The EMBO Journal Vol.16 No.13 pp.4021–4033, 1997

Cell differentiation by interaction of two HMG-box proteins: Mat1-Mc activates M cell-specific genes in *S.pombe* by recruiting the ubiquitous transcription factor Ste11 to weak binding sites

Søren Kjærulff[1], Dennis Dooijes[2], Hans Clevers[2] and Olaf Nielsen[1,3]

[1]Department of Genetics, Institute of Molecular Biology, University of Copenhagen, DK- 1353 Copenhagen K, Denmark and [2]Department of Immunology, University Hospital, PO Box 88500, 3508 GA Utrecht, The Netherlands

[3]Corresponding author
e-mail: onigen@biobase.dk

The *Schizosaccharomyces pombe mfm1* gene is expressed in an M cell-specific fashion. This regulation requires two HMG-box proteins: the ubiquitous Ste11 transcription factor and the M cell-controlling protein Mat1-Mc. Here we report that the *mfm1* promoter contains a single, weak Ste11-binding site (a so-called TR-box) that can confer M-specificity on a heterologous promoter when present in eight copies. *In vitro*, both Mat1-Mc and Ste11 can bind this box with approximately the same affinity. The Mat1-Mc protein caused a dramatic increase in the DNA-binding of Ste11 to this box, under conditions where we could not detect Mat1-Mc in the resulting protein–DNA complex. When we changed a single base in the *mfm1* TR-box, such that it resembled those boxes found in ubiquitously expressed genes, Ste11 binding was enhanced, and *in vivo* the *mfm1* gene also became expressed in P cells where Mat1-Mc is absent. These findings suggest that M-specificity results from Mat1-Mc-mediated Ste11 binding to weak TR-boxes. We have also defined a novel motif (termed M-box), adjacent to the *mfm1* TR-box, to which Mat1-Mc binds strongly. A DNA fragment containing both the TR- and the M-box allowed the formation of a complex containing both Ste11 and Mat1-Mc. A single copy of this fragment was sufficient to activate a heterologous promoter in an M-specific fashion, suggesting that these two boxes act in a synergistic manner.

Keywords: differentiation/fission yeast/HMG-box proteins/sex determination/transcriptional activation

Introduction

The HMG-box is a recently discovered DNA-binding element present in several eukaryotic proteins. Sequence analysis indicates that the HMG-box is a stretch of ~70 amino acids, with a net positive charge and an abundance of aromatic residues and prolines (Baxevanis and Landsman, 1995). Based on sequences and binding characteristics, the HMG-box protein family can be divided into two groups. One group includes proteins such as HMG1–2 and UBF that often contain multiple HMG-boxes and recognize DNA in a structure- rather than sequence-dependent fashion (reviewed by Bustin and Reeves, 1996). The second group contains proteins with a single HMG-box, and these bind DNA in a sequence-specific manner and some of them are activators of transcription. Members of this subfamily include the mammalian sex-determining factor SRY (Gubbay *et al.*, 1990; Sinclair *et al.*, 1990) and the related *SOX* gene products, the lymphoid-specific LEF-1/TCF-1 proteins (Travis *et al.*, 1991; Van de Wetering *et al.*, 1991; Waterman *et al.*, 1991) and several proteins involved in mating-type determination in fungi, e.g. Mat1-Mc and Ste11 from the fission yeast *Schizosaccharomyces pombe* (Kelly *et al.*, 1988; Sugimoto *et al.*, 1991; Dooijes *et al.*, 1993).

Although the similarity of the primary sequence is modest, all HMG-box proteins probably fold up in the same tertiary structure. Thus, in nuclear magnetic resonance (NMR) studies, the HMG-boxes of HMG1, SRY and SOX4 (Weir *et al.*, 1993; van Houte *et al.*, 1995; Werner *et al.*, 1995) were found to adopt highly similar structures consisting of three α-helixes, arranged in an L shape. This may provide a structural basis for the unusual DNA-binding characteristics of the HMG-box proteins, which include interaction with the minor groove of the DNA helix, binding to irregular DNA structures, such as the sharp bends present in four-way junction DNA molecules, and the ability to modulate the DNA helix by bending (Bustin and Reeves, 1996). The biological significance of this DNA bending is still obscure. However, it was reported recently that DNA bending induced by the LEF-1 protein facilitates the formation of a higher-order nucleoprotein complex in the T-cell receptor α (TCRα) enhancer (Giese *et al.*, 1995), suggesting that HMG-box proteins may have an architectural role in assembling such complexes. Supporting this idea, a mutant SRY protein that binds DNA with almost normal affinity but bends DNA in a different angle has been found in a sex-reversed XY patient (Pontigga *et al.*, 1994).

Although data on HMG-box proteins and their interaction with DNA have accumulated during the last few years, little is known about their *in vivo* action, and the mechanism by which the SRY protein controls male development in mammals is still largely unknown (Schafer and Goodfellow, 1996). SRY is highly homologous to the mating-type protein Mat1-Mc from *S.pombe* (Kelly *et al.*, 1988; Gubbay *et al.*, 1990; Sinclair *et al.*, 1990). SRY and Mat1-Mc bind to the same sequence, CTTTGTT, *in vitro* (Dooijes *et al.*, 1993), and Mat1-Mc has formally a function similar to SRY in establishing sex-specific gene expression.

The sexual differentiation process in *S.pombe* is activated under conditions of nitrogen starvation, where the cells are induced to exhibit either minus (M) or plus (P) mating behaviour, depending on which gene they express from the *mat1* locus. Expression of the *mat1-Mc* gene generates an M cell, whereas expression of the *mat1-Pc*

© Oxford University Press

gene gives rise to a P cell (Kelly *et al.*, 1988). The Mat1-Mc and Mat1-Pc proteins specify the mating type by activating a number of, respectively, M- or P-specific genes. The products of these are the cell type-specific components of the pheromone communication system that enables the two cell types to identify each other prior to mating (reviewed by Nielsen and Davey, 1995). The M-specific genes controlled by Mat1-Mc include three structural genes for the M-factor pheromone, *mfm1-3* (Davey, 1992; Kjærulff *et al.*, 1994), the *mam1* gene encoding an M-factor transporter (Christensen *et al.*, 1997), a gene, *mam2*, that encodes the receptor for the P-factor pheromone (Kitamura and Shimoda, 1991) and the *sxa2* gene encoding a P-factor-degrading protease (Imai and Yamamoto, 1992; Ladds *et al.*, 1996; Yabana and Yamamoto, 1996).

The Ste11 protein, which is a key transcription factor in the sexual differentiation pathway in *S.pombe*, is one of the few HMG-box proteins with known target sites. Ste11 binds to the so-called TR-box, TTCTTTGTTY (Sugimoto *et al.*, 1991), the core of which is identical to the Mat1-Mc-binding site. Ste11 is activated by nitrogen starvation (Li and McLeod, 1996), and TR-boxes are found in the promoter regions of many genes that are expressed in a Ste11-dependent manner during mating. These include the M-specific genes as well as genes that are expressed in both cell types (Sugimoto *et al.*, 1991; Kjærulff *et al.*, 1994; Petersen *et al.*, 1995).

In the present study, we investigate the mechanism by which M-specific genes are activated during sexual differentiation in *S.pombe*. We show that M-specificity is conferred on the *mfm1* gene by a special version of the TR-box that binds Ste11 poorly. Both Ste11 and Mat1-Mc can bind to this box and, under conditions of limiting amounts of Ste11, the Mat1-Mc protein can recruit Ste11 to the TR-box. We propose that ubiquitously expressed genes harbour a strong TR-box, to which Ste11 can bind on its own, whereas M-specificity results from Mat1-Mc-dependent Ste11 binding to a weak TR-box.

Results

Mat1-Mc and Ste11 bind the TR-box of the mfm1 promoter

As a representative M cell-specific gene we chose *mfm1*, one of three structural genes encoding M-factor pheromone (Davey, 1992; Kjærulff *et al.*, 1994). Expression from *mfm1* was monitored using a fusion of the *mfm1* promoter and the *Escherichia coli lacZ* gene (Figure 1A). This fusion behaves like the wild-type *mfm1* gene; expression is limited to M cells, is induced by nitrogen starvation and further stimulated by a pheromone signal from P cells (Figure 1B, wt; Kjærulff *et al.*, 1994).

The Ste11 transcription factor is required for induction of *mfm1* (Kjærulff *et al.*, 1994). While Ste11 may regulate *mfm1* expression through its control of *mat1-Mc* (Sugimoto *et al.*, 1991), it also appears to play a more direct role, since a *ste11* strain harbouring a plasmid that produces functional Mat1-Mc protein from the *nmt* promoter still fails to transcribe the *mfm1–lacZ* fusion (Figure 1B, Ste11 + pnmt-Mc). Consistent with this, a TR-box is situated 79 bp upstream of the transcription start point (*tsp*) in *mfm1*, and all other known M-specific genes contain a TR-box at a similar position (Figure 1A, Table I). To determine the functional significance of this element, we altered the conserved G of the *mfm1* TR-box to a T. This mutation prevented binding of Ste11 *in vitro* (see below) and almost completely abolished promoter function (Figure 1B, TR-mut1).

The Mat1-Mc protein was shown previously to bind the sequence CTTTGTT (Dooijes *et al.*, 1993), which constitutes the core of the TR-box, and we therefore compared the abilities of the Ste11 and Mat1-Mc proteins to bind the TR-box of the *mfm1* promoter *in vitro*. These experiments showed that *E.coli*-expressed GST–Ste11 and malE–Mat1-Mc fusion proteins bind to an oligonucleotide covering the TR-box of *mfm1* with approximately equal affinity (K_d ~10^{-8} M). In both cases, the retarded complex was competed efficiently by the TR-box, but not by the mutagenized TR-box (TR-mut1), nor by two oligonucleotides containing unrelated sequences (Figure 2A and B).

The TR-box of mfm1 confers M-specific expression on a heterologous promoter

Given the fact that Mat1-Mc binds the *mfm1* TR-box, we investigated whether this element could confer M-specific expression on a heterologous promoter. Various copies of it were inserted in the *Saccharomyces cerevisiae CYC1* minimal promoter, which was fused to the *E.coli lacZ* gene (Lowndes *et al.*, 1992). A single *mfm1* TR-box in the minimal promoter produced negligible β-galactosidase activity (Figure 3). However, when eight copies of this TR-box were present, a high level of β-galactosidase activity was induced by nitrogen starvation, and this activity was stimulated further by a pheromone signal. Most importantly, however, induction of expression was restricted to M cells.

The TR-boxes of ubiquitously expressed genes and those of M-specific genes differ in sequence

These observations indicated that the M-specificity of the *mfm1* promoter may lie in the TR-box or sequences in its immediate vicinity. This is quite surprising, since TR-boxes are also found in genes expressed in both cell types. However, we noticed a striking difference between the TR-boxes found in M-specific genes and those found in ubiquitously expressed genes (Table I). All promoters of the latter class contain at least one copy of the 10 bp motif, TTTCTTTGTT. This 10 bp motif is not found in any of the six known M-specific genes. Here the consensus is somewhat smaller, namely the 8 bp motif TCTTTGTT. To test whether this sequence difference was responsible for M-specificity, we changed the TR-box of the *mfm1–lacZ* fusion into the version found in ubiquitously expressed genes by substituting the 5′ C with a T. Interestingly, this construct was now expressed in both M and P cells (Figure 1B, TR-mut2). Hence, the addition of only one T to the 5′ end of the TR-box converts *mfm1* from an M-specific gene into a gene that is expressed in both cell types. Furthermore, we note that in P cells expression requires a pheromone signal.

Recently, we found by site selection that the preferred binding site of Ste11 *in vitro* is the 13 bp motif, TTTCTTTGTTCTC (Dooijes *et al.*, in preparation), which resembles the TR-box found in genes expressed in both cell types. Insertion of this sequence in the *mfm1* promoter

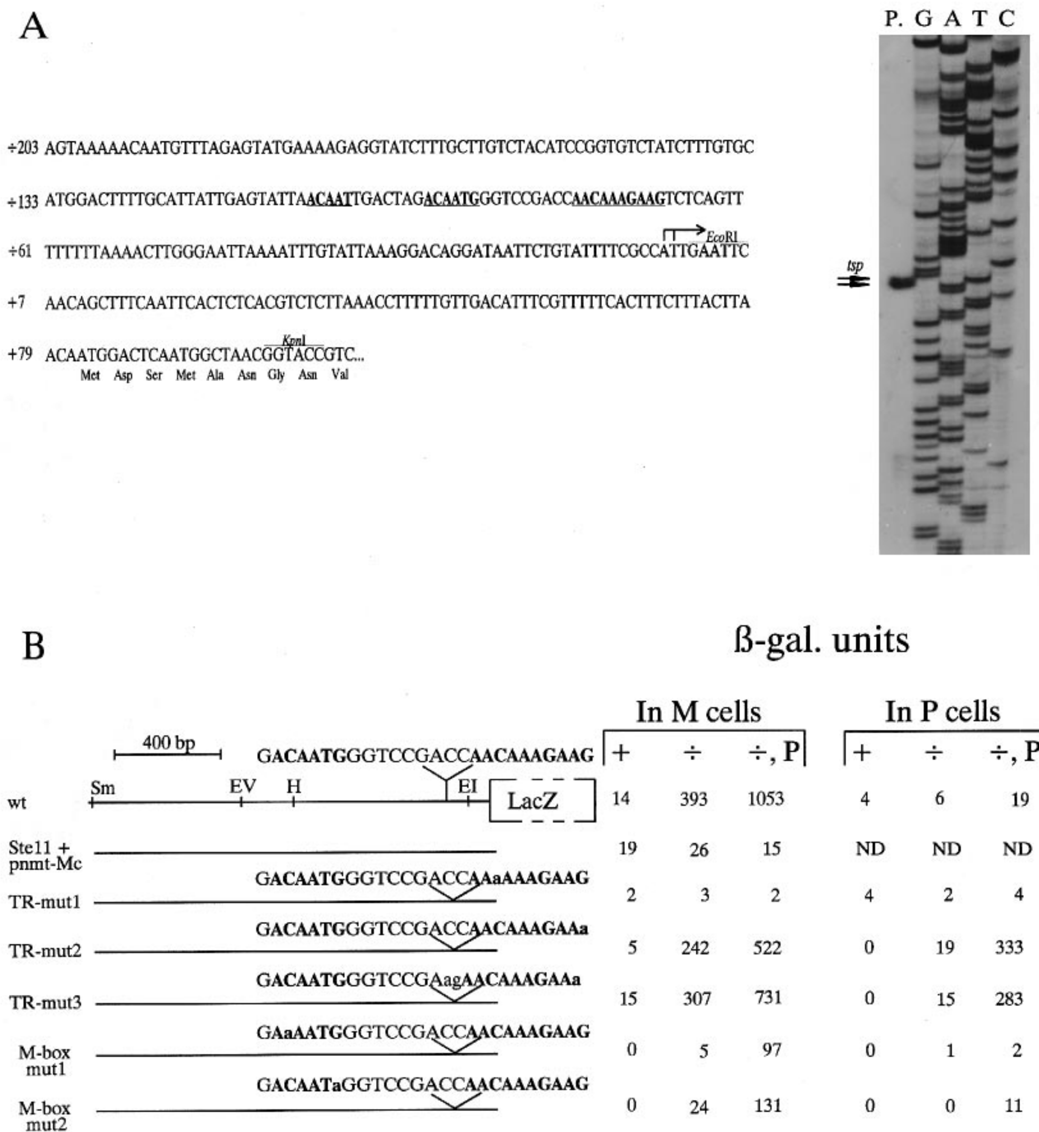


Fig. 1. Mutagenesis reveals two elements important for proper expression of the *mfm1* gene. (**A**) Sequence of the *mfm1* promoter. The TR-box and the M-boxes are in bold and underlined. Transcription start points (*tsp*) are indicated by an arrow. The *Kpn*I site is the fusion point between *mfm1* and *lacZ*. The right panel shows a primer extension analysis (P.) to determine the transcription start point of the *mfm1–lacZ* construct. The sequence of *mfm1–lacZ* was run in parallel (GATC). The arrowheads indicate the positions of the *tsp*s. (**B**) Mutational analysis of the *mfm1–lacZ* fusion. wt is the wild-type *mfm1–lacZ* construct. Ste11 + pnmt-Mc is the wild-type *mfm1–lacZ* construct integrated in an M *ste11* strain containing a plasmid overexpressing *mat1-Mc*. The sequence located 69–98 bp upstream of the transcription start point has been highlighted. Point mutations introduced in this region are indicated by lower case letters (TR-mut1-3 and M-box mut1-2). All constructs were integrated in the *mfm1* locus, at the *Sma*I site. β-Galactosidase activities were measured in exponentially growing cultures (+), in cultures starved of nitrogen for 5 h (÷) and in cultures starved of nitrogen and exposed to pheromone for 5 h (÷, P). β-Galactosidase activities are expressed in Miller units and represent the mean of three separate trials. Restriction sites: Sm, *Sma*I; EV, *Eco*RV; H, *Hin*dIII; EI, *Eco*RI.

also renders the gene ubiquitously expressed (Figure 1B, TR-mut3). Thus, the TR-box from ubiquitously expressed genes and the optimal Ste11-binding site both confer non-cell type-specific expression on the *mfm1* gene. Furthermore, methylation interference experiments have showed that the two 5′ Ts, which are missing in the M-specific TR-boxes, indeed are contacted by the Ste11 protein (Dooijes *et al.*, in preparation).

Mat1-Mc stimulates Ste11 binding to the mfm1 TR-box

How can the absence of this T residue render expression dependent on the Mat1-Mc protein? We found that Ste11 binds the TR-box of the M-specific *mfm1* gene more weakly than it binds the version found in ubiquitously expressed genes, whereas the Mat1-Mc protein seemed to bind the two different boxes equally well (Figure 4A). Western analysis showed that expression of Ste11 is induced by nitrogen starvation and that the level of Ste11 protein does not appear to be higher in M cells than in P cells (Figure 4B). We therefore speculated that Mat1-Mc may control the M-specific genes by assisting binding of Ste11 to their TR-boxes, and the following observations support this idea. Under conditions of limiting amounts of purified Ste11, where virtually no complex occurred with the *mfm1* TR-box probe, we observed that addition of small amounts of purified Mat1-Mc caused a significant increase in appearance of shifted complex (Figure 4C).

Table I. Comparison of TR-boxes found in M-cell-specific genes and in genes expressed in both cell types

Gene	Sequence	Position (C from ATG)	Orientation	Reference
In genes expressed in both cell types				
fus1[a,b]	GTA**TTTCTTTGTT**CTTTA	–247	>	1
fusl[a,b]	AAC**TTTCTTTGTT**CGGTT	–158	<	1
map1[a]	GTG**TTTCTTTGTT**ACAAA	–59	<	2,3
mei2[a]	CGA**TTTCTTTGTT**CCTAT	–1890	>	4,5
mei2[a]	AAG**TTTCTTTGTT**TTACA	–1868	>	4,5
mei2[a]	GAGA**TTCTTTGTT**TACTT	–1696	>	4,5
mei2[a,b]	TAAC**TTCTTTGTT**CTCTA	–1516	>	4,5
mei2[a,b]	TCT**TTTCTTTGTT**TGTTT	–911	>	4,5
rep1[a]	ATT**TTTCTTTGTT**TACAT	–178	>	6
rep1[a]	TACATC**CTTTGTT**TACAA	–165	>	6
spk1	AAG**TTTCTTTGTT**AATGT	–602	<	7
spk1	AAC**TTTCTTTGTT**ATTGT	–598	>	7
ste4[a]	TGC**TTTCTTTGTT**ATAAA	–84	>	8
ste6[a]	GAA**TTTCTTTGTT**TACTA	–174	<	9
ste11[a]	TTG**TTTCTTTGTT**GCAAT	–1375	>	5
zfs1	ATT**TTTCTTTGTT**TGACG	–493	>	10
zfs1	CAAC**TTCTTTGTT**CGTTT	–92	>	10
Consensus	**TTTCTTTGTT**			
In M cell-specific genes				
mam1[a]	ATGGGC**CTTTGTT**AGGTA	–203	>	11
mam2[a]	ATTCC**TCTTTGTT**TAGAA	–123	<	12
mfm1[a,b]	AGAC**TTCTTTGTT**GGTCG	–157	<	13,14
mfm2[a]	TGATC**TCTTTGTT**CATTT	–138	<	13
mfm3[a]	AGAC**TTCTTTGTT**GTTTC	–154	<	15
sxa2	GGGTG**TCTTTGTT**GCCCA	–322	>	16
Consensus	**TCTTTGTT**			

[a]Expression has been shown to be reduced in a *ste11*$^-$ mutant.
[b]TR-box has been shown to be required for expression *in vivo*.
References: 1, Petersen *et al.* (1995); 2, Yabana and Yamamoto (1996); 3, Nielsen *et al.* (1996); 4, Watanabe *et al.* (1988); 5, Sugimoto *et al.* (1991); 6, Sugiyama *et al.* (1994); 7, Toda *et al.* (1991); 8, Okazaki *et al.* (1991); 9, Hughes *et al.* (1990); 10, Kanoh *et al.* (1995); 11, Christensen *et al.* (1997); 12, Kitamura and Shimoda (1991); 13, Davey (1991); 14, this work; 15, Kjærulff *et al.* (1994); 16, Imai and Yamamoto (1992).

Interestingly, this complex co-migrated exactly with the binary Ste11–DNA complex. Moreover, addition of Ste11 antibodies to the induced complex gave rise to a supershift, whereas addition of malE antibodies (which recognize the malE–Mc fusion protein) had no effect. This indicates that the induced complex detected in this assay consists mainly of Ste11 and that Mat1-Mc is not a stable component of it. The enhancement of Ste11 binding seems to be mediated specifically by Mat1-Mc, since addition of purified human SRY protein (Sinclair *et al.*, 1990), that binds the TR-box *in vitro*, had no stimulatory effect on Ste11 binding (data not shown).

Mat1-Mc generates a DNase I-hypersensitive site in the mfm1 promoter

Taken together, the results described above strongly indicate that M-specificity is conferred on the *mfm1* gene by the presence of a special version of the TR-box, to which the binding of Ste11 is mediated by Mat1-Mc. To investigate further the mechanism by which Mat1-Mc may enhance the binding activity of Ste11, we performed an *in vitro* DNase I footprint on the *mfm1* leader (Figure 5). As expected, both Mat1-Mc and Ste11 could protect the same 12 bp region spanning the TR-box. However, Mat1-Mc created a strong hypersensitive site just 3 bp downstream of the TR-box, suggesting that binding of this protein causes a strong distortion of the *mfm1* promoter. Ste11 did not give rise to this hypersensitive site, and when Ste11 was added together with Mat1-Mc it became less apparent, consistent with a mechanism where Ste11 replaces Mat1-Mc at the TR-box. In summary, these observations confirm that Ste11 and Mat1-Mc both have the ability to bind the *mfm1* TR-box. However, they also reveal that, upon binding, the two proteins modulate the DNA helix differently. Mat1-Mc seems to produce a strong distortion of the DNA, which may be crucial for efficient binding of Ste11 to this TR-box.

Mat1-Mc binds to two different elements

Unexpectedly, we found that Mat1-Mc also protected a 21 bp A-rich region starting 8 bp upstream of the TR-box (Figure 5). In fact, Mat1-Mc seems to protect this region better than the TR-box. To a lesser extent, the Ste11 protein also protected this upstream region (Figure 5). We therefore compared the sequences next to the TR-boxes in the six known M-specific genes (Table II). Five of these genes each habour two ACAAT-boxes that are located, respectively, 14–16 bp and 24–26 bp from the inverted TR-box. The *mam2* gene is an exception: here we only found the somewhat diverged sequence, ACATA, located 26 bp from the TR-box. We next compared the abilities of Mat1-Mc and Ste11 to bind an oligonucleotide covering the most downstream ACAAT-box of *mfm1*. Purified Ste11 only forms a weak complex with this element, whereas Mat1-Mc binds strongly to the ACAAT-box (Figure 6). Actually, Mat1-Mc prefers the ACAAT-

Fig. 2. The HMG-boxes of Mat1-Mc and Ste11 bind specifically to the TR-box of the *mfm1* gene. EMSA performed on various labelled probes (see Materials and methods) using *E.coli*-expressed, purified GST–Ste11 fusion protein (**A**) or malE–Mat1-Mc fusion protein (**B**). Competitor DNA was added in 100- or 500-fold molar excess, as indicated by the triangles. FP indicates unbound probe. The mutation in TR-mut is shown in Figure 1B (TR-mut1). The TCF-box and LEF-box oligonucleotides contain binding sites for mammalian HMG-box proteins.

Fig. 3. Eight copies of the TR-box of *mfm1* confer M cell-specific expression on a minimal promoter from the *S.cerevisiae* cytochrome *c* gene fused to the *E.coli lacZ* gene. One or eight copies of an oligonucleotide containing the *mfm1* TR-box were inserted in the vector pSPΔ178 (Lowndes *et al.*, 1992), giving pcyc-TR (1/8). One or eight copies of an oligonucleotide covering the most downstream M-box of *mfm1* (see below) were inserted in pSPΔ178, giving pcyc::M-box (1/8). The *mfm1–lacZ* fusion (see Figure 1) inserted in the vector pDW232 (Weilguny *et al.*, 1991) was used as positive control (pmfm1). The vector pSPΔ178 (pcyc) was used as a negative control. The constructs were transformed into h^- (M) and h^+ (P) strains and assayed for β-galactosidase activities. The results are expressed in Miller units and each number is the average of three separate trials. Open boxes represent activities in vegetatively growing cultures, boxes hatched vertically are activities in nitrogen-starved cultures and boxes hatched horizontally are activities in nitrogen-starved cultures treated with pheromone.

box to the TR-box. Binding of Mat1-Mc was abolished by changing the conserved C of the ACAAT-box to an A. We refer to this element as an M-box, since it preferably makes complexes with Mat1-Mc.

To determine the functional significance of this element, we altered the conserved C of the *mfm1* M-box to an A. This mutation severely reduced the *mfm1* promoter function (Figure 1B, M-box mut1), demonstrating that the M-box indeed is important for expression of M-specific genes. However, the construct still supported a relatively high level of pheromone-induced expression in M cells.

Based on their expression pattern, the M-specific genes can be divided into two groups. The four genes *mam1*, *mam2*, *mfm1* and *mfm2* are highly induced by nitrogen starvation *per se*, whereas *mfm3* and *sxa2* in addition to starvation require a pheromone signal for efficient expression (Kitamura and Shimoda, 1991; Imai and Yamamoto, 1994; Kjærulff *et al.*, 1994; Christensen *et al.*, 1997). A closer examination of the M-boxes revealed that *mam1*, *mam2*, *mfm1* and *mfm2* all have a conserved G residue in the 3′ end of one of their M-boxes, whereas *mfm3* and *sxa2* lack this G (Table II). To test whether this sequence difference could explain the pheromone-induced nature of *mfm3* and *sxa2*, we changed the G in the M-box of the *mfm1* promoter to an A. This mutation also reduced binding of the Mat1-Mc protein to the M-box *in vitro*

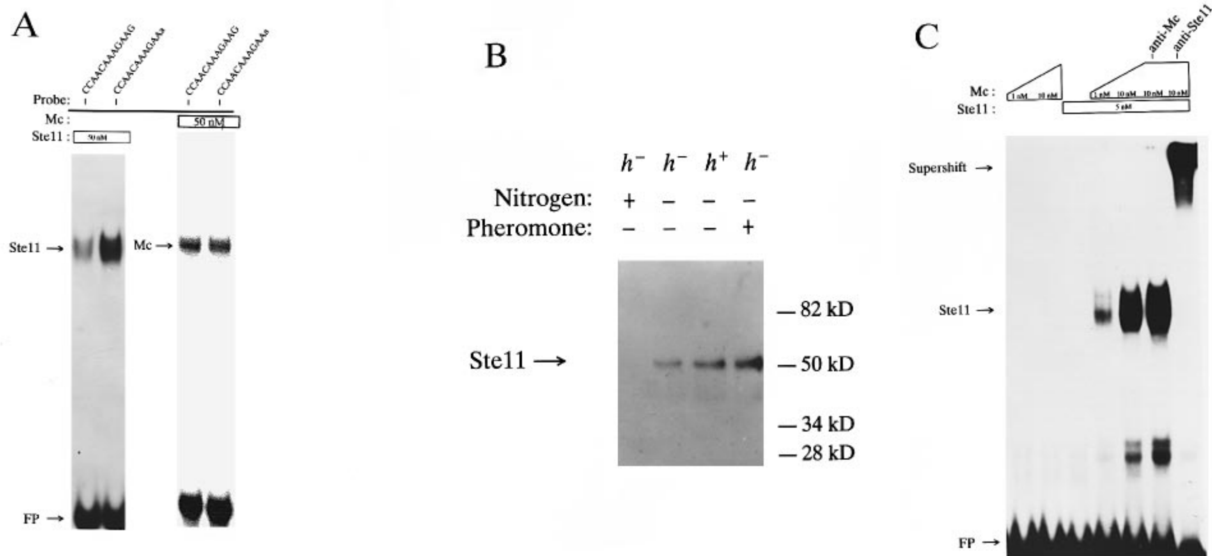


Fig. 4. (**A**) Ste11 binds better to the version of the TR-box found in ubiquitously expressed genes than to the *mfm1* TR-box, whereas mat1-Mc binds the two versions equally well. EMSA was performed with a labelled *mfm1* TR-box probe and a labelled *mfm1* TR-box mut2 probe using *E.coli*-expressed GST–Ste11 and malE–Mat1-Mc proteins. (**B**) The expression pattern of the Ste11 protein is identical in the two cell types. Western analysis of the expression of Ste11 in M cells (h^-) and in P cells (h^+). Protein extracts (100 μg) from mitotically growing cells (nitrogen +), from nitrogen-starved cells (nitrogen –) and from nitrogen-starved cells treated with pheromones (pheromone +) were Western blotted and probed with affinty-purified anti-Ste11 antibodies. (**C**) Mat1-Mc stimulates specific binding of Ste11 to the TR-box of *mfm1*. In an EMSA, the labelled TR-box was incubated with 5 nM of purified GST–Ste11 and increasing amounts of malE–Mat1-Mc (1, 10 nM). In the last two lanes, anti-malE antibodies or anti-Ste11 antibodies were added to the binding reaction after 20 min of incubation.

(Figure 6), and now expression was only moderately induced by nitrogen starvation (Figure 1B, M-box mut2). However, this construct could still be induced to a relatively high level by a pheromone signal. Hence, strong interaction of Mat1-Mc with the M-box is required for expression induced by nitrogen starvation, whereas this interaction apparently is not important for pheromone stimulation.

Synergistic function of the TR-box and the M-box of mfm1

We were unable to demonstrate any activation of the *S.cerevisiae CYC1* promoter by the M-box, even when eight copies of it were inserted (Figure 3). However, an oligonucleotide containing the M-box combined with its downstream TR-box conferred M-specific expression on the *CYC1* promoter—even in one copy (Figure 7A). Given the fact that a single TR-box had no effect on the minimal promoter (Figure 3), this result implies that the two HMG-box-binding sites work in a synergistic fashion. This synergy could reflect stable interaction between the Mat1-Mc and Ste11 proteins and DNA. To test this idea, we used the oligonucleotide containing the M-box and the TR-box in an electrophoretic mobility shift assay (EMSA). As expected, both Mat1-Mc and Ste11 bind this oligonucleotide on their own (Figure 7B). However, when both HMG-box proteins were present simultaneously, a unique complex was formed with slower mobility. Addition of antibodies against Ste11 or malE–Mat1-Mc both caused a supershift of this slow migrating complex, showing that the complex is of ternary nature containing Ste11, Mat1-Mc and DNA. Hence, Mat1-Mc may have two roles in the *mfm1* promoter; it enhances the binding of Ste11 to the TR-box (Figure 4C) and it forms a ternary complex with Ste11 when both the TR- and M-box are present (Figure 7B).

We next asked whether Mat1-Mc and Ste11 also interact *in vivo*. To address this question, we tagged the Mat1-Mc protein N-terminally with an influenza haemagglutinin (HA) epitope. The Mat1-Mc protein is expressed during vegetative growth, and further induced by nitrogen starvation (Figure 8A). Mat1-Mc appeared as a doublet band in the Western analysis irrespective of growth conditions. This doublet band is apparently due to phosphorylation (data not shown). We next overexpressed Ste11 and HA-tagged Mat1-Mc from the *nmt* promoter in M *ste11* cells and made immunoprecipitations from cleared protein extracts using Ste11 antibodies. As shown in Figure 8B, the immunoprecipitated complex contained HA–Mat1-Mc. Furthermore, the co-immunoprecipitation of HA–Mat1-Mc was dependent on the presence of Ste11.

To confirm this interaction, we tested whether malE–Mat1-Mc co-purifies with GST–Ste11 using glutathione–Sepharose chromatography. Purified GST–Ste11 protein or unfused GST were immobilized on glutathione–Sepharose beads and incubated with *E.coli* extracts containing either malE or malE–Mat1-Mc. Proteins bound to the beads were analysed by immunoblotting with anti-malE antibodies (Figure 8C, upper panel). MalE–Mat1-Mc was detected in the GST–Ste11 sample but not with GST alone, suggesting that the two HMG-box proteins also interact *in vitro*. Unfused control malE protein did not interact with any of the analysed proteins. To map the surface of interaction between Mat1-Mc and Ste11, we fused the HMG-boxes of Ste11 and Mat1-Mc to, respectively, GST and malE. GST–Ste11$_{HMG}$ protein or unfused GST was immobilized on glutathione–Sepharose beads and challenged with malE, malE–Mat1-Mc$_{HMG}$ and malE–Mat1-

Mc_{FL} (full-length Mat1-Mc). Full-length Mat1-Mc and the HMG-box of Mat1-Mc alone both bound the HMG-box of Ste11 to similar extents, indicating that the HMG-boxes of Ste11 and Mat1-Mc are sufficient for the interaction (Figure 8C, upper panel). Unfused control proteins did not interact with any of the analysed proteins. In conclusion, Mat1-Mc and Ste11 are likely to interact, and this appears to be mediated by interaction of their HMG-boxes.

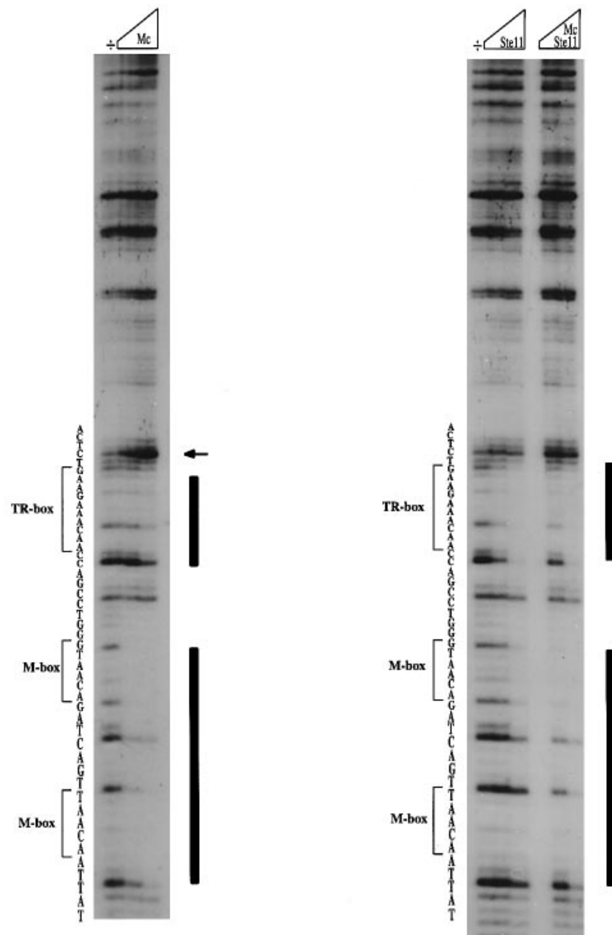


Fig. 5. Mat1-Mc creates a DNase I-hypersensitive site at the TR-box and binds several regions of the *mfm1* promoter. Solid phase DNase I footprinting analysis of a 222 bp region of the *mfm1* promoter spanning the TR-box. ÷ indicates DNase I digestion of naked DNA. Triangles indicate increasing amounts of added protein; 0.2–1.0 μM malE–Mat1-Mc; 0.1–0.5 μM Ste11. Vertical bars show the protected regions. The arrow indicates the DNase I-hypersensitive site. The sequence of the analysed region is given at the left.

Discussion

In the present study, we have examined the transcriptional regulation of M-specific genes in *S.pombe*. We have looked at the expression of the *mfm1* gene, one of three genes encoding M-factor mating pheromone. Our main conclusion is that *mfm1* is activated by the cooperative binding of the two HMG-box proteins, Mat1-Mc and Ste11. We propose that M-specificity is conferred on an *S.pombe* gene by the presence of a special version of the so-called TR-box (Figure 9). The ubiquitously expressed genes harbour a strong TR-box (TTTCTTTGTT), to which Ste11 can bind on its own, whereas M-specific genes contain a weak TR-box (TCTTTGTT) that Ste11 only binds efficiently in the presence of Mat1-Mc.

Both Mat1-Mc and Ste11 can bind to the M-specific TR-box, but we were unable to demonstrate simultaneous binding of the two HMG-box proteins. This may suggest a transient character or instability of a ternary complex under our conditions. Alternatively, Mat1-Mc and Ste11

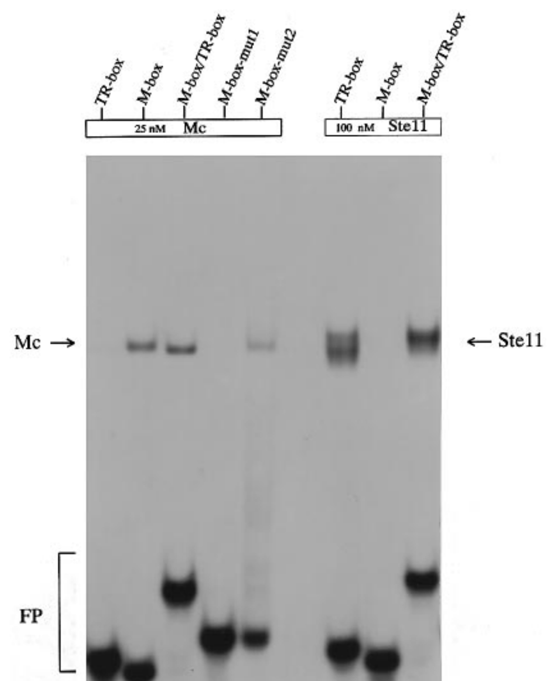


Fig. 6. Mat1-Mc binds more strongly to an M-box-containing probe than to the *mfm1* TR-box. EMSA on TR-box- or M-box-containing probes using *E.coli*-expressed malE–Mat1-Mc and GST–Ste11 proteins. The mutations M-box mut1 and M-box mut2 are shown in Figure 1.

Table II. Comparison of sequences adjacent to the TR-box in M cell-specific genes

Gene	Sequence
mam1[a]	TTCATTTTGTATGGTGGACA**ACAATG**GAGAGTACCT**AACAAAG**GCCCATTGTGTAC
mam2[a]	ACTTTTGAG**ACA**TA**G**AAGTGTTTTCTGGAAATTCTA**AACAAAGA**GGAATTATTGGC
mfm1[a]	TGAGTATTA**ACAATT**GACTAG**ACAATG**GGTCCGACC**AACAAAGAA**GTCTCAGTTTT
mfm2[a]	CCAGAATTAA**ACAATG**GGTCAA**ACAAT**AGGCAAATG**AACAAAGA**GATCACAGTTTC
mfm3[b]	TCAGTTGTA**ACAAT**TAACTAG**ACAAT**AGGCCCAACC**AACAAAGAA**GTCTCAGATTT
sxa2[b]	GTCCATTGTTT**ACAAT**CAACA**ACAAT**AGAGATGGGC**AACAAAGA**CACCCAGCGAAG

[a]Expression has been shown to be induced by nitrogen starvation.
[b]Efficient expression requires both nitrogen starvation and a pheromone signal.

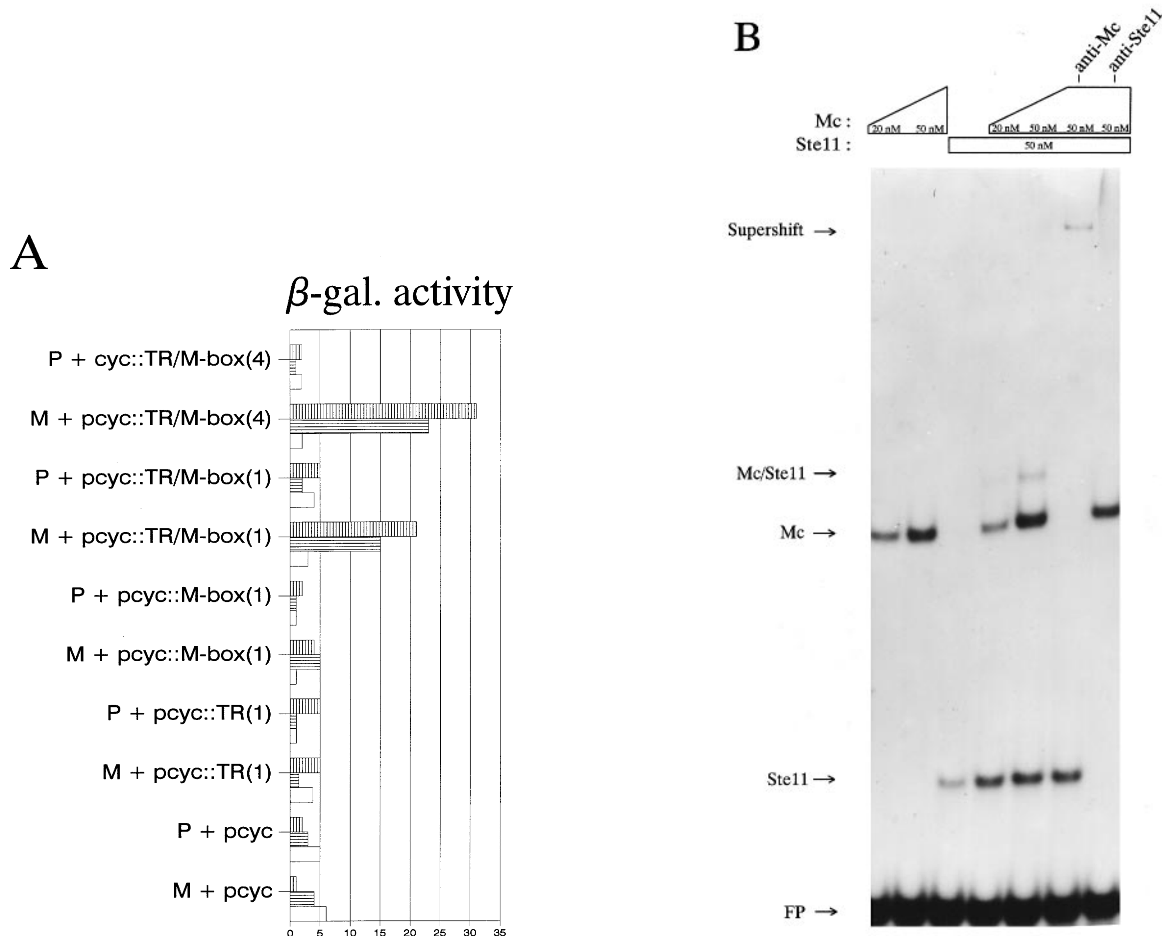


Fig. 7. (**A**) Synergistic function of the TR-box and M-box of *mfm1* on the *CYC1* minimal promoter. One or four copies of an oligonucleotide containing the most downstream M-box and the TR-box of *mfm1* were inserted in pSPΔ178, giving pcyc-TR-M-box (1/4). The constructs were transformed into h^- (M) and h^+ (P) strains and assayed for β-galactosidase activities as described in Figure 3. (**B**) Mat1-Mc and Ste11 form a stable ternary complex with the *mfm1* TR-box combined with its upstream M-box. In an EMSA, a labelled probe containing the TR-box and the most downstream M-box was incubated with purified Ste11 protein (50 nM) and increasing amounts of malE–Mat1-Mc (20–50 nM). In the last two lanes, anti-malE antibodies or anti-Ste11 antibodies were added to the binding reaction after 20 min of incubation.

may bind to the DNA in a sequential order, with Mat1-Mc binding first and changing the DNA conformation in a way that increases the affinity for Ste11, which then replaces Mat1-Mc. Supporting this model, binding of Mat1-Mc is known to induce strong bending of the DNA (Dooijes *et al.*, 1993) and, in this study, we show that binding of Mat1-Mc creates a strong hypersensitive site in the *mfm1* leader a few base pairs downstream of the TR-box, indicating a distortion of the DNA helix. Thus, one may speculate that the binding of Mat1-Mc to the TR-box prises open the minor groove in preparation for the loading of the Ste11 protein. In most cases, the HMG-box proteins actually possess considerably greater binding affinities for distorted DNA structures than they do for normal B-form DNA (Bustin and Reeves, 1996). For instance, the sequence-specific human SRY protein has even greater affinity for four-way junction DNAs than for its normal recognition sequence in B-form DNA (Ferrari *et al.*, 1992).

Several copies of this TR-box are needed for M-specific activation of a heterologous promoter. Thus, other factors are likely to be involved in enhancing the activity of the single TR-box found in the endogenous promoters. In agreement with this, we defined a region upstream of the TR-box important for full activity of the *mfm1* gene. This region is placed adjacent to the TR-box and contains two sites (ACAAT) to which Mat1-Mc can bind, hence the name M-box. Five of the six known M-specific genes habour two M-boxes and they are always situated 14–16 or 24–26 bp away from the inverted TR-box, indicating that the distance between these two elements may be important for promoter function. Mat1-Mc apparently binds much more strongly to the M-box than to the TR-box of *mfm1*. Importantly, when we combined this M-box with its downstream TR-box, we conferred M-specific expression to a heterologous promoter, even when only one copy of this construct was present (Figure 7A). Given the fact that a single TR-box or a single M-box had no effect on the same promoter, this implies that the two HMG-box-binding sites work in a synergistic fashion. Moreover, *in vitro*, Mat1-Mc and Ste11 form a stable ternary complex with an oligonucleotide containing both the M- and TR-box of *mfm1* (Figure 7B). Since the M-box is the preferred binding site of Mat1-Mc and therefore important for efficient expression from the promotor, it may seem strange that multimerization of the *mfm1* TR-box confers M-cell specificity to a heterologous promoter. However, Mat1-Mc also binds to the TR-box with a

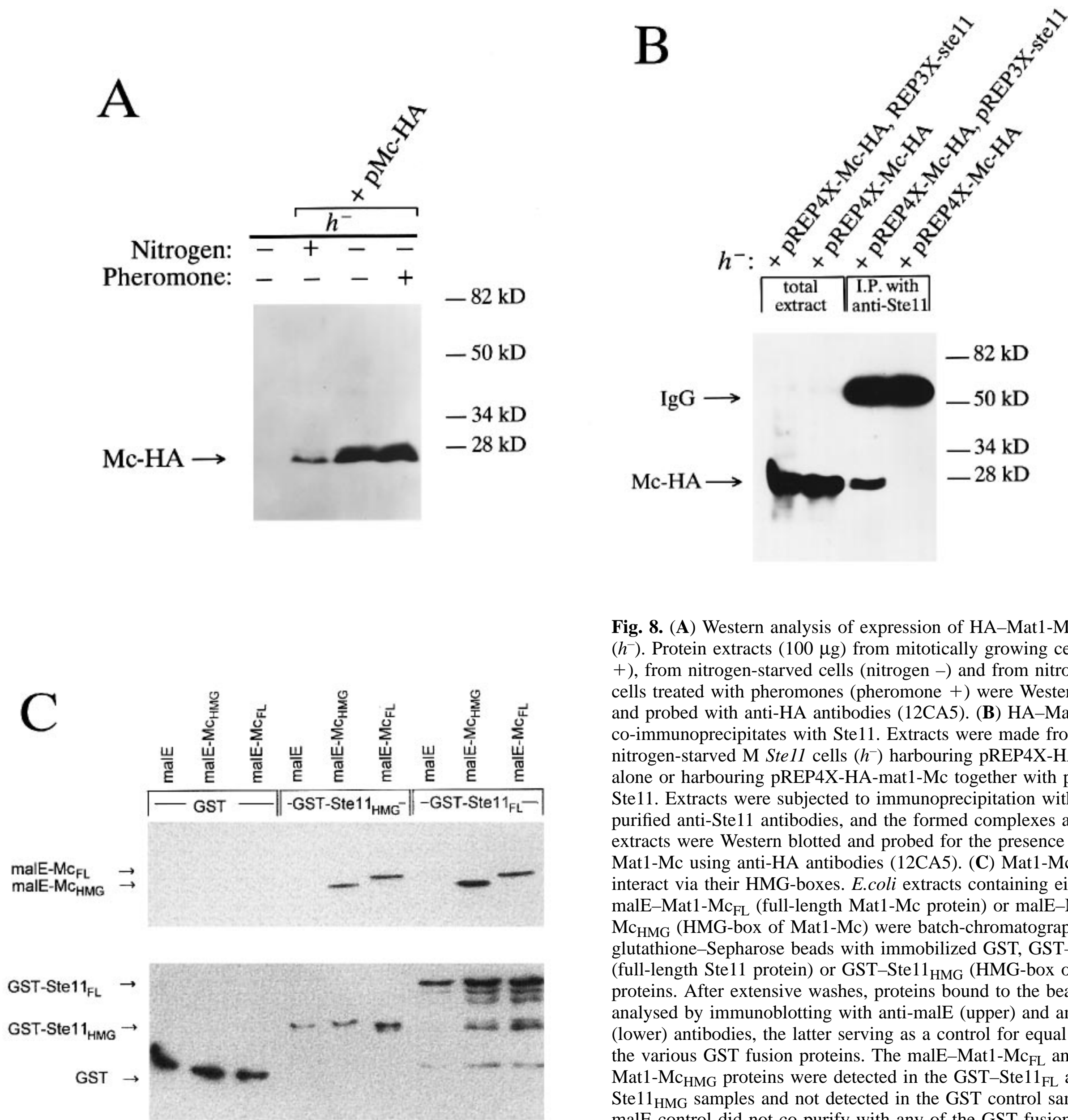


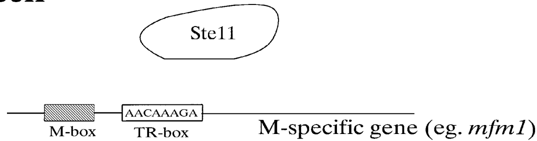
Fig. 8. (**A**) Western analysis of expression of HA–Mat1-Mc in M cells (h^-). Protein extracts (100 μg) from mitotically growing cells (nitrogen +), from nitrogen-starved cells (nitrogen –) and from nitrogen-starved cells treated with pheromones (pheromone +) were Western blotted and probed with anti-HA antibodies (12CA5). (**B**) HA–Mat1-Mc co-immunoprecipitates with Ste11. Extracts were made from nitrogen-starved M *Ste11* cells (h^-) harbouring pREP4X-HA-Mat1-Mc alone or harbouring pREP4X-HA-mat1-Mc together with pREP3X-Ste11. Extracts were subjected to immunoprecipitation with affinity-purified anti-Ste11 antibodies, and the formed complexes and total extracts were Western blotted and probed for the presence of HA-Mat1-Mc using anti-HA antibodies (12CA5). (**C**) Mat1-Mc and Ste11 interact via their HMG-boxes. *E.coli* extracts containing either malE, malE–Mat1-Mc_{FL} (full-length Mat1-Mc protein) or malE–Mat1-Mc_{HMG} (HMG-box of Mat1-Mc) were batch-chromatographed on glutathione–Sepharose beads with immobilized GST, GST–$Ste11_{FL}$ (full-length Ste11 protein) or GST–$Ste11_{HMG}$ (HMG-box of Ste11) proteins. After extensive washes, proteins bound to the beads were analysed by immunoblotting with anti-malE (upper) and anti-GST (lower) antibodies, the latter serving as a control for equal loading of the various GST fusion proteins. The malE–Mat1-Mc_{FL} and malE–Mat1-Mc_{HMG} proteins were detected in the GST–$Ste11_{FL}$ and GST–$Ste11_{HMG}$ samples and not detected in the GST control sample. The malE control did not co-purify with any of the GST fusion proteins.

relatively high affinity, and the presence of several juxtaposed TR-boxes may therefore substitute for the presence of the M-box.

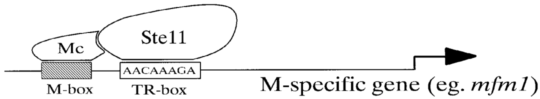
The molecular mechanism by which the two HMG-box proteins activate transcription from the *mfm1* TR-box is presently unknown. Mat1-Mc seems to have dual roles in establishing M-specific transcription. First, it can enhance the binding of Ste11 to the TR-box. Secondly, it binds to M-boxes placed adjacent to the TR-box. What is the function of this binding? One possibility is that it may serve to stabilize the binding of Ste11 to the TR-box. However, the half-life of the ternary Ste11–Mat1-Mc–M-TR-box complex, as measured by competition of binding in EMSAs, is, like that of the binary Ste11–TR-box complex, very short ($<$10 s, data not shown). Alternatively, binding of Mat1-Mc to the M-boxes may enhance the potential of Ste11 to work as a transcriptional activator. If Mat1-Mc binding is abolished or just weakened by mutagenizing one of the M-boxes, the expression of *mfm1* is clearly reduced, consistent with a contribution from M-box-bound Mat1-Mc to transcriptional activation. It was reported recently that the *S.cerevisiae* HMG-box protein NHP6A/B can potentiate promoter-specific transcriptional activation (Paull *et al.*, 1996).

Under circumstances where binding of Mat1-Mc to the M-box is reduced or abolished, transcription becomes dependent on a pheromone signal. This is true for the two M-specific genes, *sxa2* and *mfm3*, which seem to habour weaker M-boxes (Table II), and this was also the case when we mutagenized one of the *mfm1* M-boxes (Figure 1B, M-box mut2). Similarly, when we inserted a TR-box from ubiquitously expressed genes in the *mfm1* promotor, and the gene became expressed in P-cells, this transcription was dependent on a pheromone signal. In P-cells, the Mat1-Mc protein cannot bind to the M-box because it is absent. These findings suggest that Ste11 only activates transcription efficiently if Mat1-Mc is bound to an adjacent M-box or if it is modified somehow by stimulation of the

P cell



M cell



M and P cells

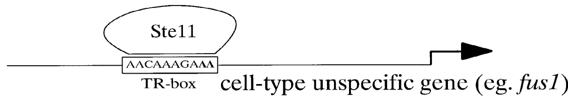


Fig. 9. Model for transcriptional control of M-specific genes. In P cells, the Ste11 transcription factor cannot bind to the particular TR-box sequences found in M-specific genes. Hence, they are not transcribed. In M cells, Mat1-Mc stimulates the binding of Ste11 to the weak TR-boxes of the M-specific genes by interaction with Ste11 and binds to the M-box(es). Consequently, the M-specific genes are transcribed. Ubiquitously expressed genes, directly controlled by Ste11, habour strong TR-boxes that Ste11 can bind without the Mat1-Mc function. Hence, these genes are transcribed in both cell types. Arrow shows that transcription is activated. A strong TR-box is indicated by the presence of two extra As. M-boxes are hatched.

pheromone signal pathway. Hence, the pheromone signal appears to contribute to transcriptional activation through the TR-box, as previously suggested by Aono *et al.* (1993). Recently, a transcription factor Map1, related to *S.cerevisae* MCM1, was identified (Nielsen *et al.*, 1996; Yabana and Yamamoto, 1996) and, although its primary function is to activate P-specific genes, the expression of M-specific genes is also severely reduced in a *map1* mutant (Yabana and Yamamoto, 1996; unpublished data). Hence, one may speculate that Ste11 interacts with Map1 to promote expression of the M-specific genes in response to a pheromone signal.

Stimulation of sequence-specific DNA binding of a transcriptional activator by interaction with an HMG-box protein may be a quite general phenomenon. In recent experiments, it was shown that HMG1 and HMG2 stimulate the binding of, respectively, the HOX9 and Oct2 transcription factors to their cognate binding sites (Zwilling *et al.*, 1995; Zappavigna *et al.*, 1996). However, to our knowledge, this is the first report describing that one sequence-specific HMG-box protein (Mat1-Mc) stimulates the binding of another sequence-specific HMG-box protein (Ste11). It has been shown that the HMG-box of HMG1 and HMG2 behaves as both a DNA-binding module and a protein–protein interaction surface (Zwilling *et al.*, 1995; Zappavigna *et al.*, 1996). We find that Mat1-Mc and Ste11 co-immunoprecipitate, and our *in vitro* observations suggest that the interaction is mediated by their HMG-boxes.

Intriguingly, it has been suggested that proper testis formation in mammals requires both of the two HMG-box proteins SRY and SOX9, and that interaction between these two proteins could be crucial for male-specific gene expression (Foster *et al.*, 1994; Wagner *et al.*, 1994; Wright *et al.*, 1995; Südbeck *et al.*, 1996). The way in which Mat1-Mc and Ste11 activate cell-specific transcription in fission yeast may provide a model for how SRY and SOX9 in cooperation turn on the male-specific genes. Ste11 and SOX9 both have functions additional to sex determination; Ste11 regulates generally expressed genes during mating (Sugimoto *et al.*, 1991; Petersen *et al.*, 1995) and SOX9 is involved in bone formation (Foster *et al.*, 1994; Wagner *et al.*, 1994; Wright *et al.*, 1995). SRY may have a role equivalent to Mat1-Mc, recruiting the SOX9 transcription factor to male-specific genes.

Materials and methods

Yeast strains, genetic procedures and media

The *S.pombe* strains used had the following genotypes. EG328 h^{90} *smt-O ura4-D18* (Styrkársdóttir *et al.*, 1993); EG432 h^{+} *Δmat2,3::LEU2 ura4-D18* (this study); EG494 h^{90} *ste11 leu1 ade6-M21* (Sugimoto *et al.*, 1991); EG544 h^{-} *Δmat2,3::LEU2 leu1*$^{+}$ (Kjærulff *et al.*, 1994); EG545 h^{+} *Δmat2,3::LEU2 leu1*$^{+}$ (Nielsen *et al.*, 1992); EG850 h^{90} *Δste11::LEU2 leu1 ura4-D18* (this study); and EG950 h^{-} *ste11 leu1 ura4-D18* (this study). All experiments in liquid culture were carried out in MSL (Egel *et al.*, 1994) at 30°C, starting with a cell density of 2×10^{6} cells/ml. To induce sexual differentiation, cells were shifted to nitrogen-deficient MSL (MSL lacking arginine) and incubated for 3–5 h. Pheromone signals were provided by mixing the transformants with either P or M cells, depending on the mating type of the transformants, followed by 3–5 h starvation in nitrogen-deficient MSL. To induce expression from the *nmt* promoter (Maundrell, 1990, 1993), transformants were grown in MSL containing 6 μM thiamine, then shifted to fresh medium lacking thiamine and grown for 14 h before starting induction of sexual differentiation.

Standard genetic procedures were carried out as described by Moreno *et al.* (1991). DNA manipulations were performed according to Sambrook *et al.* (1989), and amplification by PCR was performed as described by Kocher *et al.* (1989), using *Taq* or *Pfu* polymerase.

Construction of the mfm1–lacZ fusion gene

A 2.2 kb portion of the upstream region of the *mfm1* gene was amplified by PCR, using the oligonucleotides 5′GGCGGTACCGGTAGCCAT-TGAGTCCAT and 5′CGCGGATCCAGCTTATCGATACCG. Reaction products were cut with *Bam*HI and *Kpn*I, and ligated into *Bam*HI–*Kpn*I-digested pDW232 (Weilguny *et al.*, 1991) harbouring the *E.coli lacZ* gene, making a fusion of the sixth codon of *mfm1* in-frame with the ninth codon of *lacZ*. Most of the *ars1* sequence of this plasmid was removed by partial digestion with *Mlu*I and *Xba*I, producing pSK18. pSK18 was integrated at the *mfm1* locus via *Sma*I-directed integration. Correct integration was confirmed by PCR.

Mapping the transcriptional start point of the mfm1–lacZ fusion gene

The oligoenucleotide 5′CAGCACTGACCCTTTTGGGACCGC (#1224, New England, Biolabs), which is complementary to the *lacZ* mRNA sequence –47 bp downstream of the initiating ATG, was end-labelled with ^{32}P using T4 polynucleotide kinase. Ten μg of total RNA, extracted from nitrogen-starved cells transformed with pSK18, were mixed with an excess of labelled oligonucleotide. Subsequently, primer extension analysis was performed as described by Petersen (1987). Reaction products were analysed by electrophoresis on a 6% polyacrylamide

sequencing gel, adjacent to the sequence of the *mfm1–lacZ* fusion primed from the same oligonucleotide.

Mutational analysis of the mfm1–lacZ fusion gene

Point mutations were introduced in the *mfm1–lacZ* fusion gene by sequence overlap extension PCR (Ho *et al.*, 1989). The following oligonucleotides were used: TR-mut1, 5′CTGAGACTTCTTTTTTG-GTCGGAC together with 5′GTCCGACCAAAAAAGAAGTCTCAG; TR-mut2, 5′CCAACAAAGAAATCTCAGTTTTTT together with 5′AAAAAACTGAGATTTCTTTGTTGG; TR-mut3, 5′AGAACAAA-GAAATCTCAGTTTTTTTAAAAC together with 5′GAGATTTCTT-TGTTCTTCGGACCCATTGTCT; M-box-mut1, 5′CAATTGACTA-GAAAATGGGTCCGACCAACA together with 5′TGTTGGTCGGAC-CCATTTTCTAGTCAATTG; and M-box-mut2, 5′CAATTGACTAGAC-AATAGGTCCGACCAACA together with 5′TGTTGGTCGGACC-TATTGTCTAGTCAATTG. All *mfm1–lacZ* constructs were integrated in the *mfm1* locus at the *Sma*I site.

CYC1–lacZ constructs

The *CYC1–lacZ* reporter plasmids were constructed by inserting various copies of annealed oligonucleotides into the *Xho*I restriction site of the vector pSPΔ178 (Lowndes *et al.*, 1992). The following oligonucleotides were used: pcyc::TR-box, 5′TCGAGCCGACCAACAAAGAAGTCTCG annealed to 5′TCGACGAGACTTCTTTGTTGGTCGGC; pcyc::M-box, 5′TCGAGACTAGACAATGGGTCCGAC annealed to 5′TCGAGTCG-GACCCATTGTCTAGTC; and pcyc::TR/M-box, 5′TCGAGACTAGAC-AATGGGTCCGACCAACAAAGAAGTCTCG annealed to 5′TCGA-CGAGACTTCTTTGTTGGTCGGACCCATTGTCTAGTC.

HA tagging of Mat1-Mc

The open reading frame (ORF) of *ma1-Mc* was amplified by PCR using p*McNde*I (Kelly *et al.*, 1988) as template and the two oligonucleotides 5′GGTCTTCTGCTGACATTATTAGTA and 5′CGCGGATCCTACCA-TAAATATTAC. Reaction products were digested with *Nde*I, end-filled with Klenow fragment, cut with *Bam*HI and ligated into *Hin*cI–*Bam*HI-digested Bluescript II SK (Stratagene). From this plasmid, an *Xho*I–*Bam*HI fragment, containing the ORF of *mat1-Mc*, was ligated into *Sal*I–*Bgl*II-digested pSLF173 (kindly provided by S.Forsburg), producing a triple HA-tagged version of Mat1-Mc expressed from the *nmt* promoter (pSK94). This construct complements mutations in the *mat1-Mc* gene. A triple HA-tagged version of Mat1-Mc expressed from its own promoter was obtained in the following way. pSK94 was digested with *Xho*I, end-filled with Klenow fragment, cut with *Hin*dIII and the resulting fragment was ligated into p*McNde*I partially digested with *Nde*I, end-filled with Klenow fragment and cut with *Hin*dIII. Subsequently, a *Hin*dIII–*Sac*I fragment from this plasmid was ligated into *Hin*dIII–*Sac*I-digested pDW232.

Recombinant HMG-box proteins

The ORF of *ste11* was amplified by PCR using the oligonucleotides 5′GCACCCGGGTCTGCTTCTTTAACAGCC and 5′GCCGAATTCTA-TAAGTTTCTTTGT. Reaction products were digested with *Sma*I and *Eco*RI and ligated into *Sma*I–*Eco*RI-digested pGEX-2T (Pharmacia), producing pGEX-Ste$11_{1–468}$. A fragment encoding amino acids 1–113 of Ste11 was obtained by *Eco*RI–partial *Pst*I digestion of pGEX-Ste$11_{1–468}$. It was end-filled with Klenow fragment and religated, giving pGEX-Ste$11_{1–113}$. The plasmids were transformed into the DH5 *E.coli* strain. Cells were grown in NY containing 100 μg/ml ampicillin. Mid-log phase cultures were induced by addition of 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). After 3 h of incubation, the cells were harvested and resuspended in phosphate-buffered saline (PBS). Samples were freeze/thawed twice in liquid N_2, and sonicated six times for 1 min. Cellular debris was removed by centrifugation, and soluble GST–Ste11 protein was bulk-purified using glutathione–Sepharose (Pharmacia) as described by Ausubel *et al.* (1993). Purification and the size of proteins were tested by SDS–PAGE followed by Coomassie staining.

Production of recombinant malE–Mat1-Mc fusion proteins has been described elsewhere (Dooijes *et al.*, 1993).

Purified bacterially produced GST–Ste$11_{1–468}$ fusion protein (50 μg) was used to raise polyclonal antibodies in mice according to Harlow and Lane (1988). Polyclonal anti-Ste11 antibodies were affinity purified by incubating the serum with a nitrocellulose membrane blotted with the antigen, as described by Olmsted (1981).

Protein extraction

Soluble protein extracts were prepared from 1×10^8 cells, collected by centrifugation and resuspended in lysis buffer [50 mM Tris, pH 7.5, 100 mM NaCl, 1 mM EDTA, 5 mM $MgCl_2$, 1 mM dithiothreitol (DTT), 0.25% NP-40, 10 mM NaF, 40 mM β-glycerophosphate, 200 μM orthovanadate, 40 μM pNNP, 0.1 mM phenylmethylsulfonyl fluoride (PMSF), 3.4 μg/ml aprotinin]. One volume of acid-washed glass beads was then added, and cells were vortexed in six 30 s cycles or until ~90% lysis was obtained. After cell breakage, 400 μl of lysis buffer was added and the soluble protein fraction was recovered by three centrifugations of 10 min at 20 000 *g*. The protein concentration of the samples was determined by the BCA assay kit (Pierce).

Western blots and immunoprecipitations

Protein extracts and immunoprecipitates were electrophoresed using a 10% SDS–PAGE. For Western analysis, 100 μg of protein extract from each sample was electroblotted to nitrocellulose (Amersham) and incubated with either 12CA5 (anti-HA antibodies) or affinity-purified anti-Ste11 antibodies. Proteins were detected using enhanced chemiluminescence (ECL; Amersham). Ste11 was immunoprecipitated from 5 mg of soluble protein extracts with 2 μg of affinity-purified anti-Ste11 antibodies by incubating for 1 h at 4°C. The formed complexes were collected with protein A–Sepharose beads (Pharmacia) by incubating for 1 h at 4°C. The pellet was washed three times with lysis buffer and resuspended in SDS sample buffer.

Glutathione–Sepharose chromatography

Ten μg of GST, GST–Ste11_{FL} and GST–Ste11_{HMG} (GST–Ste$11_{1–113}$) were each incubated with 200 μl of a 1:1 slurry of glutathione–Sepharose (Pharmacia) in 1 ml of lysis buffer (50 mM Tris, pH 7.5, 100 mM NaCl, 1 mM EDTA, 5 mM $MgCl_2$, 1 mM DTT, 0.25% NP-40, 10 mM NaF, 40 mM β-glycerophosphate, 200 μM orthovanadate, 40 μM pNNP, 0.1 mM PMSF, 3.4 μg/ml aprotinin) for 1 h and subsequently washed three times in lysis buffer. Beads bound to GST–proteins were resuspended in 500 μl of lysis buffer and incubated with 250 μg of *E.coli* extracts containing either malE, malE–Mat1-Mc_{FL} or malE–Mat1-Mc_{HMG} for 1 h at 4°C. Sepharose beads were washed five times with 1 ml of lysis buffer. Proteins bound to the beads were analysed by electrophoresis in a 10% SDS–PAGE and immunoblotted with anti-malE (New England Biolab) and anti-GST antibodies.

Electrophoretic mobility shift assay

T4 polynucleotide kinase was used to label annealed oligonucleotides with [γ-^{32}P]ATP. In a binding reaction, the recombinant GST–Ste11, Ste11 and malE–Mat1-Mc proteins (1–250 nM) were incubated in a volume of 20 μl containing 10 mM Tris (pH 8.0), 5 mM $MgCl_2$, 50 mM NaCl, 0.1% NP-40, 0.1 mM EDTA, 5% glycerol, 5 μg of bovine serum albumin (BSA) and 200 ng of poly(dIdC). After addition of 0.5 ng of probe, the reactions were left at room temperature for 20 min. Samples were electrophoresed through a 6% non-denaturing polyacrylamide gel in 0.25% TBE at room temperature.

The following oligonucleotides were used: TCF-1 box, 5′GGGAG-ACTGAGAACAAAGCGCTCTCACAC annealed to 5′CCCGTGTG-AGAGCGCTTTGTTCTCAGTCT; LEF-1 box, 5′CCCAGAGCTTC-AAAGGGTGCCCTACTTG annealed to 5′GGGCAAGTAGGGCACC-CTTTGAAGCTCT; GC-rich sequence, 5′GGGAGACTGAGCCGCG-GTCGCTCTCACAC annealed to 5′CCCGTGTGAGAGCGACCG-CGGCTCAGTCT; and TR-like box, 5′TCGAGAAAAGAGGTAT-CTTTGCTTGTCTACAC annealed to 5′TCGAGTGTAGACAAGCAA-AGATACCTCTTTTC. The oligonucleotides covering the TR-box, TR-mut1, TR-mut2, M-box, M-box-mut1, M-box-mut2 and TR/M-box are described in other sections of Materials and methods.

DNase I footprint

Solid phase DNase I footprinting was carried out as described by Sandaltzopoulos and Becker (1994). Briefly, a fragment containing the *mfm1* promoter was generated by PCR using a biotinylated oligonucleotide 5′GTTTAAGAGACGTGAGAG (annealing to position +42) and a radioactively labelled oligonucleotide 5′GAAAAGAGGTATCTTTGCT-TGTCTACA (annealing to position –180). PCR products were immobilized without prior purification on 1 mg of Dynabeads (Dynal) as instructed by the manufacturer. DNA-bound beads were washed twice in PBS and resuspended in binding buffer (10 mM Tris, pH 8.0, 5 mM $MgCl_2$, 50 mM NaCl, 0.1% NP-40, 0.1 mM EDTA, 5% glycerol, 5 μg BSA). For each reation, 50 μl of DNA-bound beads (200 c.p.s.) were mixed with 2 μg of Bluescript II SK (Strategene) and 0.1–1 μM recombinant Ste11/MalE–Mc protein. After incubation for 20 min at room temperature, 0.04 U of DNase I in 50 μl of 5 mM $CaCl_2$/10 mM $MgCl_2$ were added and, after 1 min at room temperature, the reaction was quenched with 100 μl of 4 M NaCl/100 mM EDTA. The nicked

fragments on the beads were washed once in 100 µl of 2 M NaCl/20 mM EDTA and 100 µl of TE. The beads were mixed with 4 µl of loading buffer [72% formamide, 0.04% xylene cyanol, 0.04% bromophenol blue, 7.5 mM EDTA, 40 mM NaOH (freshly diluted)]. The samples were denatured for 5 min at 76°C and electrophoresed through a 6% sequencing gel.

Acknowledgements

We thank Richard Egel for comments on the manuscript and Lee Johnston and Susan Forsburg for the generous donation of plasmids. We acknowledge Jan Castrop for his advice regarding antibody production. This work was supported by The Danish Center for Microbiology and The Danish Natural Science Research Council.

References

Aono,T., Yanai,H., Miki,F., Davey,J. and Shimoda,C. (1993) Mating pheromone-induced expression of the *mat1-Pm* gene of *Schizosaccharomyces pombe*: identification of signalling components and characterization of upstream controlling elements. *Yeast*, **10**, 757–770.

Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D, Seidman,J.G., Smith,J.A. and Struhl,K. (1993) *Current Protocols in Molecular Biology.* John Wiley & Sons, NY.

Baxevanis,A.D. and Landsman,D. (1995) The HMG-1 box protein family: classification and functional relationships. *Nucleic Acids Res.*, **23**, 1604–1613.

Bustin,M. and Reeves,R. (1996) High-mobility-group chromosomal proteins: architectural components that facilitate chromatin function. *Prog. Nucleic Acid Res. Mol. Biol.*, **54**, 35–100.

Christensen,P.U., Davey,J. and Nielsen,O. (1997) The *Schizosaccharomyces pombe mam1* gene encodes an ABC transporter mediating secretion of M-factor. *Mol. Gen. Genet.*, in press.

Davey,J. (1992) Mating pheromones of the fission yeast *Schizosaccharomyces pombe*: purification and structural characterization of M-factor and isolation and analysis of two genes encoding the pheromone. *EMBO J.*, **11**, 951–960.

Dooijes,D., van de Wetering,M., Knippels,L. and Clevers,H. (1993) The *Schizosaccharomyces pombe* mating-type gene *mat-Mc* encodes a sequence-specific DNA-binding high mobility group box protein. *J. Biol. Chem.*, **268**, 24813–24817.

Egel,R., Willer,M., Kjærulff,S., Davey,J. and Nielsen,O. (1994) Assessment of pheromone production and response in fission yeast by a halo test of induced sporulation. *Yeast*, **10**, 1347–1354.

Ferrari,S., Harley,V.R., Pontiggia,A., Goodfellow,P.N., Lovell Badge,R. and Bianchi,M.E. (1992) SRY, like HMG1, recognizes sharp angles in DNA. *EMBO J.*, **11**, 4497–4506.

Foster,J.W. *et al.* (1994) Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. *Nature*, **372**, 525–530.

Giese,K., Kingsley,C., Kirshner,J.R. and Grosschedl,R. (1995) Assembly and function of a TCR alpha enhancer complex is dependent on LEF-1-induced DNA bending and multiple protein–protein interactions. *Genes Dev.*, **9**, 995–1008.

Gubbay,J., Collignon,J., Koopman,P., Capel,B., Economou,A., Munsterberg,A., Vivian,N., Goodfellow,P. and Lovell Badge,R. (1990) A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature*, **346**, 245–250.

Harlow,E. and Lane,D. (1988) *Antibodies: A Laboratory Manual*. Cold Spring Habor Laboratory Press, Cold Spring Habor, NY.

Ho,S.N., Hunt,H.D., Horton,R.M., Pullen,J.K. and Pease,L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, **77**, 51–59.

Hughes,D.A., Fukui,Y. and Yamamoto,M. (1990) Homologous activators of ras in fission and budding yeast. *Nature*, **344**, 355–357.

Imai,Y. and Yamamoto,M. (1992) *Schizosaccharomyces pombe sxa1+* and *sxa2+* encode putative proteases involved in the mating response. *Mol. Cell. Biol.*, **12**, 1827–1834.

Imai,Y. and Yamamoto,M. (1994) The fission yeast mating pheromone P-factor: its molecular structure, gene structure, and ability to induce gene expression and G1 arrest in the mating partner. *Genes Dev.*, **8**, 328–338.

Kanoh,J., Sugimoto,A. and Yamamoto,M. (1995) *Schizosaccharomyces pombe zfs1+* encoding a zinc-finger protein functions in the mating pheromone recognition pathway. *Mol. Biol. Cell*, **6**, 1185–1195.

Kelly,M., Burke,J., Smith,M., Klar,A. and Beach,D. (1988) Four mating-type genes control sexual differentiation in the fission yeast. *EMBO J.*, **7**, 1537–1547.

Kitamura,K. and Shimoda,C. (1991) The *Schizosaccharomyces pombe mam2* gene encodes a putative pheromone receptor which has a significant homology with the *Saccharomyces cerevisiae* Ste2 protein. *EMBO J.*, **10**, 3743–3751.

Kjærulff,S., Davey,J. and Nielsen,O. (1994) Analysis of the structural genes encoding M-factor in the fission yeast *Schizosaccharomyces pombe*: identification of a third gene, *mfm3*. *Mol. Cell. Biol.*, **14**, 3895–3905.

Kocher,T.D., Thomas,W.K., Meyer,A., Edwards,S.V., Paabo,S., Villablanca,F.X. and Wilson,A.C. (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc. Natl Acad. Sci. USA*, **86**, 6196–6200.

Ladds,G., Rasmussen,E.M., Young,T., Nielsen,O. and Davey,J. (1996) The *sxa2*-dependent inactivation of the P-factor mating pheromone in the fission yeast *Schizosaccharomyces pombe*. *Mol. Microbiol.*, **20**, 35–42.

Li,P. and McLeod,M. (1996) Molecular mimicry in development: identification of *ste11*$^+$ as a substrate and *mei3*$^+$ as a pseudosubstrate inhibitor of *ran1*$^+$ kinase. *Cell*, **87**, 869–880.

Lowndes,N.F., McInerny,C.J., Johnson,A.L., Fantes,P.A. and Johnston,L.H. (1992) Control of DNA synthesis genes in fission yeast by the cell-cycle gene *cdc10+*. *Nature*, **355**, 449–453.

Maundrell,K. (1990) *nmt1* of fission yeast. A highly transcribed gene completely repressed by thiamine. *J. Biol. Chem.*, **265**, 10857–10864.

Maundrell,K. (1993) Thiamine-repressible expression vectors pREP and pRIP for fission yeast. *Gene*, **123**, 127–130.

Moreno,S., Klar,A. and Nurse,P. (1991) Molecular genetic analysis of fission yeast *Schizosaccharomyces pombe. Methods Enzymol.*, **194**, 795–823.

Nielsen,O. and Davey,J. (1995) Pheromone communication in the fission yeast *Schizosaccharomyces pombe*. *Semin. Cell Biol.*, **6**, 95–104.

Nielsen,O., Davey,J. and Egel,R. (1992) The *ras1* function of *Schizosaccharomyces pombe* mediates pheromone-induced transcription. *EMBO J.*, **11**, 1391–1395.

Nielsen,O., Friis,T. and Kjaerulff,S. (1996) The *Schizosaccharomyces pombe map1* gene encodes an SRF/MCM1-related protein required for P-cell specific gene expression. *Mol. Gen. Genet.*, **253**, 387–392.

Okazaki,N., Okazaki,K., Tanaka,K. and Okayama,H. (1991) The *ste4+* gene, essential for sexual differentiation of *Schizosaccharomyces pombe*, encodes a protein with a leucine zipper motif. *Nucleic Acids Res.*, **19**, 7043–7047.

Olmsted,J.B. (1981) Affinity purification of antibodies from diazotized paper blots of heterogeneous protein samples. *J. Biol. Chem.*, **256**, 11955–11957.

Paull,T.T., Carey,M. and Johnsen,R.C. (1996) Yeast HMG proteins NHP6A/B potentiate promoter-specific transcriptional activation *in vivo* and assembly of preinitiation complexes *in vitro*. *Genes Dev.*, **10**, 2769–2781.

Petersen,C. (1987) The functional stability of the *lacZ* transcript is sensitive towards sequence alterations immediately downstream of the ribosomal binding site. *Mol. Gen. Genet.*, **209**, 179–187.

Petersen,J., Weilguny,D., Egel,R. and Nielsen,O. (1995) Characterization of *fus1* of *Schizosaccharomyces pombe*: a developmentally controlled function needed for conjugation. *Mol. Cell. Biol.*, **15**, 3697–3707.

Pontiggia,A., Rimini,R., Harley,V.R., Goodfellow,P.N., Lovell Badge,R. and Bianchi,M.E. (1994) Sex-reversing mutations affect the architecture of SRY–DNA complexes. *EMBO J.*, **13**, 6115–6124.

Sambrook,J., Fritsch,E. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd edn. Cold Spring Habor Laboratory Press, Cold Spring Habor, NY.

Sandaltzopoulos,R. and Becker,P.B. (1994) Solid phase DNase I footprinting: quick and versatile. *Nucleic Acids Res.*, **22**, 1511–1512.

Schafer,A.J. and Goodfellow,P.N. (1996) Sex determination in humans. *Bioassays*, **18**, 955–963.

Sinclair,A.H. *et al.* (1990) A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature*, **346**, 240–244.

Styrkársdottir,U., Egel,R. and Nielsen,O. (1993) The smt-0 mutation which abolishes mating-type switching in fission yeast is a deletion. *Curr. Genet.*, **23**, 184–186.

Sudbeck,P., Schmitz,M.L., Baeuerle,P.A. and Scherer,G. (1996) Sex reversal by loss of the C-terminal transactivation domain of human SOX9. *Nature Genet.*, **13**, 230–232.

Sugimoto,A., Iino,Y., Maeda,T., Watanabe,Y. and Yamamoto,M. (1991) *Schizosaccharomyces pombe ste11+* encodes a transcription factor with an HMG motif that is a critical regulator of sexual development. *Genes Dev.*, **5**, 1990–1999.

Sugiyama,A., Tanaka,K., Okazaki,K., Nojima,H. and Okayama,H. (1994) A zinc finger protein controls the onset of premeiotic DNA synthesis of fission yeast in a Mei2-independent cascade. *EMBO J.*, **13**, 1881–1887.

Toda,T., Shimanuki,M. and Yanagida,M. (1991) Fission yeast genes that confer resistance to staurosporine encode an AP-1-like transcription factor and a protein kinase related to the mammalian ERK1/MAP2 and budding yeast FUS3 and KSS1 kinases. *Genes Dev.*, **5**, 60–73.

Travis,A., Amsterdam,A., Belanger,C. and Grosschedl,R. (1991) LEF-1, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor alpha enhancer function. *Genes Dev.*, **5**, 880–894.

van de Wetering,M., Oosterwegel,M., Dooijes,D. and Clevers,H. (1991) Identification and cloning of TCF-1, a T lymphocyte-specific transcription factor containing a sequence-specific HMG box. *EMBO J.*, **10**, 123–132.

van Houte,L.P., Chuprina,V.P., van der Wetering,M., Boelens,R., Kaptein,R. and Clevers,H. (1995) Solution structure of the sequence-specific HMG box of the lymphocyte transcriptional activator Sox-4. *J. Biol. Chem.*, **270**, 30516–30524.

Wagner,T. *et al.* (1994) Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. *Cell*, **79**, 1111–1120.

Watanabe,Y., Lino,Y., Furuhata,K., Shimoda,C. and Yamamoto,M. (1988) The *S.pombe mei2* gene encoding a crucial molecule for commitment to meiosis is under the regulation of cAMP. *EMBO J.*, **7**, 761–767.

Waterman,M.L., Fischer,W.H. and Jones,K.A. (1991) A thymus-specific member of the HMG protein family regulates the human T cell receptor C alpha enhancer. *Genes Dev.*, **5**, 656–669.

Weilguny,D., Praetorius,M., Carr,A., Egel,R. and Nielsen,O. (1991) New vectors in fission yeast: application for cloning the *his2* gene. *Gene*, **99**, 47–54.

Weir,H.M., Kraulis,P.J., Hill,C.S., Raine,A.R., Laue,E.D. and Thomas,J.O. (1993) Structure of the HMG box motif in the B-domain of HMG1. *EMBO J.*, **12**, 1311–1319.

Werner,M.H., Huth,J.R., Gronenborn,A.M. and Clore,G.M. (1995) Molecular basis of human 46X,Y sex reversal revealed from the three-dimensional solution structure of the human SRY–DNA complex. *Cell*, **81**, 705–714.

Wright,E., Hargrave,M.R., Christiansen,J., Cooper,L., Kun,J., Evans,T., Gangadharan,U., Greenfield,A. and Koopman,P. (1995) The Sry-related gene Sox9 is expressed during chondrogenesis in mouse embryos. *Nature Genet.*, **9**, 15–20.

Yabana,N. and Yamamoto,M. (1996) *Schizosaccharomyces pombe map1+* encodes a MADS-box-family protein required for cell-type-specific gene expression. *Mol. Cell. Biol.*, **16**, 3420–3428.

Zappavigna,V., Falciola,L., Citterich,M.H., Mavilio,F. and Bianchi,M.E. (1996) HMG1 interacts with HOX proteins and enhances their DNA binding and transcriptional activation. *EMBO J.*, **15**, 4981–4991.

Zwilling,S., Konig,H. and Wirth,T. (1995) High mobility group protein 2 functionally interacts with the POU domains of octamer transcription factors. *EMBO J.*, **14**, 1198–1208.

Received on February 10, 1997; revised on April 2, 1997