BIOINFORMATICS

Vol. 27 ISMB 2011, pages i283–i287
doi:10.1093/bioinformatics/btr211

The role of proteosome-mediated proteolysis in modulating potentially harmful transcription factor activity in *Saccharomyces cerevisiae*

Nicola Bonzanni[1,†], Nianshu Zhang[2,†], Stephen G. Oliver[2,‡] and Jasmin Fisher[3,‡,*]

[1]Centre for Integrative Bioinformatics VU, VU University Amsterdam, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands, [2]Department of Biochemistry, Cambridge Systems Biology Centre, University of Cambridge, Sanger Building, 80 Tennis Court Road, Cambridge CB2 1GA and [3]Microsoft Research Cambridge, Roger Needham Building, 7 JJ Thomson Avenue, Cambridge CB3 0FB, UK

ABSTRACT

Motivation: The appropriate modulation of the stress response to variable environmental conditions is necessary to maintain sustained viability in *Saccharomyces cerevisiae*. Particularly, controlling the abundance of proteins that may have detrimental effects on cell growth is crucial for rapid recovery from stress-induced quiescence.

Results: Prompted by qualitative modeling of the nutrient starvation response in yeast, we investigated *in vivo* the effect of proteolysis after nutrient starvation showing that, for the Gis1 transcription factor at least, proteasome-mediated control is crucial for a rapid return to growth. Additional bioinformatics analyses show that potentially toxic transcriptional regulators have a significantly lower protein half-life, a higher fraction of unstructured regions and more potential PEST motifs than the non-detrimental ones. Furthermore, inhibiting proteasome activity tends to increase the expression of genes induced during the Environmental Stress Response more than those in the rest of the genome. Our combined results suggest that proteasome-mediated proteolysis of potentially toxic transcription factors tightly modulates the stress response in yeast.

Contact: jasmin.fisher@microsoft.com

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 INTRODUCTION

A prompt and appropriate response to abrupt fluctuations in external conditions is crucial to survive stressful environmental changes, especially in unicellular organisms such as the yeast *Saccharomyces cerevisiae*. During the nutrient starvation, in order to ensure extended survival, *S.cerevisiae* cells exit the cell cycle at G_1 and enter the quiescent state (called G_0), but rapidly resume growth and proliferation when nutrient conditions turn favorable. Two conserved signaling pathways Ras/cAMP and TOR are known to coordinate the entry into and exit from the quiescent phase (Wilson and Roach, 2002). These two pathways regulate the entry into the stationary phase, converging on the protein kinase Rim15 (Pedruzzi *et al.*, 2003) and downstream transcriptional activators, including the stress response (STRE) transcription factors (TFs) Msn2/Msn4 and the post-diauxic shift (PDS) transcription factor Gis1 (Zhang *et al.*, 2009). The nutrient starvation response is an intensively studied process, but the exact molecular mechanisms involved have not yet been fully elucidated. On the one hand, the scarcity of quantitative data pose a problem for the construction of quantitative models; on the other hand, the current understanding of the causal regulatory wiring encourages the use of qualitative computational models to gain new insights.

Executable Biology (Fisher and Henzinger, 2007; Fisher and Piterman, 2010) is an evolving paradigm that focuses on the design of executable computer algorithms that mimic biological phenomena through the use of formal methods from engineering and computer science. Biological knowledge can be captured in mathematically sound formalisms, and then easily translated into executable algorithms for dynamical analysis and automatic reasoning. Here, we show that formalizing the available knowledge on the nutrient starvation response as a qualitative model highlighted the different modulation of Gis1 availability, encouraging further *in vivo* investigations on the role of proteosome-mediated proteolysis.

Proteasome-mediated proteolysis is essential for many cellular processes in yeast and other eukaryotes, including regulation of protein concentrations and degradation of misfolded proteins. Integrating our computational insights and the *in vivo* experiments with genome-wide bioinformatics analyses lead us to suggest that proteasome-mediated proteolysis of potentially toxic transcription factors tightly modulates the stress response in yeast.

2 METHODS

2.1 Petri nets

We have built a qualitative logical model of nutrient starvation based on Petri nets. Petri nets are mathematically sound formalisms that can be graphically represented (Reisig and Rozenberg, 1998). Recently, Petri nets have been used in systems biology to build and analyze coarse-grained models of complex processes (Bonzanni *et al.*, 2009), taking advantage of the intuitiveness of their representation and the soundness of their foundation. The Petri net modeling framework used in this work has been derived from the seminal work of Chaouiya *et al.* (2006). The states predicted by the model can be found in Supplementary Material. Statistical analyses of bioinformatics data were performed using *R*.

2.2 Gis1 overexpression at the transition phase

Wild-type (BY4742) cells were transformed with pCM190 (Gari *et al.*, 1997) and pCM190-GIS1 (Zhang and Oliver, 2010). Transformants were grown on

*To whom correspondence should be addressed.

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

‡The authors wish it to be known that, in their opinion, the last two authors should be regarded as joint Second Authors.

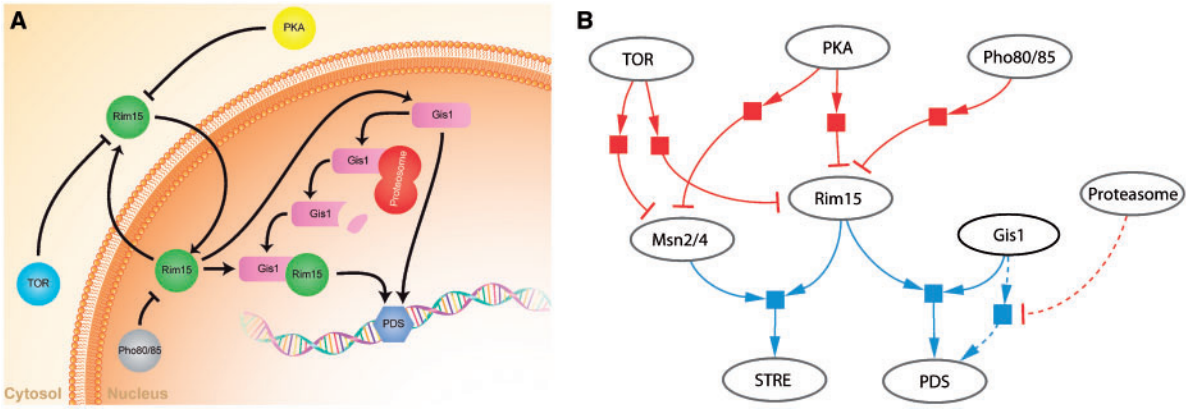

Fig. 1. Model of nutrient starvation response in yeast. (**A**) Diagrammatic model depicting the proteolytic control over Gis1 and the regulation of Rim15 by TOR, PKA and Pho80/Pho85. (**B**) Partial formal model of nutrient starvation response. Ovals = nodes that represent 'places'—proteins (e.g. PKA, Rim15, Gis1) and genes (PDS and STRE); colored squares = interactions. Arcs ending with an arrowhead (in blue) represent positive interactions (e.g. activations), while arcs ending with bars (in red) represent negative interactions (e.g. inhibitions). Note that if multiple arrows target the same square, all the sources are required at the same time. Dashed lines represent the interaction responsible for the discrepancy between the modeled and observed behaviors.

an SMM (Amberg *et al.*, 2005) containing 20 μg/ml of doxycycline (Sigma-Aldrich) and 2% glucose to glucose starvation. Cells were harvested, washed once in sterile water and resuspended in the SMM medium containing no doxycycline or glucose for 36 h to allow Gis1 overexpression. Growth was resumed by adding 2% glucose and doxycycline. Cell viability was checked by staining cells with phloxine B (Sigma-Aldrich).

3 RESULTS

3.1 Model construction and analysis

In order to investigate the consistency and explanatory power of the available knowledge about the nutrient starvation response in yeast, we have constructed a dynamic computational model based on Petri nets (Reisig and Rozenberg, 1998). Petri nets can be depicted as graphs that contain two kinds of nodes: *places*, which represent resources and correspond to proteins and genes, and *transitions*, which represent interactions between *places*. Interactions can be either activations or inhibitions (Fig. 1B) and, during the course of the execution, each resource can change its state (in a Boolean fashion) from active to inactive (and vice versa) based on the surrounding interactions. Given a network topology, it is possible to execute the model and compare its behavior with the one observed empirically.

Due to the lack of fine-grained quantitative data, we captured the coarse-grained descriptive knowledge available in the form of a qualitative model firmly based on published experimental evidence. This model includes the inhibition of Msn2/4 activity by TOR and PKA (Beck and Hall, 1999; Görner *et al.*, 1998), which is represented in Figure 1B by the red transitions connecting the TOR and PKA nodes to Msn2/4. Notice that the arc connecting TOR to the transition ends with an arrowhead, while the arc connecting the transition to Msn2/4 ends with a bar. This means that the availability of TOR is a necessary precondition for Msn2/4 repression. Similarly, we have represented Rim15 inhibition by TOR, PKA and Pho80/85 (Pedruzzi *et al.*, 2003 and Wanke *et al.*, 2005), the expression of STRE and the PDS genes upon Rim15 activation of Msn2/4 and Gis1, as well as the recently discovered proteolytic control over Gis1 (Zhang and Oliver, 2010).

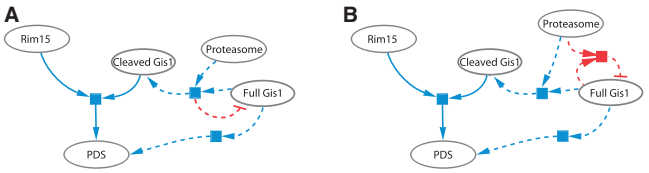

Fig. 2. Multiple possible wiring choices allow refinement of the model. Fragment of the model under refinement. The dashed interactions are more accurate alternatives than the dashed interaction in Figure 1B. Two alternative options are presented: (**A**) proteolytic activity induces complete degradation of the full-length Gis1 protein and simultaneous availability of cleaved Gis1 fragments. (**B**) Decoupling the production of cleaved Gis1 fragments and degradation of full-length protein allows partial depletion of the full-length Gis1.

After the construction of the network model, we analyzed its dynamics. By comparing our model with the experimental observations (Zhang and Oliver, 2010), we discovered a significant discrepancy in the behavior of Gis1 reproduced by the model. Our computational results (see Supplementary Material) suggested that only the full-length Gis1 was necessary for the activation of PDS genes. However, upon nutrient starvation or TORC1 inhibition, the abundance of full-length Gis1 decreases, which does not correspond to the increase of transcription activation of PDS genes (Zhang and Oliver, 2010). Moreover, although full-length Gis1 is essential for PDS gene expression, the smaller Gis1 fragments, resulting from constitutive proteolysis by the proteasome, are also able to initiate transcription upon Rim15 activation (Zhang and Oliver, 2010). These data suggested that full-length Gis1 and its smaller variants activate the transcription of PDS genes cooperatively. Therefore, we concluded that our model needed to be refined by including the full-length protein and the smaller fragments separately, in order to fully capture the biological observations and increase the model's accuracy. Different wiring choices were possible. One possibility, shown in Figure 2A, is to allow proteolytic activity to induce *complete* degradation of full-length Gis1. This is the behavior

observed during nutrient starvation; however, Gis1 is also subject to a constitutive, but *partial* degradation by the proteasome (Zhang and Oliver, 2010) during exponential growth. Therefore, an alternative modeling choice is to allow partial depletion of full-length Gis1. This can be accomplished by decoupling the availability of the cleaved Gis1 fragments from the complete degradation of the full-length protein (Fig. 2B). By refining our model as shown in Figure 2B, it qualitatively reproduced (see Supplementary Material) the behavior observed in Zhang and Oliver (2010).

3.2 Proteolytic control over Gis1 allows fast recovery from lag phase

The different causal wirings imply differences in the model behavior and may therefore suggest different roles for the proteolytic control. In order to understand the evolutionary advantages of the different proteolytic controls over Gis1 in the context of nutrient response, we were prompted to investigate its physiological role. *GIS1* overexpression leads to accumulation of the full-length protein and is toxic to cell growth (Pedruzzi *et al.*, 2000; Zhang and Oliver, 2010). Inhibition of the proteasome function results in hyperactivation of PDS genes in nutrient-starved conditions (Zhang and Oliver, 2010). Knowing that growth and budding are suspended in stationary phase, we performed an experiment to determine whether the proteolytic control over Gis1 is necessary for survival of cells entering the stationary phase, the recovery of cells from glucose starvation or both.

Wild-type yeast cells were transformed with plasmid pCM190 or the same plasmid bearing the *GIS1* gene under the control of the repressible promoter, *tet*O. Cells were grown in the presence of doxycycline to early stationary phase, washed and resuspended in medium with no glucose or doxycycline for 36 h. There is no difference in viability between cells bearing the empty plasmid and those carrying the *tet*O-*GIS1* plasmid (data not shown). Glucose and doxycycline were added to allow cells to resume growth. As shown in Figure 3, cells harboring the *tet*O-*GIS1* plasmid display a 15% longer lag phase than those bearing the empty plasmid, suggesting that *GIS1* overexpression during the transition to quiescence, delays the subsequent resumption of exponential growth on re-addition of nutrients. These data indicate that proteolytic degradation of Gis1 by the proteasome may provide cells with an important evolutionary advantage, since periods of nutrient availability and starvation are commonly experienced by microorganisms (Gasch and Werner-Washburne, 2002).

3.3 Predicting that toxic transcriptional regulators are subject to tighter proteolytic control

Prompted by the proteolytic regulation of Gis1 and its physiological implications, we went on to inquire if, in general, the stress response is restrained by the proteasome. We adopted two strategies: the first to discover whether toxic transcription factors are likely to be controlled post-translationally by the proteasome, and the second to find out whether proteasome inhibition allows transcription factors normally targeted by the proteasome to elicit a stress response.

3.3.1 Toxic transcriptional regulators have lower half-life To monitor the validity of our hypothesis, we performed a sequence of bioinformatics analyses. First, we partitioned the known yeast transcriptional regulators into two disjoint sets. The first set

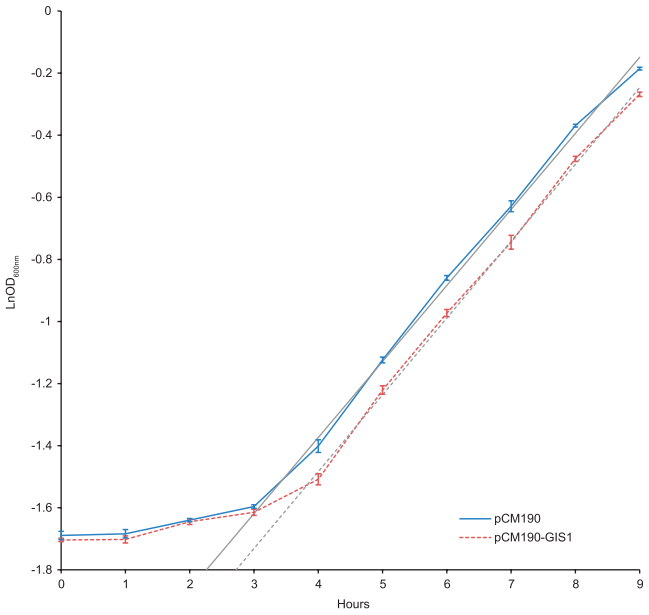

Fig. 3. Cells over-expressing Gis1 at starvation have a longer lag phase than wild-type cells. Cells bearing either pCM190-GIS1 or the empty vector, pCM190, were grown (in the presence of doxycycline) for 24 h to glucose starvation, at which point >90% of cells have no buds. Aliquots (2 ml) of cell cultures were washed twice in SMM medium without glucose or doxycycline, resuspended in 40 ml of SMM, and incubated for 36 h to allow *GIS1* expression. At this point, glucose (2%) and doxycycline (20 μg/ml) were added to the cultures. Growth was monitored as OD_{600nm}.

contained 75 potentially toxic regulators and was created by filtering the set of 796 genes, whose overexpression was found to be detrimental for cell growth (Sopko *et al.*, 2006) using the GO annotation 'transcription regulator activity' (GO:0030528). The second set contained 251 non-toxic regulators and was built by filtering the whole yeast genome with the same GO annotation after removing the toxic genes contained in the first set. Detailed data are available as Supplementary Material.

With our first analysis, we assessed whether the protein half-lives of toxic regulators are shorter than those of non-toxic regulators, using the protein half-life measurements of Belle *et al.* (2006). Since the measurements are not normally distributed ($P < 10^{-15}$; Shapiro–Wilk test), we computed the Wilcoxon rank sum test under the null hypothesis that the median difference between the two measurement sets is zero and the alternative hypothesis that the median half-life of the toxic transcription factors is less than that of the non-toxic ones. The null hypothesis has been discarded with the statistically significant value of $P = 5.54 \times 10^{-3}$ (Fig. 4A). Note that it was not possible to find measurements for all the proteins in the two sets. We also analyzed the mRNA half-life data (Wang *et al.*, 2002) for the transcripts of the toxic and the non-toxic TFs and found no significant difference between the two ($P = 0.256$; Wilcoxon test), supporting the hypothesis that a significant portion of the control over the toxic TFs is exerted post-transcriptionally (Fig. 4B).

3.3.2 Toxic transcriptional regulators have a higher fraction of unstructured regions The availability of many intrinsically unstructured proteins (IUPs) is regulated via proteolytic degradation

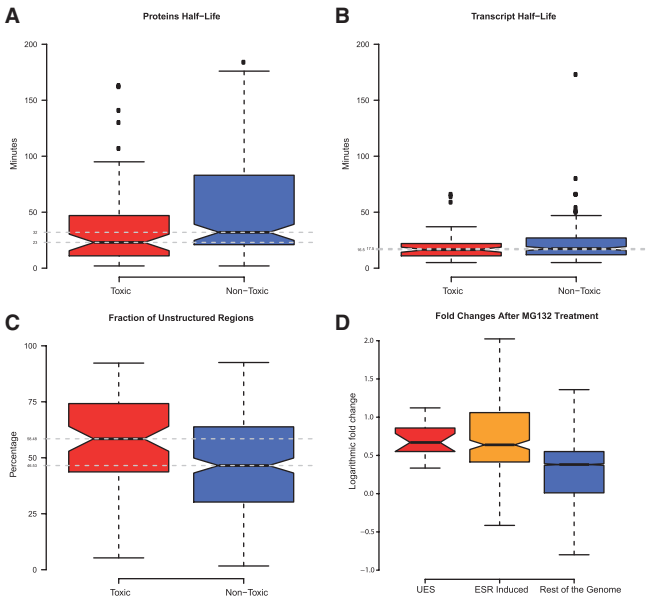

Fig. 4. Comparison between toxic and non-toxic regulators. (**A**) The half-lives of toxic regulators (red) are significantly lower ($P = 5.54 \times 10^{-3}$; Wilcoxon test) than those of non-toxic ones (blue), while (**B**) the median half-life of the transcripts of toxic regulators is not significantly different from that of the non-toxic ones ($P = 0.256$; Wilcoxon test). (**C**) The fraction of amino acids predicted to form unstructured regions is significantly higher in toxic than in non-toxic proteins ($P = 2.48 \times 10^{-4}$; Wilcoxon test). (**D**) After 120 min of proteasome inhibition by MG132, transcription rates of UES genes ($P = 8.26 \times 10^{-05}$; Wilcoxon test) and ESR induced genes ($P < 2.2 \times 10^{-16}$; Wilcoxon test) tend to be higher than those for the rest of the genome.

(Gsponer *et al.*, 2008). Therefore, for both the toxic and non-toxic regulators, we computed (using Disopred2; Ward *et al.*, 2004) the fraction of the amino acids in each protein that is predicted to lie within unstructured regions. We found (Fig. 4C) that the median content of unstructured regions is higher for toxic transcription factors than that for non-toxic regulators ($P = 2.48 \times 10^{-4}$; Wilcoxon test), supporting the hypothesis that proteasome-mediated degradation plays a significant role in the regulation of the activity of potentially detrimental TFs.

3.3.3 Toxic TFs contain more potential PEST motifs Sequence regions rich in proline (P), glutamic acid (E), serine (S) and threonine (T) are found in many rapidly degraded proteins and have been suggested to serve as signals for proteolysis (Rogers *et al.*, 1986). We analyzed the number of potential PEST motifs in the protein sequences of the two classes. Using the epestfind algorithm from the EMBOSS package (Rice *et al.*, 2000), we predicted the number of potential PEST motifs for both sets of proteins. While 44/75 (59%) toxic regulators contain at least one PEST motif, the ratio is 109/251 (43%) for the non-toxic ones (P value of 3.7×10^{-2}; Fisher's exact test). This, again, provides some support for our hypothesis on the role of proteolysis in regulating the activity of potentially toxic TFs.

3.3.4 The proteosome modulates the expression of a significant fraction of genes induced by environmental stress Finally, we investigated whether proteolytic control could contribute to modulating the stress response by checking transcriptional changes after proteasome inhibition. A previous study has shown that 23% of all yeast genes (1386 mRNAs) increase their rate of transcription by a factor of 1.5 or more (6% increase more than 2 times) after 120 min treatment with the proteosome inhibitor MG132 (Dembla-Rajpal *et al.*, 2004). We extracted the data for the Universally Expressed at Starvation (UES) genes (Wu *et al.*, 2004); these genes are controlled by Gis1 and Msn2—two TFs known to be under proteolytic control. We found that the fold changes of the UES genes tend to be higher than for the rest of the genome ($P = 8.26 \times 10^{-5}$; Wilcoxon test). More interestingly, we observe a significant fold increase with respect to the rest of the genome ($P < 2.2 \times 10^{-16}$; Wilcoxon test), further extending the analysis of the effect of inhibiting proteasome activity on the induction of gene transcription in the Environmental Stress Response (ESR; Gasch *et al.*, 2000; see Figure 4D).

To summarize, our work suggests that proteasome-mediated proteolysis of TFs tightly modulates the stress response in yeast. This hypothesis is the result of the integration of computational and *in vivo* analysis. Our computational model highlighted the particular behavior of the proteolytic control, suggesting further *in vivo* investigations. Our *in vivo* experiments showed that, for the Gis1 transcription factor at least, proteasome-mediated control is crucial for a rapid return to growth after nutrient starvation, which may give yeast cells an important selective advantage over their competitors. Finally, our bioinformatics analyses generalized our *in vivo* observations to the class of potentially toxic transcription factors that control the stress response in yeast.

ACKNOWLEDGEMENTS

We would like to thank A. Feenstra, W. Fokkink and J. Heringa for helpful discussions. Part of this work was done, while NB was an intern at Microsoft Research Cambridge.

Funding: Work in the Cambridge Systems Biology Centre was supported by BBSRC (Grant BB/C505140/2 awarded to S.G.O.); Work in the Centre for Integrative Bioinformatics VU was supported in part by ENFIN; a Network of Excellence funded by the European Commission within its FP6 Program, under the thematic area 'Life Sciences, genomics and biotechnology for health', contract number LSHG-CT-2005-518254.

Conflict of Interest: none declared.

REFERENCES

Amberg,D.C. *et al.* (2005) *Methods in Yeast Genetics*. Cold Spring Harbor Laboratory Press, New York.

Beck,T. and Hall,M.N. (1999) The TOR signalling pathway controls nuclear localization of nutrient-regulated transcription factors. *Nature*, **402**, 689–692.

Belle,A. *et al.* (2006) Quantification of protein half-lives in the budding yeast proteome. *Proc. Natl Acad. Sci. USA*, **103**, 13004–13009.

Bonzanni,N. *et al.* (2009) Executing multicellular differentiation: quantitative predictive modelling of *C. elegans* vulval development. *Bioinformatics*, **25**, 2624–2624.

Chaouiya,C. *et al.* (2006) Qualitative Petri Net modelling of genetic networks. *Trans. Comput. Sys. Biol.*, **4220**, 95–112.

Dembla-Rajpal,N. *et al.* (2004) Proteasome inhibition alters the transcription of multiple yeast genes. *Biochim. Biophys. Acta*, **1680**, 34–45.

Fisher,J. and Henzinger,T.A. (2007) Executable cell biology. *Nat. Biotechnol.* **25**, 1239–1249.

Fisher,J. and Piterman,N. (2010) The executable pathway to biological networks. *Brief. Funct. Genomic Proteomic*, **9**, 79–92.

Gari,E. *et al.* (1997) A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in *Saccharomyces cerevisiae*. *Yeast*, **13**, 837–848.

Gasch,A. and Werner-Washburne,M. (2002) The genomics of yeast responses to environmental stress and starvation. *Funct. Integr. Genomics*, **2**, 181–192.

Gasch,A.P. *et al.* (2000) Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell*, **11**, 4241–4257.

Görner,W. *et al.* (1998) Nuclear localization of the C_2H_2 zinc finger protein Msn2p is regulated by stress and protein kinase A activity. *Genes Dev.*, **12**, 586–597.

Gsponer,J. *et al.* (2008) Tight regulation of unstructured proteins: from transcript synthesis to protein degradation. *Science*, **322**, 1365–1368.

Pedruzzi,I. *et al.* (2003) TOR and PKA signaling pathways converge on the protein kinase Rim15 to control entry into G0. *Mol. Cell*, **12**, 1607–1613.

Reisig,W. and Rozenberg,G. (1998) *Lectures on Petri Nets I: Basic Models*. Vol. 1491, Springer, Berlin.

Rice,P. *et al.* (2000) EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.*, **16**, 276–277.

Rogers,S. *et al.* (1986) Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science*, **234**, 364–368.

Sopko,R. *et al.* (2006) Mapping pathways and phenotypes by systematic gene overexpression. *Mol. Cell*, **21**, 319–330.

Wang,Y. *et al.* (2002) Precision and functional specificity in mRNA decay. *Proc. Natl Acad. Sci. USA*, **99**, 5860–5865.

Wanke,V. *et al.* (2005) Regulation of G0 entry by the Pho80-Pho85 cyclin-CDK complex. *EMBO J.* **24**, 4271–4278.

Ward,J.J. *et al.* (2004) Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.*, **337**, 635–645.

Wilson,W.A. and Roach,P.J. (2002) Nutrient-regulated protein kinases in budding yeast. *Cell*, **111**, 155–158.

Zhang,N. and Oliver,S.G. (2010) The transcription activity of Gis1 is negatively modulated by proteasome-mediated limited proteolysis. *J. Biol. Chem.* **285**, 6465–6476.

Zhang,N. *et al.* (2009) Gis1 is required for transcriptional reprogramming of carbon metabolism and the stress response during transition into stationary phase in yeast. *Microbiology*, **155**, 1690–1698.