Design principles and specificity in biological networks with cross-activation

Bo Hu,[1] Herbert Levine,[1] and Wouter-Jan Rappel[1]

[1] *Center for Theoretical Biological Physics University of California San Diego La Jolla, CA 92093, USA*

Cells sense and respond to diverse environmental stimuli using a set of intracellular signaling components. Often, the signal transduction pathways contain shared components which leads to cross-activation at different levels of the pathway. To discover the design principles that ensure signaling specificity is a challenging task, especially for pathways that contain numerous components. Here, we present an analysis of cross-activating pathways and show that a general inhibitory scheme, asymmetric hierarchical inhibition, is sufficient to ensure signaling specificity. Based on this inhibitory scheme, we are able to enumerate all possible network topologies containing two inhibitory links that guarantee specificity. Furthermore, we apply our methodology to the mating and filamentous growth pathways of the yeast model system *Saccharomyces cerevisiae*. We enumerate the possible ways to wire this model system and determine which topology is consistent with experimental data.

INTRODUCTION

Cells are able to reliably execute various functions in response to different environmental conditions. Each function is mediated by a signaling transduction pathway, which converts the external stimulus into specific intracellular signals. These signaling cascades involve intracellular biochemical reactions, carried out by a multitude of second messengers [1, 2]. Often, these cascades cross-activate or share signaling components, raising the possibility of erroneous activation of certain pathways [3]. Examples include pathways involving the second messenger cyclic adenosine monophosphate (cAMP) and mitogen activated protein kinase (MAPK) cascades that regulate multiple functions in eukaryotic cells [4–6].

To ensure signaling specificity, a number of strategies can be employed [3]. Physically separating the signaling components, for example through sequestration or compartmentalization, is conceptually the simplest way to avoid undesirable cross-talk [7–9]. Alternatively, cross-pathway inhibition of components can increase the signaling specificity [10, 11]. This cross-talk occurs often on multiple levels of the signaling cascade and it remains a challenge to understand the intricate wiring required to ensure signaling specificity. In particular, it is difficult to determine the minimum number of inhibitory links that are needed to suppress undesired crosstalk, how to organize these links, and how to enumerate the number of possible network configurations that guarantee signaling specificity.

In this study, we address these questions using simple theoretical and numerical tools that allow us to reduce complex networks to essential signaling modules. In particular, we show that we can capture the essential behavior of two cross-activating signaling cascades by reducing them to a simplified module that contains only four nodes. Using a genetic algorithm [12, 13], we probe all possible configurations of this module that contain inhibitory links. We find that the unique topology that ensures signaling specificity employs the asymmetric hierarchical inhibition mechanism in which two inhibitory links operate at different levels of the signaling cascade. We apply this finding to general and complex signaling networks and, after discussing a reduction algorithm, enumerate the possible topologies that ensure specificity.

Our approach and results are then applied to the budding yeast *Saccharomyces cerevisiae*, a model system for cross-talk [3]. We focus on the mating and filamentation pathways which are characterized by a number of shared components. We enumerate the number of possible networks and explicitly determine their network architecture. Using existing experimental data, we show that only two of these topologies are plausible.

RESULTS

A four-node module underlies signaling pathways with cross-activation

As a first step, we present a coarse-graining method for signaling networks that contain cross-activating links. An example of such a network is presented in Figure. 1A and shows two branches, each with a number of nodes connected with links. Each node corresponds to biochemical components and can represent either a single component or a group of components. The reaction kinetics of these components might be complicated but, since we are only going to be

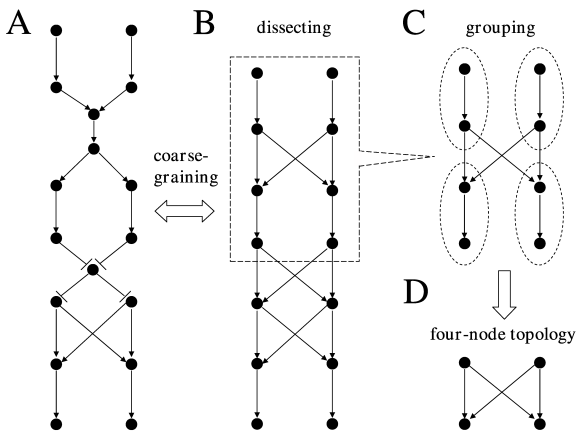


FIG. 1: A simple four-node topology underlies cross-activating signaling networks. (A) A graphical representation of two typical signaling transduction pathways with shared components at multiple levels. The nodes represent single or groups of biochemical components while the links represent either activation (arrows) or inhibition (bars). (B) After coarse-graining this network, one can dissect the network into several subsystems, each of which contains only one pair of cross-activating links. (C) The nodes in each subsystem can be grouped such that the components have the same logical states (on or off). (D) The final result is a four-node topology, containing two input nodes and two output nodes, along with two cross-activating links.

interested in the steady-state behavior of the network, will be ignored. In what follows, a link can either represent the activation (arrows) or the inhibition (bars) of one component by another component. The nodes are drawn in a causal way such that time flows from top to bottom and shared components can easily be identified. A level can then be defined as the hierarchical place in the cascade with the top node corresponding to the input level and the bottom node to the output level.

To reduce the complexity of this signaling cascade, we first identify shared components, responsible for the cross-activation. Next, we replace the shared node(s) with a set of cross-activation and activation links. These links connect the node upstream of the shared node(s) to the node downstream of the share nodes (Fig. 1B). This results in two signaling cascades with, possibly, multilevel crosstalk that is represented by activation links between the two cascades. Several subsystems, each of which contains only one set of cross-activating links at a single level, can be identified. Concentrating on one of the subsystems we see that the signaling components of each pathway can be classified into two groups. One group correspond to pathway-specific components that are activated before cross-activation and the second group contains the potentially nonspecific components after cross-activation (Fig. 1C).

To facilitate the analysis of this network, we will employ Boolean logic where a node is either on (1) or off (0) [14, 15]. Using this logic, it is easy to see that the elements in each group are always in the same logical state for all input conditions and can be replaced by a single node. This enables us to consider a reduced "four-node system", in which each input component can activate both of the output components (Fig. 1D). The architecture of this coarse-grained four-node network underlies the original complex network. Of course, the precise dynamics of the reduced network is not likely to be identical to the dynamics of the full network. However, the qualitative steady-state behavior of the simplified network will be identical to the one of the more complicated network. Since it is the steady-state behavior that is determining non-transient signaling specificity, analyzing this simple network will provide insight into the signaling specificity of the more complex cascade.

Asymmetric hierarchical inhibition ensures signaling specificity

The four-node system contains two pathways which we will denote by X and Y (Fig. 2A). With only activation links, the truth table, reporting the output states as a function of the input states, of this system does not exhibit signaling specificity (Fig. 2A). After all, signaling specificity is achieved when the output of Y is negligible compare to the output of X when only X is stimulated and vice versa. As a specific example, we will consider here the case where one of the pathways, taken to be X, is "dominant", resulting in $X_{out} = 1$ and $Y_{out} = 0$ when both pathways are activated ($X_{in} = 1$ and $Y_{in} = 1$). The corresponding desired truth table is presented in Fig. 2B.

Mathematically, the four-node system can be described by four ordinary differential equations; see Methods for details. In this four-node topology, we can enumerate eight possible inhibitory links (for instance, X_{out} suppressing Y_{out}, X_{in} suppressing Y_{out}, etc.). To investigate which inhibitory links lead to the desired signaling specificity, we employed a genetic algorithm. The algorithm starts from a population of random realizations of the eight inhibitory links, with the strength of each link being controlled by a parameter that is chosen at random within a certain range. For each network topology, we evaluate its fitness score F, which measures the network's cost-effective capacity to perform the desired signaling function (for details, see Methods). The algorithm then keeps the top 50% best-fit networks as seeds for reproduction in the next generation, where the eight inhibitory links in each seed network are subject to random mutations. This process is repeated until a stable fitness value is achieved.

The maximum fitness score, along with the average fitness, is shown in Fig. 2C as a function of the generation number. After roughly 30 generations, the algorithm converges and produces a unique stable network topology. This topology is shown in 2B and contains two inhibitory links. One link suppresses Y_{out} directly from X_{in} while the other connects Y_{out} with X_{out}. This inhibition scheme can be classified as asymmetric and hierarchical [11]. Asymmetric since the inhibitory links break the symmetry between the two pathways and hierarchical since pathway X exerts its inhibitory effect before Pathway Y.

Generalization of our results

The lessons learned from the four-node model can be used in the analysis of more general and complex topologies. In particular, the insight that asymmetric hierarchical inhibition provides a mechanism for specificity can be applied to topologies where cross-activation occurs at multiple levels, An example of such a topology is shown in Fig. 2D, where we have applied the coarse-graining technique described above for a dominant pathway X and a non-dominant pathway Y. In the simplest case, the nodes are linked through direct cross-activation and activation, again represented by arrows between nodes. More complicated cases, including nodes linked through scaffolding and dimerization, can also be investigated, as we will discuss below.

Our analysis starts with partitioning the signaling components of each pathway into three sets. The first set contains all pathway-specific components and are contained in the solid boxes of Fig. 2D. Since these components are not affected by crosstalk and are solely controlled by the input signal, they can be viewed as the input set. The second set, within the dashed-dotted boxes, contains components that are downstream from the input level and upstream from the last component in the cascade that is affected by cross-talk. This intermediate set may contain many cross-activating links and it is important to recognize that signaling specificity is not measured at this level. After all, signaling specificity needs to be ensured at the final cross-talk level, independent of the pathway activation dynamics within the intermediate set. The third and final set is the output set, drawn within the dashed boxes, and includes the final component subject to cross-activation, along with all its downstream components. Obviously, signaling specificity has to be achieved at this level of the cascades.

The asymmetric hierarchical inhibition mechanism discovered in the analysis of the four-node module suggests that adding just two inhibitory links is sufficient to achieve signaling specificity. Furthermore, one of these links needs to connect a pathway-specific component from the first set of the dominant pathway to a component in the output set of the other pathway. We refer to this link as the dominant inhibitory link. The second link, the non-dominant inhibitory link, connects a component in the output set of pathway Y occurring at the same or lower level as the component receiving the dominant inhibition, to a component in the output set of the competing pathway X.

Unlike the simple four-node topology of Fig. 2, there is, in general, no longer a unique way to place the inhibitory links. The number of possible topologies, however, can be easily enumerated. Let m_1 and n_1 (m_2 and n_2) be the number of components in the first and third sets of pathway X (pathway Y), respectively. Then, the number of ways, N, we can organize the inhibitory links is given by

$$N = m_1 \sum_{j=1}^{n_2} (n_2 - j + 1) n_1 = m_1 n_1 n_2 (n_2 + 1)/2 \tag{1}$$

We can compare N to the total number of topologies with two inhibitory links, with one connecting X to Y and one connecting Y to X: $(m_1+n_1)^2(m_2+n_2)^2$. We see that the asymmetric hierarchical inhibition topologies constitute a small fraction in the entire topology space, especially for a large number of components in the input and output sets. We have verified through the use of genetic algorithms that the topologies we discuss above are the ones ensuring signaling specificity.

Extensions and limitations

The discussion above is limited to the case of two inhibitory links. Additional inhibitory links, however, can be easily incorporated. The enumeration of the topologies is still possible but might now involve the number of components in the intermediate sets. Clearly, adding more inhibitory links increases the number of possible topologies. An example of a topology that exhibits specificity and that involves three inhibitory links is shown in Fig. 2E. The most upstream inhibitory link creates an input specific component in the intermediate set of the dominant pathway. This component then inhibits a component in the output set of the non-dominant pathway. As before, a non-dominant link is required to ensure specificity when the Y pathway is activated.

The topologies of Fig. 2D and E only contain feedforward links. The addition of feedback loops can be analyzed in a straightforward matter. Examples of these loops are shown in Fig. 2F where they are indicated in blue. It is easy to verify that the feedback loops indicated as solid links do not interfere with the specificity. Specificity can be impaired, however, by linking a component that is in the intermediate set or in the output set and upstream from the link receiving the non-dominant inhibitory link to an upstream input set component. An example of such a feedback loop is shown as a dashed blue line in Fig. 2F.

Application to yeast signaling pathways

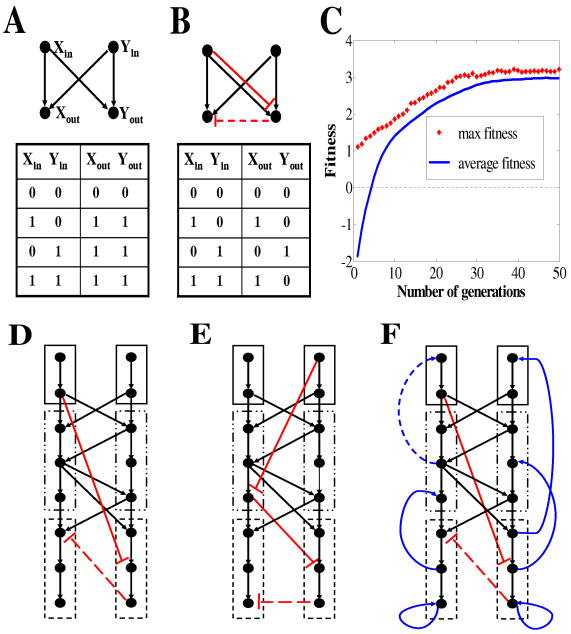
We will now apply our analysis to the yeast cell *Saccharomyces cerevisiae*, where a common set of protein kinases is activated during mating and the filamentous growth response [3, 11, 16, 17]. The mating program is initiated by the binding of pheromones to G protein-coupled receptors on the cell membrane. This triggers the dissociation of G protein into two subunits: G_α and $G_{\beta\gamma}$. The latter recruits the scaffold protein Ste5 to the membrane that co-localizes and activates in order the MAPK kinase kinase (MAPKKK) Ste11, the MAPK kinase (MAPKK) Ste7, and the MAPK Fus3. Several upstream kinases, including Ste20 and Ste50, also facilitate this sequential reaction. The activated MAPK Fus3 is a mating-specific component that not only controls mating gene expression but also promotes cell cycle arrest and the morphological projection toward the mating partner. In particular, Fus3 activates the transcription factor Ste12 in the nucleus by phosphorylating Ste12 inhibitors, Dig1/Dig2, and active Ste12 can bind cooperatively to the promoters of mating genes.

The switch to a filamentous growth form in yeast is initiated under the starvation of certain nutrients [3]. This filamentation pathway is mediated by a MAPK module that shares the MAPKKK and MAPKK with the mating pathway. The MAPK (Kss1) of the filamentation pathway is highly homologous to Fus3 and activates Ste12 as well as a filamentation-specific transcription factor Tec1. Cooperative binding of active Ste12 and Tec1 at the promoters drives the expression of filamentation genes. Remarkably, Fus3 following pheromone stimulation inhibits Tec1 by accelerating its degradation [18, 19], and thus suppresses the crosstalk with the filamentation pathway and results in a dominant mating pathway. Without such negative regulatory property, Fus3 appears functionally redundant with Kss1 [18]. Both mating and filamentation pathways have genetic positive feedback loops, as both Ste12 and Tec1 up-regulate themselves [20].

A variety of mechanisms have been proposed and studied to explain the signaling specificity in yeast cells, including the sequestration of shared components by scaffold proteins [21, 22], the experimental identification of Fus3-dependent inhibition on Tec1 [18, 19, 23], and the resource draining of Ste12 by Tec1 binding and Ste12-Tec1 degradation [11]. Since there is no absolute physical separation of the MAPK-dependent pathways, cross-pathway inhibition is the likely solution to promote signaling specificity.

The yeast mating and filamentous growth pathways can be graphically represented as in Fig. 3A. Here, some of the nodes are not connected through a direct activation link. Nevertheless, these links can still be treated as activation links since they represent a causal connection between the nodes and the network can be coarse-grained using the techniques described above, leading to the signaling cascades in Fig. 3B. For example, the scaffold protein Ste5 is required for the activation of Fus3. This requirement can be represented in the coarse-grained cascade by an activation link connecting Ste5 and Fus3. Similarly, the formation of the Ste12-Ste12 and Ste12-Tec1 dimers can be represented as activation links.

The core topology of Fig. 3B contains the minimal set of nodes and links that represent the skeleton underlying the full network. In Fig. 3B we have also identified the input, intermediate and output sets for both pathways as we did in Fig. 2. For the dominant mating pathway, the input set contains Fus3 (part of this set since Ste5 is needed for its activation), Ste5 and a node representing all pathway-specific components upstream of Ste5. The output set contains both Ste12 and Ste12-Ste12, while the intermediate set is empty. Note that Ste7, Ste11 and Ste20/Ste50 are not input specific and are therefore not in the input level. The input level for the non-dominant filamentation



X_{in}	Y_{in}	X_{out}	Y_{out}
0	0	0	0
1	0	1	1
0	1	1	1
1	1	1	1

X_{in}	Y_{in}	X_{out}	Y_{out}
0	0	0	0
1	0	1	0
0	1	0	1
1	1	1	0

FIG. 2: (A) The truth table associated with the basic four-node topology with crosstalk. (B) The inclusion of two inhibitory links (a dominant (solid line) and a non-dominant (dashed line)) leads to an asymmetric hierarchical inhibition mechanism and a truth table that exhibits signaling specificity and mutually exclusive outputs. (C) The maximum and the average fitness scores for the top 50% best-fit topologies in our genetic algorithm as a function of the generation number. After roughly 30 generations the genetic algorithm returns a stable configuration. (D) Example of a more complex signaling network with multilevel crosstalk exhibiting asymmetric hierarchical inhibition. The solid red line represents the dominant inhibitory link and dashed red line represents the non-dominant inhibitory link. (E) Example of a signaling network exhibiting specificity with three inhibitory links. (F) The introduction of most feedback loops (indicated as solid blue lines) does not impair the signaling specificity. However, specific feedback loops, such as the one drawn as a dashed blue line, can comprise signaling specificity.

pathway does not play a role in determining the suitable topologies while the intermediate set consists of Kss1 and the output set contains both Tec1 and Ste12-Tec1.

Following our enumeration scheme discussed above, we can identify $m_1 = 3$, $n_1 = 2$ and $n_2 = 2$, and, according to Eq.1, there are 18 ways to add inhibitory links to the signaling cascade that lead to asymmetric hierarchical inhibition. Six of these exhibit a dominant inhibitory link connecting a component upstream from Ste5 to a node in the output set of the filamentation pathway. Likewise, there are six possible topologies in which the dominant inhibitory link involves the scaffolding protein Ste5 and a node in the output set of the filamentation pathway. Motivated by experimental evidence that shows a direct inhibitory link between Fus3 and Tec1 [18, 19], we will focus here on the topologies with an inhibitory link originating from the Fus3 node in the mating pathway. Assuming mating dominance, we can draw six possible topologies as shown in Fig. 3C.

Of the six possible pathway topologies we have found, only two appear to be consistent with existing experimental data. Topology 1 requires the inhibition of Ste12 by Tec1. However, the formation of the heterodimer Tec1 and Ste12 is critical for the successful activation of the filamentation pathway [3, 20], making this possibility unlikely. Topology 2 and 4 involve the inhibition of the mating pathway genes by the transcription factor Ste12-Tec1. Experiments, however, indicate that this heterodimer binds specifically to filamentation gene promoters [20], making such a direct inhibition unlikely. Of course, there remains a possibility that gene products of the filamentation pathway inhibit mating gene transcription but this will require faster expression of filamentation products compared to that of mating products. The dominant inhibitory link in topology 3 and 4 comes from the direct inhibition of Ste12-Tec1 by Fus3 or the blocking of promoter sites for the filamentation genes by Fus3. Such an inhibitory interaction has not been found experimentally, making these topologies implausible.

The two most probable topologies are number 5 and 6 in which Fus3 directly inhibits Tec1, as demonstrated in experiments [18, 19]. In topology 5, Tec1 inhibits the activation of the mating genes, either through the blocking of the promoter activity of Ste12 or the DNA binding sites of Ste12-Ste12. Topology 6 relies on the inhibition of Ste12 by the heterodimer Ste12-Tec1. This inhibition can be viewed as a resource draining since the formation of

the heterodimer depletes the available Ste12 monomers. This topology was also argued for in our previous study, where we showed that this resource depletion might not be sufficient to ensure specificity. There, we argued for the additional degradation of the heterodimer will be able to increase this resource depletion such that a desired level of signaling specificity can be achieved.

DISCUSSION

Finding network topologies that can ensure signaling specificity in the presence of shared components can, in principle, be accomplished through a careful analysis of the experimental data. Such a search, however, is most likely very time consuming. Furthermore, it does not guarantee finding all possible topologies. Thus, it is useful to determine the basic design principles underlying the allowed topologies for cross-activating pathways. Such a search for basic principles in pathway designs is not limited to the problem of signaling specificity and has recently been carried out for adaptive pathways [24].

In this study, we were able to determine a four-node module as the basic building block of cross-activating pathways (Fig. 1). This simple module can be analyzed easily, either by direct simulation and parameter sweeps [24, 25], or, as we have presented here, using a genetic algorithm. The result of this analysis shows that signaling specificity can be ensured by implementing two inhibitory links. These links break the symmetry and have to be placed at different hierarchical levels of the network. Hence, the resulting inhibition scheme can be classified as asymmetric hierarchical inhibition [11].

At first glance, it seems surprising that the genetic algorithm does not return a symmetric topology in which the output components inhibit each other. This reciprocal mutual inhibition module has been suggested as a mechanism to ensure specificity in several previous studies [7, 10, 26]. For cross-activation rates that are comparable to direct activation rates, however, this topology cannot ensure specificity. For example, specificity for the X branch requires a stronger inhibition of Y_{out} by X_{out} than vice versa, comprising the specificity of the Y branch. Thus, this scheme may be more applicable for the case where cross-activation is much weaker than direct activation of downstream proteins.

The analysis of the four-node module allowed us to construct a general approach to cross-activating pathways. Crucial in this approach is the recognition that the nodes of the pathway can be classified in one of three sets: an input set, an intermediate set, and an output set (Fig. 2). The input set is made up of those components that are only activated by the input signal of a single pathway. The output set contains the component that is most downstream in the pathway that is subject to cross-activation, along with any other components that are further downstream from this node. All components that are downstream of the input set but upstream from the output set fall into the intermediate set. On the basis of our analysis of the four-node module, we argue that the implementation of just two inhibitory links is sufficient to guarantee pathway specificity. The number of possible network topologies can then be calculated exactly and depends on the number of components in the different sets of the two pathways. The number of topologies that lead to signal specificity is in general much smaller than the total number of all topologies with two inhibitory links, especially for large numbers of components in the pathways.

Our results suggest a general approach to unravel the design principles in cross-activating signaling networks. The first step requires the identification of all pathway-specific signaling components for each signaling transduction pathway (i.e., the input sets). A component is pathway-specific if it is only present or significantly activated in one particular pathway. The second step consists of identifying all components in the output sets. Then, a coarse-grained diagram can be drawn, grouping the components into the three sets. Once such a diagram is drawn, one can enumerate all possible pathways and determine all specificity-ensuring topologies. These topologies can then be examined and compared to existing experimental data. Clearly, this process can be significantly easier than starting with a detailed analysis of the full network. This is true in particular for networks in which the number of pathways-specific components are small. In that case, only a limited number of faithful components with proper wiring can impart specificity and a small set of topologies needs to be examined.

We applied this general approach to one of the most carefully studied cross-activating system, the mating and filamentation pathways in yeast. After coarse-graining of the pathways, we were able to classify the components in each of the three sets using existing experimental data. For example, the MAPK Fus3 is mating-specific since its activation depends on the scaffold protein Ste5 and is thus part of the input set of the mating pathway. The filamentation MAPK Kss1, on the other hand, is nonspecific because it is activated in both mating and filamentation pathways. Furthermore, it is not the most downstream component that is cross-activated and is thus in the intermediate set. The most downstream cross-activated component of the filamentation pathway is Tec1 and is thus, together with Fus3, key in determining the possible network topologies that deliver specificity.

Our analysis resulted in the candidate network topologies presented in Fig. 3C. In principle, since the input set of

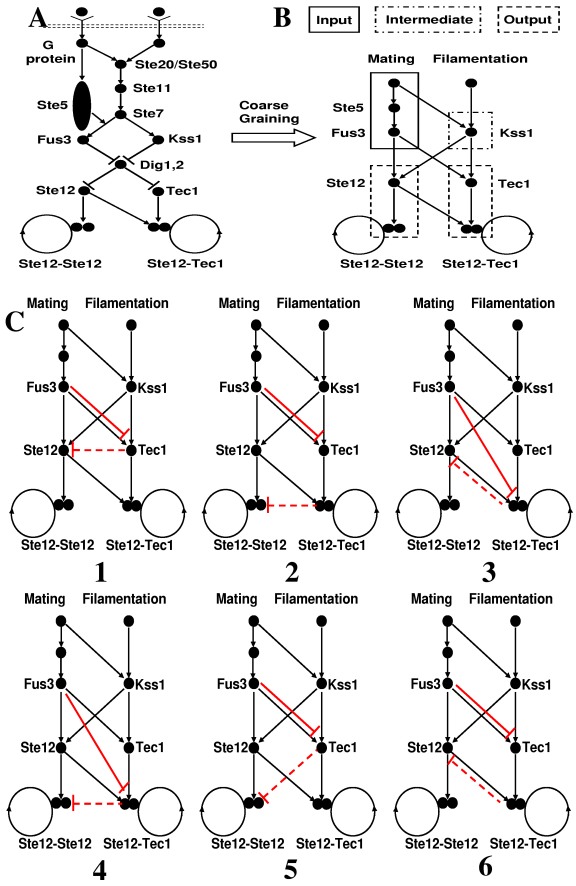


FIG. 3: (A) The MAPK-dependent mating and filamentous growth pathways in yeast, including loops representing the feedback in gene transcription. (B) The result of coarse-graining and grouping of the components of the pathways in A. (C) All possible topologies implementing the asymmetric hierarchical inhibition mechanism. The dominant inhibitory links is shown as a solid red line while the non-dominant inhibitory link is drawn as a dashed red line. Only topology 5 and 6 are possibly consistent with existing experimental data.

the dominant pathway contains more than one component, there are additional topologies that can ensure specificity. These pathways rely on an inhibitory link that connect an upstream component from Fus3 to a component in the output set of the filamentation pathway. An example of such an inhibitory link could involve a direct linking between the G protein and Tec1 or Ste12-Tec1. Since such links have, to our knowledge, not been reported in the experimental literature, we have omitted them in Fig. 3C. In fact, experiments have shown that a Fus3 mutant that is defective in suppressing Tec1 comprises signaling specificity [18, 19]. This suggests that the dominant inhibitory link, connecting the mating and filamentation pathway, cannot originate upstream from Fus3.

Of the six possible topologies, we were able to rule out four. In the two remaining candidates, number 5 and 6 in Fig. 3C, the dominant inhibitory link, ensuring filamentation inhibition upon pheromone stimulation, is between the mating pathway specific Fus3 and the filamentation pathway specific Tec1. Such inhibition has been demonstrated explicitly in experiments [18, 19]. The non-dominant link is either between the heterodimer Ste12-Tec1 and Ste12 (case 5) or between Tec1 and the homodimer Ste12-Ste12 (case 6). The inhibition in case 6 can be accomplished by a resource draining mechanism: the formation of the heterodimer Ste12-Tec1 uses almost all available Ste12 monomers such that the production of the homodimer Ste12-Ste12 is greatly suppressed. This mechanism was indeed proposed in our previous work, where we analyzed the network through detailed simulations of all components. The non-dominant inhibitory link in the additional allowed topology connects Tec1 with Ste12-Ste12. There is no direct evidence for such an interaction and it would be interesting to search for it through experimental studies.

Our analysis can be extended in several ways. Our current study only considered deterministic networks, ignoring possible stochastic effects. It would be interesting to explore how our design principles are affected by the presence of fluctuations. Furthermore, we have focused on cross-activating pathways where one of the pathways is dominant.

The case in which dual activation leads to high output levels of both pathways can be analyzed in a similar manner as discussed here. Finally, we have restricted ourselves to two parallel pathways but our analysis can be easily extended to more than two pathways. An example of such an extension could be the inclusion of the high osmolarity glycerol (HOG) pathway in yeast [3, 26].

METHODS

The four-node ODE model

The basic four-node topology with cross-activation is represented by four general ordinary differential equations (ODEs) for the four variables describing the activity of the network:

$$\frac{dX_{in}}{dt} = I_1 - [d_{xin} + g_1(Y_{in}) + g_2(Y_{out})]f_1(X_{in})$$
$$\frac{dY_{in}}{dt} = I_2 - [d_{yin} + g_3(X_{in}) + g_4(X_{out})]f_2(Y_{in})$$
$$\frac{dX_{out}}{dt} = f_3(X_{in}) + f_4(Y_{in}) - [d_{xout} + g_5(Y_{in}) + g_6(Y_{out})]f_5(X_{out})$$
$$\frac{dY_{out}}{dt} = f_6(Y_{in}) + f_7(X_{in}) - [d_{yout} + g_7(X_{in}) + g_8(X_{out})]f_8(Y_{out})$$

where I_1 and I_2 represent the input signals and where the parameters $\{d_{xin}, d_{xout}, d_{yin}, d_{yout}\}$ denote the basal degradation rates of the four components. Here, the activation and inhibition functions, f_i and g_i, are either linear, $f_i(x) = s_i x$ and $g_i(x) = k_i x$, or sigmoidal Hill functions, $f_i(x) = s_i x^n/(x^n + s_{th}^n)$ and $g_i(x) = k_i x^n/(x^n + k_{th}^n)$. Here, n represents the Hill coefficient and we should note here that, according to a recent study, ultrasensitive responses do not guarantee signaling specificity but can enhance the performance of certain networks [27]. For all values of n we have examined ($n \leq 5$), the qualitative results remain unchanged. For both functional forms, the inhibitory links are determined by one nonnegative parameter k_i ($i = 1, 2, ..., 8$) while the activation is determined by s_i ($i = 1, 2, 3, 4$), chosen to be equal to 1. The parameters s_{th} and k_{th} determine the threshold values of the inhibition and activation.

The network genetic algorithm

A network version of genetic algorithm is designed to search for optimal topologies for certain functional goal [12, 13]. Based on our four-node ODE model, the algorithm starts from a population of random realizations of the eight strength parameters $\{k_1, k_2, ..., k_8\}$. For every realized topology, we use Matlab 7.0 to solve the corresponding ODEs under three different input conditions: (a) $I_1 = 1$ and $I_2 = 0$; (b) $I_1 = 0$ and $I_2 = 1$; (c) $I_1 = I_2 = 1$. The steady-state solutions are then used to evaluate a fitness score, defined as $F = w(S_X + S_Y + C_{XY}) - \sum_{i=1}^{8} k_i$, which measures the functional performance of a realized topology. Here, we have introduced two metrics for signaling specificity (S_X and S_Y) [7, 28] and one metric for pathway dominance (C_{XY}); they are $S_X = \log(X_{out}/Y_{out})$ under condition (a), $S_Y = \log(Y_{out}/X_{out})$ under condition (b), and $C_{XY} = \log(X_{out}/Y_{out})$ under condition (c). To filter out dispensable links, the fitness function also incorporates a cost term represented by the total strength of all the inhibitory links, $\sum_{i=1}^{8} k_i$. To further balance cost and performance, a weight parameter w was found to be helpful.

All parameters are fixed throughout except $\{k_1, k_2, ..., k_8\}$ which are allowed to vary independently in the space $[0, k_{max1}] \times [0, k_{max2}] ... \times [0, k_{max8}]$. The simulation is started using 1000 topologies with randomly chosen values of k_i. The algorithm then sorts all the topologies based on their fitness scores. A fixed fraction (e.g. 50%) of best-fit topologies are kept as seed configurations for reproduction in the next generation, where the parameters $\{k_1, k_2, ..., k_8\}$ encoding every seed are subject to random mutation, that is, increasing or decreasing each k_i by a small amount δk with equal probability. We repeat the procedure by evaluating the fitness scores for the newly generated topologies and selecting the top 50% best-fit ones as seeds for another round of reproduction. To monitor the process, we calculate the average fitness for the top 50% networks. The algorithm stops when it reaches a stable average fitness value. We expect that the optimal network solutions are encoded by the parameters $\{k_1, k_2, ..., k_8\}$ that correspond to this stable fitness value. For the case where Pathway X is dominant, the algorithm returns a unique solution $\{k_1, k_2, ..., k_8\}^* = \{0, 0, 0, 0, 0, +, +, 0\}$, corresponding to the scheme of asymmetric hierarchical inhibition in Fig. 2B. We have verified that this solution is independent on the specific choice of the parameter values or on the functional forms of the biochemical interactions.

ACKNOWLEDGMENTS

This research has been supported by the NSF-sponsored Center for Theoretical Biological Physics (NSF PHY-0822283).

[1] Schwartz, M. A. & Baron, V. Interactions between mitogenic stimuli, or, a thousand and one connections. *Curr Opin Cell Biol* **11**, 197–202 (1999).
[2] White, M. A. & Anderson, R. G. Signaling networks in living cells. *Annu Rev Pharmacol Toxicol* **45**, 587–603 (2005).
[3] Schwartz, M. A. & Madhani, H. D. Principles of MAP kinase signaling specificity in Saccharomyces cerevisiae. *Annu. Rev. Genet.* **38**, 725–748 (2004).
[4] Seger, R. & Krebs, E. G. The mapk signaling cascade. *FASEB J* **9**, 726–735 (1995).
[5] Chang, L. & Karin, M. Mammalian map kinase signalling cascades. *Nature* **410**, 37–40 (2001).
[6] Stork, P. J. & Schmitt, J. M. Crosstalk between camp and map kinase signaling in the regulation of cell proliferation. *Trends Cell Biol* **12**, 258–266 (2002).
[7] Bardwell, L., Zou, X., Nie, Q. & Komarova, N. L. Mathematical models of specificity in cell signaling. *Biophys. J.* **92**, 3425–3441 (2007).
[8] Chen, W., Levine, H. & Rappel, W. J. A mathematical analysis of second messenger compartmentalization. *Phys Biol* **5**, 046006 (2008).
[9] Chen, W., Levine, H. & Rappel, W. J. Compartmentalization of second messengers in neurons: a mathematical analysis. *Phys Rev E* **80**, 041901 (2009).
[10] McClean, M. N., Mody, A., Broach, J. R. & Ramanathan, S. Cross-talk and decision making in MAP kinase pathways. *Nat. Genet.* **39(Suppl.)**, 409–414 (2007).
[11] Hu, B., Rappel, W. J. & Levine, H. Mechanisms and constraints on yeast mapk signaling specificity. *Biophys J* **96**, 4755–4763 (2009).
[12] Francois, P. & Hakim, V. Design of genetic networks with specified functions by evolution in silico. *Proc Natl Acad Sci U S A* **101**, 580–585 (2004).
[13] Francois, P., Hakim, V. & Siggia, E. D. Deriving structure from evolution: metazoan segmentation. *Mol Syst Biol* **3**, 154 (2007).
[14] Kauffman, S. *The Origins of Order: Self-Organization and Selection in Evolution* (Oxford University Press, New York, 1993).
[15] Lau, K.-Y., Ganguli, S. & Tang, C. Function constrains network architecture and dynamics: A case study on the yeast cell cycle boolean network. *Phys Rev E* **75**, 051907 (2007).
[16] Liu, H., Styles, C. A. & Fink, G. R. Elements of the yeast pheromone response pathway required for filamentous growth of diploids. *Science.* **262**, 1741–1744 (1993).
[17] Roberts, R. L. & Fink, G. R. Elements of a single MAP kinase cascade in Saccharomyces cerevisiae mediate two developmental programs in the same cell type: mating and invasive growth. *Genes Dev.* **8**, 2974–2985 (1994).
[18] Chou, S., Huang, L. & Liu, H. Fus3-regulated Tec1 degradation through SCFCdc4 determines MAPK signaling specificity during mating in yeast. *Cell.* **119**, 981–990 (2004).
[19] Bao, M. Z., Schwartz, M. A., Cantin, G. T., Yates III, J. R. & Madhani, H. D. Pheromone-dependent destruction of the Tec1 transcription factor is required for MAP kinase signaling specificity in yeast. *Cell.* **119**, 991–1000 (2004).
[20] Zeitlinger, J. *et al.* Program-specific distribution of a transcription factor dependent on partner transcription factor and MAPK signaling. *Cell.* **113**, 395–404 (2003).
[21] Flatauer, L. J., Zadeh, S. F. & Bardwell, L. Mitogen-activated protein kinases with distinct requirements for Ste5 scaffolding influence signaling specificity in Saccharomyces cerevisiae. *Mol. Cell Biol.* **25**, 1793–1803 (2005).
[22] Good, M., Tang, G., Singleton, J., Remenyi, A. & Lim, W. A. The ste5 scaffold directs mating signaling by catalytically unlocking the fus3 map kinase for activation. *Cell* **136**, 1085–1097 (2009).
[23] Chou, S., Zhao, S., Song, Y., Liu, H. & Nie, Q. Fus3-triggered Tec1 degradation modulates mating transcriptional output during the pheromone response. *Mol. Syst. Biol.* **4**, 212 (2008).
[24] Ma, W., Trusina, A., El-Samad, H., Lim, W. A. & Tang, C. Defining network topologies that can achieve biochemical adaptation. *Cell* **138**, 760–773 (2009).
[25] Ma, W., Lai, L., Ouyang, Q. & Tang, C. Robustness and modular design of the drosophila segment polarity network. *Mol Syst Biol* **2**, 70 (2006).
[26] Yang, H. Y., Tatebayashi, K., Yamamoto, K. & Saito, H. Glycosylation defects activate filamentous growth kss1 mapk and inhibit osmoregulatory hog1 mapk. *EMBO J* **28**, 1380–1391 (2009).
[27] Haney, S., Bardwell, L. & Nie, Q. Ultrasensitive responses and specificity in cell signaling. *BMC Syst Biol* **4**, 119 (2010).
[28] Komarova, N. L., Zou, X., Nie, Q. & Bardwell, L. A theoretical framework for specificity in cell signaling. *Mol. Syst. Biol.* **1**, 2005.0023 (2005).