



Full wwPDB NMR Structure Validation Report ⓘ

May 29, 2024 – 10:15 AM EDT

PDB ID : 1MBH
Title : MOUSE C-MYB DNA-BINDING DOMAIN REPEAT 2
Authors : Ogata, K.; Morikawa, S.; Nakamura, H.; Hojo, H.; Yoshimura, S.; Zhang, R.;
Aimoto, S.; Ametani, Y.; Hirata, Z.; Sarai, A.; Ishii, S.; Nishimura, Y.
Deposited on : 1995-05-19

This is a Full wwPDB NMR Structure Validation Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/NMRValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI : v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV : Wang et al. (2010)
wwPDB-ShiftChecker : v1.2
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36.2

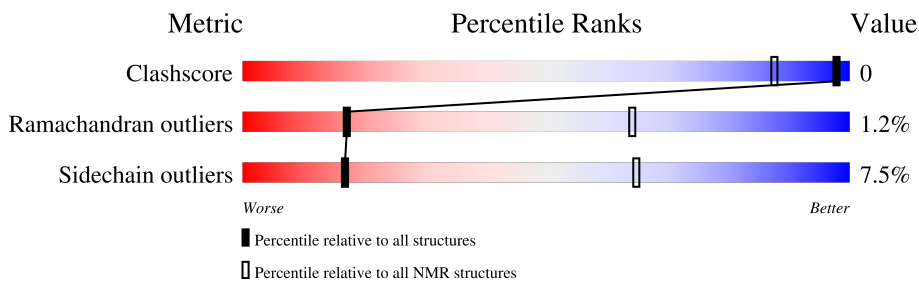
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR


The overall completeness of chemical shifts assignment was not calculated.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	53	 79% 17%

2 Ensemble composition and analysis i

This entry contains 50 models. Model 33 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:95-A:138 (44)	0.49	33

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 8 clusters and 3 single-model clusters were found.

Cluster number	Models
1	1, 2, 5, 17, 19, 21, 25, 26, 30, 34, 36, 41, 44, 45
2	6, 7, 8, 13, 23, 27, 29, 31, 32, 35, 37, 40, 42
3	3, 4, 11, 15, 18, 24
4	28, 38, 46, 49
5	10, 39, 48
6	33, 47, 50
7	16, 22
8	14, 20
Single-model clusters	9; 12; 43

3 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 905 atoms, of which 460 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called C-MYB.

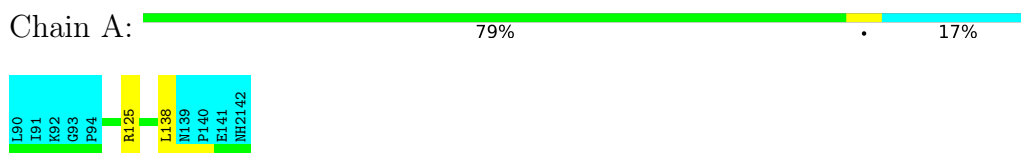
Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	53	905	283	460	89	72	1	1

4 Residue-property plots

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide 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: C-MYB

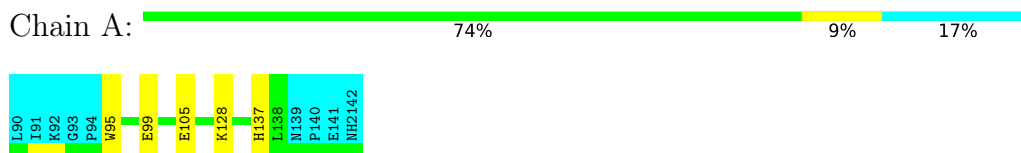


4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

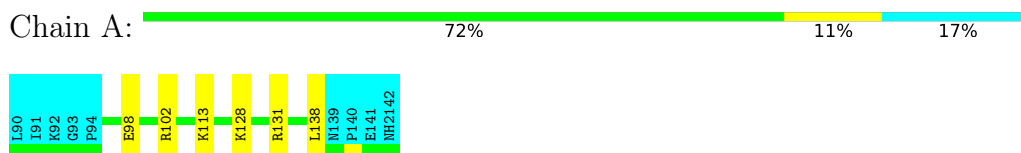
4.2.1 Score per residue for model 1

- Molecule 1: C-MYB



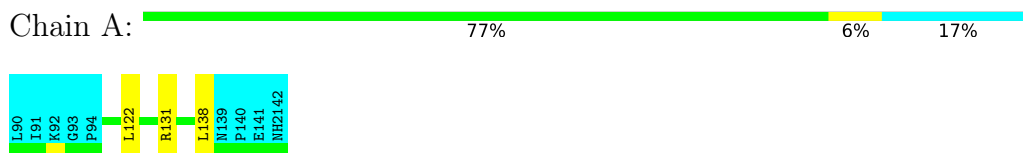
4.2.2 Score per residue for model 2

- Molecule 1: C-MYB



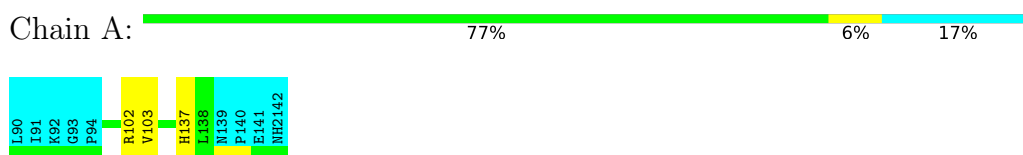
4.2.3 Score per residue for model 3

- Molecule 1: C-MYB



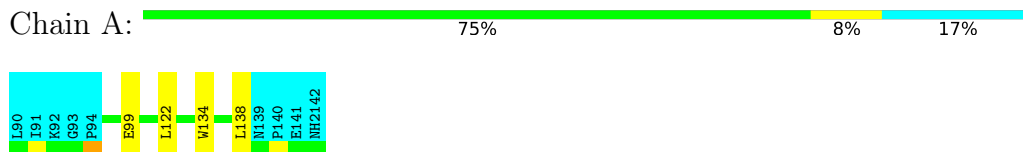
4.2.4 Score per residue for model 4

- Molecule 1: C-MYB



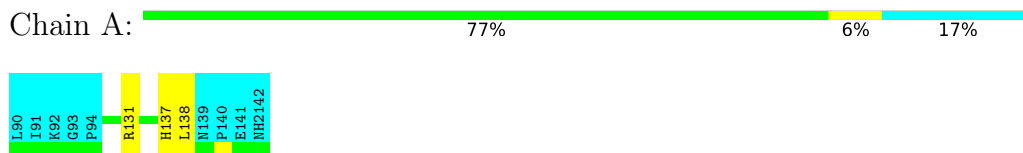
4.2.5 Score per residue for model 5

- Molecule 1: C-MYB



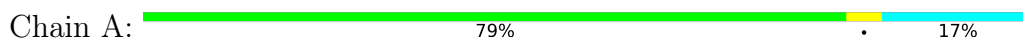
4.2.6 Score per residue for model 6

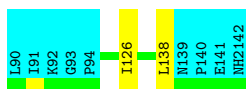
- Molecule 1: C-MYB



4.2.7 Score per residue for model 7

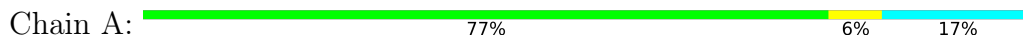
- Molecule 1: C-MYB





4.2.8 Score per residue for model 8

- Molecule 1: C-MYB



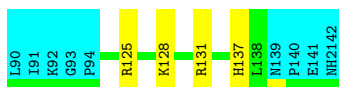
4.2.9 Score per residue for model 9

- Molecule 1: C-MYB



4.2.10 Score per residue for model 10

- Molecule 1: C-MYB



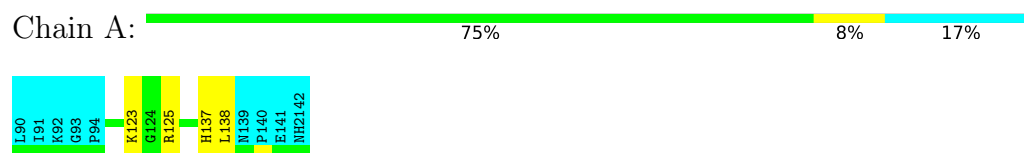
4.2.11 Score per residue for model 11

- Molecule 1: C-MYB



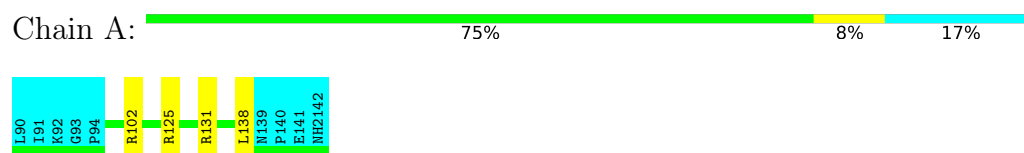
4.2.12 Score per residue for model 12

- Molecule 1: C-MYB



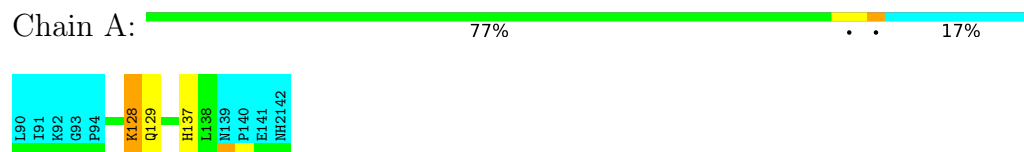
4.2.13 Score per residue for model 13

- Molecule 1: C-MYB



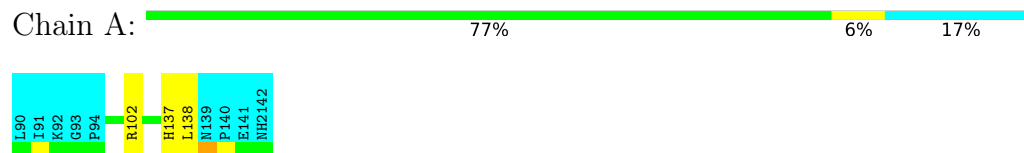
4.2.14 Score per residue for model 14

- Molecule 1: C-MYB



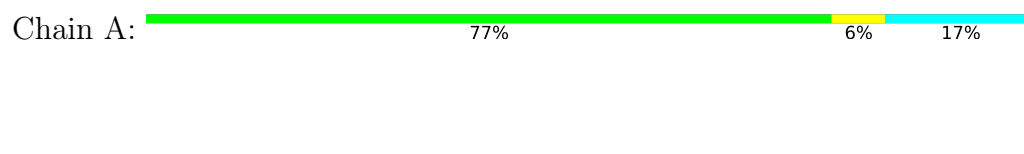
4.2.15 Score per residue for model 15

- Molecule 1: C-MYB



4.2.16 Score per residue for model 16

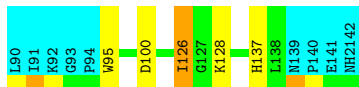
- Molecule 1: C-MYB





4.2.17 Score per residue for model 17

- Molecule 1: C-MYB



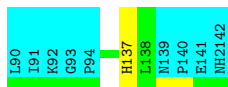
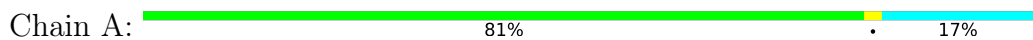
4.2.18 Score per residue for model 18

- Molecule 1: C-MYB



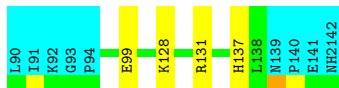
4.2.19 Score per residue for model 19

- Molecule 1: C-MYB



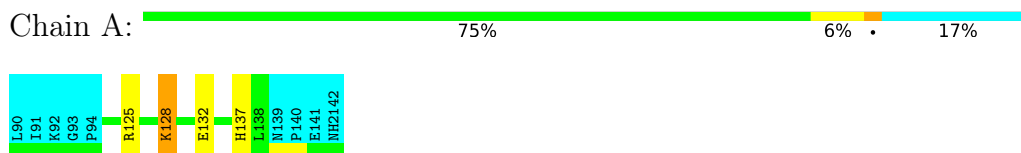
4.2.20 Score per residue for model 20

- Molecule 1: C-MYB



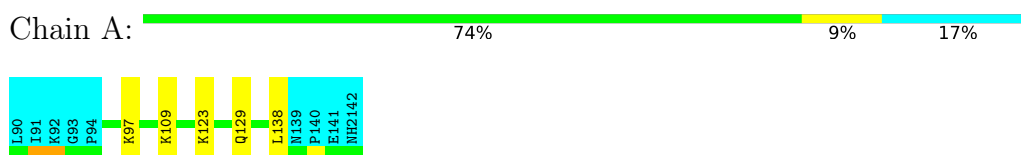
4.2.21 Score per residue for model 21

- Molecule 1: C-MYB



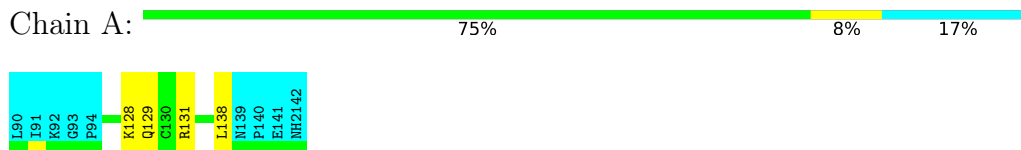
4.2.22 Score per residue for model 22

- Molecule 1: C-MYB



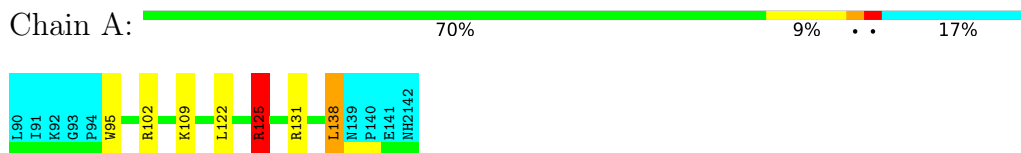
4.2.23 Score per residue for model 23

- Molecule 1: C-MYB



4.2.24 Score per residue for model 24

- Molecule 1: C-MYB



4.2.25 Score per residue for model 25

- Molecule 1: C-MYB

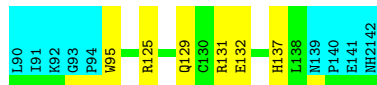




4.2.26 Score per residue for model 26

- Molecule 1: C-MYB

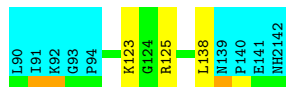
Chain A: 72% 11% 17%



4.2.27 Score per residue for model 27

- Molecule 1: C-MYB

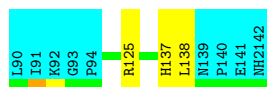
Chain A: 77% 6% 17%



4.2.28 Score per residue for model 28

- Molecule 1: C-MYB

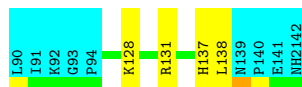
Chain A: 77% 6% 17%



4.2.29 Score per residue for model 29

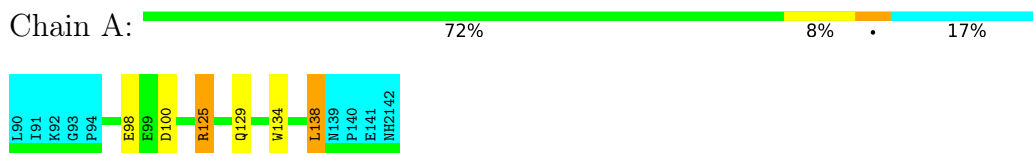
- Molecule 1: C-MYB

Chain A: 75% 8% 17%



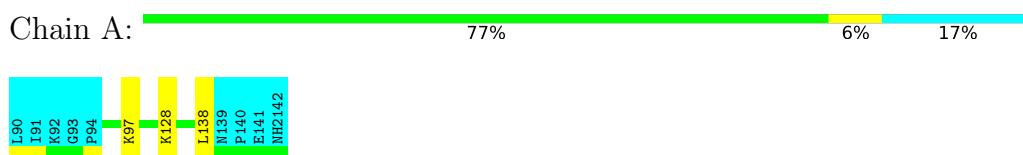
4.2.30 Score per residue for model 30

- Molecule 1: C-MYB



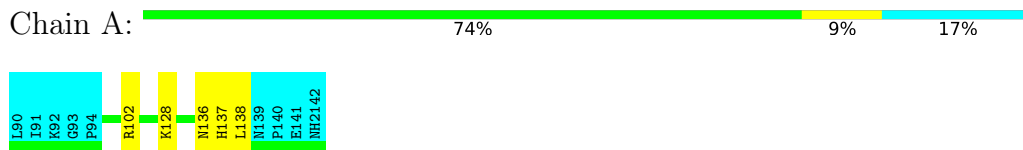
4.2.31 Score per residue for model 31

- Molecule 1: C-MYB



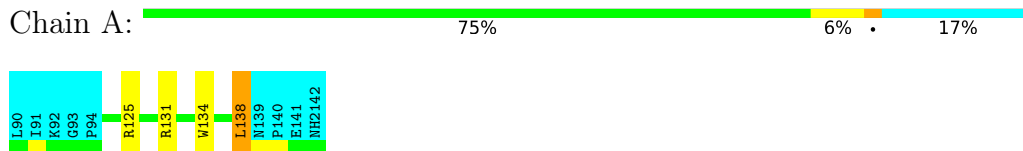
4.2.32 Score per residue for model 32

- Molecule 1: C-MYB



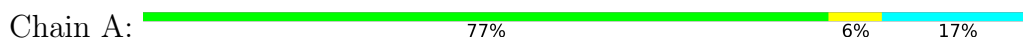
4.2.33 Score per residue for model 33 (medoid)

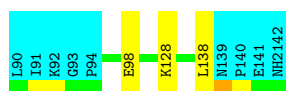
- Molecule 1: C-MYB



4.2.34 Score per residue for model 34

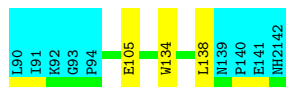
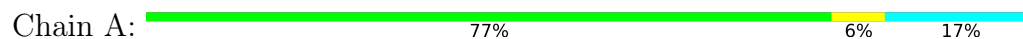
- Molecule 1: C-MYB





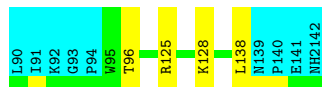
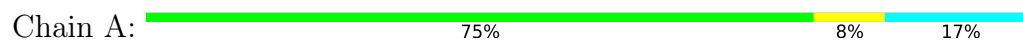
4.2.35 Score per residue for model 35

- Molecule 1: C-MYB



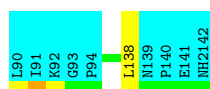
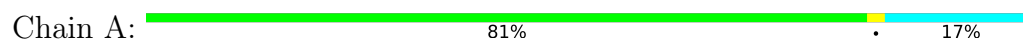
4.2.36 Score per residue for model 36

- Molecule 1: C-MYB



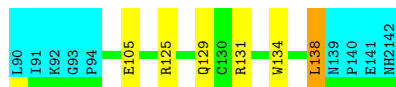
4.2.37 Score per residue for model 37

- Molecule 1: C-MYB



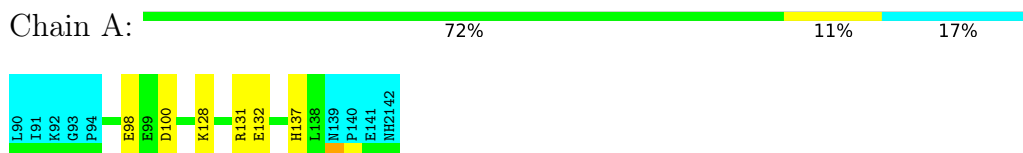
4.2.38 Score per residue for model 38

- Molecule 1: C-MYB



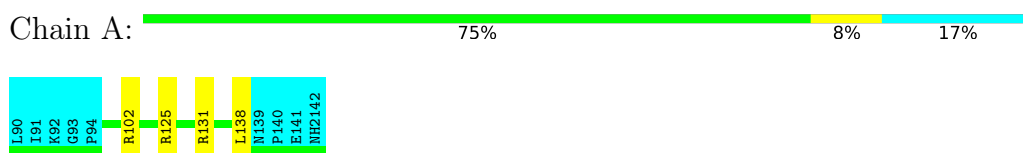
4.2.39 Score per residue for model 39

- Molecule 1: C-MYB



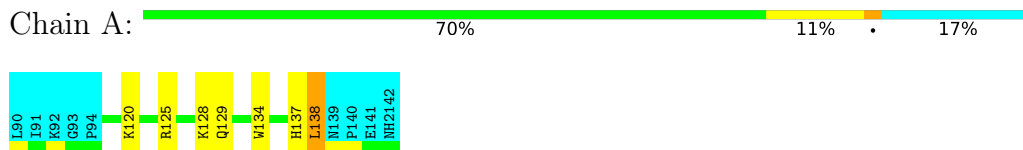
4.2.40 Score per residue for model 40

- Molecule 1: C-MYB



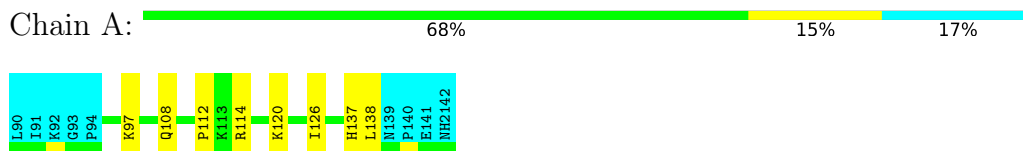
4.2.41 Score per residue for model 41

- Molecule 1: C-MYB



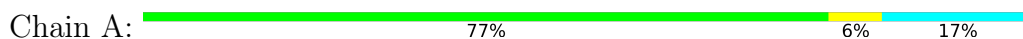
4.2.42 Score per residue for model 42

- Molecule 1: C-MYB



4.2.43 Score per residue for model 43

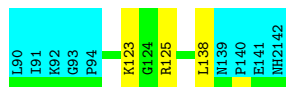
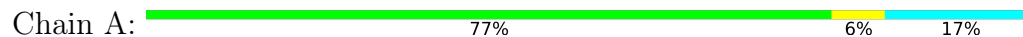
- Molecule 1: C-MYB





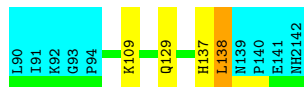
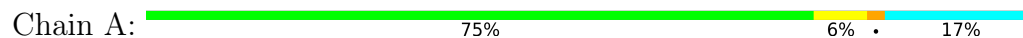
4.2.44 Score per residue for model 44

- Molecule 1: C-MYB



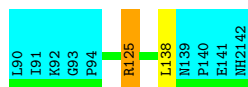
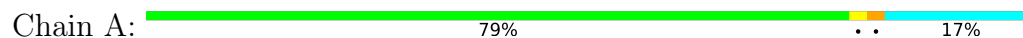
4.2.45 Score per residue for model 45

- Molecule 1: C-MYB



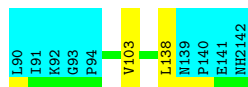
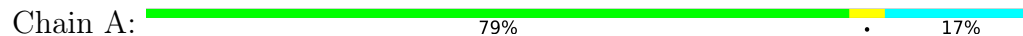
4.2.46 Score per residue for model 46

- Molecule 1: C-MYB



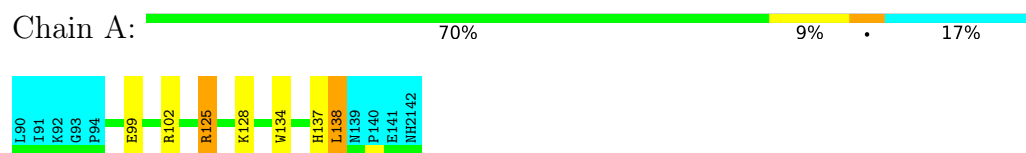
4.2.47 Score per residue for model 47

- Molecule 1: C-MYB



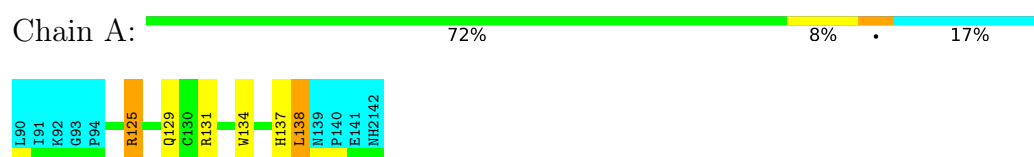
4.2.48 Score per residue for model 48

- Molecule 1: C-MYB



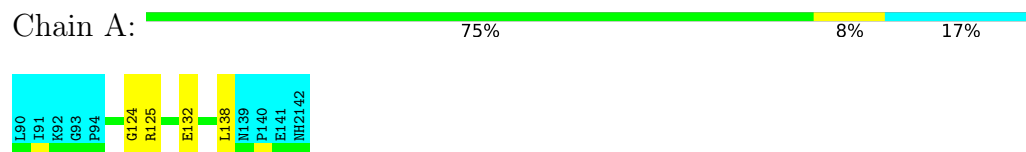
4.2.49 Score per residue for model 49

- Molecule 1: C-MYB



4.2.50 Score per residue for model 50

- Molecule 1: C-MYB



5 Refinement protocol and experimental data overview

The models were refined using the following method: ?.

Of the ? calculated structures, 50 were deposited, based on the following criterion: ?.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
EMBOSS	refinement	
PRESTO	refinement	

No chemical shift data was provided.

6 Model quality [i](#)

6.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NH2

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

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	Chirality	Planarity
1	A	0.0±0.0	0.4±0.5
All	All	0	21

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

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

Mol	Chain	Res	Type	Group	Models (Total)
1	A	125	ARG	Sidechain	19
1	A	114	ARG	Sidechain	1
1	A	102	ARG	Sidechain	1

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	384	392	392	0±1
All	All	19200	19600	19600	17

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:134:TRP:CZ3	1:A:138:LEU:HD12	0.58	2.33	5	5
1:A:134:TRP:CZ3	1:A:138:LEU:HD23	0.47	2.45	38	3
1:A:125:ARG:HH11	1:A:125:ARG:CG	0.46	2.23	49	3
1:A:114:ARG:O	1:A:118:ILE:HD12	0.45	2.11	43	1
1:A:126:ILE:HD13	1:A:126:ILE:H	0.45	1.72	17	1
1:A:138:LEU:N	1:A:138:LEU:HD23	0.44	2.27	37	1
1:A:95:TRP:CZ2	1:A:125:ARG:CD	0.41	3.04	24	1
1:A:138:LEU:N	1:A:138:LEU:CD2	0.41	2.83	24	1
1:A:125:ARG:HH11	1:A:129:GLN:NE2	0.40	2.14	9	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	44/53 (83%)	35±2 (80±5%)	8±2 (19±5%)	1±1 (1±1%)	17	64
All	All	2200/2650 (83%)	1761 (80%)	413 (19%)	26 (1%)	17	64

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)
1	A	128	LYS	14
1	A	95	TRP	4
1	A	138	LEU	4
1	A	123	LYS	3
1	A	124	GLY	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	40/47 (85%)	37±1 (92±4%)	3±1 (8±4%)	17 65
All	All	2000/2350 (85%)	1849 (92%)	151 (8%)	17 65

All 24 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	138	LEU	32
1	A	137	HIS	23
1	A	131	ARG	17
1	A	102	ARG	9
1	A	129	GLN	9
1	A	128	LYS	6
1	A	98	GLU	6
1	A	125	ARG	6
1	A	122	LEU	5
1	A	99	GLU	4
1	A	105	GLU	4
1	A	126	ILE	4
1	A	97	LYS	4
1	A	132	GLU	4
1	A	109	LYS	4
1	A	103	VAL	3
1	A	100	ASP	3
1	A	120	LYS	2
1	A	113	LYS	1
1	A	123	LYS	1
1	A	136	ASN	1
1	A	96	THR	1
1	A	108	GLN	1
1	A	112	PRO	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 monosaccharides 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

No chemical shift data were provided