Full wwPDB NMR Structure Validation Report (i)

May 31, 2020 – 10:11 pm BST

PDB ID	:	6FEH
Title	:	Solution Structure of CaM/Kv7.2-hAB Complex
Authors	:	Bernardo-Seisdedos, G.; Villarroel, A.; Millet, O.
Deposited on	:	2018-01-02

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 (1)) were used in the production of this report:

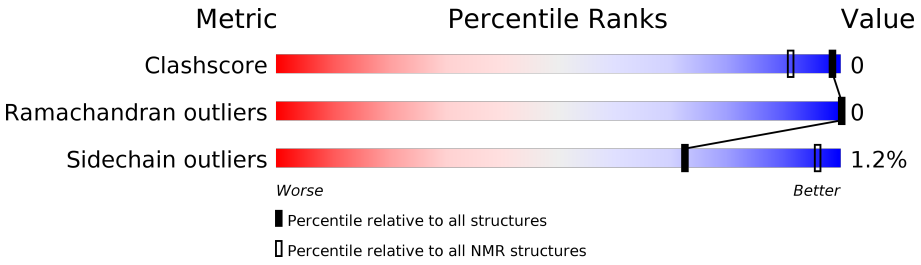
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 82%.

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	115	43% · 57%
2	B	149	93% 7% ·

2 Ensemble composition and analysis

This entry contains 10 models. Model 1 is the overall representative, medoid model (most similar to other models).

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:329-A:351, A:506-A:532, B:5-B:74, B:81-B:148 (188)	0.65	1

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 4 single-model clusters were found.

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

3 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4164 atoms, of which 2035 are hydrogens and 0 are deuteriums.

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

Mol	Chain	Residues	Atoms						Trace
1	A	115	Total	C	H	N	O	S	0
			1900	611	941	176	167	5	

There are 37 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP O43526
A	2	SER	-	expression tag	UNP O43526
A	3	TYR	-	expression tag	UNP O43526
A	4	TYR	-	expression tag	UNP O43526
A	5	HIS	-	expression tag	UNP O43526
A	6	HIS	-	expression tag	UNP O43526
A	7	HIS	-	expression tag	UNP O43526
A	8	HIS	-	expression tag	UNP O43526
A	9	HIS	-	expression tag	UNP O43526
A	10	HIS	-	expression tag	UNP O43526
A	11	ASP	-	expression tag	UNP O43526
A	12	TYR	-	expression tag	UNP O43526
A	13	ASP	-	expression tag	UNP O43526
A	14	ILE	-	expression tag	UNP O43526
A	15	PRO	-	expression tag	UNP O43526
A	16	THR	-	expression tag	UNP O43526
A	17	THR	-	expression tag	UNP O43526
A	18	GLU	-	expression tag	UNP O43526
A	19	ASN	-	expression tag	UNP O43526
A	20	LEU	-	expression tag	UNP O43526
A	21	TYR	-	expression tag	UNP O43526
A	22	PHE	-	expression tag	UNP O43526
A	23	GLN	-	expression tag	UNP O43526
A	24	GLY	-	expression tag	UNP O43526
A	25	ALA	-	expression tag	UNP O43526
A	26	MET	-	expression tag	UNP O43526
A	310	GLY	-	linker	UNP O43526
A	311	ILE	-	linker	UNP O43526
A	312	LEU	-	linker	UNP O43526
A	313	GLY	-	linker	UNP O43526

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	314	SER	-	linker	UNP O43526
A	315	GLY	-	linker	UNP O43526
A	373	ARG	-	linker	UNP O43526
A	392	GLY	-	linker	UNP O43526
A	393	LEU	-	linker	UNP O43526
A	548	LEU	-	linker	UNP O43526
A	549	ASP	-	linker	UNP O43526

- Molecule 2 is a protein called Calmodulin-1.

Mol	Chain	Residues	Atoms						Trace
2	B	148	Total	C	H	N	O	S	0
			2260	714	1094	188	255	9	

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	
3	B	4	Total	Ca
			4	4

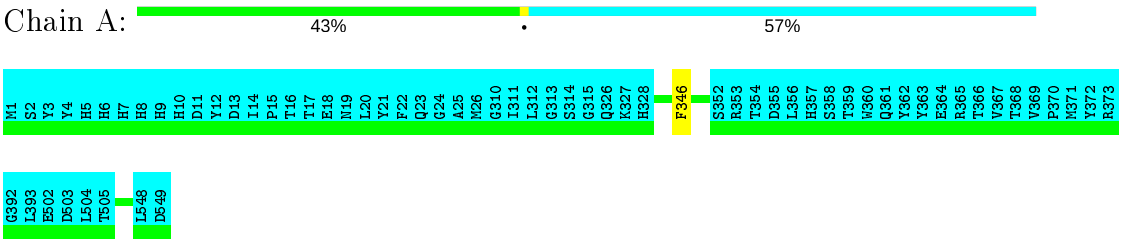
4 Residue-property plots

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.

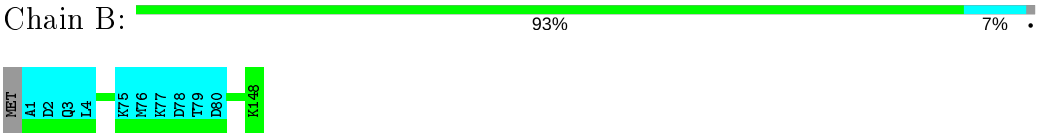
• Molecule 1: Potassium voltage-gated channel subfamily KQT member 2,Potassium voltage-gated channel subfamily KQT member 2

Chain A:



• Molecule 2: Calmodulin-1

Chain B:



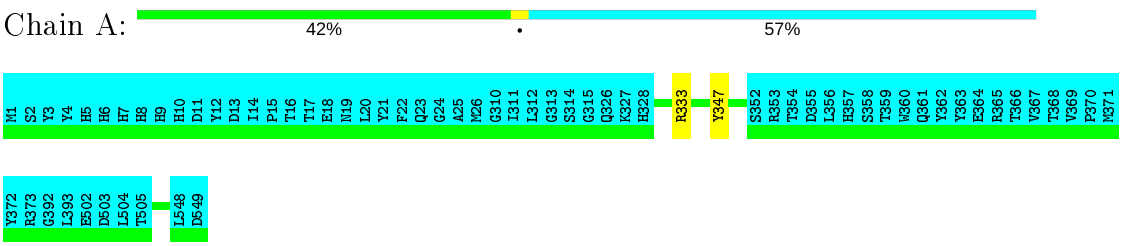
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1 (medoid)

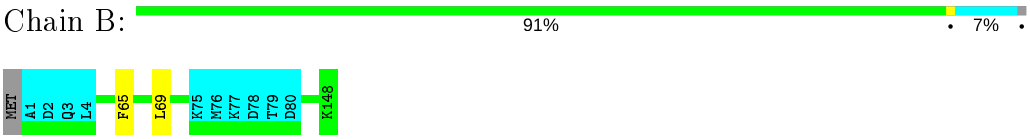
• Molecule 1: Potassium voltage-gated channel subfamily KQT member 2,Potassium voltage-gated channel subfamily KQT member 2

Chain A:



• Molecule 2: Calmodulin-1

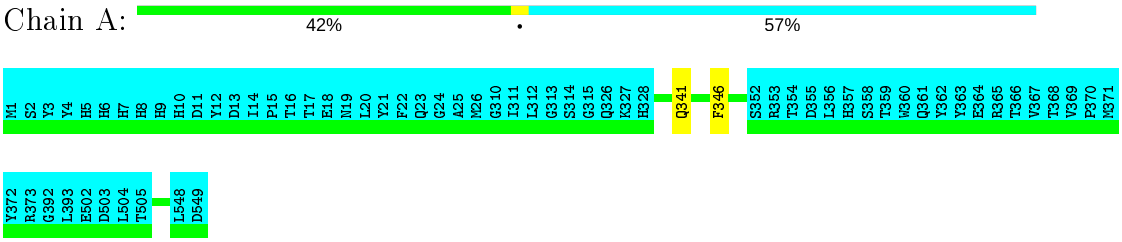
Chain B:



4.2.2 Score per residue for model 2

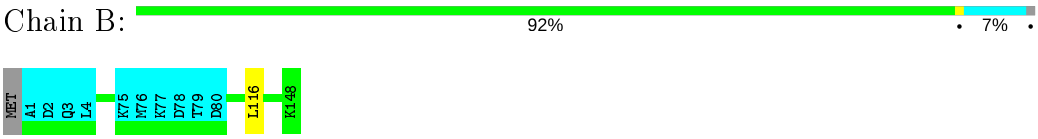
• Molecule 1: Potassium voltage-gated channel subfamily KQT member 2,Potassium voltage-gated channel subfamily KQT member 2

Chain A:



• Molecule 2: Calmodulin-1

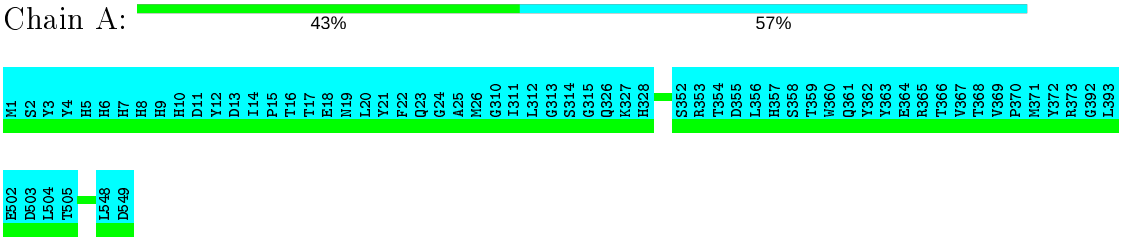
Chain B:



4.2.3 Score per residue for model 3

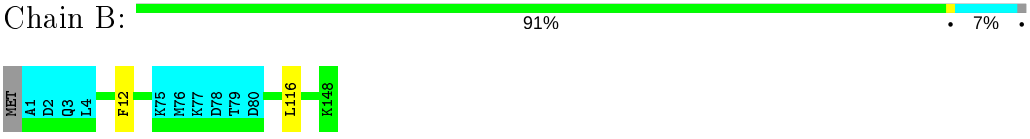
• Molecule 1: Potassium voltage-gated channel subfamily KQT member 2,Potassium voltage-gated channel subfamily KQT member 2

Chain A:



• Molecule 2: Calmodulin-1

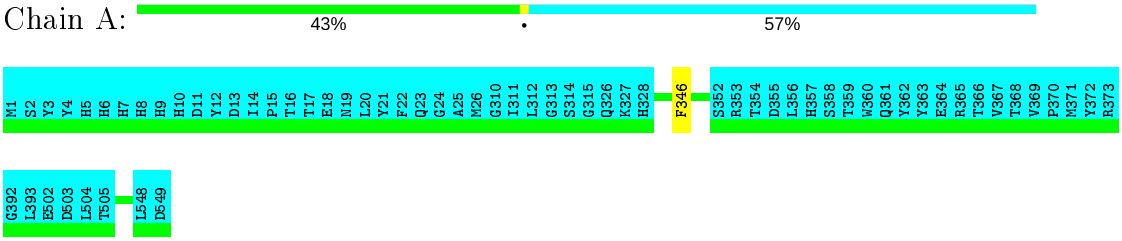
Chain B:



4.2.4 Score per residue for model 4

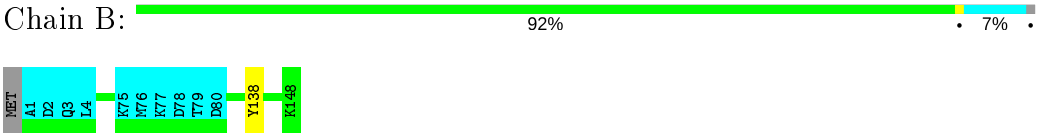
• Molecule 1: Potassium voltage-gated channel subfamily KQT member 2,Potassium voltage-gated channel subfamily KQT member 2

Chain A:



• Molecule 2: Calmodulin-1

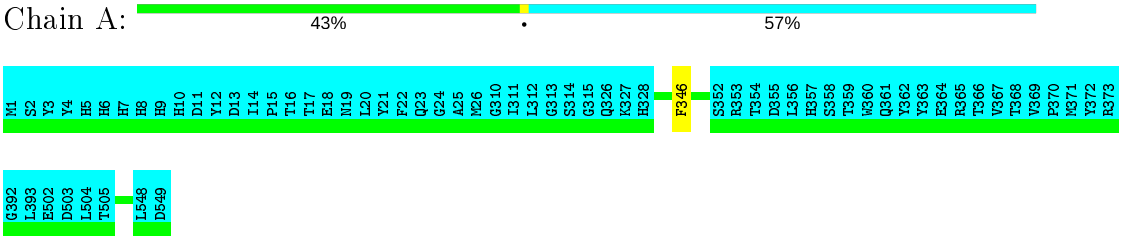
Chain B:



4.2.5 Score per residue for model 5

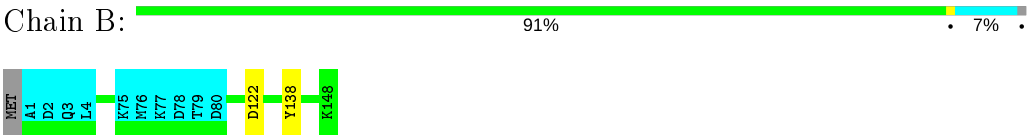
• Molecule 1: Potassium voltage-gated channel subfamily KQT member 2,Potassium voltage-gated channel subfamily KQT member 2

Chain A:



• Molecule 2: Calmodulin-1

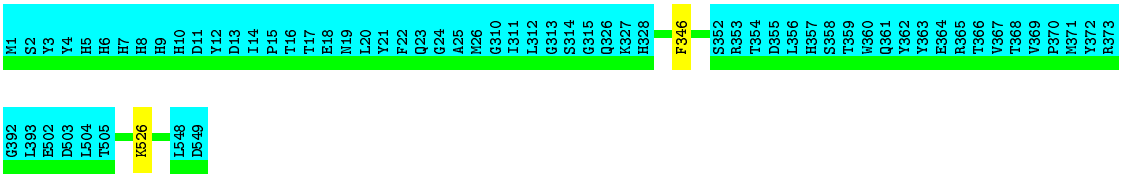
Chain B:



4.2.6 Score per residue for model 6

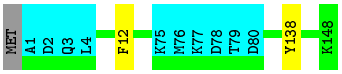
• Molecule 1: Potassium voltage-gated channel subfamily KQT member 2,Potassium voltage-gated channel subfamily KQT member 2

Chain A:



• Molecule 2: Calmodulin-1

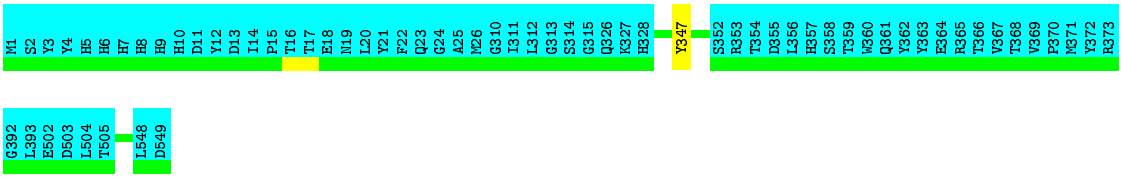
Chain B:



4.2.7 Score per residue for model 7

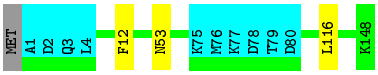
• Molecule 1: Potassium voltage-gated channel subfamily KQT member 2,Potassium voltage-gated channel subfamily KQT member 2

Chain A:



• Molecule 2: Calmodulin-1

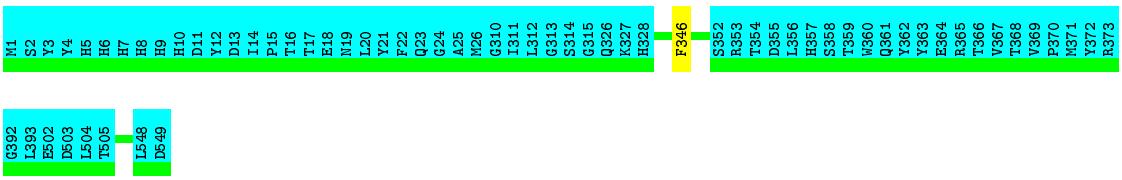
Chain B:



4.2.8 Score per residue for model 8

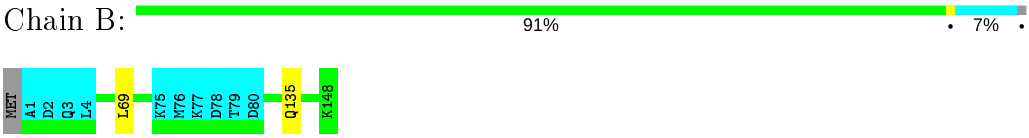
• Molecule 1: Potassium voltage-gated channel subfamily KQT member 2,Potassium voltage-gated channel subfamily KQT member 2

Chain A:



• Molecule 2: Calmodulin-1

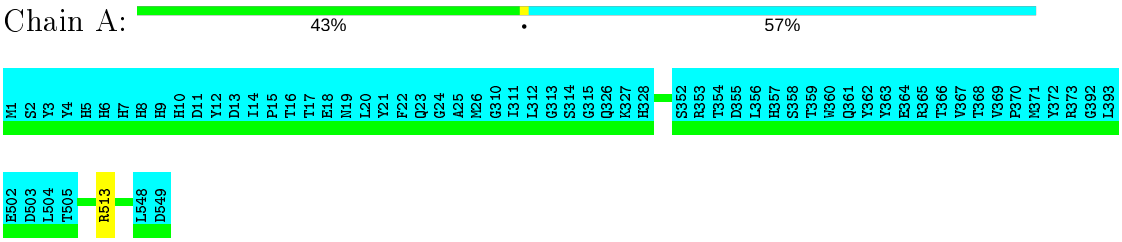
Chain B:



4.2.9 Score per residue for model 9

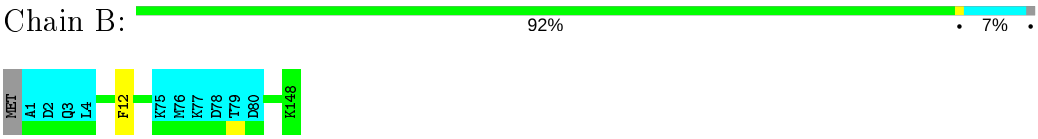
• Molecule 1: Potassium voltage-gated channel subfamily KQT member 2,Potassium voltage-gated channel subfamily KQT member 2

Chain A:



• Molecule 2: Calmodulin-1

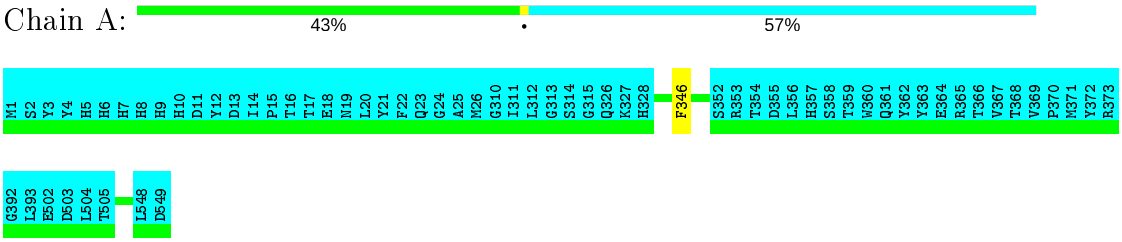
Chain B:



4.2.10 Score per residue for model 10

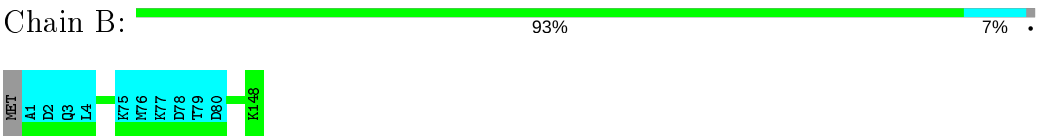
• Molecule 1: Potassium voltage-gated channel subfamily KQT member 2,Potassium voltage-gated channel subfamily KQT member 2

Chain A:



• Molecule 2: Calmodulin-1

Chain B:



5 Refinement protocol and experimental data overview

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

Of the 200 calculated structures, 10 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
ARIA	structure calculation	2.3.1
ARIA	refinement	2.3.1

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 6 of this report.

Chemical shift file(s)	input_cs.cif
Number of chemical shift lists	1
Total number of shifts	2851
Number of shifts mapped to atoms	2851
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	82%

No validations of the models with respect to experimental NMR restraints is performed at this time.

COVALENT-GEOMETRY INFOmissingINFO

5.1 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	411	442	442	0±0
2	B	1087	1017	1016	0±0
All	All	15020	14590	14579	2

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 0.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:B:65:PHE:O	2:B:69:LEU:N	0.45	2.45	1	1
1:A:341:GLN:HG3	2:B:116:LEU:HD22	0.41	1.93	2	1

5.2 Torsion angles (i)

5.2.1 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	50/115 (43%)	50±0 (99±1%)	0±0 (1±1%)	0±0 (0±0%)	100	100
2	B	137/149 (92%)	135±2 (99±1%)	2±2 (1±1%)	0±0 (0±0%)	100	100
All	All	1870/2640 (71%)	1849 (99%)	21 (1%)	0 (0%)	100	100

There are no Ramachandran outliers.

5.2.2 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	43/102 (42%)	42±1 (97±1%)	1±1 (3±1%)	49	91
2	B	117/127 (92%)	116±1 (99±1%)	1±1 (1±1%)	84	97
All	All	1600/2290 (70%)	1581 (99%)	19 (1%)	72	96

All 10 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	346	PHE	6
2	B	138	TYR	3
2	B	116	LEU	2
1	A	347	TYR	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	513	ARG	1
1	A	526	LYS	1
2	B	122	ASP	1
2	B	53	ASN	1
1	A	333	ARG	1
2	B	135	GLN	1

5.2.3 RNA

There are no RNA molecules in this entry.

5.3 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.4 Carbohydrates

There are no carbohydrates in this entry.

5.5 Ligand geometry

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

5.6 Other polymers

There are no such molecules in this entry.

5.7 Polymer linkage issues

There are no chain breaks in this entry.

6 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 82% for the well-defined parts and 77% for the entire structure.

6.1 Chemical shift list 1

File name: input_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

6.1.1 Bookkeeping

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	2851
Number of shifts mapped to atoms	2851
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	1

6.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	252	0.01 ± 0.06	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	234	0.01 ± 0.05	None needed (< 0.5 ppm)
$^{13}C'$	250	0.04 ± 0.06	None needed (< 0.5 ppm)
^{15}N	239	0.04 ± 0.26	None needed (< 0.5 ppm)

6.1.3 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 82%, i.e. 1901 atoms were assigned a chemical shift out of a possible 2331. 15 out of 24 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	919/932 (99%)	367/372 (99%)	373/376 (99%)	179/184 (97%)
Sidechain	914/1247 (73%)	559/727 (77%)	355/456 (78%)	0/64 (0%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	68/152 (45%)	68/82 (83%)	0/67 (0%)	0/3 (0%)
Overall	1901/2331 (82%)	994/1181 (84%)	728/899 (81%)	179/251 (71%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 77%, i.e. 2535 atoms were assigned a chemical shift out of a possible 3273. 21 out of 33 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	1234/1303 (95%)	493/520 (95%)	502/526 (95%)	239/257 (93%)
Sidechain	1207/1682 (72%)	737/983 (75%)	470/618 (76%)	0/81 (0%)
Aromatic	94/288 (33%)	94/153 (61%)	0/120 (0%)	0/15 (0%)
Overall	2535/3273 (77%)	1324/1656 (80%)	972/1264 (77%)	239/353 (68%)

6.1.4 Statistically unusual chemical shifts (i)

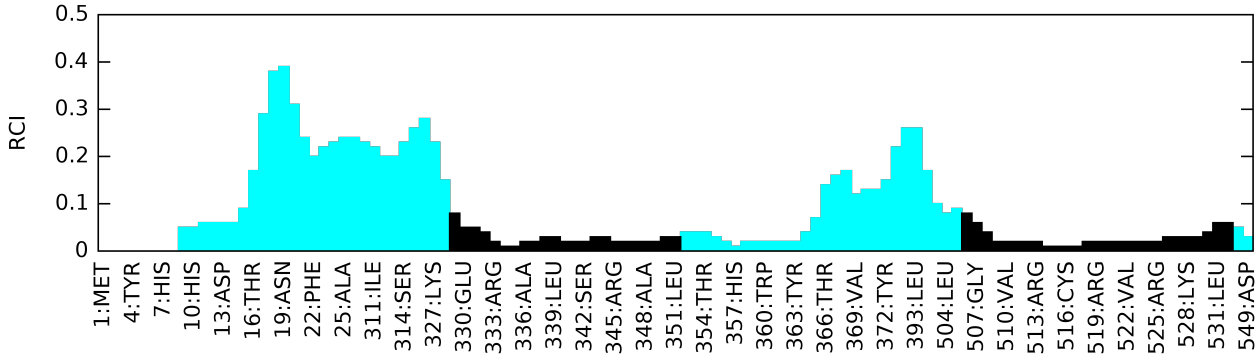
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
2	B	93	ASP	HB2	1.15	4.07 – 1.37	-5.8

6.1.5 Random Coil Index (RCI) plots (i)

The images below report *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:



Random coil index (RCI) for chain B: