Gene 526 (2013) 87–95

Contents lists available at SciVerse ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Localization and functional analysis of HmgB3p, a novel protein containing high-mobility-group-box domain from *Tetrahymena thermophila*

Jing Xu [a], Huaru Tian [a], Xingyin Liu [b], Wei Wang [a,*], Aihua Liang [a]

[a] Key Laboratory of Chemical Biology and Molecular Engineering of Ministry of Education, Institute of Biotechnology, Shanxi University, Taiyuan, China
[b] Albert Einstein College of Medicine, Department of Genetics, New York, USA

ARTICLE INFO

Article history:
Accepted 1 May 2013
Available online 15 May 2013

Keywords:
HMG-box
HMGB3
Micronuclei
Tetrahymena thermophila

ABSTRACT

The high-mobility-group (HMG)-box domain represents a very versatile protein domain that mediates the DNA-binding of non-sequence-specific and sequence-specific proteins. HMG-box proteins are involved in various nuclear functions, including modulating chromatin structure and genomic stability. In this study, we identified the gene *HMGB3* in *Tetrahymena thermophila*. The predicted HmgB3p contained a single HMG-box, an SK-rich-repeat domain and a neutral phosphorylated C-terminal. *HMGB3* was expressed in the growth and starvation stages. Furthermore, *HMGB3* showed a higher expression levels during the conjugation stage. *HMGB3* knockout strains showed no obvious cytological defects, although initiation of *HMGB3* knockout strain mating was delayed and maximum mating was decreased. HA-HmgB3p localized on the micronucleus (MIC) during the vegetative growth and starvation stages. Furthermore, HA-HmgB3p specially decorated the meiotic and mitotic functional MIC during the conjugation stage. Truncated *HMGB3* lacking the HMG box domain disappeared from MICs and diffused in the cytoplasm. Overexpressed HmgB3p was abnormally maintained in newly developing macronuclei and affected the viability of progeny. Taken together, these results show that HmgB3p is a germline micronuclear-specific marker protein. It may bind to micronucleus-specific DNA sequences or structures and is likely to have some function specific to micronuclei of *T. thermophila*.

1. Introduction

In eukaryotes, the primary protein components of chromatin are histones that compact the genome DNA. High-mobility-group (HMG) proteins are the second most abundant family of chromatin proteins (Bustin and Reeves, 1996) believed to have important functions in remodeling the assembly of chromatin and regulating gene transcription by distorting, bending, or modifying the structure of DNA, which is bound with histones and transcriptional factors (Agresti and Bianchi, 2003). Three distinct families of HMG proteins have been defined and named based on the structure of their DNA-binding domains and their substrate-binding specificity, including HMGA (the HMG-AT-hook family), HMGN (the HMG-nucleosome binding family), and HMGB (the HMG-box family) (Bustin, 2001). The HMG box family is well conserved and has a characteristic domain of about 80 amino acid (aa) residues consisting of three α-helices arranged in an "L" shape that can fit into the minor groove of duplex DNA (Thomas and Travers, 2001; Travers, 2000). Most HMG-box proteins contain a single HMG box. Chromatin HMGB proteins represent a small and specific subset of all HMG-box proteins and invariably contain two tandem HMG boxes and an acidic tail. HMGB proteins are conserved in all animals studied. Plants and unicellular organisms remarkably do not appear to contain canonical HMGB proteins (Sessa and Bianchi, 2007). Mammalian HMGB proteins are usually classified into two major groups distinguished by their abundance, function, and DNA specificity. HMGB1 and other related HMGB-type proteins typically contain two HMG-boxes that recognize distorted DNA structures with very weak or no sequence specificity, whereas the transcription factors of the HMG-box family contain only one HMG-box that binds both structure and sequence specifically to DNA (Stros et al., 2007). In yeast, loss of the HMGB-type proteins results in increased genomic instability and hypersensitivity to DNA-damaging agents (Giavara et al., 2005). In mouse, the Hmgb1 gene is essential but the lack of the Hmgb2 or Hmgb-3 leads to very slight defects (Sessa and Bianchi, 2007). In plants, HMGB proteins bind certain DNA structures, such as DNA-minicircles, four-way junctions, and supercoiled DNA, with higher affinity than the corresponding linear DNA (Grasser et al., 2007). In the dipteran insect *Chironomus*, cHMG1a protein is uniformly distributed along the chromosome, whereas the cHMG1b protein is specifically localized in chromosomal puffs of interphase giant chromosomes (Ghidelli et al., 1997). HMGB gene was also characterized from stored mRNA in resting cysts of the ciliate *Oxytricha nova* and the expressed HMG-box protein

Abbreviations: HMG, high-mobility-group; MAC, macronucleus; MIC, micronucleus; HMGA, HMG-AT-hook family; HMGN, HMG-nucleosome binding family; HMGB, HMG-box family; MLH, micronuclear linker histone; tsHMG, testis-specific high-mobility-group; SPP, super proteose peptone; RT, reverse transcriptase; qRT-PCR, real time quantitative PCR; DAPI, 4′,6-diamidino-2-phenylindole dihydrochloride.

* Corresponding author at: Institute of Biotechnology, Shanxi University, Wucheng Rd. 92, Taiyuan, Shanxi 030006, China. Tel./fax: +86 351 7011499.
E-mail address: gene@sxu.edu.cn (W. Wang).

0378-1119/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2013.05.006

might be involved in the chromatin condensation (Callejas and Gutierrez, 2003). HMG protein LG-2-like gradually accumulated in the nucleus on starvation in *T. pyriformis* (Suda and Hayashi, 1989). HMG-box proteins are found in a variety of eukaryotic organisms, and can be broadly divided into two groups, based on sequence-dependent and sequence-independent DNA recognition; the former usually contain one HMG-box motif, while the latter can contain multiple HMG-box motifs. Although some HMGB proteins in a variety of organisms have been extensively studied, many aspects of their function remain to be explored.

The ciliated protozoan *Tetrahymena thermophila* is an outstanding subject for exploration of the function of the HMG-box protein. *Tetrahymena* contains two nuclei, a germ line micronucleus (MIC) and a somatic macronucleus (MAC), in a single cell (Karrer, 2000). The MIC and MAC differ in their structure and function. The MIC is diploid and contains five pairs of chromosomes; it is the germline, the storage of genetic information for the progeny produced by conjugation in the sexual stage of the *T. thermophila* life cycle. During vegetative growth, the MIC is transcriptionally inert, whereas the MAC is transcriptionally active; gene expression occurs in the MAC, which is thus considered the somatic nucleus (Eisen et al., 2006). During conjugation, MICs undergo meiosis, mitosis, pronuclear exchange, fertilization, and postzygotic division giving rise to new MACs and new MICs (Cole et al., 1997). Although derived from the same zygotic nucleus, new MACs and new MICs differ in the organization of their genomes (Martindale et al., 1982; Yao et al., 1984). Four major high mobility group (HMG) proteins designated A, B, C, and D, have been isolated from *Tetrahymena* MACs (Levy-Wilson et al., 1983). HMGs B (LG-2) and C (LG-1) resemble vertebrate HMGs 14 and 17 in size and in the ability to interact with nucleosome core particles (Schulman et al., 1987). HMGs B and C participate in the packaging of MAC chromatin into a transcriptionally competent conformation (Schulman et al., 1991). In *Tetrahymena pyriformis*, Chromodomain protein LG-2-like protein accumulates in the macronuclei of starved cells and plays a role in the repression of chromatin template activity (Suda and Hayashi, 1989). *T. thermophila* MICs contain four linker-associated proteins, namely, α, β, γ, and δ. They are the ultimate products derived from the micronuclear linker histone (MLH). δ protein is strikingly similar to HMG protein HMG 1/2 of vertebrates (Wu et al., 1994). In mammalian spermiogenesis, the testis-specific high-mobility-roup (tsHMG) protein, similar to the δ protein of *Tetrahymena* MICs, appears in elongating mouse spermatids when they condense and cease transcription (Alami-Ouahabi et al., 1996).

In this study, we identified a novel HMG box gene *HMGB3* from the *Tetrahymena*. The structure and localization pattern of HmgB3p are different from those of HMGs B and C. HmgB3p specifically decorates the functional germline MIC throughout the vegetative growth, starvation and conjugation stages. The HMG-box domain is indispensable for HmgB3p micronuclear specific localization. Overexpression of HmgB3p affects the viability of progeny but knockout of the *HMGB3* gene has no visible effect on *Tetrahymena*. HmgB3p is a micronuclear-specific marker protein and might be involved in germline DNA meiosis and mitosis in *T. thermophila*.

2. Materials and methods

2.1. Tetrahymena strains and culture conditions

B2086 II, CU428 {Mpr/Mpr [6-methylpurinesensitive (mp-s), VII]}, and CU427 {Chx/Chx [cycloheximide sensitive (cy-s), VI]} were provided by Dr. Peter. J. Bruns (Cornell University, Ithaca, NY, USA, now available through the National Tetrahymena Stock Center, http://tetrahymena.vet.cornell.edu/index.html). *Tetrahymena* cells were grown in super proteose peptone (SPP) medium containing 1% proteose peptone, 0.2% glucose, 0.1% yeast extract, and 0.003% EDTA ferric sodium salt (1 × SPP) at 30 °C (Gorovsky et al., 1975). For conjugation, log-phase growing cells of different mating types were washed, starved in 10 mM Tris–HCl (pH 7.4) for at least 16 h without shaking, and mixed in equal amounts (~2.0×10^5 cells/mL), as previously described (Allis and Dennison, 1982).

2.2. Cloning and analysis of HMGB3 gene

HMGB3 (TTHERM_00155590) was retrieved from the *Tetrahymena* Genome Database (http://www.ciliate.org) by searching for matches with the mouse HMG Box Domain. Multiple sequence alignments were carried out with MegAlign application of the DNAStar 5.03 software package (DNASTAR, Inc., Madison, WI). RNA samples were extracted from *Tetrahymena* cells using Trizol (Takara) and then treated with RNase-free DNase I (Takara). First-strand cDNA was synthesized using a PrimeScript™ reverse transcriptase (RT) and random hexamer primers. *HMGB3* cDNA was amplified with primers HMGB3FW and HMGB3PRV (Table 1). The PCR product was cloned into T vectors and verified by sequencing. The expressed profile of *HMGB3* was analyzed by real time quantitative PCR (qRT-PCR). qRT-PCR was performed with the SYBR Premix Ex TaqTM (Takara) on an ABI StepOne Plus system (Applied Biosystems, USA). Each reaction was performed in triplicate. The values were normalized to the expression of the 17S rRNA as the internal control. The primers RT-HMGB-FW and RT-HMGB3-RV were used (Table 1). The following parameters were used for PCR: 10 min at 95 °C, followed by 40 cycles of 95 °C for 15 s, 53 °C for 30 s, and 68 °C for 35 s. The data from the real-time PCR experiments were analyzed using the $2^{-\Delta\Delta Ct}$ method, which allows for the calculation of relative changes in gene expression (Livak and Schmittgen, 2001).

2.3. Construction of HMG3 knockout strains

To make the targeting construct, the 5′ region flanking the *HMGB3* coding sequence in genomic DNA was first amplified using primers 5′ HMGB3-FW and 5′HMGB3-RFT (Table 1). The 3′ flanking region was amplified with primers 3′HMGB3-FW and 3′HMGB3-FW (Table 1). Neo4 cassette was amplified with primers Neo4-FW and Neo4-RV (Table 1). Then, the three PCR fragments were joined by overlapping PCR with primers 5′HMGB3-IFW and 3′HMGB3-IRV (Table 1) (Loidl

Table 1
Primers used in the paper.

Primer name	Sequence
5′HMGB3-FW	5′-TGCTTGCTTGCTTGTTTGTTTCCT-3′
5′HMGB3-RF	5′-GTGTATTTAAATTAAAGGAGTTATTCTGTTTT GAATTCAAATTCTTTGATAA-3′
3′HMGB3-FW	5′-CCTCTTCACATACATGTTAGCTCTGTAACTAAAAA CCTATTTGAAATATA-3′
3′HMGB3-RV	5′-ATGCAAAGCCAAAAGAAACTAGTTCT-3′
5′HMGB3-IFW	5′-CAGCATCCATTTGGTTATTTACTAA-3′
3′HMGB3-IRV	5′-TCTATTATGATTGCTGCCCTTTCTT-3′
Neo4-FW	5′-GAATAACTCCTTTAATTTAAATACAC-3′
Neo4-RV	5′-AGAGCTAACATGTATGTGAAGAGG -3′
HMGB3-FW	5′-ATGAGCCACTAAAGAAAAACTCAAAA-3′
HMGB3-RV	5′-TCATTGTTCAGCATTCATTTATTGTA-3′
RT-HMGB3-FW	5′-GAGTGGAATGAGTAGCATGCTCCTG-3′
RT-HMGB3-RV	5′-GCTGAAGCCTTCTTTTAAGAGGCCC-3′
17S-FW	5′-CCTGAGAAACGGCTACTACAACTA-3′
17S-RV	5′-AAATGTTTACTCCCTAAGTCGAAC-3′
HA-HMGB3-FW	5′-GTGGATCCAGCCACTAAAGAAAAACTCAAAAG-3′ (*BamH* I site is underlined)
HA-HMGB3-RV	5′-CAGGCGCGCCTCATTGTTCAGCATTCATTTATTGT-3′ (*Asc* I site is underlined)
HA-HMGB3ΔN-FW	5′-GTGGATCCGAGTAGCATGCTCCTGCTGATAA-3′ (*BamH* I site is underlined)
HA-HMGB3ΔC-RV	5′-AGGCGCGCCTCATTTTGAAGATTTAGAAGAATTATGG-3′ (*Asc* I site is underlined)
MTT1 FW	5′-GCTACGTGATTCACGATTTATGCAATG-3′
MTT1 RV	5′-CGAAACTGATTTTATGCAATTATGAATTAC-3′