Title
Reconstituting Signaling Systems to Interrogate the Molecular and Evolutionary Logic of Cellular Decision Making

Author
Coyle, Scott Michael

Publication Date
2015

Peer reviewed|Thesis/dissertation

Reconstituting Signaling Systems to Interrogate the Molecular and Evolutionary Logic of Cellular Decision Making

by

Scott Michael Coyle

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry and Molecular Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Acknowledgements.

My Ph.D. and dissertation would not have been possible without the help and support of many incredible people whom I have been so fortunate to have help me in this challenging journey in my life. My initial interests in becoming a scientist were nurtured by a number of extraordinarily dedicated high school teachers: Mr. Landfear, Mr. Bortz, and Ms. Tate. In particular, Ms. Tate's AP biology class cemented my interest in becoming a biochemist, and she helped me get started on my first research experiences by encouraging me to enter a local science fair and eventually apply for a research internship after graduating from high school.

As an undergraduate, I was incredibly fortunate to be mentored by remarkably talented scientists. While working at the Molecular Sciences Institute, Evgueny "Jenya" Kroll was my first serious scientific mentor. A self-described "nasty Russian", Jenya nonetheless treated me with an unusual amount of respect for someone so inexperienced, and was instrumental in me learning to think like a scientist. I also credit him with my longstanding addition to espresso. After leaving the Molecular Sciences Institute I began to work in Jennifer Doudna's lab with Wendy Gilbert, a postdoc at the time. Wendy is one of the most brilliant people I've ever had the good fortune to work with, and her hands-on training helped me to become both the talented thinker and experimentalist I am today. Getting to work with Jennifer was also probably one of the most exciting opportunities of my life. She is truly an awe-inspiring scientist, not only in terms of her success, but in her earnestness about the niche of biology she loves the most.

All of these individuals prepared me for success in graduate school, where I also "hit the jackpot" in getting to work with Wendell Lim. Wendell is truly one-of-a-kind: he's a jack of all trades that I could talk to on a serious level about evolution, biochemistry, protein structure, the immune system, mathematics (and more!). His ability to move seamlessly between fields, modes of thinking, and scales has had a huge impact on how I think about science. He also has treated me with a tremendous amount of respect, allowing me to pursue research that I thought was interesting even if it wasn't always immediately aligned with what he was most interested in. He also taught me the value of visuals and aesthetics in communicating research, and has turned me into an Adobe Illustrator master who knows the true value of the golden ratio. Outside of lab, he has always been supportive about my personal relationships and my life. To paraphrase Wendell's own evaluation of his Ph.D. Advisor Bob Sauer, I would say that Wendell has been, more or less, a great mentor.

Wendell also did a great job in building a lab full of diverse and exciting people to work with and to learn from. My time in the lab would have been a whole lot less interesting had the following people not been with me on my journey: Jessica Walter, Reid Williams, Andy "The Chairman" "Tongxhi" Ping Wei, Elias Puchner, Thomas Stevens, Jesse Zalatan, Russell Gordley, Geoff "Screw McIlroy" O'Donoghue, Joan Garbarino, Amir "Coffee Czar" Mitchell, Kole Royal, Leo Morsut, Jaline Gerardin, and our incredible administrator Noleine Blizzard.

Arsenia deGuzman, our media technician, made so much media for me during my Ph.D. it's not even funny. I could probably fill several Olympic sized swimming pools with LB thanks to her, and she never once complained about my usage. This work would have almost surely taken another year without her generous help.

I'm also grateful for all my classmates who have helped make my time here at UCSF so much more pleasant. When I first started graduate school I was very grumpy and hated everything. That's why it was so critical to meet people like Lindsey Pack, Stefan Isaac, and Phillip Dumesic that also hated everything, and more specifically the same everything. These are the people I always felt like I could talk to about anything, and most importantly, talk about things that weren't science related. Beyond these top 3, the class on the whole was great, and we were always finding fun things to do like intramural volleyball. I will miss seeing all of these people every day as I move forward in my life.

Outside of UCSF, my family has been so supportive of my adventures in science. My mom, dad and sister have not always understood what I've been up to, but they've never stopped encouraging me to play with fungus and work seven days a week year-round. They force me to unplug from work around the holidays which always ends up being way more fun than I think it's going to be. Beyond that they are incredible kind and generous people who I am so proud to be related to. I love you guys!

Last but absolutely positively not least, I am so so so thankful to have Amy Weeks in my life. It's truly remarkable to think that after some big explosion 14 billion years ago, two living beings so perfectly in tune with one another could arise from primordial soup and find one another. We are truly on the same wavelength on everything in our lives (including our unusual fondness for the Cheetah), which is truly the most wonderful and incredible feeling. Amy has been the best supporter in my life I could ever ask for. She has been the best friend in my life I could ever ask for. And on top of it all, she's an incredible scientist who I can turn to for help and understands the struggle of the scientific life. I love you so much Amy! Happiness *is* indeed building our life together! I dedicate this dissertation to you!

Acknowledgment of previously published work in this dissertation.

Portions of Chapter 2 were published in the following scientific journal article:

Zalatan, J.G.*, Coyle, S.M.*, Rajan, S., Sidhu, S.S., and Lim, W.A. "Conformational control of the Ste5 scaffold protein insulates against MAP kinase misactivation." *Science* 337 (6099): 1218-1222 (2012).

*denotes equal contribution of the authors.

Portions of Chapter 3 were published in the following scientific journal article:

Coyle, S.M., Flores, J., and Lim, W.A. "Exploitation of Latent Allostery Enables the Evolution of New Modes of MAP Kinase Regulation." *Cell* 154, 875-887 (2013).

Portions of Chapter 4 have been submitted for publication as a scientific journal article.

Abstract.

Reconstituting Signaling Systems to Interrogate the Molecular and Evolutionary Logic of Cellular Decision Making

by

Scott Michael Coyle

The ability to make sense of the complex sea of ever-changing molecules in the environment is one of the cell's most remarkable abilities. The molecular mechanisms by which these external cues are processed are well understood for many individual signaling pathways. However, fundamental questions for cell biology remain about how the collection of all such pathways within a cell—the signaling network—correctly responds to many inputs it receives. This dissertation focuses not only on determining the detailed biochemical mechanisms that resolve the complexities and ambiguities of individual cell signaling networks, but also on how these

networks can change or grow to accommodate new pathways during evolution. To this end, reconstituted signaling networks are developed to probe the behavior of signaling systems under different environmental conditions or network compositions, as well as to rigorously compare evolutionarily related systems. Two different ambiguous network structures from model systems are explored from this perspective: overlapping signaling pathways arising from duplication/divergence events, derived from fungal ERK kinase signaling; and central signaling nodes that respond to multiple inputs and fan out to many possible outputs, derived from small Ras GTPase signaling. In the former case, allosteric dependencies on pathway-specific scaffold proteins were found to distinguish evolutionarily related molecules from one another and facilitate the use of homologous molecules in distinct signaling pathways. These allosteric dependencies appear to have evolved by exploitation of pre-existing differences in the conformational landscapes of otherwise-equivalent redundant signaling molecules. In the latter case, the behavior of Ras systems was explored in a multi-turnover *in vitro* setting for the first time. This revealed that Ras systems could transmit both sustained and transient signals and that the concentration and identity of signaling components strongly impacted the timing, duration, shape and amplitude of the output. Moreover, the extent to which oncogenic mutations in Ras distorted outputs was highly dependent on this underlying network configuration. Together, these studies demonstrate the utility of examining signaling systems not only from their current configuration, but the paths that led to that configuration during evolution, and the paths that might be taken through perturbation in the future.

Table of Contents.

List of Tables.

List of Figures.

Chapter 1. Introduction.

The remarkable ability of cells to sense their environment, process that information, and execute specific biological responses is one of the most fascinating aspects of living systems (Fig. 1A). It is now well understood that this information processing that takes place inside cells is performed by signal transduction pathways: cascades of enzymatic activities that modulate how molecules function and that eventually target the transcriptional or translational machinery to cause changes in gene expression and cellular behavior (Fig. 1B) (*1*–*3*). While the molecular mechanisms of information processing are well understood for many individual pathways, fundamental questions for cell biology remain about how the collection of all such pathways within a cell—the signaling network—correctly responds to the plethora of competing inputs it receives.

This would not be an issue if every individual signaling pathway was built from unique, completely orthogonal components with no possibility for cross-talk, but this is not the case. In fact, the evolutionary processes that have built and shaped these networks have produced signaling networks in which pathways overlap, divergent components have high similarity, and for which the connectivity of components alone cannot predict signaling behaviors *prima facie* (*4*–*9*). For such networks, additional biochemical mechanisms must exist that resolve the ambiguities that are a consequence of the system's evolution (*10*, *11*). This highlights the important relationship between the evolutionary mechanisms which give rise to the molecules cells use to make or expand decision making systems and the molecular mechanisms that are required to ensure systems built from these processes are functional.

The aim of this thesis is to characterize and understand the networks that underlie cellular decision making from the perspective of these two complimentary and interdependent vantage

points: biochemical mechanism and evolutionary processes. To gain insight into these systems, decision making networks are reconstituted *in vitro* and these purified systems are used to systematically explore what types of outputs interacting molecular components produce under different environmental conditions, with different component concentrations, or in the presence of adaptor or scaffold molecules. This approach also enables one to rigorously examine and compare the behavior of evolutionarily related signaling systems, as well as hybrid systems built from a mixture of components from highly divergent species. Together, these approaches are used to gain deeper insights into the nature of these complex systems than is achieved from a single perspective.

Below, the present understanding of cellular decision making networks is reviewed from biochemical, systems, and evolutionary perspectives. This reveals deficiencies in our understanding at the network level that motivate the particular questions and projects that make up the experimental work of the remainder of the thesis.

Biochemical and Molecular Mechanisms Underlying Information Transmission.

The cell signaling literature frequently speaks of "information transmission", but what does this actually mean and how, at a molecular level, is it achieved? If a cell is in some environment and that environment changes, how does a cell acquire information about that change? That environmental change must be converted into a change in state of a molecule that alters its behavior. How can a macromolecule change state and how can such changes lead to a change in behavior that allows for propagation of information?

Changing states to change activity.

The two primary ways in which macromolecules like proteins can change state are through *covalent modification* and *interaction with other molecules*, which are not at all mutually exclusive and indeed are often coupled. Covalent modification involves the direct chemical alteration of target protein. There are many different forms of covalent modification, including phosphorylation, methylation or ubiquitylation (*12–14*). These modifications can be installed or removed by enzymes, and the modification cycle almost invariably is coupled to the consumption of energy. Proteins can also change state through interaction with other molecules. These can be other macromolecules like a protein binding partner, but can also take the form of interaction with small molecules like nucleotides or metabolites. In this case, these changes are not inherently coupled to consumption of energy.

These distinct ways of changing a protein's state can result in a change in molecular activity through *changes in localization* and *conformational rearrangements*, which again are not mutually exclusive and are often coupled. These concepts are further detailed below.

Change in localization

One of the simplest ways that a change in state of a protein can lead to a change in activity is through a relocalization of the target. Such relocalization events can change what substrates are accessible to an enzyme as well as what other potential regulatory interactions are available to further control its activity (*15*).

In the simplest case, relocalization can be mediated directly by interaction with another molecule. For example, if protein X is normally cytoplasmic, but can interact with protein Y which is only present at the plasma membrane, then the subcellular localization of X will be regulated by the presence of Y. These changes in localization can also be mediated by covalent

modifications. For example, if protein X becomes phosphorylated and this modification can be recognized by phosphopeptide binding modules that are present only at a specific subcellular location (such as the plasma membrane), then this phosphorylation event will control the subcellular distribution of X.

Change in internal protein state: conformational rearrangements

Macromolecules are complex dynamic polymers that can sample many different conformational states within their conformational energy landscape (*16*, *17*). Thus, which of these states dominates the ensemble can be altered by external forces. This enables both covalent modifications as well as interactions with other molecules to influence protein conformational state.

For example, if protein X has a binding partner that favors interaction with a minor conformational state of X, this can shift the distribution in such a way that the minor state now predominates the ensemble of X conformers. In this case, the energetic cost of shifting X into its less favored conformation is offset by the favorable energy of binding, and the maintenance of this change in conformational state of X is dependent on the sustained interaction with the binding partner. Numerous examples of this type of regulation have been observed, including the regulation of Cyclin-dependent kinases by Cyclins, the regulation of the Fus3 MAPK conformation by the scaffold protein Ste5, and the regulation of kinase activity in complexes of Raf kinase and the pseudokinase KSR (*18–20*).

Conformational rearrangements can also be mediated by covalent modifications (*21*, *22*). This is because the modified form of the protein may possess a different conformational energy landscape compared to the unmodified form. Thus, following modification the protein can

explore the new energy landscape and eventually settle into a new conformational state. It is worth noting that in contrast to a rearrangement induced by interaction with a binding partner, this mode of creating a conformational rearrangement does not depend on sustained interaction with other molecules.

Each of these mechanisms that can produce conformational rearrangements can be used to influence a target protein's activity, a process broadly referred to as *allosteric regulation*. This is possible because different conformations of a molecule can have differing degrees of activity owing to different arrangements of the active site or different abilities of alternative conformers to interact with other molecules (*23–26*).

Assembly of Primary Signaling Complexes and Propagation by Enzymatic Cascades

Having seen the primary modes by which macromolecular states can be altered as well as the ways in which those state changes can result in a change in activity, how do cells build signaling systems with this toolbox? Ultimately a change in some extracellular parameter must culminate in a change in the activity of a particular subset of internal molecules, such as transcription factors, that will in turn mount the intended cellular response. It is the job of the signaling pathway to relay that change in external state to the appropriate terminal players. There are two main modes by which such systems function: the assembly of primary signaling complexes and propagation by enzymatic cascades.

The assembly of primary signaling complexes is a critical first step in many signaling systems (*15*). During this phase, the binding of an extracellular factor to a transmembrane receptor results in a change of state in the cytoplasmic side of the receptor. In many cases, this

change of state either directly or indirectly facilitates the recruitment and assembly of signaling molecules at the plasma membrane. Once assembled, these signaling molecules become active and can begin to alter the state of additional downstream signaling molecules by the mechanisms outlined previously. Examples of this type of process include the assembly of Grb2/SOS/Ras/Raf signaling centers in response to EGFR signaling in mammalian systems as well as the formation of Ste4/Ste18/Cdc42/Ste20/Ste11/Ste5 signaling centers in response to pheromone stimulation in budding yeast (*27*, *28*). One potential function of these assembly processes is to act as "kinetic filters" against noise by requiring multiple factors to simultaneously be assembled in the center before secondary propagation of the signals begins (*29*).

Propagation of signals by enzymatic cascades is the other classical mode by which signaling systems transmit information (*3*). In this scenario, an enzyme that becomes activated by any of the previously discussed mechanisms can go on to catalyze the activation of multiple copies of downstream signaling components. These multiply activated molecules can then go on to catalyze the activation of multiple copies of their downstream targets, and so on and so forth. MAP Kinase signaling cascades are classic examples of this type of signal propagation (*30*). An interesting feature of these schemes is that they act to amplify the initial signaling event. For example, if a single MAPKKK is activated, it might go on to activate 1000 MAPKKs, each of which could in turn activate 1000 MAPKs. Thus a single MAPKKK activation could potentially lead to the activation of 1,000,000 MAPKs. This is in contrast to the assembly processes discussed previously in which, for example, an activated molecule of Ras can only template the assembly of a single effector molecule at any given time.

Both assembly and propagation processes can work together in the processing of extracellular signaling and transmission of information. For example, in EGFR signaling,

assembly processes template the formation of Grb2/SOS/Ras/Raf signaling centers in response to extracellular ligand binding, and the activated Raf kinase propagates these signals via an enzymatic cascade in which it activates the MAPKK MEK, which in turn can activate the MAPK ERK (*31*, *32*).

Having seen the basic mechanisms by which the state of molecules can be changed, how those changes can result in changes in activity, and the basic processes by which changes in activity can be used to transmit and propagate information, the stage is now set to understand how these mechanisms can work together in the context of networks of molecules to execute specific signal processing behaviors and dynamic programs.

The Dynamics of Biochemical Signaling Networks and Systems.

In an homage to electrical circuits, signaling networks are typically schematized as a collection of *nodes* and *arrows* that connect a given extracellular input to some output (Fig. 1B). The nodes in these "wiring diagrams" correspond to any number of molecular actors that can function to transmit information – such as protein kinases or small GTPases – and the arrows correspond to regulatory relationships between these components (e.g. Node A activates Node B). For any particular signaling phenomenon of interest, a variety of classical genetic and biochemical approaches can be employed to produce these wiring diagrams that showcase the connectivity of the components.

However, unlike with electrical circuits– for which it is often possible to determine the time-evolution system behavior analytically or numerically from the diagram– the predictive utility of abstract molecular wiring diagrams is not obvious. This is owing to (i) a lack of

sufficient mechanistic understanding of the regulatory relationships that are the actual content of the arrows in these diagrams, (ii) a lack of attention to the embedding of individual signaling pathways in the more complicated, often overlapping, signaling network, (iii) treatment of the wiring diagram as a single static structure as opposed to a representation of a collection of related signaling pathways with different possible parameter configurations that can be realized in different ways across different cell types or different species.

Mechanism is Critical to Understand System Dynamics

Many classical methods that are used to determine network wiring diagrams result in an incomplete understanding of the mechanistic underpinnings of the relationships between components that are required to predict dynamic systems behaviors. For example, a genetic knockout might imply that Protein A activates Protein B, but provides no quantitative information about the kinetic scheme underlying that activation. Even when quantitative data is available, it often takes the form of macroscopic parameters derived from steady-state rate measurements, such as K_M or k_{cat}. While useful, these macroscopic parameters do not contain information about the underlying microscopic kinetic scheme associated with each process that is needed to understand how a biochemical system will time-evolve.

It is in fact quite easy to see why this is the case. For example, consider a simple 1-node signaling system in which the node N can be activated by an enzyme A and N deactivates by some intrinsic mechanism (Fig. 2A). If [A] is taken to be the input, and some amount of [A] is applied to the system, what will the dynamics of the system output be? The answer is in fact highly dependent on the nature of the mechanisms by which node N can be activated and

deactivated. In the simplest scenario, N can exists in exactly one active or one inactive state. In this case, activated N will accumulate monotonically to some steady-state at some rate governed by the balance of the activation rate by A and intrinsic deactivation rate (Fig. 2B). In a second scenario, deactivation takes place in two steps and can only be reactivated by A after both steps are complete. In this case, there exist many parameter sets in which activation of N can appear to show adaptive overshoot behavior (Fig. 2C).

The reason for this is the presence of more than two states in the N molecule and the coupling of exchange between those states to the consumption of energy by enzymatic transformation (*33*). This results in a pre-steady state behavior to the system in which intermediate states initially accumulate during the initial reaction cycle. This is analogous to trajectory of states within in an enzyme that shows burst-phase kinetics when substrate is added (*34*).

It can be argued of course that what this really means is that the original one-node diagram needed to be expanded to include these additional complexities that were inside the original signal node. While that is certainly true, this only reinforces the point being made that the original diagram based on connectivity and steady-state parameters was in fact compatible with wildly different systems dynamics and behaviors. This suggests that such a diagram should be treated with extreme caution when used as the basis for any sort of quantitative modeling, and instead should be best viewed as a blueprint for more detailed mechanistic experiments that can reveal the actual contents of the diagrams and the resulting behaviors that the system produces.

Intersection with other pathways within the network influences output behaviors

Abstract wiring diagrams are often constructed in the narrow context of a single input/output relationship, but real cellular decision making systems must function correctly in the context of many possible inputs, many possible signaling pathways, and many possible outputs. To see how neglecting other interactions can be problematic, consider the following scenario. Suppose an assembly-driven signaling pathway produces an exponentially decaying upstream signal S that a downstream target T must assemble on to produce an output. In isolation, the TS signaling complex will show some dynamics based on the amount of T present in the system and its assembly and disassembly kinetics.

Now suppose S can also interact with other downstream targets T' and T'' from pathways we were not initially considering. How does the presence of T' and T'' impact the apparent signaling dynamics to T? Many outcomes are possible depending on the relative abundances of T, T', and T" and their relative assembly and disassembly kinetics. These parameters will determine how the *competition* between the various targets time-evolve, and will involve first a kinetically-determined phase followed by a thermodynamically-determined phase. In particular, it is possible for a target that will not be the thermodynamic "winner" of the competition to nonetheless outperform other targets in the early pre-steady stage of the equilibration.

Network behaviors must be understood with respect to all possible configurations

It is common to see statements in the literature that treat signaling pathways as static, singular entities such as "the Ras/ERK pathway" or "the p38 signal transduction pathway" (*35*, *36*). Yet, two different cell types can contain all of the molecular machinery necessary to transmit signals from the same receptor to Ras then to MEK then to MEKK then to ERK, they may do so with

drastically different dose-dependent behavior and dynamics. This can easily be a consequence of the configuration of these components as well as their relationship to other signaling components that may or may not be present in the cell. In this sense there is no single "Ras/Erk pathway" but rather a collection of possible pathways that can be generated from these components. Thus, when one aims to understand how the basis by which some signaling network functions, it is important to consider not only the particular configuration of interest, but how the system functions in alternative configurations that may be realized in different cell types, species, or in response to other types of perturbation. This approach becomes even more critical as one begins to consider the systems of interest in the context of the evolutionary processes from which they arose.

Evolutionary Processes that Shape Signaling Networks.

The final perspective with which to consider signaling networks from is that of the evolutionary processes that have shaped their structure and function. Indeed, given that designed systems show almost none of the ambiguous or overlapping architectures that are readily observed in biological systems, this perspective has the potential to illuminate the tricks that have been employed to expand signaling repertoires and cellular decision making prowess.

Evolutionary expansion of signaling systems involves two main processes. The production of new signaling components through *duplication/divergence mechanisms*, and *the repurposing of existing components* for new applications (*37*, *38*). Moreover, these processes must operate in such a way such that the existing systems remain functional during the transition to a new form, placing constraints on the accessible paths through evolutionary space. This is an

important distinction between engineered or designed systems, in which multiple changes can be implemented simultaneously to produce a functional form, even if any piecemeal form would be completely non-functional.

Duplication/divergence: generating new components

The duplication of existing biological components and subsequent divergence into components with new function is a widely recognized evolutionary mechanism by which novelty is generated in biological systems (*38*). Cellular decision making systems are no different, in that a relatively small number of useful activities that can transmit information have been duplicated countless times to produce the staggering array of signaling machinery we observe in nature. This is evident from the enormous number of serine/threonine protein kinases, tyrosine kinases and small GTPases in metazoan genomes (*4*, *5*, *7*).

However, how cells can use component duplication to produce new signaling components to evolve new signaling pathways is not immediately obvious. This is because, as discussed in the previous sections, it is not merely the activity of a signaling molecule that allows it to be a useful conduit for information transmission, but the context with which its activity is regulated by interaction with other molecules. This means that when a component is duplicated, it inherits the pre-existing regulatory connections of the component from which is derived. Indeed, *bona fide* new signaling components that arise from duplication will have to shed their existing regulatory relationships and acquire new regulatory relationships with other signaling components in order for new orthogonal signaling pathways to form (*10*, *11*).

Repurposing of existing components

Another way that cellular signaling systems can acquire new behavior is through repurposing of existing components (*37*). Unlike with duplication/divergence, these mechanisms do not seek to produce new, orthogonal components. Instead, the same components are used to produce a different type of behavior owing to a change in the context with which those components execute their functions. This could be something as simple as a change in the expression of level of the components, for example a change in the concentration of activator and deactivator that changes the rate limiting step for information transmission. In other cases, a substrate or downstream effector that normally is not expressed in a network might now be expressed, altering the interplay between network components.

This mode of altering signaling system behavior is, in a sense, simpler than the duplication/divergence paradigm given that far fewer events are required to take place in order for a system to explore new possibilities. Changes in gene expression can be realized by point mutations in promoter elements that alter mRNA expression or in primary protein sequence that increase or decrease stability. However, it is not immediately obvious how much diversity a system can generate using simple titration of network components, and the space of signal processing behaviors available to a system likely depends on the specific molecular and mechanistic features of the components being used. Nevertheless, there are numerous examples of the same components being reused and repurposed across numerous signaling pathways over and over again, for example the small Ras GTPases (*4*, *7*). It remains to be explored what features about these types of systems have made them particularly adept at being adapted for so many signaling applications, and whether there are any trade-offs associated with this versatility.

Drift and selection work together to shape signaling networks

Having seen the two main modes by which signaling networks can change and acquire new functionality, how are those modes realized during the evolution of cellular decision making networks. Both *selection* as well as *neutral drift* play critical roles in how networks can be transformed.

It is easy to see how selective processes can refine a network that has a desirable feature. For example, if a signaling network could previously only distinguish between one extracellular input, but now has some primitive ability to distinguish between two extracellular inputs, selective processes can refine this system to improve discrimination. Similarly, if a signaling system shows a mildly adaptive response, and that adaptively provides a selective advantage to the organism, selection can improve that system to be more adaptive. In each of these two cases, however, selective processes had a toehold with which to operate: the system already had *some* ability to sense a new extracellular input; the output dynamics already had *some* adaptive behavior.

This raises the question as to how these toeholds for new behaviors first appear in the system. This is particularly salient given that many systems with an interesting behavior require a variety of additional regulatory controls to be functional. For example, if a system requires positive feedback to be functional, how could the system ever find a viable path to arrive at the target state? Or, if new regulatory relationships are required to orthogonalize evolutionarily related signaling components, how could the duplicants ever initially provide a path to generating new signaling pathways?

Neutral evolutionary processes, like drift, may provide some of the answer. In neutral drift, components may change, their expression levels might change, or the network context might change, but without any immediate consequences to system behavior. For example, if a

steady state level of activation is held in place by 1 unit of activator X and 1 unit deactivator Y, the same steady state might be held in place by 10 X and 10 Y, 100 X and 100 Y, or 1000 X and 1000Y. This means that the system may be able to drift between these different, but apparently equivalent, configurations without any immediate consequences to the system. However, while these configurations may only be equivalent *locally*, they may be highly divergent in how they respond to some other type of perturbation to the system like a point mutation or a change in expression of some other factor. In particular, one configuration might be much closer to a novel behavior (for example adaptation) than another. Once a perturbation provides access to even a minor novelty, selective processes can begin to operate. It is important to note that these selective processes can move the system very far from the initial configuration that bordered the old and new behaviors, and thus it may not be obvious from the extant configuration how it was possible to travel from one behavior to another.

These observations highlight the importance of being able to explore how components and networks behave under completely arbitrary conditions or configurations. Duplicated components that appear equivalent might be very different in ways that have not yet been realized by both scientists as well as nature and the cell. Two signaling networks that appear identical may be poised to respond divergently to a point mutation. Understanding how signaling systems can change and how they respond to perturbation, what the structure of that space is, and what the maneuverability within that space is like can provide insights not only into how extant signaling systems have evolved, but also the ways in which these systems can change moving forward, whether by evolutionary processes, disease causing mutations, or engineered components that alter decision making behaviors in new and unprecedented ways.

Outline of Thesis.

Our review of the state of the cell signaling field from biochemical, systems, and evolutionary mechanisms has revealed a need to connect ideas from each discipline to fully understand how cellular decision making is mediated by signaling networks. The work that follows focuses not only on determining the detailed biochemical mechanisms that resolve the complexities and ambiguities of individual cell signaling networks, but also on how these networks can change or grow to accommodate new pathways during evolution. To this end, reconstituted signaling networks are developed to probe the behavior of signaling systems under different environmental conditions or network compositions, as well as to rigorously compare evolutionarily related systems. Two different ambiguous network structures from model systems are explored from this perspective: overlapping signaling pathways that arise from component duplication/divergence, and central signaling nodes that are recycle and repurposed to respond to multiple inputs and activate many possible outputs.

From one pathway to two: generating orthogonal signaling pathways through component duplication

In Chapters 2 and 3, we interrogate the mechanistic and evolutionary bases by which orthogonal signaling pathways can arise from duplication/divergence events by employing a highly quantitative in vitro reconstitution of a model MAP kinase (MAPK) signaling network from budding yeast. This network contains two MAPKs from distinct signaling pathways that arose from a duplication event: the pheromone-responsive MAPK Fus3, and the starvation-responsive MAPK Kss1. Both kinases are activated by the same upstream MAPK kinase (MAPKK), Ste7, which can be activated by either pheromone or starvation. This represents a common conundrum in complex signaling networks: given that both Fus3 and Kss1 are MAPK substrates for Ste7,

how does Ste7 “know” which input it was activated by so that it can go on to activate the appropriate target?

A widely invoked model for pathway insulation is that scaffold proteins can sequester signaling components into separate pools with distinct functions. In budding yeast, the mating-specific scaffold protein Ste5 was treated as the archetype for this mechanism, but when we measured the half-life of Ste5 signaling complexes in vitro, components exchanged on the order of seconds—far too fast for the sequestration paradigm to hold. Instead, I identified a conformational control insulation mechanism that is broadly transferable to other scaffold systems. The Ste5 scaffold contains a domain that is required to allosterically unlock Fus3 for phosphorylation by Ste7, but is dispensable for Ste7 to phosphorylate Kss1. Through biochemical and structural dissection, I showed that access to this essential domain in Ste5 is occluded by an autoinhibitory interaction in unstimulated cells. In response to mating pheromone, recruitment of Ste5 to the plasma membrane relieves this autoinhibition, ensuring Fus3 is only activated under appropriate conditions.

These results show that evolutionarily related signaling molecules with common upstream activators can be effectively insulated from one another via unique allosteric dependencies on other input-regulated macromolecules. However, because effector and target must simultaneously possess their complementary relationship for functional insulation, how do such systems evolve? This question is answered in Chapter 3, in which I performed a biochemistry-driven comparative study of MAPK signaling components spanning a billion years of organismal diversity. Surprisingly, I observed that Ste5 could allosterically activate MAPKs from species that diverged prior to the emergence of Ste5, implying these relationships arose by exploitation of latent regulatory features already present in the MAPK ancestor. Remarkably, the

magnitude of this latent allosteric potential drifts widely among pre-Ste5 MAPKs, showing that seemingly equivalent signaling components can be surprisingly diverse along hidden, allosteric dimensions of phenotypic space.

Getting the most out of what's already present: probing the versatility and fragility of Ras GTPase signaling systems.

In Chapter 4 of this thesis, we explore how a central signaling node that can be activated by many possible inputs and fan out to many possible outputs can be repurposed in different ways to generate a wide range of complex dynamic signaling programs. In particular, the behavior of Ras GTPase systems was explored in a multi-turnover *in vitro* setting for the first time. This revealed that Ras systems could transmit both sustained and transient signals and that the concentration and identity of signaling components strongly impacted the timing, duration, shape and amplitude of the output. Moreover, different effectors interpreted the same inputs with unique dynamics, enabling multiple distinct temporal outputs to be encoded in the system response. Interestingly, the extent to which oncogenic mutations in Ras distorted outputs was found to be highly dependent on this underlying network configuration. This mapping of the space of output behaviors accessible to Ras systems and its associated structure reveals that these systems are readily adapted to produce an array of dynamic behaviors well-suited for diverse cell-signaling functions, but this comes with a trade-off of increased fragility as there exist numerous paths to system states with dramatically altered signaling behaviors that can cause disease and cancer.

Methods.

KinTek simulations.

The time courses of activation shown in Figure 2 were derived using the KinTek Explorer student software package. The node N was modeled based on small GTPases (G), in which GEF serves as an activator by binding to GTPase vacating the nucleotide binding pocket of the GTPase (GT-> G or GD->G). This allows nucleotide (T or D) to bind. Intrinsic GTPase inactivation occurs by hydrolysis of GTP (T) to GDP (D). In the 3-state model, an intermediate GTPase inactivation state (GI) exists, which is not subject to reactivation by the GEF. The following schemes were used for simulation:

Two-state model:

G+T=GT

GT=GD

GD=G+D

GD+GEF=GD_GEF

GD_GEF=G+GEF_D

GEF_D=GEF+D

Three-state model:

G + T = GT

GT = GI

GI = GD

GD = G + D

GD + GEF = GD_GEF

GD_GEF = G + GEF_D

GEF_D = GEF + D

The following parameter sets were used for simulation:

Two-state model

	k_+	k_-
G+T=GT	10	0
GT=GD	0.1	0
GD=G+D	0.01	0
GD+GEF=GD_GEF	10	100
GD_GEF=G+GEF_D	10	0
GEF_D=GEF+D	100000	0

Three-state model

	k_+	k_-
G+T=GT	10	0
GT=GI	0.1	0

GI=GD	0.01	0
GD=G+D	0.01	0
GD+GEF=GD_GEF	10	100
GD_GEF=G+GEF_D	10	0
GEF_D=GEF+D	100000	0

Data were simulated for 100 time units and exported to a CSV format for further analysis and preparation of the graphs in Figure 2.

References.

1. G. Krauss, *Biochemistry of Signal Transduction and Regulation* (John Wiley & Sons, 2006).

2. W. Lim, B. Mayer, T. Pawson, *Cell Signaling: principles and mechanisms* (Taylor & Francis, 2014).

3. B. Alberts *et al.*, *Molecular Biology of the Cell* (Garland Science, ed. 4th, 2002).

4. A. Boureux, E. Vignal, S. Faure, P. Fort, Evolution of the Rho Family of Ras-Like GTPases in Eukaryotes. *Mol. Biol. Evol.* **24**, 203–216 (2007).

5. D. R. Caffrey, L. A. J. O'Neill, D. C. Shields, The Evolution of the MAP Kinase Pathways: Coduplication of Interacting Proteins Leads to New Signaling Cascades. *J. Mol. Evol.* **49**, 567–582 (1999).

6. T. Matozaki, H. Nakanishi, Y. Takai, Small G-protein networks:: Their crosstalk and signal cascades. *Cell. Signal.* **12**, 515–524 (2000).

7. T. J. P. van Dam, J. Bos, B. Snel, Evolution of the Ras-like small GTPases and their regulators. *Small GTPases*. **2**, 4–16 (2011).

8. C. Chothia, J. Gough, C. Vogel, S. A. Teichmann, Evolution of the Protein Repertoire. *Science*. **300**, 1701–1703 (2003).

9. M. Li, J. Liu, C. Zhang, Evolutionary History of the Vertebrate Mitogen Activated Protein Kinases Family. *PLoS ONE*. **6**, e26999 (2011).

10. M. C. Good, J. G. Zalatan, W. A. Lim, Scaffold Proteins: Hubs for Controlling the Flow of Cellular Information. *Science*. **332**, 680–686 (2011).

11. H. D. Madhani, G. R. Fink, The riddle of MAP kinase signaling specificity. *Trends Genet.* **14**, 151–155 (1998).

12. T. Hunter, Protein kinases and phosphatases: The Yin and Yang of protein phosphorylation and signaling. *Cell*. **80**, 225–236 (1995).

13. J. M. Aletta, T. R. Cimato, M. J. Ettinger, Protein methylation: a signal event in post-translational modification. *Trends Biochem. Sci.* **23**, 89–91 (1998).

14. K. Haglund, I. Dikic, Ubiquitylation and cell signaling. *EMBO J.* **24**, 3353–3359 (2005).

15. J. D. Scott, T. Pawson, Cell Signaling in Space and Time: Where Proteins Come Together and When They're Apart. *Science*. **326**, 1220–1224 (2009).

16. J. N. Onuchic, Z. Luthey-Schulten, P. G. Wolynes, THEORY OF PROTEIN FOLDING: The Energy Landscape Perspective. *Annu. Rev. Phys. Chem.* **48**, 545–600 (1997).

17. H. Frauenfelder, S. G. Sligar, P. G. Wolynes, The energy landscapes and motions of proteins. *Science*. **254**, 1598–1603 (1991).

18. N. P. Pavletich, Mechanisms of cyclin-dependent kinase regulation: structures of cdks, their cyclin activators, and cip and INK4 inhibitors1,2. *J. Mol. Biol.* **287**, 821–828 (1999).

19. M. Good, G. Tang, J. Singleton, A. Reményi, W. A. Lim, The Ste5 scaffold directs mating signaling by catalytically unlocking the Fus3 MAP kinase for activation. *Cell*. **136**, 1085–1097 (2009).

20. D. F. Brennan *et al.*, A Raf-induced allosteric transition of KSR stimulates phosphorylation of MEK. *Nature*. **472**, 366–369 (2011).

21. S. R. Hubbard, Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *EMBO J.* **16**, 5572–5581 (1997).

22. S. R. Hubbard, L. Wei, W. A. Hendrickson, Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature*. **372**, 746–754 (1994).

23. N. Popovych, S. Sun, R. H. Ebright, C. G. Kalodimos, Dynamically driven protein allostery. *Nat. Struct. Mol. Biol.* **13**, 831–838 (2006).

24. J. F. Swain, L. M. Gierasch, The changing landscape of protein allostery. *Curr. Opin. Struct. Biol.* **16**, 102–108 (2006).

25. C. G. Kalodimos, Protein function and allostery: a dynamic relationship. *Ann. N. Y. Acad. Sci.* **1260**, 81–86 (2012).

26. C.-J. Tsai, A. del Sol, R. Nussinov, Protein allostery, signal transmission and dynamics: a classification scheme of allosteric mechanisms. *Mol. Biosyst.* **5**, 207 (2009).

27. P. M. Pryciak, F. A. Huntress, Membrane recruitment of the kinase cascade scaffold protein Ste5 by the Gβγ complex underlies activation of the yeast pheromone response pathway. *Genes Dev.* **12**, 2684–2697 (1998).

28. E. J. Lowenstein *et al.*, The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling. *Cell*. **70**, 431–442 (1992).

29. H. Wu, Higher-Order Assemblies in a New Paradigm of Signal Transduction. *Cell*. **153**, 287–292 (2013).

30. L. Chang, M. Karin, Mammalian MAP kinase signalling cascades. *Nature*. **410**, 37–40 (2001).

31. J. Schlessinger, Cell Signaling by Receptor Tyrosine Kinases. *Cell*. **103**, 211–225 (2000).

32. S. Bogdan, C. Klämbt, Epidermal growth factor receptor signaling. *Curr. Biol.* **11**, R292–R295 (2001).

33. C. Jia, M. Qian, D. Jiang, Overshoot in biological systems modelled by Markov chains: a non-equilibrium dynamic phenomenon. *IET Syst. Biol.* **8**, 138–145 (2014).

34. A. Fersht, *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (Macmillan, 1999).

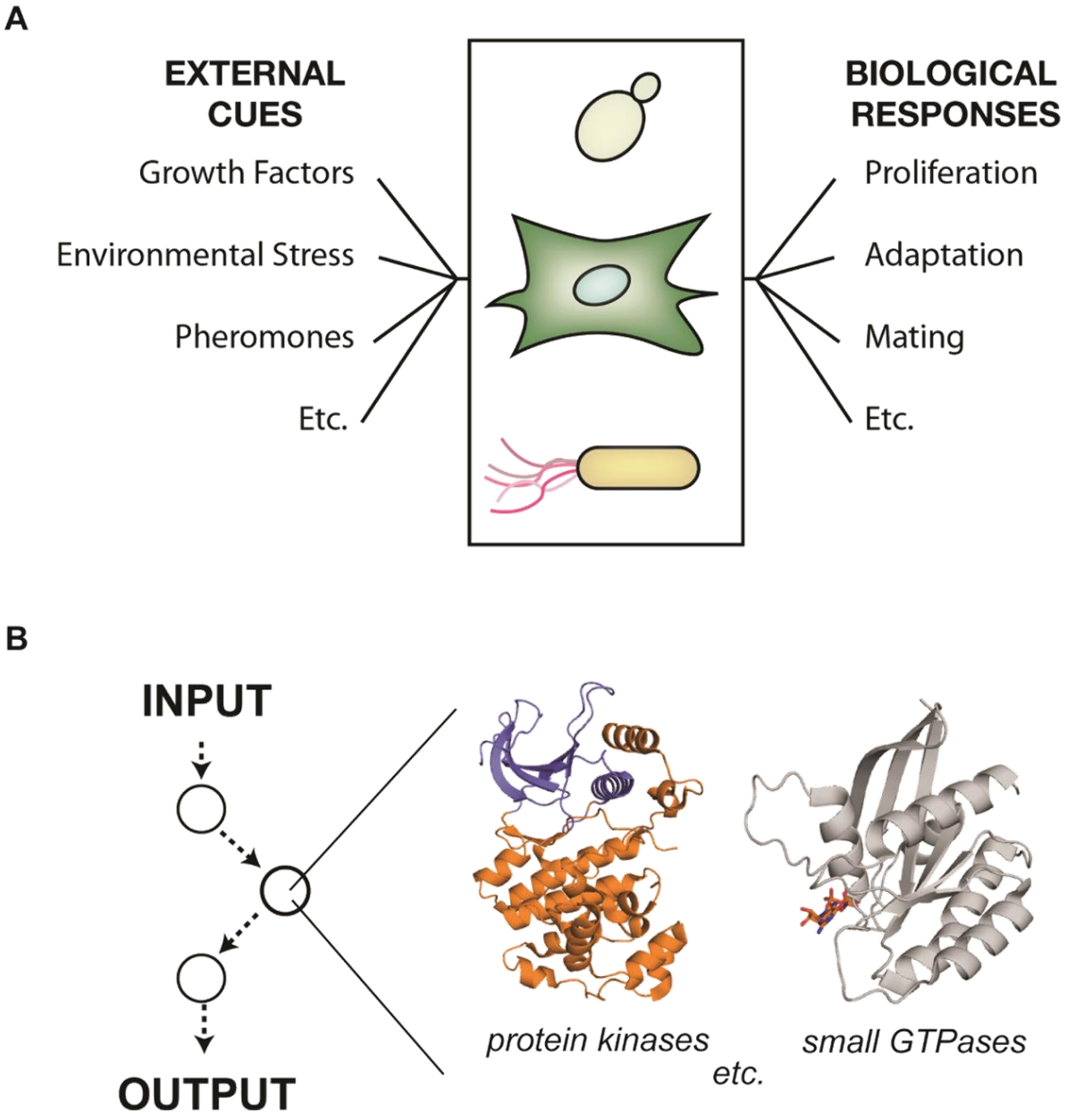
35. J. E. Toettcher, O. D. Weiner, W. A. Lim, Using Optogenetics to Interrogate the Dynamic Control of Signal Transmission by the Ras/Erk Module. *Cell*. **155**, 1422–1434 (2013).

36. K. Ono, J. Han, The p38 signal transduction pathway Activation and function. *Cell. Signal.* **12**, 1–13 (2000).

37. S. J. Gould, Exaptation: A Crucial Tool for an Evolutionary Psychology. *J. Soc. Issues*. **47**, 43–65 (1991).

38. S. Ohno, *Evolution by Gene Duplication* (Springer Science & Business Media, 2013).

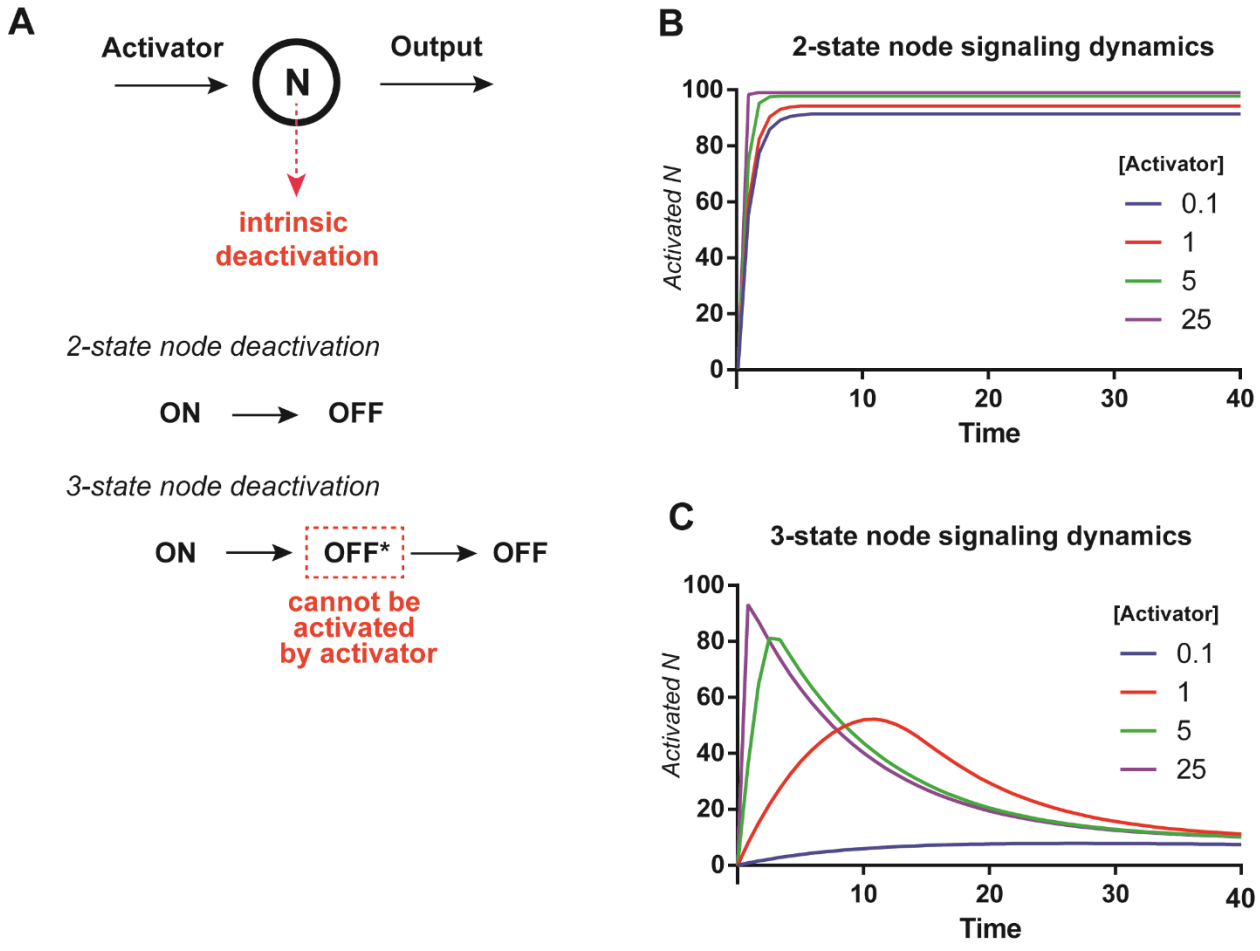
Figure 1. Cells sense external information and process that into biological responses using signal transduction networks.



(a) Illustration of the fact that all living cells must sense cues from their environment, process that information, and execute specific biological responses appropriate to that cue. (b) Diagram

showing both the abstract “nodes and arrows” representation of signal transduction pathways as well as the molecular content of these nodes. Individual signaling nodes correspond to molecules like protein kinases or small GTPases that can be effectively regulated and are well-suited to transmit information to other molecules.

Figure 2. The detailed mechanism by which a node is activated or deactivated impacts signaling behavior.



(a) Simple 1-node signaling system with an input that results in activation of N by an activator and intrinsic deactivation by N. Two different deactivation schemes for the N node are shown. In a 2-state scheme, the ON state of N proceeds directly to an OFF state that can be activated by activator. In a 3-state scheme, the ON state of N proceeds to an intermediate OFF* state which is not capable of producing output but cannot be activated by activator either, Activation of N requires that the OFF* state decay further to an OFF state, at which point activator can activate

N. (b) KinTek simulation output showing N's activation dynamics with a 2-state deactivation scheme (detailed in methods). Increasing amounts of activator lead to increasing final steady state levels of N, and the approach to that steady state always proceeds monotonically. (c) KinTek simulation output showing N's activation dynamics with a 3-state deactivation scheme. Unlike in the 2-state scheme, differing amounts of activator lead to very different dynamic behaviors. At the lowest concentration, activated N accumulates monotonically, whereas higher concentrations of activator lead to adaptive overshoot behavior with different peak and decay times.

Chapter 2. Conformational Control of the Ste5 Scaffold Protein Insulates Against MAP Kinase Misactivation.

Abstract.

Cells re-use signaling proteins in multiple pathways, raising the potential for improper crosstalk. Scaffold proteins are thought to insulate against such miscommunication by sequestering proteins into distinct physical complexes. We show that the scaffold protein Ste5, which organizes the yeast mating MAP kinase pathway, does not use sequestration to prevent misactivation of the mating response. Instead, Ste5 appears to use a conformation mechanism: under basal conditions, intramolecular interaction of the PH domain with the VWA domain blocks its the ability to co-activate the mating-specific MAPK, Fus3. Pheromone-induced membrane binding of Ste5 triggers release of this autoinhibition. Thus, in addition to serving as a conduit guiding kinase communication, Ste5 directly receives input information to decide if and when signal can be transmitted to mating output.

Main Text.

Cells use a complex network of signaling proteins to respond to diverse signals and stresses. Execution of proper decisions is complicated by the fact that individual cells contain many closely related signaling proteins (*1*). In fact, the same proteins are often reused in multiple signaling pathways (*2, 3*). The resulting interlinked networks could lead to inappropriate crosstalk between signaling pathways.

Scaffold proteins, which physically assemble components of a signaling pathway (*4-6*), provide a possible solution to this problem. By binding and organizing pathway components into complexes, scaffold proteins promote efficient signaling along a particular pathway. Scaffold proteins may also insulate against improper communication by physically sequestering signaling proteins into distinct pools (*7-15*). However, to prevent shared proteins from exchanging

between pools, a scaffold must bind its partners with dissociation rates that are slow compared to the timescale for signaling. Direct evidence for this prevailing view of scaffold-based insulation is limited.

A prototypical scaffold protein is Ste5, which coordinates the yeast mating mitogen-activated protein kinase (MAPK) response by binding to all three components of the MAPK cascade and serving as a required co-activator of the mating-specific MAPK, Fus3 (*16, 17*). The Ste5 scaffold is thought to insulate the mating response from other MAPK pathways in yeast, such as the starvation response, which uses the identical MAPK kinase (MAPKK), Ste7, and MAPKK kinase (MAPKKK), Ste11, proteins, but activates a distinct starvation-specific MAPK Kss1 to produce an invasive growth response (Fig. 1A) (*2, 17*). How the common MAPKK, Ste7, when activated by a specific input, is directed to the correct downstream MAPK is only partially understood. With mating input, both Fus3 and Kss1 are activated (binding to the Ste5 scaffold does not prevent the MAPKK Ste7 from activating Kss1) (*16, 18*). However, activation of Kss1 by mating input does not lead to crosstalk because activated Fus3 overrides the Kss1-induced starvation response by phosphorylating and downregulating a starvation-specific transcription factor (*19, 20*). Thus, proper starvation response hinges upon preventing Fus3 misactivation by starvation inputs, which would both launch the mating program and directly inhibit the starvation response.

For Ste5 to act as a sequestration-based insulator would require exchange rates for the scaffold-bound shared kinases (Ste11 or Ste7) to be slow relative to the timescale of signaling. Otherwise, shared kinases activated by non-mating inputs would be able to exchange onto the Ste5 scaffold protein and activate the mating response. We measured the dissociation of purified Ste7 from Ste5 to have a $t_{1/2}$ of <5 seconds (Fig. 1B), faster by several orders of magnitude than

the typical ~5 minute timescale of MAPK signaling pathways, and far faster than the timescale of days on which the yeast starvation response operates. Thus, physical sequestration is unlikely to be the primary mechanism that prevents activation of the mating MAPK Fus3 by non-mating inputs.

An alternative model for insulation is that, in the absence of mating input, Ste5 adopts an inactive conformation that blocks its ability to co-catalyze Fus3 phosphorylation (*17*). To test this possibility, we measured rates of Fus3 phosphorylation by the MAPKK Ste7 with full-length Ste5 and with a minimal Ste5 fragment ($Ste5_{VWA-C}$, containing a von-Willebrand Type-A (VWA) domain that is required for Fus3 co-activation together with active MAPKK Ste7 (*16*)). Under maximal rate (k_{cat}) conditions (saturating concentrations of all components), the rate of Fus3 phosphorylation with full-length Ste5 was one-tenth that in the presence of $Ste5_{VWA-C}$ (Fig. 2A). Assembly of the Ste5-Ste7-Fus3 complex was similar with either Ste5 construct (Fig. 2A), suggesting that the dominant contribution to the difference in activity is not a binding effect (disruption of kinase complex assembly) but rather a disruption of the catalytic co-activator function of the VWA domain.

To test the possibility that this activity difference contributes to insulation of the mating pathway *in vivo*, we introduced the fully active $Ste5_{VWA-C}$ fragment into yeast. In these cells, starvation led to substantial activation of the mating MAPK Fus3, whereas cells with full-length Ste5 predominantly activated the starvation MAPK Kss1 (Fig. 2B). Thus the minimal $Ste5_{VWA-C}$ fragment appears to promote Fus3 activation when the MAPKK Ste7 is activated, regardless of whether cells have received the mating signal or not. Because Fus3 activation inhibits the invasive growth response at the transcriptional level (*19, 20*), misactivation of Fus3 by cells expressing $Ste5_{VWA-C}$ overrides the invasive growth phenotype (Fig. 2C). Further, under

starvation conditions, $Ste5_{VWA-C}$ restores a partial mating phenotype in cells that lack the mating receptor, Ste2 (Fig. 2D); full rescue of mating likely requires mating pathway components upstream of the kinase cascade that are not activated by starvation (*21*). Thus when Fus3 activation is promiscuous, cells misinterpret the starvation condition as a signal to initiate the mating program.

By deletion analysis, we identified two regions of Ste5 essential for autoinhibition of the VWA domain *in vitro*: the PH domain, which binds to phosphoinositol 4,5-bisphosphate (PIP_2) to facilitate membrane binding (*22*) and binds to the MAPKKK Ste11 (*23*), and an N-terminal extension to the VWA domain (residues 544-592), within the linker that connects the PH and VWA domains (Fig. 3A and fig. S9). When Ste5 was replaced by Ste5Δ(544-592) *in vivo*, activation of Fus3 in response to mating pheromone was normal (Fig. S7), but Fus3 was misactivated in response to starvation (Fig. 3B). Because all kinase binding sites are intact in Ste5Δ(544-592), this result supports the idea that physical sequestration of kinases by Ste5 is not sufficient for pathway insulation.

We determined the crystal structure of an extended VWA fragment (residues 582-786; we were unable to obtain crystals for a PH-VWA complex) (Fig. 3C); this construct includes the minimal N-terminal extension that binds the PH domain (Fig. S10). This extension forms an N-terminal α-helix lying directly adjacent to the VWA domain "coactivator loop" that contains residues essential for Fus3 coactivation (*16*). The spatial proximity of the autoinhibitory PH domain binding site (the N-terminal extension) and the Fus3 coactivator loop indicates that PH domain binding and Fus3 activation might be mutually exclusive, providing a molecular mechanism for Ste5 autoinhibition (Fig. 3D). Indeed, the isolated PH domain of Ste5 inhibited the Fus3 co-activator function of the VWA domain *in trans* (Fig. 3E). Further, a Fab antibody

fragment that binds the PH domain competitively relieved autoinhibition (Fig. 3F and fig. S11). Also, an allele of Ste5 (S770N) that was previously found to constitutively activate the mating pathway (*24*) is not autoinhibited *in vitro* (Fig. S12).

An early step in mating pathway activation is pheromone-induced membrane recruitment of Ste5, which requires a cooperative set of membrane interactions that includes the PH domain binding to PIP_2 lipids (*22*). Thus binding of Ste5 to PIP_2-containing membranes might disrupt the PH-VWA interaction and relieve autoinhibition. We designed a minimal, membrane-binding Ste5 construct that is autoinhibited, but PIP_2-containing lipid vesicles did not bind or activate this construct *in vitro* (Fig. S13). Because pheromone-induced membrane recruitment of Ste5 is a cooperative process that requires several membrane-binding motifs (*21*), we induced association of the autoinhibited Ste5 construct to the lipid vesicles using other cooperative membrane interactions (Fig. 3G and Fig. S13). Under these conditions, PIP_2 caused a 3-fold activation of Ste5 (Fig. 3G), suggesting that membrane recruitment of Ste5 and its interaction with PIP_2 contributes to relief of autoinhibition of Ste5. The inability of such membrane association to completely relieve autoinhibition of Ste5 (a 10-fold effect, Fig. 2A) could result from incomplete binding to lipid vesicles *in vitro* (Fig. S13), or because complete activation requires additional interactions present *in vivo*. Ste5 oligomerization has been suggested to contribute to pathway activation (*25*), but we find no evidence that oligomerization plays a direct role in relief of Ste5 autoinhibition (Fig. S14).

We propose that although the shared upstream kinases (MAPKKK Ste11 and MAPKK Ste7) can be activated by other inputs, only mating input activates both the kinase cascade and the Ste5 scaffold protein to permit Fus3 activation (Fig. 4A). Activation of the mating pathway recruits Ste5 to the membrane (*21*), thus activating the MAP kinase cascade by bringing the

MAPKKK Ste11 in proximity to its upstream kinase Ste20. Membrane recruitment may also relieve autoinhibition in Ste5 when the PH domain interacts with PIP_2 at the membrane (Fig. 4B).

To further test this model, we decoupled the two functions of Ste5 by deleting the upstream kinase Ste20 (preventing normal activation of the MAPK cascade), and introducing a constitutively-active allele of the MAPKKK Ste11, thus rendering activation of the kinase cascade independent of the mating signal. Previous experiments of this type demonstrated that full pathway activation still requires the mating input, suggesting that the input acts on a step downstream of kinase cascade activation (*24*, *26*). Here we take this approach one step further by using a constitutively active allele, Ste11ΔN (*27*), which lacks the Ste5 binding site (*28*), so that any observed effects of Ste5 activation are likely to arise from promoting the Ste7→Fus3 reaction rather than the Ste11→Ste7 reaction. When wild type Ste11 was replaced by Ste11ΔN in a yeast strain lacking Ste20, the MAPK Kss1 was preferentially phosphorylated, but when this strain was treated with α-factor, activation of Fus3 was observed (Fig. 4), supporting the idea that pheromone-induced membrane recruitment of Ste5 has two distinct and separable functions: to activate the MAPKKK Ste11 and to relieve autoinhibition in Ste5 to permit Fus3 activation.

Our data do not support the prevailing model that scaffold proteins primarily insulate signaling by sequestration of proteins. Instead, Ste5 appears to function as a conformational switch to gate the flow of information between two distinct signaling outcomes. This mechanism provides a potentially general means to control information flow in complex signaling networks with shared components.

Main Text References.

1. S. S. Taylor *et al.*, A template for the protein kinase family. *Trends Biochem. Sci.* **18**, 84 (1993).
2. M. A. Schwartz, H. D. Madhani, Principles of MAP kinase signaling specificity in *Saccharomyces cerevisiae*. *Annu. Rev. Genet.* **38**, 725 (2004).
3. M. S. Qi, E. A. Elion, MAP kinase pathways. *J. Cell Sci.* **118**, 3569 (2005).
4. M. C. Good, J. G. Zalatan, W. A. Lim, Scaffold proteins: Hubs for controlling the flow of cellular information. *Science* **332**, 680 (2011).
5. D. K. Morrison, R. J. Davis, Regulation of MAP kinase signaling modules by scaffold proteins in mammals. *Annu. Rev. Cell Dev. Biol.* **19**, 91 (2003).
6. A. S. Shaw, E. L. Filbert, Scaffold proteins and immune-cell signalling. *Nat. Rev. Immunol.* **9**, 47 (2009).
7. H. Saito, Regulation of cross-talk in yeast MAPK signaling pathways. *Curr Opin Microbiol* **13**, 677 (2010).
8. D. N. Dhanasekaran, K. Kashef, C. M. Lee, H. Xu, E. P. Reddy, Scaffold proteins of MAP-kinase modules. *Oncogene* **26**, 3185 (2007).
9. K. Harris *et al.*, Role of scaffolds in MAP kinase pathway specificity revealed by custom design of pathway-dedicated signaling proteins. *Curr. Biol.* **11**, 1815 (2001).
10. W. R. Burack, A. S. Shaw, Signal transduction: hanging on a scaffold. *Curr. Opin. Cell Biol.* **12**, 211 (2000).
11. T. P. Garrington, G. L. Johnson, Organization and regulation of mitogen-activated protein kinase signaling pathways. *Curr. Opin. Cell Biol.* **11**, 211 (1999).

12. A. J. Whitmarsh, R. J. Davis, Structural organization of MAP-kinase signaling modules by scaffold proteins in yeast and mammals. *Trends Biochem. Sci.* **23**, 481 (1998).

13. T. Pawson, J. D. Scott, Signaling through scaffold, anchoring, and adaptor proteins. *Science* **278**, 2075 (1997).

14. S. Marcus, A. Polverino, M. Barr, M. Wigler, Complexes between Ste5 and Components of the Pheromone-Responsive Mitogen-Activated Protein-Kinase Module. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7762 (1994).

15. K. Y. Choi, B. Satterberg, D. M. Lyons, E. A. Elion, Ste5 tethers multiple protein kinases in the MAP kinase cascade required for mating in S. cerevisiae. *Cell* **78**, 499 (1994).

16. M. Good, G. Tang, J. Singleton, A. Remenyi, W. A. Lim, The Ste5 scaffold directs mating signaling by catalytically unlocking the Fus3 MAP kinase for activation. *Cell* **136**, 1085 (2009).

17. L. J. Flatauer, S. F. Zadeh, L. Bardwell, Mitogen-activated protein kinases with distinct requirements for Ste5 scaffolding influence signaling specificity in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **25**, 1793 (2005).

18. E. A. Elion, J. A. Brill, G. R. Fink, Fus3 represses Cln1 and Cln2 and in concert with Kss1 promotes signal transduction. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 9392 (1991).

19. M. Z. Bao, M. A. Schwartz, G. T. Cantin, J. R. Yates, H. D. Madhani, Pheromone-dependent destruction of the Tec1 transcription factor is required for MAP kinase signaling specificity in yeast. *Cell* **119**, 991 (2004).

20. S. Chou, L. Huang, H. P. Liu, Fus3-regulated Tec1 degradation through SCF Cdc4 determines MAPK signaling specificity during mating in yeast. *Cell* **119**, 981 (2004).

21. L. Bardwell, A walk-through of the yeast mating pheromone response pathway. *Peptides* **26**, 339 (2005).

22. L. S. Garrenton, S. L. Young, J. Thorner, Function of the MAPK scaffold protein, Ste5, requires a cryptic PH domain. *Gene Dev* **20**, 1946 (2006).

23. C. Inouye, N. Dhillon, T. Durfee, P. C. Zambryski, J. Thorner, Mutational analysis of STE5 in the yeast *Saccharomyces cerevisiae*: Application of a differential interaction trap assay for examining protein-protein interactions. *Genetics* **147**, 479 (1997).

24. R. E. Lamson, S. Takahashi, M. J. Winters, P. M. Pryciak, Dual role for membrane localization in yeast MAP kinase cascade activation and its contribution to signaling fidelity. *Curr. Biol.* **16**, 618 (2006).

25. C. Inouye, N. Dhillon, J. Thorner, Ste5 RING-H2 domain: Role in Ste4-promoted oligomerization for yeast pheromone signaling. *Science* **278**, 103 (1997).

26. D. M. Lyons, S. K. Mahanty, K. Y. Choi, M. Manandhar, E. A. Elion, The SH3-domain protein Bem1 coordinates mitogen-activated protein kinase cascade activation with cell cycle control in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **16**, 4095 (1996).

27. B. R. Cairns, S. W. Ramer, R. D. Kornberg, Order of action of components in the yeast pheromone response pathway revealed with a dominant allele of the STE11 kinase and the multiple phosphorylation of the STE7 kinase. *Genes Dev.* **6**, 1305 (1992).

28. G. Jansen, F. Buhring, C. P. Hollenberg, M. Ramezani Rad, Mutations in the SAM domain of *STE50* differentially influence the MAPK-mediated pathways for mating, filamentous growth and osmotolerance in *Saccharomyces cerevisiae*. *Mol. Genet. Genomics* **265**, 102 (2001).

Supporting Text.

Materials and Methods

Yeast Strain Construction and Manipulation

The parent yeast strain for all experiments was F1950 (Σ1278b; *MAT***a** *ura3 leu2 trp1 his3*) (1). Strains derived from the Σ1278b lineage exhibit a haploid invasive growth phenotype (2-4). Yeast knockout strains were prepared using PCR-mediated gene deletion (5, 6), and transformations were performed with the lithium acetate method (7).

Expression constructs were integrated in single copies into the yeast genome using a set of custom-designed vectors (gifts from N. Helman and S. Vidal). These vectors consist of a multiple cloning site to introduce a promoter and ORF, the *Candida albicans* Adh1 terminator, auxotrophic complementation markers from *Candida glabrata* or *Candida albicans*, and 5' and 3' flanking homology regions (~500 bp) to target the corresponding auxotrophic locus (603, His3; 604, Trp1; Leu2; 606, Ura3). Prior to transformation, the vector is linearized by restriction digestion at sites just outside of the 5' and 3' flanking homology regions. Using this approach, only a single copy of the target gene is introduced into the yeast genome, unlike the pRS300 series of vectors (8), which can integrate multiple copies. Yeast strains and plasmids used in this work are listed in Tables S2 and S3. All constructs were expressed from their native promoters except the following: Ste11ΔN was overexpressed from the Adh1 promoter and Ste7EE was overexpressed from the Gal10 promoter (in experiments with Ste7EE, Ste5 and $Ste5_{VWA-C}$ were overexpressed from Adh1 promoters, fig. S8).

For analysis of MAPK phosphorylation in yeast subjected to starvation, yeast cells were grown in YPD (yeast peptone dextrose) liquid cultures to mid-log phase (OD_{600} 0.4-0.6). 250 μL of cells were plated on YPD and incubated at 30 °C for 18 hours (the yeast extract in YPD

stimulates the invasive growth response (4)). Similar results were obtained when cells growing on YPD plates were streaked onto fresh YPD plates and incubated at 30 °C for 18 hours. Unstarved cells were obtained by streaking cells growing on SCD (synthetic complete dextrose) plates onto fresh SCD plates and incubating at 30 °C for 18 hours. Approximately 10 mg of cells were harvested directly from the plates. Yeast lysates were prepared using an alkaline lysis extraction (9). Samples were run on Novex 10% Tris-glycine gels (Life Technologies) to separate Fus3 and Kss1, and transferred to nitrocellulose membranes (Whatman Protran BA83). Western blots were carried out using a primary anti-phospho p44/p42 MAPK antibody (Cell Signaling Technology #4370) and a secondary IRDye 800CW Goat Anti-Rabbit IgG antibody (Li-Cor #926-32211) with Li-Cor Odyssey Blocking Buffer. Phosphoglycerate kinase (PGK) was used as a loading control with a primary anti-PGK antibody (Life Technologies #459250) and a secondary IRDye 680LT Goat Anti-Mouse IgG antibody (Li-Cor #926-68020). Blots were visualized using the Li-Cor Odyssey Imaging System.

For α-factor induction experiments, yeast cells were grown in YPD liquid cultures to mid-log phase (OD_{600} 0.4-0.6). 10 mL liquid cultures were treated with 2 μM α-factor (final concentration), incubated at 30 °C for 15 minutes, centrifuged at 1600g for 2 minutes at 4 °C, and frozen in liquid nitrogen. Yeast lysates were prepared and western blots were performed as described above for starvation assays.

The yeast invasive growth phenotype (2-4) was assayed by growing yeast strains in liquid cultures (YPD or SCD) to OD_{600} 0.4-0.6, then spotting 20 μL onto YPD (+ starvation) or SCD (- starvation) plates and incubating for 1-2 days at 30 °C. Yeast extract in the YPD medium stimulates the invasive growth response (4). Invasive growth was assayed by washing under a stream of tap water.

For galactose inductions, yeast cells were grown in SCS (synthetic complete sucrose) media to OD_{600} 0.2, then induced with 2% galactose (final concentration), incubated at 30 °C for 4 hours, centrifuged at 1600g for 2 minutes at 4 °C, and frozen in liquid nitrogen. Yeast lysates were prepared and western blots were performed as described above for starvation assays.

Quantitative and qualitative (patch) yeast mating assays were performed in YPD media as described (10) using a *MATα his1* mating tester (1). Quantitative mating efficiency was calculated from the titer of **a**/α cells divided by the titer of **a** cells.

Protein Expression and Purification

Constructs to express recombinant proteins were from previous work (11) or were generated by standard molecular biology methods. We used a catalytically-dead allele of Fus3 (K42R) in order to eliminate background autophosphorylation that is observed with the wild-type protein. For kinetic characterization, Ste5 constructs were expressed as maltose binding protein (MBP) fusions.

Fus3 and Ste5 constructs were expressed in Rosetta (DE3) pLysS *E. coli* cells by inducing with 0.5 mM IPTG overnight at 18 °C. Fus3 K42R was expressed using the pBH4 vector, a derivative of pET15 (Novagen) that was modified to produce an N-terminal His_6-tagged, TEV-cleavable protein. Full length Ste5 and truncation mapping constructs were expressed with an N-terminal TEV-cleavable MBP and a C-terminal His-tag. For Ste5Δ(544-592) and all other internal truncations (Fig. 3A and fig. S9), the deleted residues were replaced with a 5X(GAGS) linker. Ste5-ms (the minimal VWA domain, residues 593-786), the Ste5 PH domain fragment (Fig. 3E), and the $Ste5_{582-786}$ fragment used for crystallography were expressed from the pBH4 vector.

Constructs expressed from pBH4 were affinity purified on Ni-NTA (Qiagen), cleaved with TEV protease to release the N-terminal His_6 tag, and further purified by ion exchange (RESOURCE 15Q [Amersham], 20 mM Tris·HCl pH 8.0, 2 mM DTT, 10% glycerol, and 0-1 M NaCl). Constructs expressed from pMBP were affinity purified on Ni-NTA (Qiagen), followed by amylose resin (New England Biolabs). Purified proteins were dialyzed into 20 mM Tris Tris·HCl pH 8.0, 150 mM NaCl, 10% glycerol, and 2 mM DTT at 4 °C, aliqouted and stored at -80 °C.

Ste7 (constitutively active Ste7EE, bearing S359E and T363E phosphomimic mutations in the activation loop (11, 12)) and Ste11 (wild type) were expressed with an N-terminal TEV-cleavable MBP and C-terminal His-tag in *Spodoptera frugiperda* (SF9) cells, using the Bac-to-Bac Baculoviral Expression System (Life Technologies) at 27 °C. Proteins were affinity purified on Ni-NTA (Qiagen), followed by amylose resin (New England Biolabs). Purified proteins were dialyzed into 20 mM Tris Tris·HCl pH 8.0, 300 mM NaCl, 10% glycerol, and 2 mM DTT at 4 °C, aliqouted and stored at -80 °C. Prior work expressed Ste7 constructs with an N-terminal GST tag (11). In this work, we switched to an MBP tag to obtain higher purity protein and to eliminate any potential artifacts from GST dimerization (13, 14). MBP-Ste7EE and GST-Ste7EE were found to have a similar k_{cat} for Fus3 phosphorylation (in reactions with Ste5), and the K_{act} for Ste5 binding was ~5-fold weaker for MBP-Ste7EE than for GST-Ste7EE. In the previous work, a variant of Ste7 with a single docking site for Fus3 was used (mutant Ste7EE-ND2 has the second, weaker docking site removed (15)). There was no difference in the behavior of Ste7EE and Ste7EE-ND2 for the kinetic parameters reported in this work.

Selection of a Ste5 PH Domain-Specific Fab

A minimal PH domain construct of Ste5 ($Ste5_{367-525}$) was expressed as a GST fusion protein with a C-terminal His-tag and affinity purified on Ni-NTA (Qiagen), followed by glutathione agarose (Sigma). As a negative control for selection, free GST with the C-terminal His tag (GST_{His}) was also purified as above.

We used a phage-displayed scFv library with a diversity of greater than 10^{11} unique clones and randomized at all 6 complementarity determining regions (CDRs). The library was built from an scFv scaffold that binds *E. coli* maltose binding protein, using methods previously described (16). Selection of domain-specific scFv sequences was performed as described (17). Briefly, 96-well Maxisorp Immunoplates (Fisher Scientific) were coated with 5 μg/mL GST-tagged $Ste5_{367-525}$ or GST_{His} in phosphate buffered saline (PBS). At each round of selection, phage that bound non-specifically were cleared by incubation on GST_{His} and unbound phage was transferred to the antigen-coated wells and allowed to incubate for 2 hours at room temperature. Wells were washed 10 times with PBS + 0.05% Tween-20. Bound phage was eluted with 0.1 N HCl and neutralized with 1 M Tris·HCl (pH 8.0) and used to infect actively growing *E. coli* XL1-Blue cells (Agilent Technologies) for overnight propagation. After five rounds of selection, phage particles were produced from individual clones and used in phage ELISAs to detect specifically binding clones. Positive clones were sequenced and the CDRs for unique clones were grafted onto an IPTG-inducible Fab scaffold using previously described methods (18).

The SR13 Fab protein was expressed in *E. coli* 55244 cells in 2YT with 100 μg/mL carbenicillin. Cells were grown to OD_{600} 0.8 before inducing with 1 mM IPTG for 16 h at 30 °C. Cells were resuspended in lysis buffer (1 mg/mL lysozyme, 37.5 U/mL benzonase, 2 mM $MgCl_2$, 0.2 mM PMSF) and incubated on ice for 1 h. The crude lysate was spun down and the supernatant was applied to an rProtein A affinity column (GE Healthcare), washed with 25

column volumes of PBS, eluted with elution buffer (50 mM NaH_2PO_4, 100 mM H_3PO_4, 140 mM NaCl, pH 2.0), and neutralized with neutralization buffer (1 M Na_2HPO_4, 140 mM NaCl, pH 8.6). Purified protein was concentrated and buffer exchanged to PBS on Amicon Ultracel 10K filter units (Millipore).

In Vitro Kinetic Assays

In vitro kinetic assays were conducted in 25 mM Tris·HCl (pH 8.0), 100 mM NaCl, 2.5 mM $MgCl_2$, 0.05% IGEPAL, and 2 mM TCEP at 25 °C. Fus3 phosphorylation reactions were initiated by addition of ATP to a final concentration of 500 μM.

To determine the rate of Ste7 dissociation from Ste5 (k_{off}), 100 nM Ste5, 100 nM Ste11, 100 nM Ste7EE, and 500 nM Fus3 K42R were pre-incubated in reaction buffer on ice. Dissociation reactions were initiated by introducing a large excess (20 μM) of a Ste7 binding domain (the "sink", a minimal Ste7 binding domain from Ste5 [residues 759-810]), which captures Ste7 as it dissociates from Ste5. After varying times, ATP was added, and reaction timepoints were collected at 15 second intervals by quenching in 4X LDS sample loading buffer (Life Technologies) supplemented with 50 mM EDTA and 2 mM DTT and placing the samples in a boiling heat block (>100 °C) for 90 seconds. Samples were run on NuPAGE 4-12% Bis-Tris gels (Life Technologies) and transferred to nitrocellulose membranes (Whatman Protran BA83).

Quantitative western blots were carried out using a primary anti-phospho p44/p42 MAPK antibody (Cell Signaling Technology #4370) and a secondary IRDye 800CW Goat Anti-Rabbit IgG antibody (Li-Cor #926-32211) using 5% milk as a blocking solution. Blots were visualized using the Li-Cor Odyssey Imaging System and quantified using Odyssey 2.1 software as described previously (11).

Kinetic parameters and their standard errors were determined by nonlinear least-squares fitting to initial rate data. The initial rate of Fus3 phosphorylation at each timepoint after addition of the Ste7 binding domain sink corresponds to the amount of Ste5-Ste7-Fus3 complex remaining at each timepoint. k_{off} was determined by fitting to the equation $V_{obs} = V_0 \cdot e^{-k_{off} \cdot t} + C$, where t is the time after addition of the sink, V_0 is the initial rate without any sink added and C is a constant to account for weak residual activity due to incomplete capture of Ste7 by the sink. The observed k_{off} of 0.2 s^{-1} is a lower limit – dissociation occurs on a timescale faster than can be measured with mixing by hand. No change in dissociation rates was observed in the presence or absence of Ste11 (Fig. S1).

To determine steady state rate constants (k_{cat} and K_M), the concentration of Fus3 K42R was varied from 10 nM to 1 μM at a constant, saturating concentration of 1 μM Ste5. To determine K_{act}, the concentration of Ste5 was varied from 10 nM to 3 μM at a constant, saturating concentration of 500 μM Fus3 K42R. Reactions were typically conducted at 50 nM Ste7EE. In separate experiments, the concentration of Ste7EE was varied over a >10-fold range from 20-250 nM to confirm that reaction rates scaled linearly with total enzyme concentration.

Plots of V_{obs} vs. [Fus3] were fit to the Michaelis-Menten equation ($V_{obs} = k_{cat}[E]_0[S]/(K_M+[S])$) where $[E]_0$ is the total Ste7EE concentration and [S] is the Fus3 concentration. To determine K_{act}, plots of V_{obs} vs. [Ste5] were fit to a simple binding equation ($V_{obs} = k_{cat}[E]_0/(1+(K_{act}/[Ste5]))$.

Structure Determination

Crystals of $Ste5_{582-786}$ were obtained by mixing 10 mg/mL of protein (1:1 vol:vol) with a solution containing 20% PEG 3350, 0.1 M $CaCl_2$, 25 mM Tris·HCl (pH 8.0), in hanging drops at 16 °C.

Crystals grew as long rods that reached a maximum size in two days. The cryopreservant for X-ray data collection consisted of mother liquor supplemented with 25% glycerol. Diffraction data was collected in-house on a Rigaku R-Axis IV. Diffraction data was processed using XDS (19). Phases were obtained by molecular replacement with PHASER (20) using Ste5-ms (PDB 3FZE) as a search model. The resulting maps were of excellent quality and the additional density for the N-terminal extension (residues 583-592) was readily visible in difference maps. The preliminary model obtained from molecular replacement was extended to include the additional residues, manually rebuilt in COOT (21), and refined using PHENIX (22). X-ray diffraction data and refinement statistics are shown in Table S4.

PH Domain Inhibition Experiments

For the *in trans* inhibition experiments, purified tag-free Ste5 PH domain (367-525) was exchanged into 25 mM Tris·HCl (pH 8.0), 100 mM NaCl, 2.5 mM $MgCl_2$, 0.05% IGEPAL, and 2 mM TCEP at 25 °C with a 7K MWCO Zebra 0.5 mL desalting column (Thermo Scientific). Varying concentrations of the PH domain were incubated with Ste5 VWA domain constructs (1 μM) for 30 minutes at room temperature prior to the addition of 50 nM Ste7EE and saturating (1 μM) Fus3 K42R. Reactions were initiated with 500 μM ATP at 25 °C and initial rates were determined by quantitative western blot as described above. The curves shown in Fig. 3E are described by the equation $k_{obs}=k_{cat}/(1+([PH]/K_I)$, where K_I was determined by curve fitting.

Fab Activation Experiments

For Fab activation experiments, purified SR13 Fab was exchanged into 25 mM Tris·HCl (pH 8.0), 100 mM NaCl, 2.5 mM $MgCl_2$, and 0.05% IGEPAL. SR13 Fab was incubated at varying

concentrations with 1 μM full-length Ste5 for 30 minutes at room temperature before the addition of 50 nM Ste7EE and saturating (1 μM) Fus3. Reactions were initiated with 500 μM ATP at 25 °C and initial rates were determined by quantitative western blot as described above.

Reconstituted Vesicle Experiments

Phosphatidyl choline (PC, egg yolk), PI(4,5)P_2 (brain), and DGS-NTA(Ni) were purchased from Avanti. Unilamellar vesicles were generated by mixing PC with varying quantities of PI(4,5)P_2 and/or DGS-NTA(Ni) dissolved in chloroform in the molar ratios as indicated in Fig. 3G. Mixed lipids were dried under argon, placed under vacuum for 12 hours, and resuspended in 25 mM Tris·HCl (pH 8.0), 100 mM NaCl, and 2.5 mM $MgCl_2$ to a final concentration of 3.2 mM total lipid. Hydrated lipids were subjected to ten cycles of freezing in liquid nitrogen, thawing, and 1 min bath sonication. Lipids were then extruded through a 0.2 μM filter (Avanti). Approximate concentrations of an indicated lipid species were estimated as the mole fraction of the total lipid concentration divided by 2, assuming 50% of the total was on the inside layer of the vesicles. The resulting vesicles were stored at 4 °C and used in assays within five days of preparation.

For vesicle activation experiments, 500 nM (final reaction concentration) of the $Ste5_{His\text{-}PM\text{-}PH\text{-}VWA}$ construct (Fig. S13) was incubated with lipid vesicles (at ~160 μM lipid concentration) in 25 mM Tris·HCl (pH 8.0), 100 mM NaCl, 2.5 mM $MgCl_2$, and 2 mM TCEP for 1 hour at 4 °C. The final lipid concentrations in reactions with 90:5:5 PC:PI(4,5)P_2:DGS-NTA(Ni) vesicles were approximately 144 μM PC, 8 μM PI(4,5)P_2, and 8 μM DGS-NTA(Ni). Following incubation, 50 nM Ste7EE and saturating (1 μM) Fus3 were added, the reactions were initiated with 500 μM

ATP at 25 °C, and initial rates were determined by quantitative western blot as described in the previous paragraph.

Supplemental Text References.

1. M. A. Schwartz, H. D. Madhani, Control of MAPK signaling specificity by a conserved residue in the MEK-binding domain of the yeast scaffold protein Ste5. *Curr. Genet.* **49**, 351 (2006).
2. J. G. Cook, L. Bardwell, J. Thorner, Inhibitory and activating functions for MAPK Kss1 in the *S. cerevisiae* filamentous-growth signalling pathway. *Nature* **390**, 85 (1997).
3. R. L. Roberts, G. R. Fink, Elements of a single MAP kinase cascade in *Saccharomyces cerevisiae* mediate two developmental programs in the same cell type: mating and invasive growth. *Gene Dev* **8**, 2974 (1994).
4. P. J. Cullen, G. F. Sprague, Jr., Glucose depletion causes haploid invasive growth in yeast. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 13619 (2000).
5. M. S. Longtine *et al.*, Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* **14**, 953 (1998).
6. A. L. Goldstein, J. H. McCusker, Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. *Yeast* **15**, 1541 (1999).
7. R. D. Gietz, R. H. Schiestl, High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protocols* **2**, 31 (2007).
8. R. S. Sikorski, P. Hieter, A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**, 19 (1989).
9. T. von der Haar, Optimized protein extraction for quantitative proteomics of yeasts. *PLoS One* **2**, e1078 (2007).

10. G. F. Sprague, Jr., Assay of yeast mating reaction. *Methods Enzymol.* **194**, 77 (1991).

11. M. Good, G. Tang, J. Singleton, A. Remenyi, W. A. Lim, The Ste5 scaffold directs mating signaling by catalytically unlocking the Fus3 MAP kinase for activation. *Cell* **136**, 1085 (2009).

12. S. Maleri *et al.*, Persistent activation by constitutive Ste7 promotes Kss1-mediated invasive growth but fails to support Fus3-dependent mating in yeast. *Mol. Cell. Biol.* **24**, 9221 (2004).

13. J. Walker, P. Crowley, A. D. Moreman, J. Barrett, Biochemical properties of cloned glutathione *S*-transferases from *Schistosoma mansoni* and *Schistosoma japonicum*. *Mol. Biochem. Parasitol.* **61**, 255 (1993).

14. R. Fabrini *et al.*, Monomer-dimer equilibrium in glutathione transferases: a critical re-examination. *Biochemistry* **48**, 10473 (2009).

15. A. Remenyi, M. C. Good, W. A. Lim, Docking interactions in protein kinase and phosphatase networks. *Curr. Opin. Struct. Biol.* **16**, 676 (2006).

16. F. A. Fellouse *et al.*, High-throughput generation of synthetic antibodies from highly functional minimalist phage-displayed libraries. *J. Mol. Biol.* **373**, 924 (2007).

17. F. A. Fellouse, S. S. Sidhu, in *Making and using antibodies,* G. C. Howard, M. R. Kaser, Eds. (CRC Press, Boca Raton, FL, 2007), pp. 157-180.

18. T. A. Kunkel, J. D. Roberts, R. A. Zakour, Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* **154**, 367 (1987).

19. W. Kabsch, XDS. *Acta Cryst.* **D66**, 125 (2010).

20. A. J. McCoy *et al.*, Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658 (2007).

21. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Cryst.* **D66**, 486 (2010).

22. P. D. Adams *et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst.* **D66**, 213 (2010).

23. R. P. Bhattacharyya *et al.*, The Ste5 scaffold allosterically modulates signaling output of the yeast mating pathway. *Science* **311**, 822 (2006).

24. C. Inouye, N. Dhillon, T. Durfee, P. C. Zambryski, J. Thorner, Mutational analysis of STE5 in the yeast *Saccharomyces cerevisiae*: Application of a differential interaction trap assay for examining protein-protein interactions. *Genetics* **147**, 479 (1997).

25. W. Sabbagh, L. J. Flatauer, A. J. Bardwell, L. Bardwell, Specificity of MAP kinase signaling in yeast differentiation involves transient versus sustained MAPK activation. *Mol. Cell* **8**, 683 (2001).

26. S. C. Strickfaden *et al.*, A mechanism for cell-cycle regulation of MAP kinase signaling in a yeast differentiation pathway. *Cell* **128**, 519 (2007).

27. M. J. Winters, R. E. Lamson, H. Nakanishi, A. M. Neiman, P. M. Pryciak, A membrane binding domain in the Ste5 scaffold synergizes with Gβγ binding to control localization and signaling in pheromone response. *Mol. Cell* **20**, 21 (2005).

28. C. Inouye, N. Dhillon, J. Thorner, Ste5 RING-H2 domain: Role in Ste4-promoted oligomerization for yeast pheromone signaling. *Science* **278**, 103 (1997).

29. L. S. Garrenton, S. L. Young, J. Thorner, Function of the MAPK scaffold protein, Ste5, requires a cryptic PH domain. *Gene Dev* **20**, 1946 (2006).

30. B. J. Stevenson, N. Rhodes, B. Errede, G. F. Sprague, Jr., Constitutive mutants of the protein kinase STE11 activate the yeast pheromone response pathway in the absence of the G protein. *Genes Dev.* **6**, 1293 (1992).

31. L. J. Flatauer, S. F. Zadeh, L. Bardwell, Mitogen-activated protein kinases with distinct requirements for Ste5 scaffolding influence signaling specificity in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **25**, 1793 (2005).

32. J. Andersson, D. M. Simpson, M. S. Qi, Y. M. Wang, E. A. Elion, Differential input by Ste5 scaffold and Msg5 phosphatase route a MAPK cascade to multiple outcomes. *EMBO J.* **23**, 2564 (2004).

33. M. K. Malleshaiah, V. Shahrezaei, P. S. Swain, S. W. Michnick, The scaffold protein Ste5 directly controls a switch-like mating decision in yeast. *Nature* **465**, 101 (2010).

34. R. E. Lamson, S. Takahashi, M. J. Winters, P. M. Pryciak, Dual role for membrane localization in yeast MAP kinase cascade activation and its contribution to signaling fidelity. *Curr. Biol.* **16**, 618 (2006).

35. C. Sette, C. J. Inouye, S. L. Stroschein, P. J. Iaquinta, J. Thorner, Mutational analysis suggests that activation of the yeast pheromone response mitogen-activated protein kinase pathway involves conformational changes in the Ste5 scaffold protein. *Mol. Biol. Cell* **11**, 4033 (2000).

36. Y. Wang, E. A. Elion, Nuclear export and plasma membrane recruitment of the Ste5 scaffold are coordinated with oligomerization and association with signal transduction components. *Mol. Biol. Cell* **14**, 2543 (2003).

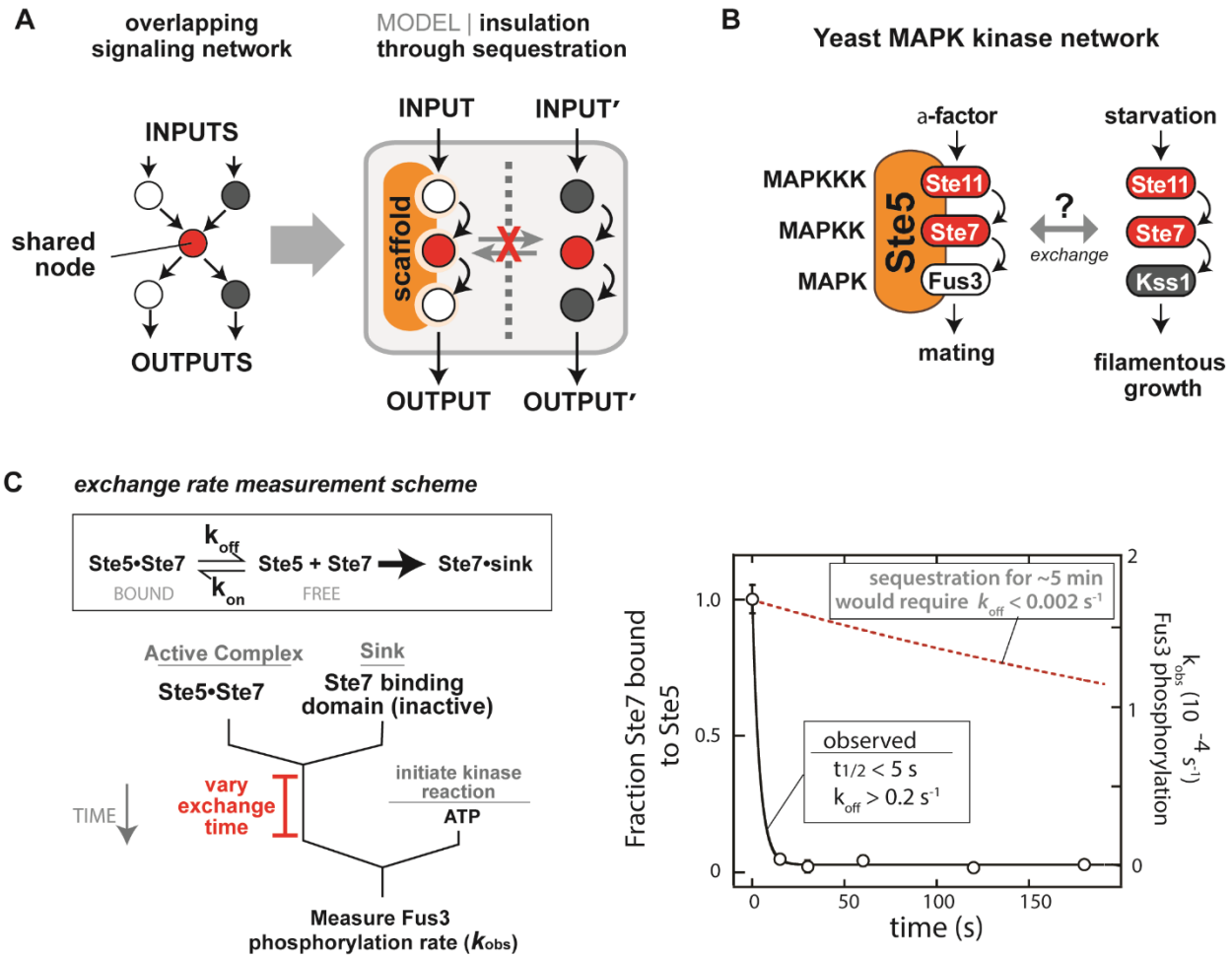
37. D. Yablonski, I. Marbach, A. Levitzki, Dimerization of Ste5, a mitogen-activated protein kinase cascade scaffold protein, is required for signal transduction. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13864 (1996).

38. B. R. Cairns, S. W. Ramer, R. D. Kornberg, Order of action of components in the yeast pheromone response pathway revealed with a dominant allele of the STE11 kinase and the multiple phosphorylation of the STE7 kinase. *Genes Dev.* **6**, 1305 (1992).

39. F. van Drogen *et al.*, Phosphorylation of the MEKK Ste11p by the PAK-like kinase Ste20p is required for MAP kinase signaling *in vivo*. *Curr. Biol.* **10**, 630 (2000).

40. G. Jansen, F. Buhring, C. P. Hollenberg, M. Ramezani Rad, Mutations in the SAM domain of *STE50* differentially influence the MAPK-mediated pathways for mating, filamentous growth and osmotolerance in *Saccharomyces cerevisiae*. *Mol. Genet. Genomics* **265**, 102 (2001).

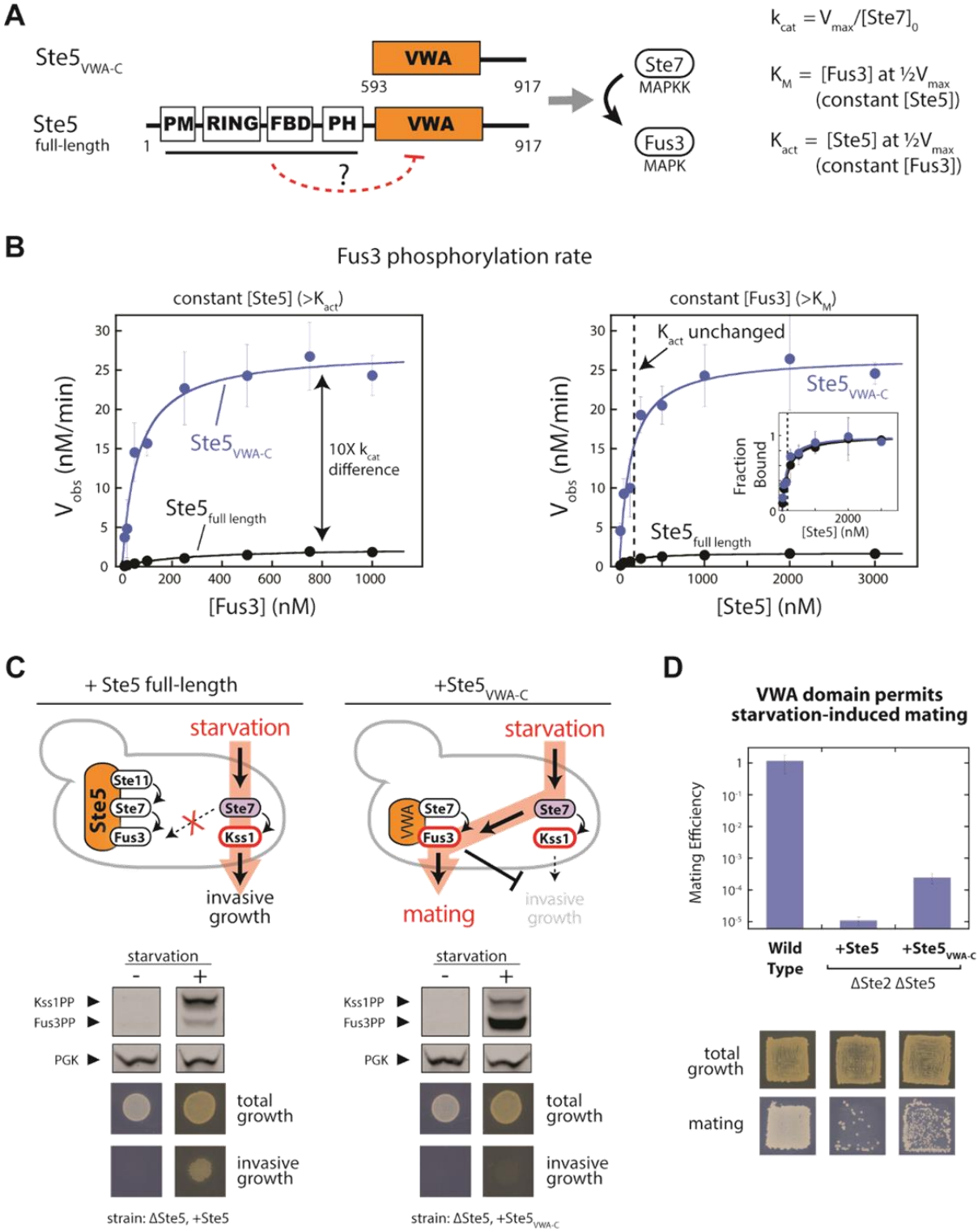
Figure 1. Exchange of the Ste7 MAPKK from the Ste5 scaffold protein is too fast for signaling pathway insulation by physical sequestration



(A) Shared components of the yeast mating and invasive growth pathways yield physiologically distinct input-output responses. (B) Dissociation rate of the MAPKK Ste7 from the Ste5 scaffold protein measured with purified recombinant Ste5, the MAPKKK Ste11, the MAPK Fus3, and a constitutively active form of the MAPKK Ste7 (Ste7EE, bearing phosphomimic mutations in the Ste7 activation loop (*16*)). To a preassembled Ste5-Ste11-Ste7-Fus3 complex, an excess of a Ste7 binding domain (a minimal Ste7 binding domain from Ste5 [residues 759-810]) was added to capture Ste7 as it dissociated from Ste5 (Fig. S1). At various times, ATP was added, and the

initial rate of Fus3 phosphorylation was measured (the amount of Ste5-Ste7-Fus3 complex remaining at each timepoint). Error bars are standard deviations. The observed k_{off} of 0.2 s^{-1} is a lower limit – dissociation occurred on a timescale faster than could be measured with mixing by hand.

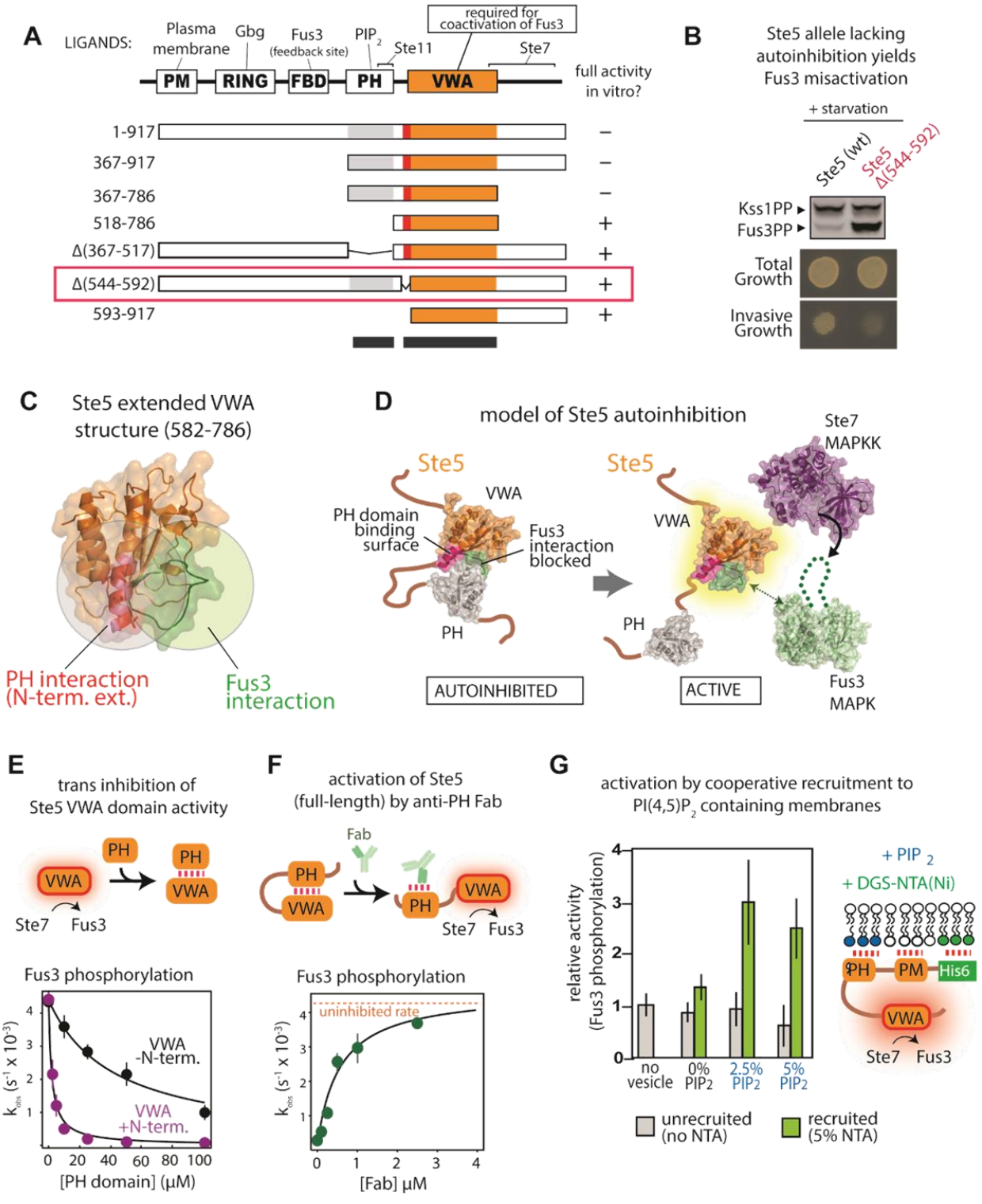
Figure 2. Autoinhibition of the Ste5 scaffold protein insulates the MAPK Fus3 from activation by incorrect inputs.



(A) Full-length Ste5 (residues 1-917) and Ste5$_{VWA-C}$ (residues 593-917) used for *in vitro* kinetic

assays for phosphorylation of Fus3, Michaelis-Menten plot of V_{obs} vs. [Fus3], and plot of V_{obs} vs. [Ste5]. K_{act} corresponds to the midpoint of the V_{obs} vs. [Ste5] plot and represents the dissociation constant for Ste5-Ste7. Error bars are standard deviations. It was unnecessary to measure binding affinity for Fus3 assembly into the ternary complex because the Ste5 VWA domain does not bind with any detectable affinity to Fus3 (Fus3 is recruited to this ternary catalytic complex via binding to Ste7; the interaction between the Ste5 VWA domain and Fus3 is a transient catalytic interaction (*16*)). See fig. S2-S4 and table S1 for values of fitted kinetic constants. (B) Fus3 misactivation in response to starvation in yeast cells expressing $Ste5_{VWA-C}$ (see Methods for growth conditions). Fus3 and Kss1 phosphorylation was monitored with an antibody to phosphorylated MAPKs by protein immunoblotting, and phosphoglycerate kinase (PGK) is shown as a loading control for yeast lysates (see fig. S5-S7 for quantitative analysis). The invasive growth response was assayed with yeast cells grown on solid agar plates (see Methods). Similar results were obtained with constitutively active alleles of Ste11 and Ste7 (Fig. S8). (C) $Ste5_{VWA-C}$ misdirects signaling to Fus3 and allows cells to mate in response to starvation. Mating efficiencies were determined using a quantitative mating assay, and patch assays were done as described (see Methods). Error bars are standard deviations.

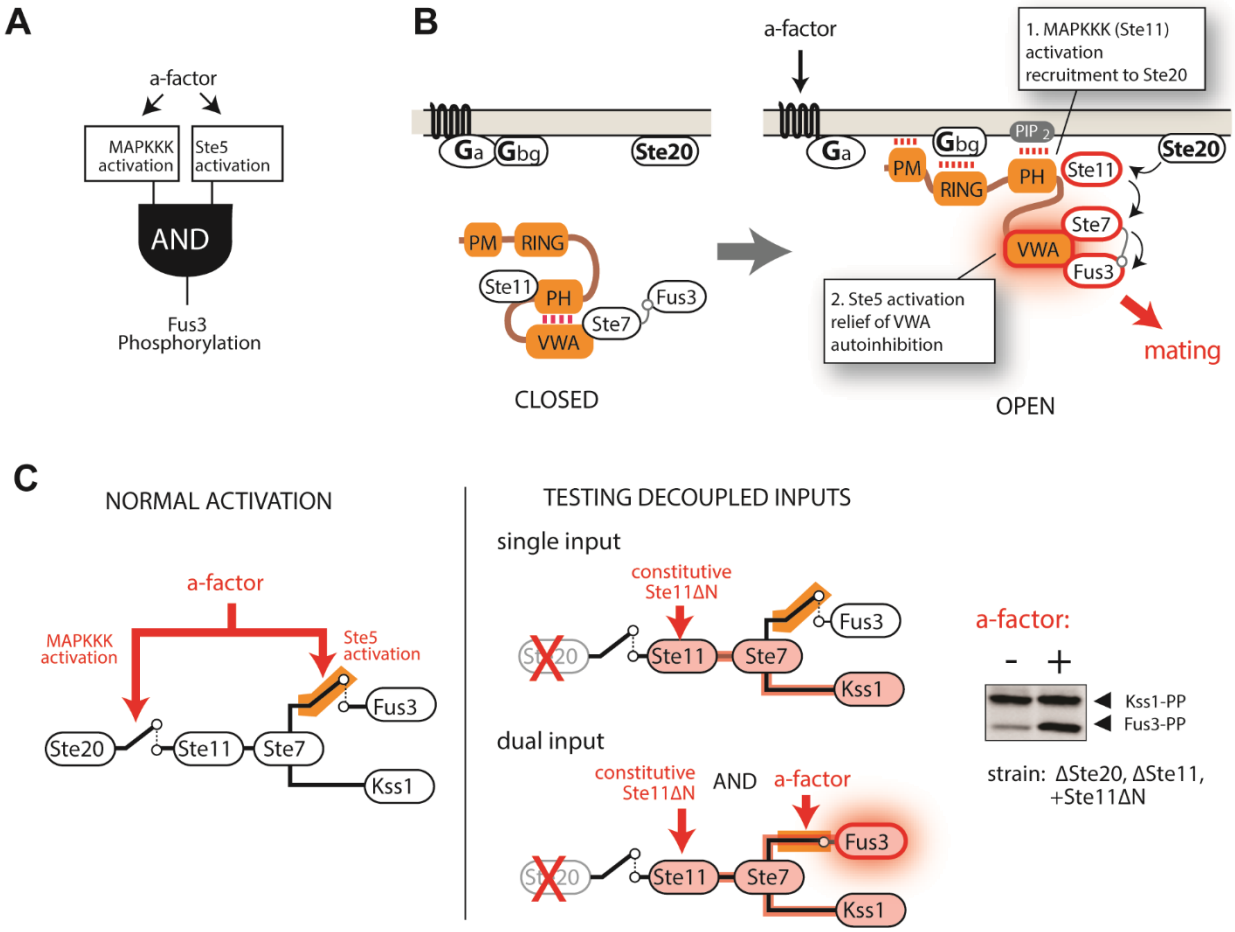
Figure 3. Model for the mechanism of Ste5 autoinhibition and activation by mating-specific input.



(A) Diagram of truncation mapping of Ste5 (red is residues 544-592). The minimal autoinhibited

fragment (367-786) contains all elements necessary to assemble the three-tiered MAPK cascade (Fus3 is recruited to the VWA domain by Ste7 (*16*)). See fig. S9 for additional constructs. (B) Effect of deletion of the N-terminal extension of the VWA domain in Ste5 (residues 544-592) to disrupt pathway insulation *in vivo*. Activation of Fus3 measured by protein immunoblotting, and invasive growth assayed on solid agar plates (assays conducted as in fig. 2C, see fig. S6). (C) Crystal structure of $Ste5_{582\text{-}786}$ (see Table S4 for crystallographic statistics). The N-terminal extension (582-592) is shown in pink, with surface residues necessary for trans inhibition by the PH domain (Fig. S10) shown as sticks in red. The Fus3-coactivating loop of Ste5 (743-756) is shown in green. The spatial proximity of the PH-domain interface and the site of Fus3 coactivation is illustrated by overlapping circles. (D) Model for Ste5 autoinhibition inferred from truncation mapping data and $Ste5_{582\text{-}786}$ crystal structure. (E) Titration of the VWA domain with the PH domain results in inhibition of the Fus3 coactivator activity (k_{obs}) of the VWA construct bearing the N-terminal extension (582-786, shown in pink), or lacking the N-terminal extension (593-786, shown in black). Error bars are standard deviations. (F) Relief of autoinhibition in full-length Ste5 by a Fab antibody (SR13) that can bind the PH domain. Error bars are standard deviations. (G) Recruitment of Ste5 to membranes with PIP_2 stimulates coactivation of Fus3. A minimal, autoinhibited Ste5 fragment bearing a hexahistidine tag can be recruited to small unilamellar vesicles of varying lipid compositions by the DGS-NTA(Ni) lipid (see fig. S13 for exact details of the Ste5 construct used here). Error bars are standard deviations.

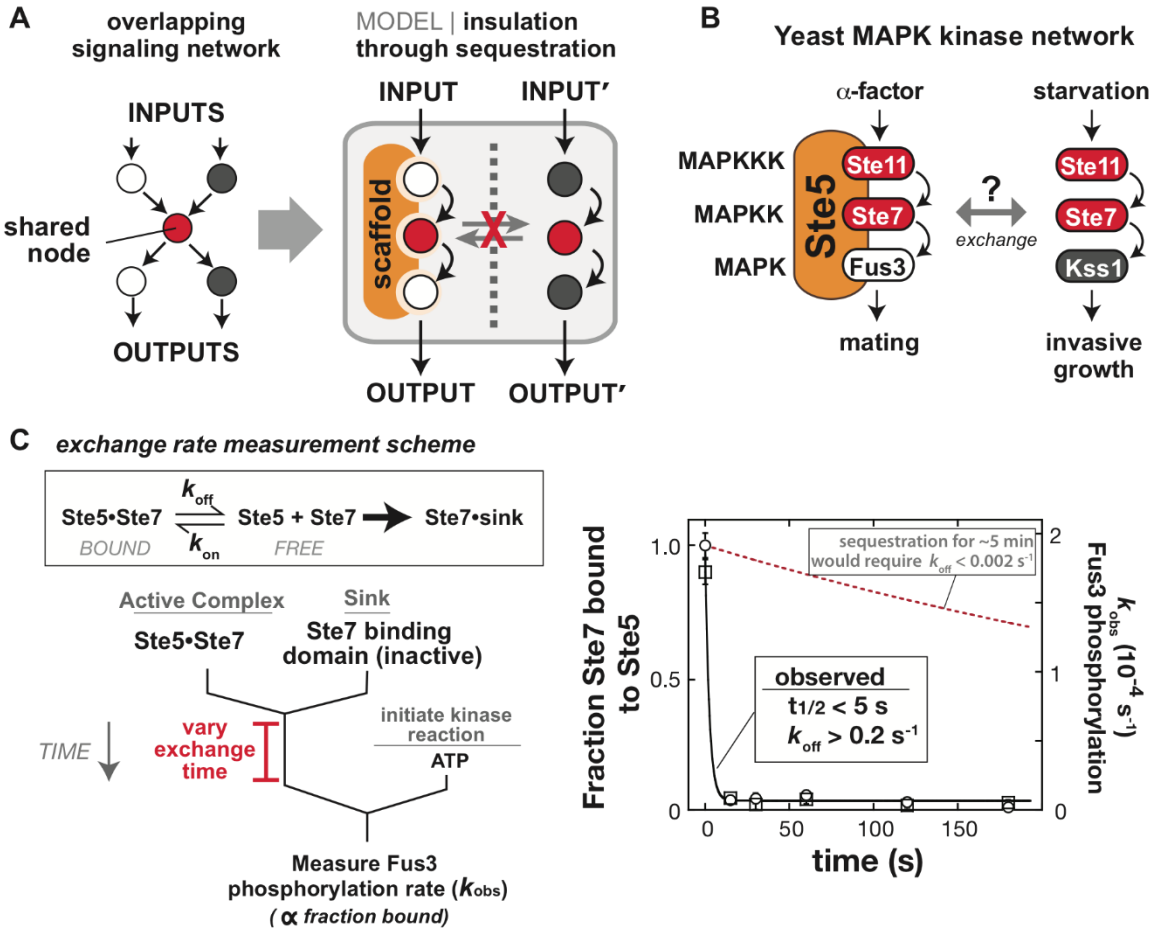
Figure 4. Model for membrane recruitment-triggered conformational activation of Ste5 and activation of the mating MAPK Fus3.



(A) Simple AND-gate model for specific mating pathway activation. Non-mating inputs that activate the shared MAPKKK do not activate Fus3. (B) Revised molecular model for mating pathway activation mediated by the Ste5 scaffold protein. Mating pheromone (α-factor) activates a heterotrimeric G protein, leading to release of the Gβγ subunit from Gα and recruitment of Ste5 to free Gβγ at the membrane (*21*). Membrane recruitment triggers activation of the MAPKKK Ste11 and PH domain binding to PIP_2, leading to release of the VWA domain and relief of autoinhibition. (C) Fus3 activation *in vivo* when kinase cascade activation is decoupled

from the mating signal (α-factor), measured by protein immunoblotting. See fig. S15 for additional Ste11 alleles and controls.

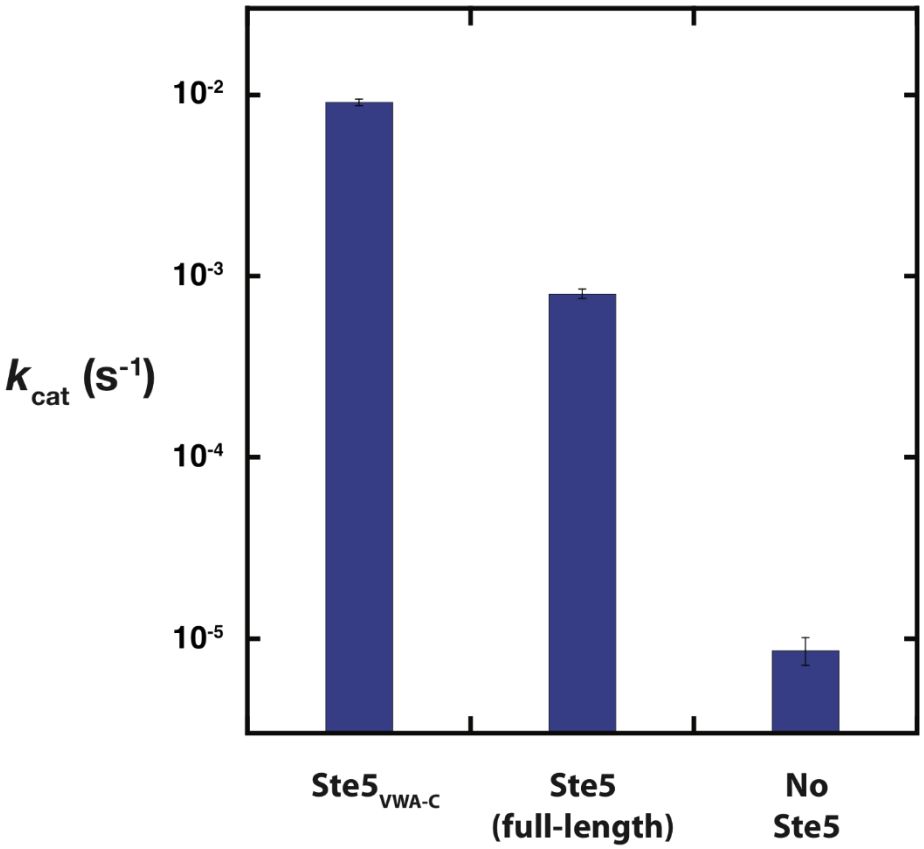
Figure S1. Exchange of the Ste7 MAPKK from the Ste5 scaffold protein is too fast for signaling pathway insulation by physical sequestration



(A) Signaling networks frequently share components. Physical sequestration by scaffold proteins has been proposed as a mechanism for pathway insulation by preventing exchange between distinct pools of signaling proteins. (B) The yeast mating and invasive growth pathways are physiologically distinct input-output responses that share two common components – the MAPKKK Ste11 and the MAPKK Ste7. The Ste5 scaffold is essential for signaling through the mating pathway. (C) The dissociation rate of the MAPKK Ste7 from the Ste5 scaffold protein is fast. Dissociation rates were measured with purified recombinant Ste5, the MAPKKK Ste11, the MAPK Fus3, and a constitutively active form of the MAPKK Ste7 (Ste7EE, bearing phosphomimic mutations in the Ste7 activation loop). To a preassembled Ste5-Ste7-Fus3

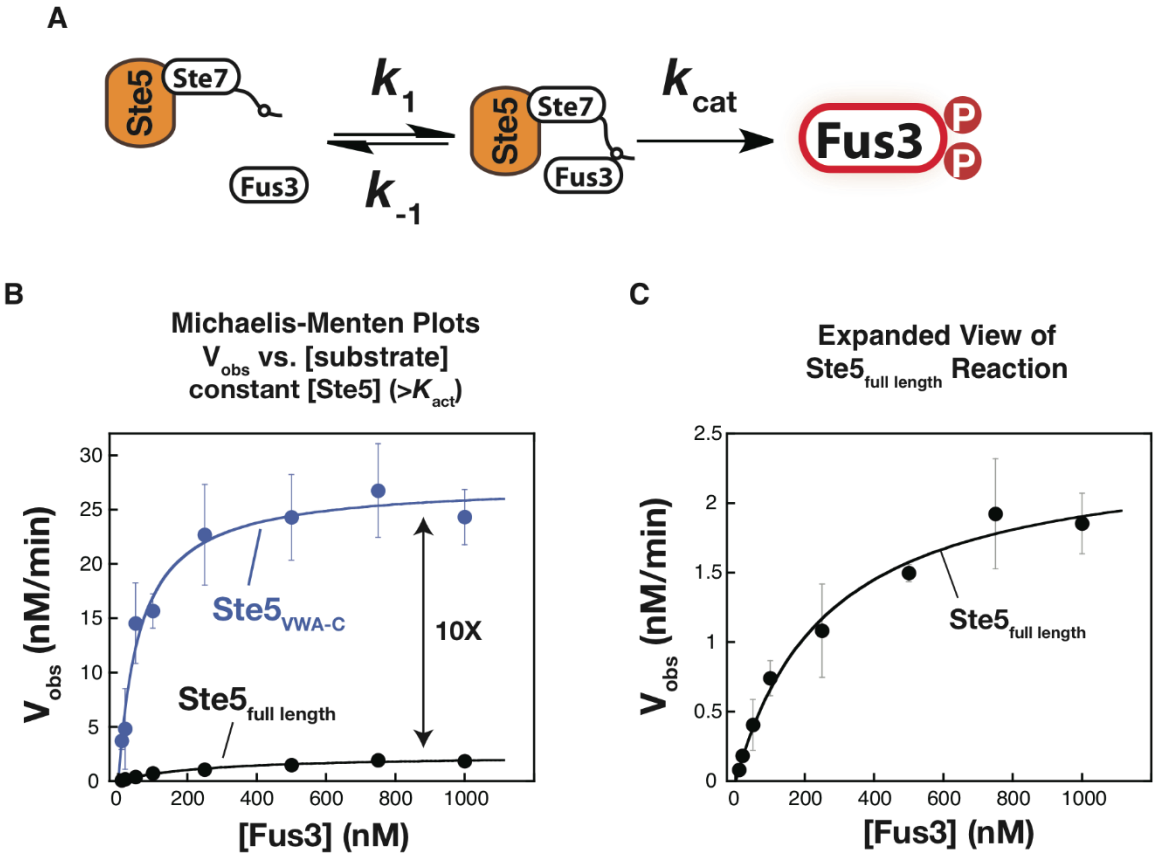
complex, an excess of a Ste7 binding domain (the “sink”, a minimal Ste7 binding domain from Ste5 [residues 759-810]) was added to capture Ste7 as it dissociates from Ste5. After varying times, ATP was added, and the initial rate of Fus3 phosphorylation corresponds to the amount of Ste5-Ste7-Fus3 complex remaining at each timepoint. The observed k_{off} of 0.2 s^{-1} is a lower limit – dissociation occurs on a timescale faster than can be measured with mixing by hand. There was no detectable difference in the Ste5-Ste7 dissociation rate measured in the presence (○) or absence (□) of Ste11.

Figure S2. Ste5 has large effects on k_{cat} for Fus3 phosphorylation *in vitro*



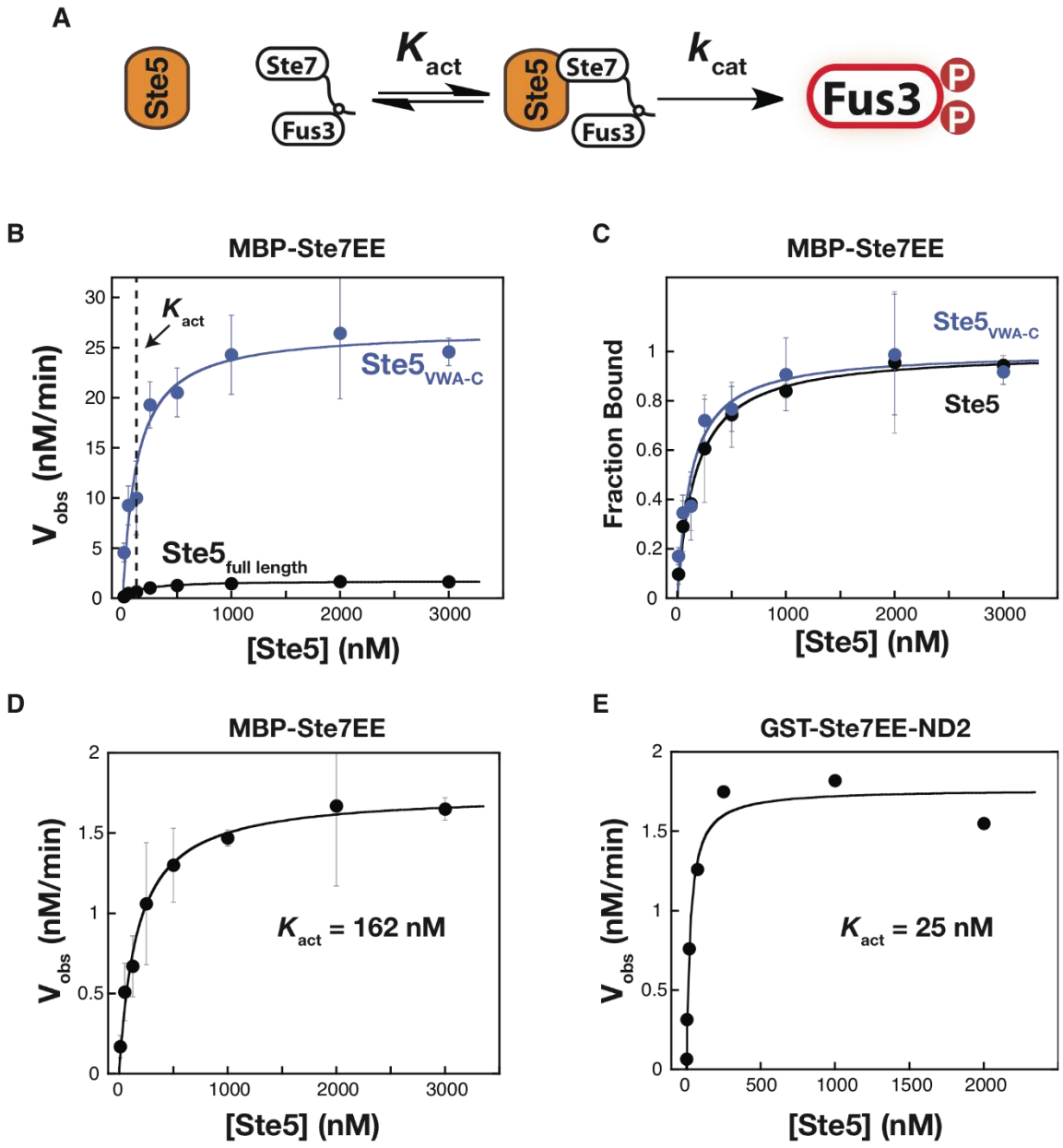
Values of k_{cat} were measured for MBP-Ste7EE-catalyzed Fus3 phosphorylation in the absence of Ste5, with full-length Ste5, and with $Ste5_{VWA-C}$. $Ste5_{VWA-C}$ increases k_{cat} by ~10^3-fold, while full-length Ste5 increases k_{cat} by ~10^2-fold relative to the reaction in the absence of Ste5. The value of k_{cat} for MBP-Ste7EE in the absence of Ste5 of $(8.6 \pm 1.5) \times 10^{-6}$ is ~10-fold greater than that measured previously with GST-Ste7EE-ND2 (*11*).

Figure S3. The VWA-C domain is a significantly better coactivator of Fus3 phosphorylation than full-length Ste5



(A) Kinetic scheme for Ste7→Fus3 phosphorylation in the presence of saturating Ste5. (B) Michaelis-Menten plots (V_{obs} vs. [Fus3]) for MBP-Ste7EE with $Ste5_{VWA-C}$ and full-length Ste5. (C) Expanded view of data for full-length Ste5. The value of k_{cat} is 10-fold slower for full-length Ste5 relative to the free VWA-C domain. The K_M is approximately 4-fold weaker for full-length Ste5, likely due to an additional binding site for the Fus3 substrate within the full-length Ste5, which can titrate out substrate (*15*, *23*). Fitted values of kinetic constants are reported in Table S1.

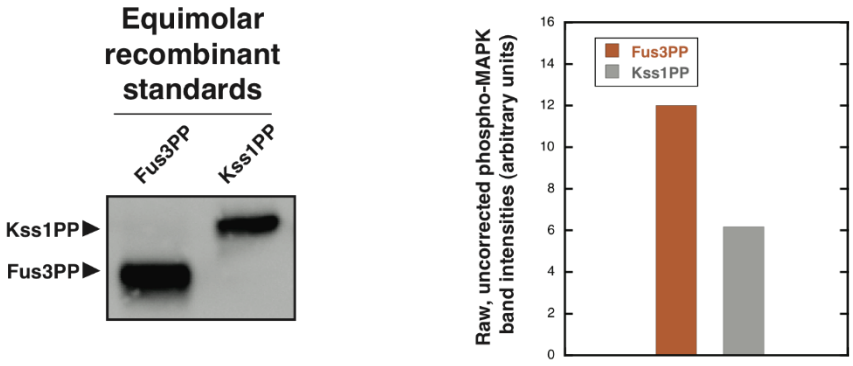
Figure S4. Full-length Ste5 and the VWA-C domain have similar binding affinities for Ste7



(A) Kinetic scheme for Ste7→Fus3 phosphorylation at saturating Fus3 and varying concentrations of Ste5. K_{act} is the midpoint of the activation curve and corresponds to the dissociation constant for Ste5-Ste7. (B) Plot of V_{obs} vs. [Ste5] for reactions of full-length Ste5 and $Ste5_{VWA-C}$ indicate that Ste5-Ste7 affinity binding is the same for these Ste5 constructs. (C) Same as (B) but with the data normalized to demonstrate that the values of K_{act} for full-length

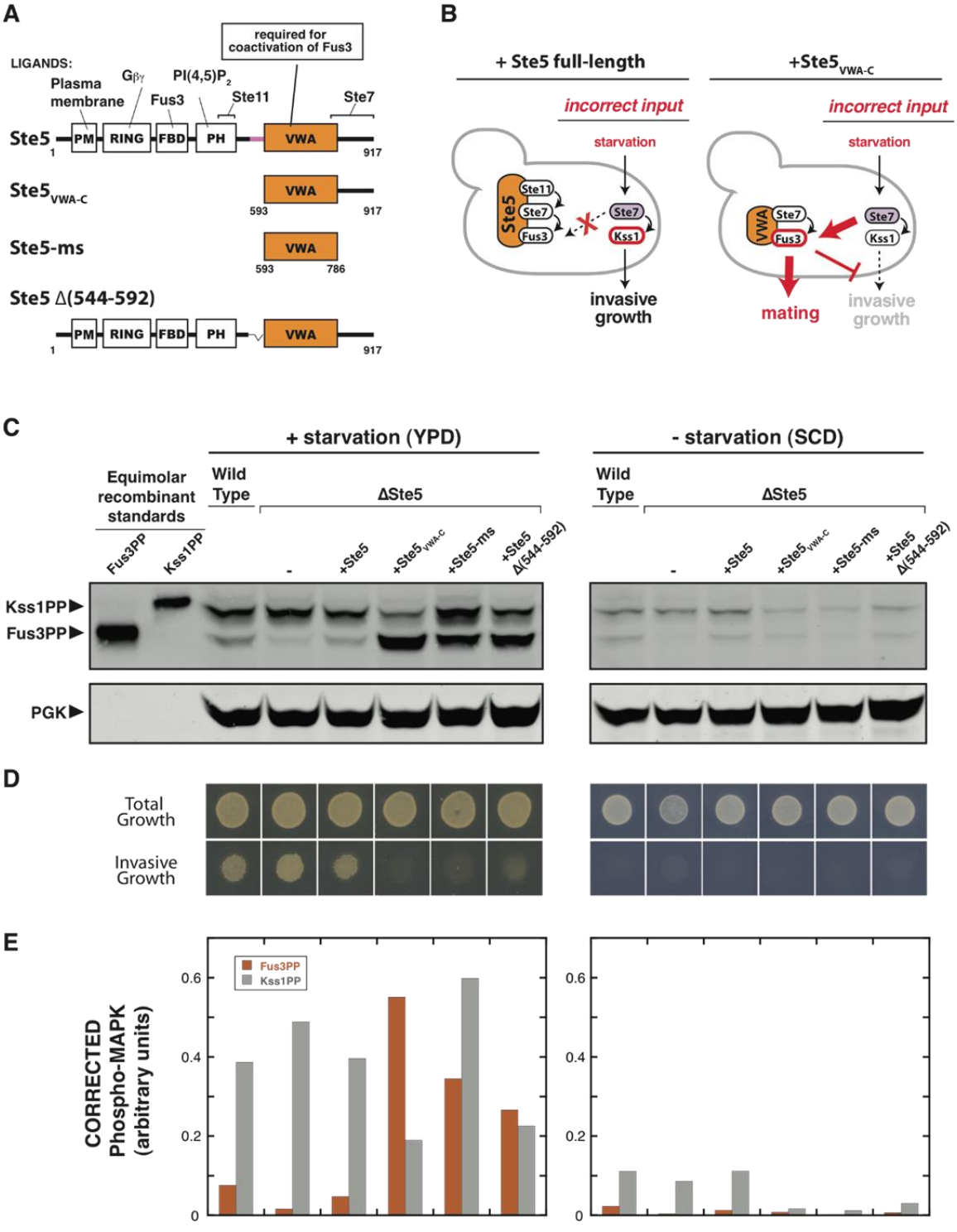
Ste5 and Ste5$_{VWA-C}$ are indistinguishable. (D and E) Comparison of K_{act} for MBP-Ste7EE and GST-Ste7EE-ND2 (with full-length Ste5, using 50 nM Ste7). Although GST-Ste7EE-ND2 was used in previous work (*11*), MBP-Ste7EE was used for all data reported in this work because it produced higher purity protein and avoids any potential artifacts arising from GST dimerization (*13*, *14*). The value of K_{act} of 25 ± 7 nM for GST-Ste7EE-ND2 is an upper limit for K_{act} because it is approximately half of the concentration of Ste7 in the reaction and thus may reflect titration conditions. This value is substantially tighter than the value of 161 nM previously reported for GST-Ste7EE-ND2 binding to Ste5-ms (residues 593-786) (*11*) because full-length Ste5 contains the complete binding site for Ste7 that extends into the C-terminal tail (*24*). The tighter apparent binding by GST-Ste7EE-ND2 could be related to GST dimerization. The ND2 designation indicates that the second, weak docking site for Fus3 has been removed (*15*). There is no difference in the behavior of Ste7EE and Ste7EE-ND2 for the kinetic parameters reported in this work.

Figure S5. Calibration of recognition of phospho-Fus3 and phospho-Kss1 on western blot allows for quantitative analysis of Fus3PP/Kss1PP levels *in vivo*



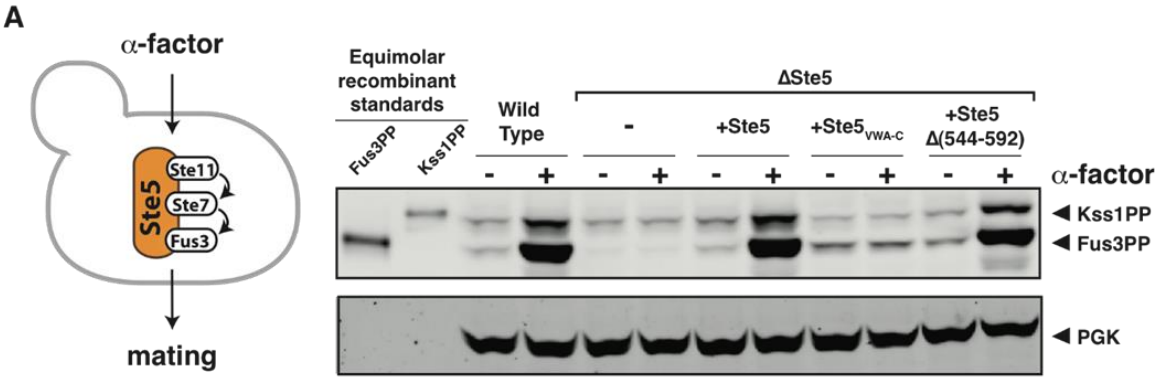
The anti-phospho p44/p42 MAPK antibody (Cell Signaling Technology #4370) preferentially recognizes Fus3PP relative to Kss1PP. The average ratio of Fus3PP to Kss1PP is 2.1 ± 0.2 (± standard deviation, 10 measurements). Recombinant standards of Fus3PP and Kss1PP were prepared from proteins purified as described (*11*) and phosphorylated to completion with Ste7EE (and $Ste5_{VWA-C}$ for Fus3). The standards were loaded in equimolar quantities and analyzed by western blot using the anti-phospho p44/42 MAPK antibody (Cell Signaling Technology #4370). Band intensities were quantified using Li-Cor Odyssey 2.1 software.

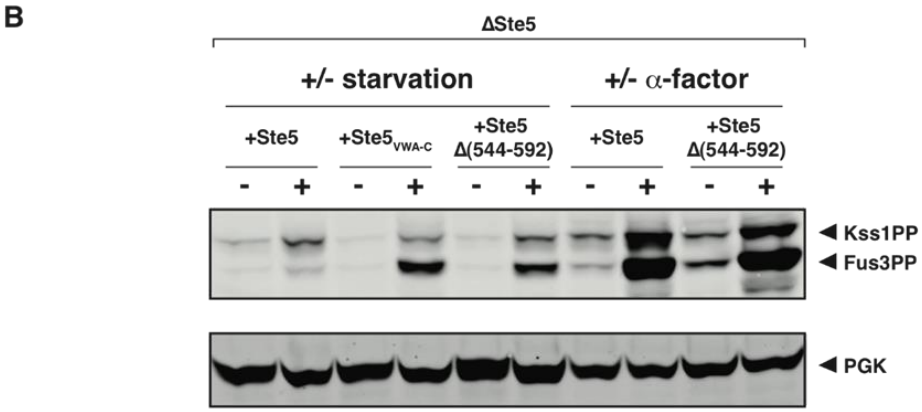
Figure S6. The Ste5 VWA-C domain promotes misactivation of the mating MAPK Fus3 in response to starvation

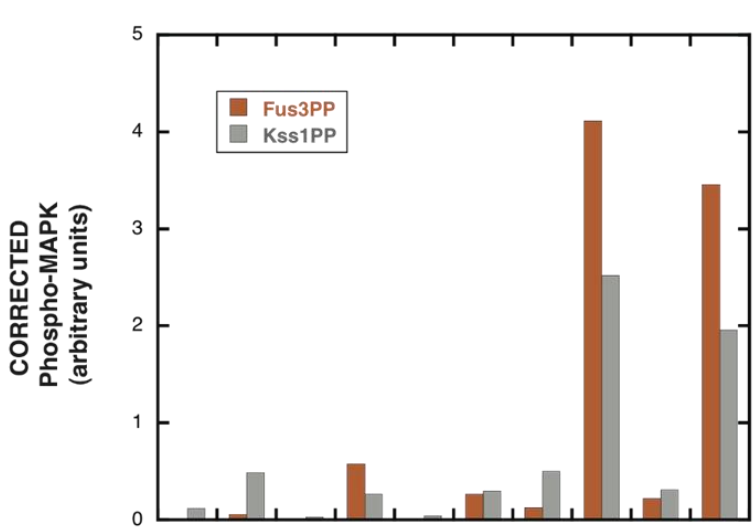


(A) Schematic diagram of full-length Ste5 and truncation mutants. $Ste5_{VWA-C}$ is the minimal VWA domain plus the C-terminal tail that includes the complete Ste7 binding region (*24*). Ste5-ms is the minimal VWA domain (residues 593-786, previously described in (*11*)). Ste5Δ(544-592) is an internal truncation that is autoinhibited *in vitro* (Fig. 3A and 3B). (B) Schematic for how starvation of yeast leads to misactivation of Fus3 when autoinhibition is disrupted in Ste5. (C) Phosphorylation of MAP kinases in starvation conditions (growth on YPD plates) or unstarved (growth on SCD plates) monitored by anti-phospho western blotting. In the presence or absence of full-length Ste5, starvation activates Kss1. When Ste5 is replaced by $Ste5_{VWA-C}$, Ste5-ms, or Ste5Δ(544-592), starvation leads to Fus3 activation. Fus3 activation is strongest for $Ste5_{VWA-C}$. Active Fus3 has been reported to promote Kss1 dephosphorylation (*25*), leading to the decrease in Kss1 activation observed with $Ste5_{VWA-C}$. Phosphoglycerate kinase (PGK) is shown as a loading control for yeast lysates. All Ste5 constructs were expressed from the Ste5 promoter. (D) Activation of Fus3 disrupts the invasive growth phenotype on YPD plates. (E) Corrected relative amounts of Fus3PP and Kss1PP, performed as described in fig. S5 and normalized to PGK loading control.

Figure S7. Comparison of α-factor and starvation responses

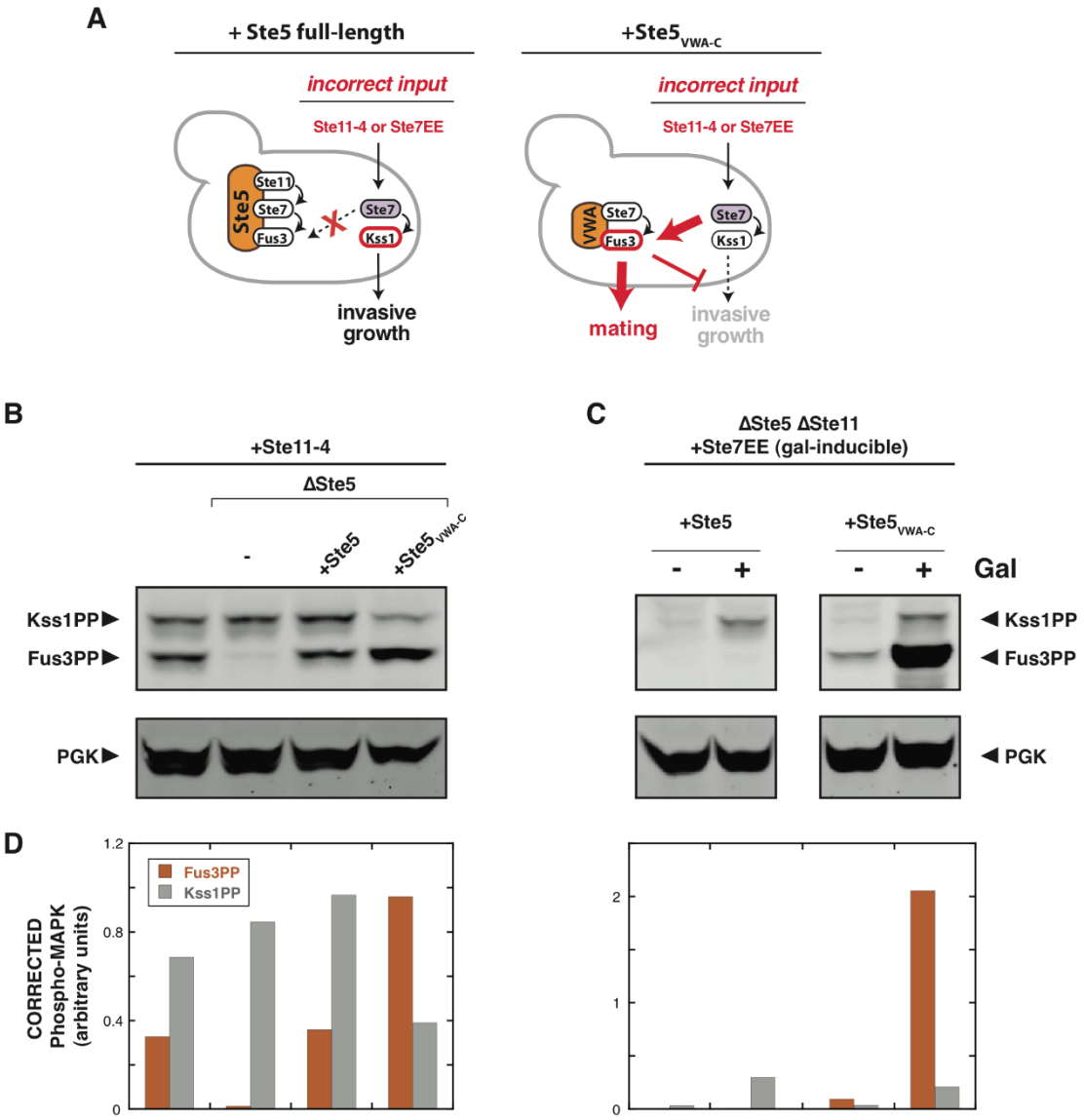






(A) α-factor treatment activates the yeast mating MAP kinase pathway. Phosphorylation of MAP kinases in response to α-factor treatment monitored by anti-phospho western blotting. Cells with full-length Ste5 display robust activation of Fus3 and Kss1, and cells with Ste5Δ(544-592) are indistinguishable. The $Ste5_{VWA-C}$ construct lacks key domains necessary for membrane recruitment in response to α-factor (the PM (*26, 27*), RING (*28*), and PH (*29*) domains, see schematic in fig. S6) and for Ste11 binding (the PH domain (*24*)). All Ste5 constructs were expressed from the Ste5 promoter. (B) Comparison of the starvation response and the mating response. (C) Corrected relative amounts of Fus3PP and Kss1PP, performed as described in fig. S5 and normalized to PGK loading control. Cells produce significantly more MAPK phosphorylation in response to saturating α-factor (2 μM) than in response to starvation.

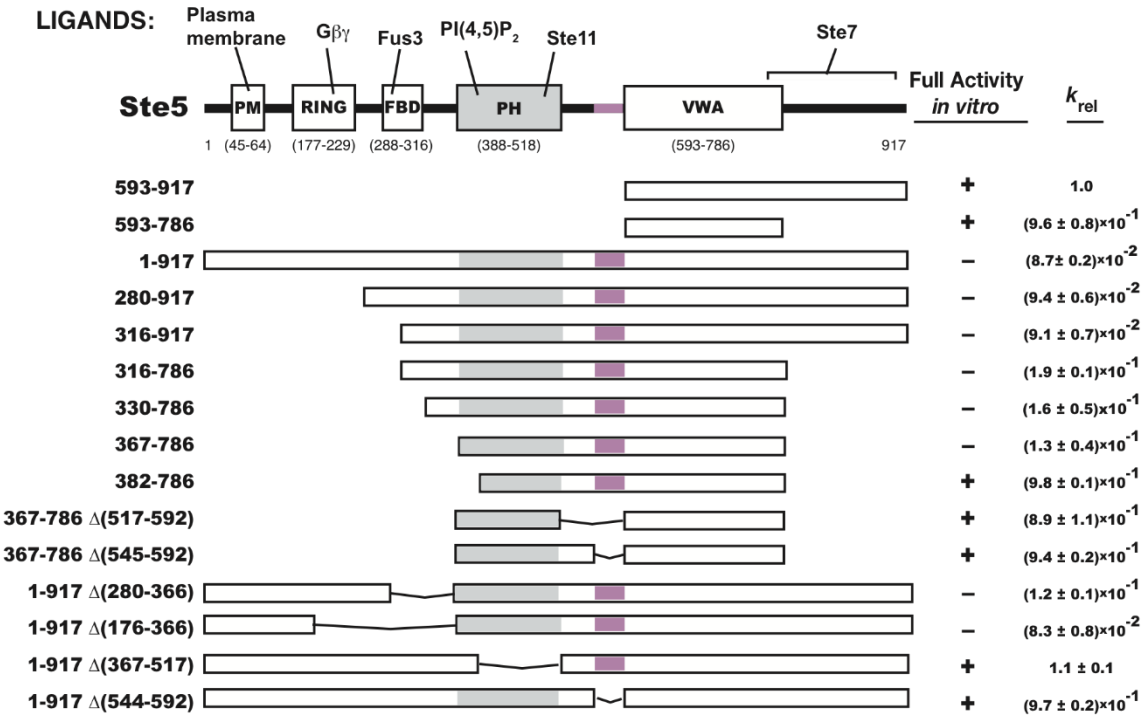
Figure S8. Ste5$_{VWA-C}$ promotes misactivation of the mating MAPK Fus3 in response to constitutively-active alleles of Ste11 and Ste7



(A) Schematic diagram for the effect of constitutive alleles of Ste11 and Ste7 in cells with full-length Ste5 and Ste5$_{VWA-C}$. Ste11-4 is a constitutively active allele of Ste11 (*30*) that preferentially activates Kss1 (*31, 32*). Ste7EE is a constitutively active allele of Ste7 (*11, 12*). (B) Phosphorylation of MAP kinases monitored by anti-phospho western blotting. When Ste5 is replaced by Ste5$_{VWA-C}$, Ste11-4 leads to preferential Fus3 phosphorylation. The decrease in Kss1

phosphorylation is a consequence of active Fus3 promoting Kss1 dephosphorylation (*25*). Yeast cells were harvested from YPD liquid cultures grown to mid-log phase (OD_{600} = 0.5). PGK is shown as a loading control for yeast lysates. Ste5 and Ste11 constructs in this panel were expressed from their native promoters. (C) Ste7EE preferentially phosphorylates Fus3 only in the presence of $Ste5_{VWA-C}$. To observe significant MAPK phosphorylation, Ste7EE was overexpressed from a galactose-inducible promoter, and Ste5 constructs were overexpressed from Adh promoters. These experiments were conducted in a strain background that is unable to activate the MAP kinase cascade (ΔSte5 ΔSte11) so that the observed MAPK phosphorylation is solely the product of Ste7EE. (D) Corrected relative amounts of Fus3PP and Kss1PP, performed as described in fig. S5 and normalized to PGK loading control.

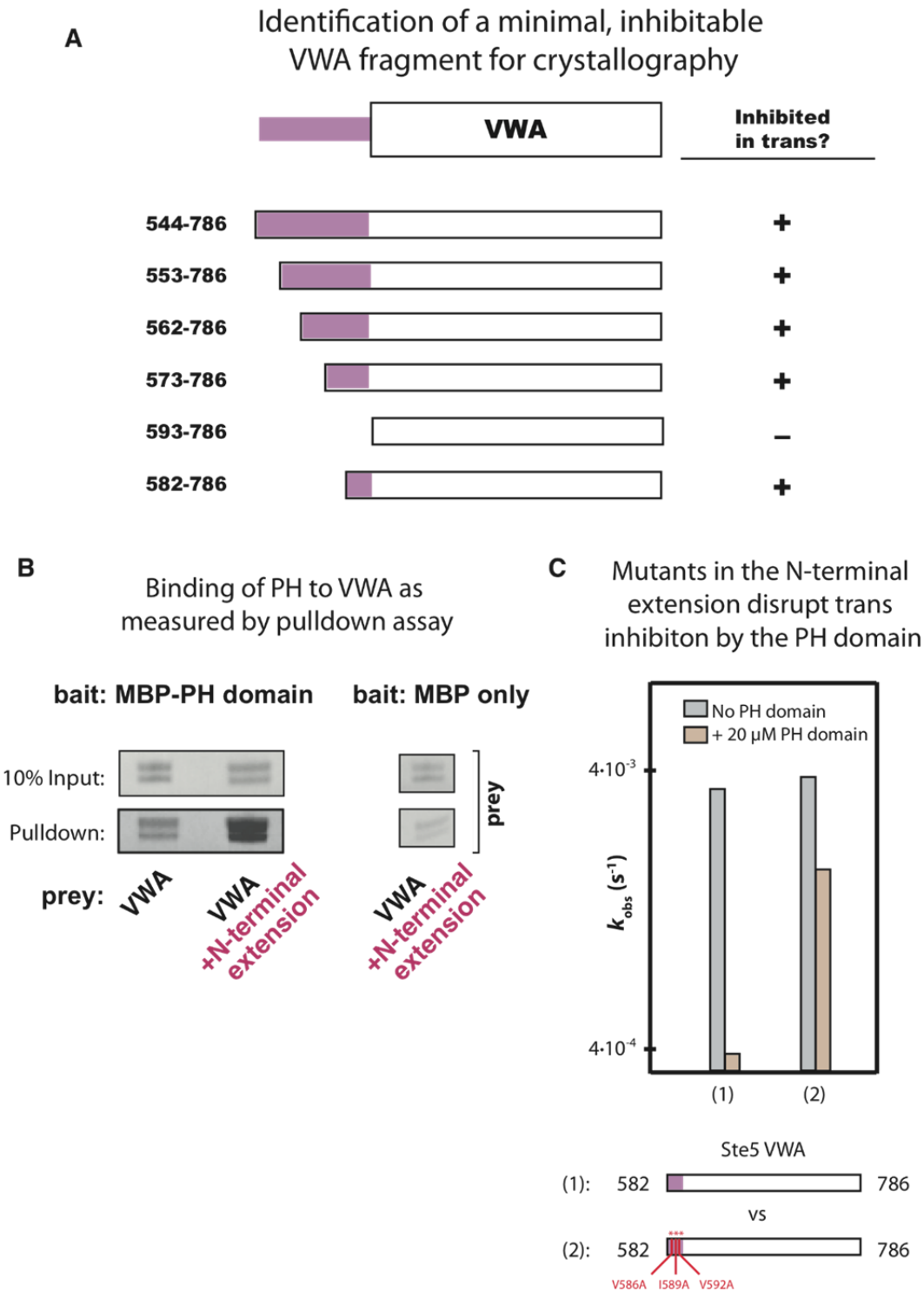
Figure S9. Detailed truncation mapping suggests that the PH domain of Ste5 interacts with and inhibits the VWA domain



Ste5 contains several well-characterized domains, including the PM helix, which facilitates membrane binding and mediates cell-cycle regulation of Ste5 activity (*26*, *27*); the RING-H2 domain, which binds Gβ (Ste4) to localize Ste5 to the membrane when the pathway is stimulated (*28*); the Fus3-binding domain (FBD), which is not required for Fus3 activation but instead serves a feedback regulatory role (*23, 33*); the PH domain, which binds to phosphoinositol 4,5-bisphosphate (PI(4,5)P_2) to facilitate Ste5 membrane binding (*29*) and also includes the binding site for the MAPKKK Ste11 (*24*); and the VWA domain, which binds Ste7 and co-catalyzes Fus3 phosphorylation (*11*). The Ste7 binding site extends past the VWA domain into the C-terminal tail (*24*). We found that the PM-helix, RING-H2 domain, and FBD were dispensable for Ste5 autoinhibition. In contrast, truncation beyond the PH domain or internal deletion of the PH domain eliminated inhibition.

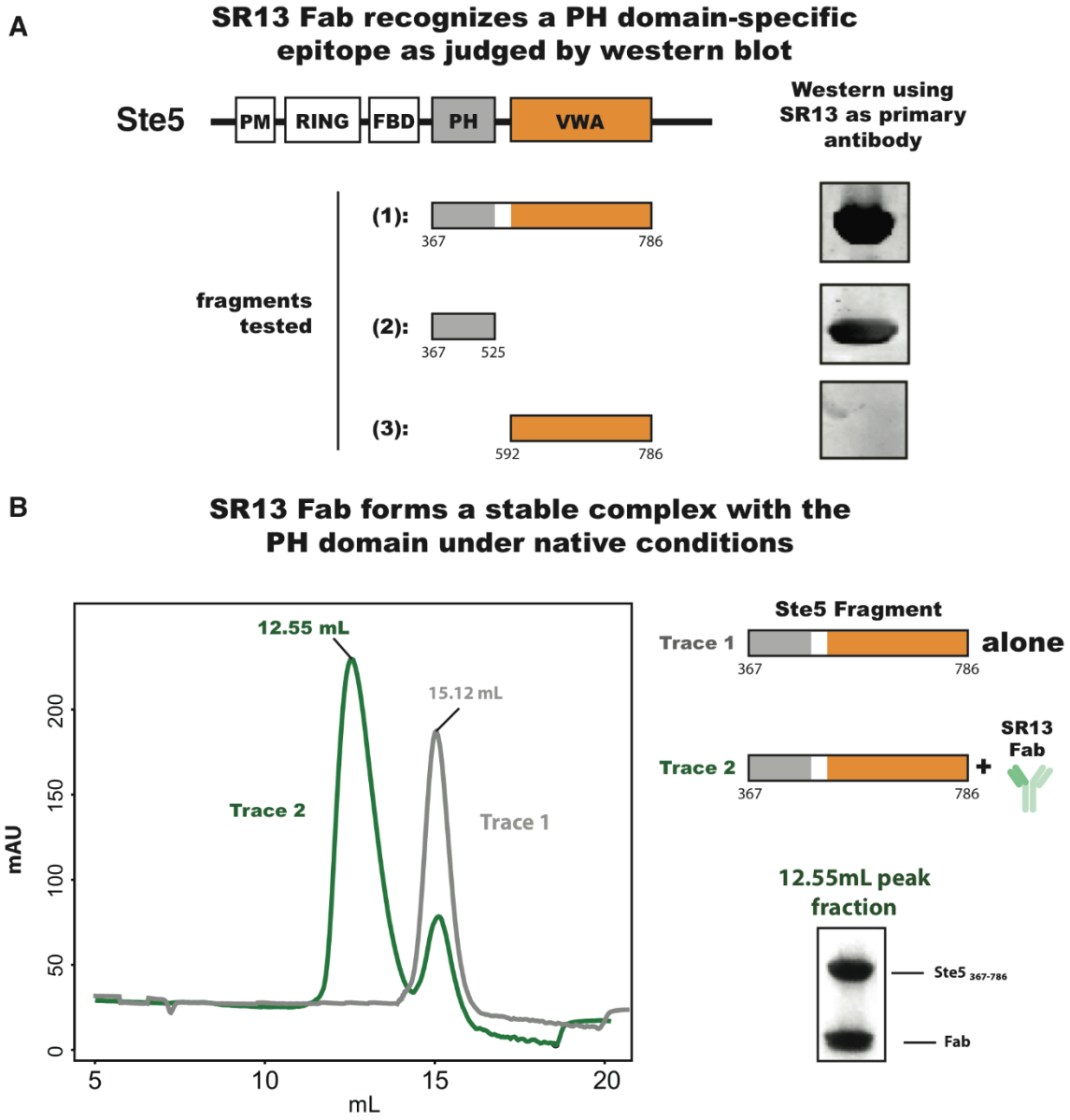
Truncation and internal deletions of Ste5 were purified to homogeneity and kinetic parameters for *in vitro* coactivator function were determined for each variant. A Ste5 variant was deemed autoinhibited if the observed k_{cat} (i.e. k_{obs} at saturating concentrations of Ste5 and Fus3) was within 2-fold of the full-length protein and deemed uninhibited if the observed k_{cat} was within 75% of the isolated VWA domain. Regions that result in loss of autoinhibition when disrupted in the context of the full-length protein are the PH/Ste11 binding domain (colored in gray), and an N-terminal extension of the VWA domain (colored in pink). The results suggest that the PH domain (367-525) and an N-terminal extension the VWA (544-592) are required to maintain Ste5 in an autoinhibited state.

Figure S10. The N-terminal extension of the Ste5 VWA domain displays a surface that is essential for strong binding of the PH domain to the VWA domain and inhibition of its Fus3 coactivator function



(A) The minimal region of the N-terminal extension (residues 544-592) required for *in trans* inhibition by the PH domain was mapped to residues 582-592. This construct contains the minimal N-terminal extension (582-592) and the minimal VWA domain (593-786). (B) 20 μM MBP-PH domain (367-525) was incubated with 20 μM of either VWA (593-786) or VWA+N-terminal extension (582-786) for 30 minutes at room-temperature before incubation with 15 μL of amylose resin at 4 °C for one hour. Samples were washed twice and eluted with 2X SDS loading buffer and analyzed by SDS-PAGE. The VWA domain runs as a doublet under these conditions. Only the VWA+N-terminal extension construct binds the PH domain. (C) The coactivating activity of $Ste5_{582-786}$ (VWA with an N-terminal extension) or a construct bearing mutations along the surface of the N-terminal extension was measured in the absence or presence of 20 μM PH domain. Mutations in the N-terminal extension disrupt inhibition by the PH domain.

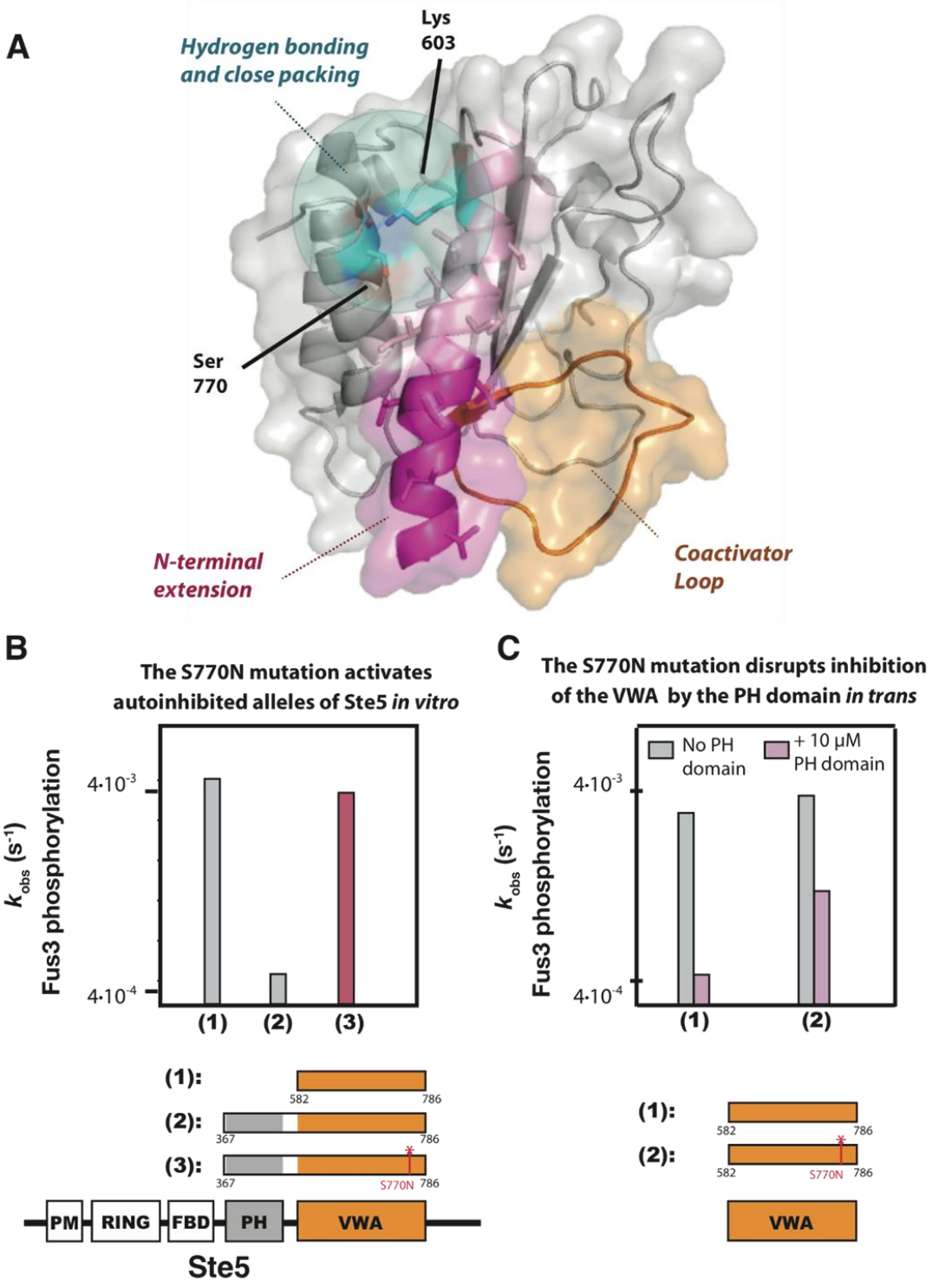
Figure S11. The SR13 Fab recognizes a PH domain-specific epitope and binds to the PH domain under native conditions



(A) Western blot using the SR13 Fab as primary antibody against a panel of Ste5 fragments. SR13 recognizes the PH domain-containing fragments (1 and 2), but does not recognize a Ste5 fragment that lacks the PH domain (fragment 3). (B) Size-exclusion chromatograms (Superdex 200 10/300; GE Healthcare) of $Ste5_{367\text{-}786}$ (grey) or $Ste5_{367\text{-}786}$ in the presence of the SR13 Fab. $Ste5_{367\text{-}786}$ elutes at 15.12 mL whereas in the presence of the Fab a substantial portion elutes in an

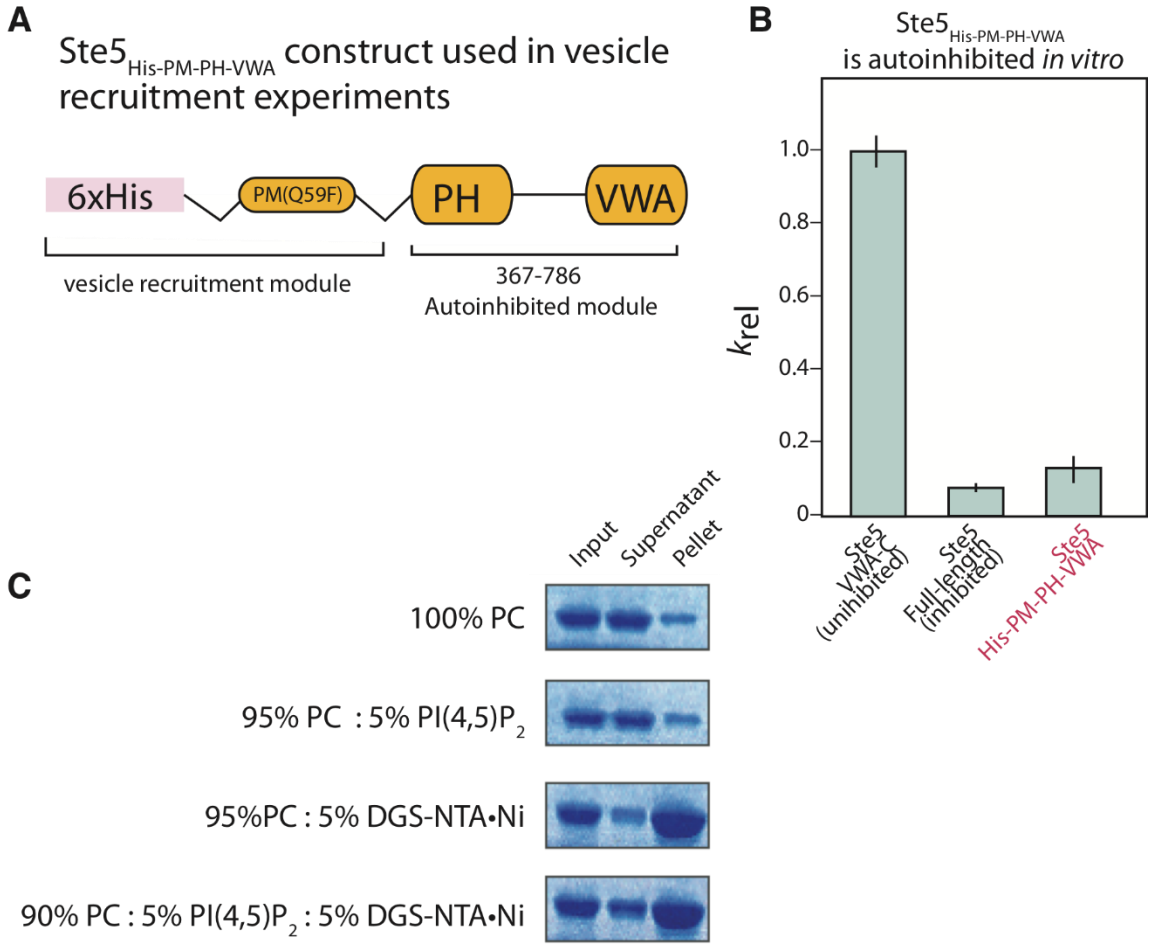
earlier, 12.55 mL peak. The peak fraction of the 12.55 mL peak was analyzed by SDS-PAGE and revealed approximately equivalent amounts of $Ste5_{367-786}$ and the Fab.

Figure S12. The "hyperactive" S770N Ste5 mutation disrupts autoinhibition of the Fus3 coactivator function of Ste5



Our model for inhibition is consistent with and suggests an explanation for the phenotype of a previously identified allele of Ste5 (S770N) that activates Fus3 in the absence of pheromone (*34, 35*). Ser770 lies on the same face of the VWA as the N-terminal helix, packs up against it, and orients its backbone carbonyl to make a hydrogen bond with Lys603. Disruption of this interaction in the S770N mutant likely destabilizes the binding site for the PH domain, and prevents Ste5 from adopting an autoinhibited conformation. Consistent with this prediction, we find that the S770N allele of Ste5 displays uninhibited coactivator activity *in vitro* and, furthermore, is inhibited less potently by the PH domain *in trans*. (A) Structure of $Ste5_{582\text{-}786}$ highlighting the position of serine 770 and lysine 603 in blue, the entire N-terminal helix in pink, and the N-terminal extension required for high-affinity PH domain binding in magenta. The backbone carbonyl of serine 770 forms a hydrogen bond with lysine 603, and the side chain of serine 770 is closely packed with the N-terminal helix. (B) Coactivator activity of $Ste5_{367\text{-}786}$ S770N at saturating [Ste5] (1 μM). Wild type $Ste5_{367\text{-}786}$ is autoinhibited. The activity of $Ste5_{367\text{-}786}$ S770N is similar to that of the isolated VWA domain. (C) Coactivator activities of either wild type $Ste5_{582\text{-}786}$ or $Ste5_{582\text{-}786}$ (S770N) in the absence (grey) or presence (pink) of 10 μM PH domain inhibitor.

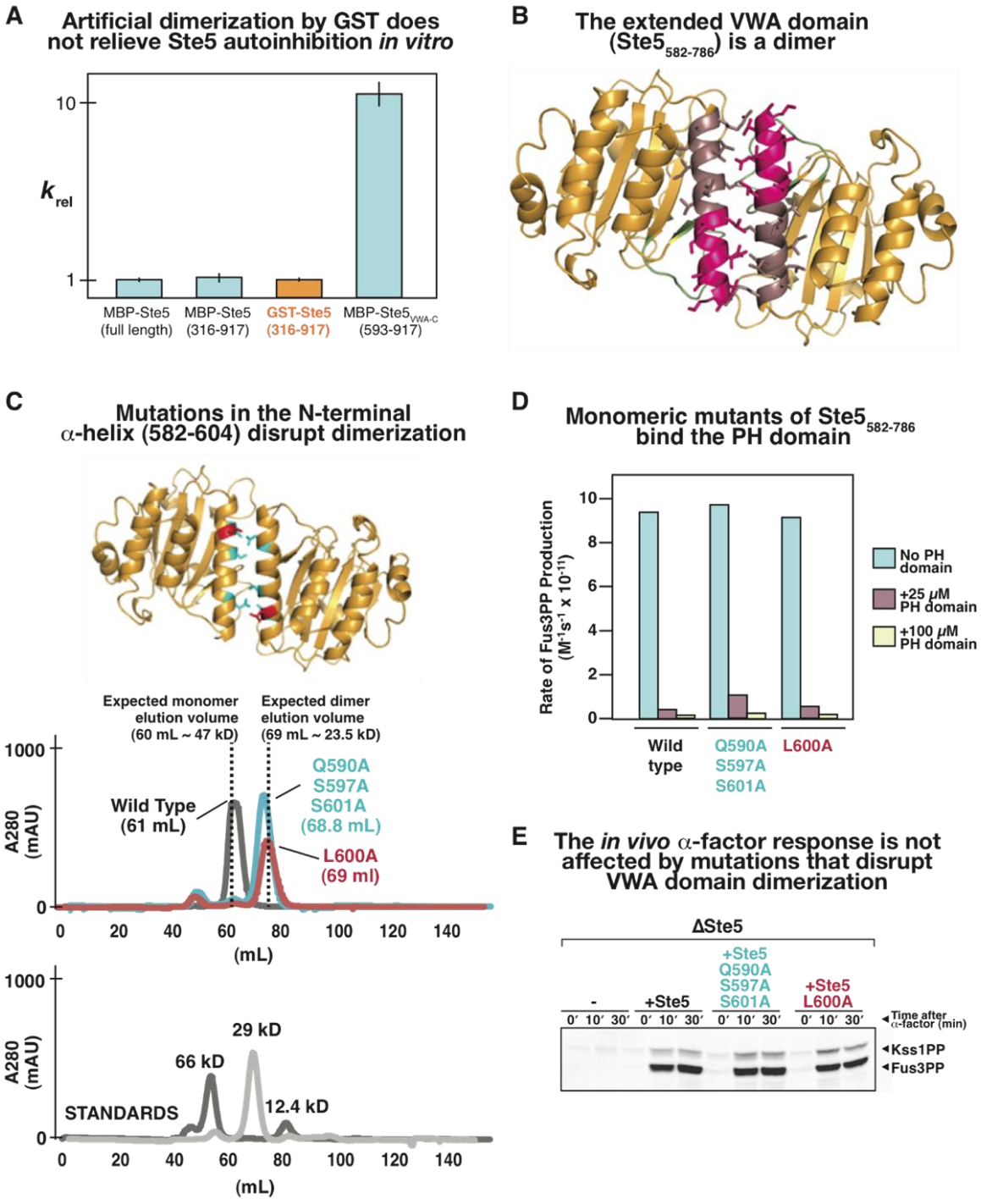
Figure S13. A minimal, autoinhibited membrane-binding Ste5 construct ($Ste5_{His\text{-}PM\text{-}PH\text{-}VWA}$) binds to PC:PI(4,5)$P_2$:DGS-NTA(Ni) vesicles



(A) Schematic of the minimal, membrane-binding $Ste5_{His\text{-}PM\text{-}PH\text{-}VWA}$ construct used in vesicle binding and activation experiments. The construct contains a 6x-His tag to allow binding to DGS-NTA(Ni) lipids and a mutant form of the PM helix (Q59F; (*27*)) that raises the pI of the protein and improves binding to vesicles under assay conditions. (B) The $Ste5_{His\text{-}PM\text{-}PH\text{-}VWA}$ construct is autoinhibited *in vitro*, similar to full-length Ste5. k_{rel} is the normalized value of k_{cat} (measured at saturating concentrations of Ste5 and Fus3). (C) Binding of $Ste5_{His\text{-}PM\text{-}PH\text{-}VWA}$ to vesicles of varying $PI(4,5)P_2$ and DGS-NTA(Ni) composition. 10 μM protein was incubated with 400 μM vesicle for 1 hour at 4 °C and centrifuged for 1 hour at 51K RPM in a TLA-100 rotor.

The amount of $Ste5_{His\text{-}PM\text{-}PH\text{-}VWA}$ in the input, supernatant, and pellet is shown as judged by Coomassie staining. No significant binding to PC or PC:PI(4,5)P_2 vesicles is observed, but significant binding is observed for vesicles with DGS-NTA(Ni). These vesicles recruit approximately half of the input protein under these conditions, which may explain why only partial activation of Ste5 is observed *in vitro* (Fig. 3G).

Figure S14. Oligomerization is not essential for relief of Ste5 autoinhibition



Dimerization of Ste5 has been suggested to be important for mating pathway activation *in vivo* (*28*, *34-37*), and this effect may act at the level of Ste11 activation (*28*, *34-36*), upstream of activation of the mating MAPK Fus3. Although we cannot exclude a role for Ste5 dimerization

in the relief of autoinhibition of the VWA domain, we find no evidence supporting this possibility.

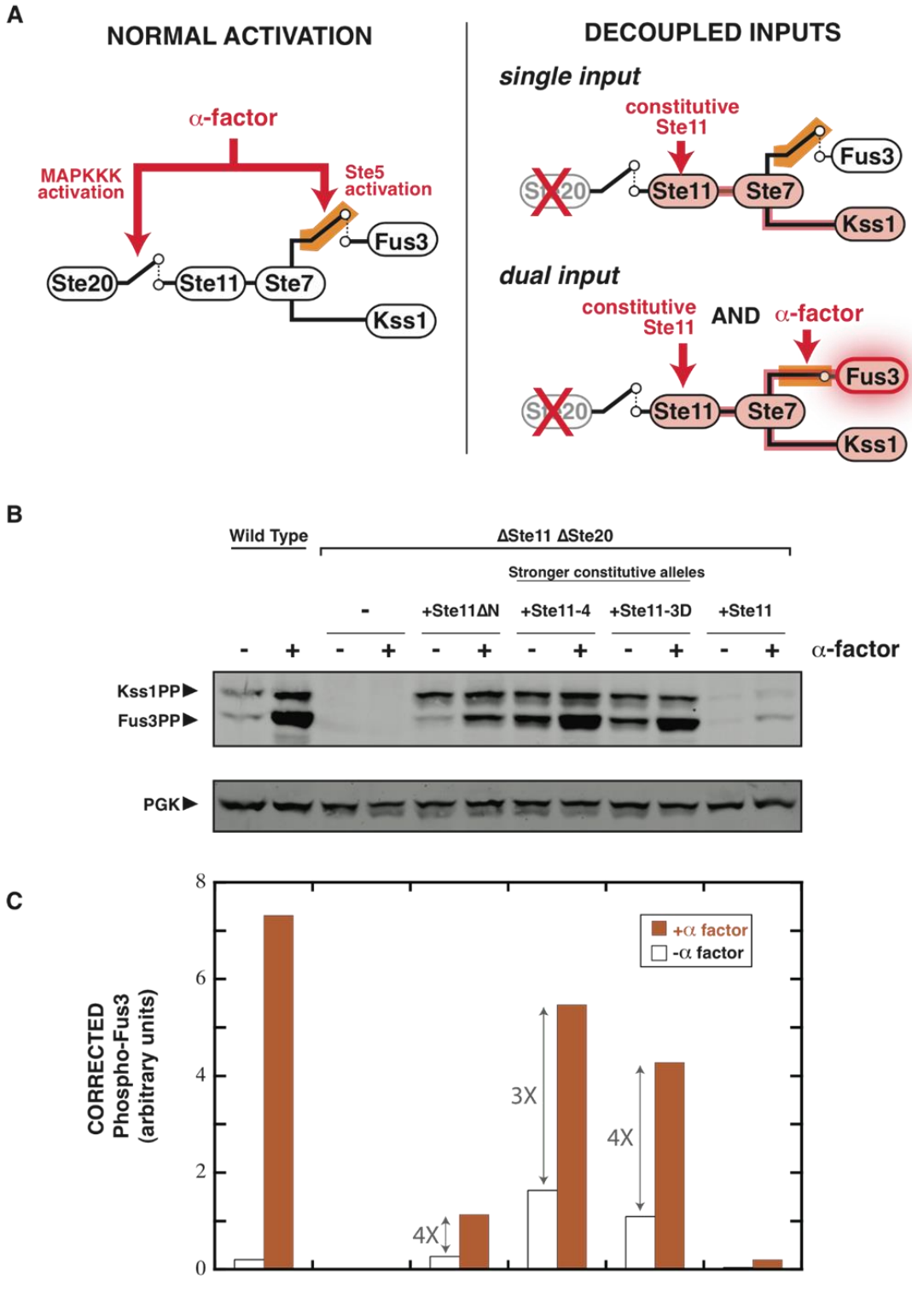
Key evidence supporting a role for dimerization of Ste5 in mating pathway activation comes from the observation that GST-tagged Ste5 leads to constitutive pathway activation in the absence of the mating input (α-factor) (*28, 34-36*). GST is known to form a dimer (*13, 14*). To test whether artificial dimerization of Ste5 with a GST tag could relieve autoinhibition of the VWA domain, we measured rate constants for a GST-tagged Ste5 construct *in vitro*, but no relief of autoinhibition was observed (Fig. S14A).

The extended Ste5 VWA domain ($Ste5_{582-786}$) is a dimer (Fig. S14B and C). Because the dimerization interface partially overlaps with features required for PH domain binding and inhibition (Fig. 3), dimerization along this interface could play a role in activating the protein. Specifically, dimerization may compete with autoinhibition by the PH domain and thereby stabilize the uninhibited conformation of Ste5. This model makes the prediction that if the dimerization interface is disrupted, Ste5 should remain inhibited even when cells are stimulated by the mating input. To test this possibility, we identified mutations that abolish dimerization (Fig. S14C) but were still subject to inhibition by the PH domain (Fig. S14D). We observed no significant changes in the mating response in yeast cells harboring these mutant versions of Ste5 (Fig. S14E), implying that dimerization along this interface is not essential for relief of Ste5 autoinhibition.

(A) The observed rate constant for Ste7EE-catalyzed phosphorylation of Fus3 with GST-$Ste5_{316-917}$ (we were unable to express full-length GST-Ste5) was indistinguishable from those for MBP-Ste5 or MBP-$Ste5_{316-917}$, constructs that are autoinhibited *in vitro* (Fig. 3A and S9). k_{cat} was measured at saturating concentrations of Ste5 and Fus3 and normalized to k_{cat} for MBP-

$Ste5_{316-917}$. (B) Structure of a novel crystallographic dimer of $Ste5_{582-786}$. The position of the N-terminal extension (residues 582-592) required for PH domain binding is colored in pink. The remaining portion of the α-helix that makes up the dimer interface (residues 593-604) is colored in burgundy. (C) $Ste5_{582-786}$ is a dimer in solution, and dimer-interface mutants are monomers based on size-exclusion chromatograms (Superdex 75 16/60; GE Healthcare) for wild type (black) or mutants (cyan or red) of $Ste5_{582-786}$. Chromatograms for three molecular weight standards on the same column are also shown, and the position that a monomeric or dimeric form of $Ste5_{582-786}$ would be expected to run is indicated. Mutations that disrupt dimerization are shown on the structure and color-coded as cyan (Q590A, S597A, S601A) or red (L600A). (D) Monomeric mutants of $Ste5_{582-786}$ retain the ability to bind to the PH domain. The PH domain of Ste5 inhibits the Fus3 co-activator function of both wild type and mutant forms of $Ste5_{582-786}$ *in trans*. (E) Phosphorylation of MAP kinases in response to α-factor treatment monitored by anti-phospho western blotting. There is no significant difference between wild type Ste5 and dimerization-defective mutants. Ste5 constructs were expressed from the Ste5 promoter.

Figure S15. Pheromone triggers activation of the Ste5 scaffold protein



(A) In wild type cells, α-factor treatment activates the MAPK cascade and the Ste5 scaffold protein. Decoupling the two functions of Ste5 can be accomplished by deleting the upstream kinase Ste20, which prevents activation of the MAPK cascade upon membrane recruitment of Ste5, and introducing a constitutively-active allele of Ste11. (B) Phosphorylation of MAP kinases in response to α-factor monitored by anti-phospho western blotting. There is no α-factor response in a ΔSte11ΔSte20 strain, and restoring wild type Ste11 has no significant effect. Introduction of Ste11ΔN, Ste11-4, or Ste11-3D, which are constitutively-active alleles of the MAPKKK Ste11 (*30, 38, 39*), leads to significant Fus3 activation. Ste11ΔN is an amino-terminal truncation that renders the kinase constitutively-active (*38*) and also lacks the Ste5 binding site (*40*). This allele is therefore the most informative for assessing whether α-factor treatment activates Ste5 for the Ste7→Fus3 reaction (rather than the Ste11→Ste7 reaction). Ste11-4 and Ste11-3D contain point mutants that activate the kinase (*30, 39*). Ste11ΔN was overexpressed from an Adh1 promoter to compensate for the relatively weak activity of this allele, and all other Ste11 constructs were expressed from the Ste11 promoter (Table S3). The stronger basal Fus3 activation observed with Ste11-4 and Ste11-3D is likely to be a consequence of the comparatively strong constitutive activity of these alleles relative to that of Ste11ΔN. (C) The change in Fus3PP levels +/- α-factor is similar for all three constitutive Ste11 alleles.

Table S1.

Kinetic constants for Fus3 phosphorylation reactions.[a]

	k_{cat} (s^{-1})	K_M (nM)	k_{cat}/K_M ($M^{-1}s^{-1}$)	K_{act} (nM)
$Ste5_{VWA-C}$	$(9.1 \pm 0.4) \times 10^{-3}$	62 ± 11	$(1.5 \pm 0.3) \times 10^{5}$	127 ± 32
Ste5 (full-length)	$(8.0 \pm 0.5) \times 10^{-4}$	263 ± 44	$(3.0 \pm 0.5) \times 10^{3}$	165 ± 21

[a] See Fig. 2 for data and Methods for kinetic models used to fit the data. The k_{cat} with $Ste5_{VWA-C}$ is 10-fold larger than with full-length Ste5, while K_{act} is unchanged. The K_M is approximately 4-fold weaker for full-length Ste5, but this is likely due to the presence of an additional binding site for the Fus3 substrate within the full-length Ste5 which can titrate out substrate (*15*, *23*). Standard errors are from non-linear least squares fits to the initial rate data.

Table S2.

Yeast strains used in this study.[a]

Strain	**Description**	**Genotype**
F1950	Σ1278b derivative	*MAT***a** *ura3 leu2 trp1 his3*
JZ003	F1950 ΔSte5	*MAT***a** *ste5::kan^R ura3 leu2 trp1 his3*
JZ011	F1950 ΔSte5 ΔSte2	*MAT***a** *ste5::kan^R ste2::nat^R ura3 leu2 trp1 his3*
JZ012	F1950 ΔSte5 ΔSte11	*MAT***a** *ste5::kan^R ste11::nat^R ura3 leu2 trp1 his3*
JZ018	F1950 ΔSte11 ΔSte20	*MAT***a** *ste11::nat^R Ste20::hph^R ura3 leu2 trp1 his3*
YM2369	Σ1278b derivative	*MATα his1*

[a] Strains F1950 and YM2369 were provided by H. Madhani (*1*).

Table S3.

Yeast expression plasmids used in this study.[a]

Plasmid	Parent Vector	Marker	Promoter[b]	Gene
pJZ525	pNH605	*leu2*	*pSte5*	Ste5
pJZ526	pNH605	*leu2*	*pSte5*	Ste5-ms (593-786)
pJZ550	pNH605	*leu2*	*pSte5*	$Ste5_{VWA-C}$ (593-917)
pJZ559[c]	pNH605	*leu2*	*pSte5*	Ste5Δ(544-592)
pJZ584	pNH605	*leu2*	*pAdh*	Ste5
pJZ586	pNH605	*leu2*	*pAdh*	$Ste5_{VWA-C}$ (593-917)
pJZ577[d]	pSV606	*ura3*	*pGal*	Ste7EE
pJZ527	pNH603	*his3*	*pSte11*	Ste11
pJZ528[d]	pNH603	*his3*	*pSte11*	Ste11-4
pJZ529[d]	pNH603	*his3*	*pSte11*	Ste11-3D
pJZ590[d]	pNH603	*his3*	*pAdh1*	Ste11ΔN
pSC179	pNH605	*leu2*	*pSte5*	$Ste5_{Q590A, S597A, S601A}$
pSC185	pNH605	*leu2*	*pSte5*	$Ste5_{L600A}$

[a] These vectors integrate a single copy into the yeast genome at the corresponding auxotrophic marker site (see Methods).

[b] Ste5 and Ste11 promoters consist of 500 bases immediately upstream of the start site for the corresponding gene. The Adh1 promoter consists of 1500 bases immediately upstream of the Adh1 gene. The Gal promoter consists of 689 bases immediately upstream of the Gal10 gene.

[c] The deleted residues in this Ste5 construct were replaced with a 5X(GAGS) linker.

[d] Constitutively active alleles of Ste7 (*11*, *12*) and Ste11 (*30*, *38*, *39*) have been described previously.

Table S4.

Crystallographic Statistics for $Ste5_{582-786}$ structure (PDB ID: 4F2H).

Data Statistics	**Ste5 582-786**
Space Group	I222
Unit Cell	63.84 87.35 100.10 90.000 90.000 90.000
Wavelength	1.5418 Å
Resolution (last shell)	19.7-3.19 Å (3.28-3.19 Å)
Unique Reflections	25304
Redundancy	5.75 (4.91)
Completeness	90% (90%)
I/ σ	25.4 (10.56)
R_{sym}	10.9% (23.6%)

Refinement Statistics	
Resolution Range	19.7-3.19 Å
Reflections Used Work (Test)	4395 (220)
R_{cryst}/R_{free}	18.3%/ 25.7%
Overall figure of merit	.81333
r.m.s.d. bonds/angles	.043 Å /1.241º
Average B factor	34.25 $Å^2$

Ramachandran Analysis (Preferred / Allowed)	93.12% (6.35%)

r.m.s.d. is the root-mean squared deviation from ideal geometry.

$R_{sym} = \Sigma_{hkl}\Sigma_i |I_{hkl,i} - \langle I_{hkl,i} \rangle| / \Sigma_{hkl}\Sigma_i |I_{hkl,i}|$

R_{cyst} and $R_{free} = \Sigma||F_{obs}| - |F_{calc}|| / \Sigma|F_{obs}|$. F_{obs} and F_{calc} are observed and calculated structure factors. R_{free} is calculated from a set of randomly chosen 5% of reflections, and R_{cyst} is calculated with the remaining 95% of reflections.

Chapter 3. Exploitation of Latent Protein Allostery Enables the Evolution of Novel and Divergent MAP Kinase Regulation.

Abstract.

Allosteric interactions provide precise spatiotemporal control over signaling proteins, but how allosteric activators and their targets co-evolve is poorly understood. Here, we trace the evolution of two allosteric activator motifs within the yeast scaffold protein Ste5 that specifically target the mating MAP kinase Fus3. One activator (Ste5-VWA) provides pathway insulation and dates to the divergence of Fus3 from its paralog, Kss1; a second activator (Ste5-FBD) that tunes mating behavior is, in contrast, not conserved in most lineages. Surprisingly, both Ste5 activator motifs could regulate MAP kinases that diverged from Fus3 prior to the emergence of Ste5, suggesting that Ste5 activators arose by exploiting latent regulatory features already present in the MAPK ancestor. The magnitude of this latent allosteric potential drifts widely among pre-Ste5 MAP kinases, providing a pool of hidden phenotypic diversity that, when revealed by new activators, could lead to functional divergence and the evolution of distinct signaling behaviors.

Main Text.

Introduction.

Eukaryotic signaling proteins display highly diverse and divergent allosteric regulation. Although any one genome might contain many evolutionarily related signaling molecules, such as protein kinases, individual family members usually display divergent substrate specificity and unique allosteric regulation by various partner proteins. By controlling when and where signaling proteins are activated, these allosteric regulatory interactions play a central role in

determining the specific "wiring" of the molecular networks that control cellular behavior (Fig. 1A).

Despite their importance, little is known about how these complex allosteric regulatory partnerships in signaling networks evolve. The molecular complexity of these systems represents a challenge for evolution: allosteric activators and the target proteins that they act on must seemingly acquire their complementary regulatory properties simultaneously for these systems to be functional and provide a selective advantage. These allosteric activators must also be specific enough to ensure that they do not inadvertently target homologous signaling components in the cell. The viable paths by which such multicomponent regulatory systems can evolve are therefore unclear.

In other complex systems, many new features appear to evolve by taking advantage of pre-existing or latent behavior: an active site that catalyzes a particular reaction can, with increased promiscuity, perform similar reactions on other substrates; a binding pocket that favors binding of one nuclear hormone can be adapted accommodate a yet-to-be-evolved hormone with somewhat similar structural features (*1*-*6*). While such latent capacities provide clear toeholds for new enzymatic activities or ligand binding capacities, these changes represent a shift in an already well-established and constitutive molecular activity. It is thus unclear the extent to which these evolutionary models apply to allosteric systems in which new protein partnerships must develop that are unrelated to any existing form of regulation and that must produce complex structural reorganization. Computational and protein engineering studies suggest that certain features of protein structure and dynamics may endow proteins with some latent capacity for allosteric regulation (*7*). Whether natural systems have harnessed such latent features to produce new allosteric regulation during evolution, however, has not been established.

Comparative studies that track the appearance of specific molecular properties across related species were instrumental in uncovering the role of latent protein features in the evolution of other systems and have provided great insights into how new enzymatic activities, receptor/ligand pairs, and transcriptional circuits evolve (*8-14*). However, applying these approaches to multi-component allosteric regulation of signaling proteins has been hindered by a lack of model systems that can be biochemically interrogated over species spanning a considerable window of evolutionary time.

The budding yeast MAP kinase network presents a unique model system with which to take a comparative approach to understand how complex multi-component allosteric regulation might have evolved. Prior biochemical studies have shown that, in *S. cerevisiae*, the function of the mating-pathway specific MAP kinase Fus3 requires its allosteric activation by the scaffold protein Ste5 (Fig. 1B). This scaffold-mediated allosteric activation ensures that Fus3 is only activated in signaling complexes that are organized in response to pheromone stimulation, thus preventing inappropriate cross-talk in which distinct MAP kinase mediated pathways trigger mating (*15*). Interestingly, the closely related starvation-responsive MAP kinase, Kss1, functions independent of Ste5 regulation, despite the fact that Fus3 and Kss1 are 55 % identical, are both targets of the MAPKK Ste7, and likely arose from duplication of the same Erk-like MAP kinase ancestor (Fig. 1C) (*16*).

Given their common MAPK ancestor, how did Fus3 become dependent on allosteric regulation while Kss1 did not? The availability of a large number of sequenced fungal genomes provides an opportunity to gain insights into this evolutionary question by exploring the regulatory properties of orthologs from scaffold and MAP kinase species throughout the fungal tree. Comparison of Erk-like MAP kinase sequences from across the Ascomycota fungi (to

which *S. cer.* belongs) indicates that these kinases are highly divergent and fall into distinct classes that are associated with specific fungal lineages (Fig. S1A-C). Interestingly, only those species that have both a Fus3 and Kss1 ortholog also have a Ste5 scaffold ortholog (Fig. 1D; detailed in supplement). It is unclear how both a potent allosteric activator (Ste5) and its regulated target (Fus3) could simultaneously evolve as a two-part complementary system. However, because we have access to signaling repertoires from species that clearly diverged from the *S. cer.* lineage prior to the appearance of Fus3, Kss1 and Ste5, we have the potential to uncover the mechanism by which this allosteric partnership evolved.

Here we expressed and purified Erk-like MAP kinases and Ste5 orthologs (if present) from 13 diverse fungal species that span from *S. pombe* to *S. cer.* (~1 billion years of divergence—comparable to the divergence between sea squirt and human). Using an *in vitro* reconstituted system, we determined the ability of these orthologs to cross-activate one another, even for species that do not contain a Ste5 protein. These quantitative data allowed us to determine when specific kinase and scaffold biochemical features arose during evolution and to formulate a model for the evolution of the allosteric regulatory schemes observed in *S. cer.*

First, we find that the Ste5 allosteric interaction required for Fus3 activation by the MAPKK Ste7 (Ste5-VWA) is a conserved scaffold feature of all Fus3/Kss1 containing species, while a second allosteric region in Ste5 (Ste5-FBD) that tunes the ultrasensitiviy of the mating response is, in general, not conserved outside of *S. cer*. This is consistent with a model in which a core function of the Ste5 scaffold protein has been to functionally insulate Fus3 and Kss1 since their divergence, but also suggests that Ste5/Fus3 interactions might continue to evolve to meet specific organismal needs. Second, and surprisingly, we find that the Ste5 scaffold can allosterically activate orthologous MAP kinases from species that diverged prior to the evolution

of Ste5, i.e. kinases that are likely to never have co-existed with the Ste5 scaffold. This result suggests that the Ste5 allosteric interactions evolved by tapping into latent, pre-existing dynamic properties of the MAP kinase. The magnitude of this latent allostery appears to drift significantly within the pre-Ste5 MAP kinases—some orthologs are primed for Fus3-like regulation (strong allosteric response) while others are primed for Kss1-like regulation (inability to respond). We propose that hidden diversity in these latent allosteric properties provides a toehold that new partner molecules can exploit to develop novel, component-specific allosteric regulatory relationships, simplifying the evolutionary paths to allosteric controls that shape pathway behavior and distinguish functional identity.

Results.

The *S.cer.* Ste5 scaffold protein allosterically activates the Fus3 MAP kinase via two mechanisms

In prior work we identified two modes by which the Ste5 scaffold protein allosterically activates the Fus3 MAP kinase in budding yeast *S. cer.* (Fig. 1B). The first allosteric interaction involves a Von-Wildebrand Type A (VWA) domain in the Ste5 scaffold protein that is required to allosterically unlock Fus3 to allow for its dual phosphorylation and activation by the upstream MAP kinase kinase (MAPKK), Ste7 (Fig. 2A). This VWA allosteric co-activation is essential for the transmission of the mating signal but has no influence on activation of the paralagous starvation-specific MAP kinase Kss1, which is also a substrate for the MAPKK Ste7 (*17*). In the resting Ste5 molecule, Ste5-VWA activity is autoinhibited by other domains in Ste5. This autoinhibition prevents Fus3 from being activated until mating inputs relieve this inhibition,

providing insulation from alternative inputs that activate the upstream MAPKK Ste7, such as starvation (*15*).

The second allosteric interaction involves a linear motif in Ste5 called the "Fus3 binding domain" (Ste5-FBD), which binds Fus3 and allosterically activates autophosphorylation of the MAP kinase on its activation loop tyrosine (Fig. 2B) (*18*). This partially activated form of Fus3 back-phosphorylates Ste5 to down-regulate mating pathway output and reshapes the morphological response of cells to α-factor ("shmooing") to be switch-like (ultrasensitive) instead of graded (*19*). The FBD allosteric activation is not essential for mating signaling, but instead appears to fine-tune the quantitative aspects of the mating response.

The Ste5-VWA allosteric interaction dates back to Fus3/ Kss1 divergence, while the Ste5-FBD allosteric interaction is a recent innovation that tunes mating behavior in a few specific lineages.

We first examined when Ste5-VWA allosteric activity appeared relative to the emergence of Fus3 and Kss1 kinases families. We purified Ste5-VWA domain orthologs from diverse fungal species that contain the Ste5 scaffold, and determined if they could allosterically co-activate *S. cer.* Fus3 phosphorylation by the *S. cer.* Ste7 MAPKK (henceforth, Ste7) (Fig. 2A and Fig. S2A). As observed previously, phosphorylation of *S. cer.* Fus3 by Ste7 is very slow in the absence of *S. cer.* Ste5-VWA (k_{cat} = 6.0±0.4 x10^{-7} s^{-1}) and the addition of saturating *S. cer.* Ste5-VWA stimulates this rate by greater than 3 orders of magnitude (6250 ± 610 fold). When saturating amounts of other Ste5-VWA orthologs were provided instead, rate enhancements were nearly identical to that of *S. cer.* Ste5-VWA. The most parsimonious interpretation of these data

is that the Ste5-VWA domain possessed potent allosteric activity towards Fus3 in the last common ancestor of these species (Fig. 2C). Consistent with this, chimeric *S. cer.* Ste5 molecules in which the native VWA domain was replaced with the VWA domain from other Ste5 orthologs were able to support robust mating *in vivo* (Fig. S2E).

Is the Ste5-VWA domain of other orthologs subject to autoinhibition, as in *S. cer.*? A simple diagnostic for Ste5 autoinhibition is that full-length Ste5 provides a smaller rate enhancement for Fus3 phosphorylation than the isolated VWA domain (*15*). Thus we compared the rate enhancement provided by the longest Ste5 construct we could express for each ortholog to that of the corresponding isolated VWA domain (Fig. S2A, S2C). As in *S. cer.*, every ortholog we examined was less effective than the isolated VWA domain in enhancing the Ste7•Fus3 reaction. The extent of this autoinhibition ranged from values comparable to the *S. cer.* inhibition (~10-fold) to values that were as much as 80-fold inhibited. Additional experiments indicate that the molecular mechanism of this inhibition is likely the same as in *S. cer.* Ste5 (Fig. S2D), and thus the simplest explanation for these data is that this mechanism to control Ste5 VWA allosteric activation was a conserved Ste5 feature present in the last common ancestor of these species.

We next examined the evolutionary history of the *S. cer.* Ste5-FBD allosteric regulatory interaction. Orthologous Ste5-FBD sequences (detailed in supplement) were purified and assayed for the ability to stimulate *S. cer.* Fus3 autophosphorylation (Fig. 2B and Fig. S3A-B). As observed previously, *S. cer.* Ste5-FBD potently stimulated the rate of *S. cer.* Fus3 autophosphorylation (149.7 ± 13.5 fold rate enhancement). In contrast, the FBD region from all but one Ste5 ortholog failed to provide a detectable rate enhancement for *S. cer.* Fus3 autophosphorylation. The one exception was the FBD sequence from *V. pol.*, which provided an

intermediate effect (12.8 ± 0.3 fold rate enhancement). One possible explanation for the lack of allosteric activity we observed for most Ste5-FBD sequences is that perhaps each FBD motif is optimized for its corresponding Fus3 ortholog. However, no differences were observed when the Fus3 ortholog from the same source species was used as a target instead of *S. cer.* Fus3 (Fig. S3D). Thus, unlike Ste5-VWA regulation, the Ste5-FBD regulation found in *S. cer.* is not conserved in every organism that contains Fus3, Kss1, and Ste5 (Fig. 2C).

Most likely, the FBD interaction evolved as a recent lineage specific feature to tune the mating behavior of *S. cer.* It has previously been shown that mutating the Ste5-FBD in *S.cer*. converts a switch-like (ultrasensitive) shmooing response to α-factor into a graded (linear) response (*19*). This model would predict that species lacking an active FBD motif would show a linear shmooing response. To test this model, we quantitatively examined the morphological responses of *K. lactis* – a species that lacks an active FBD motif but retains an active VWA domain (Fig. 2D). As predicted, the morphological dose response of *K. lactis* to α-factor was graded (n_H = 0.9 +/- 0.2) in comparison to the switch-like response observed in *S. cer.* (n_H = 7.7 +/- 0.7). The fact that *K. lactis* cultures must undergo prolonged phosphate starvation to be mating competent (*9, 20*) may complicate a direct comparison of these two profiles. Nonetheless, together with our biochemical analyses, these data suggest that the Ste5-FBD interaction arose well after the divergence of Fus3 and Kss1 as a mechanism to fine tune quantitative mating responses. We note, however, that we cannot definitively rule out repeated loss of the Ste5-FBD from multiple lineages as an alternative explanation of these data.

Together, the simplest evolutionary model for these data is that a potent but tightly regulated Ste5-VWA activity was present in the last common ancestor of the species that contain

both Fus3 and Kss1 MAP kinase types, while the Ste5-FBD activity was likely layered on top of the core conserved Ste5 activities to reshape the morphological response to mating pheromone in only certain species (Fig. 2E). This suggests that a core function of the Ste5 scaffold protein has been to functionally insulate Fus3 and Kss1 since their divergence but also suggests that Ste5 scaffold interactions with the Fus3 kinase might continue to evolve to meet specific organismal signaling needs.

Latent Allostery: Ste5 allosteric activator domains can stimulate MAP kinases that diverged prior to the evolution Ste5

We then turned to the converse question of understanding how the Fus3 MAPK acquired the necessary features to serve as a target of these two Ste5 allosteric interactions. Here we reversed our *in vivo* cross-reaction components and tested the extent to which Fus3 and Kss1 orthologs from other species could be regulated by the *S. cer.* Ste5 scaffold activities. Starting with the *S. cer.* Ste5-VWA domain (Fig. 3A), we found that all Fus3 orthologs were strongly allosterically regulated by the VWA domain: they were poor substrates for Ste7 in the absence of *S. cer.* Ste5-VWA ($k_{cat} < 5x10^{-6}$ s^{-1}) but the addition of *S. cer.* Ste5-VWA enhanced phosphorylation of each MAPK by greater than 2000 fold. This strong allosteric activation was identical when other Ste5-VWA orthologs were used in place of *S. cer.* Ste5-VWA (Fig. S2B).

In contrast, all of the Kss1 orthologs we tested were not targets for VWA activation — these MAPKs were ideal substrates for Ste7 ($k_{cat} > 1x10^{-3}s^{-1}$) in the absence of any other molecules, and were unaffected by the addition of *S. cer.* Ste5-VWA (rate-enhancement < 1.5 fold). From these data, we infer that Fus3 and Kss1 likely possessed their divergent responses to

Ste5-VWA regulation in the last common ancestor of the species that contain these kinases. The ability of Fus3 orthologs to be activated by the VWA domain, thus, appears to be tightly conserved after the functional divergence of the Fus3 and Kss1 MAPKs.

We then tested whether Erk-like kinases from species that diverged from *S. cer*. prior to the evolution of Ste5—henceforth referred to as "pre-Ste5" Erk-like kinases— had the capacity to be regulated by the modern *S.cer* Ste5 VWA domain (Fig. 3A). Unlike *S.cer* Fus3, these pre-Ste5 kinases were intrinsically good substrates for Ste7 catalyzed phosphorylation *in vitro* (k_{cat} $>7x10^{-5}$ s^{-1}). However, addition of *S. cer.* Ste5-VWA surprisingly stimulated phosphorylation of many of these kinases by as much as a 42-fold rate enhancement. Thus, these pre-Ste5 MAP kinases are similar to Fus3 in that they have a modest capability to serve as a target for allosterically activation by the Ste5 VWA domain, despite the fact that the species from which they come lack Ste5.

We then analogously examined when the ability to serve as a target for the Ste5-FBD interaction arose within the MAP kinase family (Fig. 3B). Although only the *S. cer* Ste5 ortholog possessed potent Ste5-FBD activity, we surprisingly found that the Fus3 orthologs from nearly every species that we examined were targets for FBD activation – like *S.cer* Fus3, they all displayed a FBD enhanced rate autophosphorylation of greater than 100-fold. In contrast, *S. cer.* Ste5-FBD did not significantly enhance the rate of autophosphorylation of the Kss1 orthologs we tested. We conclude that Fus3 was primed for regulation by a *S. cer.* Ste5-FBD mechanism in the common ancestor of these species, even before the FBD activity had evolved in the Ste5 scaffold; little change to the kinase was necessary for the *S. cer.* Ste5-FBD to be able to influence the rate of autophosphorylation.

We then tested whether *S. cer.* Ste5 FBD could enhance the rate of autophosphorylation of the pre-Ste5 Erk-like kinases (Fig. 3B). We observed a broad range of capacities for regulation by *S. cer.* Ste5-FBD. Several kinases were not allosterically affected by the *S. cer.* Ste5-FBD (*D.han.* Cek2, *C. alb.* Cek1, *A. nid.* MpkB, *N. cra.* MpkB) even though these kinases readily bound to *S. cer.* Ste5-FBD (Fig. S4A). Some kinases, however, showed intermediate effects (*S. pom.* Spk1, *C. alb. Cek2*); and still others showed allosteric responses that approached or even exceeded the enhancement in autophosphorylation that is seen for *S. cer.* Fus3 (*C. tro.* Cek2, *L. elo.* Cek2). Thus, many of the pre-Ste5 MAPKs display the ability to serve as a target for both VWA and FBD mediated allosteric activation.

These findings suggest that both the VWA and FBD allosteric interactions evolved by tapping into *latent allosteric features* that pre-existed within this family of kinases. Because the pre-Ste5 Erk-like kinases — including the Spk1 kinase from *S. pombe*, which is the most distantly related to *S. cer.* Fus3 — broadly show modest regulation by both Ste5-VWA and Ste5-FBD, the most parsimonious explanation of these data is that some capacity for both of these forms of allosteric regulation was likely present in the ancestral kinase of all the orthologs we inspected (Fig. 3C). Although it is formally possible that there exist other alternative allosteric regulators that capitalize on these modest features in the pre-Ste5 lineages, several lines of reasoning argue against this. First, we tested several likely candidate proteins present in these organisms for such activity and found no evidence in support of this (Fig. S4B, Fig. S4D-E). Second, the extensive variation in the latent allosteric features of the pre-Ste5 MAPKs—including the absence of these features in particular orthologs—suggests that these features are not under selective pressure. That is, these particular allosteric regulations of the MAPK substrate have not been fixed in all of the pre-Ste5 branches of the Ascomycota to the extent that

they have been fixed in the post-Ste5 species, casting doubt on the existence of other critical allosteric regulators that are using the latent allosteric features. Third, given that the Ste5-VWA regulation is functionally required for pathway specificity—i.e. discriminating between the Fus3 and Kss1 kinases—it is unclear why such allosteric effectors would exist in lineages that contain only a single Erk-like kinase. Finally, the observed allosteric effects on the pre-Ste5 kinases are in most cases relatively small — all of the kinases were adequate MAPKK substrates in the absence of any additional Ste5 regulation. Thus, it is unlikely that these species functionally *require* such allosteric effectors. As such, we favor a model in which the capacity for the allosteric regulation we observed was already present in the ancestral kinases, providing a toe-hold for the emergence of new forms allosteric regulation.

Drift in latent allostery produces evolutionary related kinases that are primed for divergent responses to new allosteric activators.

A model of latent allostery within the MAP kinase family provides a simple framework for how new allosteric regulators such as the Ste5-FBD and Ste5-VWA domain may have evolved. However, it also raises an important question in terms of divergent regulation: how then is it that Fus3 orthologs are targets for this allosteric regulation, while Kss1 orthologs are not?

To gain insights into this question, we examined the diversity in the distribution of properties observed for the pre-Ste5 Erk-like kinases (Fig. 4A). When only considered as substrates for the MAPKK Ste7, pre-Ste5 Erk-like kinases generally cluster together and appear to be similar quality substrates in the absence of any scaffold coactivator. However, the latent capacity for allosteric regulation in each of these substrates results in additional dimensions of

MAPK phenotypic diversity beyond their basic properties as substrates for phosphorylation by the MAPKK Ste7. This diversity is easily visualized by plotting each of the kinases on phenotypic morphospace plots in which one dimension is the rate of Ste7→Fus3 phosphorylation—the apparent kinase diversity in the absence of any allosteric activators—and a second dimension is the allosteric enhancement of either of the two Ste5 allosteric interactions—the hidden phenotypic diversity that is only revealed upon interaction with scaffold effectors (Fig. 4A—see also Fig. S4C). For both Ste5-VWA and Ste5-FBD activities, the highly divergent regulation of Fus3 and Kss1 orthologs places them in opposite regions of this space, while most of the pre-Ste5 Erk-like kinases are 'hybrids' that, as a set, occupy a region of space in between Fus3 and Kss1. Importantly, these plots reveal that kinases that may appear close together in the one-dimensional perspective as substrates for MAPKK phosphorylation can be far apart along these hidden allosteric dimensions. Thus, drift in these hidden phenotypic properties (latent allostery) results in a distribution of family members, with some much closer to Fus3 in behavior, and others much closer to Kss1.

These findings suggest a simple and general mechanism for the evolution of novel and divergent allosteric regulation of paralogous signaling components such as the Kss1 and Fus3 MAP kinases (Fig. 4B and Fig. 4C). Neutral drift in a latent capacity for allosteric regulation produces paralagous variants that are primed for divergent responses to regulation. Appearance of a new interaction partner with weak activity against this latent allosteric feature 'reveals' the pre-existing diversity and provides a toehold for Darwinian processes to exploit these differences and drive these kinases into divergent regulatory modes by selection, as was observed for the divergent responses of Fus3 and Kss1 to the Ste5-VWA domain. Such selection events have the potential to fix other latent allosteric properties within the newly selected lineage owing to

founder-effects or hitch-hiking, which could explain why all of the Fus3 and Kss1 orthologs we tested also display divergent responses to Ste5-FBD regulation even before this activity evolved.

Dissection of the *V.pol.* Ste5-FBD with intermediate allosteric activity reveals alternative paths for coopting the same latent regulatory features

Our data demonstrate that a significant capacity for allosteric regulation is present in kinases prior to the evolution of the effectors that provide that regulation in *S. cer*. How, at a molecular level, does evolution discover activators that can tap into these hidden allosteric features and coopt this pre-existing capacity for new regulation? Our biochemical screen of Ste5-FBD motifs, identified a sequence with intermediate allosteric activity from *V. polyspora* (*V. pol.* Ste5-FBD) that gives us an opportunity to biochemically dissect how this FBD-mediated allosteric activity may have arisen (Fig. 5A) (Addressing this question for Ste5-VWA domain regulation is difficult because of the lack of any forms of Ste5 that show intermediate VWA activities).

We wanted to determine whether the *V.pol.* Ste5-FBD functions through a related mechanism to that used by the *S. cer.* Ste5-FBD. We previously showed that the *S. cer.* Ste5-FBD sequence binds to Fus3 in a bipartite manner to allosterically activate the Fus3 kinase: an "A-site" motif binds to the N-lobe of the kinase, while a second "B-site" motif binds to a canonical docking groove on the C-lobe of the MAP kinase (albeit in a non-canonical reverse C-to-N terminal orientation); linking these two binding sites is thought to constrain the two kinase domain into a more active conformation that promotes autophosphorylation (Fig. 5C) (Bhattacharyya et al., 2006). Inspection of the *V. pol.* Ste5-FBD sequence reveals a sequence that

resembles the "A-site" motif of *S. cer.* Ste5-FBD, but there is no obvious sequence that resembles the "B-site" motif.

To better understand the mechanism of the *V. pol.* Ste5-FBD interaction, we used deletion analysis to map the regions of this sequence that were required for its allosteric activity (Fig. 5B and Fig. S5A). Like *S. cer.* Ste5-FBD, we found that two distinct regions were required for activity. One of these regions contained the motif that resembles the "A-site" of the *S. cer.* Ste5-FBD, suggesting that both *V. pol.* Ste5-FBD and *S. cer.* Ste5-FBD use this "A-site" sequence to engage the N-lobe of Fus3. Unlike in the *S. cer.* Ste5-FBD, however, the second region of *V. pol.* Ste5-FBD required for allosteric activity was on the opposite side of the "A site" (N-terminal to it, i.e. the opposite orientation relative to the *S. cer.* Ste5-FBD). This second required region in the *V. pol.* Ste5-FBD motif fits the consensus MAPK docking motif ($[R/K]_{1-2}$-X_{2-6}-Φ–x-Φ-x-Φ) that is used by many signaling partners to interact with MAPKs (*21*,*22*). Consistent with this, mutation of the residues within this motif that would disrupt a MAPK docking interaction completely abolished the allosteric activity of *V. pol.* Ste5-FBD (Fig. 5B).

From these data, we infer a model for how *V. pol.* Ste5-FBD interacts with Fus3 to exert allosteric influence (Fig. 5C). At low resolution, both the *S. cer.* and *V. pol.* FBD mechanisms appear very similar – they both bind at the same two sites on the MAPK, potentially constraining the kinase N and C lobes relative to one another in a manner than increases autophosphorylation. Nonetheless, while the *S. cer.* Ste5-FBD binds Fus3 with an "A-site"-"B-site" bipartite polypeptide, the *V. pol.* Ste5-FBD appears to binds to Fus3 with a "docking motif"- "A-site" bipartite polypeptide (where the docking motif functionally replaces the B-site motif). In both cases, functionally analogous motifs that bind the C-lobe docking groove, cooperate with binding of the "A-site" motif to the N-lobe of the kinase to achieve allosteric activation. We postulate

that this distinct but analogous bipartite binding represents a case of convergent evolution – both bipartite peptides can constrain the kinase lobes required to stimulate autophosphorylation, albeit to different degrees (Fig. S5B).

Is the *V. pol.* Ste5-FBD motif, despite its detailed differences, tapping into the same latent allosteric features present in the fungal MAPK family as those exploited by the *S.cer.* Ste5-FBD? If so, then we predict that the effects of the *V. pol.* Ste5-FBD motif on diverse members of the Erk-like fungal kinase family should mirror those observed for the *S.cer* Ste5-FBD motif. Indeed, we observe a linear relationship across fungal species between the degree to which the *V. pol.* Ste5-FBD and the *S.cer* Ste5-FBD motifs can activate individual MAPK family members (Fig. 5E; see also Fig. S3D). Thus, even this weak activator appears to reveal the same latent potential for allosteric regulation that is present in many fungal MAP kinases, including the pre-Ste5 Erk-like kinases. During the course of evolution, once a weak effector like the *V. pol.* Ste5-FBD uncovers these latent kinase regulatory features, Darwinian processes can proceed to optimize this allosteric regulation. Because different yeast species occupy distinct environments and exhibit different mating preferences (*9*), the outcomes of these Darwinian processes will differ depending on local selective pressures and organismal niche: the activity can be optimized to increase potency (as in *S. cer.* Ste5-FBD), it can be maintained as a weak effector (as in *V. pol.* Ste5-FBD), or it can be turned over to a state in which Ste5-FBD regulation is lost (as observed in *C. gla.*—see Fig. 2B).

Discussion.

Our analysis of kinase and scaffold properties from across Ascomycota fungi allowed us to determine when particular Ste5 scaffold allosteric activator functions most likely arose, as well as when the capacity of a MAP kinase to serve as a target for such allosteric regulation arose. From this analysis, we made the surprising observation that many kinases from species that diverged from *S. cer*. prior to the evolution of the Ste5 scaffold can still be regulated by the allosteric motifs within Ste5. These findings suggest that a latent capacity for allosteric regulation was present within this MAP kinase family long before the evolution of effectors that target this allostery for regulation.

Exploitation of an existing latent capacity to derive a new molecular regulatory relationship is similar to proposed models for the evolution of new catalytic activities by catalytic promiscuity (*5,11*) and new hormone receptor signaling responses by molecular exploitation (*3*). In this case, however, the latent allosteric kinase features we have described in this study are not obviously similar to some pre-existing regulatory interaction, but represent new regulatory connections that can redirect and reshape information flow in cell signaling pathways.

The dynamic protein kinase structure as a source of latent and diverse allosteric behaviors.

We postulate that the dynamic nature of the protein kinase structure itself may provide the latent allosteric potential and diversity observed in these fungal MAP kinases. Indeed, both Ste5-VWA domain and Ste5-FBD motif are thought to allosterically activate the Fus3 MAP kinase by altering the kinase flexibility (*17,18*), and such flexibility is an innate but variable feature of protein kinases (*23*). Thus, rather than requiring that evolution create unprecedented structural features to produce allosteric innovation, the ruggedness of the MAP kinase conformational

landscape itself may provide toeholds for weak, but specific, activators that act by selecting and stabilizing particular kinase conformations (*24*) (Fig. 4C). These relationships can then be strengthened by evolutionary mechanisms that tune and modulate the stability of certain states of the molecule or widen the difference in activity between alternative states. These findings are consistent with patterns observed in many other members of the protein kinase family: all kinases appear to require the proper assembly of the same core catalytic and structural elements in order to adopt an active state; but different kinases adopt a wide array of distinct inactive conformations, each of which requires a different set of inputs to stabilize the conserved active conformation (*23*,*25*). More generally, modes of flexibility intrinsic to particular protein folds may provide the starting point for future regulatory evolution (*27*).

The hidden regulatory diversity that we find in the fungal MAP kinases may be a more general feature of many protein kinases as well as other dynamically regulated macromolecules. In fact, many drugs may act as effectors that uncover this regulatory potential. Indeed, such hidden conformational toe-holds serve as the basis of action of the Abl-specific tyrosine kinase inhibitor Gleevac, which stabilizes a inactive conformation that is uniquely accessible to that kinase (*26*). Similarly some small molecules have been found to allosterically activate regulatory proteins, despite the lack of a clear physiologic analog that normally targets that site (*27*).

Co-localization may facilitate the evolution of new allosteric regulation.

The latent allosteric properties in MAP kinases we have described must be 'revealed' by an effector in order for selection to be possible. How do such primitive allosteric regulators evolve?

The mechanism of the Ste5 VWA and FBD motifs suggests that co-localization may facilitate this process, as both interactions involve the interplay between allosteric interactions and co-localization interactions. The VWA domain allosterically regulates Fus3 in the context of a higher order molecular complex (a Ste5•Fus3•Ste7 ternary complex) that is assembled by non-allosteric co-localization interactions that are sufficient for tight complex formation (Fig. 6A). Similarly, binding of the Ste5-FBD A-site motif to the Fus3 kinase depends on a second-site interaction with the docking groove of MAP kinase which is sufficient for complex formation on its own (Fig. 6B). A simple model is that co-localization of the future allosteric target and regulator was an early step in the evolution of these allosteric relationships. Such co-localization establishes effective concentrations of the components in the milimolar range in which fleeting and weak interactions occur more readily (*7*,*28*), thus enhancing the likelihood of uncovering a weak interaction that reveals a latent allosteric feature in a target.

This evolutionary 'co-localization first' strategy is similar to the novel 'tethering' approach used for developing small molecule allosteric effectors, in which a library of disulfide-containing small molecules is localized to a particular cysteine residue on the drug target (29) allowing for the identification of weak effectors that bind to surprising new allosteric protein sites (27). The synergy between evolutionary and engineering approaches suggests that genetically encoded libraries that 'tether' a variable protein or RNA library to a target might offer an effective *in vivo* screening approach for identifying new allosteric effectors.

Finally, these observations indicate that the co-localization of signaling components on scaffolds or at the membrane may play a more active role in the evolution of new signaling pathways and behaviors than previously appreciated, by producing local environments in which hidden allosteric diversity and the ruggedness of conformational landscapes are revealed by high

effective concentrations and potential new effector interactions. Indeed, many primitive signaling pathways may have initially simply consisted of components that became co-localized upon stimulation with an input signal (Fig. 6C). These assemblies, however, might then provide a context that would facilitate the evolution of allosteric regulation, as described here, that yielded the diverse forms of precision control that we observe in modern pathways. An analogous progression of regulatory evolution is suggested to take place among DNA binding factors that are tethered at a promoter (*2*,*30*).

Methods.

Identification, sequence analysis, and cloning of kinase and scaffold orthologs

S. cer. Fus3, Kss1 and Ste5 sequences were used to query the fungal orthogroups database to identify orthologous sequences in the Ascomycota (additional details in supplemental methods) which were subsequently cloned from gDNA or synthesized directly. Phylogenetic analysis of these sequences is detailed in supplementary materials. A complete list of all constructs used in this study is in Table S1.

Protein purification

MAP kinases, Ste5 scaffold fragments, and the SR13 Fab antibody were expressed in BL21(T1R) *E. coli* cells. The *S. cer.* Kss1 ortholog and the constitutively active form of the MEK Ste7 (Ste7EE) were expressed from *S. frugiperda* (SF9) cells. Proteins were purified similarly as described previously (*15, 17, 21*) with minor modification as detailed in supplemental methods.

In vitro kinase activity assays

Initial rates for Ste7 catalyzed phosphorylation of a MAPK as well as MAPK autophosphorylation were measured by quantitative western blotting as described and detailed in supplemental methods. Under saturating conditions, VWA reactions contained 50 nM of MBP-Ste7EE, 5 µM MAPK substrate, and (if present) 5 µM Ste5-VWA ortholog; saturating FBD reactions contained 10 µM MAPK and, if present 25 µM of a Ste5-FBD sequence.

Morphological dose response to α-factor

Morpohological responses to α -factor were performed for *S. cer.* (strain W303) and *K. lac.* (strain yLB17a (*9*)) as described previously (*19*). For *K. lac.*, the response was measured after 6 hours of growth in SCD media lacking phosphate to ensure cells were mating competent (detailed in supplement). The percentage of cells shmooing at a given concentration of pheromone was determined by microscopy and the resulting dose response curves were fit to a Hill-equation to extract the hill-coefficient parameter n_H.

Main Text References.

1. Aharoni, A., Gaidukov, L., Khersonsky, O., Gould, S.M., Roodveldt, C., and Tawfik, D.S. (2005). The 'evolvability' of promiscuous protein functions. Nature Genetics *37*, 73-76

2. Baker, C.R., Booth, L.N., Sorrells, T.R., and Johnson, A.D. (2012). Protein Modularity, Cooperative Binding, and Hybrid Regulatory States Underlie Transcriptional Network Diversification. Cell *151*, 80-95

3. Bridgham, J.T., Carroll, S.M., and Thornton, J.W. (2006). Evolution of Hormone-Receptor Complexity by Molecular Exploitation. Science *312*, 97-101

4. O'Brien, P.J., Herschlag, D. (1999). Functional Interrelationships in the Alkaline Phosphatase Superfamily: Phosphodiesterase Activity of Escherichia coli Alkaline Phosphatase. Biochemistry *40*, 5691-5699

5. Khersonsky O., and Tawfik D.S. (2010). Enzyme Promiscuity: A Mechanistic and Evolutionary Perspective. Annual Review of Biochemistry *79*, 471-505

6. Wise, E.L., Yew, W.S., Akana, J., Gerlt, J.A., and Rayment, I. (2005). Evolution of enzymatic activities in the orotidine 5'-monophosphate decarboxylase suprafamily: structural basis for catalytic promiscuity in wild-type and designed mutants of 3-keto-L-gulonate 6-phosphate decarboxylase. Biochemistry *44*, 1816-1823

7. Reynolds, K.A., McLaughlin, R.N., and Ranganathan, R. (2011). Hot Spots for Allosteric Regulation on Protein Surfaces. Cell *147*, 1564-1575

8. Afriat, L., Roodveldt, C., Manco, G., and Tawfik, D.S. (2006). The Latent Promiscuity of Newly Identified Microbial Lactonases Is Linked to a Recently Diverged Phosphotriesterase. Biochemistry *45*, 13677-13686

9. Booth, L.N., Tuch, B.B., and Johnson, A.D. (2010). Intercalation of a new tier of transcription regulation into an ancient circuit. Nature *468*, 959-963

10. Gerlt, J.A., and Babbitt, P.C. (2001). Divergent evolution of enzymatic function: mechanistically diverse superfamilies and functionally distinct suprafamilies. Annu. Rev. Biochem. *70*, 209-246

11. O'Brien, P.J., Herschlag, D. (1999). Catalytic promiscuity and the evolution of new enzymatic activities. Chem. Biol. *6*, R91-R105

12. Roodveldt, C., and Tawfik, D.S. (2005). Shared promiscuous activities and evolutionary features in various members of the amidohydrolase superfamily. Biochemistry *44*, 12728-12736

13. Taylor, J.W., and Berbee, M.L. (2006). Dating divergences in the Fungal Tree of Life: review and new analyses. Mycologia *98*, 838-849

14. Thornton, J.W., Need, E., and Crews, D. (2003). Resurrecting the Ancestral Steroid Receptor: Ancient Origin of Estrogen Signaling. Science *301*, 1714-1717

15. Zalatan, J.G., Coyle, S.M., Rajan, S., Sidhu, S.S., and Lim, W.A. (2012). Conformational Control of the Ste5 Scaffold Protein Insulates Against MAP Kinase Misactivation. Science *337*, 1218-1222

16. Madhani, H.D., and Fink, G.R. (1998). The riddle of MAP kinase signaling specificity. Trends in Genetics *14*, 151-155

17. Good, M., Tang, G., Singleton, J., Reményi, A., Lim, W.A. (2009) The Ste5 Scaffold Directs Mating Signaling by Catalytically Unlocking the Fus3 MAP Kinase for Activation. Cell *136*, 1085-1097

18. Bhattacharyya, R.P., Remenyi, A., Good, M.C., Bashor, C.J., Falick, A.M., and Lim, W.A. (2006). The Ste5 Scaffold Allosterically Modulates Signaling Output of the Yeast Mating Pathway. Science *311*, 822-826

19. Malleshaiah, M.K., Shahrezaei, V., Swain, P.S., and Michnick, S.W. (2010). The scaffold protein Ste5 directly controls a switch-like mating decision in yeast. Nature *465*, 101-105

20. Tuch, B.B., Galgoczy, D.J., Hernday, A.D., Li, H., and Johnson, A.D. (2008a). The Evolution of Combinatorial Gene Regulation in Fungi. PLoS Biol *6*, e38.

21. Reményi, A. Good, M.C., Bhattacharyya, R.P., Lim, W.A. (2005). The Role of Docking Interactions in Mediating Signaling Input, Output, and Discrimination in the Yeast MAPK Network. Molecular Cell *20*, 951-962

22. Tanoue, T., Adachi, M., Moriguchi, T., and Nishida, E. (2000). A conserved docking motif in MAP kinases common to substrates, activators and regulators. Nature Cell Biology *2*, 110-116

23. Huse, M., and Kuriyan, J. (2002). The Conformational Plasticity of Protein Kinases. Cell *109*, 275-282.

24. Ma B., Kumar, S., Tsai, C.J., Nussinov, R. (1999) Folding funnels and binding mechanisms. Protein Eng. *12*, 713-720

25. Kornev, A.P., Haste, N.M., Taylor, S.S., and Eyck, L.F.T. (2006). Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. PNAS *103*, 17783-17788

26. Schindler, T., Bornmann, W., Pellicena, P., Miller, W.T., Clarkson, B., and Kuriyan, J. (2000). Structural Mechanism for STI-571 Inhibition of Abelson Tyrosine Kinase. Science *289*, 1938-1942

27. Hardy, J.A., and Wells, J.A. (2004). Searching for new allosteric sites in enzymes. Current Opinion in Structural Biology *14*, 706-715

28. Kuriyan, J., and Eisenberg, D. (2007). The origin of protein interactions and allostery in colocalization. Nature *450*, 983-990

29. Erlanson, D.A., Wells, J.A., and Braisted, A.C. (2004). TETHERING: Fragment-Based Drug Discovery. Annual Review of Biophysics and Biomolecular Structure *33*, 199-223

30. Tuch, B.B., Li, H., and Johnson, A.D. (2008b). Evolution of Eukaryotic Transcription Circuits. Science *319*, 1797-1799

Supplemental Text.

Extended Methods.

Identification, sequence analysis and cloning of kinase and scaffold orthologs

S. cerevisiae Fus3 and Kss1 sequences were used to query the fungal orthogroups database to identify 39 Erk-like kinase sequences from across the Ascomycota. The resulting sequences were aligned using MUSCLE and adjusted manually to ensure proper alignment of critical conserved kinase features such as the DFG motif. The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model using the MEGA5 software

package to produce the tree in Fig. S1A. Sequences were cloned directly from genomic DNAs or, in some cases, codon optimized forms were synthesized by Integrated DNA Technologies.

The *S. cerevisiae* Ste5 sequence was used to query the fungal orthologroups database to identify orthologous scaffold sequences. We also used BLAST to look for additional Ste5 sequences in organisms outside of those identified by the fungal orthogroups database, but did not find any additional sequences. Full-length sequences for the Ste5 orthologs we identified were cloned from genomic DNAs and served as the basis for subsequent subcloning of Ste5-FBD and Ste5-VWA sequences. A complete list of all kinase and scaffold constructs used in this study can be found in Table S1.

For identification of the Far1 and Cst5 orthologs indicated in Figure S4B, the same procedure was used as for identifying Ste5 orthologs, using the *S. cerevisiae* Far1 sequence to query Far1 orthologs and the *C. albicans* Cst5 sequence to query Cst5 orthologs.

Selection of Ste5 fragment sequence boundaries

Ste5-VWA sequence boundaries were determined based on (1) sequence alignment to *S. cer.* Ste5-VWA and (2) using secondary structure prediction with SSpred to identify regions topologically congruent to Ste5-VWA.

Unlike the *S. cer.* Ste5-VWA, the *S. cer.* Ste5-FBD is a short linear motif with no obvious sequence homology present in other Ste5 orthologs. Despite this, active FBDs could hypothetically be present in these orthologs, given that linear motifs are of low sequence complexity and so can be cryptically present even in the absence of any obvious homology. We

reasoned that because the *S. cer.* Ste5-FBD is found in between two easily identifiable Ste5 domains (the RING and PH domains), any functional FBD that was evolutionarily related to the *S. cer.* Ste5-FBD should also be found between these domains. As such, we identified both RING and PH domains in Ste5 scaffold sequences on the basis of both sequence alignment and secondary structure prediction with SSpred, and considered the region between the RING and PH domains of diverse Ste5 orthologs to serve as candidate FBD sequences.

Phylogenetic Analysis

The evolutionary history of Ascomycota Erk-like kinases was inferred by using the Maximum Likelihood method based on the JTT matrix-based model (*1*). The bootstrap consensus tree inferred from 500 replicates (*2*) is taken to represent the evolutionary history of the taxa analyzed (*2*). Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches (*2*). Initial tree(s) for the heuristic search were obtained automatically as follows. When the number of common sites was < 100 or less than one fourth of the total number of sites, the maximum parsimony method was used; otherwise BIONJ method with MCL distance matrix was used. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 39 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 342 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (*3*).

Protein purification

Recombinant proteins were expressed from constructs derived from previous work (pMBP, pETARA: (*4*); pSV272: (*5*); pFastBac-MBP: (*6*)) and were cloned by standard molecular biological techniques. Point mutants were generated by site-directed mutagenesis.

S. cerevisiae Kss1 and MBP-Ste7EE were expressed from *Spodoptera frugiperda* (SF9) cells, using the Bac-to-Bac Baculoviral Expression System (Invitrogen) at 27° as described previously (Reményi et al., 2005; Zalatan et al., 2012). Other Kss1 orthologs, Fus3 orthologs, Erk-like kinase orthologs, and Ste5 fragments were expressed and purified from *E. coli* BL21(T1R). Wildtype kinases were coexpressed with the bacterial tyrosine phosphatase YopH to ensure a homogeneously unphosphorylated activation loop tyrosine.

All constructs were purified by Ni-NTA chromatography, followed by amylose resin (for pMBP, pSV272, and pFastBac-MBP constructs) or glutathione resin (for pETARA constructs). Ion exchange (RESOURCE 15Q [Amersham], 20 mM Tris•HCl pH 8.0, 0-1M NaCl) and size exclusion chromatography (Superdex 75 10/300) were used if additional purification steps were necessary following affinity purification. Purified proteins were stored in standard buffer (20 mM Tris•HCl pH 8.0, 150 mM NaCl, 10% glycerol, and 2 mM TCEP), flash frozen, and stored at -80°.

In vitro kinase reactions

Ste7→MAPK reactions

Reactions were performed in Standard Kinase Buffer (SKB: 100 mM NaCl, 25 mM Tris pH 8.0, 0.05% NP-40, and 2 mM TCEP, 2mM $MgCl_2$) and initiated with addition of ATP to 2.0 mM. Reactions were performed under saturating concentrations of MAPK substrate (K_M for Ste7

phosphorylation of MAPKs typically ranges between 100-250 nM) and (if present) Ste5-VWA ortholog and contained 50 nM of MBP-Ste7EE, 5 µM MAPK substrate, and (if present) 5 µM Ste5-VWA ortholog. We further verified the [MAPK] was saturating by measuring the reaction rates in the presence of 10µM MAPK; no changes to the rates were observed. Time points were quenched by the addition of 4x SDS Loading buffer that contained 50 mM EDTA. Samples were separated on 4-12% Bis-Tris gels run in MES buffer and transferred to nitrocellulose membranes by standard methods. Western blots were performed using the 4370 anti-phospho p44/42 MAPK antibody as a primary and IRDye 800CW Goat Anti-Rabbit IgG antibody (Li-Cor #926-32211) as a secondary. Blots were visualized using the Li-Cor Odyssey Imaging System, normalized against recombinant standards for each MAPK substrate, and quantified using Odyssey 2.1 software as described previously. Kinetic constants and their errors were determined by non-linear least squares fitting of initial rate data in R.

For the Ste5 VWA titration experiments in Figure S2A, initial rates of MAPK autophosphorylation were determined as above but contained variable amounts of the Ste5 VWA domain.

Ste5 autoinhibition and FAB activation experiments

To determine the extent of autoinhibition in full-length or 'long' Ste5 fragments (constructs detailed in supplement), we measured initial rates for Ste7→*S. cer.* Fus3 phosphorylation (under the same conditions as described for the Ste5-VWA assays) in the presence of increasing amounts of the Ste5 scaffold protein to produce an 'activation curve' (see Fig. 2A) that could be fit to determine the maximal rate enhancement provided at saturation, as described previously

(*6*). Activation of the autoinhibited *A. gossypii* Ste5 ortholog (*Agos*Ste5) by the SR13 Fab antibody fragment was tested by measuring initial rates for Ste7→*S. cer.* Fus3 phosphorylation in the presence of 250 nM AgSte5 ± 2 μM SR13 Fab.

MAPK autophosphorylation

Reactions were performed in SKB and contained 10 μM MAPK and, if present, 25 μM of a Ste5-FBD sequence. Prior to initiation with 2 mM ATP, the MAPK and the Ste5-FBD were allowed to incubate for 30 minutes at room temperature. Time points were quenched with 4x SDS Loading buffer that contained 50 mM EDTA and samples were separated by SDS-PAGE and transferred to nitrocellulose. Western blots were performed using the 4G10 platinum anti-phosphotyrosine antibody as a primary (Milipore) and IRDye 800CW Goat Anti-Mouse IgG antibody (Li-Cor #827-08364) as a secondary. Blots were visualized using the Li-Cor Odyssey Imaging System, and kinetic constants and their errors for reactions in the presence or absence of Ste5-FBD fragment were determined by non-linear least squares fitting of initial rate data in R.

For the Ste5 FBD titration experiments in Figure S3B and S5B, initial rates of MAPK autophosphorylation were determined as before but reactions contained only 1 μM MAP kinase (to ensure [MAPK] < K_{act}) and variable amounts of the Ste5 FBD fragment.

Pulldowns of Erk-like kinases by S. cerevisiae Ste5-FBD

Pulldowns were performed using GST-tagged *S. cer*. Ste5-FBD as bait (or GST alone as a control) to capture ERK-like kinase preys. 100 μL of 50 μM GST-*S. cer*.-Ste5-FBD or GST was

incubated with 15 µL of glutathione agarose beads (Sigma) in SKB for 30 minutes at room temperature to saturate the beads with bait. The unbound protein was removed, and 100 µL of 10 µM kinase prey in SKB was applied to the beads and incubated for 1 hour at 4° with gentle agitation. Following incubation, beads were washed three times with 500 µL of SKB. The bound sample was eluted by boiling the beads in the presence of 30 µL 2x SDS loading buffer. Samples corresponding to the load, unbound, and pull-down fractions were analyzed by SDS-PAGE.

Morphological response to α-factor

S. cerevisiae strain W303 and *K. lactis* strain yLB17a (*7*) were used for morphological analysis as performed previously (*8*). For *S. cerevisiae*, an overnight culture was washed twice with distilled water to dilute Bar1 from the media and cultures were diluted to OD 0.05 in SC and grown for 4 hours at 30°. The culture was split into separate tubes and α-factor peptide (*S. cer.*-α: N-WHWLQLKPGQPMY-C) was added to the indicated concentration. After 180 minutes, the morphology of cells was determined by light microscopy. For *K. lactis*, an overnight culture was grown in SC, washed twice with distilled water, and grown for an additional 6 hours in SC media lacking phosphate to allow cells to become mating competent (*7,9*). The culture was split into separate tubes and α-factor peptide (*Klac*-α: N-WSWITLRPGQPIF-C) was added to the indicated concentration. After 6 hours, the morphology of cells was determined by light microscopy. For both *S. cerevisiae* and *K. lactis*, morphologies were manually determined by inspection of images in ImageJ. Unlike S. *cerevisiae*, the majority of *K. lactis* cells do not produce mating projections even under saturating doses of α-factor peptide (<30%) although the percentage of cells in the culture shmooing does indeed plateau with increasing dose. For this

reason, we normalized the percentage of cells shmooing for each species to the maximum percentage of cells shmooing we observed to facilitate comparison of the two profiles.

Analysis of in vivo mating behavior of Ste5 chimeras harboring VWA sequences from different species.

The ability of Ste5 chimeras harboring VWA sequences from different species to support wildtype mating *in vivo* was determined by a growth complementation assay. Transformants of *S. cerevisiae* strain RB201 (a ΔSte5 W303 derivative) harboring CEN/ARS plasmids expressing wildtype or chimeric Ste5 protein sequences from the native Ste5 promoter were grown overnight in SD–Ura media. Cultures were diluted to OD 0.5 and 10 µL of each culture was spotted onto SC-Ura plates and grown overnight at 30° to produce a uniform patch. To test for mating, a 10 mL culture of the MAYA12 mating tester strain (α mating type, *lys*-) was grown to OD 1.0, pelleted, and resuspended in 500 µL of YPD and plated as a lawn on a YPD agar plate. Once dry, the RB201 derivative patches were replica-plated to the MAYA12 lawn YPD plates and allowed to mate overnight at 30°. These plates were then replica-plated to synthetical minimal media plates on which only diploids produced from MAYA12 x RB201 matings are able to grow. After 2 days, the plates were photographed for analysis.

Supplemental Text References.

1. Jones, D.T., Taylor, W.R., and Thornton, J.M. (1992). The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci *8*, 275–282.

2. Felsenstein, J. (1985). Confidence limits on phylogenies: an approach unsing the bootstrap. Evolution *39*, 783–791.

3. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol Biol Evol.

4. Reményi, A., Good, M.C., Bhattacharyya, R.P., and Lim, W.A. (2005). The Role of Docking Interactions in Mediating Signaling Input, Output, and Discrimination in the Yeast MAPK Network. Molecular Cell *20*, 951–962.

5. Coyle, S.M., Gilbert, W.V., and Doudna, J.A. (2009). Direct Link between RACK1 Function and Localization at the Ribosome In Vivo. Mol. Cell. Biol. *29*, 1626–1634.

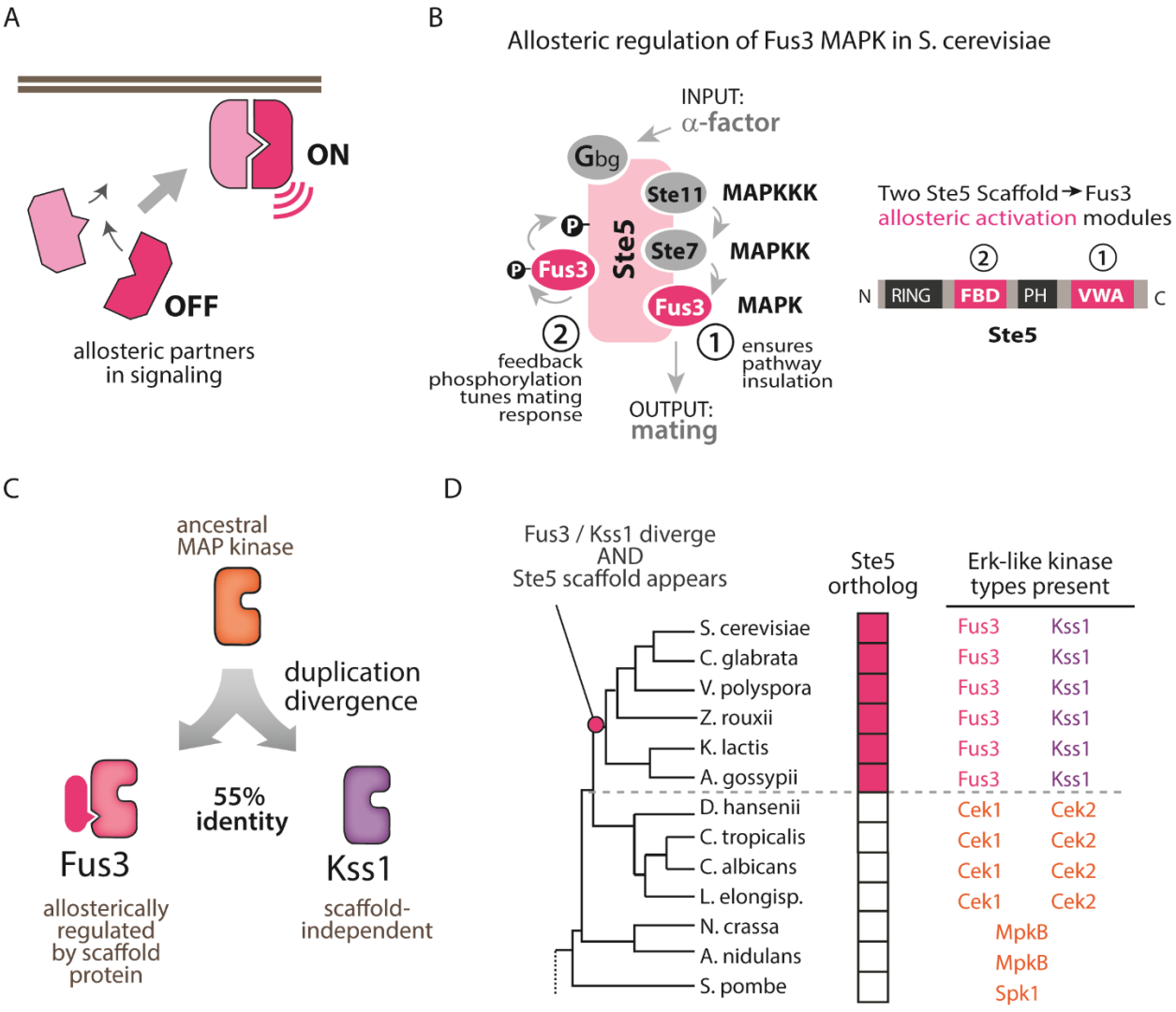
6. Zalatan, J.G., Coyle, S.M., Rajan, S., Sidhu, S.S., and Lim, W.A. (2012). Conformational Control of the Ste5 Scaffold Protein Insulates Against MAP Kinase Misactivation. Science *337*, 1218–1222.

7. Booth, L.N., Tuch, B.B., and Johnson, A.D. (2010). Intercalation of a new tier of transcription regulation into an ancient circuit. Nature *468*, 959–963.

8. Malleshaiah, M.K., Shahrezaei, V., Swain, P.S., and Michnick, S.W. (2010). The scaffold protein Ste5 directly controls a switch-like mating decision in yeast. Nature *465*, 101–105.

9. Tuch, B.B., Galgoczy, D.J., Hernday, A.D., Li, H., and Johnson, A.D. (2008). The Evolution of Combinatorial Gene Regulation in Fungi. PLoS Biol *6*, e38.

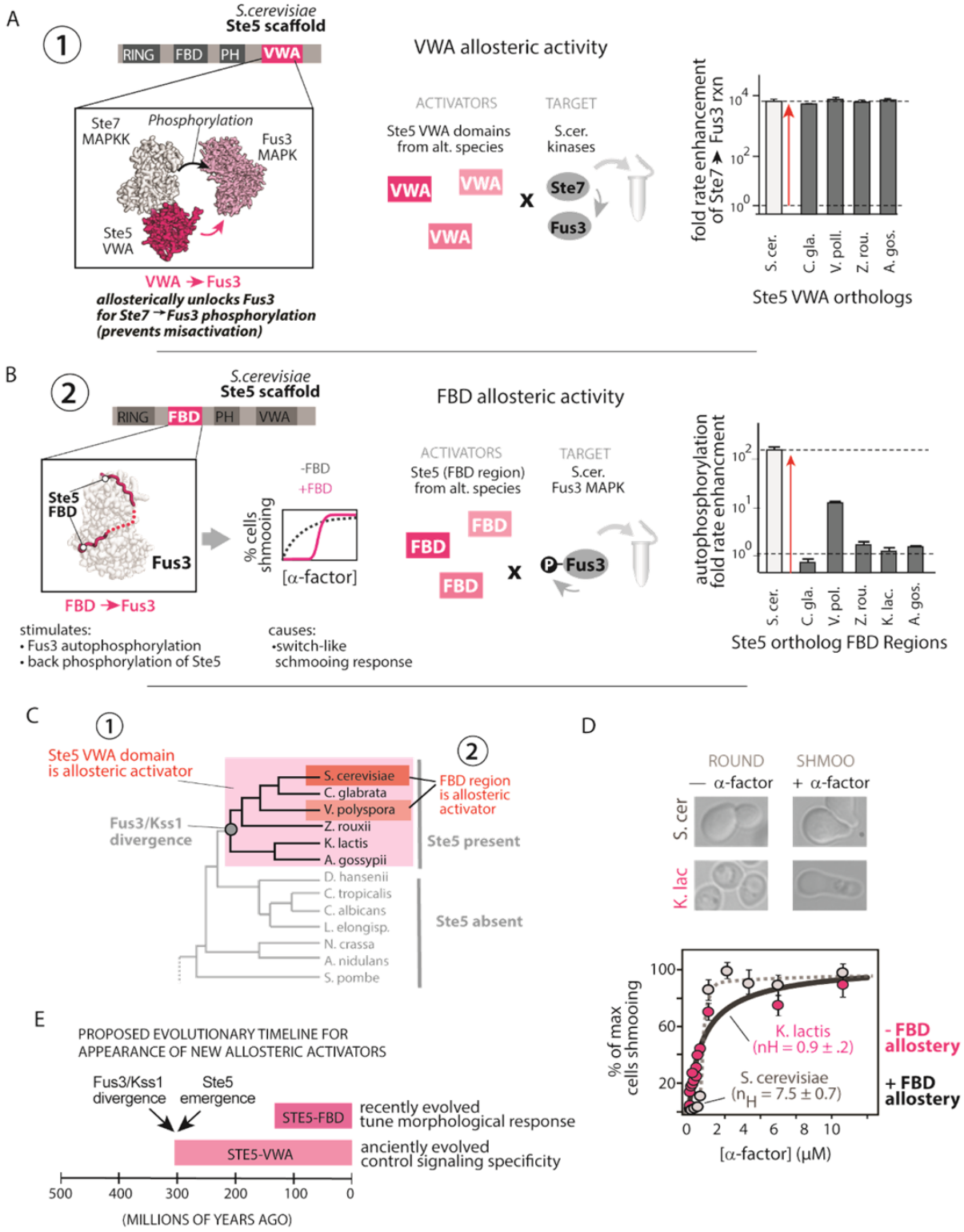
Figure 1. Fungal Erk kinase signaling repertoires provide a model system for biochemically interrogating the evolution of novel and divergent allosteric activation mechanisms.



(A) Allosteric interactions between signaling partners control when and where signaling molecules are activated in cells. (B) In the *S. cer.* mating MAP kinase pathway, two unique allosteric activities of the Ste5 scaffold regulate the MAP kinase Fus3 but not its paralog Kss1: (1) a VWA domain in Ste5 (Ste5-VWA) that is required to allosterically prime Fus3 for phosphorylation by the upstream MAPKK Ste7; and (2) a Fus3-binding domain (FBD) which

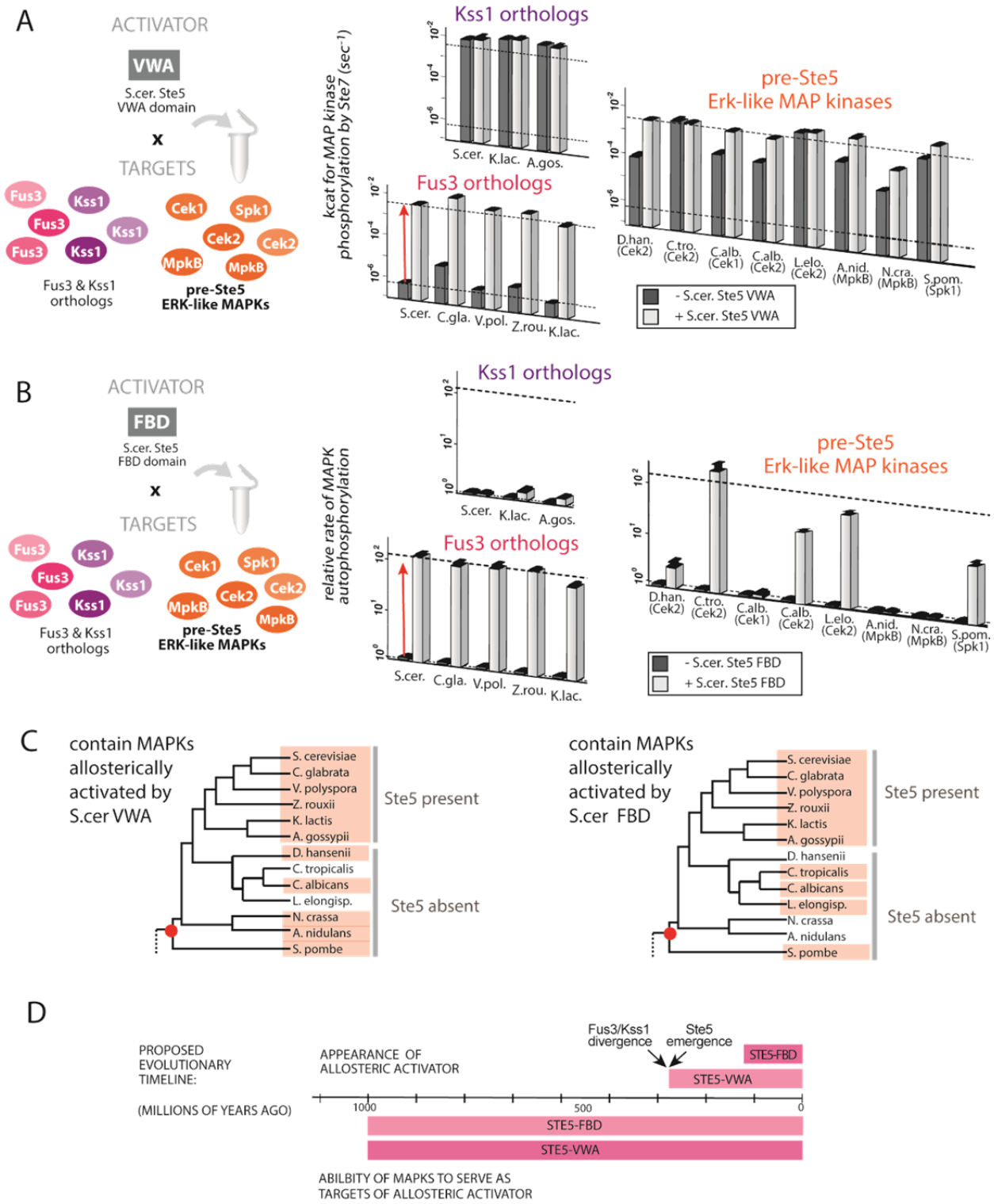
stimulates Fus3 autophosphorylation as part of a negative feedback loop that shapes the morphological response of cells to mating pheromone. (C) Fus3 and Kss1 are Erk-like kinases that are 55% identical and arose from a duplication of an ancestral MAP kinase. (D) Abbreviated phylogeny of fungal species from Ascomycota with the signaling repertoire (number and types of ERK-like kinases present; presence or absence of Ste5 scaffold) indicated for each species (see also Fig. S1A-C).

Figure 2. Tracking the emergence of MAPK allosteric activating domains within the Ste5 scaffold protein.



(A) In *S. cer.*, Ste5-VWA is required to allosterically unlock Fus3 for phosphorylation by Ste7; Ste7 cannot effectively phosphorylate Fus3 in the absence of this domain. Fold rate enhancements (mean ± SEM) for Ste7 catalyzed phosphorylation of *S. cer.* Fus3 in the presence of saturating amounts of the indicated Ste5-VWA ortholog are shown (see also titration curves in Fig. S2A). (B) In *S. cer.*, the Ste5-FBD is a linear motif between the Ste5 RING and PH domains that binds Fus3 and stimulates its autophosphorylation activity as part of a mechanism that results in a switch-like morphological dose response profile to α-factor. Rate-enhancements for *S. cer.* Fus3 autophosphorylation (mean ± SEM) provided by addition of 25 μM of the indicated Ste5-FBD are shown (see also titration curves in Fig. S3B). (C) Phylogeny of Ascomycota indicating the appearance of Ste5-VWA and Ste5-FBD scaffold activities as inferred from (A) and (B). (D) Morphological response to α-factor (mean ± SEM for % of maximum cells shmooing, n ≥ 500 cells) of *S. cer.* (grey dots; dashed line) and *K. lactis* (pink dots; solid line). Data were fit to a hill-equation to extract the parameter n_H. (E) Timeline indicating the proposed appearance of Ste5 allosteric activators relative to the appearance of the Ste5 and the Fus3/Kss1.

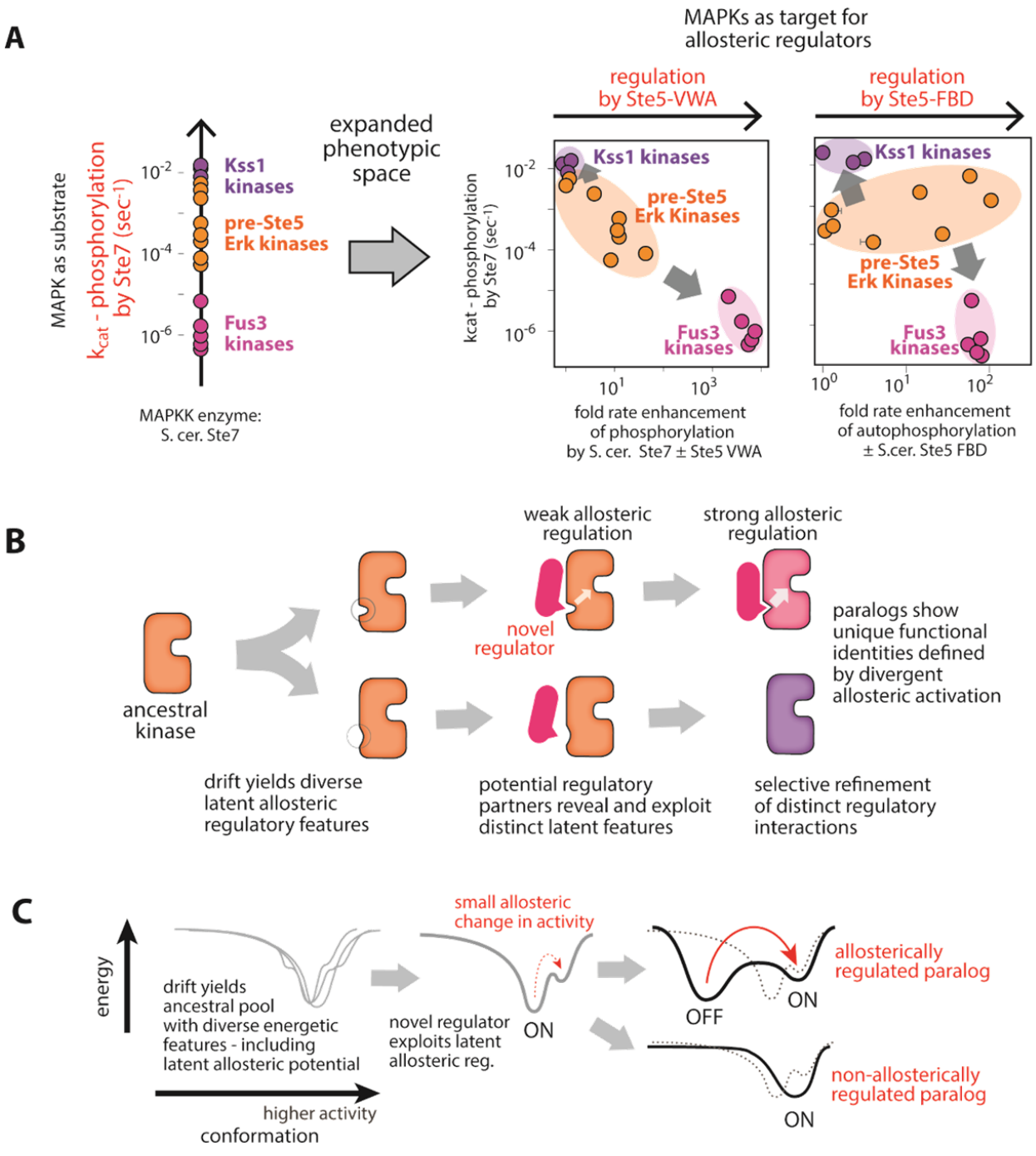
Figure 3. Ste5 scaffold domains can allosterically activate Erk-like MAP kinases that diverged prior to the evolution of Ste5.



(A) Rate constants (mean ± SEM) for Ste7 catalyzed phosphorylation of the indicated MAP

kinase substrate in the presence (white) or absence (grey) of saturating amounts of *S. cer.* Ste5-VWA for the indicated Fus3 orthologs, Kss1 orthologs or pre-Ste5 Erk-like MAP kinases. (B) Relative rates (mean ± SEM) for MAP kinase autophosphorylation (normalized to the no Ste5-FBD rate) in the presence (white) or absence (grey) of saturating amounts of *S. cer.* Ste5-FBD for the indicated Fus3 orthologs, Kss1 orthologs or pre-Ste5 Erk-like MAP kinases. (C) Phylogeny of Ascomycota fungi with species that contains an Erk-like kinase that can be regulated by the indicated Ste5 activator domain marked in pink. The red dot represents the last common ancestor that likely contained an Erk-like kinase that could be regulated by the indicated Ste5 activator. (D) Proposed timeline indicating when the capacity for regulation by a Ste5 activator appeared in the fungal Erk-like MAP kinase family relative to the appearance of Ste5 scaffold activator domains that target that capacity for regulation.

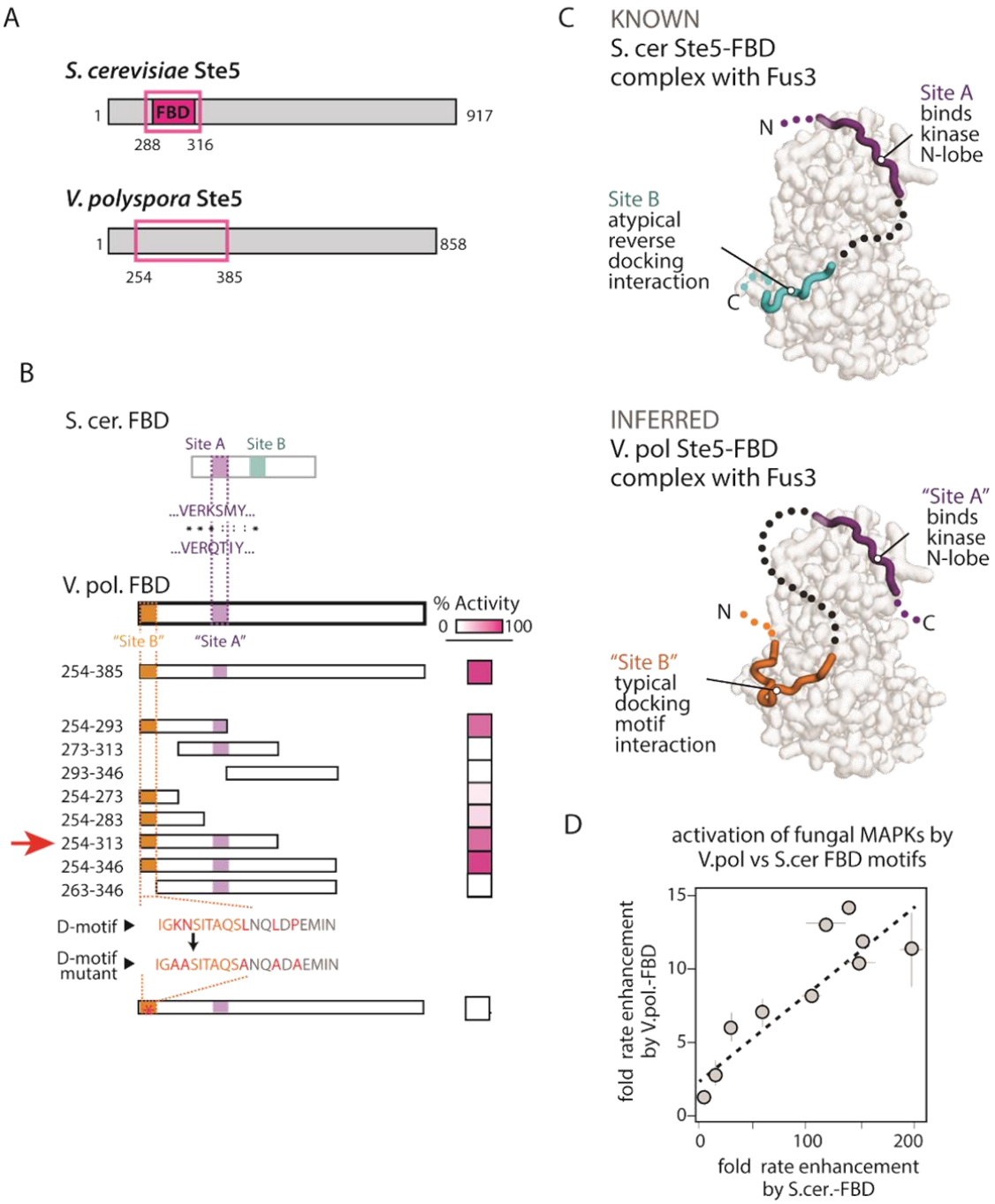
Figure 4. Drift in latent protein allostery provides a path for evolution of divergent regulatory phenotypes within seemingly equivalent kinase paralogs.



(A) Morphospace visualizations of MAP kinase biochemical diversity found in this study, either only as substrates for phosphorylation by the upstream MAPKK, or taking into account the

additional hidden dimensions of diversity in substrates upon interaction with the Ste5 scaffold. Circles in the plots correspond to individual MAP kinases (Fus3-type kinases: pink; Kss1-type kinases: purple; pre-Ste5 Erk-like kinases: orange) that we examined and indicate their associated properties. See also Fig. S4C. (B) Evolutionary model for novel and divergent regulation by exploitation of latent allosteric diversity. An ancestral kinase (orange) with some capacity for allosteric regulation is duplicated. In the absence of an effector, drift yields paralagous kinases with distinct latent regulatory features. Potential regulatory partners can reveal and exploit differences in these distinct latent features, providing a foothold for selection to refine and optimize the targets and effectors by coevolution to produce paralagous kinases with divergent allosteric responses to an effector molecule. (C) The above model illustrated in terms of the conformational energy landscape of the proteins.

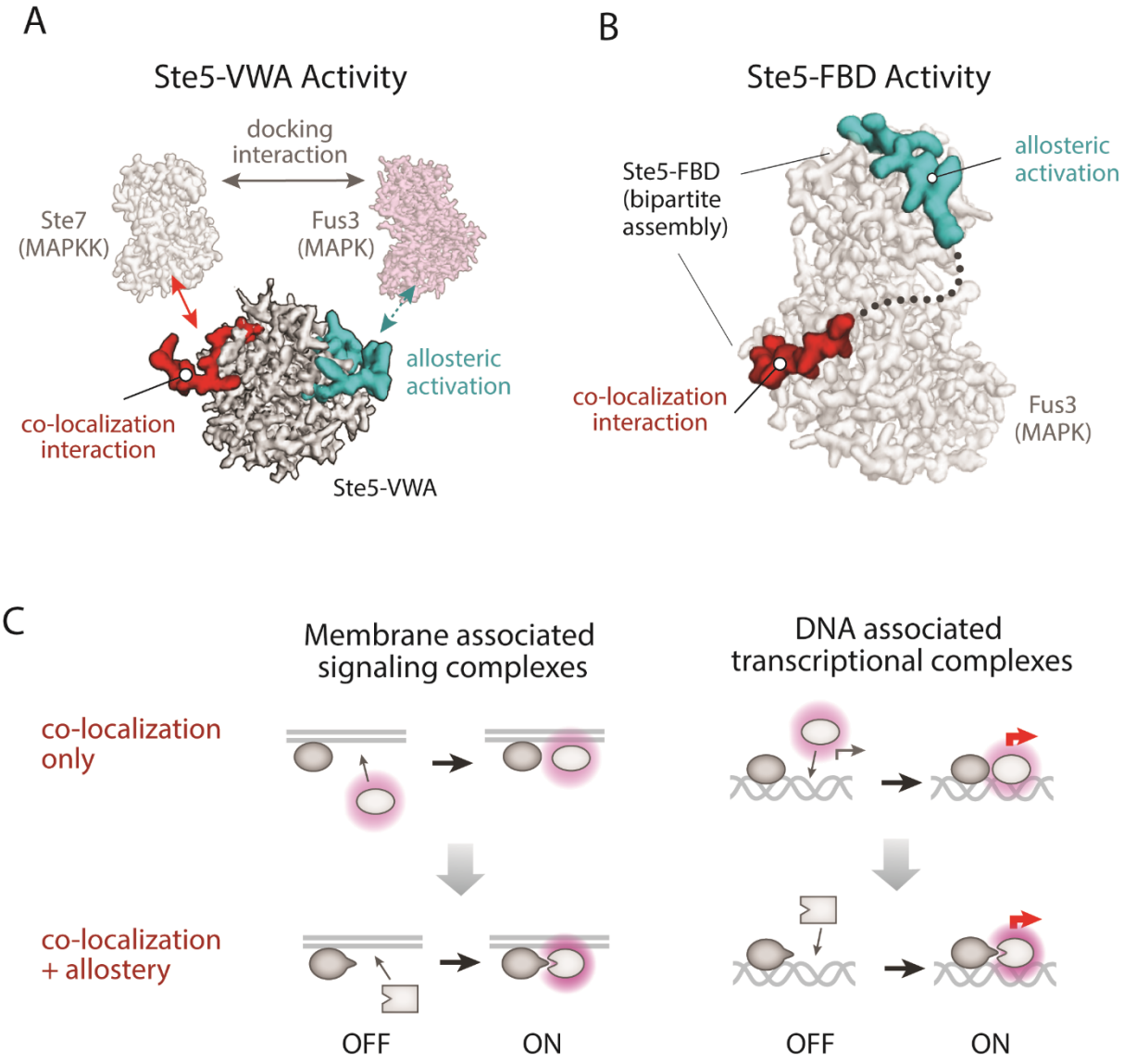
Figure 5. Dissection of the V.pol. Ste5-FBD motif with intermediate allosteric activity reveals alternative paths for coopting the same latent regulatory features.



(A) Position of minimal *S. cer.* Ste5-FBD sequence mapped previously (Bhattacharyya et al., 2006) and initial region of *V. pol.* Ste5 that showed FBD activity that serves as the starting point

for further analysis. (B) Truncation mapping of the *V. pol.* Ste5-FBD fragment, showing the relative activity of an indicated truncation or fragment (see also Fig. S5). This analysis identifies two sites that are required for activity. One site ("Site A") is similar to a sequence required for activity in the *S. cer.* Ste5-FBD fragment. A second site ("Site B") resembles a traditional MAPK docking peptide sequence. (C) Comparison of the known structure of the *S. cer.* Ste5-FBD•Fus3 complex to the inferred structure based on homology and the truncation mapping from (B). The two kinases appear to use similar "Site A" sequences to bind the N-lobe of the kinase but use distinct mechanisms to engage the docking groove of the MAP kinase C-lobe. (D) A plot indicating the magnitude (mean ± SEM) of the *S. cer.* FBD effect and *V. pol.* FBD effect; each point corresponds to an individual kinases we inspected. The tight linear relationship between these effects suggests that both the *S. cer.* and *V. pol.* Ste5-FBD sequences target the same allosteric features and diversity present in the MAP kinases we inspected.

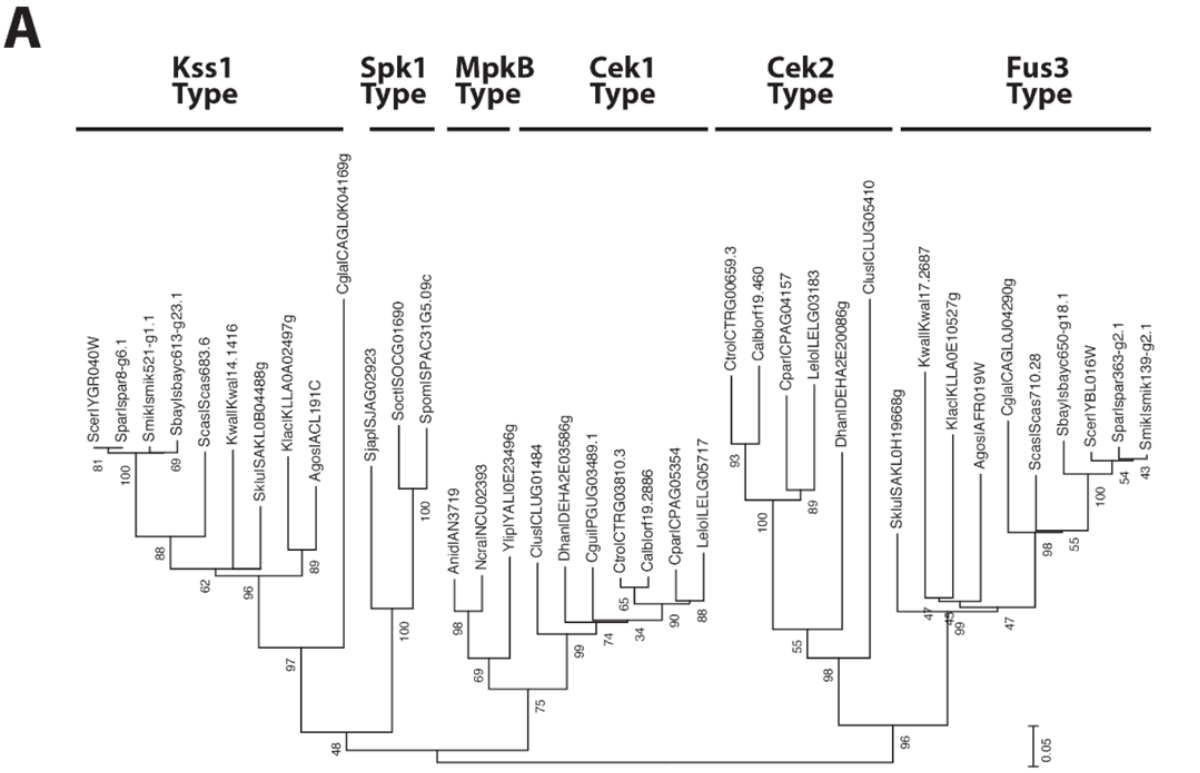
Figure 6. The role of co-localization in the evolution of new allosteric regulation.



(A) Colocalization interactions that are distinct from the essential allosteric surface of the Ste5-VWA assemble a Ste5•Fus3•Ste7 ternary complex that is essential for the Ste5-VWA to allosterically activate Fus3. These colocalization interactions are sufficient for tight complex assembly on their own. (B) Binding of the Ste5-FBD A-site motif to the Fus3 MAP kinase, which is essential for allosteric activation of autophosphorylation, requires a second-site interaction with the docking groove of MAP kinase. This docking interaction is, by itself, a non-

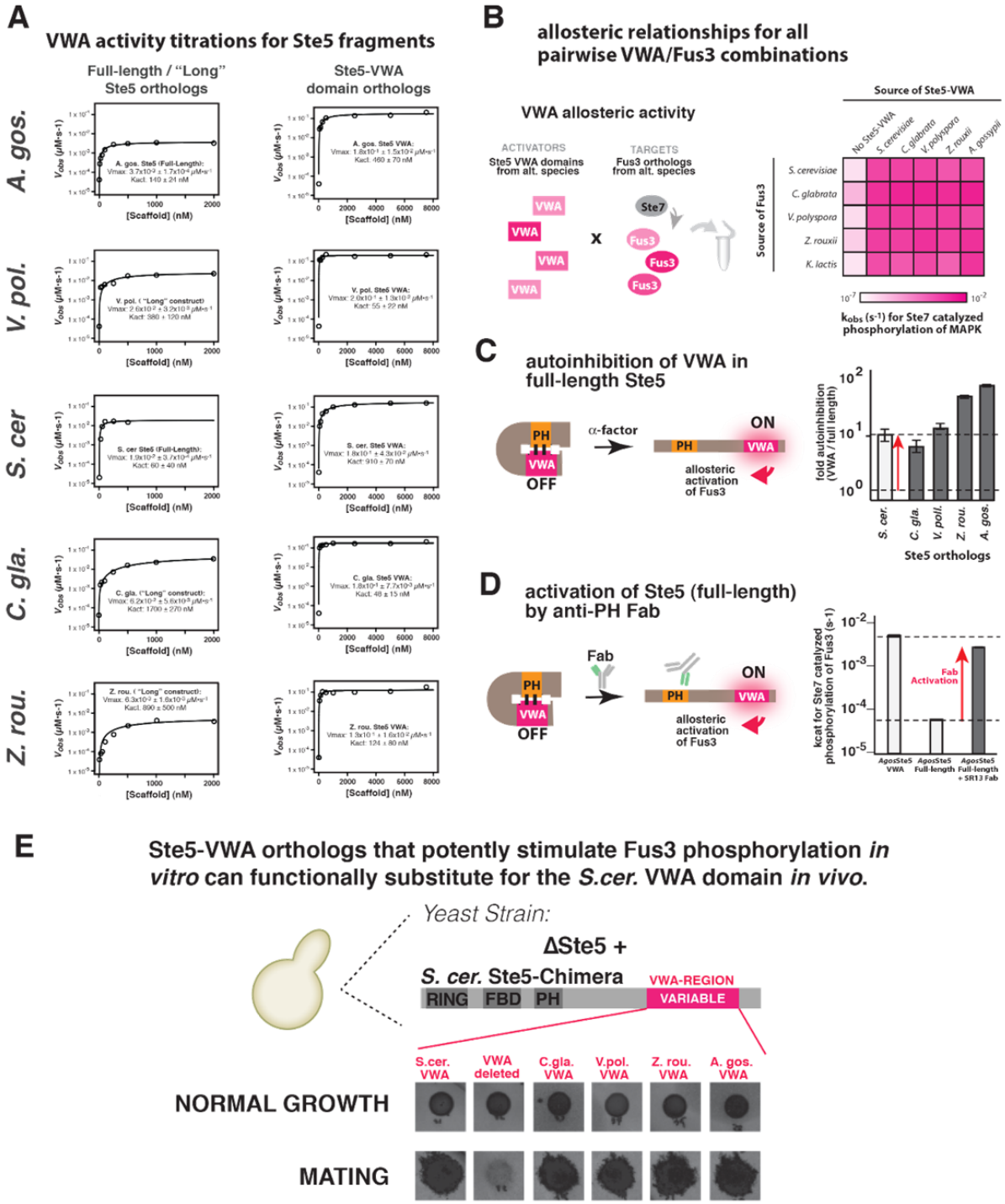
allosteric colocalization interaction that can be sufficient for complex formation. (C) Co-localization-based activation mechanisms—whether on scaffolds, membranes, or DNA—can facilitate the evolution of allosteric interactions between the co-localized components and yield the tighter, precise spatiotemporal control of activation that is observed in modern pathways.

Figure S1. Additional support that the Erk-like kinases are divergent across the Ascomycota and that distinct Erk-like kinase types are associated with specific fungal lineages; related to Figure 1.



(A) The bootstrap consensus tree inferred from 500 replicates is shown and is taken to represent the evolutionary history of the ERK-like kinase taxa analyzed, as detailed in the supplemental methods. Major branches with good bootstraps values correspond to particular kinase types found in distinct fungal lineages: Fus3 and Kss1 type kinases, found in the *Saccharomyces* and the *Kluveromyces*; Cek1 and Cek2 type kinases, found in the *Candida*; MpkB-like kinases, found in the filamentous fungi (pezizomycotina) such as *Neurospora crassa*; and Spk1-like kinases, found in taphrinomycotina such as *Schizosaccharomyces pombe*. (B) Simplified representation of tree in (A), showing the relationship of the major kinase types and the fungal lineages that contain those kinase types. (C) Fus3 crystal structure colored by the extent of conservation at each sequence position using the ClustalW similarity scoring system: (*) conserved, white; (:) similar, pink; (.) less similar, burgundy; ( ) no conservation, red. This coloring shows that while some elements of the hydrophobic core as well as key residues involved in catalysis are highly conserved, overall there is considerable divergence across the sequences we queried at most amino acid positions.

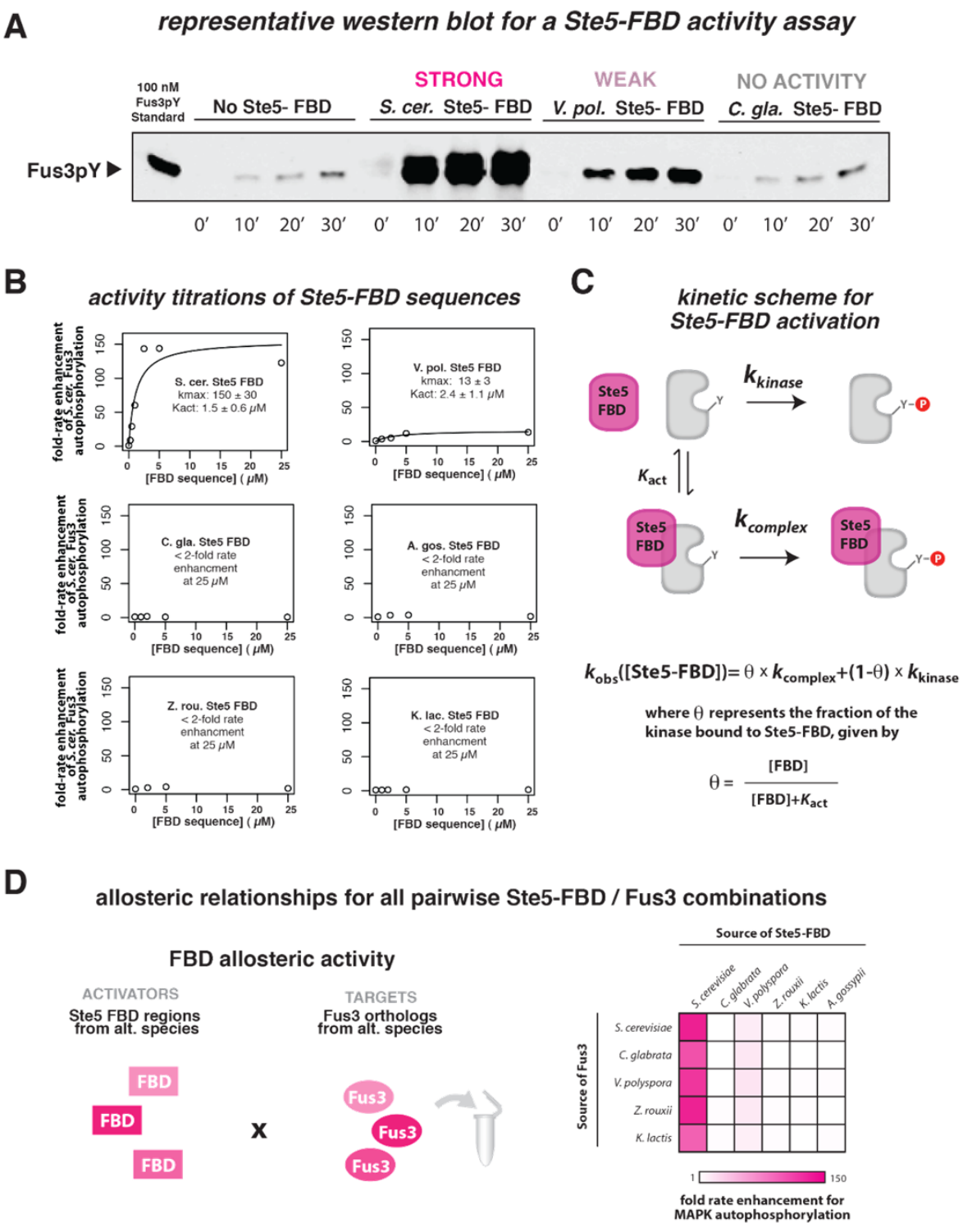
Figure S2. Additional support that a potent autoinhibited Ste5-VWA activity is conserved in all Fus3/Kss1/Ste5 containing species; related to Figure 2.



(A) Titration curves indicating the rate-enhancement for Ste7 catalyzed phosphorylation of Fus3

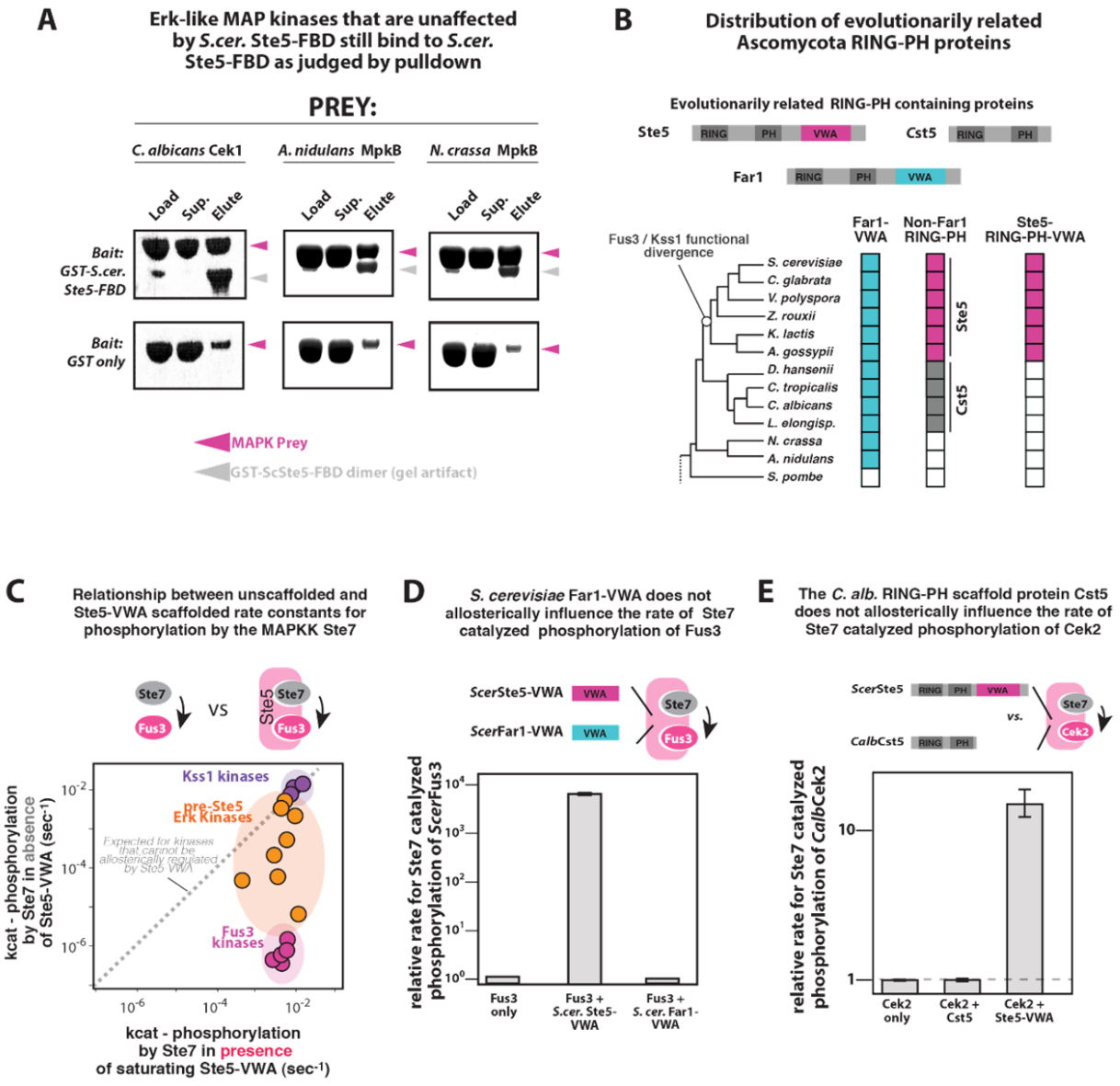
as a function of increasing concentrations of scaffold effector. The associated data, curves (fit to $V_{obs}=V_{max}$*[Scaffold]/([Scaffold]+K_{act})) and fit parameters for each construct are indicated. (B) Matrix indicating the k_{cat} for Ste7 catalyzed phosphorylation of an indicated Ste5-VWA and Fus3 ortholog pairing. (C) Rate enhancements provided by the longest Ste5 construct we could express for each ortholog relative to that of the corresponding isolated VWA domain are indicated. (D) A Fab antibody fragment that binds to the PH domain of *S. cerevisiae* Ste5 relieves Ste5 autoinhibition. The observed rate constant for Ste7-catalyzed phosphorylation of Fus3 in the presence of 250 nM *A. gossypii* Ste5-VWA domain, 250 nM *A. gossypii* full-length Ste5, and 250 nM *A. gossypii* full-length + 2 μM Sr13 Fab antibody are indicated. (E) Total growth and mating-dependent growth of RB201 (a ΔSte5 W303 derivative) *S. cerevisiae* strains containing Ste5 chimeras harboring VWA sequences from other species is shown.

Figure S3. Additional support that potent Ste5-FBD activity is a unique feature of only select Ste5-containing lineages; related to Figure 2.



(A) Representative western blot showing time-courses for Fus3 autophosphorylation in either the absence of any Ste5-FBD sequence or in the presence of 25 μM of the indicated Ste5-FBD fragment. (B) Titration curves indicating the rate-enhancement for Fus3 autophosphorylation as a function of increasing concentrations of scaffold effector. The associated data and—when appropriate—curves (fit to the equation described in (C)) and associated fit-paramters are indicated. *S. cer.* and *V. pol.* Ste5-FBD sequences differ in the maximum allosteric effect that they could exert on the kinase (maximal rate of autophosphorylation: *S. cer.*-k_{max} =154 ± 16 fold and *V. pol.*-k_{max} =12.8 ± 1.1 fold) and not in their apparent affinities (K_{act}, the concentration of activator required for half-maximal activation: *S. cer.*-K_{act} = 1.5 ± 0.6 μM and *V. pol.*-K_{act} = 2.4 ± 1.1 μM). (C) Scheme for Ste5-FBD activation of Fus3 used to fit data in (B) and extract a binding parameter (K_{act}) as well as a maximal rate enhancement parameter (k_{max}). (D) Matrix indicating the fold-rate enhancement for MAPK autophosphorylation of the corresponding Ste5-FBD and Fus3 orthologs. Ste5-FBD activity appears to be absent from the majority of species tested and the allosteric features that *S. cer.* Ste5-FBD and *V. pol.* Ste5-FBD act on is present in all Fus3 orthologs to a comparable degree, even when no Ste5-FBD activity is present in that species.

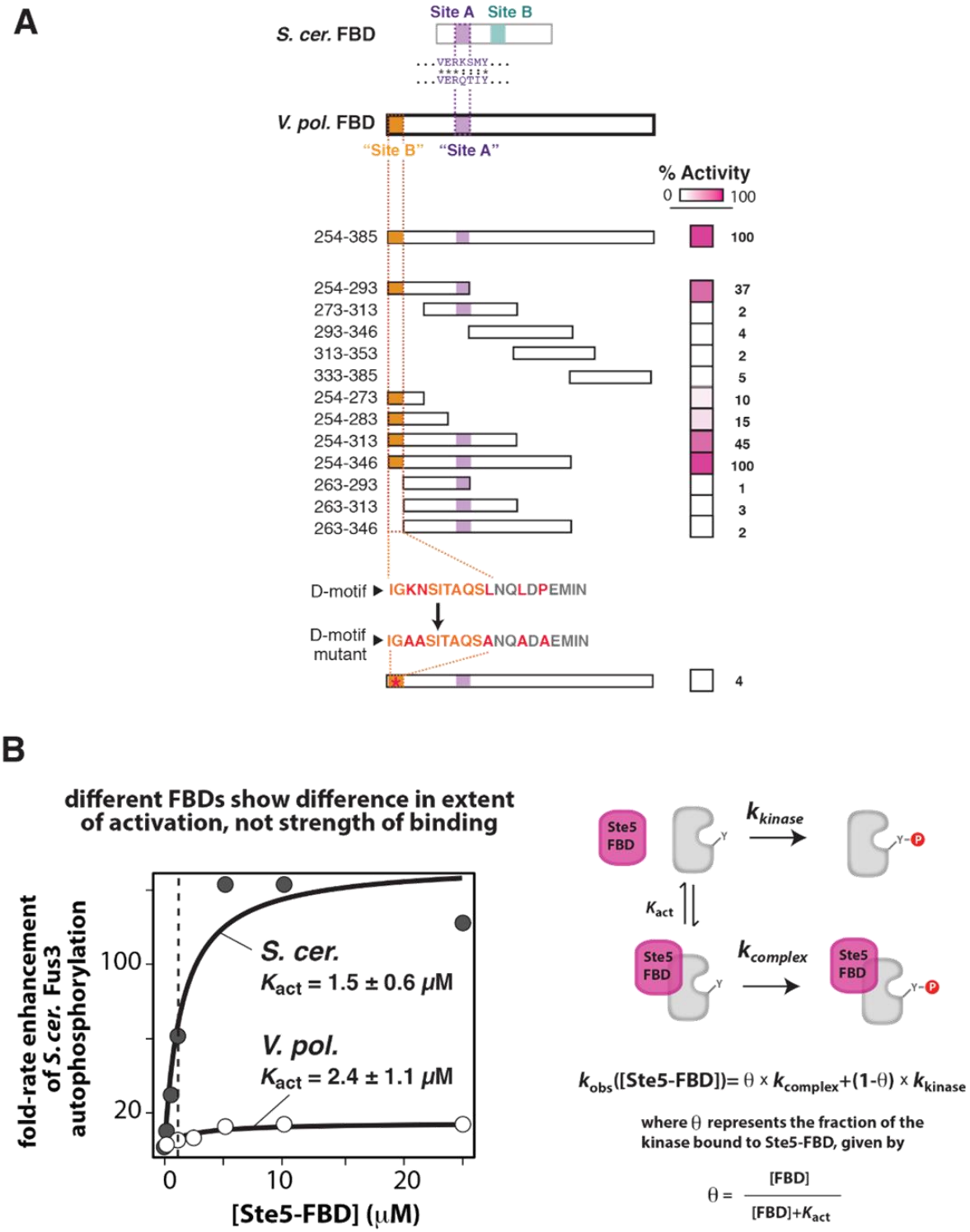
Figure S4. Additional support for a latent and drifting capacity for allosteric regulation by Ste5-FBD and Ste5-VWA in the pre-Ste5 Erk-like MAP kinases; related to Figure 4.



(A) GST pulldowns showing the extent of *S. cer.* Ste5-FBD binding to the indicated Erk-like kinases that did not respond to allosteric regulation by *S. cer.* Ste5-FBD. (B) Distribution of RING-PH containing proteins that are evolutionarily related to Ste5 across the Ascomycota. Far1

orthologs were identified in all species except *S. pombe*; Cst5 orthologs were identified only in species within the *Candida* (CTG) clade; only the *Saccharomyces* and *Kluveromyces* clades were found to contain Ste5 RING-PH-VWA orthologs. (C) Relationship between rate constants for Ste7 catalyzed phosphorylation of an Erk-like MAP kinase in the presence or absence of the *S. cer.* Ste5-VWA domain. Circles in the plots correspond to individual MAP kinases (Fus3-type kinases: pink; Kss1-type kinases: purple; pre-Ste5 Erk-like kinases: orange) that we examined and indicate their associated properties. The grey dotted line indicates where kinases would be expected to be distributed if they were not responsive to Ste5-VWA allosteric regulation. (D) Rate constants for Ste7 catalyzed phosphorylation of Cek2 under the indicated conditions are shown. *S. cer.* Ste5-VWA, but not Cst5, stimulated phosphorylation of *C. alb.* Cek2 by ~12-fold. (D) The rate constants for Ste7 catalyzed phosphorylation of *S. cer.* Fus3 under the indicated conditions are shown. *S. cer.* Far1-VWA does not appear to stimulate phosphorylation of Fus3, suggesting VWA domains from Far1 do not possess cryptic allosteric activity.

Figure S5. Additional biochemical dissection of the V.pol. Ste5-FBD activation mechanism; related to Figure 5.



(A) Additional truncation mapping data for the *V. pol.* Ste5-FBD fragment, showing the relative activity of an indicated truncation or fragment. (B) Comparison of *S. cer.* and *V.pol.* Ste5-FBD titration profiles for the rate-enhancement of Fus3 autophosphorylation. Data were fit to the indicated model to extract K_{act} and k_{max} parameters. These two motifs differed in the maximum allosteric effect that they could exert on the kinase (maximal rate of autophosphorylation: *S. cer.*-k_{max} =154 ± 16 fold and *V. pol.*-k_{max} =12.8 ± 1.1 fold) and not in their apparent affinities (K_{act}, the concentration of activator required for half-maximal activation: *S. cer.*-K_{act} = 1.5 ± 0.6 μM and *V. pol.*-K_{act} = 2.4 ± 1.1 μM).

Table S1. Plasmid constructs used in this study, related to Experimental Procedures.

		Bacterial Resistance
Ste5 "Long" PH-VWA containing constructs		
pJZ001	pMal-S.cerevisiae-Ste5-Full-length	amp
pSC158	pMal-A.gossypii-Ste5-Full-length	amp
pSC220	pMal Ste5 PH-VWA C. glabrata (143-600)	amp
pSC222	pMal Ste5 PH-VWA V. pol.(366-847)	amp
pSC223	pMal Ste5 PH-VWA Z. rouxii (341-881)	amp
Ste5-VWA constructs		
vMG750	pMal-S.cerevisiae-VWA	amp
pSC14	pSV272-C.glabrata-VWA	kan
pSC16	pSV272-K.lactis-VWA	kan
pSC18	pSV272-Z.rouxii-VWA	kan
pSC20	pSV272-V.poysporusl-VWA	kan
pSC22	pSV272-A.gossypii-VWA	kan
Ste5-FBD constructs		
pSC230	pETARA-S.cerevisiae-Ste5-FBD	amp
pSC227	pETARA-A.gossypii-Ste5-FBD	amp
pSc228	pETARA-C.glabrata-Ste5-FBD	amp

pSC229	pETARA-K.lactis-Ste5FBD	amp
pJF1	pMal-S. cerevisiae-Ste5-FBD	amp
pJF2	pMal-A.gossypii-Ste5-FBD	amp
pJF3	pMal-C.glabrata-Ste5-FBD	amp
pJF4	pMal-K.lactis-Ste5FBD	amp
pJF5	pMal-Z.rouxii-Ste5-FBD	amp
pJF6	pMal-V.polyspora-Ste5-FBD	amp
pSC268	pMal V.p. FBD (254-385)	amp
pSC269	pMal V.p. FBD (273-313)	amp
pSC270	pMal V.p. FBD (293-346)	amp
pSC271	pMal V.p. FBD (313-353)	amp
pSC272	pMal V.p. FBD (333-385)	amp
pSC274	pMal V.p. FBD (254-313)	amp
pSC275	pMal V.p. FBD (254-346)	amp
pSC276	pMal V.p. FBD (263-283)	amp
pSC277	pMal V.p. FBD (263-293)	amp
pSC278	pMal V.p. FBD (263-313)	amp
pSC279	pMal V.p. FBD (263-346)	amp
pSC298	pMal V.p. FBD (254-273)	amp
pSC299	pMal V.p. FBD (254-283)	amp
pSC304	pMal V.p. FBD(K256A,N257A,L264A,L267A,P269A)	amp

Fus3 constructs		
vMG500	pBH4-S. cerevisiae Fus3	amp
pSC15	pSV272-C.glabrata-Fus3	kan
pSC17	pSV272-K.lactis-Fus3	kan
pSC19	pSV272-Z.rouxii-Fus3	kan
pSC21	pSV272-V. pol. Fus3	kan
vMG510	pBH4-S. cerevisiae-Fus3(K42R)	amp
pSC199	pSV272-C.glabrata-Fus3(K42R)	kan
pSC200	pSV272-K.lactis-Fus3(K42R)	kan
pSC201	pSV272-Z.rouxii-Fus3(K42R)	kan
pSC202	pSV272-V. pol. Fus3(K42R)	kan
Kss1 constructs		
vMG800	pFBHt-b-S. cerevisiae Kss1	amp
pSC318	pMal-A. gossypii Kss1	amp
pSC319	pMal-K.lactis-Kss1	amp
pSC322	pMal-A.gossypii-Kss1(K42R)	amp
pSC323	pMal-K.lactis-Kss1(K42R)	amp
Erk-like kinase constructs		
pSC213	pMal C.albicans Cek1	amp

pSC214	pMal C.albicans Cek2	amp
pSC215	pMal C. albicans Cek1 K42R	amp
pSC216	pMal C. albicans Cek2 K42R	amp
pSC226	pMal S. pombe Spk1	amp
pSC233	pMal S. pombe Spk1(K42R)	amp
pSC311	pMal-D.hansenii-Fus3(K42R)	amp
pSC312	pMal-L.elongisporus-Fus3(K42R)	amp
pSC313	pMal-C.tropicalis-Fus3(K42R)	amp
pSC315	pMal-D.hansenii-Fus3	amp
pSC316	pMal-L.elongisporus-Fus3	amp
pSC317	pMal-C.tropicalis-Fus3	amp
pSC327	pMal-A.nidulans-MpkB	amp
pSC328	pMal-N.crassa-MpkB	amp
pSC329	pMal-A.nidulans-MpkB K42R	amp
pSC330	pMal-N.crassa-MpkB K42R	amp
Ste5 Chimera Constructs for *in vivo* analysis		
pSC31	pRS316-S.cer.Ste5_HA	amp
pSC32	pRS316-S.cer.Ste5ΔVWA_HA	amp
pSC34	pRS316-S.cerSte5(A.gos.-VWA)_HA	amp
pSC35	pRS316-S.cer.Ste5(C.gla.-VWA)_HA	amp
pSC38	pRS316-S.cer.Ste5(V.pol.-VWA)_HA	amp

pSC39	pRS316-S.cer.Ste5(Z.rou.-VWA)_HA	amp
Miscellaneous constructs		
pJZ224	pFB1MBP_Ste7EE	amp
pSC194	SR13 Fab Expression Vector	amp
pSC333	pMal-*S. cer.* Far1-VWA	amp
pSC212	pMal-*Calb*Cst5	amp

Table containing the list of all plasmid constructs that were used in this study, either for the expression of recombinant protein or for *in vivo* experiments in *S. cerevisiae.* The table contains a database index for plasmid requests as well as a brief description of the associated plasmid and its bacterial resistance marker.

Chapter 4. Probing the Versatility and Fragility of Ras GTPase Systems by Reconstituting Dynamic Signal Processing Networks *In Vitro*.

Abstract.

The Ras-family small GTPases are signaling hubs that control dynamic cellular processes like proliferation and morphological change. Multiple enzymes act in concert during signaling to regulate Ras activity and the subsequent assembly of Ras effectors, and both mutational and expression level changes to these components are frequently associated with cancer and other diseases. Here, we developed a dynamic *in vitro* reconstitution of H-Ras signal processing that includes both upstream regulators and downstream effectors to define how different network configurations and perturbations shape the system response. We found that the degree to which oncogenic alleles of Ras distorted outputs was highly dependent on the balance of positive and negative regulators in the system. Under these multi-turnover conditions, Ras systems could produce either sustained or transient effector outputs in which the concentration and identity of signaling components strongly impacted the timing, duration, shape and amplitude of the response. Moreover, different effectors interpreted the same inputs with unique dynamics, enabling multiple distinct temporal outputs to be encoded in the system response. Our mapping of the space of output behaviors accessible to Ras systems and its associated structure reveals that these systems are readily adapted to produce an array of dynamic behaviors well-suited for diverse cell-signaling functions, but this comes with a trade-off of increased fragility as there exist numerous paths to system states with dramatically altered signaling behaviors that can cause disease and cancer.

Introduction.

Many dynamic processes in the cell such as proliferation, differentiation or morphological change are regulated by signaling through members of the Ras family of small GTPases (*1–5*),

and mutations in these important molecules is often associated with cancer or other diseases (Fig. 1A) (*6, 7*). These small GTPases act as macromolecular "switches" at cell membranes that cycle between an ON state when bound to GTP and an OFF state when bound to GDP (*5*). This notion of an ON and OFF state of the GTPase is manifest in differences in the conformation of the GTP bound state relative to the GDP bound state such that only the GTP bound form is permitted to interact with downstream effector molecules and assemble signaling complexes (*8–11*).

As enzymes, these GTPases are formally capable of binding GTP, hydrolyzing it to GDP+Pi, and releasing product to complete the catalytic cycle on their own, but in practice the GTPase is incredibly slow at each stage of this cycle except for the initial binding of nucleotide (*12–14*). As such, molecules that can accelerate these slow steps in the catalytic cycle function as core regulators of GTPase activity during signaling events: guanine exchange factors (GEFs), which promote product release by emptying the nucleotide pocket of the GTPase and allowing subsequent reloading of the GTPase with nucleotide (OFF->ON transition); and GTPase activating proteins (GAPs) which accelerate the hydrolysis of GTP to GDP+Pi (ON->OFF transition) (*15–18*). How Ras signaling systems process inputs into outputs, then, is dependent on Ras itself, the activity of GEFs and GAPs that regulate its activity, and the effector molecules that bind to and assemble on the activated GTPase (Fig. 1B).

Considerable *in vitro* work over several decades has provided structural and biochemical insights into how GEFs, GAPs and effectors function, as well as how their activity can be controlled through mechanisms like autoinhibition and allostery (*19–26*). To study the regulators, these reconstitutions are almost always performed under conditions in which Ras cannot productively cycle: GEF assays monitor a single exchange of fluorescent nucleotide for non-fluorescent nucleotide; GAP assays monitor a single turnover of GTP without the possibility

of the nucleotide reloading (*27*). Likewise, studies of effector interactions with activated GTPase are typically done under non-cycling conditions using non-hydrolyzable nucleotide analogs to measure equilibrium binding constants (*28–30*).

By comparison, cell-based investigations of signaling are inherently multi-turnover with respect to Ras but provide far less control over the internal system parameters and usually contain fewer observables that are often far removed from the proximal signaling events. A variety of ingenious experiments have aimed to probe Ras activation during signaling in cells using fluorescent reporter molecules or nucleotide analogs (*31–33*). However, the complications of using these tools in living cells have made it difficult to systematically probe the behavior of the Ras signaling module.

To date, there has been little exploration *in vitro* as to how the multiple activities that regulate Ras work together to dynamically cycle Ras and control the assembly of competing effectors on activated Ras during signal processing. As such, we are largely ignorant about how the concentration and identity of the components within a Ras system define its signaling properties. Given that different cell types can harbor different configurations of network components as well as the fact that many diseases are associated with perturbations to the network configuration (Fig. 1C), a systems level reconstitution of Ras signaling systems could be highly informative as to types of outputs and behaviors that these systems are capable of, as well as how these systems respond to perturbations such as mutation.

Here, we develop a microscopy-based reconstitution of dynamic signal processing by human H-Ras, a canonical member of the Ras subfamily of GTPases. Our system includes both upstream regulators of Ras activity as well as downstream effectors that bind to and perceive RasGTP signals. This allows us to follow multiple cycles of Ras turnover in real-time using the

assembly of effectors on Ras as the output - the same way cells couple RasGTP levels to signaling outputs. Using this system we explore how oncogenic substitutions in Ras impact output behavior, how the concentration of each type of network component sculpts effector outputs in response to a simple step input of GEF activity, how systems that contain multiple competing effector molecules behave, and how different mechanisms for implementing positive feedback reshape the landscape of output behaviors. More generally, the methods we develop herein provide a framework for studying other assembly driven signaling systems or more complex systems that incorporate multiple interconnected signaling nodes.

Results.

Reconstituting Ras signal processing *in vitro* to track effector output dynamics across multiple GTPase turnovers during signaling.

To gain insight into how Ras transmits signals to downstream effectors under different network configurations or perturbations, we sought a dynamic *in vitro* reconstitution of Ras signal processing that would allow us to track effector outputs across multiple Ras turnovers. We reasoned that a microsphere surface charged with Ras could serve as a platform for the assembly and disassembly of fluorescent effector molecules from solution in response to inputs, much like the native Ras system functions in cells (Fig. 2A). The amount of effector on the surface of many individual beads could then be dynamically tracked during signaling by fluorescence microscopy and averaged to provide robust measurements of the output behavior.

To test this idea, we first asked whether we could observe GTP-dependent translocation of an effector molecule to a Ras-coated bead catalyzed by a guanine nucleotide exchange factor

(GEF). For our initial studies, we chose to use the catalytic domain from the RasGRF GEF, which is constitutively active and, unlike other GEFs, contains no allosteric feedback sites (*34*). Ni-NTA microspheres were charged with H-Ras•GDP (OFF state) and incubated in the presence of 50 nM (~K_D) of a model effector: the Ras-binding domain (RBD) from the C-Raf kinase (*35*). Under these basal conditions, the amount of fluorescence on the bead was comparable to the background levels of fluorescence from the effector in solution. We then added 2 υM of the catalytic domain of the RasGRF GEF and 5 mM of either GDP or GTP and monitored the effector fluorescence of the bead (Fig. 2B). Within seconds, there was noticeable accumulation of fluorescent effector on the bead surface of the GTP containing reactions, and considerable fluorescent signal was observed by ten minutes. In contrast, no fluorescent effector accumulated on the surface of reactions containing GDP, indicating that GEF-catalyzed translocation of the effector was dependent on Ras becoming GTP loaded.

We then tested whether this assay could show *quantitative* differences in this output behavior when different amounts of GEF activity were used as inputs. When identically loaded beads were stimulated with increasing amounts of RasGRF, we observed both faster rates of effector translocation and higher steady state amplitudes of effector output (Fig. 2C). This indicates that our reconstituted Ras signal processing system produced outputs that responded quantitatively to the amount of GEF present in the system.

In addition to being able to turn on, a dynamic reconstitution of Ras signal processing must be reversible and be able to turn off. To test the reversibility and turn-off of our system, we prepared beads loaded with H-Ras•GTP (ON state), incubated them with a saturating excess amount of C-Raf RBD effector (2.5 μM total, 50 nM fluorescently labeled), and monitored the loss of effector signal from the bead over time (Fig. 2D). In the absence of any GTPase

activating protein (GAP), effector fluorescence decayed with a half-time of 0.01 min-1, which is similar to the expected rate of intrinsic hydrolysis by H-Ras under our assay conditions (*12*, *13*). The fact that even a highly saturating concentration of effector had almost no inhibitory effect on the intrinsic turn-off of Ras is consistent with previous observations about the interaction between effectors and GTP loaded Ras (*36*). When increasing amounts of the catalytic domain from the Neurofibromin-1 GAP (NF1-GAP (*37*, *38*)) were included in the reactions, effector signal disappeared from the bead at an increased rate in a dose-dependent manner. Thus, as with the turn-on of the system, these data indicate that our reconstituted Ras signal processing system displays turn-off that responds quantitatively to the amount of GAP present in the system.

Having found that our system can produce effector outputs that are turned on by GEF and turned off by GAP, we wanted to verify that the system was truly multi-turnover and that the output dynamics would respond to both of these activities working in concert. To this end, we incubated Ras•GDP loaded beads with 50 nM fluorescent effector, initiated signaling with 2 μM RasGRF GEF and 5 mM GTP, and then added 1 μM NF1-GAP at 25 minutes (Fig. 2E). As before, addition of GEF stimulated assembly of effector on the bead as the levels of Ras•GTP increased. When NF1-GAP was added to the reactions, the system responded with a rapid decrease in effector levels before stabilizing at a non-zero plateau corresponding to the non-equilibrium steady state maintained by the balance of GEF and GAP activities present in the reaction.

Taken together, these data imply that our on bead reconstitution of H-ras signal processing can track dynamic effector outputs across multiple cycles of Ras activation and deactivation during signaling. This now puts us in position to explore how different mutational

states, network configurations, protein identities or feedback mechanisms affect signal processing by Ras GTPase systems.

The extent to which oncogenic Ras alleles distort wildtype signal processing depends on the balance of positive and negative regulatory activities in the network.

Substitution at the G12, G13, or Q61 positions of Ras is frequently associated with cancer or other diseases (*39*). These alleles are primarily though to impact Ras signaling through three mechanisms: 1) decreasing the intrinsic hydrolysis rate of the GTPase, 2) blocking GAP-mediated hydrolysis of the GTPase, and 3) altering the interaction and preference of the GTPase for downstream effectors (Fig. 3A) (*16, 39–41*). Because the same mutant allele of Ras can elicit different phenotypes in different cell types and tissues, it is likely that some system configurations are more sensitive than others to this type of oncogenic perturbation.

Using our dynamic, multi-turnover reconstitution of Ras signal processing, we examined how signaling networks bearing the G12V allele of the Ras GTPase distorted effector outputs relative to the wildtype Ras GTPase. By labeling wildtype and G12V Ras GTPases with different fluorophores, we could distinguish beads loaded with each variant in a common solution of network components to see differences in effector outputs from each system side-by-side.

With this approach, we first examined the output of 50 nM C-Raf RBD effector from G12V or wildtype Ras networks without GAP activity in response to a step input of 2 μM RasGRF (Fig. 3B). Under this network configuration, wildtype and G12V Ras systems produced very similar outputs with almost no difference in the total integrated effector output and only small differences in the overall dynamics of their responses. This suggests that neither the intrinsic

hydrolysis nor changes in C-Raf RBD effector interactions of the G12V substitution are particularly perturbative to the output of the signaling system under this GAP-free network configuration.

We then looked at the system responses of wildtype and G12V Ras systems to the exact same step-input (2 μM RasGRF) but in networks that now included 1 μM basal NF1-GAP (Fig. 3C). Unlike in the GEF-only networks, both the dynamics and amplitude of the effector output were substantially distorted by the Ras-G12V allele. In this network configuration, wildtype Ras produced a transient response that peaked within an hour and declined to a steady state less than 20% its maximum value. In contrast, outputs from G12V were sustained and increased in magnitude for over 6 hours before settling at a steady state more than 40 times higher than wildtype Ras. Thus, the G12V mutation is significantly perturbative in a high-GAP network context.

Together these data imply that the balance of positive and negative regulatory activities in a signaling network impacts the severity by which Ras-G12V distorts signals (note that similar results were also observed for G12C and Q61L alleles; Fig. S2). To determine which particular configurations are most distorted by the G12V allele, we measured the effector output response across a variety of input strengths ([GEF] activity) and NF1-GAP levels. We then calculated a distortion score as the fold-change integrated output from Ras-G12V relative to wildtype Ras and interpolated these data to produce a phase diagram of signal distortion by G12V under arbitrary network conditions (Fig. 3D). This revealed that G12V alleles were most perturbative with low-GEF inputs and a high-GAP network context, conditions in which the GAP activity would, for wildtype Ras, completely dominate over the small amount of activating GEF input. These observations are consistent with models of oncogenic Ras signaling in which low-level

inputs or noise from the environment that would normally be filtered out by basal GAP-activity are misinterpreted by the cell as bona fide activating signals.

The concentration and identity of Ras network components can modulate the timing, duration, shape, or amplitude of effector outputs.

Our comparison of wildtype and G12V Ras signaling systems illustrated the importance of the network composition in shaping signal processing outputs. Each individual network component is, in essence, a separate "dial" of the Ras signaling system that can be turned by adjusting the concentration of that component (Fig. 4A). Because expression levels of signaling components vary across different cell types and are often different in oncogenic states [REFs], we wondered how the level of each network component impacted signal processing wildtype Ras signaling networks. To this end, we fixed a particular input (2 μM RasGRF GEF) and, starting from a particular initial system configuration (50 nM C-Raf effector, no GAP activity) asked how titration of individual system components modulated the effector output.

Effects of GAP Activity

Having seen dramatic effects of negative regulatory activities in our distortion analysis of G12V Ras, we first looked more generally at how GAP activity sculpted signal processing dynamics in wildtype Ras networks. We considered two distinct GAPs domains with different biochemical properties and expression patterns: the NF1-GAP and the p120-GAP. NF1-GAP expression is somewhat ubiquitous but highest in neuronal cells and leukocytes, has a tight Km for Ras (0.3 uM) and modest kcat ($1.4s^{-1}$) (*42*). In contrast, p120 has a higher Km (9.7 uM) for Ras, but also a higher kcat ($19\ s^{-1}$), and shows a much more ubiquitous expression profile (*42*).

When increasing amounts of NF1-GAP was included in networks, the same step input of GEF activity produced dramatically different effector output behaviors (Fig. 4B). First, increasing levels of NF1-GAP resulted in substantially lower end point steady state levels of bound effector. Moreover, the amount of NF1-GAP changed the dynamics of *approach* to steady state significantly: while GAP-free networks approached steady state from below, networks containing NF1-GAP showed transient overshoot that decayed back down to the steady state from above, creating an initial pulse of strong output followed by weaker levels of output in the long term. The pulse-width and peak-time of effector outputs were inversely correlated with the concentration of the NF1-GAP (Fig. 4B). Thus, NF1-GAP shapes not only the final steady-state levels of output in the system, but also the shape of the timing and duration of Ras signal processing.

We then repeated this analysis using the catalytic domain of p120GAP in our networks instead of NF1-GAP. Compared to the NF1-GAP, the impact of the p120GAP on the end-point effector output of the system was much less substantial (Fig. 4C). This may in part be owing to the much higher Km of p120GAP compared to NF1-GAP, leading to much lower apparent activity in the concentration regimes we could readily explore. Nonetheless, increasing amounts of p120-GAP levels did lead to a marked change in the output dynamics of Ras signal processing in a manner similar to that of NF1-GAP. As with NF1-GAP, transient behaviors emerged when p120GAP was present, and the pulse-width and peak-time of effector outputs were inversely correlated with the concentration of p120GAP. Thus, as with NF1-GAP, p120GAP shapes the dynamics and steady-state behavior of signal processing by Ras, but with a different dose-dependent behavior owing to its unique biochemical characteristics.

Effects of Ras Density

Given that the Ras expression levels can vary among different cell-types and that Ras distribution in the plasma membrane can be both free as well as clustered (*43*, *44*), the next system parameter we considered was the density of Ras. To explore this, we made a dilution series of Ras, loaded beads with each dilution, and then mixed these beads together to assay the responses of different Ras densities side-by-side in the exact same network solution. Because the Ras was fluorescently labeled, we could estimate the Ras density from each bead and bin the responses from similar density beads together to obtain average behaviors for different density classes.

Applying this approach to our fixed step-input, we found that increasing Ras density was correlated with increased amplitude of effector output produced by the system (Fig. 4D). This is expected as higher Ras densities provide more molecules that can form Ras-Effector complexes. Additionally, the normalized traces of these responses revealed differences in the dynamic behavior of effector output. When Ras densities were high, the step-response effector output monotonically approached its steady state and was sustained. At lower Ras densities, however, effector outputs showed an increasingly transient character. These differences likely reflect a switch from a network configuration in which Ras is in excess of the GEF to one in which the GEF is in excess of Ras. Such differences will cause a change the initial fraction of the Ras population that is activated, such that a much larger synchronous cohort is formed at lower Ras densities.

Effects of C-Raf RBD effector concentration

Finally, we considered the impact of effector concentration on signal processing, the actual molecules that are used by cells to perceive and interpret RasGTP dynamics in the cell. For these experiments, we fixed the amount of fluorescent C-Raf RBD effector at 50 nM and added

additional unlabeled C-Raf RBD to achieve a target final concentration of effector. We could then normalize the observed fluorescent effector output by its proportion in the total effector population to infer the true magnitude of the output.

When we measured the system step-response in the presence of increasing amounts of effector in the network, dramatic changes in the both the amplitude and dynamics of output were observed (Fig. 4E). At an effector concentration of 50 nM (~K_D), the system output showed the typical monotonic approach to a sustained steady-state. As effector concentrations increased, both the amplitude and the dynamics of the output response changed markedly. At 250 nM effector concentration, output increased over two hours to a level 20 times that of the 50 nM effector system at that time, before decaying down to a its final steady state level. At even higher effector concentrations (500 nM, 1000 nM), system output peaked quickly within 20 minutes and then decayed monotonically over several hours to its final steady state level. These data demonstrate that higher effector concentrations not only increase output amplitudes, but also enable the output to capture more transient features of the upstream RasGTP signal. This is because a higher effector concentration decreases the time needed to equilibrate against a fixed concentration of RasGTP.

A tunable space of diverse dynamic outputs achieved by titration of components

Our data have show that the identity and concentration of each component in a Ras signaling network can have a profound impact on the timing, duration, shape, or amplitude of effector outputs. This implies by tuning the abundance and *identity* of network components and controlling the strength of inputs, a variety of diverse dynamic effector output programs can be realized by Ras signaling system (Fig. 5A). To explore this space of output programs more thoroughly, we fixed a particular effector concentration (~K_D) and measured system output from

different Ras densities, p120GAP concentrations, and input strengths ([GEF]). Because the number of different resulting output responses is large (see Fig. S3 and Fig. S4 for complete set of responses), we extracted 3 features from each output response trace that describe the behavior – integrated signal intensity, initial rate of activation, and a transient score that reflect the amount of overshoot in the response – and interpolated these values for different Ras density bins to create a phase diagram for each feature as a function of the network configuration (Fig. 5B).

These phase diagrams illustrate how simply varying the levels of multiple components in the network, Ras signaling systems can access a surprisingly rich space of dynamic effector outputs that vary in amplitude, duration, shape and timing. Moreover, by sampling behaviors throughout this space, we gain insight into its structure and thus the impact of perturbations that move the system from one region of the space to another.

Different Ras effectors perceive the same input uniquely, enabling multiple distinct temporal outputs to be encoded in the system response of multi-effector networks.

So far our characterization of Ras signal processing has used the C-Raf RBD as the sole downstream effector, but *bona fide* Ras networks in cells typically contain multiple effectors targeting different responses, and each effector possesses its own affinity for RasGTP and expression level in the cell (*11*, *41*). Because our reconstituted signal processing is microscopy based, we can track the behavior of multiple different competing effectors processing signals on the same bead simultaneously by labeling each effector with a different color fluorophore (Fig. 6A). To this end, we purified and labeled Ras-binding domains from the A-Raf and B-Raf

kinases, which have lower and higher affinities for RasGTP respectively, to examine the signal processing behavior of two-effector systems in either GAP-free or high NF1-GAP networks.

We first considered networks containing equivalent, physiological amounts of C-Raf and B-Raf effectors, which both have high affinity for RasGTP (*41*). In response to a 2 μM RasGRF GEF step input in a GAP-free network, C-Raf and B-Raf processed these signals with completely different dynamics (Fig. 6B). Initially C-Raf and B-Raf outputs assembled at comparable rates, but within 1 hour C-Raf output peaked and began to decrease while B-Raf continued to increase in output monotonically over the entire timecourse. The differences in the dynamic character of each effector's perception of the signaling input are more obvious in the normalized traces.

When these step-responses were re-examined in a high NF1-GAP network context, we continued to see different effector responses between C-Raf and B-Raf: C-Raf output peaked within 30 minutes before sharply declining to a steady state value 25% of its maximum. In contrast, B-Raf output peaked later at one hour, and declined to a 75% its peak maximum, a much higher steady state compared to C-Raf. Thus, in this case, one effector produces a transient response while a different effector produces a more sustained output.

We then performed a similar analysis for two-effector systems containing equivalent, physiological amounts of C-Raf and A-Raf RBDs, where A-Raf has a weaker affinity for RasGTP and higher off-rate from the GTPase than C-Raf (*41*). In GAP-free networks, the step-responses for A-Raf and C-Raf differed most significantly in terms of their amplitudes, with A-Raf output 20 times lower than C-Raf output (Fig. 6D). The normalized traces of each effector output show only minor differences in the dynamics of these outputs. In high NF1-GAP networks, we could only detect an output response from the C-Raf effector; the ability of A-Raf

to assemble on the bead was either filtered out by the low affinity combined with high levels of GAP activity, or was simply too weak to detect above the background A-Raf in solution.

Clearly effector molecules with distinct identities result in differential perception of inputs, but how easy is it a given effector molecule to acquire a new dynamic behavior? Given that A-Raf, B-Raf, and C-Raf have different affinities for the GTPase, we wondered whether mutations in C-Raf that changed its affinity would be sufficient to drive new dynamic behaviors. To this end, we characterized the step-response of two-effector network containing equivalent amounts (50 nM) of the C-Raf RBD and the C-RafN64A point mutant which has decreased affinity for RasGTP (*35*). Different dynamic output behaviors for each effector were observed for both GAP-free (Fig. 6F) and high NF1-GAP (Fig. 6G) networks. In both contexts, the lower affinity C-Raf N64A mutant peaked earlier and had lower overall amplitude than the wildtype C-Raf RBD, and like before, high GAP-networks sharpened these temporal differences. This shows that new dynamic behaviors are readily realized by mutation of an effector molecule.

Taken together, this analysis shows that distinct effector molecules can perceive the same input to a Ras signaling system with different dynamics and amplitudes depending on their affinities and biochemical properties. Consequently, a single step input can be in principle be used encode multiple temporally distinct outputs that peak and decline out of phase with one another, allowing for a sequence of different activities to be organized during signal processing. Furthermore, the extent of GAP activity in the network can influence how these different dynamic responses unfold, sharpening temporal distinctions in some cases while restricting the ability of certain effectors to assemble productively at all in other cases. Finally, because even simple point mutations to an effector can dramatically alter its output dynamics, new dynamic patterns are not difficult to produce and can be easily accessed during evolution.

Positive feedback alters the landscape of effector outputs from Ras system in different ways depending on how the feedback mechanism is implemented.

The signaling networks we have examined thus far have not included any explicit feedback mechanisms, and so the behaviors we have seen are solely the product of constitutive enzymatic activities and effector assembly processes unfolding in the simplest possible RasGTPase signaling circuit. However, many cellular circuits involve additional layers of regulation and feedback control that could alter the signaling properties of the system. To gain insight into how such regulation might alter the ground state signaling behavior of Ras GTPase systems, we examined the effect of two different modes of positive feedback on effector outputs.

Recruitment-based feedback mechanism (ON -> LOCALIZED ON)

In the first feedback mode, we made a synthetic GEF in which we fused the C-Raf RBD effector domain to the Ras-GRF GEF to produce "RasGRF-RBD" (Fig. 7A). In this case, the catalytic activity of RasGRF-RBD is always constitutive, but activated Ras will assemble the synthetic GEF on the bead surface to provide a higher effective concentration of GEF and potentially increase the apparent GEF activity. Thus, this feedback mechanism takes an ON molecule and makes it MORE ON in response to RasGTP. We then measured outputs from a variety of identical network contexts in which either wildtype RasGRF or RasGRF-RBD served as the input, and determined a feedback gain as the ratio of the integrated signal from the feedback system to the non-feedback system. We then interpolated these values to produce phase diagrams showing the impact of feedback on different network configurations (Fig. 7B, see Fig. S4 and Fig. S5 for data used in interpolation).

These data show that this type of feedback mechanism can produce strong effects in certain network configurations but have little to no effect in other configurations. In particular, networks that contained high amounts of basal GAP-activity but only small amounts of input GEF showed the strongest differences in signal. These correspond to regimes in which the amount of GAP activity in the system dominates the small amount of intrinsic catalytic activity of the GEF, but cannot overcome the small amount of localized GEF activity arising from the RasGRF-RBD feedback GEF (Fig. 7C). In contrast, the behaviors of the wildtype and feedback systems were most similar under high GEF, low GAP network configurations. In this regime, the GEF activity from the catalytic domain is sufficient to provide strong activation, and any additional boost in activity for localizing the GEF provides only marginal gains (Fig. 7D). Thus, this type of feedback mechanism seems most powerful for amplifying signals arising from weak inputs in a high turnover background.

Allosteric feedback mechanism (OFF -> ON)

For the second feedback mode, we replaced the RasGRF GEF with SOScat, the catalytic domain from the Son of Sevenless GEF. SOScat has an intrinsic feedback mechanism built in, such that the molecule has very low GEF activity in the absence of RasGTP, but RasGTP binding to a distal site on SOScat allosterically stimulates GEF activity to very high levels (Fig. 7E) (*21*, *34*). Thus, in contrast to the synthetic RasGRF effector fusion (which was always ON), this feedback mechanism takes an OFF molecule and makes it ON in response to RasGTP. We then measured outputs from a variety of identical network contexts in which either RasGRF (no feedback) or SOScat (allosteric feedback) served as the input, and determined a feedback gain as the ratio of the integrated signal from the feedback system to the non-feedback system (Fig. 7F).

These data show a very different effect of SOScat feedback on signaling outcomes than was observed for the RasGRF-RBD fusion. From the perspective of integrated signaling output, we were surprised to find that there was no network configuration in which SOScat and RasGRF displayed substantially different integrated effector outputs. However, when we examined the dynamics of the output produced from each system, we did see significant differences in behavior. In particular, output dynamics from SOScat networks appeared more stable and monotonic than those from RasGRF networks, but also turned on more slowly. These effects were most dramatic in high GAP network contexts in which RasGRF produced a large transient overshoot phase whereas SOScat produced a monotonic approach to the steady state (Fig. 7H). However, this increased stability was still noticeable even for low GAP high GEF networks in which RasGRF produces a sustained response (Fig. 7G).

These differences can be attributed to the OFF->ON feedback mechanism of SOScat in the context of a step-input: for a RasGRF step-input, the system experiences a sudden increase in catalytic activity all at once; for SOScat, though, this step-input in protein levels becomes a *ramp-input* in terms of protein *activity*, given that it takes multiple cycles of GEF activity to produce enough RasGTP for SOScat molecules to realize their full activity. Thus, this type of feedback mechanism appears to play less of a role in increasing the magnitude of effector outputs in response to weak inputs and more of a role in dampening dynamic and transient features that would arise naturally in the step-response without this feedback.

Discussion.

One system, many behaviors: versatility and fragility in Ras GTPase signal processing

In this work, we developed a multi-turnover reconstitution of Ras signaling to explore the space of dynamic output behaviors that could be produced by Ras GTPase systems and to characterize how each network component contributes to these behaviors. Using these assays explored how different perturbations such as oncogenic mutation, component levels, inclusion of additional effector molecules, or introducing positive feedback altered the landscape of available outputs.

Our experiments imply that, much in the same way that a single genome can encode multiple cell types that are regulated through differential gene expression, a single signaling system like Ras can encode multiple dynamic signal processing behaviors by regulating the concentration and identity of network components. Consistent with this, a simple survey of published p120GAP, Ras, and c-Raf mRNA expression levels across a variety of tissue types reveals a staggering amount of diversity in what types of network configurations are present in different cell types and tissues (Fig. 8A, see Table S1 for original data). The true diversity in these configurations is likely even greater given the plethora of additional GEFs, GAPs, Ras variants and effectors that cells can deploy. Thus, different cells can position themselves at different points in the space of available Ras network configurations to provide versatile top-level control of the amplitude and duration of signal processing events (Fig. 8B).

This versatility is not without trade-offs, however. In particular, we observed many different paths in network-space from one signaling processing behavior to another with much higher or sustained amplitude (Fig. 8C). These paths include classic oncogenic substitutions like G12V in Ras, but can also be realized by increased GEF activity, decreased GAP activity, or inclusion of high-affinity effectors that increase, extend and sustain signaling responses. While some of these perturbations have not been definitively recognized as drivers of cancer, many are associated with other RASopathies in humans, like Noonan syndrome or type 1

Neurofibromatosis (*7*, *45*). Thus, the same flexibility that allows Ras systems to realize many different signaling behaviors creates many opportunities for misregulation in response to perturbation.

How any particular perturbation distorted signaling output was highly dependent on network configuration. This was most obvious from comparing the effect of Ras G12V perturbation on GAP-free and high-GAP networks (see Fig. 3), but is also readily apparent from inspection of the structure of our experimental maps between network configuration and signaling outputs (see Fig. 5). Even networks configurations that produced highly similar output behaviors could nonetheless respond divergently to perturbations. These are network configurations along a contour in the phase diagrams and represent neutral paths along which a cell or species can move without immediate consequence to the system. For example, a low signaling output that is maintained by a weak GEF activity alone might also be produced by a higher GEF activity balanced by a high GAP activity. However, these two configurations would respond very differently to substitution with oncogenic alleles of Ras like G12V (Fig. 8D). This observation demonstrates the limited predictive power of static steady-state measurements of cellular states and highlights the need to obtain *dynamic* data about the pre-steady state and impulse-response behavior of cellular systems using fine-grained time courses or new methods such as optogenetic pathway activation (*46*).

The dependence of a perturbation on network configuration can also afford cells new opportunities that may be positive rather than deleterious. As an example, our analysis of the feedback gain produced by the RBD-GEF fusion revealed that some configurations had almost no impact on the system output, while others were highly impacted and produced little to no signal without the presence of feedback (Fig. 7B). Thus, access to certain regions of this space is

not feasible without first acquiring permissive modifications to the feedback architecture of the system. Acquisition of this permissive architecture though, can occur in regions in which there is minimal consequence to the system output. Once this feedback mechanism is present, the space of behaviors available to the system changes and previously non-functional regions of the space can now be accessed (Fig. 8D).

By interrogating the space of available behaviors to a signaling system in an unbiased way as we have in the present work, we learn not only what the behavior of any particular system configuration is, but also how systems respond to *change* and what paths exist to travel to new configurations with new behaviors. For Ras, this space appears rich with dynamic possibilities and sufficient neutral network structure to provide evolution with ample fodder to facilitate the use of Ras for the wide array of diverse signaling roles at it plays across different cell types and species, but at the risk of harmful perturbation by diseased alleles or expression states.

Building signaling output programs by coupling dynamic RasGTP levels to effector assembly.

One striking observation from this work was the importance of effector molecules in determining how a dynamic RasGTP signal is perceived. This is, in fact, a critical aspect of how these particular signaling systems work as activated Ras itself has no *enzymatic* activity towards other molecules, but instead serves only as a platform for the recruitment of many possible effector molecules within the cell. Moreover, activated Ras cannot engage more than one effector

simultaneously and thus competition between effectors as well as upstream regulators like GAPs contributes to the system's output dynamics.

For simple one-effector systems that we studied with the C-Raf RBD effector, the concentration of effector shaped not only the amplitude of the output, but also the dynamics of that output as well (see Fig. 4D). Effector concentration is thus more than a passive "volume" knob that interprets RasGTP levels and instead is an active system component that sculpts effector-specific output dynamics.

The importance of effectors was even more apparent in two-effector Ras signaling systems, in which we found that equivalent amounts of different effectors interpreted the same system inputs with markedly different outputs that differed both in amplitudes as well as in duration and dynamics (see Fig. 6). These systems showed a variety of interesting multi-effector programs, such as one effector exhibiting a transient response while another was sustained, one effector responding while another did not at all, or two transient responses that peaked and declined with different timing and duration. Moreover, we found that different dynamic behaviors could readily be produced by introducing point mutations in an effector that altered affinity for activated Ras. These observations stress the important roles that both kinetic and thermodynamic aspects of effector assembly and competition play in shaping how an individual effector interprets dynamic RasGTP levels in the context of the rest of the network.

An interesting consequence of these different effector behaviors and dynamics is that it can naturally result in the temporal partitioning of distinct activities during a signaling response. This can allow some effector outputs to be restricted to early phases of signaling, only to decline and be displaced by other more dominant effectors at later stages. These observations extend recent observations of hierarchies of binding by different effectors to Ras under equilibrium

conditions with non-hydrolyzable analogs (*41*). Thus, the differential perception of RasGTP signals by distinct effectors may not be a flaw in the method by which cells make measurements, but a useful feature by which cells can use a single signaling currency like Ras to execute a programmed sequence of different activities during signaling.

Network level reconstitutions as tools to probe the phenomenology of biochemical signal processing systems.

The network level biochemical approach to interrogating signaling systems we employed in this study occupies a relatively underexplored area in our understanding of cellular decision making systems, but is similar to approaches used to understand dynamic *mechanical* systems in cells like microtubules. Indeed, because additional complexities can emerge when multiple energetically driven processes are coupled together to promote the dynamic assembly and disassembly of competing effectors, exploring how these systems behave *in vitro* under different configurations sheds new light on the phenomenology of how this biochemical signaling device functions and responds to perturbation.

The ability to prepare non-equilibrium steady states in which Ras is actively cycling between ON and OFF states may also prove useful in developing new strategies to ameliorate erroneous signaling associated with diseased states. For example, the fact that we could reconstitute radically different signaling behaviors for wildtype and G12V Ras under high GAP conditions is consistent with the notion that wildtype Ras ordinarily cycles quickly but G12V does not. This difference in the *lifetime* of a RasGTP molecule compared to RasG12V•GTP could potentially serve as an additional *dynamic* selectivity handle for small molecules in which

we want to only target the oncogenic form of Ras. The assays we developed are well suited to compare how small molecules differentially impact signaling through these different forms of Ras side by side under active energy-consuming conditions.

More generally, the simplicity of the approach we present here paves the way for further studies on other types of non-equilibrium signaling systems that center around the assembly of molecules from the cytoplasm on a surface such as the other members of the Ras superfamily like Rho, Rac, Cdc42 and Rab GTPases, as well as other completely different multi-turnover signaling systems like receptor tyrosine kinases. Some aspects of the H-Ras system may be shared with these systems, while other aspects may be different owing to idiosyncratic features of particular molecules. These systems could also be extended in other ways as well to explore how other biophysical constraints impact these signaling processes. Our reconstitutions could, for example, be extended to lipid-coated beads to explore how membrane fluidity or lipid identity shape effector outputs. Multi-currency networks that include multiple Ras isoforms or contain more than one type of GTPase could also be examined to look at higher-order networks and cascades. Only by building these systems, turning them on and watching them run can we begin to understand how they actually perform and operate in different signaling regimes.

Methods.

Protein Purification

Purification, labeling, and nucleotide loading of GTPases

Full-length H-Ras, H-Ras(G12V), H-Ras (G12C), and H-Ras(Q61L) were expressed as N-terminal SNAPtag-$(GAGS)_{2x}$-MBP C-terminal DoubleHisTag (10xHis-$(GAGS)_{3x}$-6xHis) fusion

proteins using custom expression plasmids (see plasmids table). The SNAP-tag facilitated labeling with high-performance inorganic dyes for imaging, and the DoubleHisTag on the C-terminus allowed Ni-NTA supports to be loaded stably ($t_{1/2}$ >24 hours) with GTPase in a configuration resembling the native C-terminal attachment mode.

To express protein, BL21(T1R) *E. coli* cells were grown to an OD of 0.4 from a fresh transformation, chilled to 18°C, induced with 0.8 mM IPTG, and allowed to express overnight. The proteins were purified by Ni-NTA affinity chromatography following the manufacturer's instructions, but were eluted in the presence of 1M Imidazole to facilitate elution of the DoubleHisTag. Proteins were subsequently purified by amylose affinity chromatography per the manufacturer's guidelines. The protein was then concentrated to ~0.5 mL and purified by gel-filtration chromatography on a Superdex S200 10/300 equilibrated in Nucleotide Exchange Buffer (5 mM EDTA, 20 mM Tris, 150 mM NaCl, pH 8.0) to remove any bound nucleotide from the GTPase.

Labeling of the SNAPtag on the GTPases was performed per the manufacturer's instructions. Briefly, samples were buffer-exchanged into a labeling buffer (150 mM NaCl, 25 mM Tris, pH 7.5) using a zebra desalting column. Following twenty minute incubation with 5 mM DTT at 37°, SNAP-Cell 430 substrate was added at a 1.1:1 dye:protein molar ratio and incubated for 1 hour at 37°C or overnight at 4°C. Unlabeled dye was removed by passing the sample through 4 zebra desalting columns.

The labeled GTPases were loaded with nucleotide using established protocols (*27*). Briefly, GTPases were exchanged into Nucelotide Exchange Buffer using a zebra desalting column and incubated with a 20-fold molar excess of nucleotide (typically GDP) for thirty minutes at room temperature. The loading reaction was quenched by the addition of $MgCl_2$ to 10

mM. Unloaded nucleotide was removed by passing the sample through 4 zebra desalting columns. Samples were finally exchanged into GTPase storage buffer (20 mM Tris, 50 mM NaCl, 10 mM $MgCl_2$, 5% glycerol, pH 7.5), concentrated to ~100 uM, alliquotted, flash-frozen, and stored at -80°C.

Purification and labeling of effectors.

The RBD effector domains of A-Raf, B-Raf, C-Raf and any associated mutants were expressed as N-terminal SNAPtag-$(GAGS)_{2x}$-MBP C-terminal Strep-II tag fusions. The SNAP-tag facilitated labeling with high-performance inorganic dyes for imaging, and the Strep-II tag provided a handle for affinity chromatography that did not interact with Ni-NTA supports.

To express protein BL21(T1R) *E. coli* cells were grown to an OD of 0.8 from a fresh transformation, chilled to 18°C, induced with 0.8 mM IPTG, and allowed to express overnight. The proteins were purified by Streptactin affinity chromatography and then amylose affinity chromatography, per the manufacturer's instructions. Proteins were concentrated to ~1 mL and further purified by gel-filtration chromatography on a Superdex S75 16/60 equilibrated in standard protein buffer (150 mM NaCl, 25 mM Tris pH 7.5).

Labeling of the SNAPtag on the effectors was performed per the manufacturer's instructions. Briefly, samples were incubated with 5 mM DTT at 37°C for twenty minutes. SNAP-Surface 488 or SNAP-Surface 549 were added at a 1.1:1 dye:protein molar ratio and incubated for 1 hour at 37 or overnight at 4 degrees. Unlabeled dye was removed by passing the sample through 4 zebra desalting columns. Samples were concentrated to 100 uM, exchanged

into storage buffer (150 mM NaCl, 25 mM Tris pH 7.5, 5% glycerol), aliquotted, flash frozen and stored at -80°C.

Purification of GEFs, GAPs, and synthetic effector-GEF fusions.

The catalytic domains of RasGRF, p120GAP, NF1-GAP, SOS and RBD-RasGRF were expressed as N-terminal MBP C-terminal Strep-II tag fusions. The Strep-II tag provided a handle for affinity chromatography that did not interact with Ni-NTA supports.

To express protein BL21(T1R) *E. coli* cells were grown to an OD of 0.4 from a fresh transformation, chilled to 18°C, induced with 0.8 mM IPTG, and allowed to express overnight. The proteins were purified by Streptactin affinity chromatography and then amylose affinity chromatography, per the manufacturer's instructions. Proteins were concentrated to ~0.5 mL and purified by gel-filtration chromatography on a Superdex S200 10/300 equilibrated in standard protein storage buffer (150 mM NaCl, 25 mM Tris pH 7.5, 5% glycerol). Proteins were concentrated, aliquotted, flash frozen, and stored at -80°C

In Vitro Signal Processing Assays

Preparation of GTPase loaded beads

50 uL of NiSepharose High Performance beads (GE Healthcare) were washed twice with 1 mL water, twice with GTPase assay buffer (GAB: 20 mM Tris, 50 mM NaCl, 10 mM $MgCl_2$, 30 mM Imidazole), and resupended in a final volume of 1 mL GAB. To load beads, 7.5 μL of this bead slurry was mixed with 7.5 μL of GTPase in a PCR tube and incubated on ice for 1 hour with

occasional flicking. The amount of GTPase used to load the beads depending on the desired density for the specific experiments being performed downstream, but a typical bead-loading used 7.5 μL of 30 μM GTPase. Following incubation, the beads were spun in a table-top minifuge, the supernatant removed, and washe`d 3 times with GAB. The washed beads were resuspended in ~100 μL GAB, transferred to an Eppendorf tube, shielded from light, and stored on ice. The exact amount of final GAB the beads were in was adjusted for any particular experiment such that 2 μL of bead slurry contained roughly 25-100 total beads when placed in a 384 well microscopy plate.

Preparation of signal processing reactions and data collection

Signal processing reactions were set up in two stages. First, a "bead-mix" was prepared that contained fluorescent effector at the desired concentration (typically 50 nM) and beads in GAB. It should be noted that the inclusion of 20 mM Imidazole was critical for eliminating non-specific background effector staining on the bead surface and improved reproducibility dramatically. 20 μL of this bead mix was dispensed into the wells of a 384-well Costar microscopy plate. Second, an "initiation-mix" was prepared that contained fluorescent effector at the desired concentration, 5 mM GTP (or other nucleotide if used) and GEFs and GAPs at the desired concentration, all in GAB. 10 μl of this reaction mix was gently added to the 20 μl bead mix in the 384-well plate to initiate reaction. The large volume of the initiation mix was critical for getting sufficient mixing without the need to pipette up and down and disrupt the beads and improved reproducibility. Once signal processing reactions were initiated, they wells were sealed with PCR plate sealant to prevent evaporation.

All data was collected on a Nikon Eclipse TI inverted microscope equipped with a Yokogawa CSU-X1 spinning disk confocal using a 20x PlanAPO 0.75 NA objective and an electron microscopy charge-coupled deivce (EM-CCD) camera (Andor). Depending on the experiment, 405 nm, 488 nm, and/or 561 nm wavelength laser light (LMM5, Spectral applied Research) were used for excitation.

For typical experiments, 5-10 x-y positions within a given well were used to collect signal processing behaviors from 20-100 individual beads. Timepoints varied depending on the experiment, but for typical large matrix experiments of 24-60 different GEF/GAP/effector conditions, we typically imaged every 15 minutes for 6-12 hours. MicroManager software was used to design the imaging protocols and collect the actual data.

Image Analysis

A combination of standard and custom ImageJ macros were used to prepare the primary image data for further analysis. First, drift in the stage was corrected using a macro based around the MultiStackReg plugin. Two or three color multi-tiff timecourses were split into separate channels. Matrix transformations to register timecourses were obtained using the constant GTPase fluorescence on every bead from the blue channel. These matrices were then used to register the timecourses of the red or green channels. The three channels were then recombined to produce the properly registered multi-tiff used for analysis of the beads. The ImageJ macro code was tweaked depending on the particular experiment, but a representative example of the code is shown below:

```
matrixpath= "C:\\Users\\scoyl_000\\Google Drive\\MICROSCOPY\\150403\\matrix\\";
inputpath = "C:\\Users\\scoyl_000\\Google Drive\\MICROSCOPY\\150403\\raw\\";
```

```
outputpath = "C:\\Users\\scoyl_000\\Google Drive\\MICROSCOPY\\150403\\processed\\";


function scott_register(input,output,filename) {
		open(input+filename);
		run("Split Channels");
		run("MultiStackReg", "stack_1=[C3-" + filename + "] action_1=Align file_1=["+matrixpath+"matrix.txt] stack_2=None action_2=Ignore
file_2=[] transformation=[Rigid Body] save");
		run("MultiStackReg", "stack_1=[C1-" + filename + "] action_1=[Load Transformation File] file_1=["+matrixpath+"matrix.txt]
stack_2=None action_2=Ignore file_2=[] transformation=[Rigid Body]");
		run("MultiStackReg", "stack_1=[C2-" + filename + "] action_1=[Load Transformation File] file_1=["+matrixpath+"matrix.txt]
stack_2=None action_2=Ignore file_2=[] transformation=[Rigid Body]");
		run("Merge Channels...", "c1=C1-" + filename + " c2=C2-" + filename + " c3=C3-" + filename + " create");
}

list = getFileList(inputpath);
for (i = 0; i < list.length; i++)
	scott_register(inputpath, outputpath, list[i]);
```

We then corrected for uneven sample illumination or background artifacts using a rolling ball background substraction of 50 pixels. These processed images were subsequently used to analyze the signaling behavior of each bead.

To analyze the images, beads were either identified and stored as regions-of-interest (ROIs) automatically from the GTPase fluorescent signal on beads using a custom macro that makes use of ImageJ's "find particles" function, or in the case of experiments with beads that deliberately contain differing levels of Ras density, beads were identified and stored as ROIs by hand. The particular parameters of the automated bead finding varied depending on the particular experiment, but a representative ImageJ macro is shown below:

```
inputpath = "C:\\Users\\scoyl_000\\Google Drive\\MICROSCOPY\\150403\\processed\\";
outputpath = "C:\\Users\\scoyl_000\\Google Drive\\MICROSCOPY\\150403\\beads\\";
```

```
function scott_findbeads(input,output,filename) {

		open(input+filename);
		run("Split Channels");

		selectWindow("C3-" + filename);
		run("Smooth", "stack");

		//set threshold and find particles
		Stack.setPosition(1, 18,1);
		setSlice(20);
		setThreshold(3744, 23285);
		run("Analyze Particles...", "size=400-15000 circularity=0.60-1.00 display exclude clear include add slice");

		//close any and all open windows
		close();
		close();
		close();


		//reopen original file
		open(input+filename);

		//transfer ROIs to overlay and save
		run("From ROI Manager");
		saveAs("Tiff", output+filename);

		//clear manager
		roiManager("Delete");
		close();
}

list = getFileList(inputpath);
for (i = 0; i < list.length; i++)
	scott_findbeads(inputpath, outputpath, list[i]);
```

Once beads were identified and stored as ROIs, a variety of measurements were made for each bead using a custom macro. We measured the average total Ras fluorescence in the ROI, and the average fluorescent effector signal in the ROI at every timepoint in the experiment. We also measured the area and perimeter of the bead, such that we could convert from average fluorescent to total fluorescence (Total=Average*Area) and then project that total onto the perimeter of the bead surface (Surface=Total/Perimeter). This had only modest effects on the signaling traces that we measured, but did reduce variation between beads that stemmed from some beads having different diameters than others. As with other macros, the exact details of the code varied depending on the number of effectors we were simultaneously examining or other aspects of the setup, but a representative ImageJ macro is shown below:

```
inputpath = "C:\\Users\\scoyl_000\\Google Drive\\MICROSCOPY\\150318\\one_shot_processed\\";
outputpath = "C:\\Users\\scoyl_000\\Google Drive\\MICROSCOPY\\150318\\one_shot_data\\";

setBatchMode(true);
list = getFileList(inputpath);
for (i = 0; i < list.length; i++)
     crunchimage(inputpath, outputpath, list[i]);
setBatchMode(false);

//The Functions that are used in the macro are below

function crunchimage(input,output,filename) {

                    //open the image
                    open(input+filename);

                    //clear log file and reulsts table
                    run("Clear Results");
                    print("\\Clear");

                    //import ROIs from overlay
                    run("To ROI Manager");
```

```
//count ROis
count=roiManager("count");

//record Ras AMPS and RBD timecourse for each ROI
for(i=0;i<count;i++){

	// first record the AMP for the Ras field;

	recordAMP(i);

	// next calculate the time series for the data

	recordTimecourse(i);

	// print newline marker

	print("!");
}

selectWindow("Log");
saveAs("Text", output+filename);
close();


}

function recordAMP(index){
	//function that reports area, mean intensity, and perimieter
	// of an ROI used to get everything EXCEPT the timecourse data

	run("Clear Results");
	roiManager("Select",index);

	// select channel 2 and select the midpoint of hte stack

	Stack.setPosition(2, 18,18);

	// make measurements
```

```
		run("Measure");
		BeadArea=getResult("Area",0);
		BeadMean=getResult("Mean",0);
		BeadPerim=getResult("Perim.",0);
		print(BeadArea+","+BeadMean+","+BeadPerim+",");

}

function recordTimecourse(index){
		run("Clear Results");
		roiManager("Select",index);
		sliceCount=nSlices()/2;
		// print(sliceCount);

		for(k=0;k<sliceCount;k++){
			Stack.setPosition(1, 18,k+1);
			run("Measure");
			timepointK=getResult("Mean",k);
			print(timepointK+",");
		}


}
```

The output of this macro is a file that contains a list of every single bead trace in the multi-tiff image. Each trace begins with the area measurement of the bead, the mean total Ras intensity of the bead, and the perimeter measurement of the bead, followed by the effector measurement of the bead at every timepoint. Each measurement ends with a “,” and is on a newline. At the end of each bead trace, a stop marker “!” is printed. These data are then transformed by GREP and shell script into a CSV where each line contains all the relevant information about each single bead’s signaling trace. These data can then be loaded into either Matlab or Excel for further analysis.

Once in Excel, data were typically further analyzed as follows: (i) intensity measurements were normalized to perimeter instead of area, (ii) the time-series data for a given bead was normalized such that the time-zero effector measurement was zero, (iii) single bead traces were binned based on total Ras-GTPase levels to obtain statistics on the signaling behavior.

Estimation of Ras Density

We estimated the approximate Ras density of a bead in molecules × um^{-2} in the following manner. First we determined the correspondence between the concentration of labeled Ras and its fluorescence intensity by imaging serial dilutions of known concentrations of SNAP-Cell 430 labeled Ras in solution. From this we could associate a particular fluorescence intensity with a three-dimensional concentration in μM. For our imaging conditions, this relationships was:

$$\text{Concentration } (\mu M) = 4.287\ (\mu M \times AU^{-1}) \times \text{Intensity (AU)}$$

For any individual bead then, there is some maximum fluorescence intensity on the bead surface that we can associate with an apparent three-dimensional concentration. This apparent-three dimensional concentration can be used to estimate an inferred two-dimensional density by assuming that the apparent three-dimensional concentration is a consequence of molecules within some two-dimensional area exploring three-dimensional space as constrained by some confinement length *h* (47,48). Our estimation of *h* was based on previous work of tethering EGFR kinases with His-tags to a DGS-NTA(Ni) charged vesicle surface, which assumed a radius of confinement of 55 nm (48). Given that 1 μM corresponds to ~602 molecules / μm^3, this

enabled us to convert between three-dimensional and two-dimensional concentrations using the relation:

$$1\ \mu M \times (602\ (\text{molecules} \times \mu m^{-3}) \times \mu M^{-1}) \times 0.055\ \mu m = 33.1\ \text{molecules}\ \mu m^{-2}$$

$$\Rightarrow \quad \text{2-D Concentration (molecules} \times \mu m^{-2}) =$$

$$\text{3-D concentration } (\mu M) \times 33.1\ ((\text{molecules} \times \mu m^{2}) \times \mu M^{-1}))$$

We stress that this is only an *estimate* of the Ras density and should not be taken as a highly accurate assessment of the Ras density. Nonetheless, it provides a crude estimation that indicates that our experiments are not operating in a highly non-physiologic regime. Importantly, for our analysis the exact number of Ras molecules on the bead surface is not critical. Indeed, the relative abundances of Ras on different beads is more important as it enables us to compare behaviors between beads as Ras densities change.

Comparison of Network Configurations Across Cell and Tissue Types.

Relative log-transformed expression levels for p120GAP, C-Raf, and H-Ras across a variety of cell-types and tissue-types were obtained from data contained within the Genevestigator software package (see data in Table 1). The three-dimensional phenotypes associated with each cell or tissue type was plotted as a 3D scatterplot using Matlab.

References.

1. F. Chang *et al.*, Regulation of cell cycle progression and apoptosis by the Ras/Raf/MEK/ERK pathway (Review). *Int. J. Oncol.* **22**, 469–480 (2003).

2. A. Sjölander, K. Yamamoto, B. E. Huber, E. G. Lapetina, Association of p21ras with phosphatidylinositol 3-kinase. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 7908–7912 (1991).

3. F. Hofer, S. Fields, C. Schneider, G. S. Martin, Activated Ras interacts with the Ral guanine nucleotide dissociation stimulator. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 11089–11093 (1994).

4. A. B. Vojtek, C. J. Der, Increasing Complexity of the Ras Signaling Pathway. *J. Biol. Chem.* **273**, 19925–19928 (1998).

5. H. R. Bourne, D. A. Sanders, F. McCormick, The GTPase superfamily: a conserved switch for diverse cell functions. *Nature*. **348**, 125–132 (1990).

6. J. L. Bos, ras oncogenes in human cancer: a review. *Cancer Res.* **49**, 4682–4689 (1989).

7. S. Schubbert, K. Shannon, G. Bollag, Hyperactive Ras in developmental disorders and cancer. *Nat. Rev. Cancer*. **7**, 295–308 (2007).

8. U. Krengel *et al.*, Three-dimensional structures of H-ras p21 mutants: Molecular basis for their inability to function as signal switch molecules. *Cell*. **62**, 539–548 (1990).

9. M. V. Milburn *et al.*, Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins. *Science*. **247**, 939–945 (1990).

10. N. Nassar *et al.*, The 2.2 Å crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf1 in complex with RaplA and a GTP analogue. *Nature*. **375**, 554–560 (1995).

11. C. Herrmann, Ras–effector interactions: after one decade. *Curr. Opin. Struct. Biol.* **13**, 122–129 (2003).

12. S. E. Neal, J. F. Eccleston, A. Hall, M. R. Webb, Kinetic analysis of the hydrolysis of GTP by p21N-ras. The basal GTPase mechanism. *J. Biol. Chem.* **263**, 19718–19722 (1988).

13. J. B. Gibbs, I. S. Sigal, M. Poe, E. M. Scolnick, Intrinsic GTPase activity distinguishes normal and oncogenic ras p21 molecules. *Proc. Natl. Acad. Sci.* **81**, 5704–5708 (1984).

14. J. P. McGrath, D. J. Capon, D. V. Goeddel, A. D. Levinson, Comparative biochemical properties of normal and activated human ras p21 protein. *Nature*. **310**, 644–649 (1984).

15. M. S. Boguski, F. McCormick, Proteins regulating Ras and its relatives. *Nature*. **366**, 643–654 (1993).

16. M. Trahey, F. McCormick, A cytoplasmic protein stimulates normal N-ras p21 GTPase, but does not affect oncogenic mutants. *Science*. **238**, 542–545 (1987).

17. J. L. Bos, H. Rehmann, A. Wittinghofer, GEFs and GAPs: Critical Elements in the Control of Small G Proteins. *Cell*. **129**, 865–877 (2007).

18. F. McCormick, G. A. Martin, R. Clark, G. Bollag, P. Polakis, Regulation of ras p21 by GTPase Activating Proteins. *Cold Spring Harb. Symp. Quant. Biol.* **56**, 237–241 (1991).

19. C. Lenzen, R. H. Cool, H. Prinz, J. Kuhlmann, A. Wittinghofer, Kinetic Analysis by Fluorescence of the Interaction between Ras and the Catalytic Domain of the Guanine Nucleotide Exchange Factor Cdc25Mm. *Biochemistry (Mosc.)*. **37**, 7420–7430 (1998).

20. H. Sondermann *et al.*, Structural Analysis of Autoinhibition in the Ras Activator Son of Sevenless. *Cell*. **119**, 393–405 (2004).

21. S. M. Margarit *et al.*, Structural Evidence for Feedback Activation by Ras·GTP of the Ras-Specific Nucleotide Exchange Factor SOS. *Cell*. **112**, 685–695 (2003).

22. J. S. Iwig *et al.*, Structural analysis of autoinhibition in the Ras-specific exchange factor RasGRP1. *eLife*. **2**, e00813 (2013).

23. K. Scheffzek *et al.*, The Ras-RasGAP Complex: Structural Basis for GTPase Activation and Its Loss in Oncogenic Ras Mutants. *Science*. **277**, 333–339 (1997).

24. P. A. Boriack-Sjodin, S. M. Margarit, D. Bar-Sagi, J. Kuriyan, The structural basis of the activation of Ras by Sos. *Nature*. **394**, 337–343 (1998).

25. L. Feng *et al.*, PKA phosphorylation and 14-3-3 interaction regulate the function of neurofibromatosis type I tumor suppressor, neurofibromin. *FEBS Lett.* **557**, 275–282 (2004).

26. G. Bollag, F. McCormick, Differential regulation of rasGAP and neurofibromatosis gene product activities. *Nature*. **351**, 576–579 (1991).

27. A. Eberth, M. R. Ahmadian, in *Current Protocols in Cell Biology* (John Wiley & Sons, Inc., 2001; http://onlinelibrary.wiley.com/doi/10.1002/0471143030.cb1409s43/abstract).

28. M. Geyer *et al.*, Conformational Transitions in p21ras and in Its Complexes with the Effector Protein Raf-RBD and the GTPase Activating Protein GAP. *Biochemistry (Mosc.)*. **35**, 10308–10320 (1996).

29. C. Herrmann, G. Horn, M. Spaargaren, A. Wittinghofer, Differential Interaction of the Ras Family GTP-binding Proteins H-Ras, Rap1A, and R-Ras with the Putative Effector Molecules Raf Kinase and Ral-Guanine Nucleotide Exchange Factor. *J. Biol. Chem.* **271**, 6794–6800 (1996).

30. J. R. Sydor, M. Engelhard, A. Wittinghofer, R. S. Goody, C. Herrmann, Transient Kinetic Studies on the Interaction of Ras and the Ras-Binding Domain of c-Raf-1 Reveal Rapid Equilibration of the Complex. *Biochemistry (Mosc.)*. **37**, 14292–14299 (1998).

31. I. Rubio *et al.*, TCR-Induced Activation of Ras Proceeds at the Plasma Membrane and Requires Palmitoylation of N-Ras. *J. Immunol.* **185**, 3536–3543 (2010).

32. H. Murakoshi *et al.*, Single-molecule imaging analysis of Ras activation in living cells. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 7317–7322 (2004).

33. X. Nan *et al.*, Ras-GTP dimers activate the Mitogen-Activated Protein Kinase (MAPK) pathway. *Proc. Natl. Acad. Sci.* **112**, 7996–8001 (2015).

34. T. S. Freedman *et al.*, A Ras-induced conformational switch in the Ras activator Son of sevenless. *Proc. Natl. Acad. Sci.* **103**, 16692–16697 (2006).

35. C. Block, R. Janknecht, C. Herrmann, N. Nassar, A. Wittinghofer, Quantitative structure-activity analysis correlating Ras/Raf interaction in vitro to Raf activation in vivo. *Nat. Struct. Mol. Biol.* **3**, 244–251 (1996).

36. C. Herrmann, G. A. Martin, A. Wittinghofer, Quantitative Analysis of the Complex between p21 and the Ras-binding Domain of the Human Raf-1 Protein Kinase. *J. Biol. Chem.* **270**, 2901–2905 (1995).

37. K. Scheffzek *et al.*, Structural analysis of the GAP-related domain from neurofibromin and its implications. *EMBO J.* **17**, 4313–4327 (1998).

38. G. F. Xu *et al.*, The neurofibromatosis type 1 gene encodes a protein related to GAP. *Cell*. **62**, 599–608 (1990).

39. M. Barbacid, ras Genes. *Annu. Rev. Biochem.* **56**, 779–827 (1987).

40. K. Rajalingam, R. Schreck, U. R. Rapp, Š. Albert, Ras oncogenes and their downstream targets. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **1773**, 1177–1195 (2007).

41. M. J. Smith, M. Ikura, Integrated RAS signaling defined by parallel NMR detection of effectors and regulators. *Nat. Chem. Biol.* **10**, 223–230 (2014).

42. L. Wiesmüller, A. Wittinghofer, Expression of the GTPase activating domain of the neurofibromatosis type 1 (NF1) gene in Escherichia coli and role of the conserved lysine residue. *J. Biol. Chem.* **267**, 10207–10210 (1992).

43. L. Janosi, Z. Li, J. F. Hancock, A. A. Gorfe, Organization, dynamics, and segregation of Ras nanoclusters in membrane domains. *Proc. Natl. Acad. Sci.* **109**, 8097–8102 (2012).

44. S. J. Plowman, C. Muncke, R. G. Parton, J. F. Hancock, H-ras, K-ras, and inner plasma membrane raft proteins operate in nanoclusters with differential dependence on the actin cytoskeleton. *Proc. Natl. Acad. Sci.* **102**, 15500–15505 (2005).

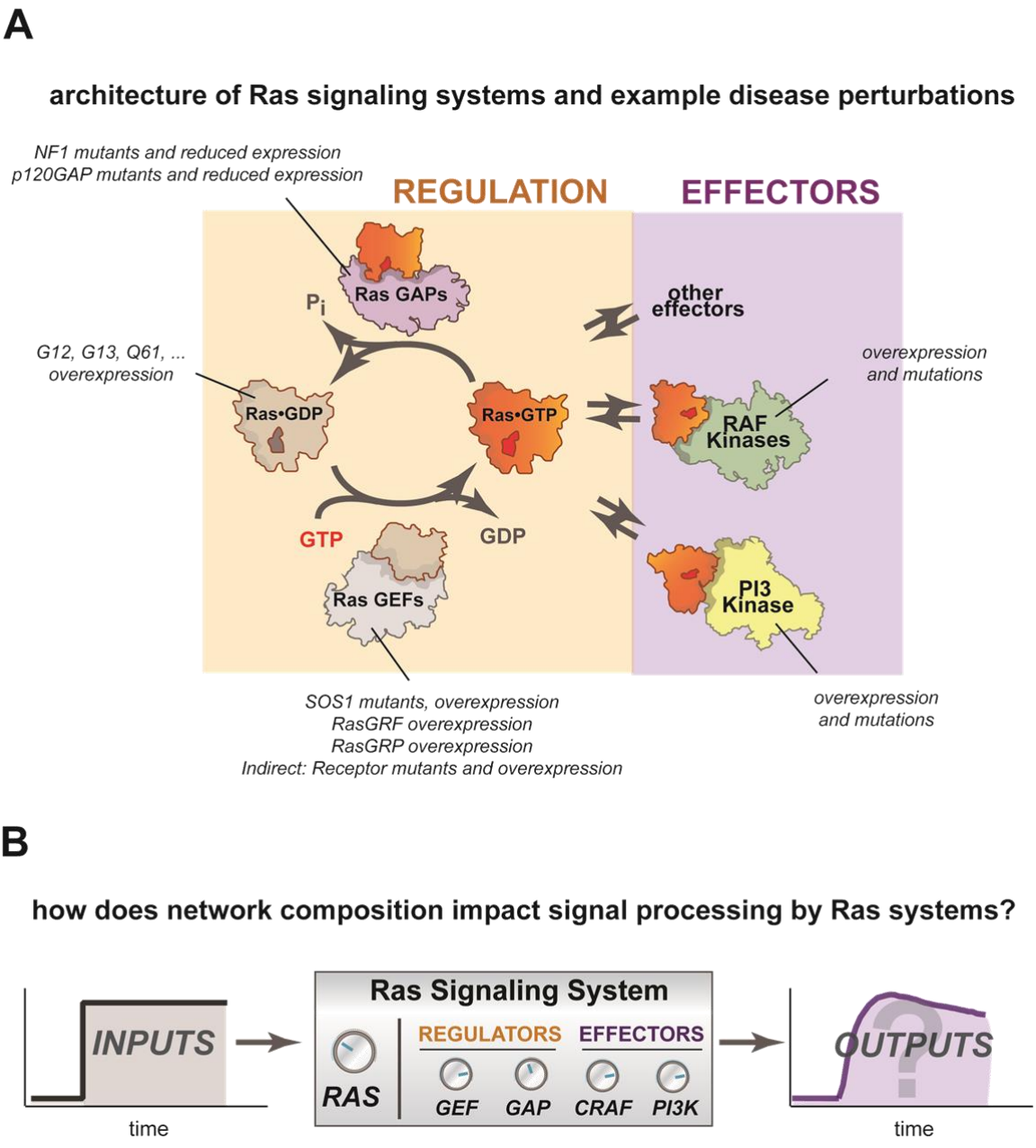
45. G. Bollag *et al.*, Loss of NF1 results in activation of the Ras signaling pathway and leads to aberrant growth in haematopoietic cells. *Nat. Genet.* **12**, 144–148 (1996).

46. J. E. Toettcher, O. D. Weiner, W. A. Lim, Using Optogenetics to Interrogate the Dynamic Control of Signal Transmission by the Ras/Erk Module. *Cell*. **155**, 1422–1434 (2013).

47. X. Zhang, J. Gureasko, K. Shen, P. A. Cole, J. Kuriyan, An Allosteric Mechanism for Activation of the Kinase Domain of Epidermal Growth Factor Receptor. *Cell*. **125**, 1137–1149 (2006).

48. Y. Wu, J. Vendome, L. Shapiro, A. Ben-Shaul, B. Honig, Transforming binding affinities from 3D to 2D with application to cadherin clustering. *Nature*. **475**, 510–513 (2011).

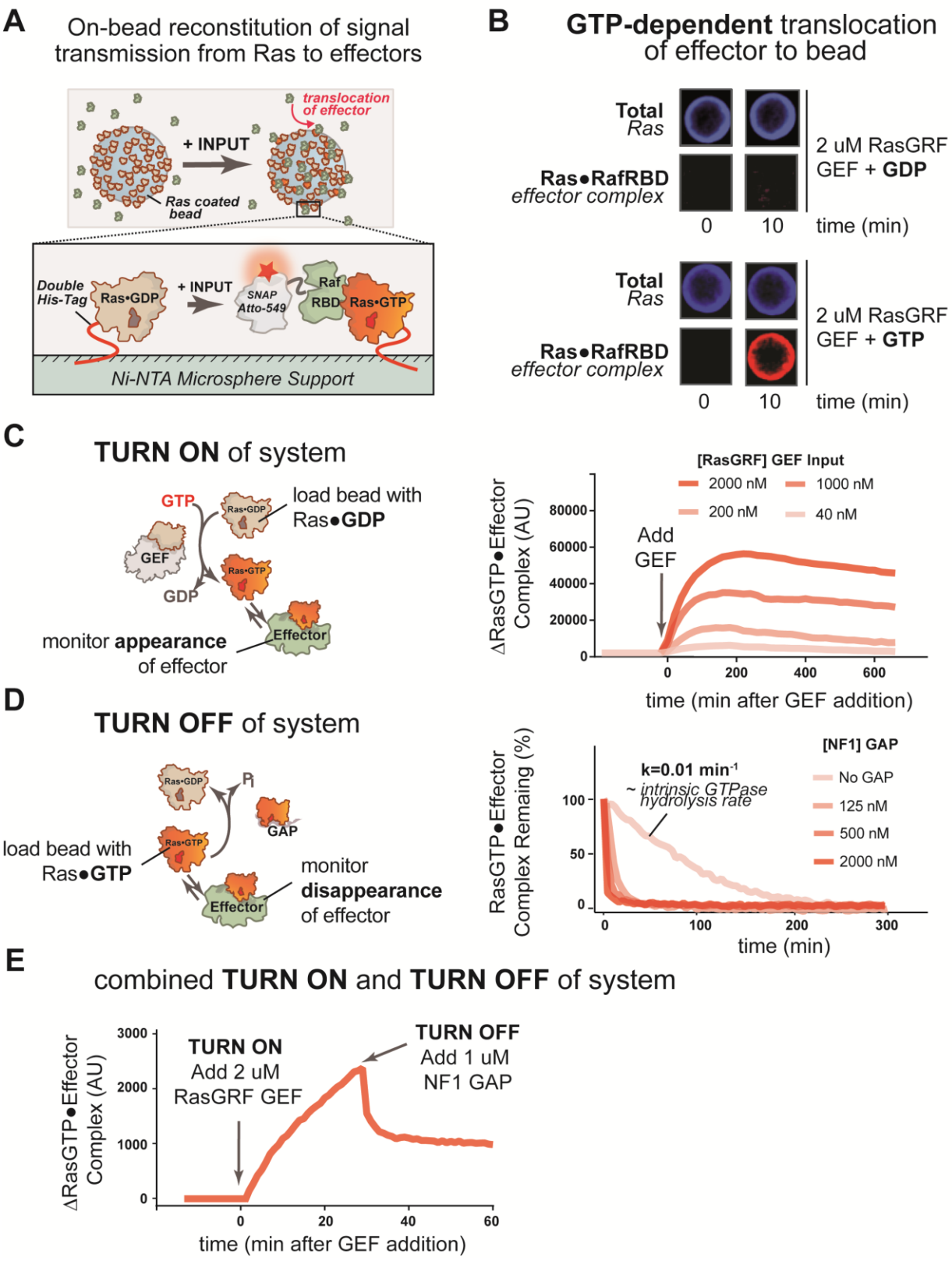
Figure 1. Multiple activities that are frequently perturbed in disease dynamically regulate Ras activity to control the assembly of downstream effectors during signal processing.



(a) Depiction of the proximal architecture of Ras signaling systems and examples of perturbations that are associated with disease. Ras is activated by guanine exchange factors

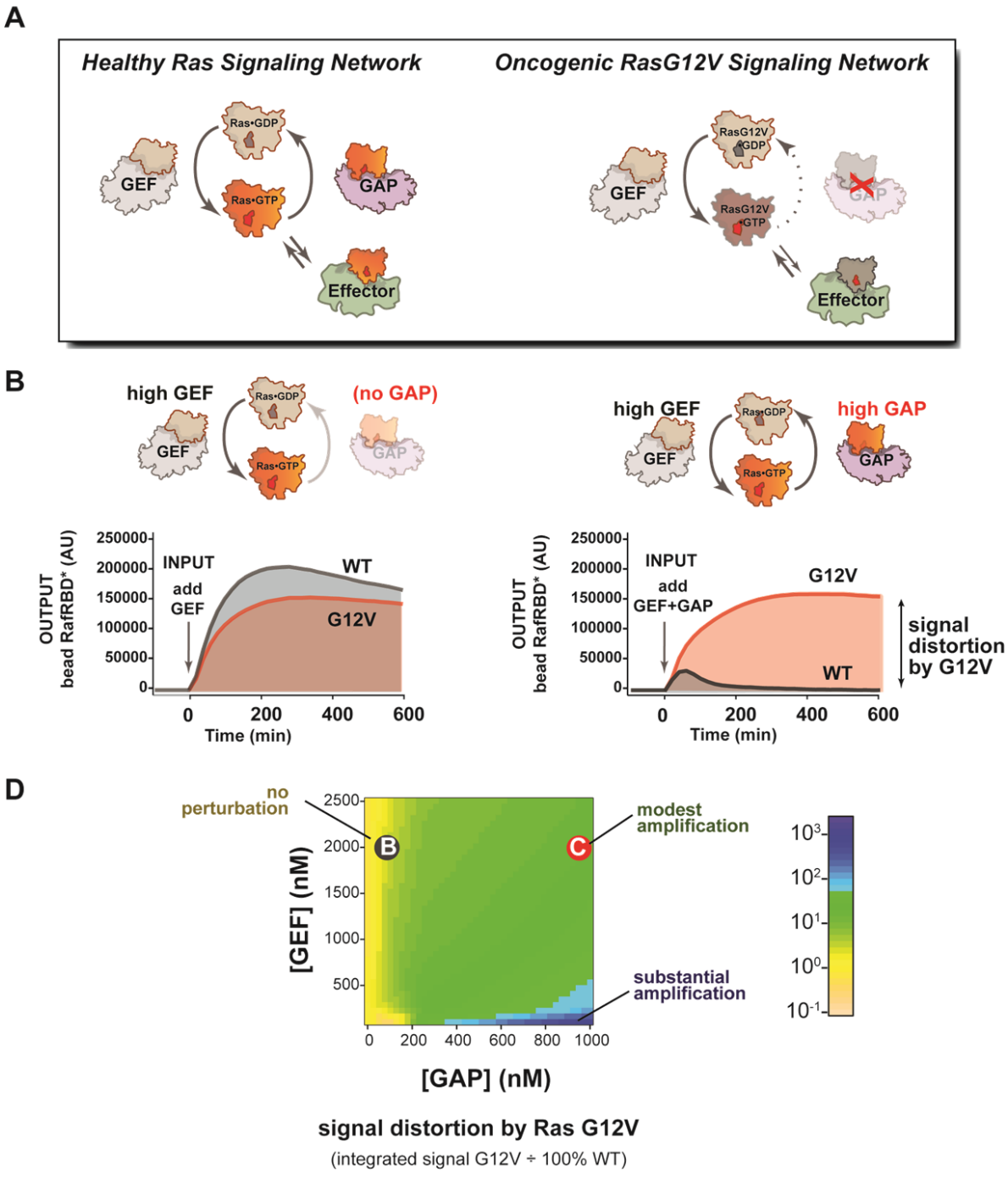
(GEFs) that exchange GDP for GTP and is inactivated by GTPase activating proteins (GAPs) that accelerate the hydrolysis of GTP. Activated Ras interacts with downstream effectors such as Raf or PI3 Kinase to assemble signaling complexes and elicit signaling outputs. (b) Abstraction of the proximal biochemical machinery underlying Ras processing of inputs into outputs, raising the question as to how the network configuration shapes signaling.

Figure 2. A network-level multi-turnover reconstitution of dynamic signal transmission from Ras to downstream effectors.



(a) Bead-based approach used to study how Ras systems assemble effector complexes in response to inputs. By incubating Ni-NTA microspheres that have been loaded with Ras in solutions containing GEFs, GAPs, and fluorescent effectors, system outputs can be observed by monitoring the accumulation of effector on the bead-bound Ras. (b) Example of GEF-catalyzed GTP-dependent translocation of fluorescent effector to Ras loaded bead. The amount of fluorescent effector bound to an individual bead before or after (10 min) addition of of 2 μM GEF +/- either GDP or GTP is shown. (c) Dose dependent signaling response of effector translocation in response to increasing amounts of RasGRF GEF activity. (d) Dose-dependent turn-off of output in the presence of saturating effector and increasing amounts of NF1 GAP activity. (e) Combined turn on and turn off behavior of effector response when the system was activated with 2 μM RasGRF GEF and after 30 min NF1-GAP was added.

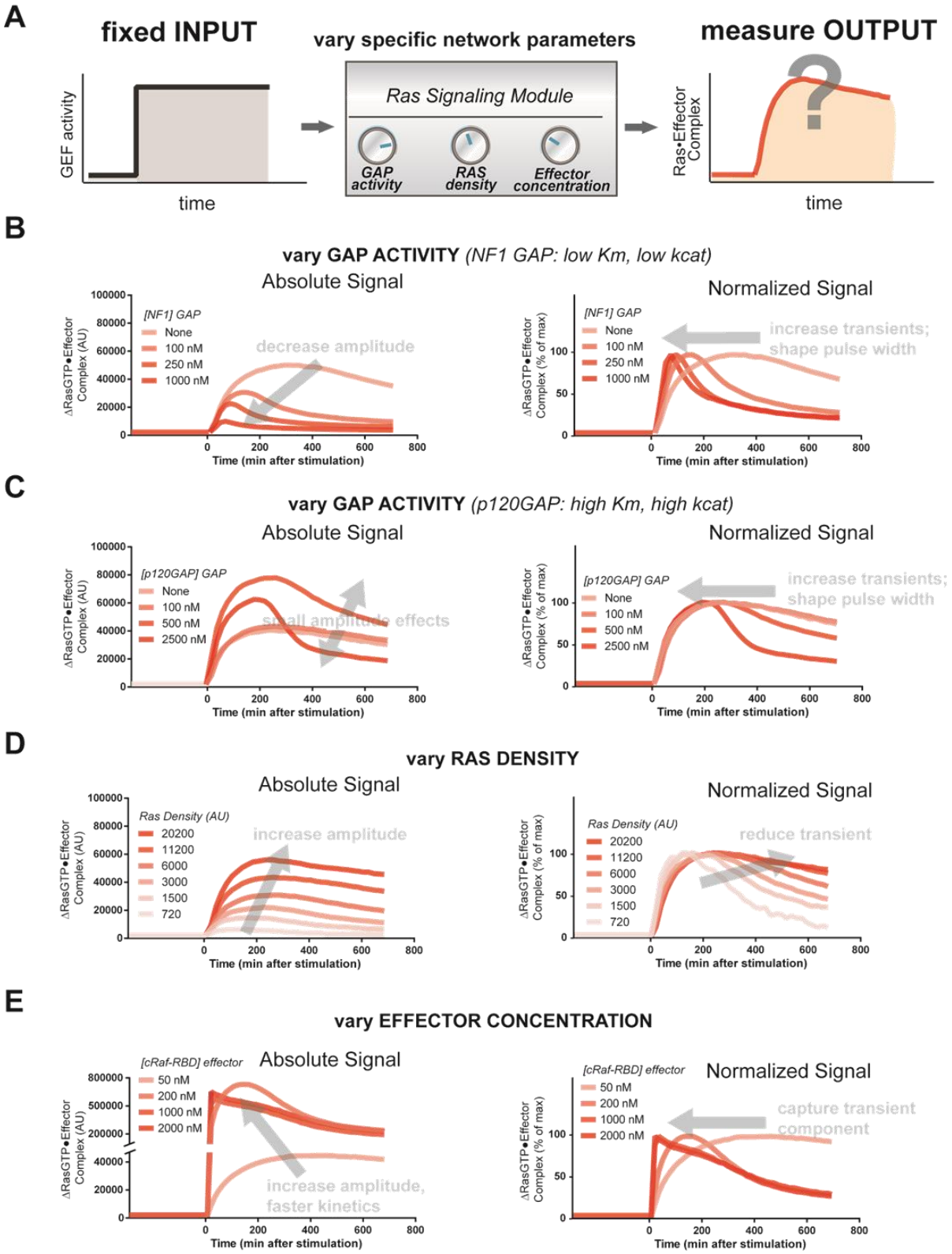
Figure 3. The extent of signal processing distortion by oncogenic alleles of Ras depends on the balance of positive and negative regulatory activities in the network.



(a) Depiction of wildtype Ras and oncogenic G12V Ras illustrating the modes by which mutation is thought to impact the network behavior: changing in intrinsic hydrolysis rate, blocking GAP mediated hydrolysis, and modulating effector interactions. (b) Absolute and (c)

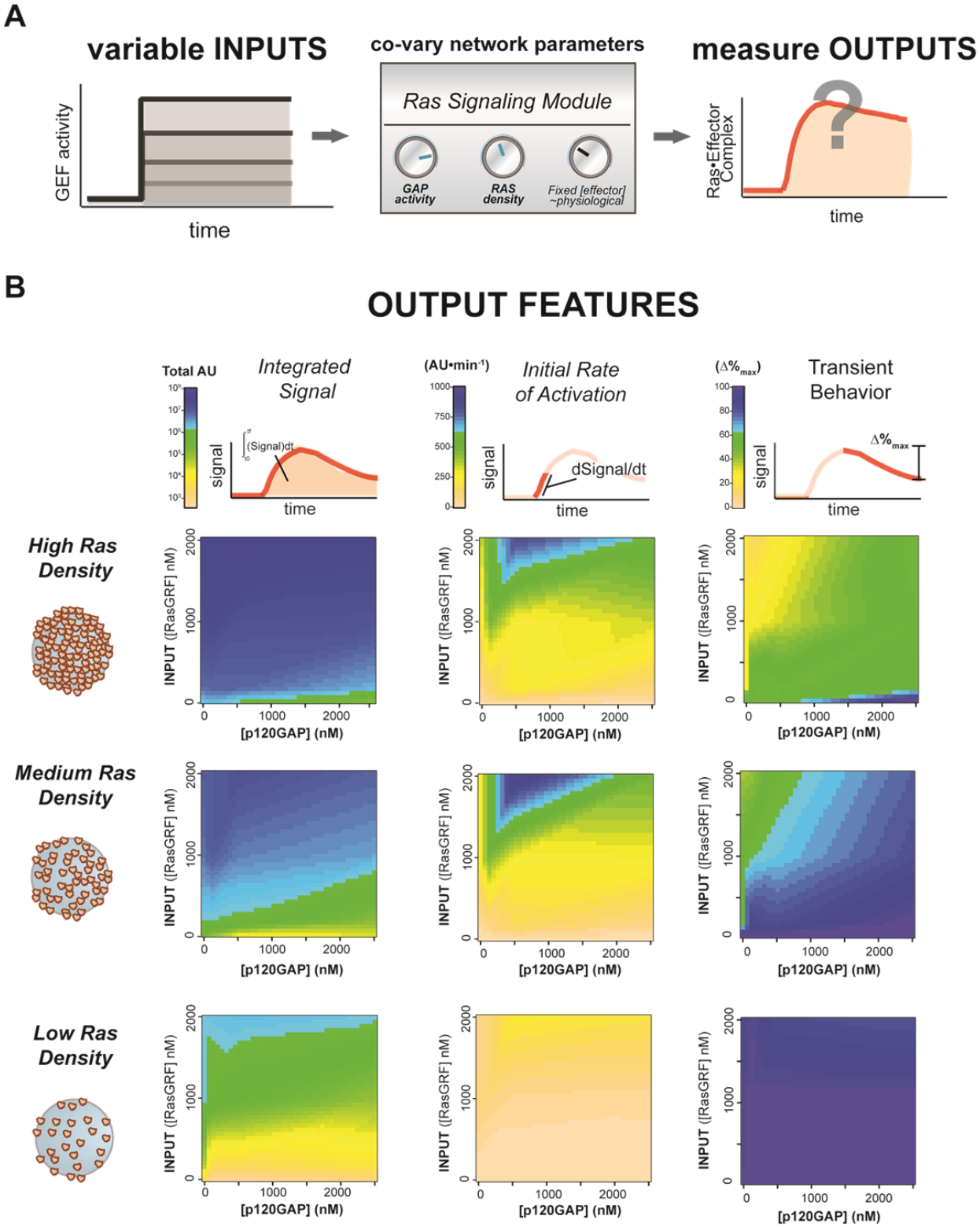
normalized effector responses to a 2 μM RasGRF GEF step input in the absence of any GAP activity. (d) Absolute and (e) normalized responses of the same step input as in (a) and (b), but with 1 μM NF1 GAP activity present in the network. (f) Phase diagram of the magnitude of signal distortion caused by G12V substitution (defined as fold-change in integrated signal of G12V relative to wildtype) in different GEF and GAP network configurations.

Figure 4. The concentration and identity of each Ras network component can modulate the timing, duration, shape, or amplitude of effector outputs.



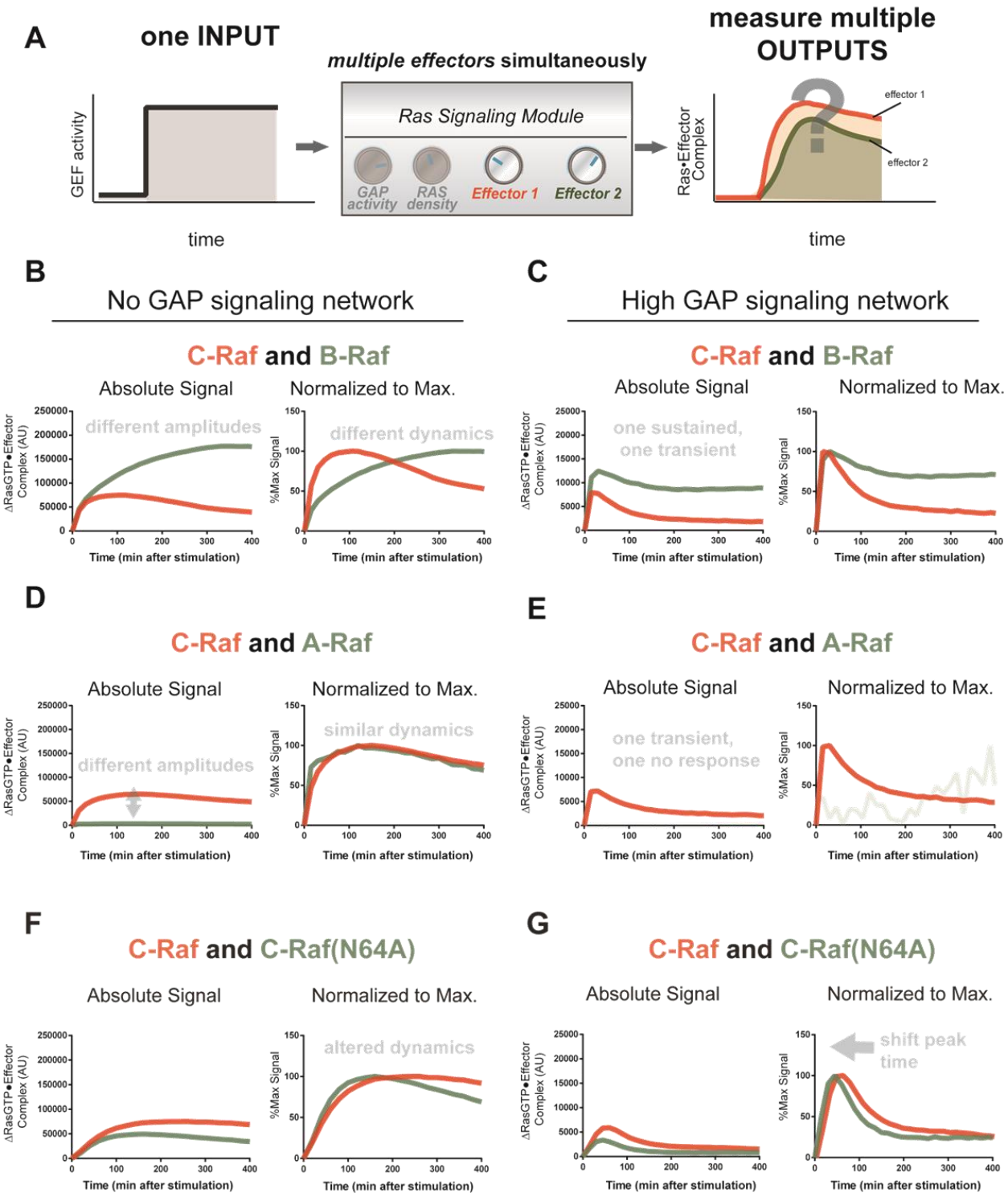
(a) Depiction of the experimental setup: a fixed step-input is applied to a panel of Ras signaling systems in which the concentration of a single network component is varied to determine how each network component individually modulates system output. (b) Absolute and normalized effector responses to step-input in the presence of increasing amounts of the NF1 gap. (c) Absolute and normalized effector responses to step-input in the presence of increasing amounts of the p120 GAP. (d) Absolute and normalized responses to step-input in the presence of different densities of Ras on the bead surface. (e) Absolute and normalized responses to step-input in the presence of increasing amounts of the C-Raf RBD effector.

Figure 5. Tuning the levels of GEF, GAP, and GTPase provide access to a rich and diverse space of possible Ras signal processing behaviors.



(a) Depiction of the experimental setup: a variety of different inputs (changes in apparent GEF activity) are applied to a panel of Ras signaling systems in which both the concentration of GAP and the density of Ras are varied to determine how different inputs are processed into outputs in different network configurations. (b) Phase diagrams for three different output features—integrated signal, initial rate of response, and overshoot behavior—at three different Ras density levels, constructed using the collection of output responses in Fig. S2.

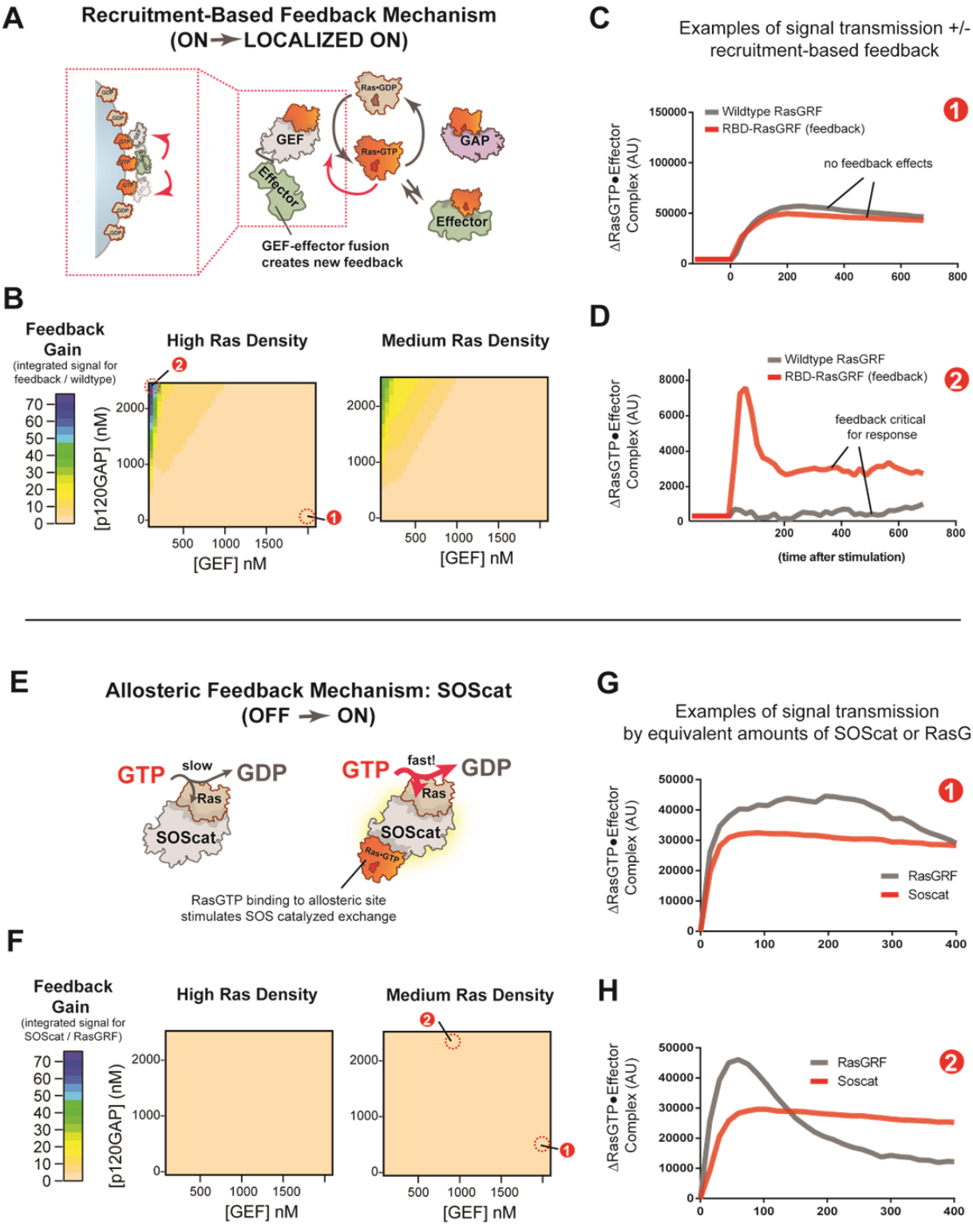
Figure 6. Unique interpretation of RasGTP signals by different effectors in multi-effector networks encodes multiple distinct temporal outputs in the system response.



(a) Depiction of the experimental design: a fixed step-input is applied to a variety of network

configurations in which more than one effector molecule is present resulting in multiple system outputs. (b) Absolute and normalized responses to step-input of C-Raf RBD and B-Raf RBD in the absence of any GAP activity. (c) as in (b) but with 1 μM NF1-GAP present in the signaling network. (d) Absolute and normalized responses to step-input of C-Raf RBD and A-Raf RBD in the absence of any GAP activity. (e) as in (d) but with 1 μM NF1-GAP present in the signaling network. (f) Absolute and normalized responses to step-input of C-Raf RBD and the C-Raf(N64A) mutant RBD in the absence of any GAP activity. (g) as with (f) but with 1 μM NF1-GAP present in the signaling network.

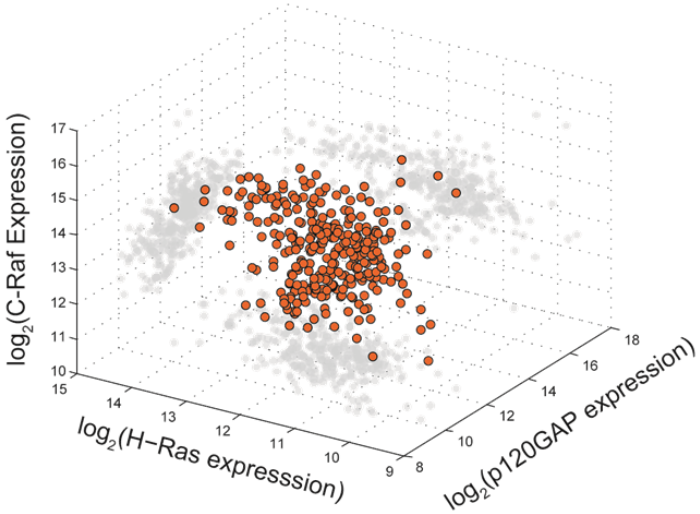
Figure 7. Differential effects of recruitment-based positive feedback or allosteric-based positive feedback on system behavior in different Ras signaling network configurations.



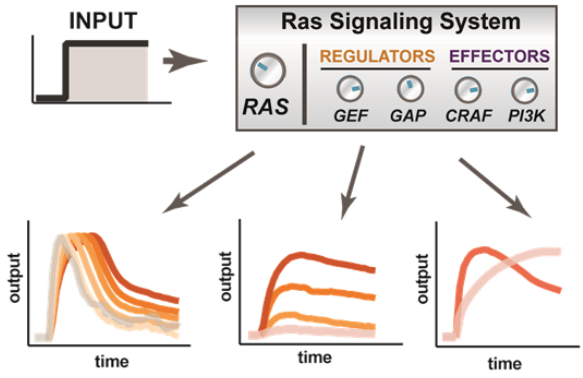
(a) Illustation of the synthetic recruitment-based positive feedback GEF (RasGRF-RBD) that was examined. (b) Phase diagram depicting the feedback gain of the recruitment-based mechanism (defined as fold-change in integrated signal of RasGRF-RBD relative to wildtype RasGRF) in different network configurations. (c) Example of output responses for wildtype or RasGRF-RBD containing systems in position (1) from figure (b). (d) Example of output responses for wildtype or RasGRF-RBD containing systems in position (2) from figure (b). (e) Illustration of the naturally occurring allosteric-based positive feedback GEF (SOScat) that was examined. (f) Phase diagram depicting the feedback gain of the recruitment-based mechanism (defined as fold-change in integrated signal of SOScat relative to wildtype RasGRF) in different network configurations. (g) Example of output responses for RasGRF or SOScat containing systems in position (1) from figure (f). (d) Example of output responses for wildtype or RasGRF-RBD containing systems in position (2) from figure (f).

Figure 8. One system, many behaviors: versatility and fragility in the space of Ras GTPase signal processing behaviors.

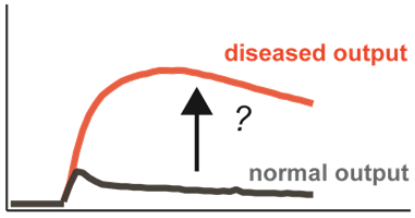
A **Ras network configurations vary substantially across different cell-types**



B **VERSATILITY** | *a diverse space of readily accessed dynamic signal processing programs*



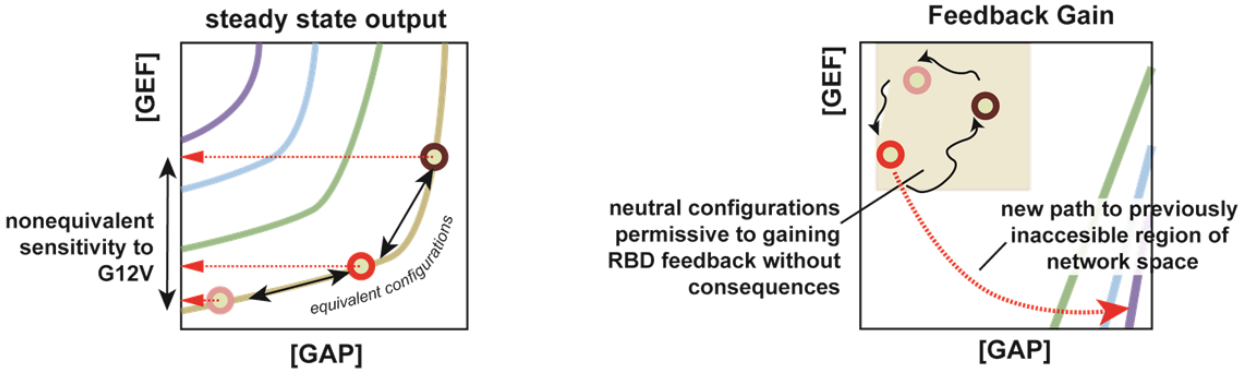
C **FRAGILITY** | *multiple paths to undesirable signaling outcomes*



? = Ras mutations,
GAP inactivation, decreased levels
GEF hyperactivation, increased levels
Effector expression, affinity

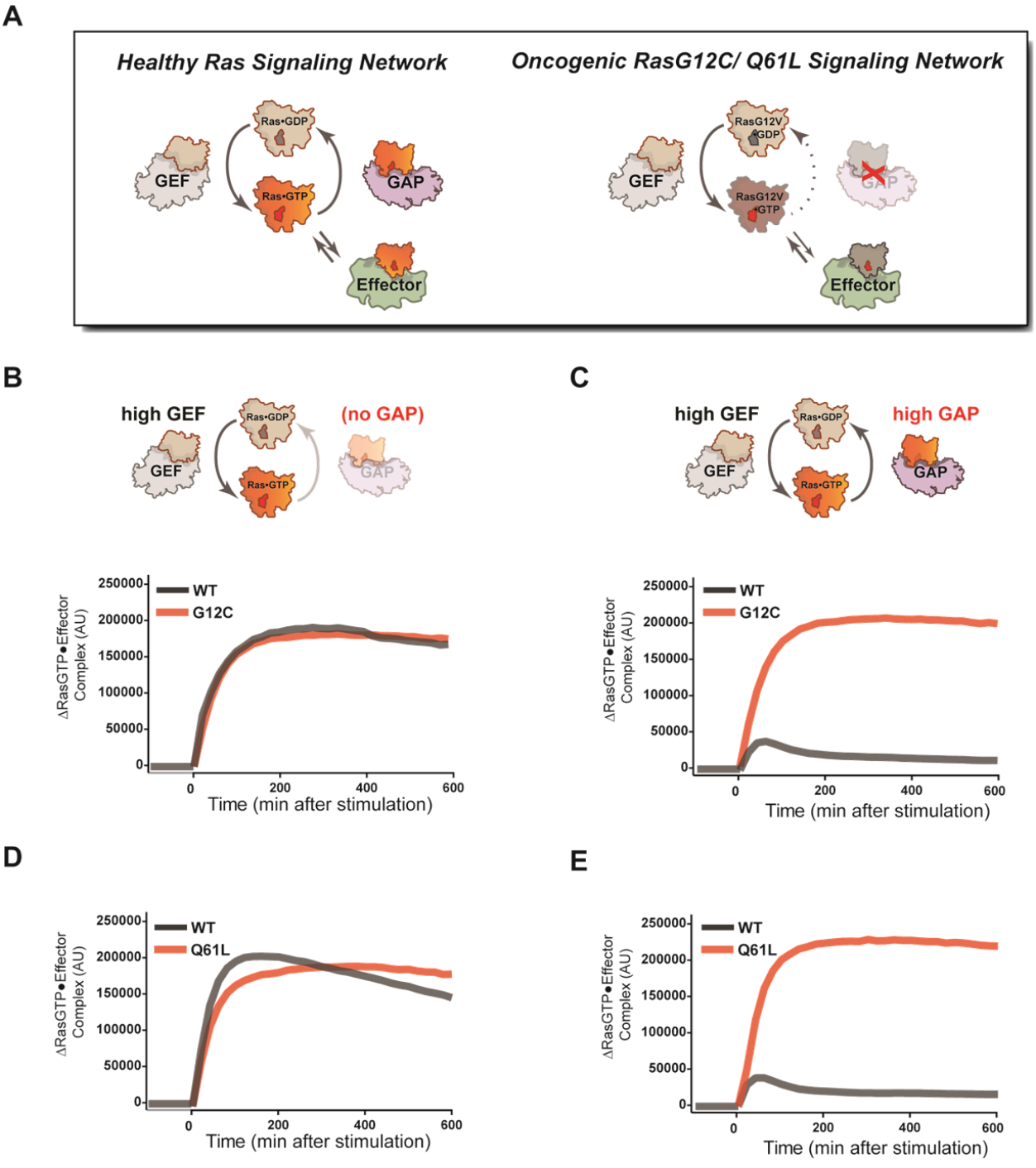
D

neutral structure implies divergent responses of "equivalent states" to pertubation



(a) Distribution of p120GAP, H-Ras, and C-Raf gene expression levels across a variety of human cell-types, synthesized from Genevestigator data (see methods). Each orange point corresponds to a cell-type and its position in the space indicates its associated expression level in each coordinate. A “shadow” of each point is projected onto each two-dimensional sub-plane to further clarify the distribution. (b) Illustration depicting the versatility of Ras GTPase signaling systems. A simple step-input can be processed into a variety of different dynamic outputs depending on the network configuration. (c) Illustration depicting the fragility of Ras GTPase signaling systems. Given a particular signaling output and a higher level disease output, there exist many paths by which the network configuration can change to produce the diseased output. (d) Phase diagram schematic showing how network configurations with equivalent output responses (i.e. those along a contour in the phase diagram) can nonetheless be primed for divergent responses to a perturbation like the G12V Ras mutation. (e) Phase diagram schematic showing how a new the RBD-GEF feedback architecture can be implemented without consequence in certain regions of the phase diagram, enabling access to previously non-functional network-configurations.

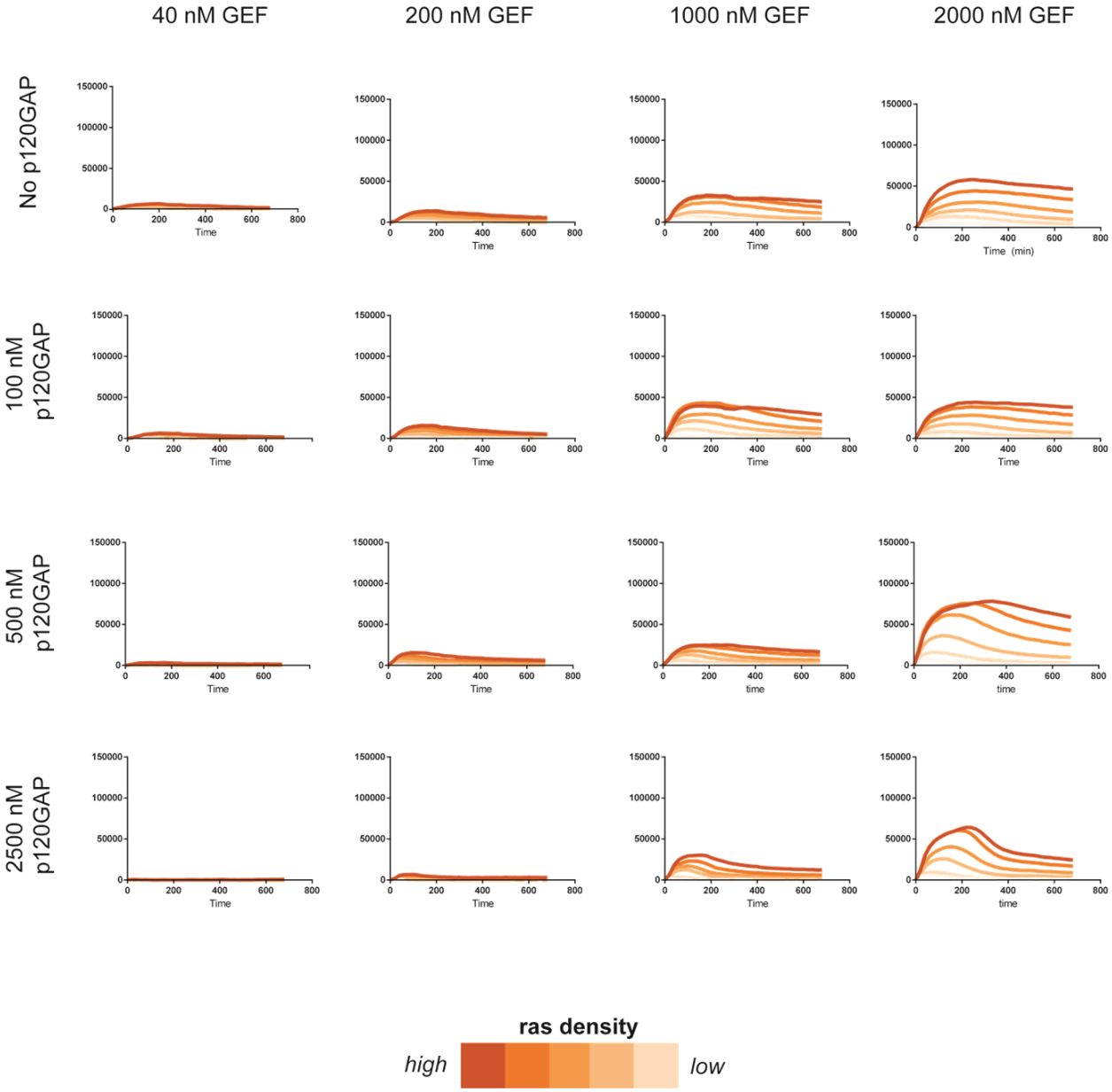
Figure S1. The extent of signal processing distortion by oncogenic G12C and Q61L alleles of Ras depends on the balance of positive and negative regulatory activities in the network.



(a) Depiction of wildtype Ras and oncogenic G12C/Q61L Ras illustrating the modes by which mutation is thought to impact the network behavior: changing in intrinsic hydrolysis rate,

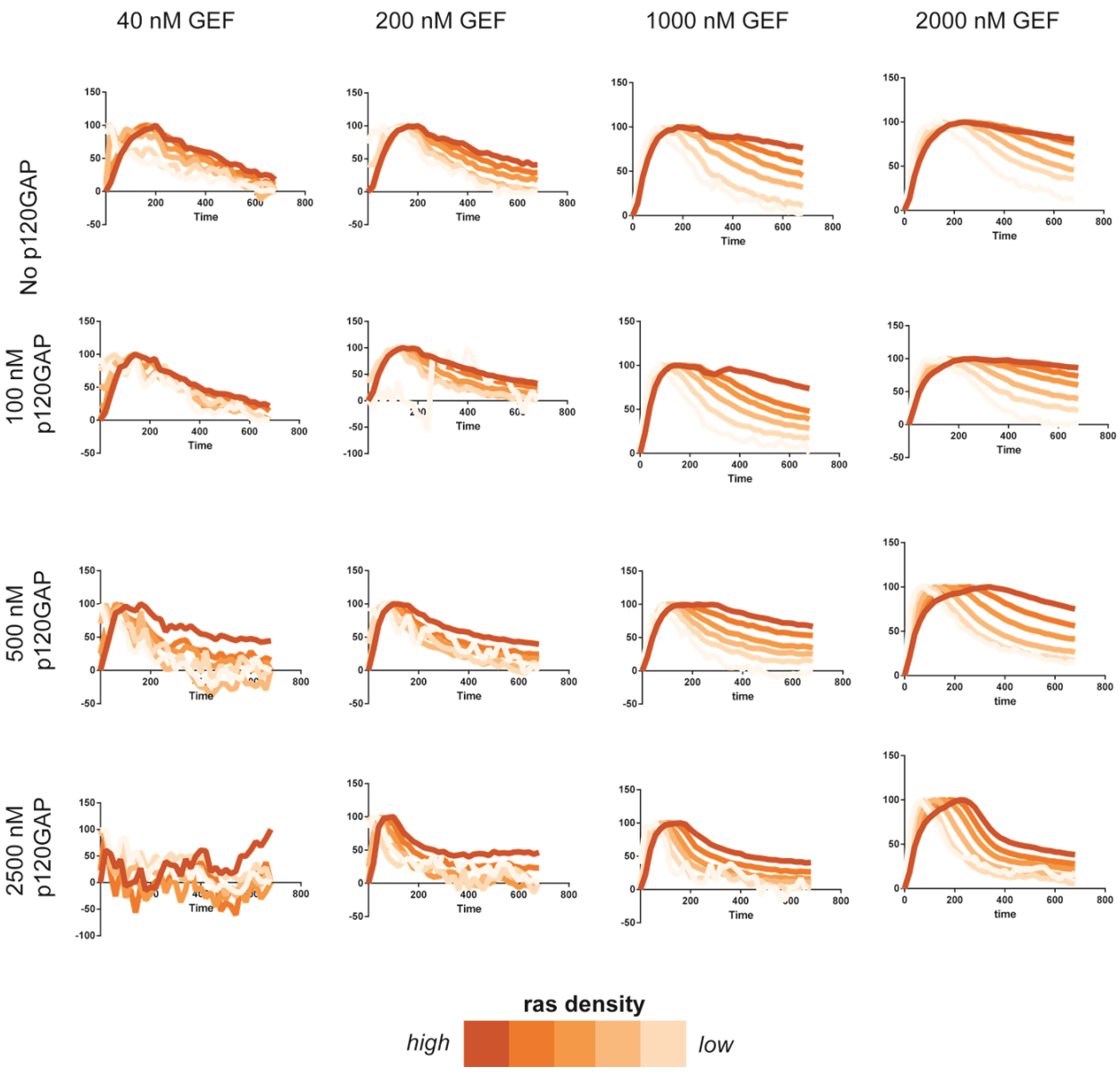
blocking GAP mediated hydrolysis, and modulating effector interactions. (b) Absolute and (c) normalized effector responses to a 2 μM RasGRF GEF step input in the absence of any GAP activity. (d) Absolute and (e) normalized responses of the same step input as in (a) and (b), but with 1 μM NF1 GAP activity present in the network.

Figure S2. Absolute Responses of Ras Signaling System Under A Variety of Network Configurations



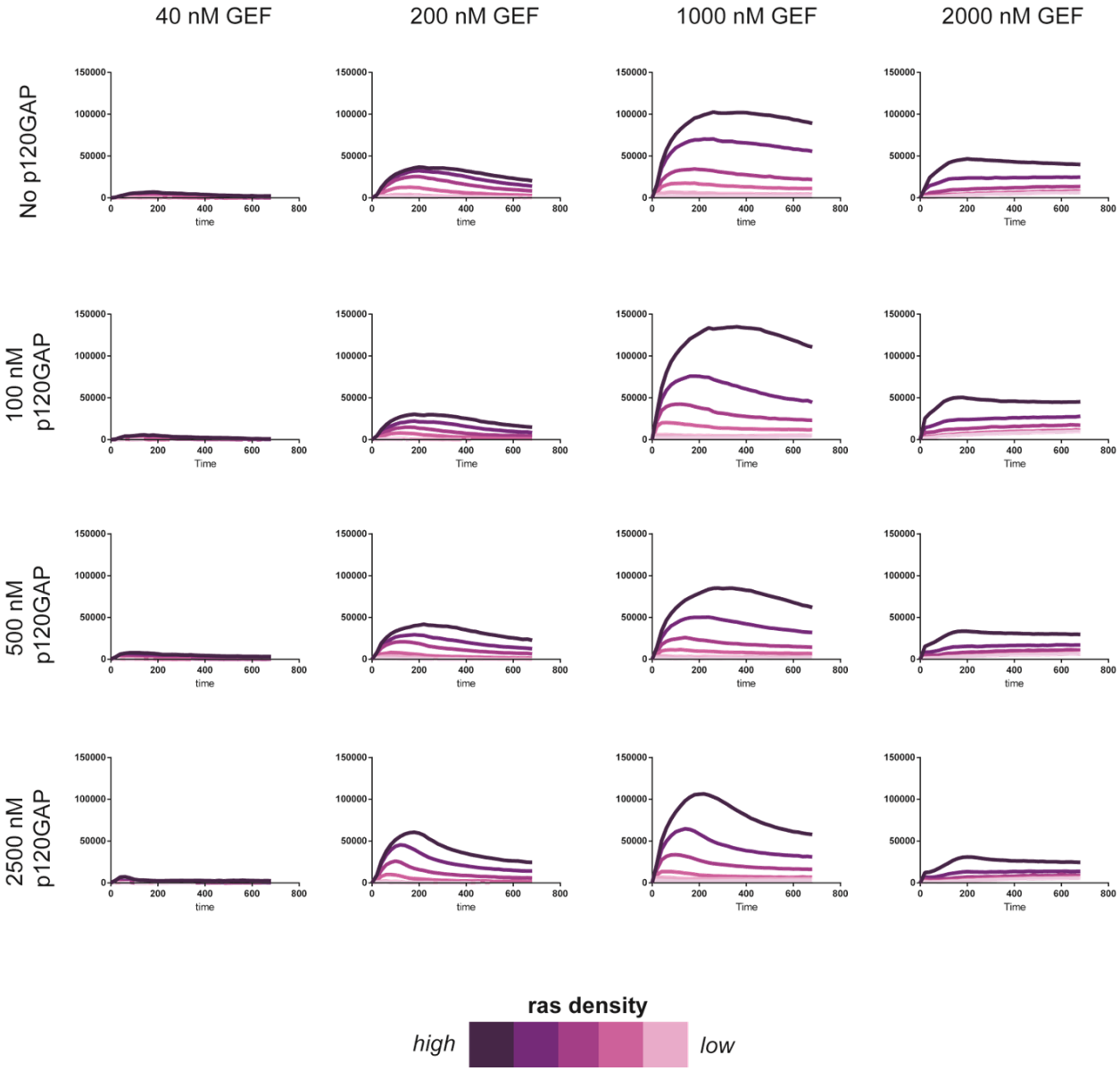
Absolute signaling responses for different network GEF/GAP/Ras density configurations. The RasGRF catalytic domain was used as the activating GEF in these experiments.

Figure S3. Normalized Responses of Ras Signaling System Under A Variety of Network Configurations



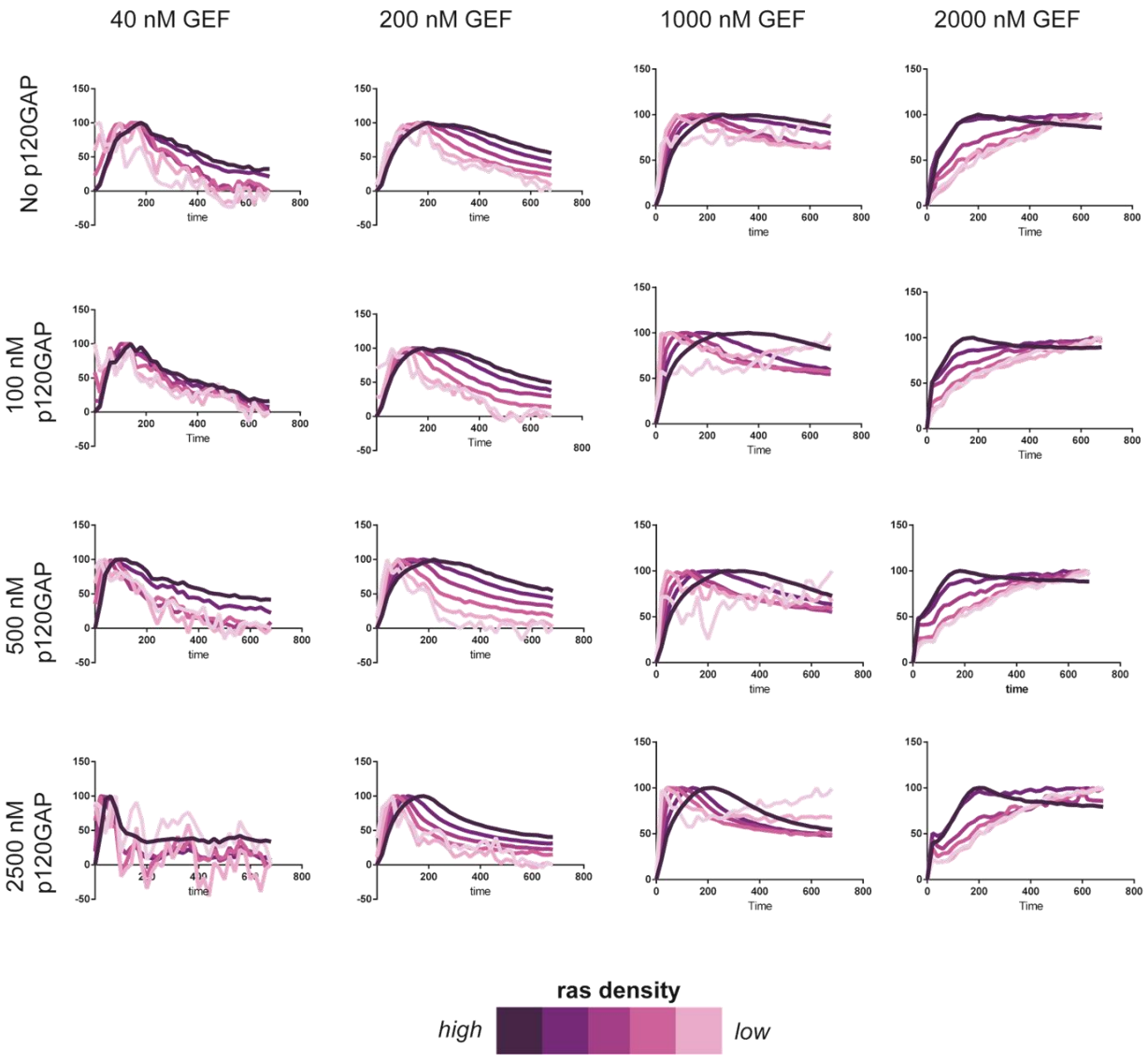
Normalized (to maximum value) signaling responses for different network GEF/GAP/Ras density configurations. The RasGRF catalytic domain was used as the activating GEF in these experiments.

Figure S4. Absolute Responses of Ras Signaling System with RBD-RasGRF Positive Feedback Under A Variety of Network Configurations



Absolute signaling responses for different network GEF/GAP/Ras density configurations. The RasGRF-RBD catalytic domain was used as the activating GEF in these experiments to add a recruitment-based positive feedback mode to the system.

Figure S5. Normalized Responses of Ras Signaling System with RBD-RasGRF Positive Feedback Under A Variety of Network Configurations



Normalized (to maximum value) signaling responses for different network GEF/GAP/Ras density configurations. The RasGRF-RBD catalytic domain was used as the activating GEF in these experiments to add a recruitment-based positive feedback mode to the system.

Table 1. Gene expression data from a variety of human tissue and cell types that was used to produce the plot in Figure 8A.

	Gene Expression Level		
Gene Symbol	**RASA1**	**HRAS**	**RAF1**
chorionic villus	17.011	12.473	13.833
placenta	16.521	12.31	13.438
urothelial cell	15.031	14.542	13.262
iliac vein endothelial cell	14.708	12.293	14
lung fibroblast	14.707	13	13.689
human osteoblast like cell	14.654	13.453	14.091
dental pulp stem cell	14.58	12.597	13.817
bone marrow stromal cell	14.422	13.304	13.877
iris microvascular endothelial cell	14.323	14.036	13.463
microvascular endothelial cell	14.312	13.057	13.49
lung myofibroblasts	14.246	13.928	13.498
osteoblast	14.245	13.085	13.94
coronary artery endothelial cell	14.2	13.58	13.747
choroidal microvascular endothelial cell	14.076	14.333	13.629
trabecular meshwork cell	14.061	12.272	13.348
chorion villus cell	14.027	12.914	12.771
hepatic vein endothelial cell	14.025	12.92	14.326
aortic endothelial cell	13.991	12.831	13.883
retinal microvascular endothelial cell	13.884	14.594	12.699
dendritic cell	13.799	11.173	13.647
neonatal foreskin-derived fibroblast	13.753	12.721	13.718
dermal fibroblast	13.743	12.857	13.675
synovial fibroblast	13.711	12.189	13.486
hepatic myofibroblast	13.697	12.706	12.826
chondrogenic progenitor cell	13.696	13.177	13.674
umbilical artery endothelial cell	13.67	12.157	14.021
alveolar macrophage	13.606	12.142	13.745
pigment epithelium	13.595	11.818	13.304
dermal papilla cell	13.529	12.602	14.143
chondrocyte	13.519	13.078	13.762
bronchial smooth muscle cell	13.515	11.583	13.335
bronchial epithelial cell 3D culture	13.507	13.712	13.739
umbilical vein endothelial cell	13.502	13.431	13.494
aortic vascular smooth muscle cell	13.484	12.665	13.025
scalp	13.457	13.236	13.55
granulosa cell	13.416	11.173	14.233

melanocyte	13.415	10.643	13.183
retina	13.383	12.9	13.92
pulmonary vein endothelial cell	13.377	13.73	14.148
iliac artery endothelial cell	13.354	13.571	13.918
Brodmann area 46	13.313	13.09	12.717
coronary smooth muscle cell	13.298	12.232	13.217
bone marrow derived mesenchymal stem cell	13.266	12.56	13.077
hepatic artery endothelial cell	13.266	12.818	14.326
amniocyte	13.226	12.712	13.603
labial salivary gland	13.214	11.482	12.992
retina pigment epithelial cell	13.177	12.473	13.577
bronchial epithelial cell	13.167	13.931	13.765
macrophage	13.135	11.465	13.578
glomerulus	13.122	13.178	14.176
peripheral blood mononuclear cell (CD14-)	13.102	11.304	13.429
omental adipose tissue	13.082	11.606	13.249
pulmonary artery endothelial cell	13.079	13.305	13.641
monocyte derived macrophage	13.022	11.546	14.486
breast epithelial cell	13.022	13.014	13.012
pulmonary microvascular endothelial cell	13.012	13.2	12.836
left internal mammary artery	13.012	13.296	13.332
gluteal subcutaneous adipose tissue (glutSAT)	12.967	12.165	13.202
limbus	12.936	13.731	13.51
memory B-lymphocyte	12.917	10.84	13.397
cornea	12.916	14.524	13.671
tracheal epithelial cell	12.903	12.292	13.316
oocyte	12.903	10.42	10.372
conjunctiva	12.872	13.883	13.773
tarsal conjunctiva	12.872	13.303	13.933
mucosa (ascending colon)	12.867	11.911	13.208
colonic crypt	12.822	11.074	13.844
abdominal subcutaneous adipose tissue (SAAT)	12.809	12.307	13.147
human bronchial organotypic tissue culture	12.8	12.433	13.336
human nasal organotypic tissue culture	12.789	12.501	13.384
tubulointerstitium	12.766	12.242	13.554
tonsillar T-lymphocyte	12.743	11.491	13.841
cumulus cell	12.738	11.882	14.188
synovial membrane	12.738	11.391	13.092
prefrontal cortex	12.731	12.795	12.217
astrocyte	12.73	12.765	13.89

postcentral gyrus	12.72	12.787	12.694
bone marrow mononuclear cell	12.717	11.131	14.457
superior frontal gyrus	12.699	12.906	12.654
aortic root wall	12.699	12.464	13.907
cuboidal alveolar type 2 cell (AT2)	12.685	11.343	13.647
synovial fluid mononuclear cell	12.66	12.477	13.658
choroid	12.652	10.924	12.973
corpus	12.652	11.612	12.686
subcutaneous adipose tissue (SAT)	12.631	11.65	12.681
transverse colon	12.613	11.181	12.9
neuronal cell	12.605	13.922	13.135
internal mammary artery wall	12.598	12.461	14.177
caput	12.567	10.842	12.768
CD8 T-lymphocyte (activated)	12.564	11.946	13.683
mediastinal adipose tissue (MAT)	12.559	12.118	13.43
CD4 T-lymphocyte (resting)	12.521	11.613	13.622
B-lymphocyte (germinal center)	12.52	11.571	13.606
frontal pole	12.506	12.621	13.096
lung	12.497	11.818	13.365
synovial fluid	12.491	11.661	13.473
lymphoblast cell	12.491	11.292	13.204
dermal microvascular lymphatic endothelial cell	12.486	13.097	13.131
inguinal lymph node	12.463	10.393	12.026
hippocampus	12.449	12.875	12.691
T-lymphocyte	12.447	11.411	13.35
CD4 T-lymphocyte (activated)	12.446	12.847	13.865
bladder	12.427	11.674	13.201
large airway epithelial cell	12.393	12.172	13.267
colon	12.392	11.881	13.261
hepatocyte	12.391	12.978	13.628
mammary gland (breast)	12.367	11.771	13.123
cauda	12.367	11.173	12.832
thrombus-derived leukocyte	12.36	11.442	14.865
minor salivary gland	12.329	11.205	13.451
small airway epithelial cell	12.326	11.931	13.035
brain (encephalon)	12.301	12.301	12.21
monocyte derived dendritic cell	12.29	11.53	13.516
colon polyp	12.288	12.154	13.358
conjunctival epithelial cell	12.288	14.313	13.664
neutrophil granulocyte	12.277	10.014	16.101
hair follicle keratinocyte (CD200-/CD49+)	12.276	13.844	13.581
peripheral blood mononuclear cell	12.268	11.181	13.809

conjunctival epithelial cell (sp.)	12.263	12.825	14.262
sigmoid colon	12.261	11.276	12.805
nasal epithelium	12.251	12.281	13.164
B-lymphocyte	12.245	11.465	13.686
endometrium	12.244	10.969	13.039
deltoid	12.229	11.951	14.251
heart left atrial appendage	12.224	11.815	13.851
lateral rectus	12.213	11.482	13.455
peripheral blood mononuclear cell (CD14+)	12.194	12.042	13.661
epidermis	12.193	13.842	14.186
B-lymphocyte (naive)	12.187	11.223	13.872
descending colon	12.185	11.233	12.73
CD8 T-lymphocyte (resting)	12.173	11.149	13.717
medial rectus	12.172	11.9	13.608
peripheral blood lymphocyte	12.171	11.836	14.842
mucosa (oral region)	12.131	13.723	13.646
spleen	12.128	11.125	13.258
cerebellum	12.116	12.492	13.172
ascending colon	12.107	11.429	12.898
jejunum	12.103	11.581	13.829
ovary (female gonad)	12.092	11.014	13.429
mucosa (sigmoideum)	12.084	12.294	13.459
hair follicle keratinocyte (CD200+/CD49+)	12.082	13.404	13.648
endometrial stromal cell	12.07	12.25	13.144
heart right ventricle	12.062	11.824	13.942
bone marrow hematopoietic stem cell	12.057	11.804	14.078
associative striatum	12.054	13.741	12.693
plasma cell	12.052	11.292	13.746
CD4+/CD25high regulatory T-cell	12.05	10.92	14.502
myeloid dendritic cell	12.046	10.647	13.683
gingiva	12.046	13.777	13.431
biceps femoris	12.044	12.596	14.612
monocyte (CD14+)	12.04	11.334	14.492
ileum	12.026	11.468	13.297
mucosa (colon)	12.025	11.579	13.397
monocyte	12.023	10.814	14.092
heart right atrial appendage	12.02	12.069	13.447
mucosa (descending colon)	12.019	11.356	12.8
thymus	12.016	11.009	13.348
oesophagus (esophagus)	12.011	12.577	13.407
testis (male gonad)	12.007	11.017	12.498
ileal epithelial cell	12.004	12.024	13.982

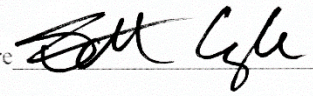
cerebrospinal fluid cell	12	11.344	14.178
colon epithelial cell (colonocyte)	11.999	12.26	13.205
occipital lobe	11.984	12.158	12.188
temporal lobe	11.977	12.13	12.115
cardiomyocyte	11.963	11.384	13.926
pancreas	11.949	11.02	12.886
skin	11.937	12.994	13.727
mucosa (rectosigmoid colon)	11.934	11.71	12.76
transiliac crest core	11.92	11.233	13.514
dorsal root ganglion	11.919	12.041	12.463
urethra	11.901	11.754	13.366
bone marrow myeloid cell (CD11b+)	11.886	11.515	14.247
pancreatic islet	11.881	12.88	12.492
joints	11.86	11.201	13.284
frontal lobe	11.86	12.442	12.147
vulva	11.854	13.088	13.291
CD4 T-lymphocyte	11.839	11.675	14.091
heart left ventricle	11.833	11.54	13.556
rectus abdominus	11.831	12.803	14.573
epidermal keratinocyte	11.817	13.156	13.691
leukocyte	11.79	10.925	14.849
nasopharyngeal epithelium	11.779	10.852	13.439
polymorphonuclear leukocyte	11.731	9.464	16.008
diaphragm	11.723	12.354	14.035
adrenal gland	11.693	11.951	13.401
nasopharynx	11.691	10.423	13.032
ovarian surface epithelium cell	11.675	10.644	13.007
entorhinal cortex	11.664	12.451	12.502
middle temporal gyrus	11.649	12.455	12.524
trigeminal ganglion	11.642	11.817	12.634
tonsillar T-lymphocyte (activated)	11.634	11.036	13.761
CD4 T-lymphocyte (naive)	11.632	11.514	14.021
penis	11.627	12.362	13.451
prostate stromal fibromuscular cell (CD49a+)	11.626	11.198	13.548
stomach	11.617	10.96	12.786
uterine cervix	11.614	11.064	13.089
myometrium	11.59	10.968	13.024
pituitary gland	11.575	12.019	12.356
quadriceps femoris	11.571	13.072	14.468
liver	11.57	11.227	13.534
parietal lobe	11.539	12.463	12.315
cerebral cortex (neopallium)	11.51	12.192	12.026

aortic valve	11.507	11.741	12.683
CD8 T-lymphocyte	11.506	11.379	14.099
T-lymphocyte (resting)	11.5	10.403	12.565
vastus lateralis	11.497	12.865	14.608
rectum	11.494	10.716	13.119
prostate gland	11.493	11.478	12.938
blood	11.491	10.848	14.468
bone marrow	11.472	11.143	13.875
cortex	11.457	11.629	13.215
spinal cord	11.446	11.825	12.194
caudate nucleus	11.446	13.109	12.534
lymphocyte	11.444	11.84	14.314
aorta	11.431	11.287	13.322
bladder stromal fibromuscular cell (CD13+)	11.429	12.195	14.173
B-lymphocyte (anergic)	11.429	11.599	14.333
cerebral hemisphere	11.425	11.796	12.991
right ventricle interventricular septum	11.424	13.097	14.582
substantia nigra	11.41	12.397	12.366
duodenum	11.406	11.045	13.087
vagina	11.402	11.65	13.05
superior vestibular nucleus	11.382	12.095	11.966
T-lymphocyte (activated)	11.366	12.135	13.065
oviduct (fallopian tube)	11.334	11.04	13.001
corpus callosum	11.322	11.675	12.475
midbrain (mesencephalon)	11.316	12.232	11.899
pars compacta	11.298	12.452	11.833
Langerhans cell	11.277	12.299	13.851
vermis	11.237	11.986	12.955
adrenal gland cortex (adrenal cortex)	11.208	11.542	12.75
adipose tissue (fat)	11.203	11.535	12.71
cardiac muscle (myocardium)	11.181	11.935	13.205
pars reticulata	11.167	12.367	11.84
amygdala	11.164	12.4	12.006
tonsil	11.16	11.645	13.062
reticular dermis cell	11.141	11.902	14.007
fundus	11.134	11.172	13.109
subthalamic nucleus	11.113	11.627	11.906
pons	11.083	12.31	11.768
salivary gland	11.065	11.123	12.601
CD4 T-large granular lymphocyte	11.024	10.692	13.441
lymph node	11.001	11.443	12.86
heart atrium	10.99	11.561	13.571

thyroid gland	10.952	11.092	13.024
peripheral blood leukocyte	10.934	10.803	13.753
medulla oblongata	10.915	12.194	12.2
hypothalamus	10.908	12.365	12.312
umbilical cord	10.895	11.993	13.327
medulla	10.892	11.449	13.165
pyramidal neuron (hippocampus)	10.891	11.728	12.667
internal globus pallidus	10.853	11.965	12.506
heart	10.831	11.349	13.558
lateral nuclei	10.817	11.847	11.526
pyloric antrum	10.815	11.186	12.894
thalamus	10.794	12.035	11.936
prostate stromal cell	10.79	11.943	13.451
large intestine	10.784	10.791	12.474
kidney	10.779	10.88	13.356
mucosa (pharynx)	10.764	12.45	13.225
breast stromal cell	10.756	11.133	13.35
pyramidal neuron (posterior cingulate)	10.742	11.505	12.255
ventral tegmental area	10.728	12.055	11.933
globus pallidus	10.722	12.035	11.893
basal epidermis cell	10.699	13.785	13.708
accumbens nucleus	10.698	12.653	11.97
pyramidal neuron (superior frontal gyrus)	10.673	11.642	12.65
bronchus	10.633	11.322	12.982
pyramidal neuron (middle temporal gyrus)	10.623	11.738	12.712
tongue	10.605	11.936	13.327
plasmacytoid dendritic cell	10.601	11.705	13.842
pyramidal neuron (primary visual cortex)	10.574	11.41	12.722
putamen	10.493	12.856	11.908
cardia	10.488	11.084	12.875
bone marrow erythroblast (CD235a+)	10.479	11.111	12.839
dermal papilla	10.45	11.778	12.935
basal ganglia	10.418	12.471	11.755
tonsillar T-lymphocyte (resting)	10.418	10.843	13.936
ovary stromal cell	10.39	9.425	12.742
saphenous vein	10.387	11.455	13.349
suprabasal epidermis cell	10.382	14.137	14.244
trachea	10.322	11.586	12.956
heart ventricle	10.297	11.528	13.127
genu	10.257	11.147	12.04
small intestine	10.202	11.341	13.065
skeletal muscle	10.2	12.176	14.043
nipple	10.171	12.533	13.17

cerebral white matter	10.163	11.919	11.8
large stellate neuron (entorhinal cortex)	10.05	11.437	12.233
coronary artery	10.031	11.854	12.724
bone marrow endothelial progenitor cell (CD133+)	10.014	9.714	13.648
pancreatic duct	9.987	11	13.286
CD30+/CD3+ T-lymphocyte	9.984	10.486	13.306
CD4+/CD3+/CD25+ activated T-lymphocyte	9.875	10.375	13.014
tongue squamous cell	9.856	10.754	13.574
uterus	9.787	11.689	13.011
CD4+/CD3+/CD25- resting T-lymphocyte	9.749	10.489	13.142
CD4+/CD3+/CD25high regulatory T-cell	9.615	10.556	13.345
platelet	9.528	9.26	12.951
bone marrow cell	9.418	11.487	13.85
CD8+/CD3+/HLA-DR+ activated T-lymphocyte	9.382	10.132	12.989
tonsillar NK cell	9.359	10.352	12.875
natural killer cell	9.183	10.023	11.999
natural killer T-cell	9.081	10.43	13.183
CD8+/CD3+/HLA-DR- resting T-lymphocyte	9.011	10.251	12.523

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature 

Date 8/8/2015