Full wwPDB NMR Structure Validation Report (i)

Apr 26, 2016 – 05:12 PM BST

PDB ID	:	1UJL
Title	:	Solution Structure of the HERG K+ channel S5-P extracellular linker
Authors	:	Torres, A.M.; Bansal, P.S.; Sunde, M.; Clarke, C.E.; Bursill, J.A.; Smith, D.J.; Bauskin, A.; Breit, S.N.; Campbell, T.J.; Alewood, P.F.; Kuchel, P.W.; Vandenberg, J.I.
Deposited on	:	2003-08-05

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
http://wwpdb.org/validation/2016/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:

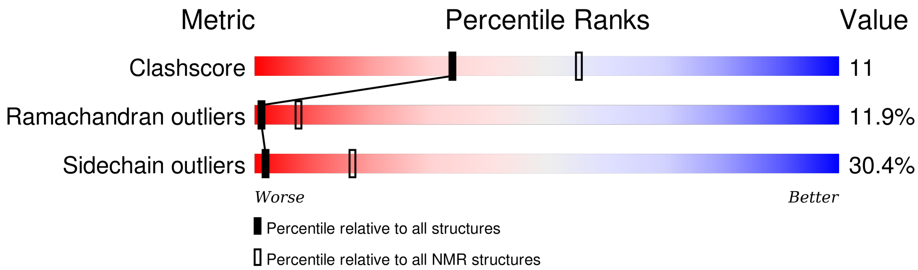
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027457

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 is 39%.

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	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

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	42	

2 Ensemble composition and analysis (i)

This entry contains 20 models.

Cyrange was unable to find well-defined residues.

Error message: Only domains with $<$ 8 residues could be identified.

NmrClust was unable to cluster the ensemble.

Error message: Wrapper check: not enough residues in core to run NmrClust

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 634 atoms, of which 312 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Potassium voltage-gated channel subfamily H member 2.

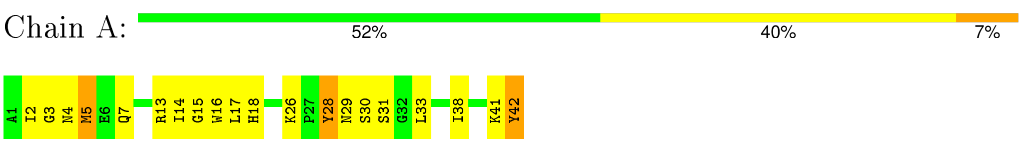
Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	42	634	200	312	58	62	2	0

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 H member 2

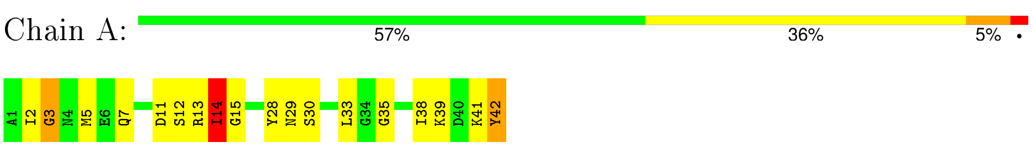


4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

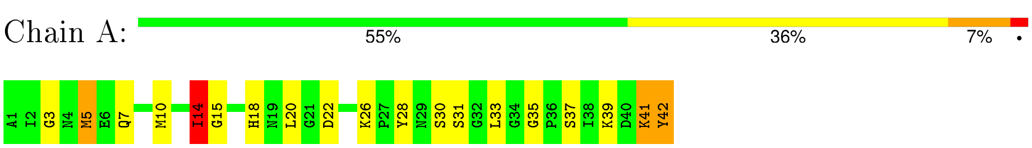
4.2.1 Score per residue for model 1

- Molecule 1: Potassium voltage-gated channel subfamily H member 2



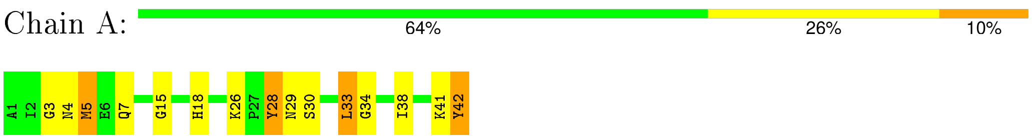
4.2.2 Score per residue for model 2

- Molecule 1: Potassium voltage-gated channel subfamily H member 2



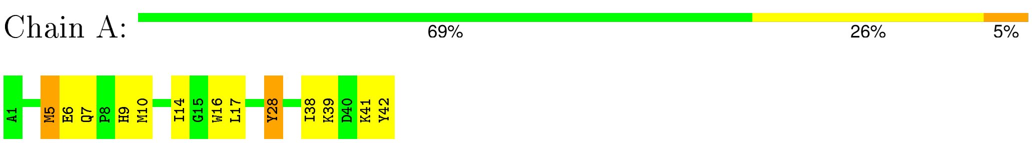
4.2.3 Score per residue for model 3

• Molecule 1: Potassium voltage-gated channel subfamily H member 2



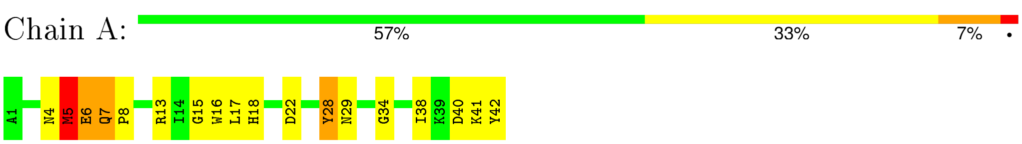
4.2.4 Score per residue for model 4

• Molecule 1: Potassium voltage-gated channel subfamily H member 2



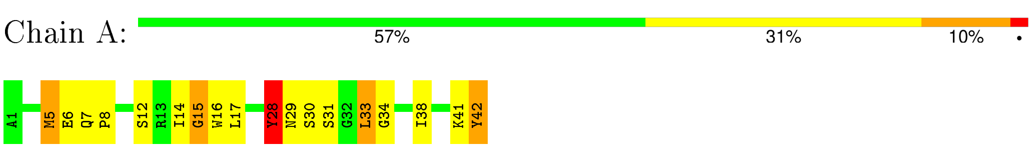
4.2.5 Score per residue for model 5

• Molecule 1: Potassium voltage-gated channel subfamily H member 2



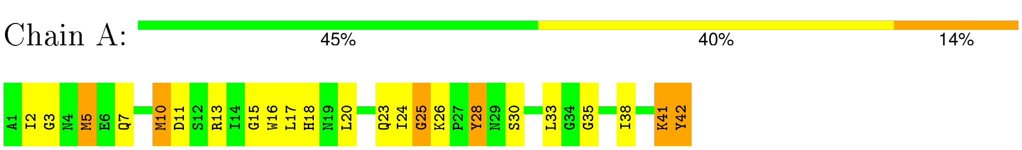
4.2.6 Score per residue for model 6

• Molecule 1: Potassium voltage-gated channel subfamily H member 2



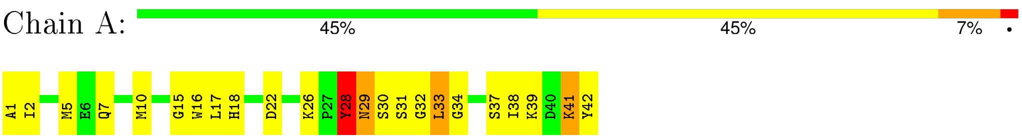
4.2.7 Score per residue for model 7

• Molecule 1: Potassium voltage-gated channel subfamily H member 2



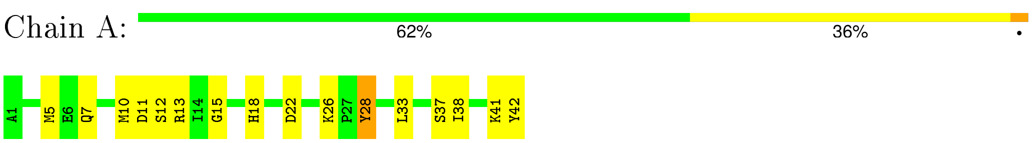
4.2.8 Score per residue for model 8

• Molecule 1: Potassium voltage-gated channel subfamily H member 2



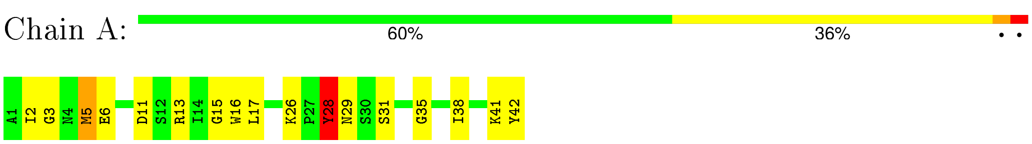
4.2.9 Score per residue for model 9

• Molecule 1: Potassium voltage-gated channel subfamily H member 2



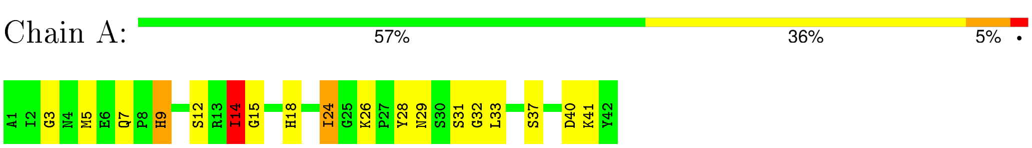
4.2.10 Score per residue for model 10

• Molecule 1: Potassium voltage-gated channel subfamily H member 2



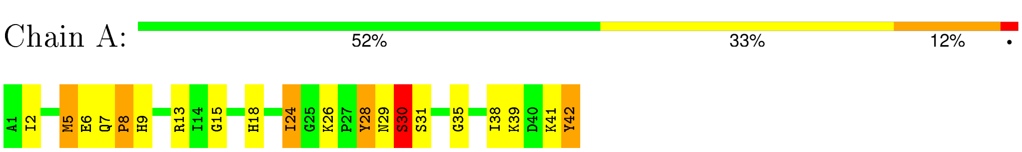
4.2.11 Score per residue for model 11

• Molecule 1: Potassium voltage-gated channel subfamily H member 2



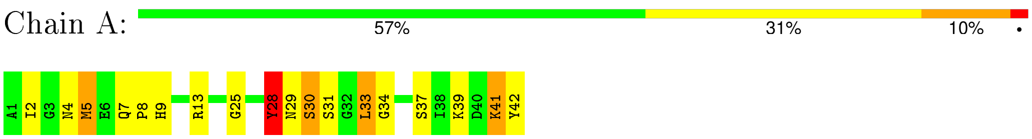
4.2.12 Score per residue for model 12

• Molecule 1: Potassium voltage-gated channel subfamily H member 2



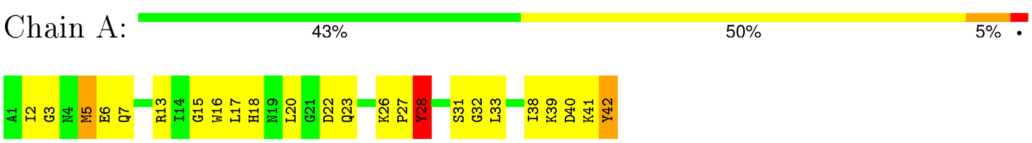
4.2.13 Score per residue for model 13

- Molecule 1: Potassium voltage-gated channel subfamily H member 2



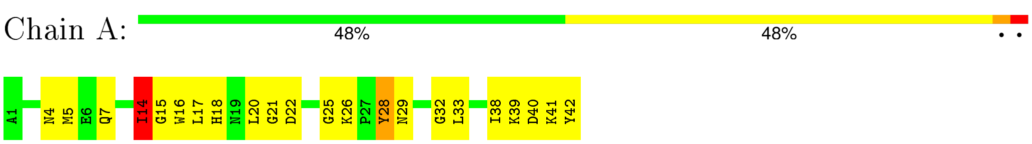
4.2.14 Score per residue for model 14

- Molecule 1: Potassium voltage-gated channel subfamily H member 2



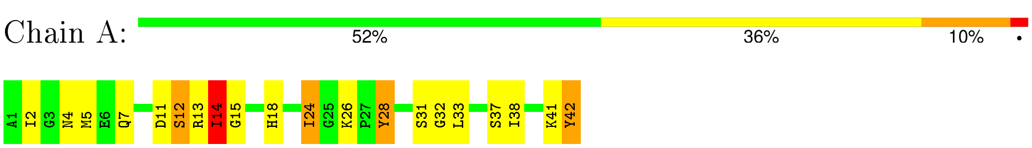
4.2.15 Score per residue for model 15

- Molecule 1: Potassium voltage-gated channel subfamily H member 2



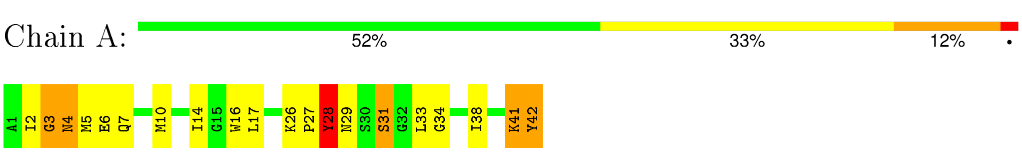
4.2.16 Score per residue for model 16

- Molecule 1: Potassium voltage-gated channel subfamily H member 2



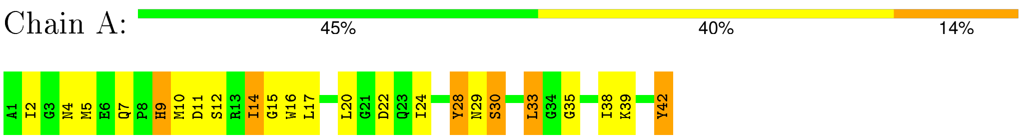
4.2.17 Score per residue for model 17

- Molecule 1: Potassium voltage-gated channel subfamily H member 2



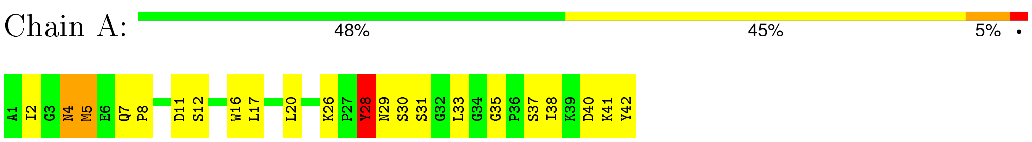
4.2.18 Score per residue for model 18

• Molecule 1: Potassium voltage-gated channel subfamily H member 2



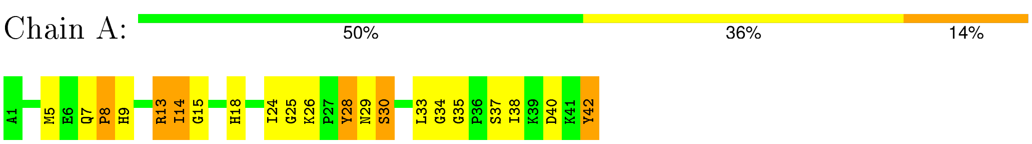
4.2.19 Score per residue for model 19

• Molecule 1: Potassium voltage-gated channel subfamily H member 2



4.2.20 Score per residue for model 20

• Molecule 1: Potassium voltage-gated channel subfamily H member 2



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *simulated annealing, molecular dynamics, distance geometry, torsion angle dynamics*.

Of the 1600 calculated structures, 20 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
DYANA	structure solution	1.5
CNS	refinement	1.1

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	BMRB entry 5922
Number of chemical shift lists	1
Total number of shifts	250
Number of shifts mapped to atoms	239
Number of unparsed shifts	0
Number of shifts with mapping errors	11
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	39%

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)

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	322	312	310	7±2
All	All	6440	6240	6200	136

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 11.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:20:LEU:O	1:A:24:ILE:HG22	0.78	1.79	18	1
1:A:20:LEU:HD12	1:A:21:GLY:N	0.76	1.95	15	1
1:A:34:GLY:C	1:A:38:ILE:HD13	0.71	2.04	6	1
1:A:24:ILE:HG22	1:A:24:ILE:O	0.64	1.92	16	3
1:A:14:ILE:HD13	1:A:15:GLY:N	0.64	2.07	18	1
1:A:28:TYR:N	1:A:28:TYR:CD1	0.62	2.66	14	4
1:A:16:TRP:CZ2	1:A:17:LEU:HD11	0.61	2.30	17	11
1:A:14:ILE:HG23	1:A:15:GLY:N	0.60	2.11	2	5
1:A:13:ARG:O	1:A:14:ILE:HG22	0.57	1.99	20	1
1:A:38:ILE:HG23	1:A:42:TYR:CE1	0.57	2.35	17	11
1:A:8:PRO:O	1:A:9:HIS:CG	0.54	2.61	12	1
1:A:2:ILE:O	1:A:2:ILE:HG22	0.53	2.04	19	3
1:A:28:TYR:O	1:A:28:TYR:CG	0.52	2.61	12	3
1:A:34:GLY:HA3	1:A:38:ILE:HD12	0.52	1.82	17	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:35:GLY:N	1:A:38:ILE:HD13	0.51	2.20	10	1
1:A:35:GLY:HA3	1:A:38:ILE:HD12	0.50	1.82	12	1
1:A:28:TYR:O	1:A:30:SER:N	0.48	2.46	13	5
1:A:16:TRP:CE2	1:A:17:LEU:HD11	0.47	2.44	19	1
1:A:16:TRP:CE2	1:A:17:LEU:CD1	0.47	2.97	19	1
1:A:38:ILE:HG23	1:A:42:TYR:HE1	0.47	1.70	3	14
1:A:24:ILE:CG2	1:A:24:ILE:O	0.47	2.63	12	3
1:A:28:TYR:O	1:A:29:ASN:C	0.46	2.53	8	1
1:A:16:TRP:CZ2	1:A:17:LEU:CD1	0.46	2.98	6	3
1:A:20:LEU:HA	1:A:23:GLN:CG	0.46	2.39	14	2
1:A:29:ASN:ND2	1:A:33:LEU:HD12	0.45	2.26	6	1
1:A:27:PRO:O	1:A:28:TYR:C	0.45	2.54	17	2
1:A:33:LEU:HD23	1:A:34:GLY:N	0.45	2.27	8	1
1:A:7:GLN:N	1:A:8:PRO:CD	0.45	2.80	20	1
1:A:14:ILE:HD13	1:A:14:ILE:C	0.45	2.31	18	1
1:A:33:LEU:HD23	1:A:34:GLY:H	0.44	1.73	6	4
1:A:2:ILE:HG22	1:A:3:GLY:N	0.44	2.28	1	5
1:A:28:TYR:CE1	1:A:31:SER:CB	0.44	3.00	17	1
1:A:24:ILE:HG23	1:A:25:GLY:N	0.44	2.28	20	2
1:A:38:ILE:O	1:A:41:LYS:N	0.43	2.51	8	2
1:A:28:TYR:CD1	1:A:28:TYR:O	0.43	2.71	12	1
1:A:7:GLN:O	1:A:7:GLN:CG	0.43	2.67	13	1
1:A:13:ARG:O	1:A:14:ILE:O	0.43	2.37	20	1
1:A:14:ILE:CG2	1:A:15:GLY:N	0.43	2.82	11	2
1:A:5:MET:O	1:A:7:GLN:N	0.43	2.52	5	1
1:A:28:TYR:CD1	1:A:28:TYR:N	0.43	2.86	16	2
1:A:41:LYS:HB3	1:A:42:TYR:CE1	0.43	2.49	17	4
1:A:28:TYR:CG	1:A:28:TYR:O	0.42	2.72	2	1
1:A:14:ILE:HG22	1:A:15:GLY:N	0.42	2.29	6	1
1:A:2:ILE:N	1:A:2:ILE:HD12	0.42	2.28	13	1
1:A:5:MET:CG	1:A:6:GLU:N	0.42	2.82	5	1
1:A:15:GLY:O	1:A:18:HIS:N	0.42	2.53	8	12
1:A:14:ILE:HG22	1:A:16:TRP:CD1	0.41	2.49	4	2
1:A:1:ALA:O	1:A:2:ILE:HD13	0.41	2.15	8	1
1:A:8:PRO:C	1:A:9:HIS:CG	0.41	2.94	20	1
1:A:7:GLN:N	1:A:8:PRO:HD3	0.41	2.31	13	1
1:A:12:SER:OG	1:A:13:ARG:N	0.40	2.54	9	1
1:A:7:GLN:CG	1:A:8:PRO:HD2	0.40	2.46	5	1
1:A:8:PRO:HG2	1:A:9:HIS:CE1	0.40	2.52	20	1
1:A:7:GLN:CB	1:A:8:PRO:CD	0.40	2.99	5	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	40/42 (95%)	25±2 (62±6%)	11±2 (26±6%)	5±1 (12±3%)	1	7
All	All	800/840 (95%)	493 (62%)	212 (26%)	95 (12%)	1	7

All 22 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	5	MET	11
1	A	28	TYR	11
1	A	29	ASN	11
1	A	30	SER	6
1	A	35	GLY	6
1	A	14	ILE	6
1	A	3	GLY	5
1	A	32	GLY	5
1	A	6	GLU	5
1	A	8	PRO	4
1	A	31	SER	3
1	A	4	ASN	3
1	A	24	ILE	3
1	A	9	HIS	3
1	A	25	GLY	3
1	A	15	GLY	2
1	A	34	GLY	2
1	A	11	ASP	2
1	A	33	LEU	1
1	A	13	ARG	1
1	A	10	MET	1
1	A	12	SER	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	34/34 (100%)	24±2 (70±7%)	10±2 (30±7%)	2	17
All	All	680/680 (100%)	473 (70%)	207 (30%)	2	17

All 25 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	5	MET	20
1	A	28	TYR	18
1	A	7	GLN	17
1	A	41	LYS	17
1	A	33	LEU	16
1	A	26	LYS	14
1	A	42	TYR	12
1	A	39	LYS	9
1	A	31	SER	8
1	A	13	ARG	8
1	A	37	SER	8
1	A	22	ASP	7
1	A	10	MET	7
1	A	4	ASN	7
1	A	14	ILE	6
1	A	40	ASP	6
1	A	12	SER	6
1	A	30	SER	5
1	A	11	ASP	5
1	A	9	HIS	3
1	A	20	LEU	2
1	A	6	GLU	2
1	A	29	ASN	2
1	A	2	ILE	1
1	A	38	ILE	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 carbohydrates 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)

The completeness of assignment taking into account all chemical shift lists is 39% for the well-defined parts and 39% for the entire structure.

7.1 Chemical shift list 1

File name: BMRB entry 5922

Chemical shift list name: *assigned_chem_shift_list_1*

7.1.1 Bookkeeping (i)

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	250
Number of shifts mapped to atoms	239
Number of unparsed shifts	0
Number of shifts with mapping errors	11
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- Residue not found in structure. All 11 occurences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	5	NLE	HB2	1.41	0.01	1
A	5	NLE	H	7.974	0.01	1
A	10	NLE	HG2	1.339	0.01	1
A	5	NLE	HD2	0.902	0.01	1
A	10	NLE	HD2	0.89	0.01	1
A	5	NLE	HB3	1.815	0.01	1
A	10	NLE	H	8.114	0.01	1
A	5	NLE	HA	4.212	0.01	1
A	5	NLE	HG2	1.342	0.01	1
A	10	NLE	HA	4.168	0.01	1
A	10	NLE	HB2	1.812	0.01	1

7.1.2 Chemical shift referencing

No chemical shift referencing corrections were calculated (not enough data).

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 39%, i.e. 195 atoms were assigned a chemical shift out of a possible 500. 0 out of 3 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	72/204 (35%)	72/81 (89%)	0/84 (0%)	0/39 (0%)
Sidechain	109/252 (43%)	109/151 (72%)	0/90 (0%)	0/11 (0%)
Aromatic	14/44 (32%)	14/22 (64%)	0/17 (0%)	0/5 (0%)
Overall	195/500 (39%)	195/254 (77%)	0/191 (0%)	0/55 (0%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 39%, i.e. 195 atoms were assigned a chemical shift out of a possible 500. 0 out of 3 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	72/204 (35%)	72/81 (89%)	0/84 (0%)	0/39 (0%)
Sidechain	109/252 (43%)	109/151 (72%)	0/90 (0%)	0/11 (0%)
Aromatic	14/44 (32%)	14/22 (64%)	0/17 (0%)	0/5 (0%)
Overall	195/500 (39%)	195/254 (77%)	0/191 (0%)	0/55 (0%)

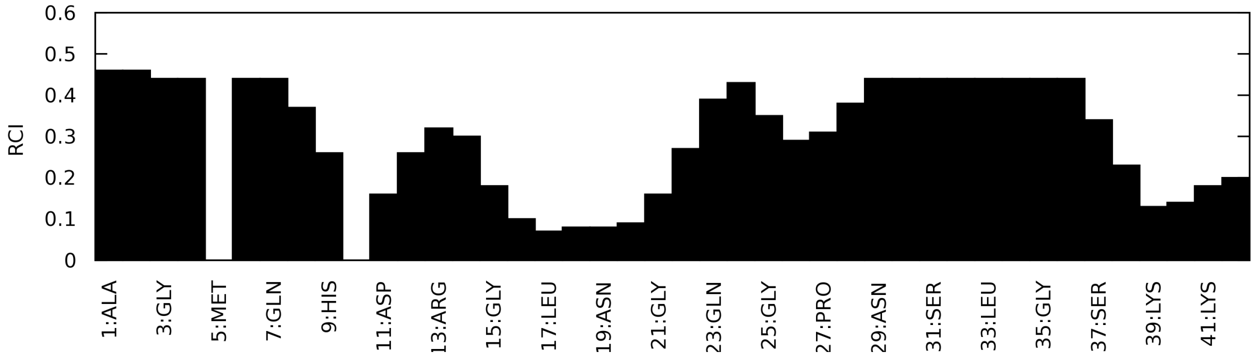
7.1.4 Statistically unusual chemical shifts

There are no statistically unusual chemical shifts.

7.1.5 Random Coil Index (RCI) plots

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.

Random coil index (RCI) for chain A:


RCI
0.6
0.5
0.4
0.3
0.2
0.1
0
1:ALA
3:GLY
5:MET
7:GLN
9:HIS
11:ASP
13:ARG
15:GLY
17:LEU
19:ASN
21:GLY
23:GLN
25:GLY
27:PRO
29:ASN
31:SER
33:LEU
35:GLY
37:SER
39:LYS
41:LYS