W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

Full wwPDB NMR Structure Validation Report (i)

May 28, 2020 – 11:26 pm BST

PDB ID	:	2LWW
Title	:	NMR structure of RelA-TAD/CBP-TAZ1 complex
Authors	:	Mukherjee, S.P.; Ghosh, G.; Wright, P.E.
Deposited on	:	2012-08-07

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

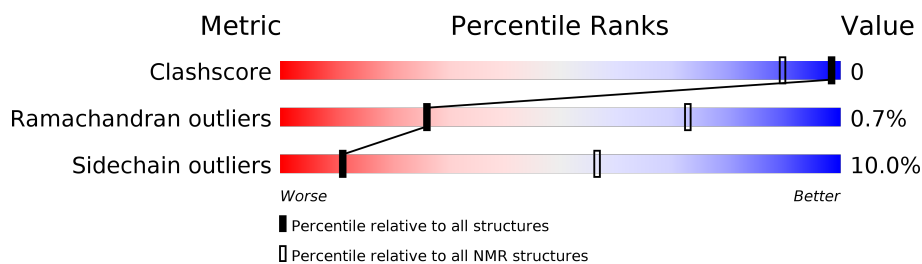|		
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.11
Ideal geometry (proteins)	:	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 81%.

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	Percentile Ranks	Value
Clashscore		0
Ramachandran outliers		0.7%
Sidechain outliers		10.0%

Worse   *Better*

■ Percentile relative to all structures
▯ Percentile relative to all NMR structures

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	100	81% 9% 10%
2	B	70	53% 10% 37%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 4 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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:347-A:436, B:434-B:453, B:457-B:480 (134)	0.53	4

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

3 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2556 atoms, of which 1261 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called CREB-binding protein.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	100	1568	472	788	162	135	11	0

- Molecule 2 is a protein called Nuclear transcription factor RelA.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	70	985	311	473	85	114	2	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	421	GLY	-	EXPRESSION TAG	UNP Q548Y4
B	422	SER	-	EXPRESSION TAG	UNP Q548Y4
B	423	HIS	-	EXPRESSION TAG	UNP Q548Y4
B	424	MET	-	EXPRESSION TAG	UNP Q548Y4

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	
			Total	Zn
3	A	3	3	3

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: CREB-binding protein
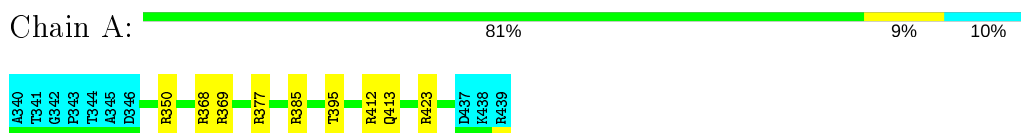
- Molecule 2: Nuclear transcription factor RelA
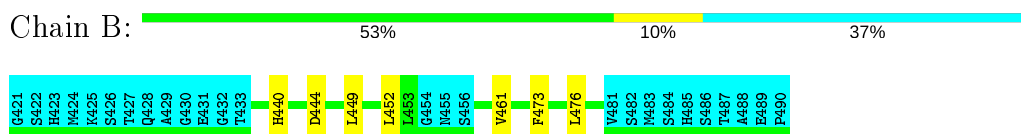
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

- Molecule 1: CREB-binding protein
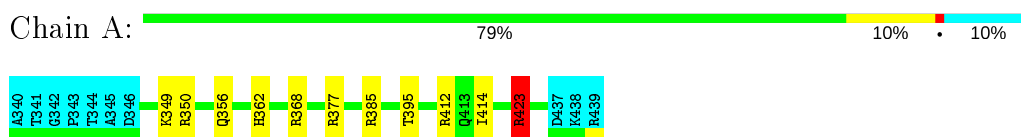
- Molecule 2: Nuclear transcription factor RelA
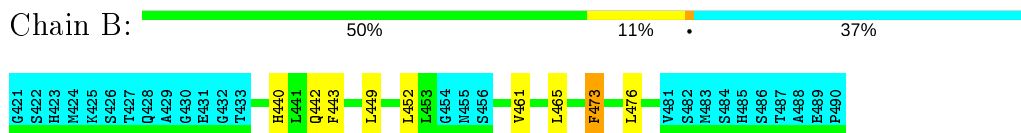
4.2.2 Score per residue for model 2
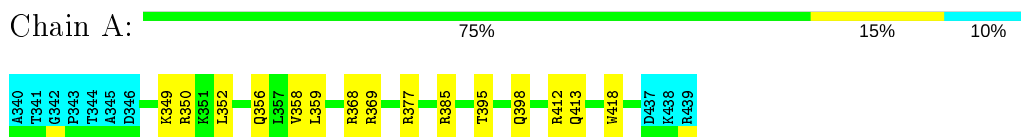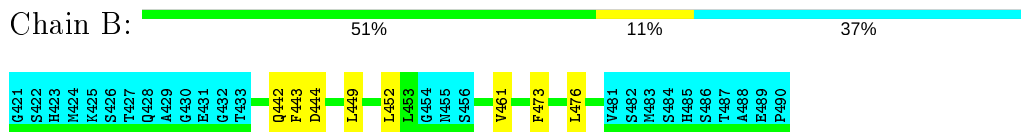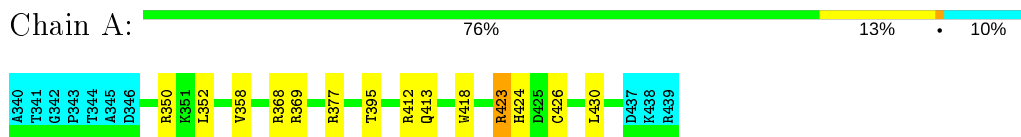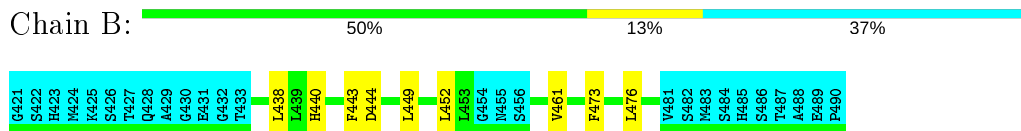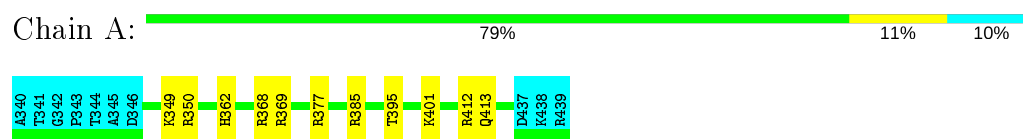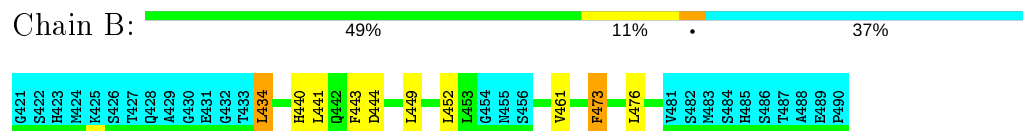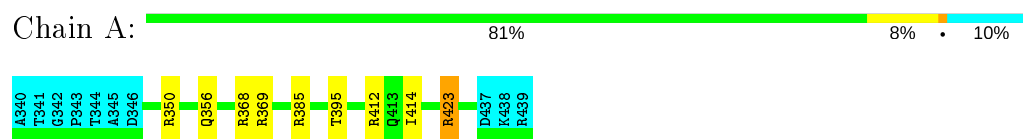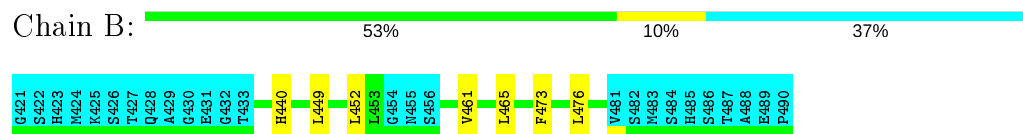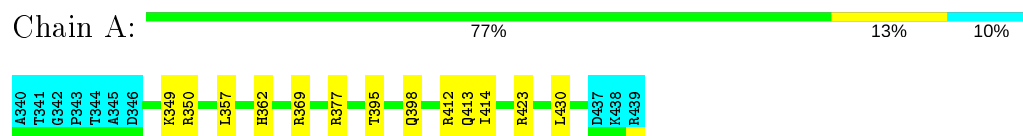
• Molecule 1: CREB-binding protein

Chain A:

• Molecule 2: Nuclear transcription factor RelA

Chain B:
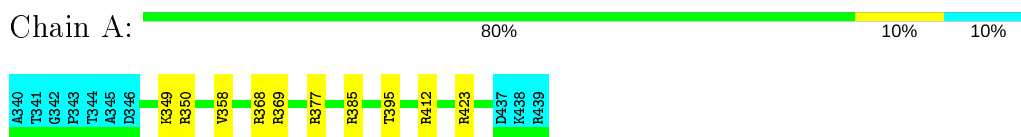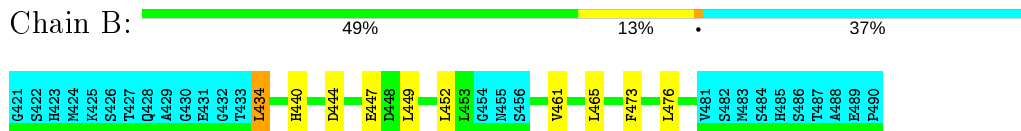
4.2.3 Score per residue for model 3

• Molecule 1: CREB-binding protein

Chain A:

• Molecule 2: Nuclear transcription factor RelA

Chain B:

4.2.4 Score per residue for model 4 (medoid)

• Molecule 1: CREB-binding protein

Chain A:

• Molecule 2: Nuclear transcription factor RelA

Chain B:

4.2.5 Score per residue for model 5

• Molecule 1: CREB-binding protein

Chain A:

• Molecule 2: Nuclear transcription factor RelA

Chain B:

4.2.6 Score per residue for model 6

• Molecule 1: CREB-binding protein

Chain A:

• Molecule 2: Nuclear transcription factor RelA
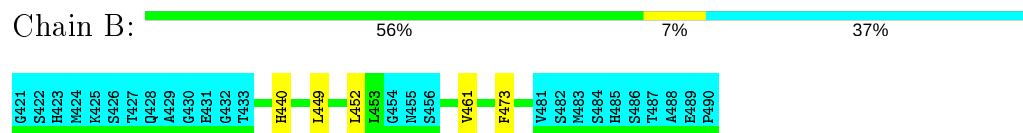
Chain B:

4.2.7 Score per residue for model 7

• Molecule 1: CREB-binding protein

Chain A:

• Molecule 2: Nuclear transcription factor RelA

Chain B:

4.2.8 Score per residue for model 8

● Molecule 1: CREB-binding protein

Chain A:

● Molecule 2: Nuclear transcription factor RelA

Chain B:
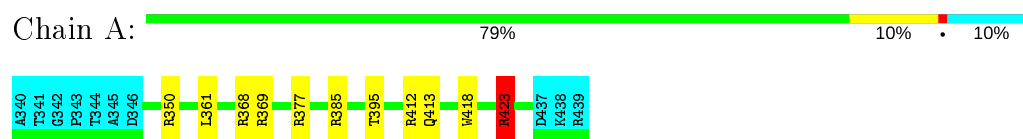
4.2.9 Score per residue for model 9

● Molecule 1: CREB-binding protein

Chain A:

● Molecule 2: Nuclear transcription factor RelA

Chain B:

4.2.10 Score per residue for model 10

● Molecule 1: CREB-binding protein

Chain A:

● Molecule 2: Nuclear transcription factor RelA
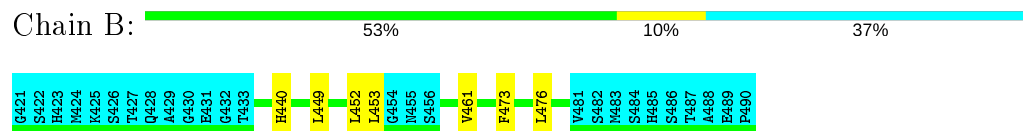
Chain B:

4.2.11 Score per residue for model 11
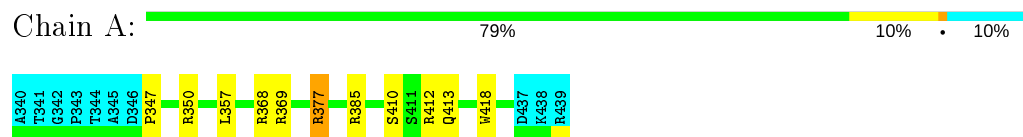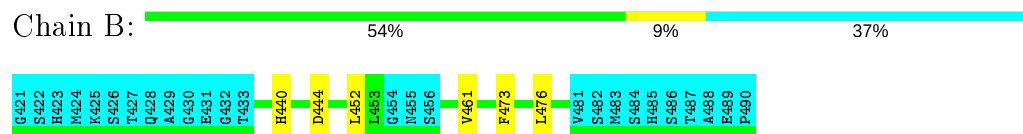
• Molecule 1: CREB-binding protein

• Molecule 2: Nuclear transcription factor RelA

4.2.12 Score per residue for model 12

• Molecule 1: CREB-binding protein

• Molecule 2: Nuclear transcription factor RelA

4.2.13 Score per residue for model 13

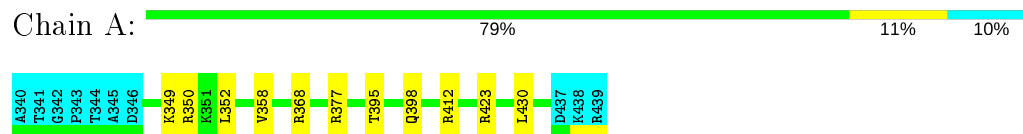• Molecule 1: CREB-binding protein
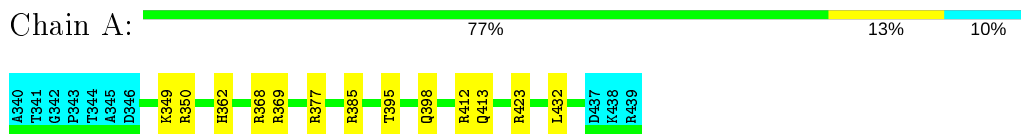
• Molecule 2: Nuclear transcription factor RelA
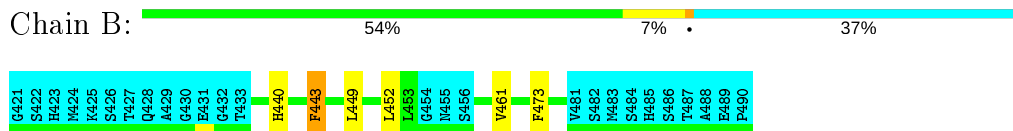
4.2.14 Score per residue for model 14

• Molecule 1: CREB-binding protein
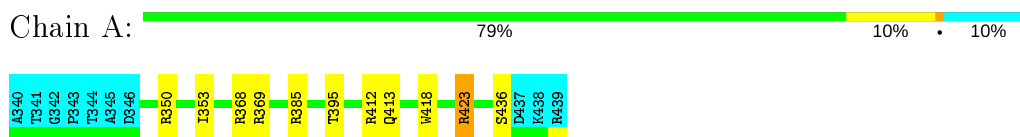
• Molecule 2: Nuclear transcription factor RelA
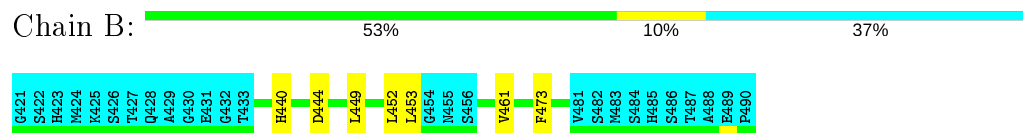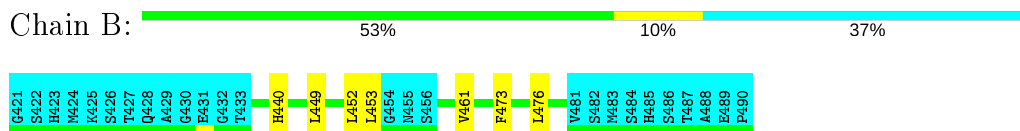

4.2.15 Score per residue for model 15

• Molecule 1: CREB-binding protein
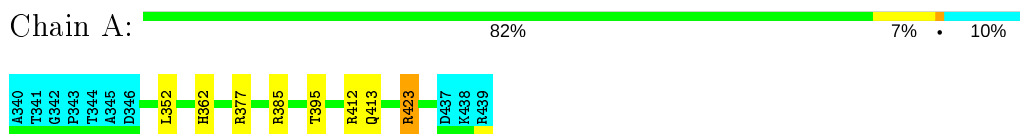
• Molecule 2: Nuclear transcription factor RelA
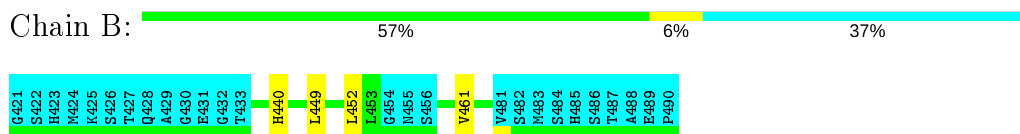
4.2.16 Score per residue for model 16

• Molecule 1: CREB-binding protein
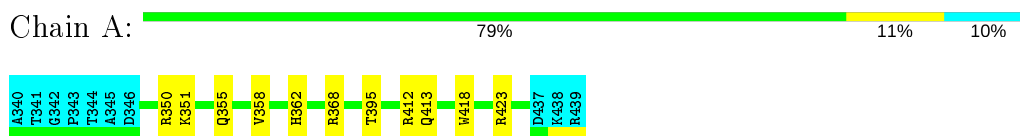
• Molecule 2: Nuclear transcription factor RelA
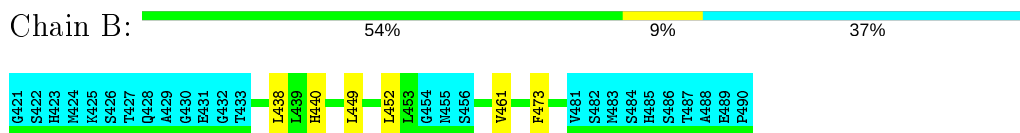
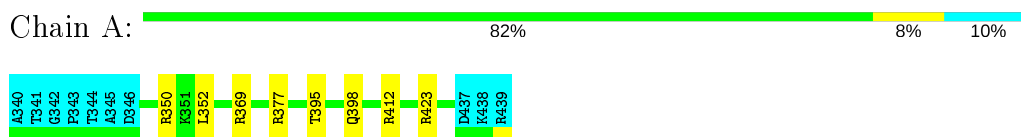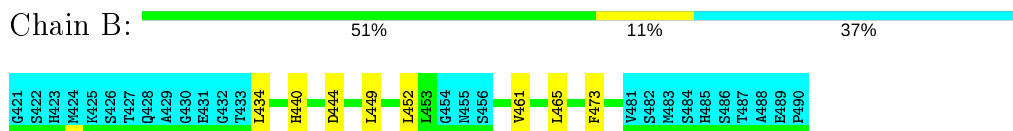### 4.2.17 Score per residue for model 17
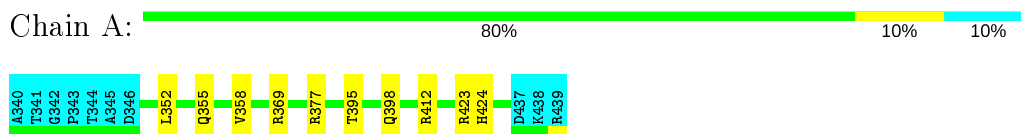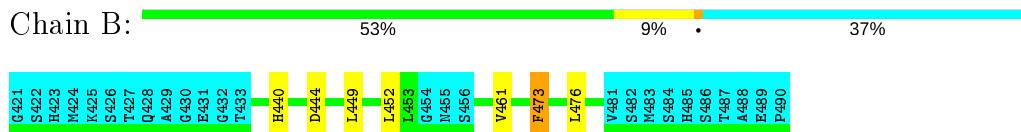
- Molecule 1: CREB-binding protein

Chain A:

- Molecule 2: Nuclear transcription factor RelA

Chain B:
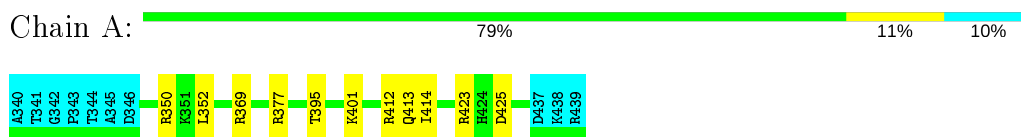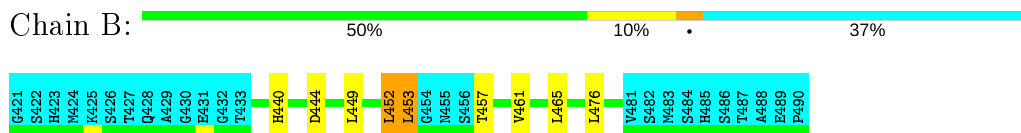
4.2.18 Score per residue for model 18

- Molecule 1: CREB-binding protein

Chain A:

- Molecule 2: Nuclear transcription factor RelA

Chain B:
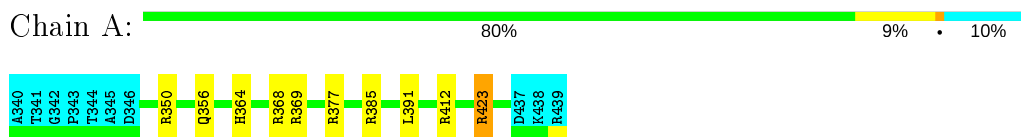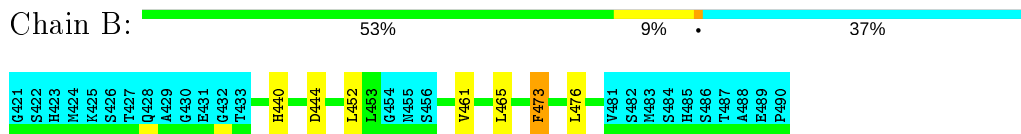
4.2.19 Score per residue for model 19

- Molecule 1: CREB-binding protein

Chain A:

- Molecule 2: Nuclear transcription factor RelA

Chain B:
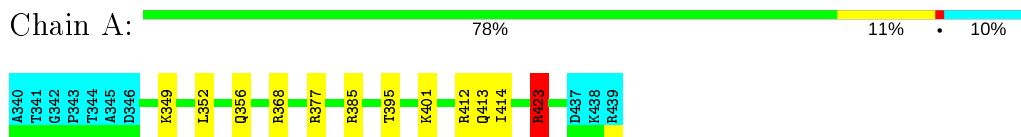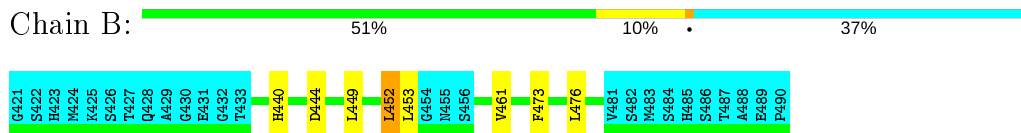
4.2.20 Score per residue for model 20

• Molecule 1: CREB-binding protein

Chain A:

• Molecule 2: Nuclear transcription factor RelA

Chain B:
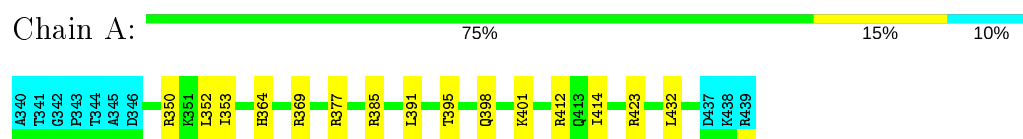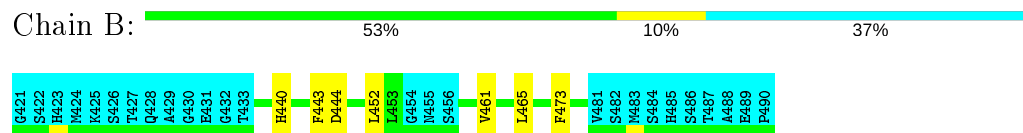
5 Refinement protocol and experimental data overview (i)

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

Of the 200 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
CYANA	structure solution	2.1
AMBER	refinement	10

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	1822
Number of shifts mapped to atoms	1822
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	81%

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

COVALENT-GEOMETRY INFOmissingINFO

5.1 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	709	720	715	0±1
2	B	335	313	313	0±1
All	All	20940	20660	20560	9

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
1:A:432:LEU:HD22	2:B:443:PHE:CZ	0.52	2.39	13	2
2:B:434:LEU:O	2:B:434:LEU:HD12	0.45	2.10	5	1
2:B:452:LEU:HD13	2:B:453:LEU:N	0.45	2.27	17	2
1:A:364:HIS:CE1	1:A:391:LEU:HD22	0.43	2.48	18	2
2:B:434:LEU:HD12	2:B:434:LEU:C	0.42	2.35	5	1
2:B:434:LEU:C	2:B:434:LEU:HD22	0.40	2.36	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	90/100 (90%)	88±1 (97±1%)	2±1 (2±1%)	0±0 (0±1%)	38	78
2	B	44/70 (63%)	41±1 (94±3%)	2±1 (5±3%)	1±0 (1±1%)	15	61
All	All	2680/3400 (79%)	2579 (96%)	82 (3%)	19 (1%)	26	73

All 5 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
2	B	444	ASP	12
1	A	423	ARG	4
1	A	436	SER	1
1	A	347	PRO	1
1	A	425	ASP	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	80/87 (92%)	75±2 (93±2%)	5±2 (7±2%)	20	68

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	37/57 (65%)	31±1 (83±3%)	6±1 (17±3%)	5	39
All	All	2340/2880 (81%)	2105 (90%)	235 (10%)	11	56

All 37 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)
2	B	452	LEU	20
2	B	461	VAL	20
2	B	440	HIS	19
2	B	473	PHE	18
1	A	395	THR	18
2	B	449	LEU	17
2	B	476	LEU	13
1	A	413	GLN	12
1	A	352	LEU	9
1	A	423	ARG	8
1	A	349	LYS	8
1	A	398	GLN	7
2	B	465	LEU	7
1	A	358	VAL	6
1	A	418	TRP	6
1	A	414	ILE	6
1	A	356	GLN	5
2	B	453	LEU	4
1	A	401	LYS	4
1	A	430	LEU	3
2	B	434	LEU	3
2	B	442	GLN	2
2	B	438	LEU	2
1	A	355	GLN	2
1	A	424	HIS	2
1	A	357	LEU	2
1	A	353	ILE	2
1	A	426	CYS	1
1	A	359	LEU	1
1	A	350	ARG	1
1	A	351	LYS	1
1	A	410	SER	1
2	B	447	GLU	1
2	B	457	THR	1
2	B	441	LEU	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	377	ARG	1
1	A	361	LEU	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 3 ligands modelled in this entry, 3 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 81% for the well-defined parts and 80% 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	1822
Number of shifts mapped to atoms	1822
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

6.1.2 Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_\alpha$	169	-0.22 ± 0.15	None needed (< 0.5 ppm)
$^{13}C_\beta$	158	0.46 ± 0.11	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	156	-0.55 ± 0.22	Should be applied

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 81%, i.e. 1338 atoms were assigned a chemical shift out of a possible 1658. 26 out of 26 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	516/660 (78%)	258/263 (98%)	134/268 (50%)	124/129 (96%)
Sidechain	766/891 (86%)	474/524 (90%)	275/320 (86%)	17/47 (36%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	56/107 (52%)	30/57 (53%)	25/35 (71%)	1/15 (7%)
Overall	1338/1658 (81%)	762/844 (90%)	434/623 (70%)	142/191 (74%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 80%, i.e. 1614 atoms were assigned a chemical shift out of a possible 2023. 27 out of 27 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

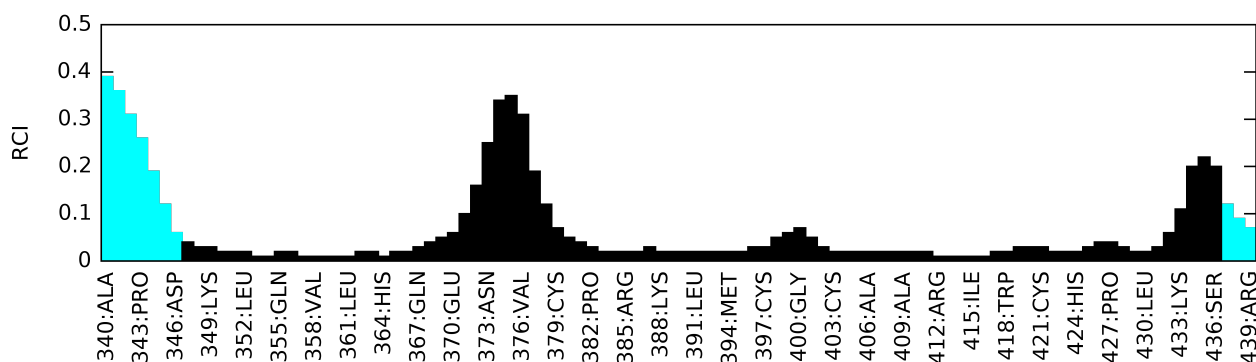|	Total	^{1}H	^{13}C	^{15}N
Backbone	650/836 (78%)	325/333 (98%)	169/340 (50%)	156/163 (96%)
Sidechain	904/1066 (85%)	563/628 (90%)	323/384 (84%)	18/54 (33%)
Aromatic	60/121 (50%)	32/65 (49%)	27/39 (69%)	1/17 (6%)
Overall	1614/2023 (80%)	920/1026 (90%)	519/763 (68%)	175/234 (75%)

6.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

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: