Supplementary data

Cloning and comparative protein modelling of two MADS-box genes, HsMADS1 and HsMADS2 isolated from Hibiscus sabdariffa L. var. UMKL (roselle)

Siti N. Othman, Yong S.Y.C[*], Roghayeh Abedi Karjiban and Adibah Shakri

Supplementary Table 1 Summary of successfully produced models by MODELLER ver 9.13 for HsMADS1 protein

Model	DOPE score	GA341 score
HsMADS1.B99990001	-12034.94531	0.74210
HsMADS1.B99990002	-11700.74121	0.79254
HsMADS1.B99990003	-11645.07715	0.94086
HsMADS1.B99990004	-11308.71875	0.78525
HsMADS1.B99990005	-11538.41016	0.78700

**Supplementary Table 2** Summary of successfully produced models by MODELLER ver 9.13 for HsMADS2 protein

Model	DOPE score	GA341score
HsMADS2.B99990001	-8986.24805	0.38457
HsMADS2.B99990002	-9444.38281	0.44596
HsMADS2.B99990003	-9234.76074	0.37662
HsMADS2.B99990004	-9098.85742	0.34488
HsMADS2.B99990005	-9229.38867	0.19162

**Supplementary Table 3** Sequences that homologous to HsMADS1 protein analysed by PSI-BLAST

PDB ID of template	Region of alignment (aa residues)	Identity (%)	E-value	Annotation of template
4OX0, chain A	1-103	61	$2e^{-33}$	Chain A, Crystal Structure Of The Keratin-like Domain From The Mads Transcription Factor Sepallata 3 [Arabidopsis thaliana]
1EGW, chain A	1-68	57	$3e^{-19}$	Chain A, Crystal Structure Of Mef2a Core Bound To Dna [Homo sapiens]
3KOV, chain A	1-68	57	$4e^{-19}$	Chain A, Structure Of Mef2a Bound To Dna Reveals A Completely Folded Mads-BoxMEF2 DOMAIN THAT RECOGNIZES DNA AND RECRUITS Transcription Co-Factors [Homo sapiens]
3 MU6, Chain A	1-68	57	$5e^{-19}$	Chain A, Inhibiting The Binding Of Class Iia Histone Deacetylases To Myocyte Enhancer Factor-2 By Small Molecules [Homo sapiens]
1TQE, Chain P	1-73	53	$9e^{-8}$	Chain P, Mechanism Of Recruitment Of Class Ii Histone Deacetylases By Myocyte

Supplementary Table 4 Sequences that homologous to HsMADS2 protein analysed by PSI-BLAST

PDB ID of template	Region of template sequence aligned (aa residues)	Identity (%)	e-value	Annotation of template
1EGW chain A	1-70	51	$4e^{-20}$	Chain A, Crystal Structure Of Mef2a Core Bound To Dna [Homo sapiens]
4OX0 chain A	7-104	47	$4e^{-20}$	Chain A, Crystal Structure Of The Keratin-like Domain From The Mads Transcription Factor Sepallata 3 [Arabidopsis thaliana]
3KOV chain A	1-70	51	$4e^{-20}$	Chain A, Structure Of Mef2a Bound To Dna Reveals A Completely Folded Mads-BoxMEF2 DOMAIN THAT RECOGNIZES DNA AND RECRUITS Transcription Co-Factors [Homo sapiens]
1TQE Chain P	1-71	51	$5e^{-20}$	Chain A, Inhibiting The Binding Of Class Iia Histone Deacetylases To Myocyte Enhancer Factor-2 By Small Molecules [Homo sapiens]
1C7U Chain A	1-70	51	$2e^{-19}$	Chain A, Complex Of The Dna Binding Core Domain Of The Transcription Factor
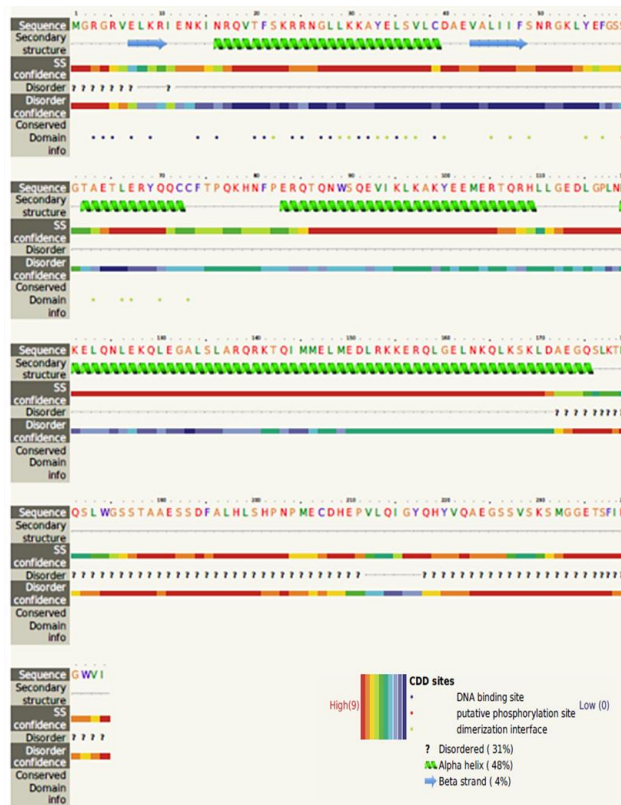
Supplementary Fig. 1 The summary of the secondary structure and disorder prediction of HsMADS1 using Phyre[2] (Kelley and Sternberg, 2009). The distribution of the disoredered, alpha helix, beta strands and binding sites were showed in 244 residues of HsMADS2 protein sequence.
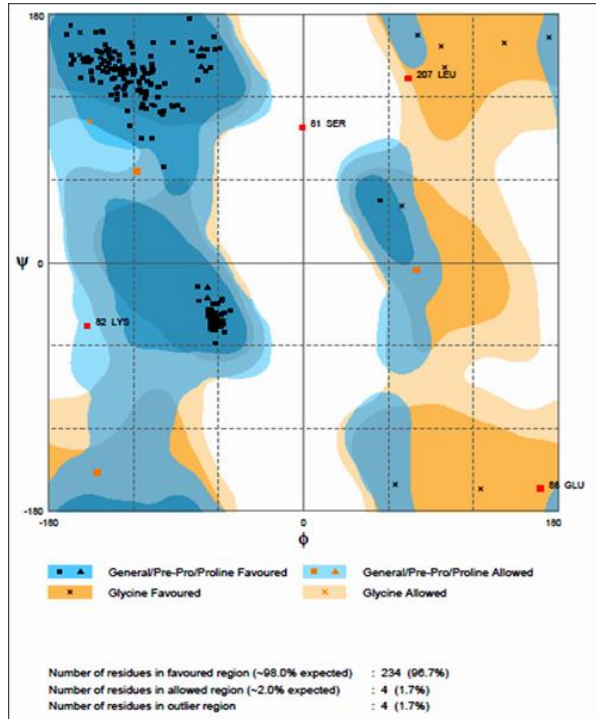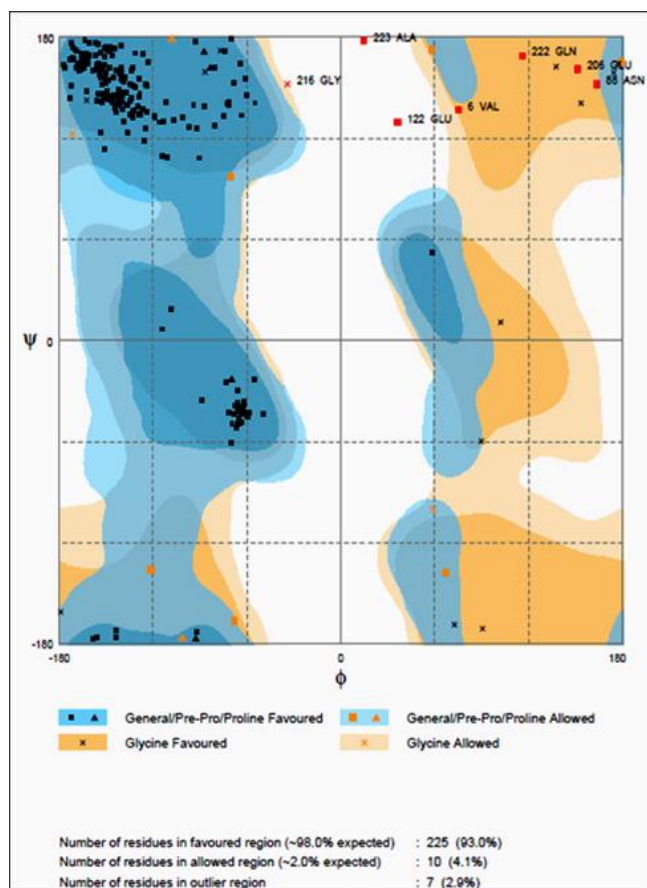
Supplementary Fig. 2 Ramachandran plot analysis for HsMADS1.B99990001 model by RAMPAGE (Lovell et al. 2003).
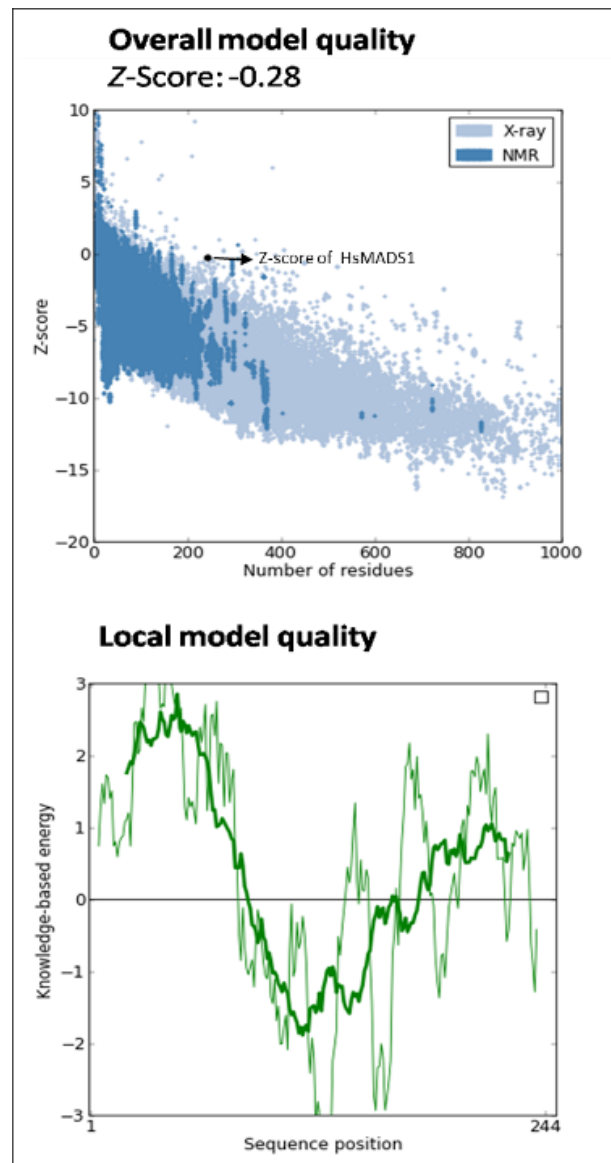
Supplementary Fig. 3 The overall modal quality and local model quality analysis for HsMADS1.B99990001 by ProSA (Sippl, 1993; Wiederstein and Sippl, 2007). The Z- scores showed by the plot contained the Z-scores of all experimentally determined protein chains in the current PDB database.

Supplementary Fig. 4 The summary of the secondary structure and disorder prediction of HsMADS2 using Phyre[2] (Kelley and Sternberg, 2009). The distribution of the disoredered, alpha helix, beta strands and binding sites were showed in 244 residues of HsMADS2 protein sequence.

Supplementary Fig. 5 Ramachandran plot for HsMADS2.B99990002 model by RAMPAGE (Lovell *et al*., 2003).
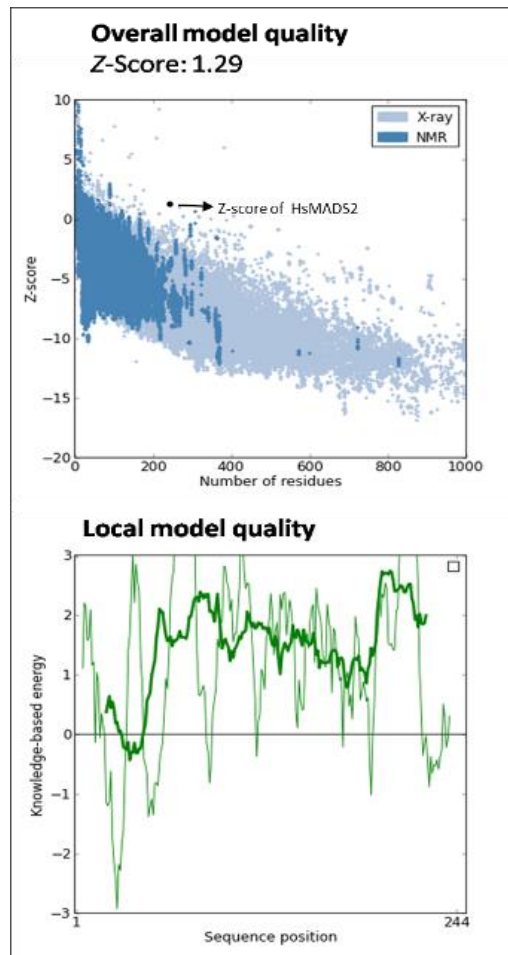
Supplementary Fig. 6 The overall modal quality and local model quality analysis for HsMADS2.B99990002 by ProSA (Sippl, 1993; Wiederstein and Sippl, 2007). The Z- scores showed by the plot contained the Z-scores of all experimentally determined protein chains in the current PDB database.

References:

Kelley LA, Sternberg MJE (2009) Protein structure prediction on the web: a case study using the Phyre server. Nat Protoc. 4: 363 – 371.

Lovell SC, Davis IW, Arendall, WB, de Bakker PI, Word JM, Prisant, MG, Richardson JS, Richardson DC (2003) Structure validation by C alpha geometry: phi, psi and C beta deviation. Proteins 50(3): 437-450.

Sippl MJ (1993) Recognition of Errors in Three-Dimensional Structures of Proteins. Proteins 17: 355-362.

Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 35: 407-410.