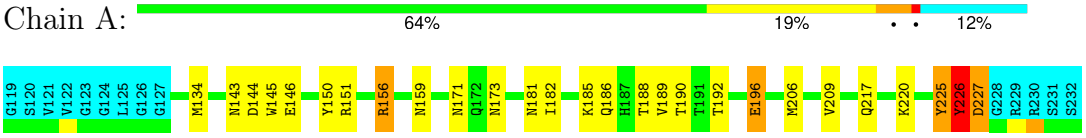
4.2.8 Score per residue for model 8

• Molecule 1: Major prion protein



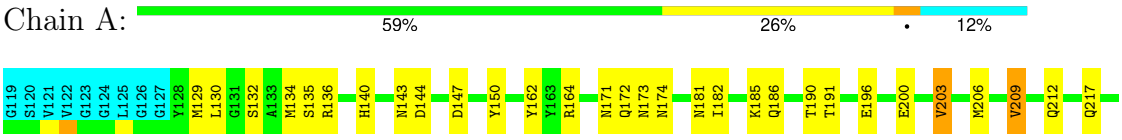
4.2.9 Score per residue for model 9 (medoid)

• Molecule 1: Major prion protein



4.2.10 Score per residue for model 10

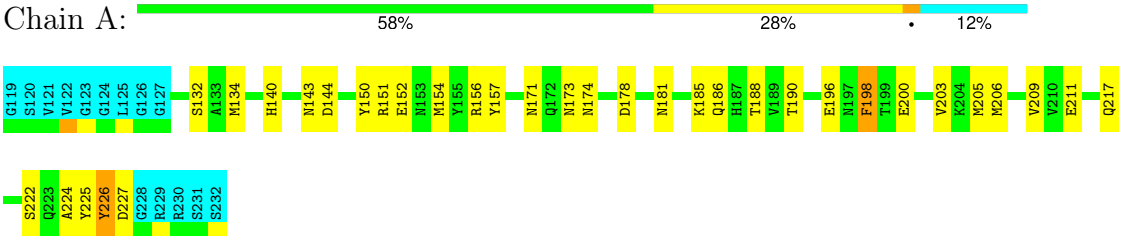
• Molecule 1: Major prion protein



4.2.11 Score per residue for model 11

• Molecule 1: Major prion protein

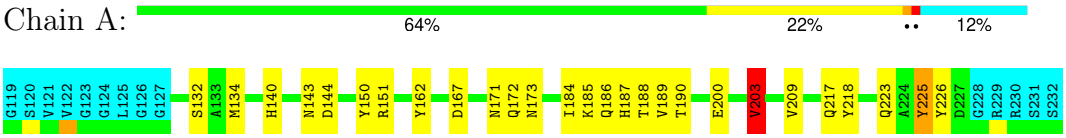
Chain A:



4.2.12 Score per residue for model 12

• Molecule 1: Major prion protein

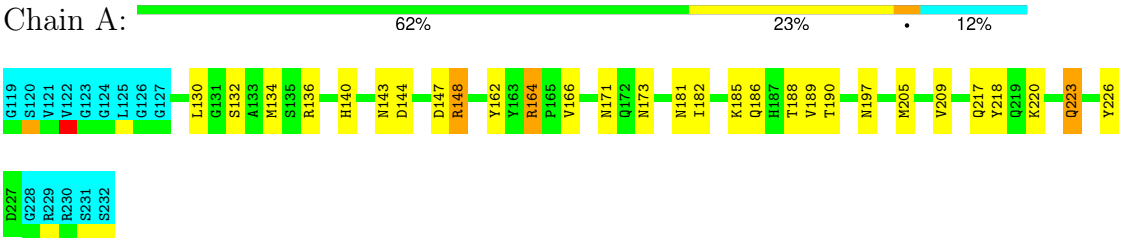
Chain A:



4.2.13 Score per residue for model 13

• Molecule 1: Major prion protein

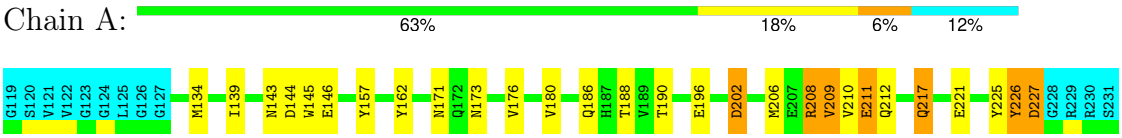
Chain A:



4.2.14 Score per residue for model 14

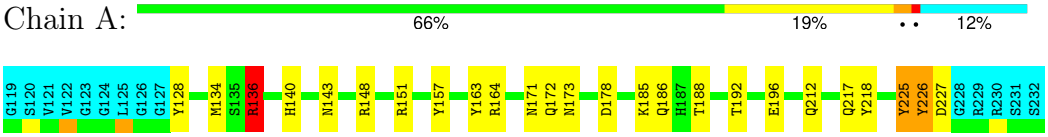
• Molecule 1: Major prion protein

Chain A:



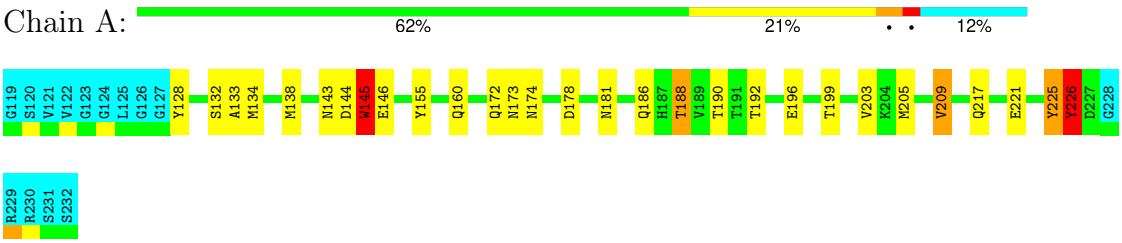
4.2.15 Score per residue for model 15

• Molecule 1: Major prion protein



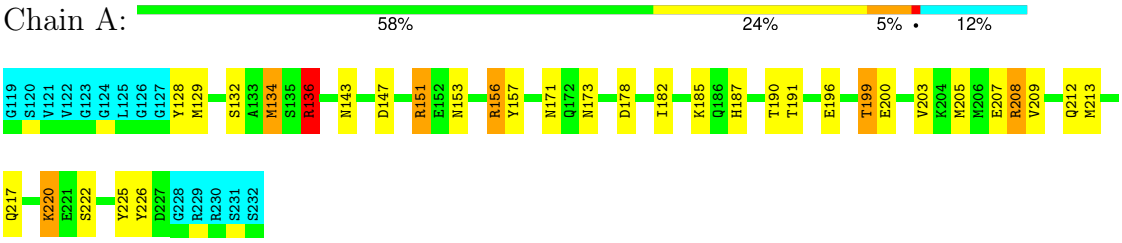
4.2.16 Score per residue for model 16

• Molecule 1: Major prion protein



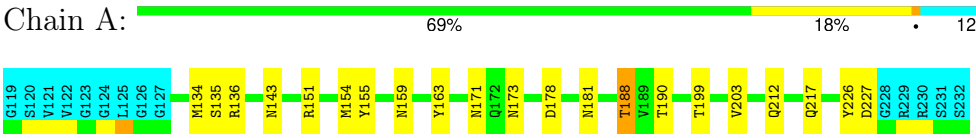
4.2.17 Score per residue for model 17

• Molecule 1: Major prion protein



4.2.18 Score per residue for model 18

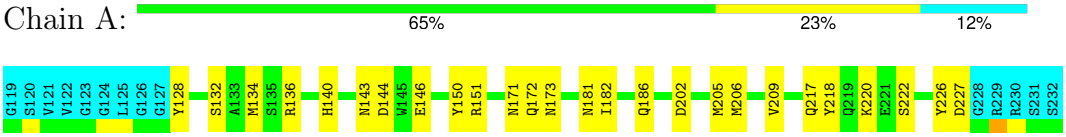
• Molecule 1: Major prion protein



4.2.19 Score per residue for model 19

• Molecule 1: Major prion protein

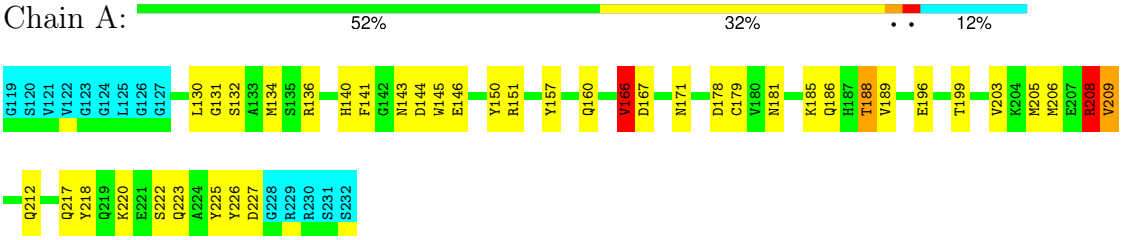
Chain A:



4.2.20 Score per residue for model 20

• Molecule 1: Major prion protein

Chain A:



5 Refinement protocol and experimental data overview

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

Of the 80 calculated structures, 20 were deposited, based on the following criterion: *target function*.

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

Software name	Classification	Version
DYANA	structure solution	1.0.3
OPALp	refinement	

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	0.69±0.01	0±0/866 (0.0± 0.0%)	1.16±0.03	2±1/1174 (0.2± 0.1%)
All	All	0.69	0/17320 (0.0%)	1.16	43/23480 (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	2.8±1.2
All	All	0	55

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	209	VAL	CA-CB-CG2	9.66	125.39	110.90	10	5
1	A	203	VAL	CA-CB-CG1	7.71	122.46	110.90	12	2
1	A	166	VAL	CG1-CB-CG2	7.20	122.41	110.90	20	1
1	A	156	ARG	NE-CZ-NH2	-7.10	116.75	120.30	1	2
1	A	156	ARG	NE-CZ-NH1	7.04	123.82	120.30	9	1
1	A	188	THR	CA-CB-CG2	6.76	121.86	112.40	18	3
1	A	151	ARG	NE-CZ-NH1	6.67	123.63	120.30	2	2
1	A	210	VAL	CA-CB-CG2	6.46	120.59	110.90	14	1
1	A	136	ARG	NE-CZ-NH2	-6.45	117.07	120.30	15	1
1	A	164	ARG	NE-CZ-NH1	6.39	123.50	120.30	15	3
1	A	148	ARG	NE-CZ-NH1	6.38	123.49	120.30	3	1
1	A	209	VAL	CA-CB-CG1	6.15	120.12	110.90	20	2
1	A	132	SER	N-CA-CB	-6.09	101.36	110.50	17	3
1	A	163	TYR	CB-CG-CD2	-6.09	117.34	121.00	7	1
1	A	180	VAL	CA-CB-CG1	5.89	119.73	110.90	8	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	162	TYR	CB-CG-CD2	-5.85	117.49	121.00	6	1
1	A	136	ARG	CD-NE-CZ	5.82	131.75	123.60	2	1
1	A	176	VAL	CA-CB-CG1	5.64	119.35	110.90	6	1
1	A	190	THR	CA-CB-CG2	5.56	120.18	112.40	13	1
1	A	218	TYR	CB-CG-CD1	-5.50	117.70	121.00	12	2
1	A	199	THR	CA-CB-CG2	-5.41	104.83	112.40	17	1
1	A	202	ASP	CB-CG-OD2	-5.33	113.51	118.30	14	1
1	A	157	TYR	CB-CG-CD2	-5.28	117.83	121.00	17	1
1	A	135	SER	C-N-CA	5.26	134.86	121.70	10	1
1	A	164	ARG	CD-NE-CZ	5.25	130.95	123.60	8	1
1	A	164	ARG	NE-CZ-NH2	-5.20	117.70	120.30	8	1
1	A	226	TYR	CB-CG-CD2	-5.12	117.93	121.00	16	1
1	A	136	ARG	NE-CZ-NH1	5.08	122.84	120.30	18	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	150	TYR	Sidechain	8
1	A	162	TYR	Sidechain	7
1	A	157	TYR	Sidechain	4
1	A	156	ARG	Sidechain	4
1	A	128	TYR	Sidechain	4
1	A	151	ARG	Sidechain	3
1	A	163	TYR	Sidechain	3
1	A	148	ARG	Sidechain	3
1	A	218	TYR	Sidechain	3
1	A	136	ARG	Sidechain	3
1	A	208	ARG	Sidechain	3
1	A	226	TYR	Sidechain	2
1	A	169	TYR	Sidechain	1
1	A	141	PHE	Sidechain	1
1	A	164	ARG	Sidechain	1
1	A	149	TYR	Sidechain	1
1	A	196	GLU	Peptide	1
1	A	198	PHE	Sidechain	1
1	A	225	TYR	Sidechain	1
1	A	155	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	845	780	780	4±2
All	All	16900	15600	15600	81

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:206:MET:HA	1:A:209:VAL:HG22	0.76	1.55	10	4
1:A:191:THR:HG22	1:A:196:GLU:HB3	0.63	1.69	17	5
1:A:185:LYS:O	1:A:189:VAL:HG23	0.61	1.94	13	3
1:A:199:THR:O	1:A:203:VAL:HG23	0.57	1.99	4	6
1:A:184:ILE:HG21	1:A:203:VAL:HG23	0.57	1.76	12	1
1:A:141:PHE:CE1	1:A:208:ARG:CZ	0.56	2.89	20	2
1:A:169:TYR:CD2	1:A:175:ALA:HB2	0.54	2.37	7	1
1:A:200:GLU:HA	1:A:203:VAL:HG12	0.54	1.79	11	4
1:A:200:GLU:HA	1:A:203:VAL:CG1	0.54	2.33	12	2
1:A:162:TYR:CE2	1:A:186:GLN:HG2	0.54	2.37	6	1
1:A:172:GLN:NE2	1:A:219:GLN:HE21	0.50	2.04	1	1
1:A:133:ALA:HB2	1:A:160:GLN:CD	0.50	2.26	16	1
1:A:191:THR:HG22	1:A:196:GLU:CB	0.49	2.37	17	2
1:A:145:TRP:CG	1:A:146:GLU:N	0.49	2.81	16	3
1:A:206:MET:HA	1:A:209:VAL:CG2	0.47	2.39	19	2
1:A:141:PHE:CE1	1:A:208:ARG:NH2	0.47	2.83	20	2
1:A:136:ARG:HH12	1:A:209:VAL:HG11	0.46	1.69	17	1
1:A:226:TYR:CD1	1:A:227:ASP:N	0.45	2.84	9	5
1:A:130:LEU:HD11	1:A:160:GLN:NE2	0.45	2.27	20	1
1:A:198:PHE:CD2	1:A:206:MET:CE	0.45	2.98	11	1
1:A:134:MET:SD	1:A:220:LYS:HE3	0.45	2.51	17	1
1:A:225:TYR:CE1	1:A:226:TYR:HB3	0.45	2.46	15	4
1:A:185:LYS:O	1:A:189:VAL:HG12	0.45	2.12	9	3
1:A:223:GLN:HE21	1:A:223:GLN:HA	0.45	1.72	13	1
1:A:180:VAL:HG22	1:A:211:GLU:HA	0.45	1.89	14	1
1:A:161:VAL:HG11	1:A:213:MET:HB2	0.44	1.88	6	1
1:A:166:VAL:CG1	1:A:218:TYR:CE1	0.44	3.01	20	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:171:ASN:HD21	1:A:174:ASN:ND2	0.43	2.12	10	1
1:A:176:VAL:HG22	1:A:218:TYR:CD2	0.43	2.49	8	2
1:A:217:GLN:HE22	1:A:221:GLU:HG3	0.43	1.73	14	1
1:A:186:GLN:HA	1:A:189:VAL:HG12	0.42	1.91	20	2
1:A:154:MET:HA	1:A:157:TYR:CD2	0.42	2.50	11	2
1:A:191:THR:HG21	1:A:198:PHE:CZ	0.42	2.50	3	1
1:A:150:TYR:CG	1:A:208:ARG:CZ	0.42	3.03	20	1
1:A:225:TYR:CE1	1:A:226:TYR:CD1	0.41	3.08	10	2
1:A:209:VAL:HB	1:A:213:MET:SD	0.41	2.55	1	1
1:A:145:TRP:CE2	1:A:146:GLU:HG2	0.41	2.50	9	2
1:A:130:LEU:C	1:A:130:LEU:HD13	0.41	2.36	10	1
1:A:180:VAL:HA	1:A:210:VAL:HG12	0.41	1.93	8	1
1:A:145:TRP:CZ2	1:A:146:GLU:CD	0.41	2.94	14	1
1:A:139:ILE:HG21	1:A:208:ARG:HD3	0.41	1.92	3	1
1:A:157:TYR:CE1	1:A:206:MET:SD	0.40	3.14	20	1
1:A:176:VAL:O	1:A:180:VAL:HG23	0.40	2.16	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	100/114 (88%)	88±2 (88±2%)	11±3 (11±3%)	1±1 (1±1%)	15	64
All	All	2000/2280 (88%)	1751 (88%)	228 (11%)	21 (1%)	15	64

All 14 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	227	ASP	4
1	A	136	ARG	2
1	A	132	SER	2
1	A	159	ASN	2
1	A	187	HIS	2
1	A	163	TYR	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	170	SER	1
1	A	171	ASN	1
1	A	224	ALA	1
1	A	144	ASP	1
1	A	166	VAL	1
1	A	185	LYS	1
1	A	145	TRP	1
1	A	131	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	94/102 (92%)	74±3 (79±3%)	20±3 (21±3%)	2 29
All	All	1880/2040 (92%)	1478 (79%)	402 (21%)	2 29

All 61 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	134	MET	20
1	A	143	ASN	20
1	A	217	GLN	20
1	A	226	TYR	20
1	A	173	ASN	19
1	A	171	ASN	18
1	A	190	THR	14
1	A	181	ASN	13
1	A	188	THR	13
1	A	220	LYS	12
1	A	151	ARG	11
1	A	186	GLN	11
1	A	140	HIS	11
1	A	225	TYR	11
1	A	178	ASP	9
1	A	185	LYS	9
1	A	182	ILE	9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	144	ASP	9
1	A	136	ARG	8
1	A	196	GLU	8
1	A	209	VAL	8
1	A	132	SER	8
1	A	192	THR	8
1	A	222	SER	8
1	A	205	MET	8
1	A	172	GLN	7
1	A	212	GLN	7
1	A	156	ARG	6
1	A	167	ASP	6
1	A	227	ASP	6
1	A	179	CYS	5
1	A	147	ASP	5
1	A	223	GLN	5
1	A	208	ARG	4
1	A	129	MET	4
1	A	221	GLU	3
1	A	164	ARG	3
1	A	145	TRP	3
1	A	148	ARG	3
1	A	174	ASN	3
1	A	139	ILE	2
1	A	199	THR	2
1	A	170	SER	2
1	A	207	GLU	2
1	A	211	GLU	2
1	A	202	ASP	2
1	A	214	CYS	1
1	A	159	ASN	1
1	A	180	VAL	1
1	A	152	GLU	1
1	A	203	VAL	1
1	A	130	LEU	1
1	A	197	ASN	1
1	A	138	MET	1
1	A	155	TYR	1
1	A	153	ASN	1
1	A	213	MET	1
1	A	135	SER	1
1	A	154	MET	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	146	GLU	1
1	A	166	VAL	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 oligosaccharides 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

No chemical shift data were provided