Full wwPDB NMR Structure Validation Report (i)

Feb 15, 2022 – 09:29 AM EST

PDB ID	:	1LG4
Title	:	NMR structure of the human doppel protein fragment 24-152
Authors	:	Luhrs, T.; Riek, R.; Guntert, P.; Wuthrich, K.
Deposited on	:	2002-04-15

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:

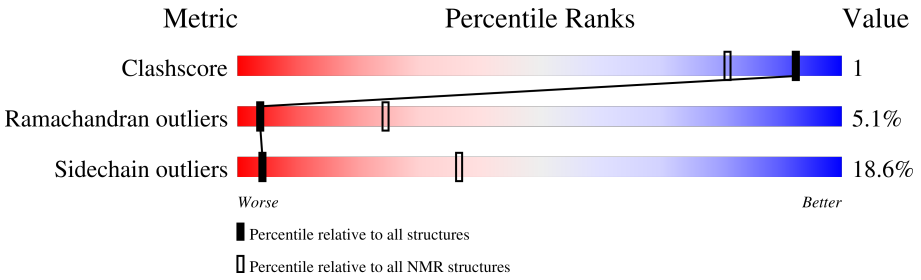
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.26
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
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	129	64% 10% • • 23%

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: *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	A:52-A:148 (97)	0.62	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 3 clusters and 2 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1551 atoms, of which 747 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	99	1551	508	747	142	150	4	0

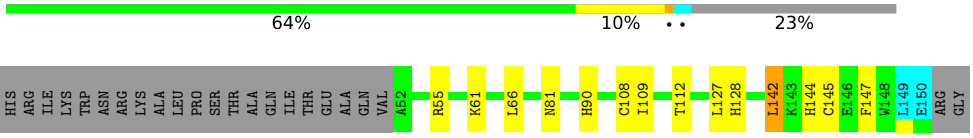
4 Residue-property plots (i)

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:



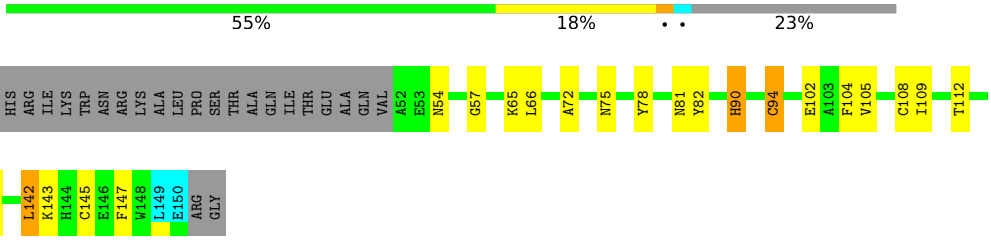
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Prion-like protein

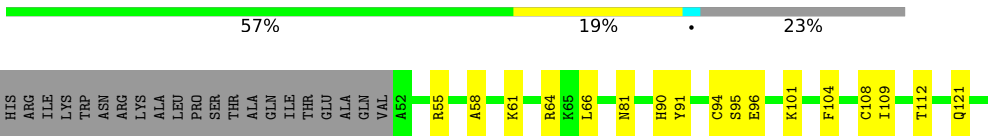
Chain A:



4.2.2 Score per residue for model 2

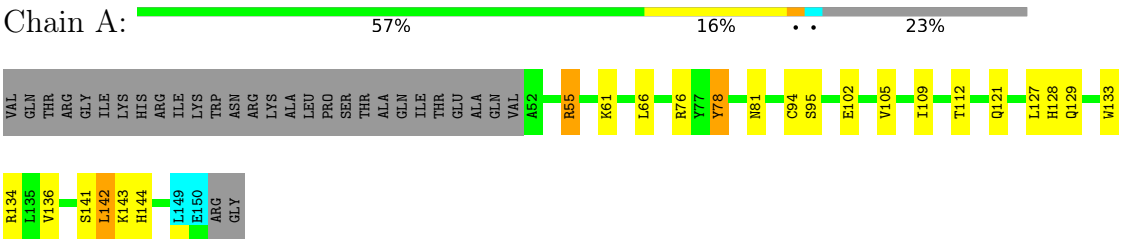
• Molecule 1: Prion-like protein

Chain A:



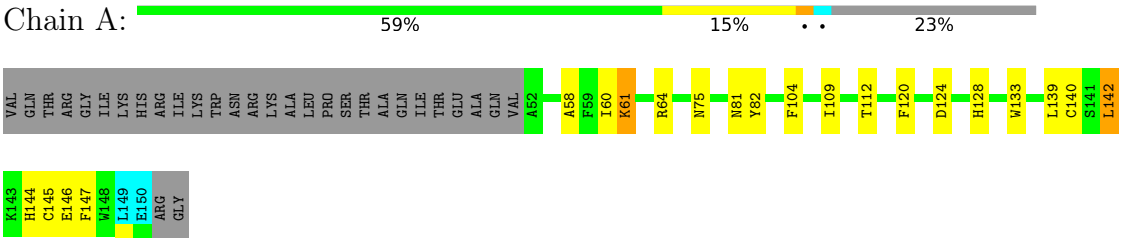
4.2.3 Score per residue for model 3

- Molecule 1: Prion-like protein



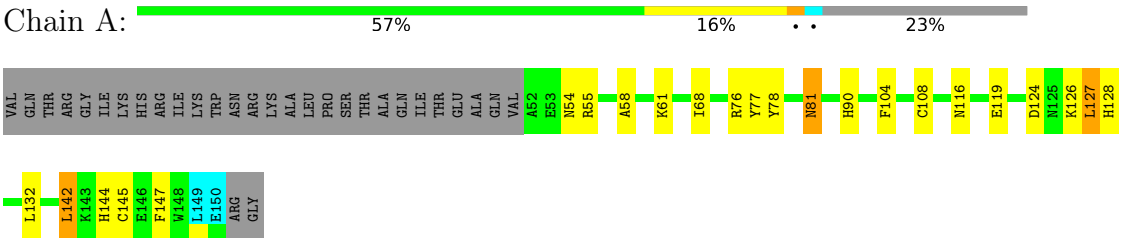
4.2.4 Score per residue for model 4

- Molecule 1: Prion-like protein



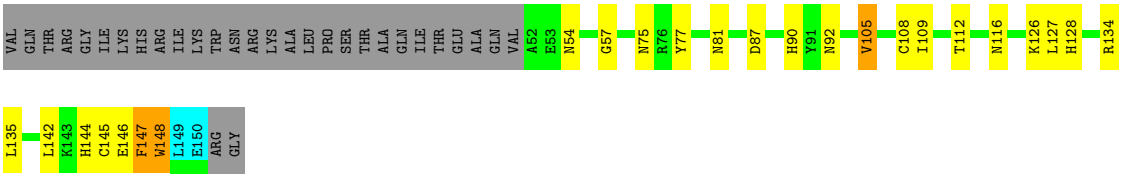
4.2.5 Score per residue for model 5

- Molecule 1: Prion-like protein



4.2.6 Score per residue for model 6

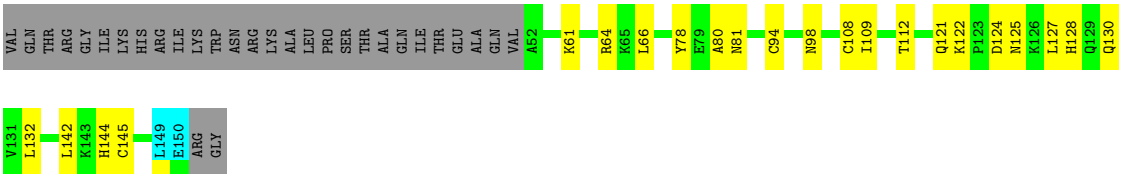
- Molecule 1: Prion-like protein



4.2.7 Score per residue for model 7

- Molecule 1: Prion-like protein

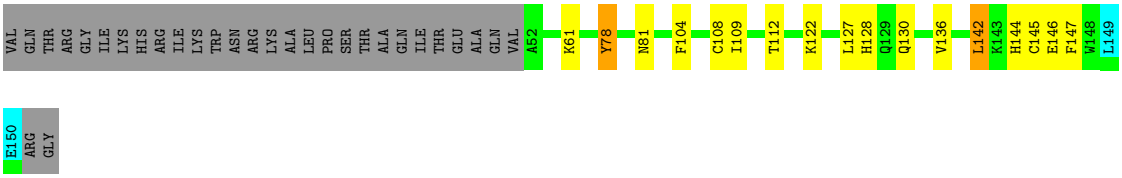
Chain A: 58% 17% • 23%



4.2.8 Score per residue for model 8 (medoid)

- Molecule 1: Prion-like protein

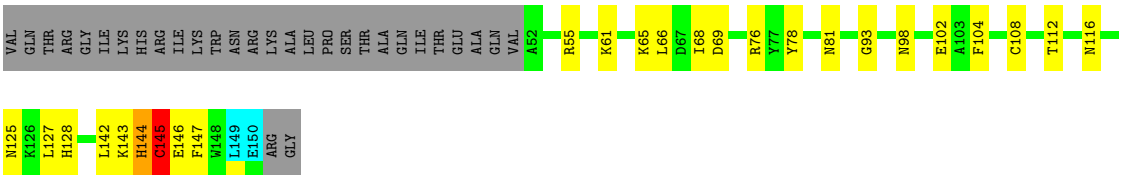
Chain A: 62% 12% • • 23%



4.2.9 Score per residue for model 9

- Molecule 1: Prion-like protein

Chain A: 56% 18% • • • 23%



4.2.10 Score per residue for model 10

- Molecule 1: Prion-like protein

Chain A: 60% 13% • • 23%

VAL GLN THR ARG GLY ILE LYS HIS ARG ILE ILE LYS TRP ASN ARG LYS ALA LEU PRO SER THR ALA GLN ILE THR GLU ALA GLN VAL A52 A58 K61 L66 Y78 N81 H90 Y91 N92 N98 I109 T112 Q121 D124 N125 K126 L127 H128 V136

L142 K143 E146 L149 E150 ARG GLY

4.2.11 Score per residue for model 11

- Molecule 1: Prion-like protein

Chain A: 56% 19% • • 23%

VAL GLN THR ARG GLY ILE LYS HIS ARG ILE ILE LYS TRP ASN ARG LYS ALA LEU PRO SER THR ALA GLN ILE THR GLU ALA GLN VAL A52 K61 L66 D67 I68 Y77 N81 N92 G93 C94 S95 E96 A97 N98 F104 C108 I109 T112 Q121 D124 N125 K126

L127 H128 W133 R134 L135 V136 L142 K143 F147 W148 L149 E150 ARG GLY

4.2.12 Score per residue for model 12

- Molecule 1: Prion-like protein

Chain A: 58% 16% • • 23%

VAL GLN THR ARG GLY ILE LYS HIS ARG ILE ILE LYS TRP ASN ARG LYS ALA LEU PRO SER THR ALA GLN ILE THR GLU ALA GLN VAL A52 E53 N54 A58 F59 I60 K61 R64 D69 Y78 N81 S95 N98 E102 I109 T112 N116 E119 L127

W133 L139 L142 K143 H144 C145 E146 F147 W148 L149 E150 ARG GLY

4.2.13 Score per residue for model 13

- Molecule 1: Prion-like protein

Chain A: 54% 19% • • 23%

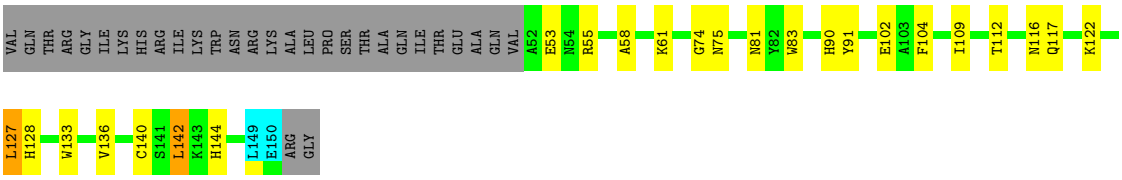
VAL GLN THR ARG GLY ILE LYS HIS ARG ILE ILE LYS TRP ASN ARG LYS ALA LEU PRO SER THR ALA GLN ILE THR GLU ALA GLN VAL A52 F59 I60 K61 R64 K65 L66 F70 R76 Y77 N81 H90 S95 N98 V105 C108 I109 T112 D124 N125

K126 L127 H128 Q129 Q130 V131 L135 L142 E146 F147 W148 L149 E150 ARG GLY

4.2.14 Score per residue for model 14

- Molecule 1: Prion-like protein

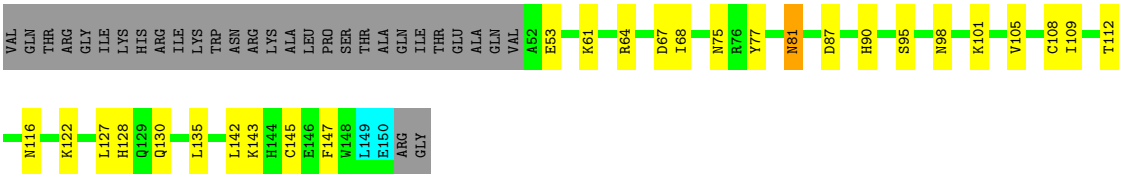
Chain A: 57% 17% • • 23%



4.2.15 Score per residue for model 15

- Molecule 1: Prion-like protein

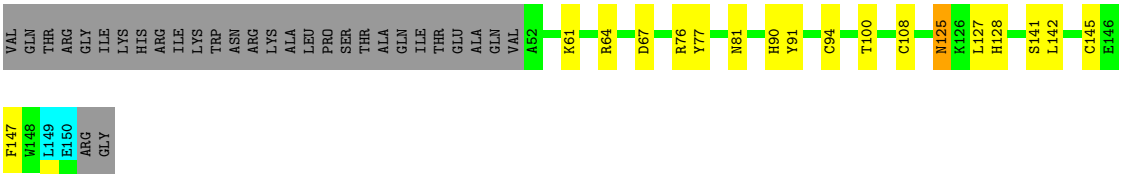
Chain A: 54% 20% • • 23%



4.2.16 Score per residue for model 16

- Molecule 1: Prion-like protein

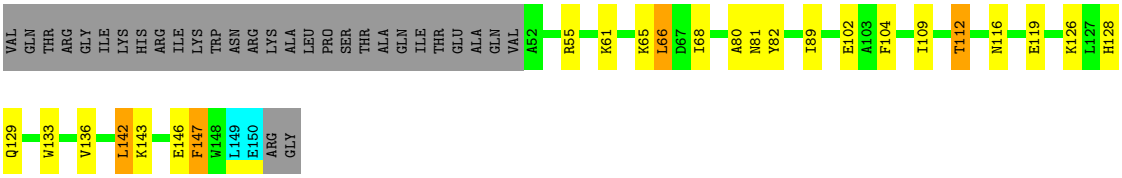
Chain A: 61% 13% • • 23%



4.2.17 Score per residue for model 17

- Molecule 1: Prion-like protein

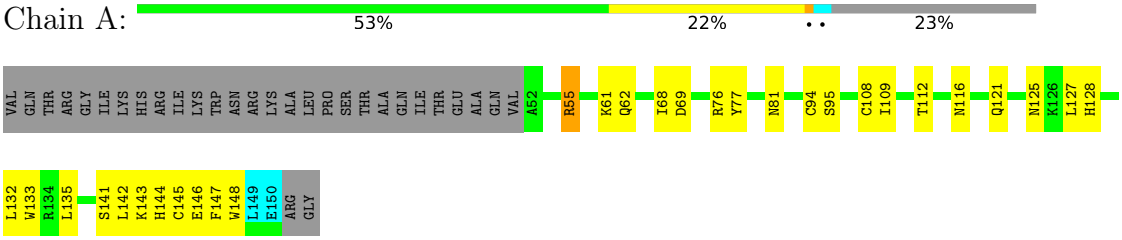
Chain A: 57% 16% • • 23%



4.2.18 Score per residue for model 18

- Molecule 1: Prion-like protein

Chain A:



4.2.19 Score per residue for model 19

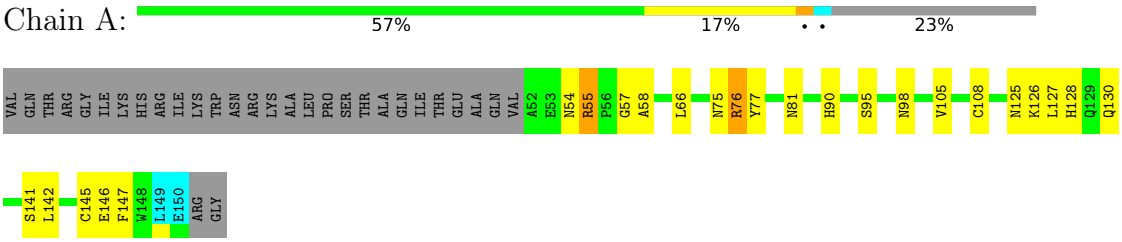
- Molecule 1: Prion-like protein

Chain A:

4.2.20 Score per residue for model 20

- Molecule 1: Prion-like protein

Chain A:



5 Refinement protocol and experimental data overview (i)

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	1.65
OPALp	refinement	1.3

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	0.64±0.01	0±0/809 (0.0± 0.0%)	1.10±0.03	1±1/1095 (0.1± 0.1%)
All	All	0.64	0/16180 (0.0%)	1.11	15/21900 (0.1%)

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	1.2±0.8
All	All	0	25

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	76	ARG	NE-CZ-NH2	-6.75	116.92	120.30	20	1
1	A	78	TYR	CB-CG-CD2	-6.71	116.97	121.00	8	2
1	A	91	TYR	CB-CG-CD2	-6.52	117.09	121.00	2	2
1	A	134	ARG	NE-CZ-NH1	6.06	123.33	120.30	1	1
1	A	105	VAL	CA-CB-CG1	6.00	119.90	110.90	6	1
1	A	136	VAL	CA-CB-CG1	5.79	119.58	110.90	8	3
1	A	55	ARG	CD-NE-CZ	5.67	131.54	123.60	9	1
1	A	77	TYR	CB-CG-CD1	-5.38	117.77	121.00	20	1
1	A	69	ASP	CB-CG-OD1	5.29	123.06	118.30	18	1
1	A	135	LEU	CB-CG-CD2	5.22	119.88	111.00	13	1
1	A	90	HIS	CA-CB-CG	5.19	122.43	113.60	1	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	76	ARG	Sidechain	5
1	A	78	TYR	Sidechain	5
1	A	55	ARG	Sidechain	4
1	A	134	ARG	Sidechain	2
1	A	82	TYR	Sidechain	2
1	A	121	GLN	Peptide	1
1	A	92	ASN	Peptide	1
1	A	148	TRP	Peptide	1
1	A	66	LEU	Peptide	1
1	A	124	ASP	Peptide	1
1	A	64	ARG	Sidechain	1
1	A	87	ASP	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	787	730	729	2±1
All	All	15740	14600	14580	42

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	Total
1:A:55:ARG:HB2	1:A:58:ALA:HB2	0.64	1.68	20	4
1:A:109:ILE:HA	1:A:112:THR:HG22	0.57	1.76	7	14
1:A:65:LYS:HE3	1:A:82:TYR:CE2	0.51	2.40	17	1
1:A:126:LYS:HD2	1:A:127:LEU:HD22	0.51	1.81	10	1
1:A:144:HIS:HA	1:A:148:TRP:CH2	0.50	2.42	18	1
1:A:94:CYS:HB2	1:A:104:PHE:CE1	0.47	2.44	1	1
1:A:77:TYR:CD2	1:A:127:LEU:HB3	0.47	2.45	5	1
1:A:112:THR:HG21	1:A:136:VAL:HG11	0.46	1.87	17	3
1:A:68:ILE:CD1	1:A:135:LEU:HD22	0.46	2.39	18	1
1:A:58:ALA:HB3	1:A:90:HIS:CD2	0.45	2.46	10	1
1:A:105:VAL:HG12	1:A:109:ILE:HD12	0.44	1.88	3	1
1:A:68:ILE:HD13	1:A:135:LEU:HD22	0.44	1.90	11	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:66:LEU:HD22	1:A:66:LEU:N	0.44	2.28	19	1
1:A:144:HIS:HA	1:A:148:TRP:CZ2	0.43	2.48	6	1
1:A:65:LYS:HE2	1:A:82:TYR:CD2	0.43	2.48	19	1
1:A:65:LYS:HE2	1:A:82:TYR:CE2	0.43	2.49	19	1
1:A:65:LYS:HG3	1:A:82:TYR:CG	0.43	2.49	1	1
1:A:70:PHE:CZ	1:A:131:VAL:HG13	0.42	2.49	13	1
1:A:74:GLY:HA2	1:A:127:LEU:HD12	0.42	1.92	14	1
1:A:144:HIS:CG	1:A:145:CYS:H	0.41	2.34	9	1
1:A:89:ILE:HG13	1:A:136:VAL:HG13	0.41	1.93	17	1
1:A:60:ILE:HG13	1:A:90:HIS:CE1	0.41	2.51	13	1
1:A:68:ILE:CD1	1:A:135:LEU:CD2	0.40	2.99	18	1
1:A:68:ILE:HD12	1:A:135:LEU:HD12	0.40	1.94	15	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	96/129 (74%)	71±3 (74±3%)	20±3 (21±3%)	5±1 (5±1%)	4 24
All	All	1920/2580 (74%)	1425 (74%)	397 (21%)	98 (5%)	4 24

All 30 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	142	LEU	13
1	A	145	CYS	8
1	A	90	HIS	7
1	A	143	LYS	6
1	A	125	ASN	5
1	A	144	HIS	5
1	A	66	LEU	4
1	A	146	GLU	4
1	A	147	PHE	4
1	A	126	LYS	4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	57	GLY	3
1	A	94	CYS	3
1	A	141	SER	3
1	A	81	ASN	3
1	A	124	ASP	3
1	A	122	LYS	3
1	A	120	PHE	2
1	A	95	SER	2
1	A	58	ALA	2
1	A	61	LYS	2
1	A	80	ALA	2
1	A	91	TYR	2
1	A	72	ALA	1
1	A	121	GLN	1
1	A	93	GLY	1
1	A	83	TRP	1
1	A	101	LYS	1
1	A	62	GLN	1
1	A	56	PRO	1
1	A	96	GLU	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	80/107 (75%)	65±3 (81±4%)	15±3 (19±4%)	4 37
All	All	1600/2140 (75%)	1303 (81%)	297 (19%)	4 37

All 56 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	81	ASN	20
1	A	128	HIS	19
1	A	127	LEU	17
1	A	142	LEU	17
1	A	61	LYS	16

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	108	CYS	14
1	A	147	PHE	13
1	A	104	PHE	9
1	A	98	ASN	9
1	A	145	CYS	8
1	A	116	ASN	8
1	A	66	LEU	7
1	A	133	TRP	7
1	A	77	TYR	7
1	A	54	ASN	6
1	A	75	ASN	6
1	A	102	GLU	6
1	A	64	ARG	6
1	A	95	SER	6
1	A	105	VAL	5
1	A	94	CYS	5
1	A	126	LYS	5
1	A	125	ASN	5
1	A	130	GLN	5
1	A	146	GLU	5
1	A	121	GLN	4
1	A	78	TYR	3
1	A	90	HIS	3
1	A	129	GLN	3
1	A	55	ARG	3
1	A	144	HIS	3
1	A	68	ILE	3
1	A	119	GLU	3
1	A	132	LEU	3
1	A	112	THR	3
1	A	67	ASP	3
1	A	96	GLU	2
1	A	76	ARG	2
1	A	124	ASP	2
1	A	139	LEU	2
1	A	140	CYS	2
1	A	135	LEU	2
1	A	65	LYS	2
1	A	69	ASP	2
1	A	143	LYS	2
1	A	92	ASN	2
1	A	59	PHE	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	53	GLU	2
1	A	101	LYS	1
1	A	60	ILE	1
1	A	87	ASP	1
1	A	117	GLN	1
1	A	122	LYS	1
1	A	100	THR	1
1	A	109	ILE	1
1	A	141	SER	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