Available online at www.sciencedirect.com

Gene 347 (2005) 183–198

GENE
SECTION
EVOLUTIONARY GENOMICS

www.elsevier.com/locate/gene

Review

MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants

Kerstin Kaufmann, Rainer Melzer, Günter Theißen*

Friedrich-Schiller-Universität Jena, Lehrstuhl für Genetik, Philosophenweg 12, D-07743 Jena, Germany

Received 10 September 2004; received in revised form 11 December 2004; accepted 16 December 2004

Received by H.E. Roman

Abstract

MIKC-type proteins represent a class of MADS-domain transcription factors and are defined by a unique domain structure: in addition to the highly conserved DNA-binding MADS-domain, they have three other domains ('I', 'K' and 'C'), with the keratin-like K-domain being the most highly conserved and characteristic one. The number and functional diversity of MIKC-type proteins increased considerably during land plant evolution, culminating in higher flowering plants, where they dominate the control of reproductive development from early to late stages. We wonder how one special class of proteins became important in the control of essentially all stages of a morphogenetic process. All MADS-domain proteins appear to bind to DNA as homo- or heterodimers and may function as part of ternary transcription factor complexes involving non-MADS-domain proteins. Only MIKC-type proteins, however, generate complex intrafamily interaction networks. These are based on the special potential of MIKC-type proteins to form complexes involving more than two homologous proteins constituting transcriptional regulators. We speculate that the potential to form heteromultimers of homologous proteins was achieved by the acquisition of the K-domain during evolution. There is emerging evidence that organismal complexity arises from progressively more elaborate regulation of gene expression. We hypothesize that combinatorial multimer formation of MIKC-type MADS-domain proteins facilitated an unusually efficient and rapid functional diversification based on gene duplication, sequence divergence and fixation. This 'networking' may have enabled a more sophisticated transcriptional control of target genes which was recruited for controlling increasingly complex and diverse developmental pathways during the rapid origin and diversification of plant reproductive structures. Therefore, MIKC-type proteins may owe their evolutionary 'success' and present-day developmental importance in part to their modular domain structure. Investigating the evolution of MIKC-type genes may thus help to better understand origin and diversification of gene regulatory networks.

Keywords: Multidomain proteins; K-domain; Regulatory network; Flower

1. Introduction: from MADS to MIKC

MADS-box genes encode transcription factors which play fundamental roles in developmental control or signal transduction processes in most, if not all, eukaryotes (for a review, see Messenguy and Dubois, 2003). All MADS-box genes have in common a highly conserved DNA sequence of about 180 base pairs length, the MADS-box, which encodes the DNA-binding domain of MADS-domain proteins. MADS-domain proteins in animals and fungi comprise two large groups, ARG80/SRF-like and MEF2-like proteins (De Bodt et al., 2003).

Abbreviations: ABS, *Arabidopsis* B_{sister}; AG, Agamous; AGL1, 2, 4, 5, 9, 11, 13, Agamous-like gene1, 2, 4, 5, 9, 11, 13; AP1, 2, 3, Apetala1, 2, 3; BEL1, Bell1; CAL, Cauliflower; CO, Constans; DEF, Deficiens; GLO, Globosa; ELF1, Early flowering1; FLC, Flowering locus c; FRET, Fluorescence resonance energy transfer; FLIM, Fluorescence lifetime imaging microscopy; FUL, Fruitful; LD, Luminidependens; LFY, Leafy; LUG, Leunig; MADS, Mcm1, agamous, deficiens, srf; MEF2, Myocyte enhancer factor 2; NAP, NAC-like, activated by AP3/PI; PI, Pistillata; SAM, Srf, ag, mcm1; SIN1, Short integuments1; SRF, Serum response factor; SQUA, Squamosa; SUP, Superman; UFO, Unusual floral organs; TFL, Terminal flower.

* Corresponding author. Tel.: +49 3641 949550; fax: +49 3641 949552.
E-mail address: guenter.theissen@uni-jena.de (G. Theißen).

0378-1119/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2004.12.014

MADS-domain proteins of animals and fungi have roles in various processes, like growth, stress response and metabolism (Messenguy and Dubois, 2003). SRF (SERUM RESPONSE FACTOR) proteins are involved in, e.g., response to growth factors and muscle-specific gene transcription in mammals. Closely related to SRF is ARG80 from brewer's yeast (*Saccharomyces cerevisiae*), a regulatory protein of arginine metabolism and encoded by the first MADS-box gene ever cloned (Messenguy and Dubois, 2003). More widely known, however, is the closely related MINICHROMOSOME MAINTENANCE1 (MCM1) protein, which is involved in the transcriptional regulation of many genes involved in the yeast cell cycle and cell growth, metabolism and specialization, and cell type determination (Messenguy and Dubois, 2003). MEF2-like proteins, the other group of MADS proteins in animals and fungi, are key factors in muscle-specific gene regulation (animals) or involved in controlling the expression of genes required for cell wall integrity or osmotic stress response (yeast) (Messenguy and Dubois, 2003).

MIKC-type proteins constitute a special class of MADS-domain proteins that has been named after its characteristic domain structure, including a MADS (M-), intervening (I-), keratin-like (K-) and C-terminal (C-) domain (Theissen et al., 1996) (Fig. 1). MIKC-type genes are exclusive to plants. Phylogeny reconstructions suggest that MIKC-type proteins are more closely related to MEF2-like proteins than to ARG80/SRF-like proteins from fungi and animals (Fig. 2) (Alvarez-Buylla et al., 2000). There is no clear homology outside the MADS-domain, however. Alvarez-Buylla et al. (2000) hypothesized that an ancient duplication event occurred prior to the plant–animal divergence which gave rise to type I and type II MADS-box genes. Type I genes comprise the ARG80/SRF-like genes of animals and fungi (Theissen et al., 1996) and various groups of plant MADS-box genes that do not possess a K-domain and which are functionally largely uncharacterized (Alvarez-Buylla et al., 2000; De Bodt et al., 2003). Type II genes include the MEF2-like genes from animals and fungi and the MIKC-type genes from plants. However, even though plant MIKC-type genes appear more closely related to MEF2-like genes than to ARG80/SRF-like MADS-box genes in some phylogeny reconstructions, the deep branching of the MADS-box gene tree in eukaryotes is not clear (Fig. 2), particularly because of the weakly supported relationships of plant type I genes (Alvarez-Buylla et al., 2000; De Bodt et al., 2003).

Gene sampling in several basal green plant lineages and close relatives has narrowed down the evolutionary origin of MIKC-type genes. Red algae are probably the sister group of the green plants, which comprise green algae and land plants (Fig. 2). In the complete genome sequence of the red alga *Cyanidioschyzon merolae* (Matsuzaki et al., 2004) just one MADS-box gene with a MEF2-like MADS-domain, but without a K-domain, was identified so far. Charophytes (stoneworts) are a group of green algae which represent the closest relatives of the land plants, while chlorophytes comprise all other green algae (Fig. 2). Up to now, no MIKC-type gene has been reported from a chlorophyte (despite the *Chlamydomonas* genome project we could only identify a MADS-box-like motif that is only distantly related to other MADS-boxes). Since charophytes do possess MIKC-type genes (Y. Tanabe, M. Hasebe, and M. Ito, personal communication) these genes may have originated near the base of streptophytes more than 450 million years ago (MYA), prior to the origin of land plants (Fig. 2).

MIKC-type MADS-box genes have undergone radiation in ferns and seed plants (=flowering plants+gymnosperms). In seed plants the gene duplications clearly led to the establishment of conserved gene subfamilies that are pivotal in the development of novel seed plant-specific reproductive organs, and they were likely recruited for the initial evolutionary establishment of these key innovations. Function and general evolutionary aspects of major clades of MIKC-type MADS-box genes have been reviewed in detail elsewhere (Theissen et al., 2000). Briefly, members of the early identified and well-characterized subfamilies, such as the AG-, DEF-, GLO-and SQUA-like genes, typically share similar expression patterns and highly related functions

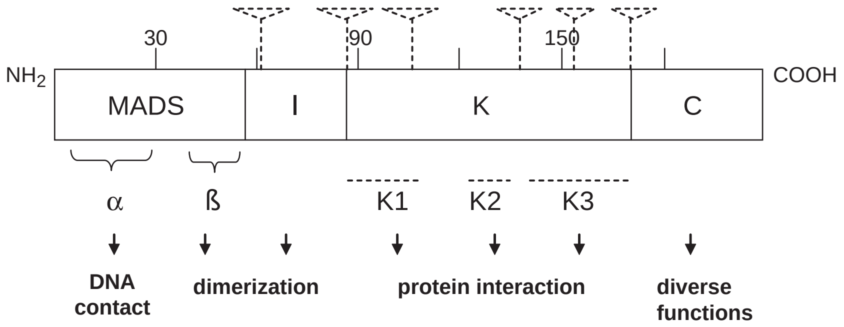

Fig. 1. Domain structure of MIKC-type MADS proteins: the MADS-, I-, K- and C-domains of APETALA3 are schematically presented. The N−terminal part of the MADS-domain forms an α-helix and binds DNA together with a partner protein. The other part constitutes a dimerization motif consisting of two antiparallel β strands. The K-domain forms probably three amphipathic α helices that may assemble into coiled-coil structures (K1, K2, K3). The numbers above the schematic drawing refer to amino acid positions, intron positions are marked as triangles (modified after Yang et al., 2003a,b).