Basal association of a transcription factor favors early gene expression

Sandrine Pinheiro, Vincent Vincenzetti, Yves Dusserre, Serge Pelet

Department of Fundamental Microbiology, University of Lausanne
1015 Lausanne, Switzerland.

*Correspondance : serge.pelet@unil.ch

Abstract

Response to extracellular signals via Mitogen-Activated Protein Kinase (MAPK) pathways regulate complex transcriptional programs where hundreds of genes are induced at a desired level with a specific timing. Gene expression regulation is largely encoded in the promoter of the gene, which harbors numerous transcription factor binding sites. In the mating MAPK pathway of *Saccharomyces cerevisiae*, one major transcription factor, Ste12, controls the chronology of gene expression necessary for the fusion of two haploid cells. Because endogenous promoters encode a wide diversity of Ste12 binding sites (PRE), synthetic promoters were engineered to decipher the rules that dictate mating gene induction. The conformation of PRE dimers that allow efficient gene expression were identified. The strength of binding of Ste12 to the PRE and the distance of the binding sites to the core promoter modulate the level of induction. The speed of activation is ensured by placing a dimer of PRE in a nucleosome depleted region favoring a basal association of Ste12 prior to the stimulus.

Introduction

De novo protein synthesis plays a central role in all cellular functions. This process can be controlled by internal regulatory input emanating from the cell cycle machinery[1,2], from fluctuations in circadian rhythms[3] or from oscillations in the metabolic state[4]. In addition, extracellular cues such as stresses, nutrients or hormones can stimulate gene expression. In all eukaryotic cells, Mitogen-Activated Protein Kinase (MAPK) pathways play an essential role in

transducing extracellular information into a cellular response, which generally includes the production of new proteins[5,6]. The gene induction process may be transient, in order to adapt to a new stressful environment[7]. However, if the protein production is sustained, it profoundly alters the cellular physiology by altering its entire proteome[8]. Cell fate decision systems implicated in cellular differentiation mechanisms rely on this *de novo* protein production to transform naive pluripotent cells into differentiated cells that will ultimately give rise to the different parts of a multicellular organism.

Despite their simplicity, unicellular organisms can also take complex decisions. The budding yeast *Saccharomyces cerevisiae* induces diverse transcriptional programs via MAPK cascades that allow this microorganism to make appropriate decisions in response to a specific stimulus. As an illustration, in low nutrient conditions, both haploid and diploid cells can alter their growth pattern to form pseudohyphae[9]. Under more drastic nutrient limitations, diploid cells begin to sporulate[10]. In rich medium and in the presence of a mating partner, haploid cells can commit to mating to ultimately generate a diploid cell[11,12].

During the mating process, cells communicate by secreting pheromones (a- or α-factor). At the cell surface, binding of pheromones stimulates a G-protein coupled receptor which in turn activates the three-tiered kinase cascade[11,12] (Supplementary Figure 1). The MAPKs, Fus3 and Kss1, release the inhibition of Dig1 and Dig2 on the transcription factor (TF) Ste12[13]. This step is critical for activating the mating transcriptional program, resulting in the up-regulation of more than 200 genes[14]. The transcriptional response promotes the arrest of the cell cycle and the formation of the mating projection, two processes which are necessary to ensure a robust mating of the partner cells.

The proteins involved in the various stages of the mating process are tightly regulated in their expression levels and dynamics by Ste12. Genes involved in the early phase of mating, such as the establishment of a pheromone gradient (BAR1), MAPK signal transduction (FUS3, STE12), cell cycle arrest (FAR1) and cell agglutination (AGA1) are expressed rapidly after the detection of the pheromone[15]. In contrast, genes involved in later stages, such as karyogamy (KAR3) and membrane fusion (FIG1) will be expressed with a delay[15,16]. The mechanisms which allow a single transcription factor (Ste12) to orchestrate this chronology of gene expression remain poorly understood.

The promoter sequence upstream of the protein-coding sequence regulates the level and dynamics of transcription. The promoter combines two distinct segments: the core promoter and the regulatory region[17,18]. With a typical length of 100-200bp, the core promoter contains the TATA box which is recognized by the TATA-binding protein. This protein recruits other general transcription factors and contributes to the formation of the Pre-Initiation Complex (PIC). In yeast, the regulatory region is typically smaller than 1kb and carries Upstream Activation Sequences (UAS) recognized by TFs. Transcriptional activation is induced via the Mediator complex, which bridges the TFs and the RNA polymerase II, thereby assembling the PIC to initiate transcription[19,20].

Studies have shown that induction of mating genes requires the formation of a Ste12 homodimer on the UAS of mating promoters[21,22]. The region is recognized by the protein via the minimal DNA consensus motif TGAAAC, commonly referred to as the Pheromone Response Element (PRE)[23–25]. In addition to these consensus PREs, Ste12 also binds with lower affinity to non-consensus sites, known as PRE-like sites. In general, multiple PREs or/and PRE-like sites can be identified on the promoters of mating genes[15,26] and the arrangement of these PRE motifs controls the expression profile of a gene[15,25]. Unfortunately, defining simple rules that would allow to predict the expression pattern of mating genes is challenging because of the wide diversity in configurations present on endogenous promoters. In addition, the identification of all possible Ste12 binding sites is difficult, since it is unclear how much a PRE-like site can deviate from the consensus while retaining affinity for Ste12.

Obviously, a complex promoter architecture is not restricted to mating genes. Essentially, all endogenous promoters harbor an intricate arrangement of TF binding sites, often combining multiple TF inputs[27]. General rules determining gene expression patterns have been obtained by analyzing libraries of synthetic promoters tested for a wide diversity of binding site organization[28,29]. In general, the number of binding sites, their affinity and their distance from the core promoter can all influence the expression output. A prolonged residence time of a TF on a promoter will increase the expression output, as the chance of initiating transcription via the recruitment of the Mediator and formation of the PIC will rise[30]. However, it is not known whether a high transcriptional output is necessarily correlated with fast gene expression of whether these two parameters (i.e., strength and speed of induction) can be decoupled.

To decipher how the promoter sequence modulates gene expression dynamics, we have engineered synthetic promoters under the control of the TF Ste12. These promoters combine different PRE conformations to understand the parameters that regulate the transcriptional program for cell fate decisions during mating. The dynamics and level of protein production were quantified to assess how the orientation and the spacing between PRE pairs, their affinity and their location along the promoter influence the transcriptional response. We demonstrate that the binding affinity of active Ste12 to the promoter controls the level of induction. However, the dynamics of the expression is limited by the ability of Ste12 to associate to a promoter prior to the stimulus and can be controlled by the Ste12 binding site conformation or its access to the DNA.

Results

To measure the dynamics of gene expression in the mating pathway, we use a dynamic Protein Synthesis Translocation Reporter (dPSTR)[31]. It consists of two transcriptional units: the first one encodes a fluorescent protein and a SynZip[32] and is expressed constitutively. The second unit encodes two NLSs and the complementary SynZip and is placed under the control of a promoter of interest. Because SynZips form strong heterodimers, upon the expression of the NLSs, the fluorescent protein relocates into the nucleus (Figure 1A and B). This sensing strategy allows a rapid quantification of protein production, which can be otherwise impaired by the slow maturation kinetics of fluorescent proteins.

In our assays we combine two dPSTRs. In the yellow channel, we have a reference p*AGA1*-dPSTR^{Y}. AGA1 encodes an agglutinin that promotes cellular adhesion of mating pairs. This gene has been shown to belong to the early gene category and is expressed at high levels, similarly to the well-established p*FUS1* reporter[15,24,33,34]. In the red channel, we monitor the dPSTR^{R} controlled by a promoter of interest (Figure 1A and B). The comparison of the nuclear enrichment dynamics between the reference dPSTR^{Y} and the test dPSTR^{R} provides a very accurate measurement of the kinetics of activation of our promoter. Notably, it abolishes cell to cell fluctuations due to the cell-cycle regulated activation of the mating (Supplementary Figure 2).

Pherormone inducible synthetic promoter

Starting from the *AGA1* promoter, we first selected an alternative core promoter based on the *CYC1* promoter (Supplementary Figure 3). Although the core ensures a sizable inducibility, the

expression dynamics are delayed by 5 minutes relative to p*AGA1*. This delay reinforces the idea that the core promoter structure plays an integral role in regulating the level and dynamics of gene expression[15].

The promoter activity is controlled by the regulatory region, which can contain multiple TF binding sites. In p*AGA1*, we have identified three consensus PREs and multiple PRE-like. We had determined that two PR Es spaced by 29 bp were important for the fast and high induction of the promoter upon α-factor treatment[15]. However, when these PREs are placed in a completely synthetic context (a modified CYC1 promoter[35]), no induction is observed (Supplementary Figure 3B and C). The inducibility of the construct is recovered when the original sequence between the two PREs from p*AGA1* is included. This difference was attributed to the presence of a PRE-like site located 3bp away from the second consensus PRE[22]. Engineering a *pCYC1* UAS containing two PRE-consensus sites 3bp apart fused to the *pCYC1* core, resulted in a rapid pheromone-inducible synthetic promoter. This initial synthetic construct (pSYN_{3TT}) will be used as a reference for the systematic alterations that will be performed on the PRE sites to generate our library of Ste12-dependent promoters.

PRE-binding conformations

The engineering of the synthetic promoter has clearly demonstrated that while two PREs are required to generate an inducible promoter, not all PRE conformations result in a functional reporter. Endogenous promoters harbor a wide diversity in PRE configurations and it is difficult to predict which ones of these PRE or PRE-like sites allow formation of a Ste12 dimer and contribute to the overall expression output. Dorrity *et al.* have described the canonical 3bp tail-to-tail conformation as the most favorable binding *in vitro* for the Ste12 DNA-binding domain (DBD) fragment[22]. This PRE conformation is found on multiple endogenous promoters such as p*AGA1*, p*STE12,* p*KAR4* and p*FAR1*. However, many endogenous promoters don't harbor this element. Therefore, it is likely that other conformations can promote a functional Ste12 binding. In addition, Su *et al.* demonstrated that a head-to-tail conformation could induce gene expression, while a head-to-head positioning prevented Ste12 binding when placed in close proximity[25].

Our measurements show that with a 3bp spacing only the tail-to-tail orientation provides a strong expression output (Figure 1C). However, the promoter becomes non-functional if the distance between the PREs in tail-to-tail orientation is extended to 5 bp. Interestingly, this 5 bp spacing

allows gene expression for the head-to-tail and head-to-head conformations. While we have not identified the head-to-head conformation in the endogenous promoters investigated, the head-to-tail conformation spaced by 5 bp is readily found in numerous promoters (p*AGA1*, p*FUS1*, p*PRM1*, p*FIG1*).

Increasing the spacing between the PRE to 7 or 10 bp abolishes the induction of the dPSTR. However, two PREs spaced by 13 bp in a tail-to-tail orientation promote a strong and fast induction (Figure 1C, D and E). This functional binding can be rationalized by the fact that the DNA-helix turn corresponds to 10.5 bp[36]. Therefore, spacing the two PRE by 13 bp positions the two Ste12 proteins in a similar orientation as with the 3 bp. This spacing can be even further extended to 23 bp. Similarly, the head-to-tail or head-to-head conformations spaced by 15 bp results in functional binding of Ste12 (Figure 1C, Supplementary Figure 4). Interestingly, this 13bp tail-to-tail conformation is found in the promoter of SST2. However, the PRE spacings of 23 or 15 bp were not identified in the endogenous promoters we scrutinized.

Overall, these findings validate the use of synthetic promoters to identify functional binding site conformations. These results can be transferred to endogenous promoters to pinpoint the sites implicated in the mating-dependent induction. The promoter sequences can be analyzed to identify short-range interaction with specific spacings of 3bp for tail-to-tail orientations and 5 bp for the other orientations while including a possible increment of 10 or 20 bp corresponding to one or two DNA helix turns.

Contribution of Ste12 DNA binding domain to the promoter association.

Ste12 activates transcription via specific PRE conformations. This extended set of functional PRE pairs indicates a surprising flexibility from Ste12 to homodimerize (Figure 2A). *In vitro* data have shown that the DBD of Ste12 can dimerize to bind to two PREs[22]. However, it is difficult to imagine how the Ste12 DBD can support the various dimerization conformations identified with our synthetic promoter, which includes diverse orientations and distances. It has also been shown that the activation domain (AD) of Ste12 can multimerize[21]. Therefore, we wanted to determine if this flexible region of the protein is implicated in the stabilization of some of the Ste12 binding conformation that we have identified, for instance on PREs separated by larger distances.

To test this hypothesis, we replaced the region coding for the activation domain in the endogenous STE12 locus by a fusion between the human estrogen receptor and the VP16 activation

domain[37,38]. Using this construct (Figure 2B, Ste12-EV), Ste12-responsive genes will be activated by stimulating the cells with β-estradiol, which promotes the relocation of the Ste12-EV to the nucleus and by-passes the mating MAPK cascade. The strength of induction of the PRE containing promoters will be governed only by the binding of the Ste12 DBD. Indeed, no interaction from the EV domain is expected (Supplementary Figure 5 A and B). Thus, we can test if the absence of the Ste12 activation domain lowers the inducibility of our synthetic promoters.

Stimulating the cells expressing Ste12-EV with β-estradiol results in a potent but delayed activation of the p*AGA1*-dPSTR^{Y} and of the p*SYN*$_{3TT}$-dPSTR^{R} (Figure 2C and D). The time required to relocate Ste12 for the cytoplasm to the nucleus can contribute to this delay[38,39]. Additionally, Ste12 controls its own expression, thus the level of the chimeric transcription factor is probably lower than the ones of the endogenous protein in absence of β-estradiol.

While the p*SYN*$_{3TT}$-dPSTR^{R} is strongly induced in the Ste12-EV background, synthetic promoters bearing no PRE or a single one fail to be induced (Supplementary Figure 5 C and D). Similarly, if the PREs are spaced by 40 bp, no relocation of the dPSTR can be observed. These control experiments confirm our hypothesis that the Ste12-EV chimeric transcription factor relies solely on the DBD domain of Ste12 for activation and if the promoter contains a single PRE or 2 PREs placed in an undesired configuration, the association of Ste12-EV to the DNA is too weak to promote transcription despite the presence of the strong VP16 activation domain.

To determine if the activation domain of Ste12 plays a role in the capacity of Ste12 to promote transcription from a wide diversity of PRE conformations, we compared the strength of induction of various synthetic promoters from wild-type Ste12 and Ste12-EV (Figure 2E). Remarkably, the normalized levels of induction are similar for the two transcriptional activators for most of the promoters. The PRE spaced by 13 bp lead to a stronger normalized induction with the Ste12-EV chimeric transcription factor. This is probably due to an important variability in the induction by β-estradiol also observed for the p*AGA1*-dPSTR^{Y} (Supplementary Figure 5E). However, one interesting finding is the lower inducibility of the promoter with PREs spaced by 15 bp in head-to-tail conformation which clearly induces at a lower level with the chimeric TF, suggesting that the AD domain of Ste12 plays a role in the stabilization of this interaction. However, for all the other PRE conformations, the strength of the Ste12 dimer binding to the DNA is governed essentially by the DBD, while the AD contributes only in a limited manner to the stabilization of the Ste12 on the promoter.

Kar4 contributes to Ste12 binding

Suboptimal PRE-arrangements can prevent Ste12 from binding to the promoter and limit the expression output. However, even in the context of PRE-dimers with satisfactory arrangements, Ste12 binding on mating promoters can be affected if the binding site carries point mutations (PRE-like). Multiple endogenous promoters contain satisfactory PRE-arrangements consisting of a combination of a consensus PRE and a PRE-like site (p*AGA1*, p*FIG1*)[22].

Therefore, to test the influence of the strength of the PRE site on the expression output, two PREs spaced by 3 base pairs were used and the sequence of one of the binding sites was modified. A clear decrease in the fraction of responding cells and the level of expression are observed for the six PRE-like variants tested (Figure 3A and B). This is in line with previous measurements where the lowering of the affinity of the DNA-binding site results in a weaker gene expression output[25,28]. We can postulate that the lower affinity of the site decreases the residence time of Ste12 on the promoter and therefore limits the expression output. Interestingly, however, the dynamics of gene expression is not influenced by the lowering of the binding site affinity and all the promoter tested here despite their low induction level are induced rapidly (Figure 3C).

As hypothesized in a previous study, the association of Ste12 on promoters containing suboptimal PRE-arrangements is enhanced by Kar4[15]. Indeed, slow mating promoters show minimal expression in *kar4*Δ cells while fast ones are independent of Kar4. Therefore, to evaluate the contribution of Kar4 to the expression output of synthetic promoters containing PRE sites with different affinities for Ste12, we measured the constructs in *kar4*Δ cells. Our measurements suggest that Kar4 has a dual role. On the one hand, Kar4 contributes to the rapid activation of Ste12-bound promoters, while, on the other hand, it limits the level of induction of the promoter (Supplementary Figure 6). In line with our previous findings made on endogenous promoters, we can postulate that Kar4 plays a role in stabilizing Ste12 on PRE dimers that contain a non-consensus site. This enhanced binding enables Ste12 to be preloaded onto the promoter prior to the stimulus, facilitating rapid induction of the promoter upon activation of the MAPK cascade. However, the role of Kar4 is more complex since in the absence of this co-factor the induction level from the promoters with non-consensus PREs is enhanced. Once the Ste12 dimer is engaged in transcription, Kar4 seems to have a negative impact on the activity of Ste12 possibly by shortening the residence time of

Ste12 on the promoter or by limiting the ability of Ste12 to interact with the core transcriptional machinery.

PRE location

An additional parameter that can be modulated is the location of the PRE sites on the promoter. We have moved the two PREs from p*SYN*$_{3TT}$ from their original position at -223 from the Start codon between -183 to -413 bp. Extending the distance between the Ste12 binding sites and the core promoter leads to a general decline of the expression output of the promoter (Figure 4 A and B). A similar behavior has been observed for the Msn2 stress response TF where moving its binding sites away from the core promoter lowered the expression output of the promoter[40]. Here we demonstrate that this lower expression output is not correlated with a decrease in the speed of gene expression, since all the p*SYN* tested remain fast (Figure 4C). The binding dynamics of Ste12 should not be modified by the location of the PREs on the promoter. Thus, we postulate that the increased distance between the Ste12 binding sites and the core promoter precludes the ability of Ste12 to activate the general transcription factors via the Mediator.

The modulation of the distance between core promoter and PRE was also tested using a modified AGA1 promoter where all the consensus binding sites were mutated, and a PRE-dimer was moved from -128 to -435bp relative to the Start codon (Figure 5A). In this context, however, both the expression level and the dynamics of induction are strongly influenced by the location of the PRE-sites (Figure 5B, C and D). We believe that this behavior can be explained by the position of the nucleosomes on the promoter. Indeed, the promoter activation is rapid and strong when the PREs fall in a nucleosome-depleted region (NDR). When the sites are located within a region protected by a nucleosome, the fraction of responding cells and the level and speed of induction are all attenuated. The most significant impact is observed with the PRE-dimers located at -320bp, where Ste12 binding is conflicting with a nucleosome centered around -349bp[41]. Interestingly, the fraction of expressing cells for this promoter is low (18%), but the few cells that express display a substantial nuclear enrichment of the dPSTR (Figure 5E). This stochastic activation suggests the presence of a dynamic interplay between the binding of the nucleosomes and the Ste12 dimer. When the nucleosome is bound, no transcription takes place. In some cells, the histones can be displaced by the binding of Ste12, which remains stably associated to the promoter to induce a sizable but delayed expression.

To test this model, we placed a nucleosome disfavoring sequence (poly dA/dT) in the vicinity of the Ste12 binding sites (Figure 6A). The fraction of expressing cells increased from 20% to 50% when the dA/dT element was placed 6 bp away from the PRE sites (Figure 6B and C). An alternative option to increase the binding efficiency of Ste12 in the nucleosome bound region is to use multiple PRE sites. We used conformations with 3 PREs present in PRM1 and FUS1. The multiple sites and the high A/T content of the element facilitated the association of Ste12 to the promoter and resulted in an increased fraction of transcribing cells (Figure 6B and C). However, for all these constructs, the dynamics of induction were slow compared to the endogenous p*AGA1*. The addition of the dA/dT sequence slightly accelerated the induction of the reporter, suggesting that in a fraction of the population Ste12 can associate to the promoter under basal conditions (Figure 6D). In contrast, the presence of the three PRE from FUS3 despite their high inducibility display a slightly delayed induction compared to the two PRE.

The endogenous p*AGA1* promoter contains three consensus PREs and at least five PRE-like sites (Supplementary Figure 7A). The PRE/PRE-like sites spaced by 3 bp at position -220 alone seem sufficient to ensure a high induction level[15]. Therefore, it is not clear why there are so many additional Ste12 binding sites present in this sequence. One of their roles might be to increase the local concentration of Ste12 in the vicinity of the locus. Interestingly, if the first PRE at -196 bp or its neighboring PRE-like at -185 bp are mutated, the level of induction of the promoter remains almost identical, while the dynamics of induction are delayed (Supplementary Figure 7B and C). Since these two PREs are 5bp apart in a head-to-tail conformation, it is likely that they allow the formation of a Ste12 dimer. One alternative explanation for the delayed expression observed is that these two binding sites contribute to define the nucleosome depleted region in the p*AGA1* promoter which favors the basal association of Ste12 to the central PRE of the promoter. Taken together, these results demonstrate that nucleosomes prevent the basal association of Ste12 outside of the NDR. Sites present in a nucleosome protected region of the promoter will display slow induction kinetics and a low fraction of inducing cells. Additional PRE sites on endogenous promoters may contribute to shaping the NDR and thereby allowing indirectly an early activation of transcription.

Discussion

The coding sequence of a protein can provide insights into many of its properties. However, predicting the level and the timing of its expression based on the promoter sequence remains a challenge. Endogenous promoters display a large and complex palette of regulation combining binding sites for multiple TFs at various positions. Even in the simpler case of the mating pathway in budding yeast, where 200 genes are under the control of the single TF Ste12, the diversity in the organization of the PREs on these promoters is large. To elucidate the fundamental rules governing the induction of mating genes by Ste12, we designed synthetic promoters where the conformations of Ste12 binding sites could be tested systematically.

Based on these results we have identified various strategies that are at play in mating dependent promoters to regulate the properties of induction of a promoter. The level of induction can be controlled by two parameters: the binding affinity of the Ste12 sites and the distance from the core promoter. p*AGA1*, p*FAR1* and p*STE12* all possess 2 PRE spaced by 3bp in tail-to-tail orientation. The PRE dimers is placed at -200 bp from the start in p*AGA1* while it is located at -300 bp and -400 bp for p*FAR1* and p*STE12* which can explain their lower inducibility[15]. The strength of Ste12 association on a promoter can be tuned by point mutation in the PREs or by changing the binding conformation. The two PREs spaced by 13 bp (-250 bp from ATG) in p*SST2* results in lower induction compared to p*AGA1*. p*KAR4* and p*AGA1* both have a PRE dimer spaced by 3bp positioned ~200 bp from the start codon with one of the PRE which contains two point-mutations[22]. The stronger level of p*AGA1* induction suggests that the binding to the p*AGA1* PRE dimer is tighter. However, both promoters harbor numerous additional PRE and PRE-like sites that could all somewhat contribute to the final inducibility of the promoter by increasing the local Ste12 concentration via local transient binding.

We have shown that the fast induction of a promoter relies on its ability to bind Ste12 in basal conditions. Kar4 contributes to this fast activation by stabilizing the Ste12 dimer on weaker PRE sites. We identified two strategies to achieve a slow induction of the promoter: either by restricting the access of Ste12 to the promoter by nucleosomes or by using an unconventional PRE dimer (23bp tail-to-tail or 15bp head-to-tail or head-to-head). In the 26 mating-dependent promoters that we have analyzed, we did not identify these specific conformations. The nucleosome protection may be the preferred option because it allows to tune both the speed and level of induction, while the large spacing between PRE can only achieve a slow and low level of gene expression. When

the PRE dimer is moved within the *pAGA1* promoter, positioning the PRE dimer under the nucleosome slows down the induction but also strictly reduces the fraction of responding cells. In contrast, placing the binding sites at the edge of the nucleosome can be sufficient to obtain a similar decrease in the speed of induction without affecting severely the expression level and the fraction of responding cells. In the promoter of the late gene FIG1, a PRE dimer (5bp, tail-to-head) essential for gene expression is also placed at the boundary of the sequence protected by the nucleosome[15]. This positioning of the PRE dimers delays the induction of FIG1 until it is required for the fusion of the two mating cells without compromising its robust expression when it is needed. Positioning the PREs further into the nucleosome protected region could lead to a stochastic activation of the protein which would be detrimental for the mating outcome.

The timing of gene expression is crucial for many processes, from the rapid induction of stress response genes to the controlled induction of proteins during the cell cycle. Cell-fate decisions are characterized with a chronology of gene expression which requires the ability to tune the both the level and the dynamics of gene expression independently of each other. Regulating the basal association of Ste12 on promoters via the positioning of nucleosomes offers the opportunity to control the chronology of gene expression during the mating process. With our improved understanding of the regulation of the mating gene expression, we are in a better position to perturb the timing of key proteins implicated in mating to verify the importance of this chronology on the mating process.

References

1. McInerny, C. J. Cell Cycle Regulated Gene Expression in Yeasts. in *Advances in Genetics* vol. 73 51–85 (Elsevier, 2011).
2. Fischer, M., Schade, A. E., Branigan, T. B., Müller, G. A. & DeCaprio, J. A. Coordinating gene expression during the cell cycle. *Trends in Biochemical Sciences* **47**, 1009–1022 (2022).
3. Takahashi, J. S. Transcriptional architecture of the mammalian circadian clock. *Nat Rev Genet* **18**, 164–179 (2017).
4. Slavov, N., Macinskas, J., Caudy, A. & Botstein, D. Metabolic cycling without cell division cycling in respiring yeast. *Proceedings of the National Academy of Sciences* **108**, 19090–19095 (2011).
5. Roux, P. P. & Blenis, J. ERK and p38 MAPK-Activated Protein Kinases: a Family of Protein Kinases with Diverse Biological Functions. *Microbiology and Molecular Biology Reviews* **68**, 320–344 (2004).
6. Chen, R. E. & Thorner, J. Function and regulation in MAPK signaling pathways: lessons learned from the yeast Saccharomyces cerevisiae. *Biochim Biophys Acta* **1773**, 1311–1340 (2007).
7. Proft, M. & Struhl, K. MAP Kinase-Mediated Stress Relief that Precedes and Regulates the Timing of Transcriptional Induction. *Cell* **118**, 351–361 (2004).
8. Murphy, L. O., Smith, S., Chen, R.-H., Fingar, D. C. & Blenis, J. Molecular interpretation of ERK signal duration by immediate early gene products. *Nat Cell Biol* **4**, 556–564 (2002).
9. Cullen, P. J. & Sprague, G. F. The Regulation of Filamentous Growth in Yeast. *Genetics* **190**, 23–49 (2012).

10. Neiman, A. M. Sporulation in the Budding Yeast *Saccharomyces cerevisiae*. *Genetics* **189**, 737–765 (2011).
11. Bardwell, L. A walk-through of the yeast mating pheromone response pathway. *Peptides* **26**, 339–350 (2005).
12. Atay, O. & Skotheim, J. M. Spatial and temporal signal processing and decision making by MAPK pathways. *J Cell Biol* **216**, 317–330 (2017).
13. Tedford, K., Kim, S., Sa, D., Stevens, K. & Tyers, M. Regulation of the mating pheromone and invasive growth responses in yeast by two MAP kinase substrates. *Curr Biol* **7**, 228–238 (1997).
14. Roberts, C. J. *et al.* Signaling and circuitry of multiple MAPK pathways revealed by a matrix of global gene expression profiles. *Science* **287**, 873–880 (2000).
15. Aymoz, D. *et al.* Timing of gene expression in a cell-fate decision system. *Molecular Systems Biology* **14**, e8024 (2018).
16. Lahav, R., Gammie, A., Tavazoie, S. & Rose, M. D. Role of transcription factor Kar4 in regulating downstream events in the Saccharomyces cerevisiae pheromone response pathway. *Mol Cell Biol* **27**, 818–829 (2007).
17. Struhl, K. YEAST TRANSCRIPTIONAL REGULATORY MECHANISMS. *Annual Review of Genetics* (1995).
18. Hahn, S. & Young, E. T. Transcriptional Regulation in Saccharomyces cerevisiae: Transcription Factor Regulation and Function, Mechanisms of Initiation, and Roles of Activators and Coactivators. *Genetics* **189**, 705–736 (2011).
19. Björklund, S. & Gustafsson, C. M. The yeast Mediator complex and its regulation. *Trends in Biochemical Sciences* **30**, 240–244 (2005).
20. Soutourina, J. Transcription regulation by the Mediator complex. *Nature Publishing Group* **19**, 262–274 (2018).
21. Olson, K. A. *et al.* Two regulators of Ste12p inhibit pheromone-responsive transcription by separate mechanisms. *Mol Cell Biol* **20**, 4199–4209 (2000).
22. Dorrity, M. W., Cuperus, J. T., Carlisle, J. A., Fields, S. & Queitsch, C. Preferences in a trait decision determined by transcription factor variants. *Proceedings of the National Academy of Sciences* **115**, E7997–E8006 (2018).
23. Errede, B. & Ammerer, G. STE12, a protein involved in cell-type-specific transcription and signal transduction in yeast, is part of protein-DNA complexes. *Genes & Development* **3**, 1349–1361 (1989).
24. Hagen, D. C., McCaffrey, G. & Sprague, G. F. Pheromone response elements are necessary and sufficient for basal and pheromone-induced transcription of the FUS1 gene of Saccharomyces cerevisiae. *Mol Cell Biol* **11**, 2952–2961 (1991).
25. Su, T.-C., Tamarkina, E. & Sadowski, I. Organizational constraints on Ste12 cis-elements for a pheromone response in Saccharomyces cerevisiae. *FEBS J.* **277**, 3235–3248 (2010).
26. Chou, S., Lane, S. & Liu, H. Regulation of mating and filamentation genes by two distinct Ste12 complexes in Saccharomyces cerevisiae. *Mol Cell Biol* **26**, 4794–4805 (2006).
27. Teixeira, M. C. *et al.* The YEASTRACT database: a tool for the analysis of transcription regulatory associations in Saccharomyces cerevisiae. *Nucleic Acids Res* **34**, D446-451 (2006).
28. Sharon, E. *et al.* Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nat Biotechnol* **30**, 521–530 (2012).
29. Levo, M. & Segal, E. In pursuit of design principles of regulatory sequences. *Nat Rev Genet* **15**, 453–468 (2014).
30. de Jonge, W. J., Patel, H. P., Meeussen, J. V. W. & Lenstra, T. L. Following the tracks: How transcription factor binding dynamics control transcription. *Biophysical Journal* **121**, 1583–1592 (2022).
31. Aymoz, D., Wosika, V., Durandau, E. & Pelet, S. Real-time quantification of protein expression at the single-cell level via dynamic protein synthesis translocation reporters. *Nature Communications* **7**, 11304 (2016).
32. Reinke, A. W., Grant, R. A. & Keating, A. E. A Synthetic Coiled-Coil Interactome Provides Heterospecific Modules for Molecular Engineering. *J. Am. Chem. Soc.* **132**, 6025–6031 (2010).
33. Roy, A., Lu, C. F., Marykwas, D. L., Lipke, P. N. & Kurjan, J. The Aga1 Product Is Involved in Cell-Surface Attachment of the Saccharomyces-Cerevisiae Cell-Adhesion Glycoprotein a-Agglutinin. *Mol Cell Biol* **11**, 4196–4206 (1991).
34. Colman-Lerner, A. *et al.* Regulated cell-to-cell variation in a cell-fate decision system. *Nature* **437**, 699–706 (2005).
35. Curran, K. A. *et al.* Design of synthetic yeast promoters via tuning of nucleosome architecture. *Nat Commun* **5**, 4002 (2014).
36. Neidle, S. Beyond the double helix: DNA structural diversity and the PDB. *Journal of Biological Chemistry* **296**, 100553 (2021).

37. Louvion, J. F., Havaux-Copf, B. & Picard, D. Fusion of GAL4-VP16 to a steroid-binding domain provides a tool for gratuitous induction of galactose-responsive genes in yeast. *GENE* **131**, 129–134 (1993).
38. McIsaac, R. S. *et al.* Fast-acting and nearly gratuitous induction of gene expression and protein depletion in Saccharomyces cerevisiae. *Mol Biol Cell* **22**, 4447–4459 (2011).
39. Zechner, C., Unger, M., Pelet, S., Peter, M. & Koeppl, H. Scalable inference of heterogeneous reaction kinetics from pooled single-cell recordings. *Nat Methods* **11**, 197–202 (2014).
40. Hansen, A. S. & O'Shea, E. K. cis Determinants of Promoter Threshold and Activation Timescale. *CellReports* **12**, 1226–1233 (2015).
41. Rossi, M. J. *et al.* A high-resolution protein architecture of the budding yeast genome. *Nature* **592**, 309–314 (2021).

End Notes

Acknowledgment

We thank members of the Pelet lab for critical comments on the project. We thank Michael Taschner and Stephan Gruber for helpful discussions, Marta Schmitt, Stella Parzanese for technical help. Work in the Pelet lab is funded by the Swiss National Science Foundation (SNSF, 31003A_182431) and the University of Lausanne.

Author Contributions

SPi and SPe designed the experiments and wrote the manuscript. SPi performed the time-lapse measurements. SPe analyzed the data and wrote the manuscript. SPi, VV and YD generated the strains and plasmids used in the study.

Declaration of Interests

The authors declare no competing interests.

Figure Legends

Figure 1. Effect of the orientation and spacing of Ste12 binding sites on the expression output.

A. Scheme of the two reporters present in the strains. The reference p*AGA1*-dPSTR^{Y} and the p*SYN*$_{3TT}$-dPSTR^{R}. Both constructs encode the production of a constitutively expressed fluorescent protein and an inducible peptide containing two nuclear localization sequences (NLS) and a SynZip (SZ1 or SZ3, which can form a strong heterodimer with SZ2 and SZ4, respectively). The

reference $dPSTR^Y$ is under the control of the AGA1 promoter in the yellow channel and the test $dPSTR^R$ is regulated by a synthetic promoter of interest based on a modified CYC1 promoter.

B. Images of cells induced with 1μM α-factor at time 0. The nuclear enrichment of the fluorescent proteins serves as a measure of promoter activity.

C. Matrix representing the mean expression output for PRE sites with four different orientations and spaced between 3 to 40 bp. The strength of the red color is proportional to the expression output of the promoter. Dark gray areas represent construct where fewer than 10% of cells are expressing. Light gray squares are PRE conformations that were not measured.

D. Time course of the nuclear enrichment of the $dPSTR^R$ for various distances of PRE placed in tail-to-tail orientation. The three functional conformations are plotted in blue (3bp) light blue (13bp) and light green (23bp). The solid lines represent the median and the shaded area the 25- to 75- percentile of the population. Gray lines represent the median of non-functional PRE conformations. The black dashed line is the median of the control synthetic promoter without PREs inserted.

E. Summary graph displaying the expression output, the speed and the fraction of responding cells for various spacings of the PRE dimer placed in the tail-to-tail orientation. The color of the marker indicates the difference in response time between the synthetic promoter and the reference p*AGA1*-$dPSTR^Y$, with fast responding promoters in red and slow ones in blue as indicated by the two small schematic graphs on the left. The size of the marker represents the fraction of responding cells represented by the two schemes on the right. The expression threshold based on the level of pSYN_{3TT} is indicated by the dashed dotted line.

Figure 2. Dominant role of Ste12 DNA binding domain for PRE conformation selection.

A. Schematic of the Ste12 transcription factor composed of a DNA-binding domain and a flexible activation domain. The small schemes describe the diversity of possible Ste12 homodimer interactions when varying the orientation or the distances of PRE sites.

B. Scheme of the Ste12-EV chimeric transcription factor where the activation domain of Ste12 has been replaced by the estrogen receptor and the VP16 activation domain resulting in an β-estradiol responsive TF.

C., D. Dynamics of nuclear relocation following stimulation with 1μM β-estradiol at time 0 for the p*AGA1*-dPSTR^{Y} (C) and various p*SYN*-dPSTR^{R} (D). The solid lines represent the median of the population and the shaded area, the 25- to 75- percentiles.

E. Comparison of the inducibility of various PRE conformation with the Ste12 WT (solid border) or the Ste12-EV (dashed borders). The Expression Outputs for the synthetic reporters induced by the Ste12 WT or the Ste12-EV were normalized relative to the Expression Output of the reference pSYN_{3TT}. The bar represents the mean response of the replicates shown by the circles. The opacity of the color is proportional to the fraction of responding cells. A significant difference between the normalized EO Ste12-WT and Ste12-EV is indicated by one or two stars (t-test: ✻: p-val < 0.05 and ✻✻: p-val <0.005)

Figure 3. Influence of the PRE affinity on the induction of the promoter.

A. Dynamics of nuclear relocation for the dPSTR^{R} controlled by two PRE spaced by 3 bp in tail-to-tail orientation where one of the PRE sequences was mutated to alter its binding affinity (colored lines median of the population and shaded area 25- to 75- percentile). The solid black line represents the median response of the consensus PRE (pSYN_{3TT}), while the dashed line represents the control promoter without PRE.

B. Fraction of responding cells for the various PRE sequences. The darker portion of the bar represents the fraction of highly expressing cells (expression output > 50% of reference expression output) and the light bar the fraction of low-expressing cells (> 20% expression output < 50%). The marker represents the fraction of responding cells for individual replicates.

C. Summary graph displaying the expression output, the speed and the fraction of responding cells for promoters with mutations in one of the two PREs. The color of the marker indicates the difference in response time between the synthetic promoter and the reference p*AGA1*-dPSTR^{Y}.

The size of the marker represents the fraction of responding cells. The dashed line represents the expression output and the dashed dotted line the expression threshold calculated based on the p*SYN*$_{3TT}$.

Figure 4. Distance of the PREs to the core promoter impacts expression output in the synthetic promoter.

A. Dynamics of nuclear relocation for the dPSTR^{R} controlled by two PREs spaced by 3 bp in tail-to-tail orientation placed at different positions on the promoter (color lines represent the median of the population). The solid black line represents the median response of the PREs placed 223 bp downstream of the start codon (p*SYN*$_{3TT}$), while the dashed line represents the control promoter without PRE.

B. Fraction of responding cells for the various PRE sequences. The darker portion of the bar represents the fraction of highly expressing cells (expression output > 50% of reference expression output) and the light bar the fraction of low-expressing cells (> 20% expression output < 50%). The marker represents the fraction of responding cells for individual replicates.

C. Summary graph displaying the expression output, the speed and the fraction of responding cells for PRE placed at various distances from the START codon. The color of the marker indicates the difference in response time between the synthetic promoter and the reference p*AGA1*-dPSTR^{Y}. The size of the marker represents the fraction of responding cells. The dashed line represents the expression output and the dashed dotted line the expression threshold calculated based on the p*SYN*$_{3TT}$.

Figure 5. Inducibility of an endogenous promoter depends on the location of the nucleosomes.

A. Scheme of the modified p*AGA1* promoters with consensus PRE sites indicated by white arrows and non-consensus PRE sites by smaller gray arrows. The regions protected by the nucleosomes positioned at -100, -349 and -516 are indicated by a lighter color.

B. Dynamics of nuclear relocation for the dPSTR^{R} controlled by modified p*AGA1* promoters. Two PRE spaced by 3 bp in tail-to-tail orientation are placed at different positions along the promoter

(colored lines median of the population and shaded area 25- to 75- percentile). The solid black line represents the median response of the reference endogenous AGA1 promoter, while the dashed line represents a mutated p*AGA1* where the three consensus PREs and one non-consensus PRE are mutated.

C. Summary graph displaying the expression output, the speed and the fraction of responding cells. The color of the marker indicates the difference in response time between the synthetic promoter and the reference p*AGA1*-dPSTR^{Y}. The size of the marker represents the fraction of responding cells. The dashed line represents the expression output and the dashed dotted line the expression threshold calculated based on the p*AGA1*-dPSTR^{R}. The gray areas represent the regions of the promoters covered by the three nucleosomes.

D. Histograms of the difference in response time between the modified p*AGA1*-dPSTR^{R} and the reference p*AGA1*-dPSTR^{Y} in cells that express both constructs.

E. Thumbnail images of cells bearing the p*AGA1*-dPSTR^{Y} and the modified p*AGA1*-dPSTR^{R} with the PR dimer positioned 330 base pairs before the START codon. While all cells relocate the dPSTR^{Y}, only one cell displays a relocation of the dPSTR^{R}. The scale bar represents 5 µm.

Figure 6. Modulation of nucleosome binding by changing PRE sites or inserting poly-dA/dT sequences.

A. Scheme of the various promoters tested. Starting from the modified p*AGA1* with 2 PREs in tail-to-tail orientation at -330, poly-dA/dT segments of 20 bp were inserted 6 or 20 bp away from the 2 PREs. Alternatively, the two PREs were replaced by three PRE sites extracted from the FUS1 or PRM1 promoters.

B. Dynamics of nuclear relocation for the dPSTR^{R} controlled by modified p*AGA1* promoters. The four different modified promoters tested are shown in color. The solid black line represents the median response of the reference endogenous AGA1 promoter, while the dashed line represents the median of the population of cells with the 2 PREs in tail-to-tail orientation at -330.

C. Summary graph displaying the expression output, the speed and the fraction of responding cells. The color of the marker indicates the difference in response time between the synthetic promoter and the reference p*AGA1*-dPSTR^{Y}. The size of the marker represents the fraction of responding cells. The dashed line represents the expression output and the dashed dotted line the expression threshold calculated based on the p*AGA1*-dPSTR^{R}. The inset represents the fraction of responding cells for each promoter. The dark portion of the bar represents the fraction of highly expressing cells (Expression output > 50% EO of the reference promoter) and the light bar the fraction of low-expressing cells (20%< EO < 50%). The marker represents the fraction of responding cells for individual replicates.

D. Histograms of the response time difference measured between the modified p*AGA1*-dPSTR^{R} and the reference p*AGA1*-dPSTR^{Y}. The solid black line corresponds to the endogenous AGA1 promoter, while the dashed line represents the 2 PREs in tail-to-tail orientation at -330. The addition of the poly-dA/dT 6bp away from the PRE dimer (light green) slightly accelerates the induction of the promoter. Placing the three PREs of FUS1 (blue) slows down the induction of the promoter.

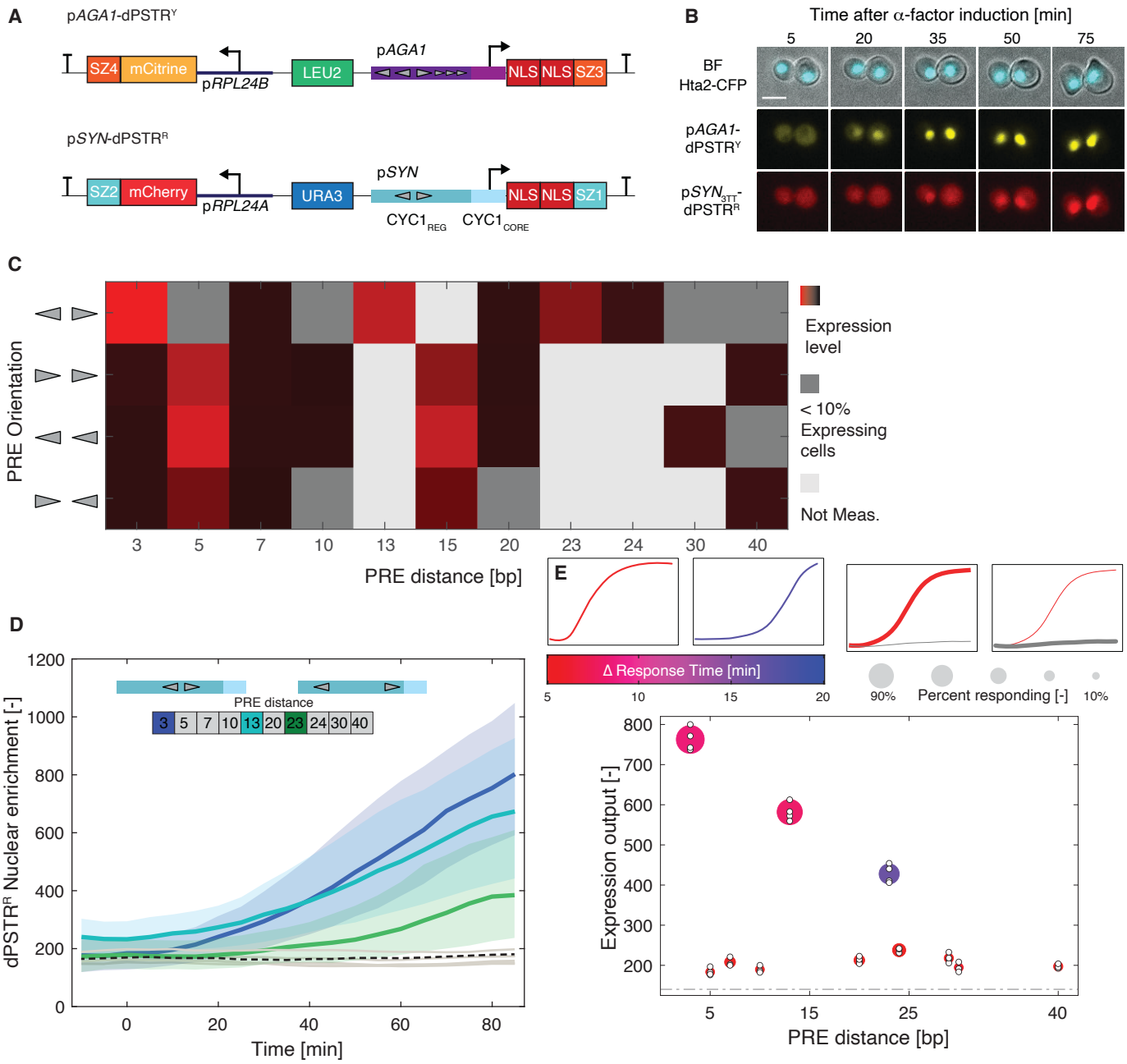
A
pAGA1-dPSTRY
SZ4
mCitrine
pRPL24B
LEU2
pAGA1
NLS
NLS
SZ3
pSYN-dPSTRR
SZ2
mCherry
pRPL24A
URA3
pSYN
CYC1REG
CYC1CORE
NLS
NLS
SZ1
B
Time after α-factor induction [min]
5
20
35
50
75
BF
Hta2-CFP
pAGA1-
dPSTRY
pSYN3TT-
dPSTRR
C
PRE Orientation
3
5
7
10
13
15
20
23
24
30
40
PRE distance [bp]
Expression
level
< 10%
Expressing
cells
Not Meas.
D
dPSTRR Nuclear enrichment [-]
1200
1000
800
600
400
200
0
PRE distance
3
5
7
10
13
20
23
24
30
40
0
20
40
60
80
Time [min]
E
Δ Response Time [min]
5
10
15
20
90%
Percent responding [-]
10%
Expression output [-]
800
700
600
500
400
300
200
5
15
25
40
PRE distance [bp]

Figure 1

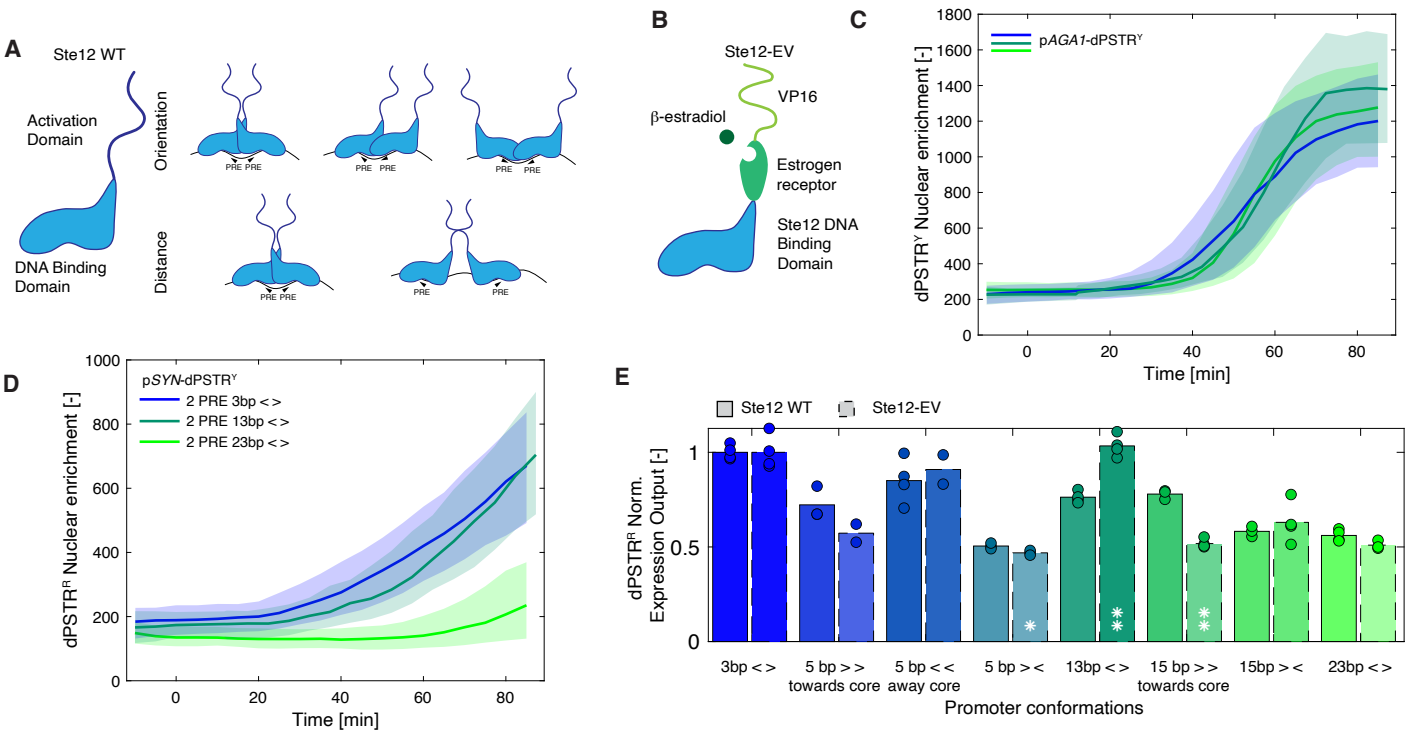
A
Ste12 WT
Activation Domain
DNA Binding Domain
Orientation
Distance
PRE PRE
B
Ste12-EV
VP16
β-estradiol
Estrogen receptor
Ste12 DNA Binding Domain
C
pAGA1-dPSTRY
dPSTRY Nuclear enrichment [-]
Time [min]
D
pSYN-dPSTRY
2 PRE 3bp < >
2 PRE 13bp < >
2 PRE 23bp < >
dPSTRR Nuclear enrichment [-]
Time [min]
E
Ste12 WT
Ste12-EV
dPSTRR Norm. Expression Output [-]
3bp < >
5 bp > > towards core
5 bp < < away core
5 bp > <
13bp < >
15 bp > > towards core
15bp > <
23bp < >
Promoter conformations

Figure 2

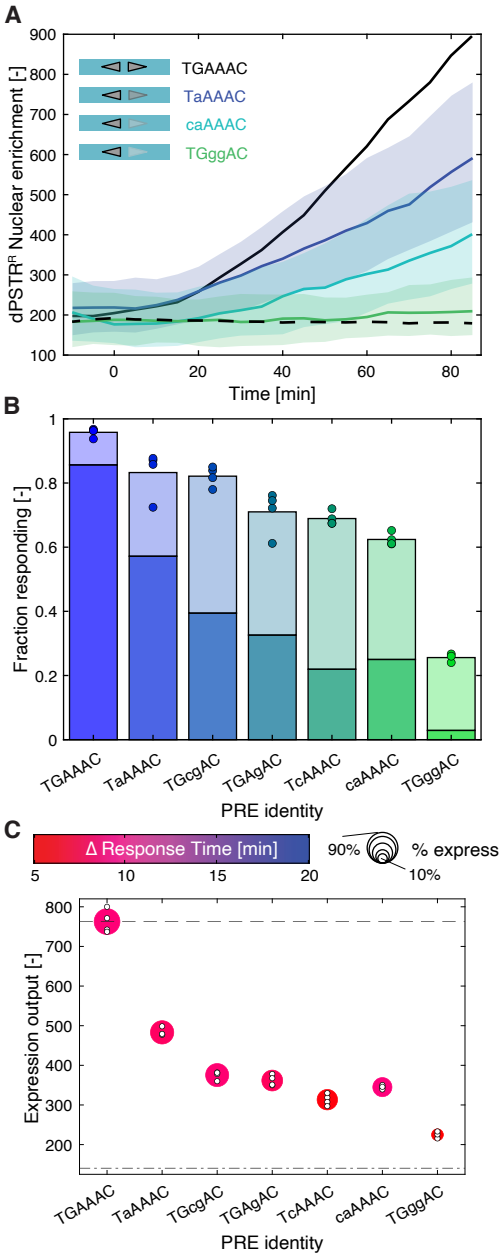
A
TGAAAC
TaAAAC
caAAAC
TGggAC
dPSTR^R Nuclear enrichment [-]
Time [min]
B
Fraction responding [-]
TGAAAC
TaAAAC
TGcgAC
TGAgAC
TcAAAC
caAAAC
TGggAC
PRE identity
C
Δ Response Time [min]
% express
90%
10%
Expression output [-]
PRE identity

Figure 3

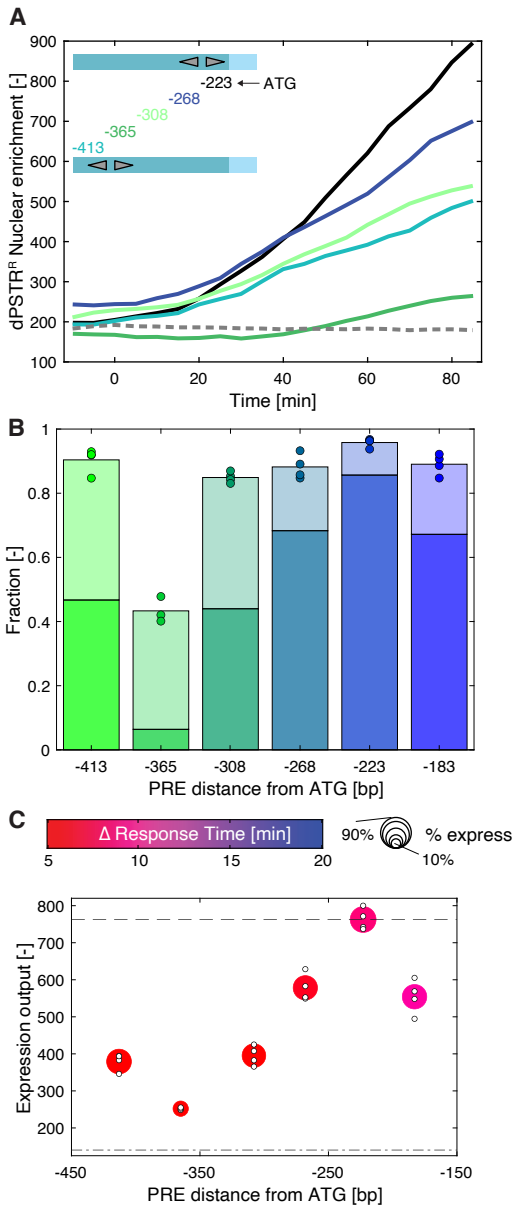
A
dPSTR^R Nuclear enrichment [-]
Time [min]
-223
ATG
-268
-308
-365
-413
B
Fraction [-]
PRE distance from ATG [bp]
-413
-365
-308
-268
-223
-183
C
Δ Response Time [min]
% express
90%
10%
Expression output [-]
PRE distance from ATG [bp]

Figure 4

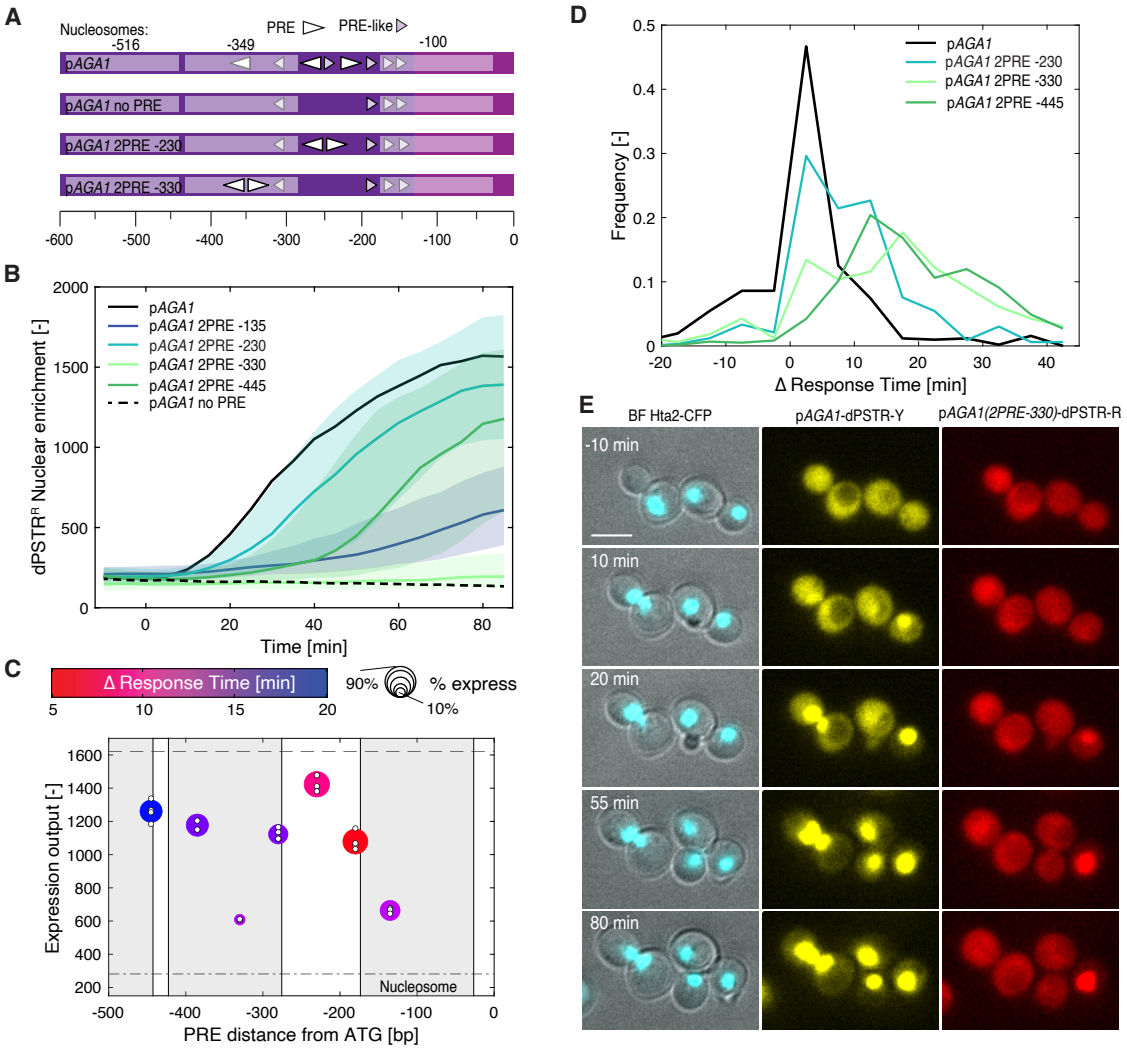
A
Nucleosomes:
-516
-349
PRE
PRE-like
-100
pAGA1
pAGA1 no PRE
pAGA1 2PRE -230
pAGA1 2PRE -330
-600
-500
-400
-300
-200
-100
0
B
pAGA1
pAGA1 2PRE -135
pAGA1 2PRE -230
pAGA1 2PRE -330
pAGA1 2PRE -445
pAGA1 no PRE
dPSTR^R Nuclear enrichment [-]
Time [min]
C
Δ Response Time [min]
90%
% express
10%
Expression output [-]
Nucleosome
PRE distance from ATG [bp]
D
pAGA1
pAGA1 2PRE -230
pAGA1 2PRE -330
pAGA1 2PRE -445
Frequency [-]
Δ Response Time [min]
E
BF Hta2-CFP
pAGA1-dPSTR-Y
pAGA1(2PRE-330)-dPSTR-R
-10 min
10 min
20 min
55 min
80 min

Figure 5

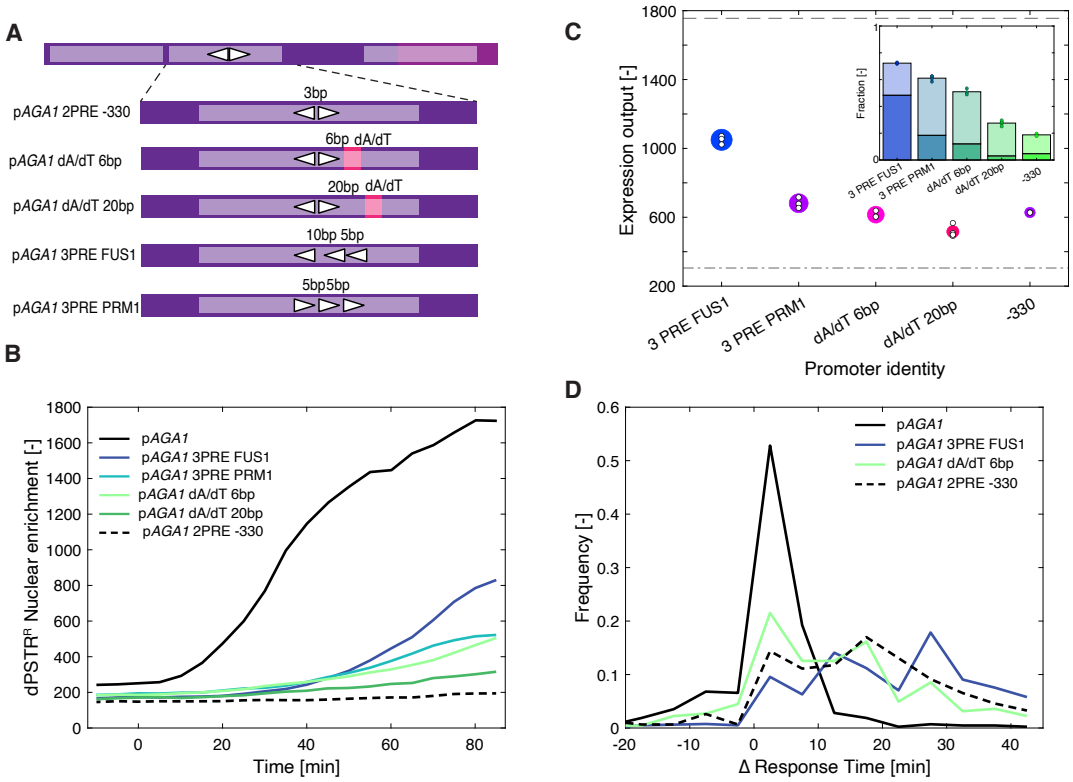
A
3bp
pAGA1 2PRE -330
6bp dA/dT
pAGA1 dA/dT 6bp
20bp dA/dT
pAGA1 dA/dT 20bp
10bp 5bp
pAGA1 3PRE FUS1
5bp5bp
pAGA1 3PRE PRM1
B
dPSTR^R Nuclear enrichment [-]
Time [min]
pAGA1
pAGA1 3PRE FUS1
pAGA1 3PRE PRM1
pAGA1 dA/dT 6bp
pAGA1 dA/dT 20bp
pAGA1 2PRE -330
C
Expression output [-]
Promoter identity
3 PRE FUS1
3 PRE PRM1
dA/dT 6bp
dA/dT 20bp
-330
Fraction [-]
D
Frequency [-]
Δ Response Time [min]
pAGA1
pAGA1 3PRE FUS1
pAGA1 dA/dT 6bp
pAGA1 2PRE -330

Figure 6