Full wwPDB NMR Structure Validation Report (i)

Feb 16, 2018 – 03:55 am GMT

PDB ID	:	2K21
Title	:	NMR structure of human KCNE1 in LMPG micelles at pH 6.0 and 40 degree C
Authors	:	Kang, C.; Tian, C.; Sonnichsen, F.D.; Smith, J.A.; Meiler, J.; George, A.L.; Vanoye, C.G.; Sanders, C.R.; Kim, H.
Deposited on	:	2008-03-19

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:

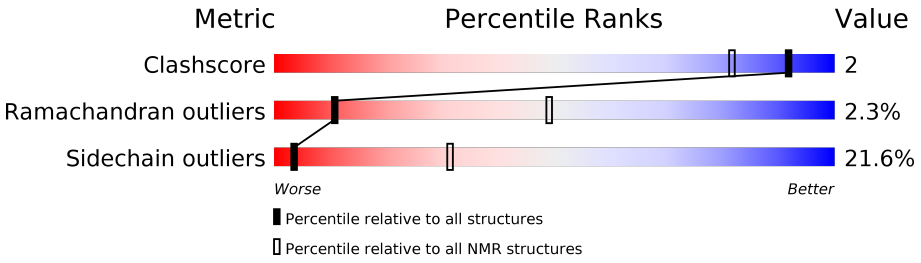
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	trunk30686
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk30686

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	136279	12091
Ramachandran outliers	132675	10835
Sidechain outliers	132484	10811

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	138	27% 41% 9% 16% 7%

2 Ensemble composition and analysis

This entry contains 10 models. Model 5 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:1-A:107 (107)	1.45	5

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 2 clusters and 2 single-model clusters were found.

Cluster number	Models
1	1, 4, 5, 7, 8
2	3, 9, 10
Single-model clusters	2; 6

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2055 atoms, of which 1025 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Potassium voltage-gated channel subfamily E member.

Mol	Chain	Residues	Atoms						Trace
1	A	129	Total	C	H	N	O	S	0
			2055	660	1025	172	193	5	

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-8	MET	-	EXPRESSION TAG	UNP Q6FHJ6
A	-7	GLY	-	EXPRESSION TAG	UNP Q6FHJ6
A	-6	HIS	-	EXPRESSION TAG	UNP Q6FHJ6
A	-5	HIS	-	EXPRESSION TAG	UNP Q6FHJ6
A	-4	HIS	-	EXPRESSION TAG	UNP Q6FHJ6
A	-3	HIS	-	EXPRESSION TAG	UNP Q6FHJ6
A	-2	HIS	-	EXPRESSION TAG	UNP Q6FHJ6
A	-1	HIS	-	EXPRESSION TAG	UNP Q6FHJ6
A	0	GLY	-	EXPRESSION TAG	UNP Q6FHJ6
A	104	LYS	ARG	ENGINEERED	UNP Q6FHJ6

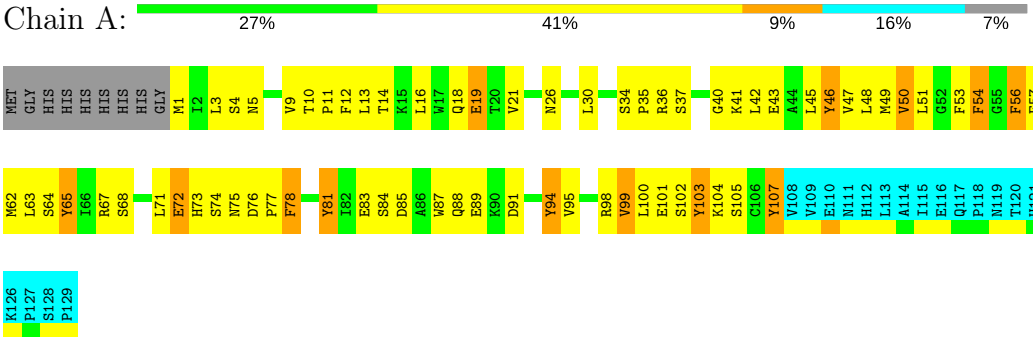
4 Residue-property plots

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA 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: Potassium voltage-gated channel subfamily E member

Chain A:



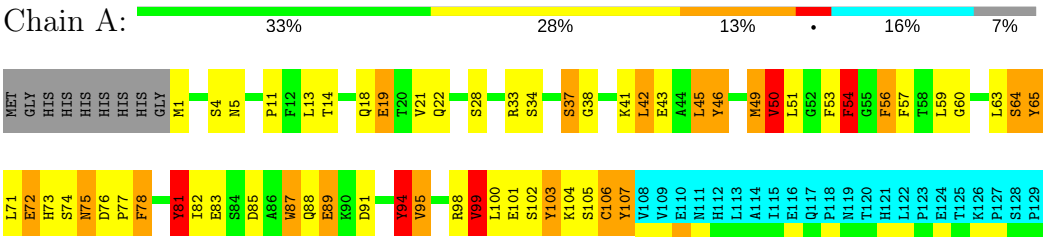
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

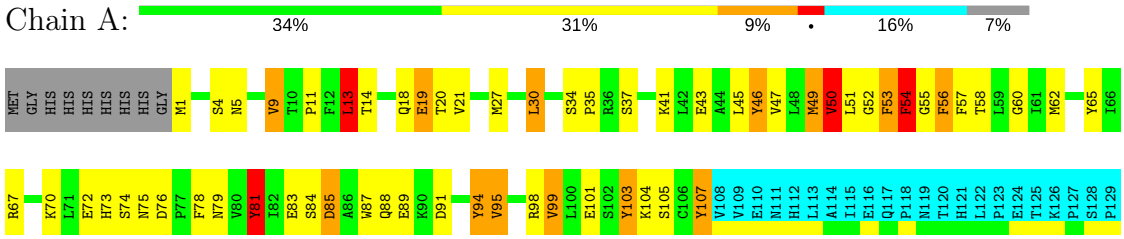
• Molecule 1: Potassium voltage-gated channel subfamily E member

Chain A:



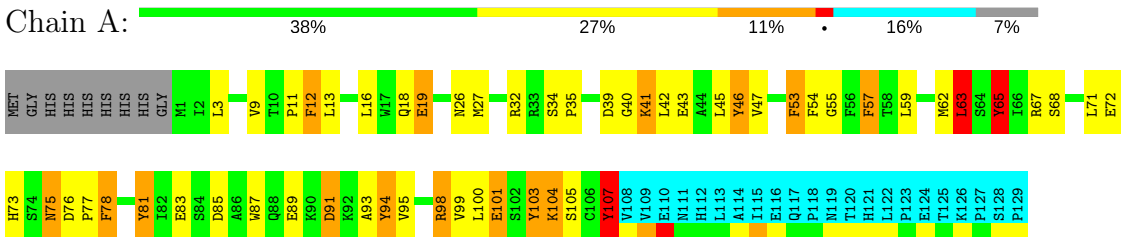
4.2.2 Score per residue for model 2

• Molecule 1: Potassium voltage-gated channel subfamily E member



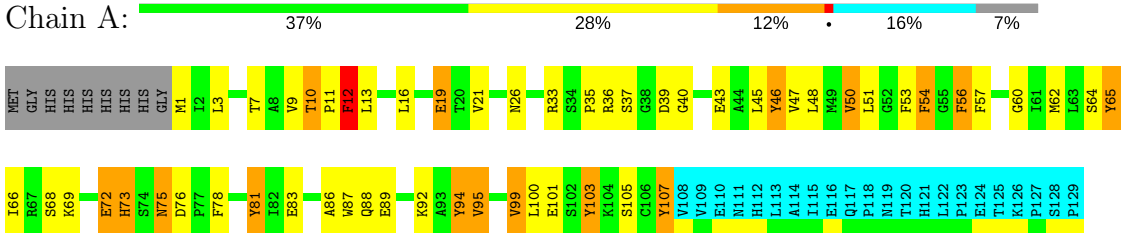
4.2.3 Score per residue for model 3

• Molecule 1: Potassium voltage-gated channel subfamily E member



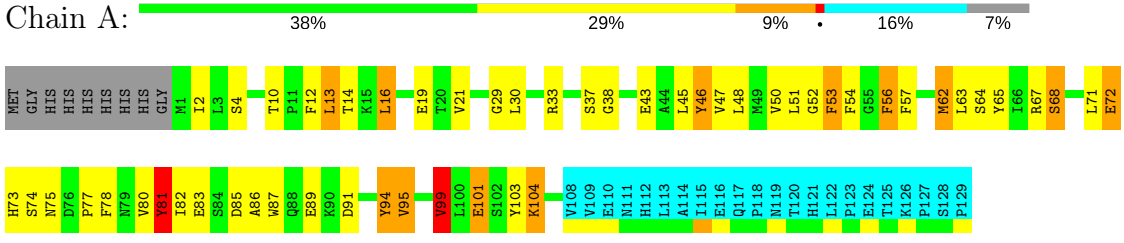
4.2.4 Score per residue for model 4

• Molecule 1: Potassium voltage-gated channel subfamily E member



4.2.5 Score per residue for model 5 (medoid)

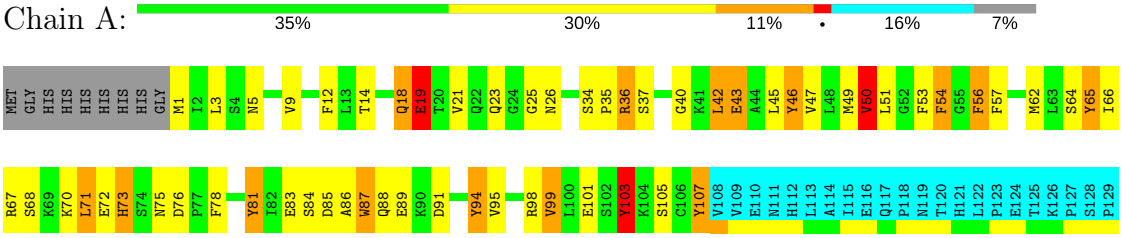
• Molecule 1: Potassium voltage-gated channel subfamily E member



4.2.6 Score per residue for model 6

• Molecule 1: Potassium voltage-gated channel subfamily E member

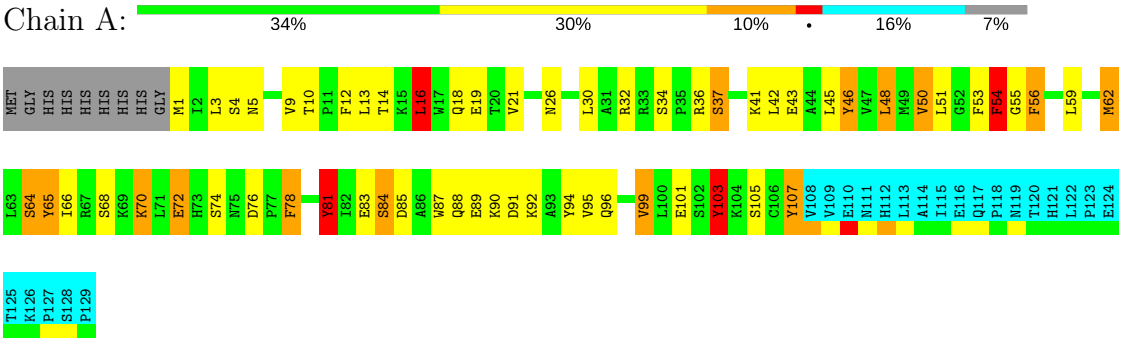
Chain A:



4.2.7 Score per residue for model 7

• Molecule 1: Potassium voltage-gated channel subfamily E member

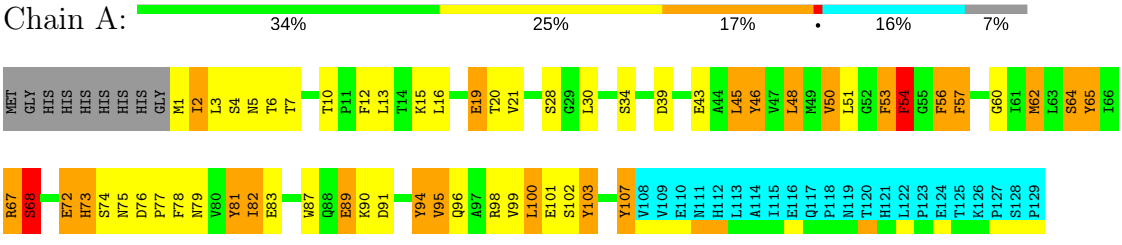
Chain A:



4.2.8 Score per residue for model 8

• Molecule 1: Potassium voltage-gated channel subfamily E member

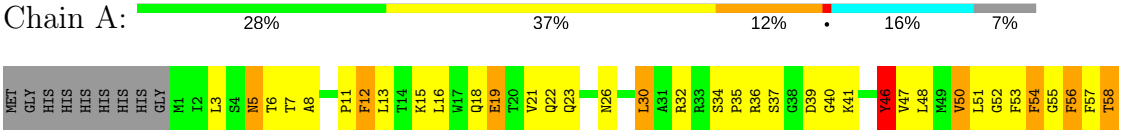
Chain A:



4.2.9 Score per residue for model 9

• Molecule 1: Potassium voltage-gated channel subfamily E member

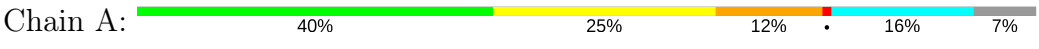
Chain A:

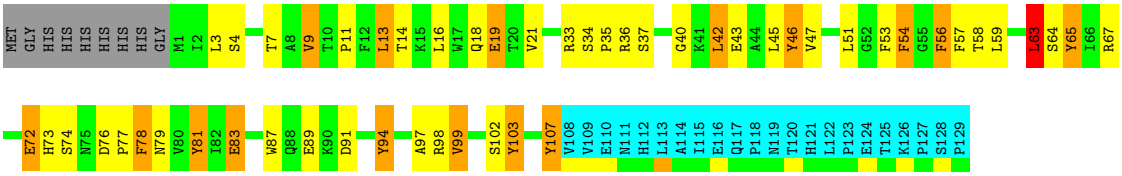


4.2.10 Score per residue for model 10

• Molecule 1: Potassium voltage-gated channel subfamily E member

Chain A:





5 Refinement protocol and experimental data overview

The models were refined using the following method: *DGSA-distance geometry simulated annealing.*

Of the 500 calculated structures, 10 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
X-PLOR NIH	refinement	2.17

No chemical shift data was provided. No validations of the models with respect to experimental NMR restraints is performed at this time.

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	2.64±0.01	27±2/877 (3.0±0.3%)	0.92±0.03	1±1/1184 (0.1±0.1%)
All	All	2.64	266/8770 (3.0%)	0.92	13/11840 (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	26.9±3.4
All	All	0	269

All unique bond 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(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	107	TYR	CB-CG	7.23	1.62	1.51	10	9
1	A	94	TYR	CB-CG	7.14	1.62	1.51	10	10
1	A	46	TYR	CB-CG	7.01	1.62	1.51	6	9
1	A	65	TYR	CB-CG	6.95	1.62	1.51	6	10
1	A	103	TYR	CB-CG	6.79	1.61	1.51	9	10
1	A	57	PHE	CB-CG	6.58	1.62	1.51	5	8
1	A	81	TYR	CB-CG	6.52	1.61	1.51	10	10
1	A	78	PHE	CB-CG	6.52	1.62	1.51	1	9
1	A	89	GLU	CB-CG	6.41	1.64	1.52	6	8
1	A	101	GLU	CB-CG	6.25	1.64	1.52	4	9
1	A	72	GLU	CB-CG	6.22	1.64	1.52	1	10
1	A	54	PHE	CB-CG	6.20	1.61	1.51	10	10
1	A	19	GLU	CB-CG	6.14	1.63	1.52	2	10
1	A	83	GLU	CB-CG	6.08	1.63	1.52	4	10
1	A	53	PHE	CB-CG	6.02	1.61	1.51	4	9
1	A	77	PRO	N-CA	5.94	1.57	1.47	1	4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	43	GLU	CB-CG	5.85	1.63	1.52	2	9
1	A	12	PHE	CB-CG	5.76	1.61	1.51	3	6
1	A	56	PHE	CB-CG	5.64	1.60	1.51	10	9
1	A	13	LEU	CB-CG	5.62	1.68	1.52	3	9
1	A	21	VAL	CA-CB	5.51	1.66	1.54	8	7
1	A	99	VAL	CA-CB	5.45	1.66	1.54	1	4
1	A	9	VAL	CA-CB	5.43	1.66	1.54	10	3
1	A	35	PRO	N-CA	5.41	1.56	1.47	3	6
1	A	59	LEU	CB-CG	5.39	1.68	1.52	9	1
1	A	11	PRO	N-CA	5.38	1.56	1.47	9	5
1	A	2	ILE	CA-CB	5.37	1.67	1.54	5	2
1	A	50	VAL	CA-CB	5.36	1.66	1.54	4	5
1	A	21	VAL	CB-CG1	5.32	1.64	1.52	8	2
1	A	51	LEU	CB-CG	5.32	1.68	1.52	9	9
1	A	63	LEU	CB-CG	5.31	1.68	1.52	3	4
1	A	16	LEU	CB-CG	5.27	1.67	1.52	4	3
1	A	83	GLU	CG-CD	5.24	1.59	1.51	4	1
1	A	37	SER	CA-CB	5.22	1.60	1.52	1	1
1	A	30	LEU	CB-CG	5.20	1.67	1.52	9	3
1	A	21	VAL	CB-CG2	5.18	1.63	1.52	2	5
1	A	66	ILE	CA-CB	5.18	1.66	1.54	9	1
1	A	50	VAL	CB-CG1	5.16	1.63	1.52	5	2
1	A	3	LEU	CB-CG	5.16	1.67	1.52	9	2
1	A	42	LEU	CB-CG	5.13	1.67	1.52	6	4
1	A	48	LEU	CB-CG	5.12	1.67	1.52	7	2
1	A	71	LEU	CB-CG	5.12	1.67	1.52	1	2
1	A	9	VAL	CB-CG1	5.11	1.63	1.52	7	3
1	A	47	VAL	CB-CG1	5.09	1.63	1.52	6	1
1	A	100	LEU	CB-CG	5.08	1.67	1.52	1	1
1	A	99	VAL	CB-CG1	5.05	1.63	1.52	9	1
1	A	47	VAL	CA-CB	5.05	1.65	1.54	4	3
1	A	50	VAL	CB-CG2	5.04	1.63	1.52	4	2
1	A	47	VAL	CB-CG2	5.03	1.63	1.52	5	1
1	A	80	VAL	CB-CG1	5.01	1.63	1.52	5	1
1	A	45	LEU	CB-CG	5.01	1.67	1.52	8	1

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	36	ARG	NE-CZ-NH2	-7.93	116.33	120.30	7	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	33	ARG	NE-CZ-NH2	-6.83	116.89	120.30	1	2
1	A	32	ARG	NE-CZ-NH1	6.45	123.53	120.30	7	2
1	A	36	ARG	NE-CZ-NH1	6.30	123.45	120.30	4	2
1	A	33	ARG	NE-CZ-NH1	5.82	123.21	120.30	4	1
1	A	67	ARG	NE-CZ-NH1	5.72	123.16	120.30	9	2
1	A	32	ARG	NE-CZ-NH2	-5.55	117.52	120.30	7	2
1	A	98	ARG	NE-CZ-NH1	5.02	122.81	120.30	6	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	46	TYR	Mainchain,Sidechain	10
1	A	81	TYR	Sidechain	9
1	A	95	VAL	Mainchain	9
1	A	94	TYR	Sidechain,Mainchain	9
1	A	107	TYR	Sidechain,Mainchain	8
1	A	99	VAL	Mainchain	8
1	A	65	TYR	Sidechain	8
1	A	19	GLU	Mainchain	8
1	A	73	HIS	Mainchain,Sidechain	8
1	A	103	TYR	Sidechain,Mainchain	8
1	A	56	PHE	Sidechain	8
1	A	62	MET	Mainchain	7
1	A	54	PHE	Mainchain,Sidechain	7
1	A	75	ASN	Mainchain	6
1	A	14	THR	Mainchain	6
1	A	18	GLN	Mainchain	6
1	A	64	SER	Mainchain	6
1	A	50	VAL	Mainchain	6
1	A	91	ASP	Mainchain	5
1	A	78	PHE	Sidechain,Mainchain	5
1	A	68	SER	Mainchain	5
1	A	74	SER	Mainchain	5
1	A	53	PHE	Sidechain,Mainchain	4
1	A	86	ALA	Mainchain	4
1	A	55	GLY	Mainchain	4
1	A	37	SER	Mainchain	4
1	A	87	TRP	Mainchain	4
1	A	12	PHE	Sidechain	4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	57	PHE	Sidechain	3
1	A	49	MET	Mainchain	3
1	A	84	SER	Mainchain	3
1	A	16	LEU	Mainchain	3
1	A	52	GLY	Mainchain	3
1	A	60	GLY	Mainchain	3
1	A	47	VAL	Mainchain	3
1	A	98	ARG	Mainchain,Sidechain	3
1	A	66	ILE	Mainchain	3
1	A	92	LYS	Mainchain	2
1	A	70	LYS	Mainchain	2
1	A	82	ILE	Mainchain	2
1	A	13	LEU	Mainchain	2
1	A	1	MET	Mainchain	2
1	A	63	LEU	Mainchain	2
1	A	88	GLN	Mainchain	2
1	A	58	THR	Mainchain	2
1	A	89	GLU	Mainchain	2
1	A	93	ALA	Mainchain	2
1	A	38	GLY	Mainchain	2
1	A	5	ASN	Mainchain	2
1	A	100	LEU	Mainchain	2
1	A	104	LYS	Mainchain	2
1	A	76	ASP	Sidechain	1
1	A	39	ASP	Mainchain	1
1	A	45	LEU	Mainchain	1
1	A	67	ARG	Sidechain	1
1	A	15	LYS	Mainchain	1
1	A	11	PRO	Mainchain	1
1	A	41	LYS	Mainchain	1
1	A	85	ASP	Mainchain	1
1	A	7	THR	Mainchain	1
1	A	40	GLY	Mainchain	1
1	A	97	ALA	Mainchain	1
1	A	8	ALA	Mainchain	1
1	A	105	SER	Mainchain	1
1	A	59	LEU	Mainchain	1
1	A	77	PRO	Mainchain	1
1	A	4	SER	Mainchain	1
1	A	102	SER	Mainchain	1
1	A	2	ILE	Mainchain	1
1	A	6	THR	Mainchain	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	857	855	856	3±1
All	All	8570	8550	8560	28

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:23:GLN:O	1:A:26:ASN:ND2	0.69	2.25	9	1
1:A:99:VAL:O	1:A:103:TYR:N	0.57	2.34	2	3
1:A:7:THR:O	1:A:10:THR:OG1	0.57	2.21	8	2
1:A:102:SER:O	1:A:106:CYS:N	0.57	2.37	9	2
1:A:50:VAL:O	1:A:54:PHE:N	0.55	2.32	6	5
1:A:95:VAL:O	1:A:99:VAL:N	0.53	2.38	1	5
1:A:98:ARG:O	1:A:102:SER:N	0.50	2.37	10	2
1:A:68:SER:O	1:A:72:GLU:N	0.48	2.47	5	3
1:A:63:LEU:O	1:A:67:ARG:N	0.45	2.46	3	2
1:A:79:ASN:O	1:A:82:ILE:O	0.45	2.35	8	1
1:A:42:LEU:HD23	1:A:42:LEU:H	0.42	1.73	1	1
1:A:56:PHE:O	1:A:60:GLY:N	0.41	2.54	9	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	106/138 (77%)	97±3 (91±2%)	7±2 (7±2%)	2±1 (2±1%)	11	49
All	All	1060/1380 (77%)	967 (91%)	69 (7%)	24 (2%)	11	49

All 10 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	45	LEU	7
1	A	87	TRP	4
1	A	40	GLY	4
1	A	70	LYS	2
1	A	73	HIS	2
1	A	88	GLN	1
1	A	25	GLY	1
1	A	37	SER	1
1	A	29	GLY	1
1	A	107	TYR	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	93/121 (77%)	73±4 (78±4%)	20±4 (22±4%)	3	31
All	All	930/1210 (77%)	729 (78%)	201 (22%)	3	31

All 73 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	76	ASP	8
1	A	34	SER	8
1	A	85	ASP	7
1	A	4	SER	5
1	A	81	TYR	5
1	A	48	LEU	5
1	A	3	LEU	5
1	A	5	ASN	5
1	A	105	SER	5
1	A	64	SER	5
1	A	87	TRP	5
1	A	91	ASP	5
1	A	104	LYS	5
1	A	41	LYS	5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	30	LEU	4
1	A	1	MET	4
1	A	16	LEU	4
1	A	75	ASN	4
1	A	68	SER	4
1	A	71	LEU	4
1	A	37	SER	4
1	A	26	ASN	4
1	A	62	MET	4
1	A	90	LYS	3
1	A	72	GLU	3
1	A	39	ASP	3
1	A	42	LEU	3
1	A	59	LEU	3
1	A	67	ARG	3
1	A	10	THR	3
1	A	36	ARG	3
1	A	18	GLN	2
1	A	58	THR	2
1	A	100	LEU	2
1	A	94	TYR	2
1	A	98	ARG	2
1	A	101	GLU	2
1	A	9	VAL	2
1	A	49	MET	2
1	A	88	GLN	2
1	A	84	SER	2
1	A	79	ASN	2
1	A	45	LEU	2
1	A	22	GLN	2
1	A	46	TYR	2
1	A	27	MET	2
1	A	13	LEU	2
1	A	69	LYS	2
1	A	20	THR	2
1	A	96	GLN	2
1	A	89	GLU	2
1	A	28	SER	2
1	A	73	HIS	1
1	A	83	GLU	1
1	A	23	GLN	1
1	A	107	TYR	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	65	TYR	1
1	A	19	GLU	1
1	A	50	VAL	1
1	A	33	ARG	1
1	A	15	LYS	1
1	A	103	TYR	1
1	A	99	VAL	1
1	A	43	GLU	1
1	A	106	CYS	1
1	A	78	PHE	1
1	A	6	THR	1
1	A	7	THR	1
1	A	74	SER	1
1	A	82	ILE	1
1	A	53	PHE	1
1	A	92	LYS	1
1	A	12	PHE	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 carbohydrates 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