Full wwPDB NMR Structure Validation Report (i)

May 28, 2020 – 11:55 pm BST

PDB ID	:	2MKI
Title	:	Solution structure of tandem RRM domains of cytoplasmic polyadenylation element binding protein 4 (CPEB4) in complex with RNA
Authors	:	Afroz, T.; Skrisovska, L.; Belloc, E.; Boixet, J.G.; Mendez, R.; Allain, F.H.-T.
Deposited on	:	2014-02-07

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:

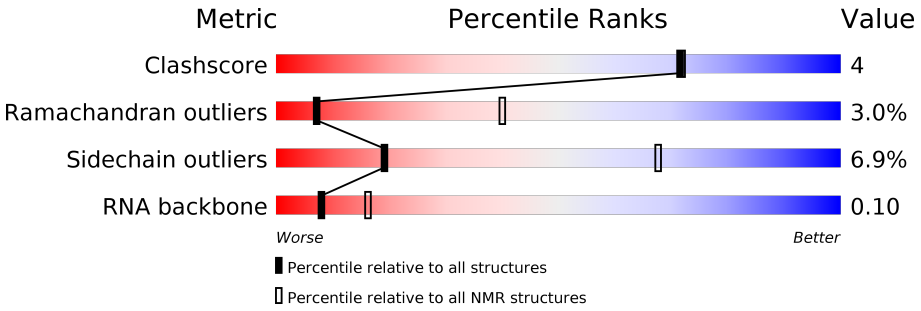
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.11
Ideal geometry (proteins)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

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment is 66%.

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
RNA backbone	4643	676

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	203	82% 8% 9%
2	B	5	60% 40%

2 Ensemble composition and analysis

This entry contains 20 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:62-A:99, A:109-A:254 (184)	0.74	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 3 clusters and 11 single-model clusters were found.

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

3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3355 atoms, of which 1647 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Cytoplasmic polyadenylation element-binding protein 4.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	203	3202	1029	1593	274	300	6	0

- Molecule 2 is a RNA chain called RNA (5'-R(*CP*UP*UP*UP*A)-3').

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
2	B	5	153	46	54	14	35	4	0

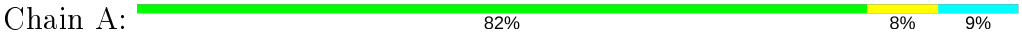
4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA 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.

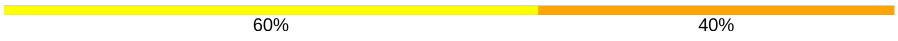
• Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:



• Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

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: Cytoplasmic polyadenylation element-binding protein 4

Chain A:

• Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

Chain B:

4.2.2 Score per residue for model 2

• Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:

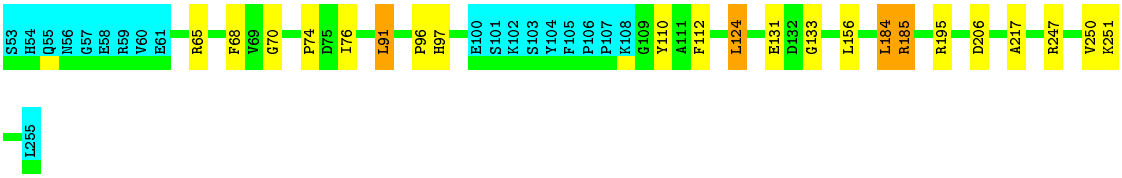
• Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

Chain B:

4.2.3 Score per residue for model 3

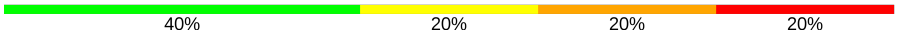
• Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:



• Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

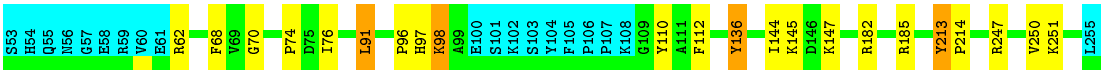
Chain B:



4.2.4 Score per residue for model 4

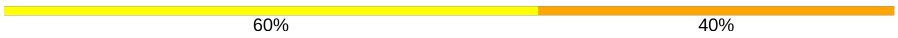
• Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:



• Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

Chain B:



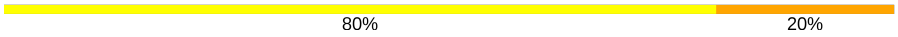
4.2.5 Score per residue for model 5

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:

- Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

Chain B:



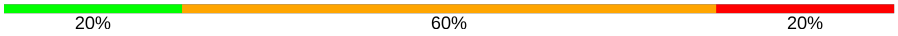
4.2.6 Score per residue for model 6

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:

- Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

Chain B:



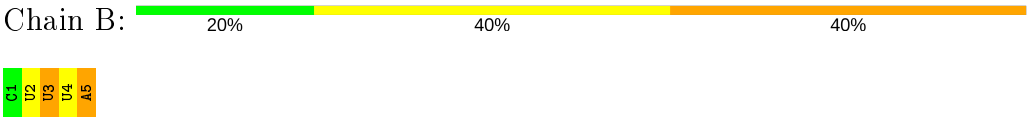
4.2.7 Score per residue for model 7

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:

- Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

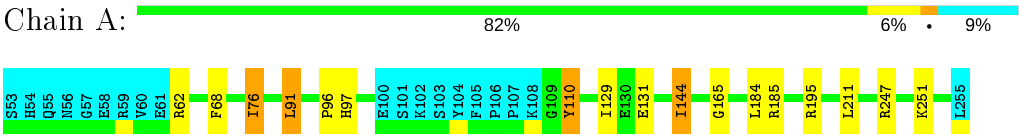
Chain B:



4.2.8 Score per residue for model 8

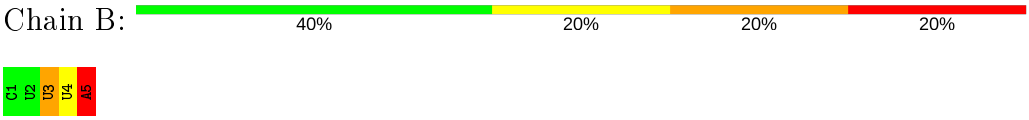
- Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:



- Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

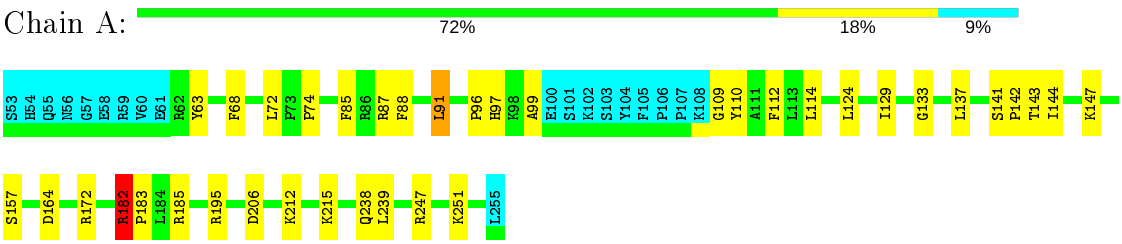
Chain B:



4.2.9 Score per residue for model 9

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:



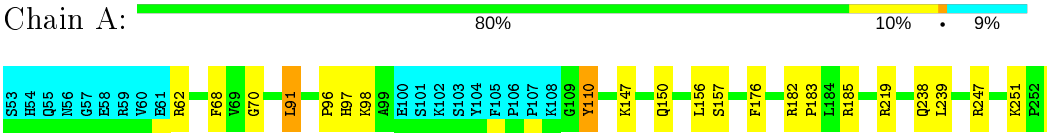
- Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

Chain B:

4.2.10 Score per residue for model 10

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:



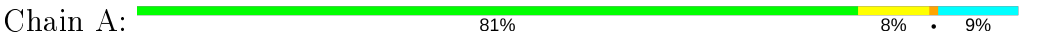
• Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

Chain B:

4.2.11 Score per residue for model 11

• Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:



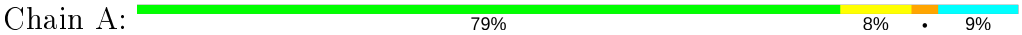
• Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

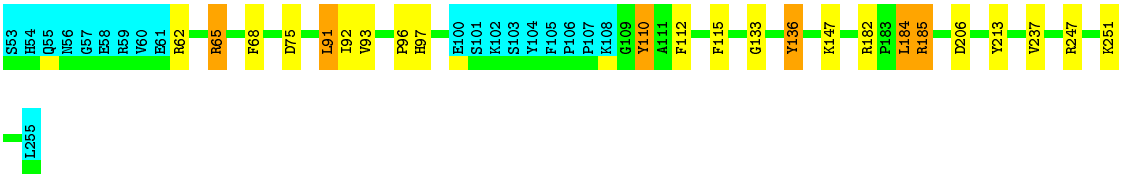
Chain B:

4.2.12 Score per residue for model 12

• Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:





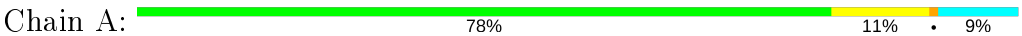
• Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

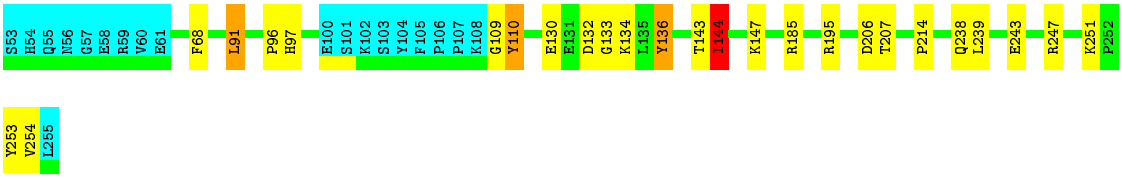
Chain B:

4.2.13 Score per residue for model 13

• Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:





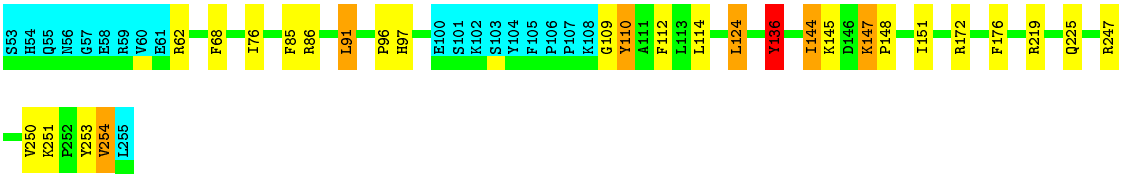
• Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

Chain B:

4.2.14 Score per residue for model 14

• Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:



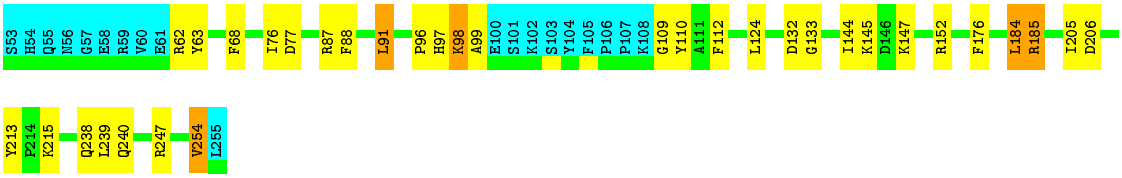
• Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

Chain B:

4.2.15 Score per residue for model 15

• Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:



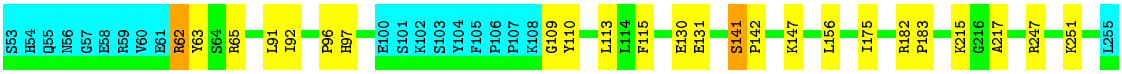
• Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

Chain B:

4.2.16 Score per residue for model 16

• Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:



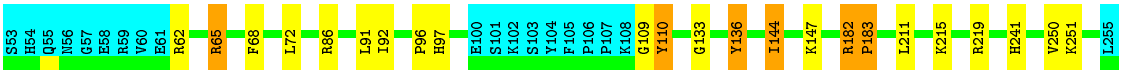
• Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

Chain B:

4.2.17 Score per residue for model 17

• Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:



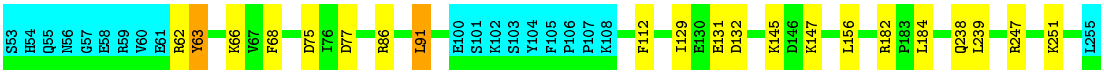
• Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

Chain B:

4.2.18 Score per residue for model 18

• Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:



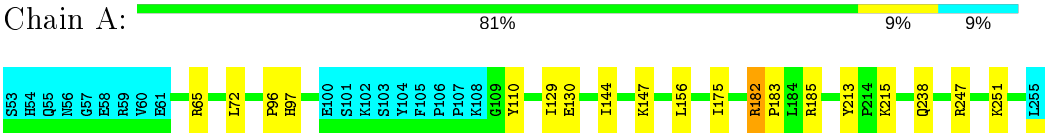
• Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

Chain B:

4.2.19 Score per residue for model 19

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:



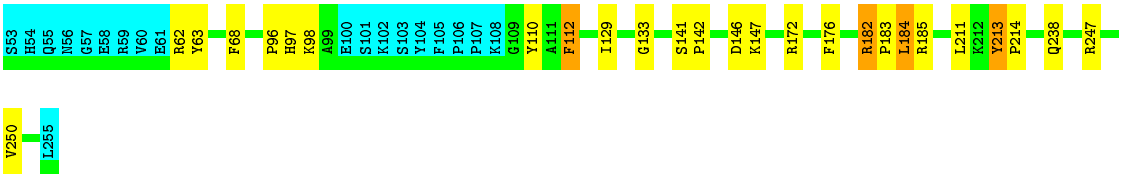
- Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

Chain B:

4.2.20 Score per residue for model 20

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 4

Chain A:



- Molecule 2: RNA (5'-R(*CP*UP*UP*UP*A)-3')

Chain B:

5 Refinement protocol and experimental data overview

The models were refined using the following method: *simulated annealing.*

Of the 30 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
AMBER	refinement	

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	1845
Number of shifts mapped to atoms	1845
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	66%

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	1451	1443	1443	10±4
2	B	99	54	54	4±2
All	All	31000	29940	29940	216

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 4.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:147:LYS:HE2	2:B:1:C:C6	0.65	2.26	5	4
1:A:156:LEU:N	1:A:156:LEU:HD12	0.63	2.09	16	2
1:A:156:LEU:HD12	1:A:156:LEU:H	0.61	1.55	16	1
1:A:182:ARG:H	1:A:183:PRO:CD	0.61	2.09	17	3
1:A:156:LEU:CD1	1:A:156:LEU:H	0.61	2.09	16	1
1:A:63:TYR:C	1:A:156:LEU:HD11	0.59	2.17	16	1
1:A:143:THR:HG23	1:A:144:ILE:HG13	0.59	1.75	9	2
1:A:96:PRO:HG2	1:A:97:HIS:CD2	0.58	2.34	12	18
1:A:156:LEU:H	1:A:156:LEU:HD13	0.57	1.60	11	1
1:A:184:LEU:CD1	1:A:185:ARG:H	0.55	2.14	12	3
1:A:147:LYS:HD3	2:B:1:C:C6	0.54	2.36	16	2
1:A:68:PHE:CE2	2:B:3:U:C6	0.53	2.96	8	13
1:A:68:PHE:CZ	2:B:3:U:C6	0.52	2.98	8	5
1:A:147:LYS:CD	2:B:1:C:H2'	0.52	2.34	13	2
1:A:147:LYS:HE3	2:B:1:C:H2'	0.52	1.80	13	2
1:A:141:SER:N	1:A:142:PRO:CD	0.52	2.72	20	2
1:A:68:PHE:CZ	1:A:70:GLY:CA	0.52	2.92	4	5
1:A:156:LEU:CD1	1:A:156:LEU:N	0.51	2.73	16	1
1:A:110:TYR:CE1	2:B:2:U:H4'	0.51	2.40	10	2
1:A:112:PHE:CE1	2:B:4:U:C5	0.51	2.98	9	5
1:A:238:GLN:C	1:A:239:LEU:HD22	0.50	2.27	1	7
1:A:147:LYS:CE	2:B:1:C:H2'	0.50	2.36	13	1
1:A:143:THR:C	1:A:145:LYS:H	0.50	2.10	7	1
1:A:176:PHE:CZ	2:B:5:A:N7	0.50	2.80	20	2
1:A:213:TYR:CB	1:A:214:PRO:CD	0.50	2.90	4	1
2:B:5:A:H5"	2:B:5:A:C8	0.49	2.43	19	6
1:A:147:LYS:HE3	2:B:1:C:C6	0.48	2.43	14	1
1:A:65:ARG:HA	1:A:115:PHE:CE1	0.48	2.42	16	1
1:A:182:ARG:HB3	1:A:183:PRO:CD	0.48	2.38	19	1
1:A:124:LEU:HD22	1:A:124:LEU:C	0.48	2.29	3	2
1:A:184:LEU:HD12	1:A:185:ARG:H	0.48	1.69	3	2
1:A:182:ARG:H	1:A:183:PRO:HD2	0.48	1.67	6	2
1:A:156:LEU:HD12	1:A:156:LEU:N	0.48	2.22	18	1
1:A:182:ARG:N	1:A:183:PRO:CD	0.48	2.76	9	4
1:A:68:PHE:CZ	2:B:3:U:O4'	0.48	2.67	4	2
1:A:72:LEU:H	1:A:72:LEU:HD23	0.48	1.68	17	1
1:A:68:PHE:CE2	2:B:3:U:C5	0.48	3.02	2	3
1:A:213:TYR:HB3	1:A:214:PRO:CD	0.48	2.39	4	4
1:A:176:PHE:CD1	2:B:5:A:N6	0.47	2.82	15	1
1:A:112:PHE:CZ	2:B:4:U:C6	0.47	3.02	15	2
1:A:85:PHE:CE1	1:A:137:LEU:HD23	0.47	2.44	9	1
1:A:147:LYS:CE	2:B:1:C:C6	0.47	2.97	17	2

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:141:SER:HB3	1:A:142:PRO:CD	0.47	2.40	16	1
1:A:144:ILE:N	1:A:144:ILE:HD12	0.47	2.25	1	1
1:A:68:PHE:CE1	2:B:3:U:O4'	0.47	2.67	4	3
1:A:91:LEU:C	1:A:91:LEU:HD22	0.46	2.31	8	5
1:A:156:LEU:N	1:A:156:LEU:CD1	0.46	2.79	18	2
1:A:66:LYS:HE2	1:A:154:TRP:CD1	0.45	2.45	5	1
1:A:215:LYS:HA	2:B:4:U:C4	0.45	2.46	9	3
1:A:144:ILE:CD1	2:B:1:C:C5	0.45	2.99	11	1
1:A:176:PHE:CD1	1:A:253:TYR:CE1	0.45	3.05	6	3
2:B:5:A:C8	2:B:5:A:H5"	0.45	2.47	3	2
1:A:144:ILE:HG22	1:A:145:LYS:H	0.45	1.71	15	3
2:B:2:U:C6	2:B:2:U:P	0.45	3.10	11	1
1:A:141:SER:HB3	1:A:142:PRO:HD2	0.45	1.88	16	1
1:A:129:ILE:HD12	1:A:129:ILE:N	0.44	2.28	6	3
1:A:147:LYS:HE3	2:B:1:C:C2'	0.44	2.43	13	1
1:A:112:PHE:CE1	2:B:4:U:C4	0.44	3.05	9	4
1:A:62:ARG:HD3	1:A:156:LEU:HD13	0.44	1.90	16	1
1:A:98:LYS:HA	1:A:98:LYS:HE2	0.44	1.89	4	2
1:A:110:TYR:CD2	2:B:3:U:H4'	0.44	2.48	8	5
1:A:65:ARG:H	1:A:65:ARG:CD	0.44	2.26	17	1
1:A:91:LEU:HD22	1:A:91:LEU:C	0.44	2.33	15	3
1:A:68:PHE:CZ	1:A:70:GLY:HA3	0.44	2.47	4	2
1:A:184:LEU:HD23	1:A:184:LEU:H	0.43	1.72	6	1
1:A:65:ARG:HA	1:A:115:PHE:CE2	0.43	2.49	12	2
1:A:91:LEU:HD23	1:A:91:LEU:N	0.43	2.29	5	2
1:A:115:PHE:CZ	1:A:121:VAL:HG11	0.43	2.49	5	1
1:A:91:LEU:CD2	1:A:91:LEU:N	0.42	2.82	5	6
1:A:213:TYR:CB	1:A:214:PRO:HD2	0.42	2.45	4	1
1:A:87:ARG:HG3	1:A:88:PHE:CD2	0.42	2.50	9	2
1:A:91:LEU:N	1:A:91:LEU:HD23	0.42	2.30	3	1
1:A:92:ILE:HG22	1:A:93:VAL:N	0.42	2.30	12	1
1:A:147:LYS:HD2	2:B:1:C:C6	0.42	2.50	18	1
1:A:136:TYR:HA	1:A:149:VAL:O	0.42	2.15	5	1
1:A:147:LYS:HD3	2:B:1:C:H2'	0.41	1.91	13	1
1:A:136:TYR:CD2	1:A:148:PRO:HB2	0.41	2.50	5	1
2:B:5:A:C5'	2:B:5:A:C8	0.41	3.03	10	2
1:A:207:THR:H	1:A:214:PRO:HA	0.41	1.75	13	1
1:A:136:TYR:CD2	1:A:148:PRO:CB	0.41	3.04	5	1
1:A:141:SER:CB	1:A:142:PRO:CD	0.41	2.98	16	1
1:A:96:PRO:HG3	2:B:4:U:C1'	0.41	2.45	8	3
1:A:216:GLY:N	2:B:4:U:C5	0.41	2.88	2	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:91:LEU:N	1:A:91:LEU:CD2	0.41	2.83	14	1
1:A:182:ARG:HB3	1:A:183:PRO:HD3	0.41	1.93	19	1
2:B:5:A:C8	2:B:5:A:C5'	0.41	3.04	3	1
2:B:2:U:O2	2:B:2:U:H2'	0.41	2.16	18	1
1:A:129:ILE:N	1:A:129:ILE:HD12	0.40	2.31	8	1
1:A:184:LEU:CD2	1:A:184:LEU:H	0.40	2.30	6	1
2:B:3:U:C6	2:B:3:U:C5'	0.40	3.04	10	1
1:A:176:PHE:CG	2:B:5:A:C6	0.40	3.10	15	1
1:A:136:TYR:CE1	1:A:148:PRO:HB2	0.40	2.51	14	1
1:A:85:PHE:CZ	1:A:151:ILE:HD11	0.40	2.51	14	1
1:A:76:ILE:HD12	1:A:76:ILE:C	0.40	2.37	15	1
1:A:63:TYR:HA	1:A:156:LEU:HD13	0.40	1.94	18	1
1:A:182:ARG:CB	1:A:183:PRO:CD	0.40	2.99	19	1
1:A:205:ILE:HD12	1:A:205:ILE:N	0.40	2.32	15	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	184/203 (91%)	163±4 (89±2%)	15±3 (8±2%)	6±2 (3±1%)	7	40
All	All	3680/4060 (91%)	3260 (89%)	309 (8%)	111 (3%)	7	40

All 37 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	185	ARG	15
1	A	133	GLY	10
1	A	182	ARG	9
1	A	144	ILE	8
1	A	109	GLY	7
1	A	132	ASP	5
1	A	206	ASP	5
1	A	75	ASP	4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	213	TYR	4
1	A	254	VAL	3
1	A	146	ASP	3
1	A	184	LEU	3
1	A	74	PRO	3
1	A	164	ASP	2
1	A	157	SER	2
1	A	76	ILE	2
1	A	77	ASP	2
1	A	183	PRO	2
1	A	217	ALA	2
1	A	110	TYR	2
1	A	99	ALA	2
1	A	181	PRO	1
1	A	145	LYS	1
1	A	66	LYS	1
1	A	239	LEU	1
1	A	240	GLN	1
1	A	212	LYS	1
1	A	63	TYR	1
1	A	211	LEU	1
1	A	71	GLY	1
1	A	142	PRO	1
1	A	165	GLY	1
1	A	141	SER	1
1	A	235	ARG	1
1	A	78	GLU	1
1	A	62	ARG	1
1	A	72	LEU	1

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	157/175 (90%)	146±2 (93±1%)	11±2 (7±1%)	19 68
All	All	3140/3500 (90%)	2924 (93%)	216 (7%)	19 68

All 46 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	247	ARG	19
1	A	251	LYS	18
1	A	91	LEU	17
1	A	110	TYR	16
1	A	62	ARG	12
1	A	147	LYS	8
1	A	136	TYR	7
1	A	250	VAL	7
1	A	144	ILE	6
1	A	213	TYR	6
1	A	98	LYS	6
1	A	184	LEU	6
1	A	65	ARG	6
1	A	131	GLU	5
1	A	112	PHE	5
1	A	195	ARG	4
1	A	124	LEU	4
1	A	182	ARG	4
1	A	156	LEU	3
1	A	215	LYS	3
1	A	211	LEU	3
1	A	130	GLU	3
1	A	175	ILE	3
1	A	129	ILE	3
1	A	219	ARG	3
1	A	63	TYR	3
1	A	76	ILE	3
1	A	86	ARG	3
1	A	152	ARG	3
1	A	92	ILE	3
1	A	150	GLN	3
1	A	172	ARG	3
1	A	254	VAL	2
1	A	253	TYR	2
1	A	114	LEU	2
1	A	238	GLN	2
1	A	113	LEU	1
1	A	240	GLN	1
1	A	185	ARG	1
1	A	243	GLU	1
1	A	72	LEU	1
1	A	241	HIS	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	225	GLN	1
1	A	134	LYS	1
1	A	169	LEU	1
1	A	237	VAL	1

5.2.3 RNA

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers	Suiteness
2	B	4/5 (80%)	2±1 (55±20%)	0±1 (10±15%)	0.10±0.08
All	All	81/100 (81%)	44 (54%)	8 (10%)	0.10

The overall RNA backbone suiteness is 0.10.

All unique RNA backbone outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
2	B	5	A	18
2	B	4	U	15
2	B	3	U	9
2	B	2	U	2

All unique RNA pucker outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
2	B	2	U	6
2	B	3	U	1
2	B	1	C	1

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

There are no ligands in this entry.

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 66% for the well-defined parts and 62% for the entire structure.

6.1 Chemical shift list 1

File name: input_cs.cif

Chemical shift list name: *CPEB4RRM12Complex*

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	1845
Number of shifts mapped to atoms	1845
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	5

6.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	151	3.01 ± 0.10	Should be applied
$^{13}C_{\beta}$	141	3.10 ± 0.10	Should be applied
$^{13}C'$	0	—	None (insufficient data)
^{15}N	167	0.18 ± 0.44	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 66%, i.e. 1563 atoms were assigned a chemical shift out of a possible 2370. 21 out of 30 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	608/892 (68%)	304/354 (86%)	146/368 (40%)	158/170 (93%)
Sidechain	790/1206 (66%)	483/711 (68%)	299/439 (68%)	8/56 (14%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	121/183 (66%)	78/97 (80%)	41/82 (50%)	2/4 (50%)
Overall	1563/2370 (66%)	902/1211 (74%)	493/924 (53%)	168/235 (71%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 62%, i.e. 1626 atoms were assigned a chemical shift out of a possible 2618. 21 out of 32 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	636/983 (65%)	318/390 (82%)	151/406 (37%)	167/187 (89%)
Sidechain	812/1339 (61%)	497/792 (63%)	307/484 (63%)	8/63 (13%)
Aromatic	134/207 (65%)	87/110 (79%)	45/92 (49%)	2/5 (40%)
Overall	1626/2618 (62%)	939/1341 (70%)	510/1017 (50%)	177/260 (68%)

6.1.4 Statistically unusual chemical shifts (i)

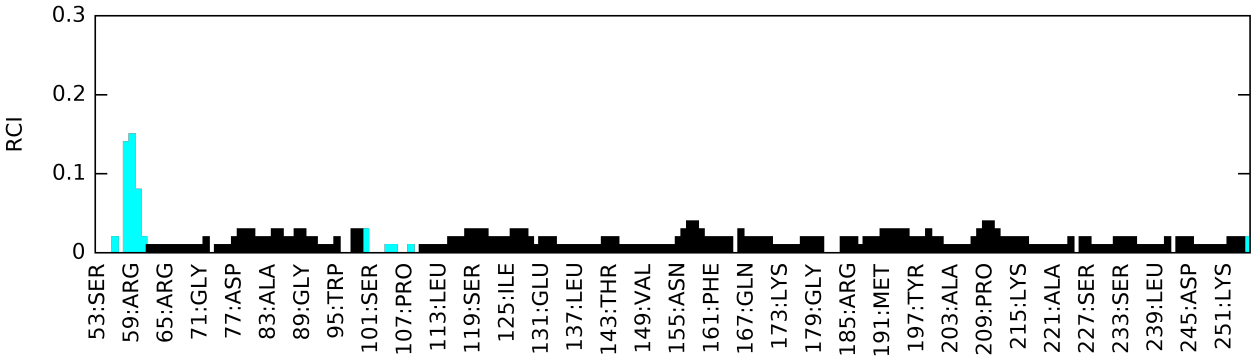
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	76	ILE	CG1	13.21	36.54 – 18.94	-8.3
1	A	214	PRO	HG3	-0.24	3.56 – 0.26	-6.5
1	A	214	PRO	HG2	0.25	3.48 – 0.38	-5.4
1	A	87	ARG	HG3	0.01	3.00 – 0.10	-5.3
1	A	173	LYS	CE	37.57	46.00 – 37.80	-5.3

6.1.5 Random Coil Index (RCI) plots (i)

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
0.3
0.2
0.1
0
53:SER
59:ARG
65:ARG
71:GLY
77:ASP
83:ALA
89:GLY
95:TRP
101:SER
107:PRO
113:LEU
119:SER
125:ILE
131:GLU
137:LEU
143:THR
149:VAL
155:ASN
161:PHE
167:GLN
173:LYS
179:GLY
185:ARG
191:MET
197:TYR
203:ALA
209:PRO
215:LYS
221:ALA
227:SER
233:SER
239:LEU
245:ASP
251:LYS