Cell Engineering Lecture 1: Protein Circuit Design

Patrick Barth

BIOENG-320

Synthetic Biology: Lecture series

From molecules to circuit to cell engineering

Protein design (Patrick Barth)

Protein circuits; cell engineering (Patrick Barth)

Gene circuits (Sahand Rahi)

Scale
Complexity

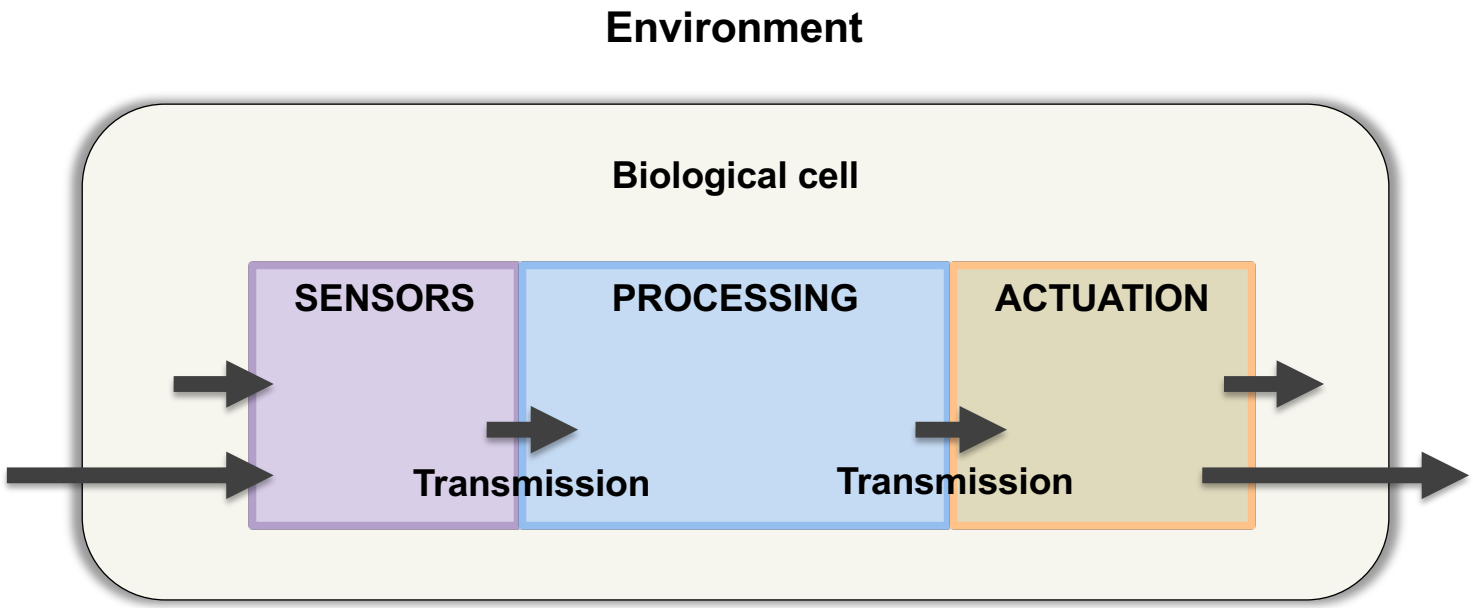
Cell Engineering is one of the ultimate goals in synthetic biology

WHY?

Minimal self replicating living system / factory

Multicellular systems / organs etc…

What functionalities do we need to implement to engineer “intelligent” cells?

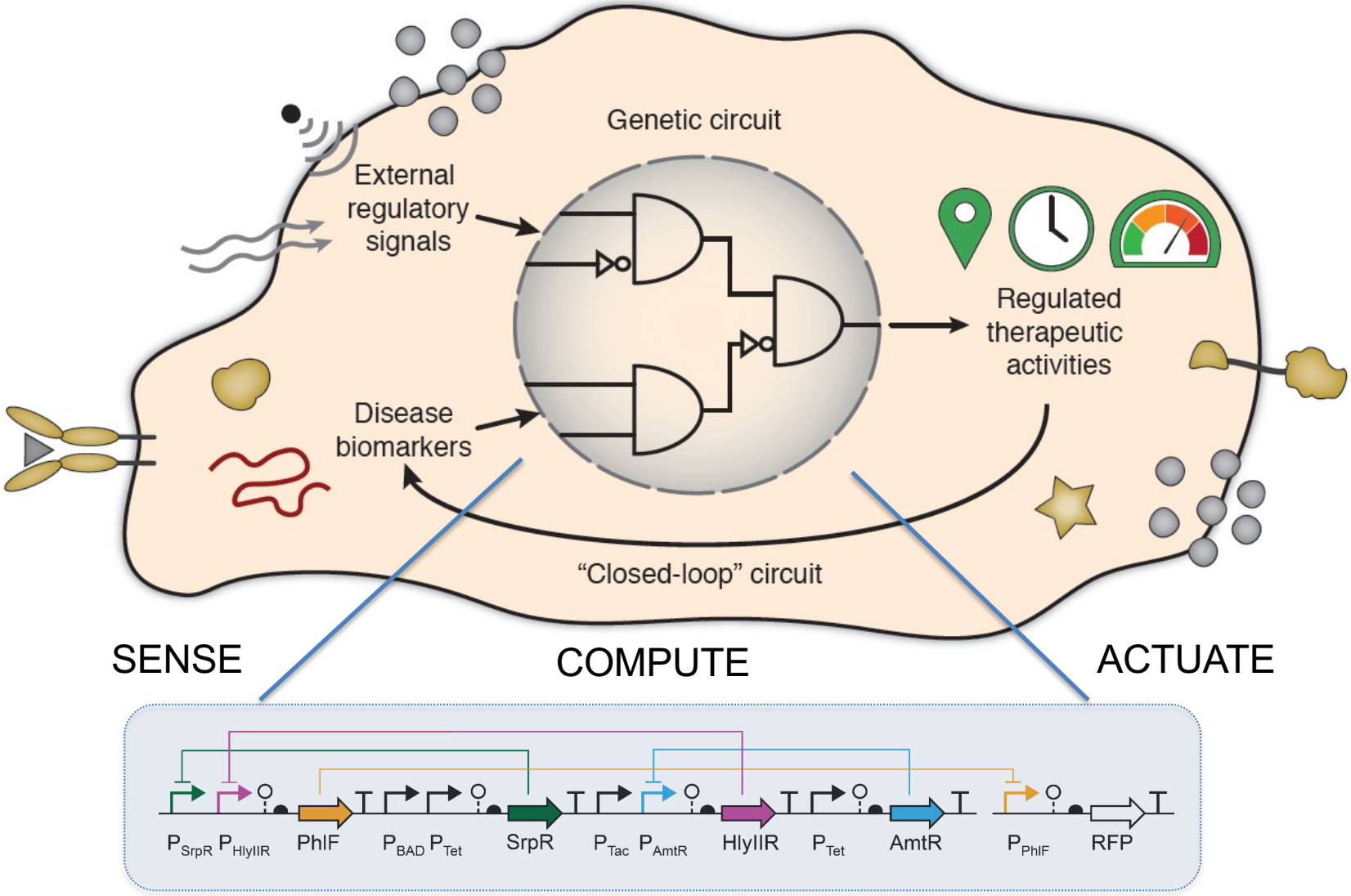


Synthetic gene circuit?

Synthetic protein circuits?

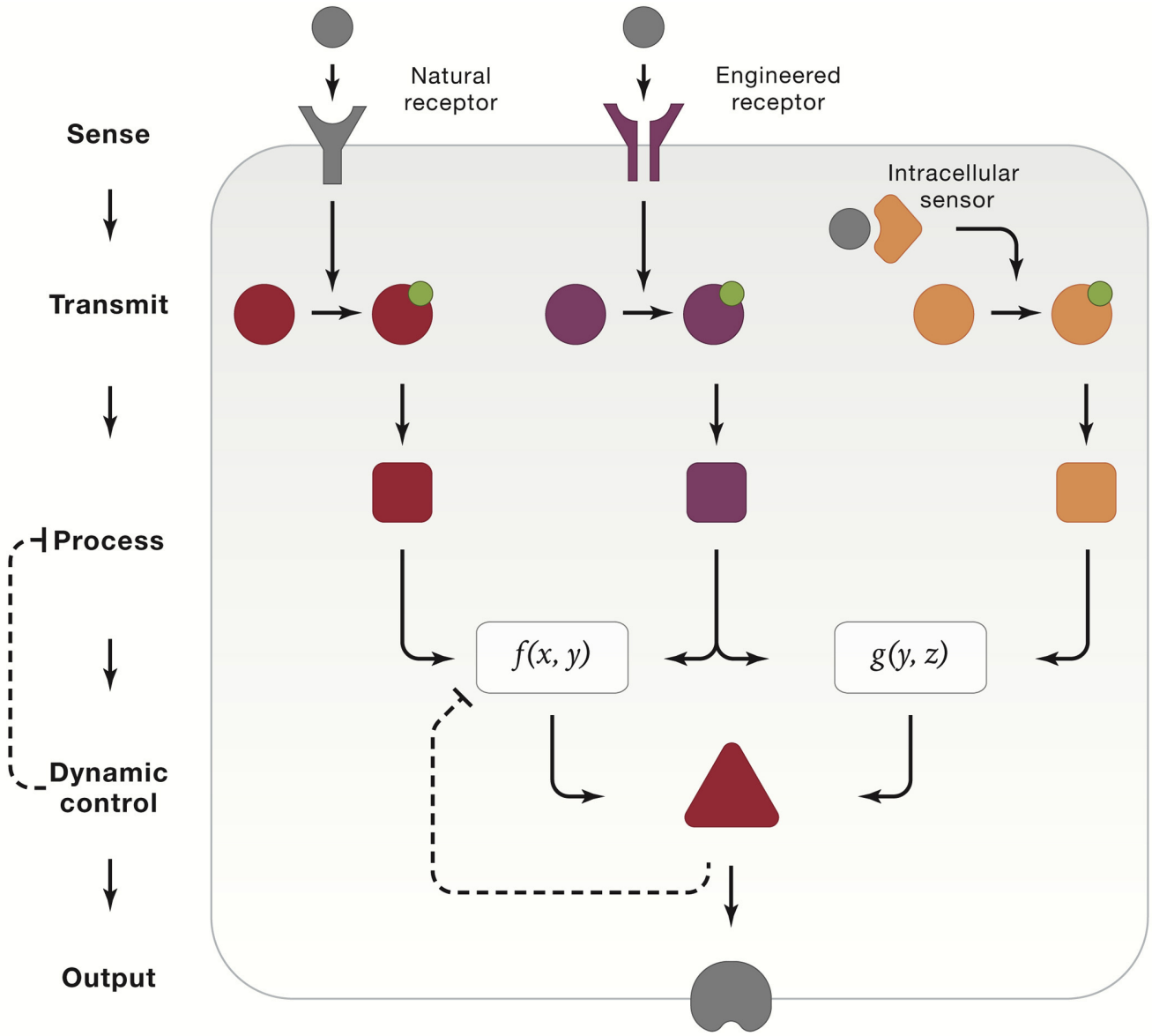
Synthetic hybrid protein/ gene circuits?

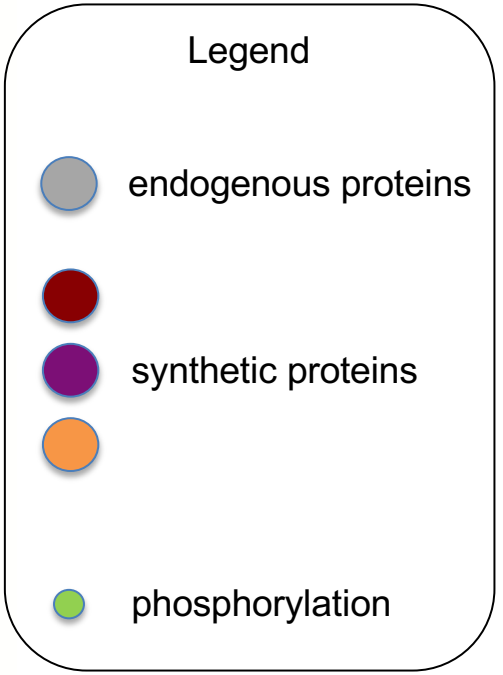
Genetic circuit



CLOSED CIRCUIT — All important computations through Transcription Factor-DNA interactions confined within the circuit

Protein circuit





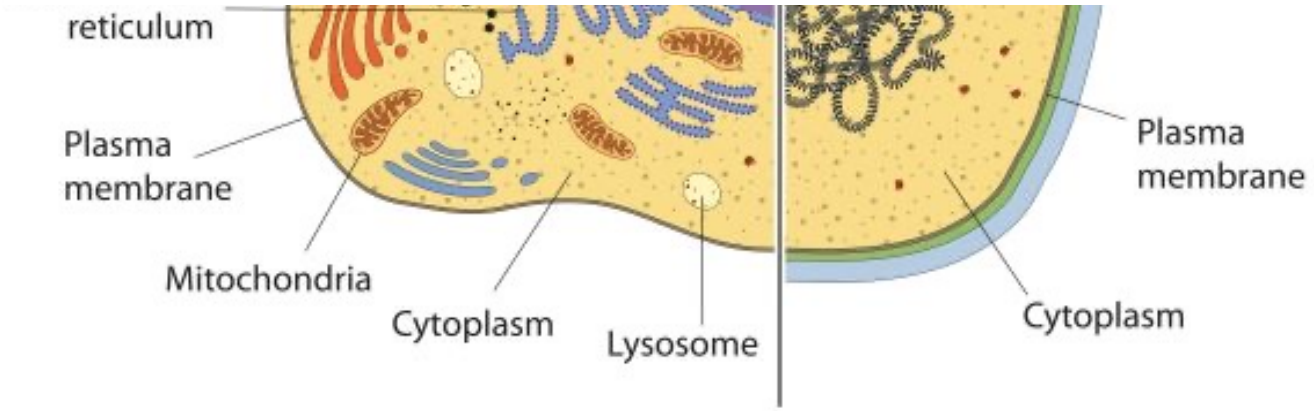
OPEN CIRCUIT

Computations occur at different locations in the cell

What kind of cells do we wish to engineer ?

EUKARYOTIC CELL | PROKARYOTIC CELL

What are the implications for the choice of genetic versus protein circuits ?



Many compartments
DNA inaccessible to the cytoplasm

No compartments
DNA accessible to the cytoplasm

Quiz: Genetic versus protein circuit

Circuit type	gene	protein
Part design		
Part interactions		
Direct interaction with endogenous pathway to reprogram cell function		
Dynamic response		
Operate across distinct cellular compartments		
Permanent genetic modification		
Potential to sense and respond to complex stimuli or states		
Which cell types are these circuits best suited for?		

Conclusion: Genetic versus protein circuit

Nucleic acids

Pros: versatile toolbox for biomolecular computation. relatively easy to program in a predictable manner, based on their ability to bind each other.

Cons: challenging to interface with endogenous protein pathways in the living cell. Slow dynamic responses.

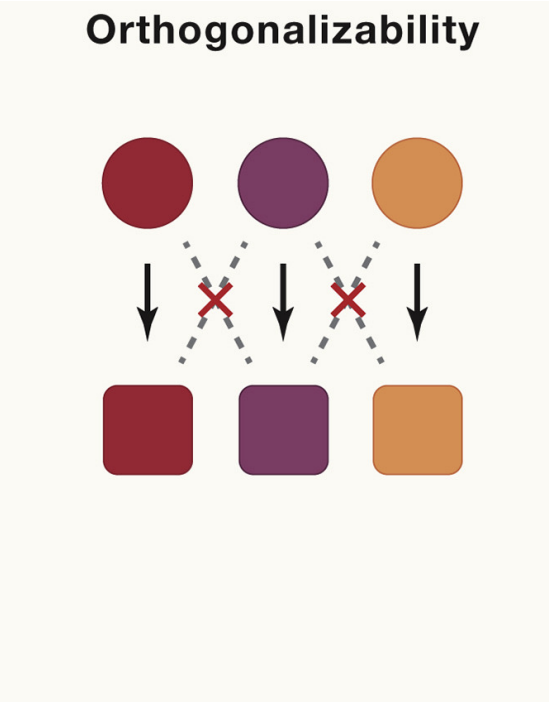
Proteins

Cons: more difficult to predictably design

Pros: possess a much larger potential repertoire of activities and interactions including binding, cleavage, ligation, allosteric modulation, and chemical modification.

Ideal protein circuit components

molecular interaction specificity similar to wiring in electronics
=> broad range of computational capabilities

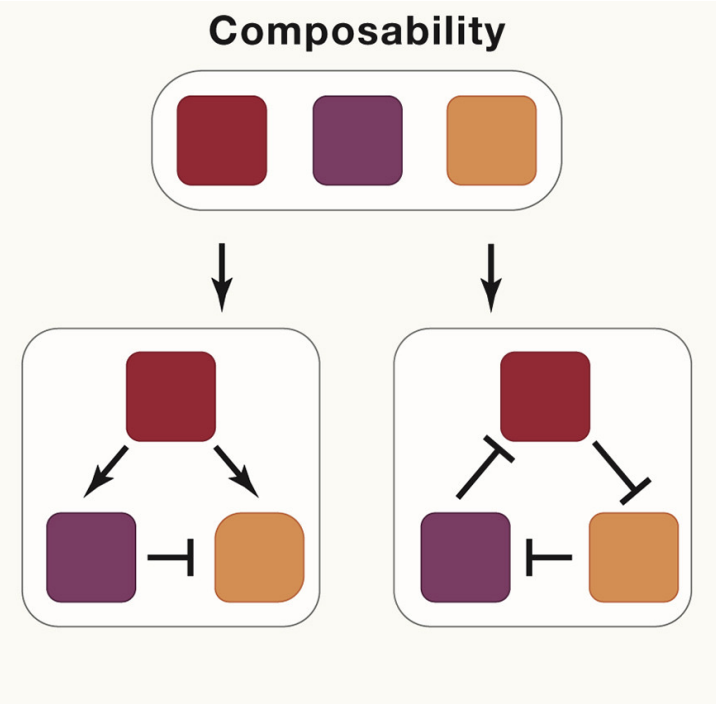


multiple variants that operate equivalently but independently with minimal crosstalk

⇒ multiple similarly functioning modules operating "in parallel"

⇒ control over complex information flow, integration and processing

Ideal protein circuit components



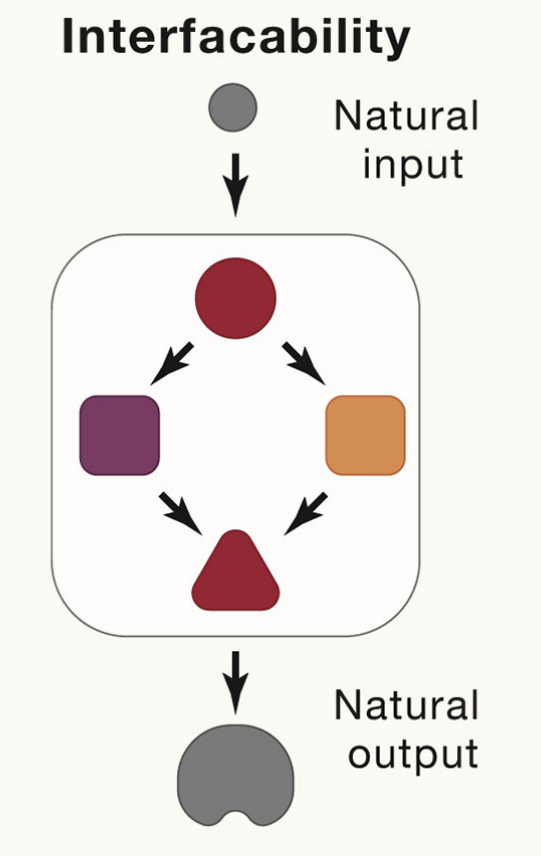
communicate with and regulate other components of their own type.

=> string components together ‘‘in series or networks’’ so that the output of one component can serve as an input for the next.

$\Rightarrow$ feedforward and feedback loops that can process signals

$\Rightarrow$ fast dynamic control

Ideal protein circuit components

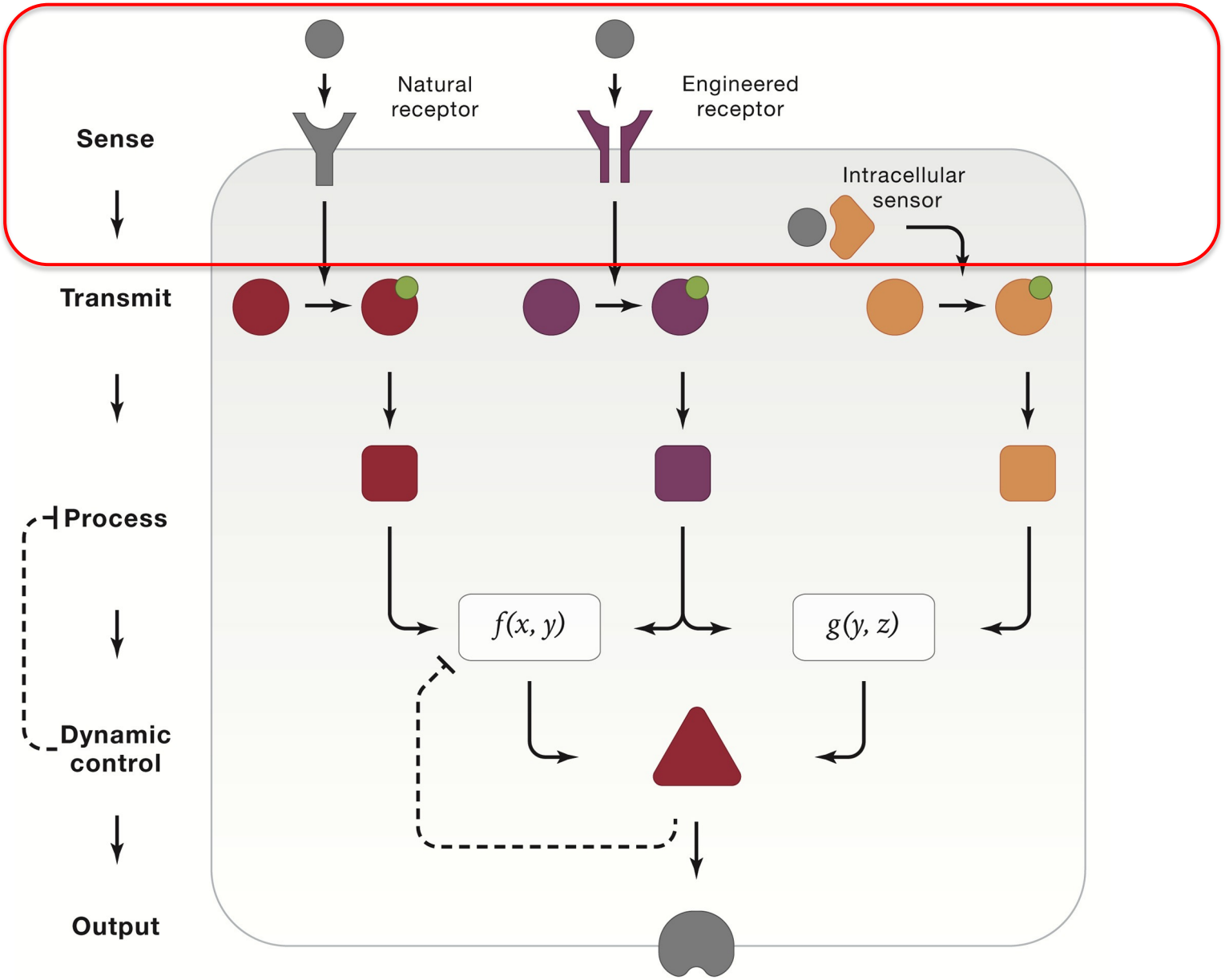


interfaceable with endogenous cellular proteins

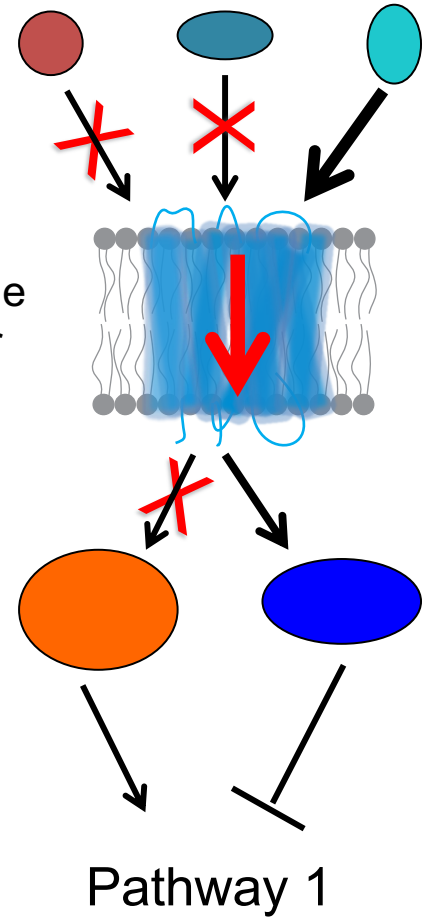
⇒ they should be able to directly interface with endogenous cellular proteins (e.g., through allosteric regulation or post-translational modification)

⇒ sensing and control of cellular pathways

Generic protein circuit operational in a living cell



First circuit layer: Sensing input signals

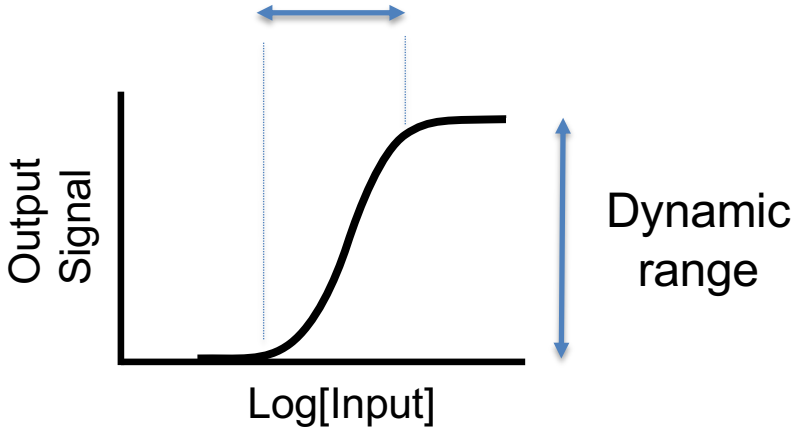


Input sensing specificity

Potent signal transmission

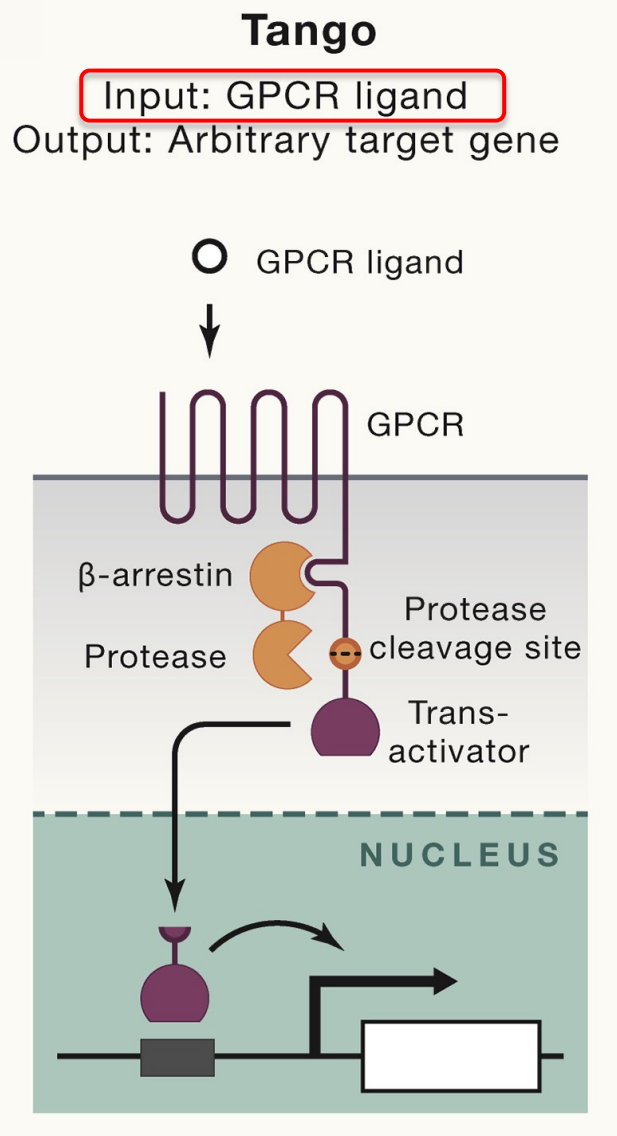
Output selectivity

Digital vs analog behavior



=> Sensitive detection and conversion of an input into an output signal

Engineered protein systems sensing extracellular targets



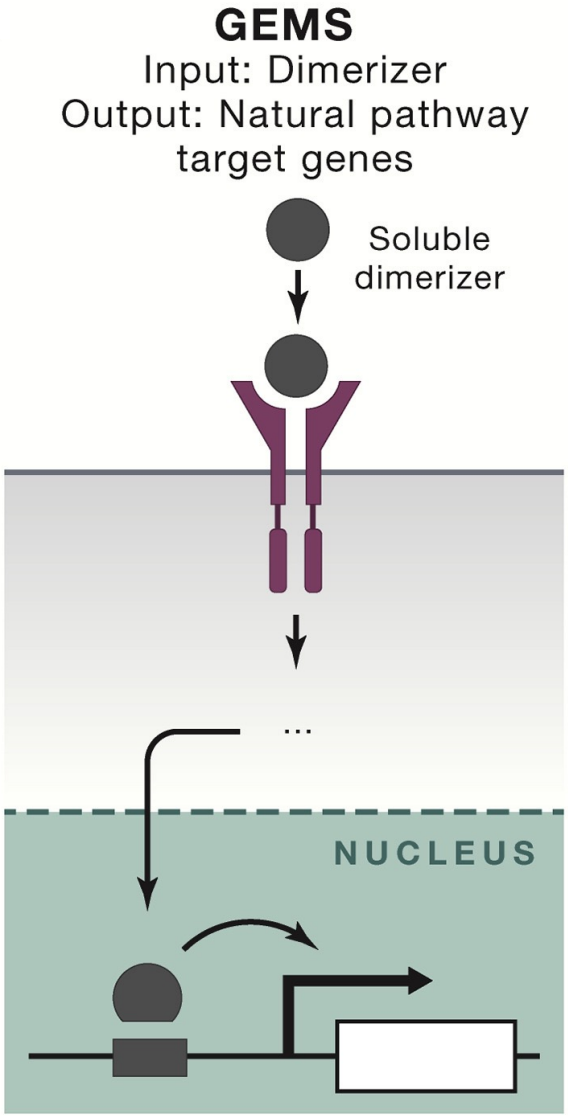
Extracellular input sensing linked to activation of intracellular viral proteases

Why viral proteases?

1. high target site specificity, orthogonality to bacterial and mammalian proteins
2. function similarly in diverse cellular contexts.

Range of input limited by natural receptor domains

Generalized extracellular molecule sensors

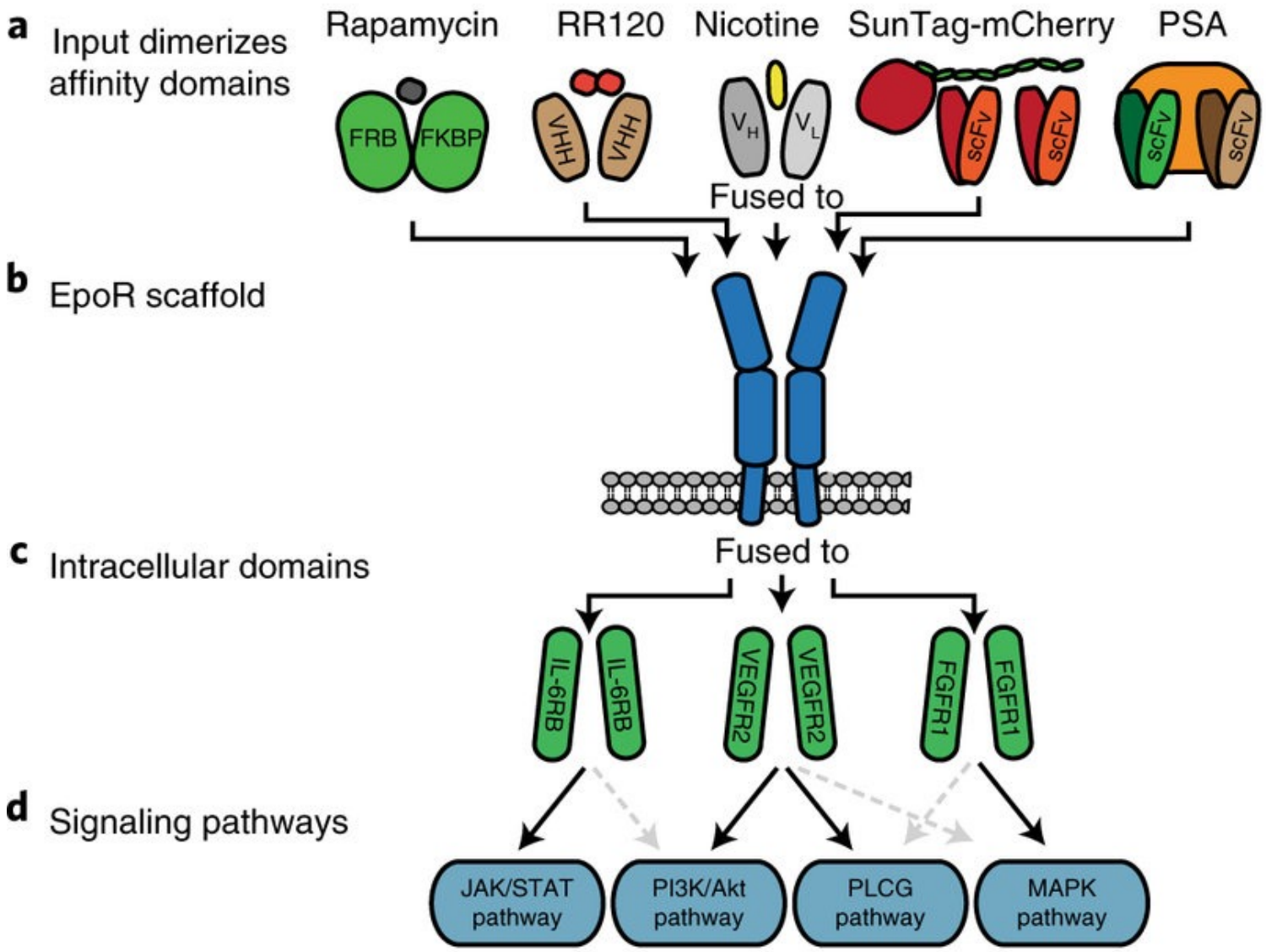


Modular design for arbitrary input and diverse output

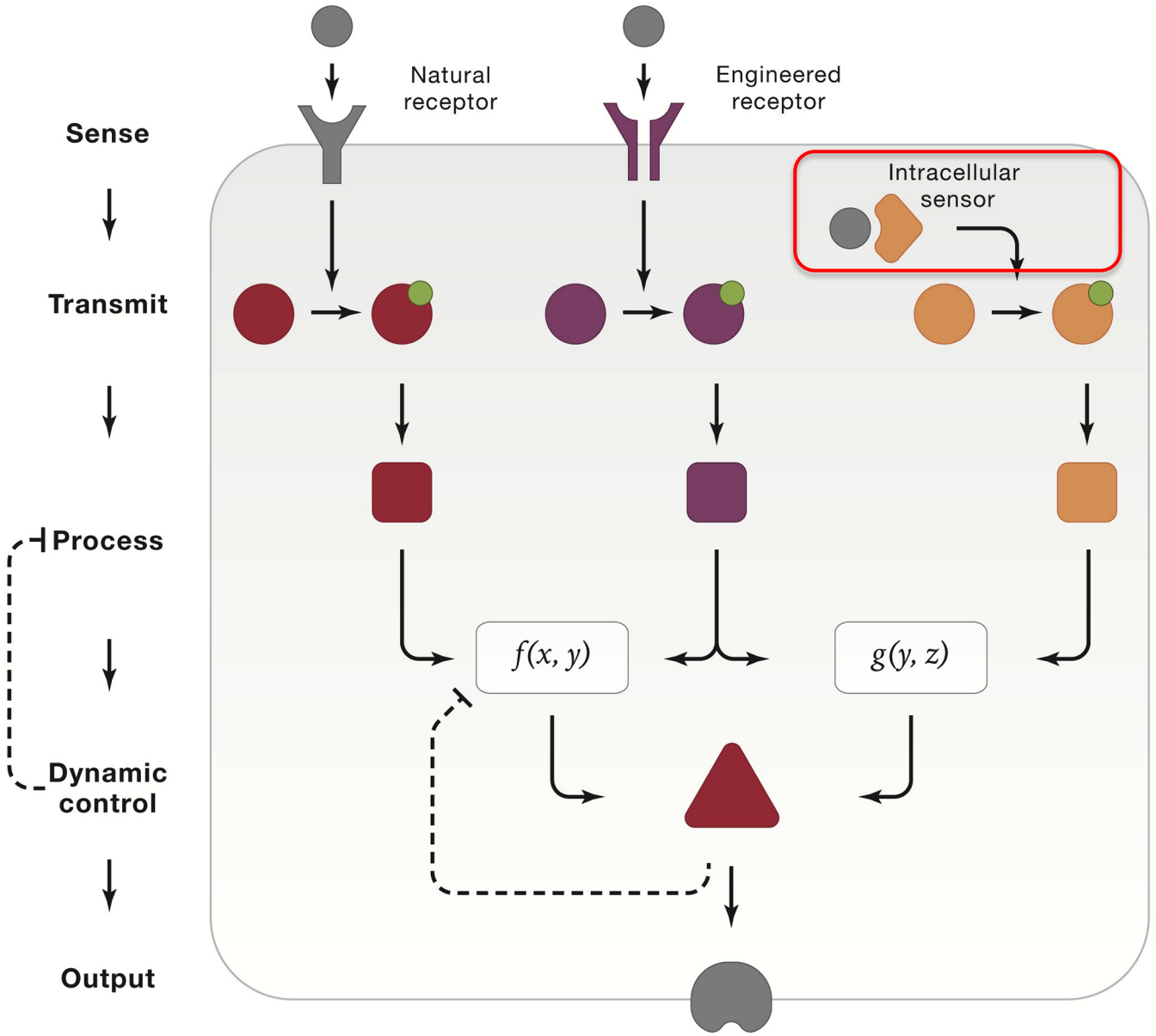
Extracellular domains dimerize upon ligand binding and trigger intracellular activation

Reminiscent of a class of native proteins?

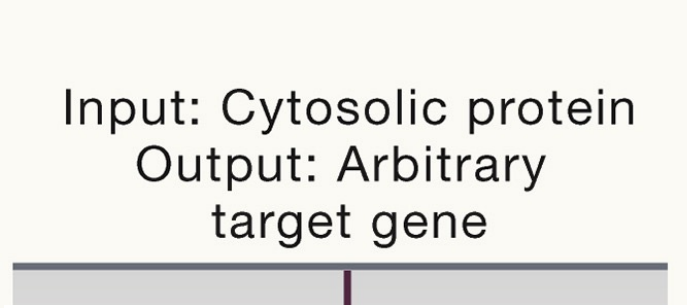
Generalized extracellular molecule sensors



First circuit layer: intracellular sensors

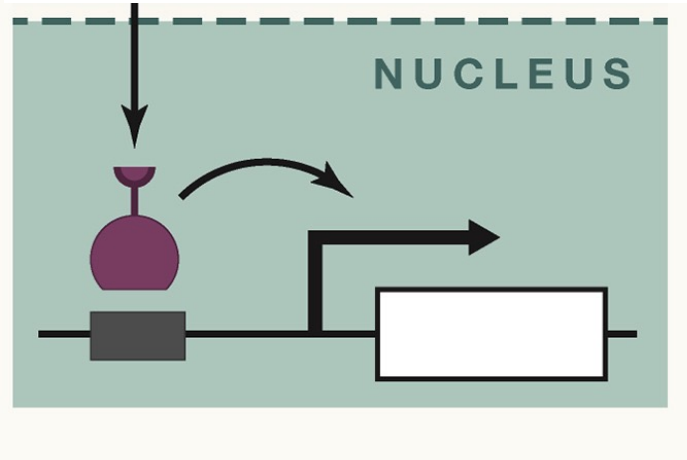


Engineered protein systems sensing intracellular proteins



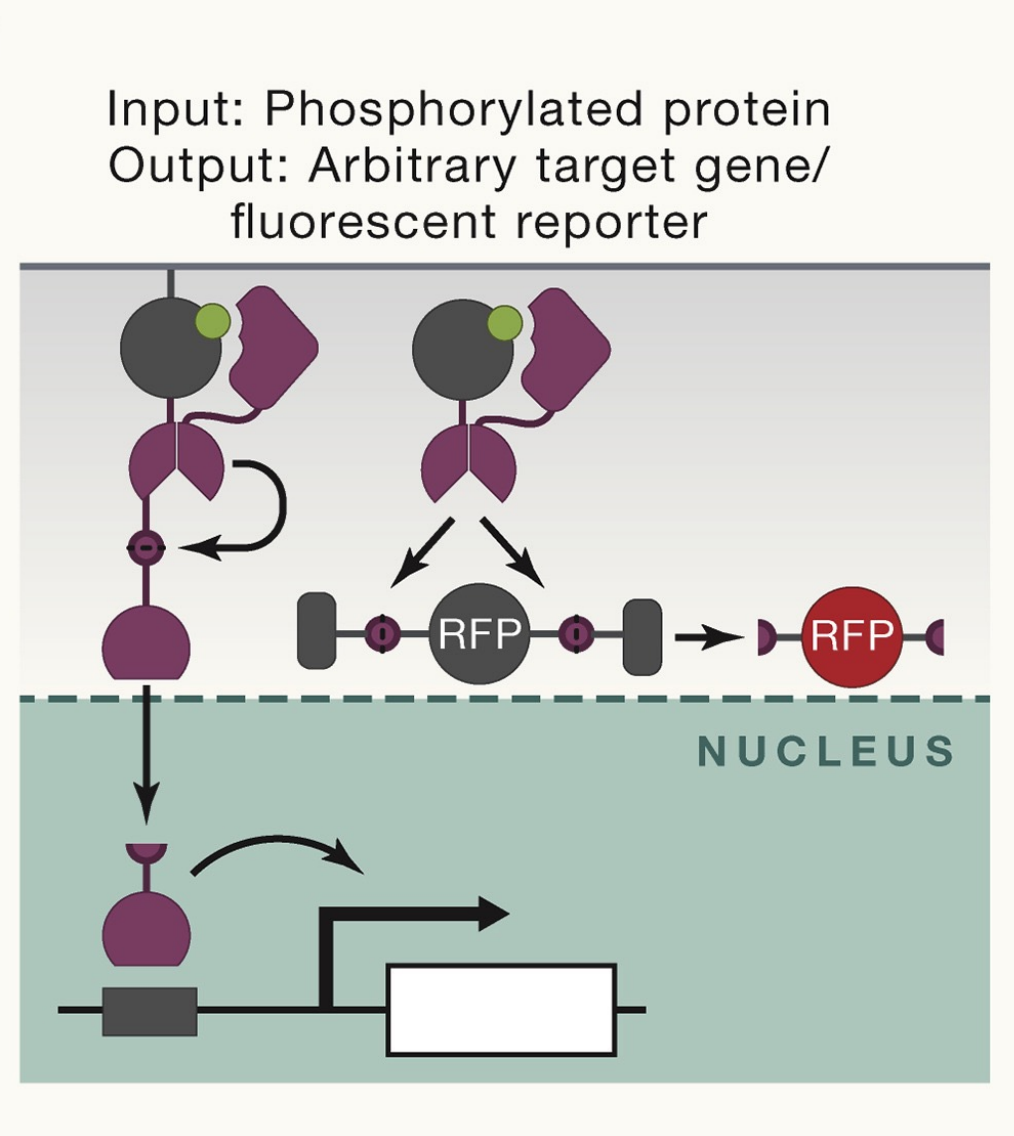
Selectivity: two nanobodies that bind distinct epitopes of the same target protein

CAN YOU THINK OF OTHER IMPORTANT INTRACELLULAR SIGNALS TO SENSE?



Limitations: requirement for two nanobodies that bind different target protein epitopes in a non-exclusive and non-perturbative manner

Engineered protein systems sensing intracellular protein states

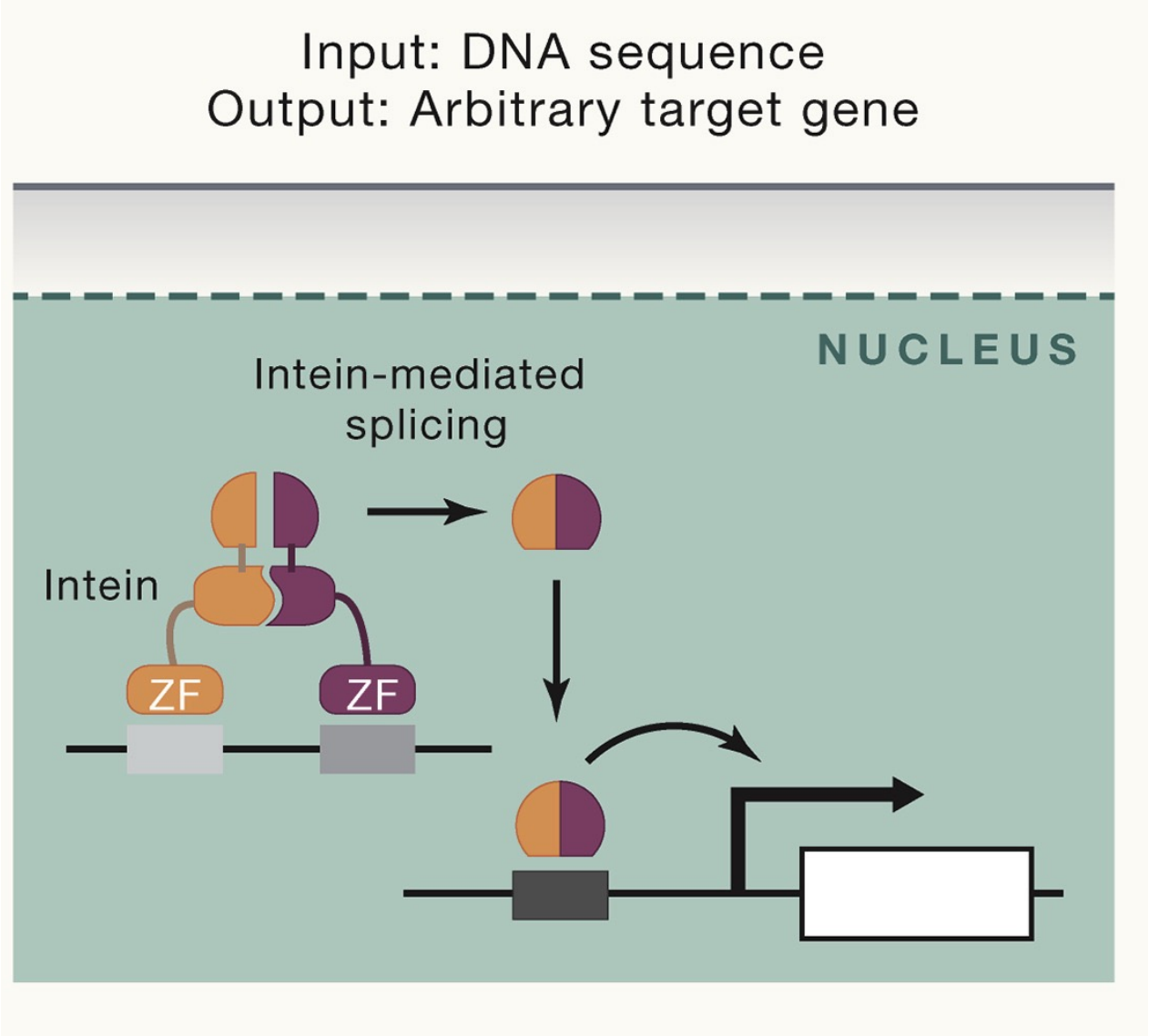


Reversible chemical modifications of proteins enable a huge range of natural signal processing capabilities inside cells

Versatile and reliable approach to detect intracellular proteins or protein states:

colocalization-dependent reconstitution of a split effector protein

Engineered protein systems sensing intracellular nucleic acids



Converting the sensing event to a customized signal:

binding of two domains of a split protein to adjacent sites on a target DNA, where they reconstitute a functional protein

Limitations of current sensing schemes

Limitations of current sensing schemes

1. Sensors rely on highly specific target-binding domains.
Hard to efficiently and reliably design binders selective in cellular environment

Limitations of current sensing schemes

1. Sensors rely on highly specific target-binding domains. Hard to efficiently and reliably design binders selective in cellular environment

2. Sensors rely on binding. How can binding be converted into an arbitrary downstream protein signal (degradation, post-translational modifications) ? Current sensors are not very modular and produce a single type of output (e.g., reconstitution or release of an effector protein). Need more flexibility in signal conversion.

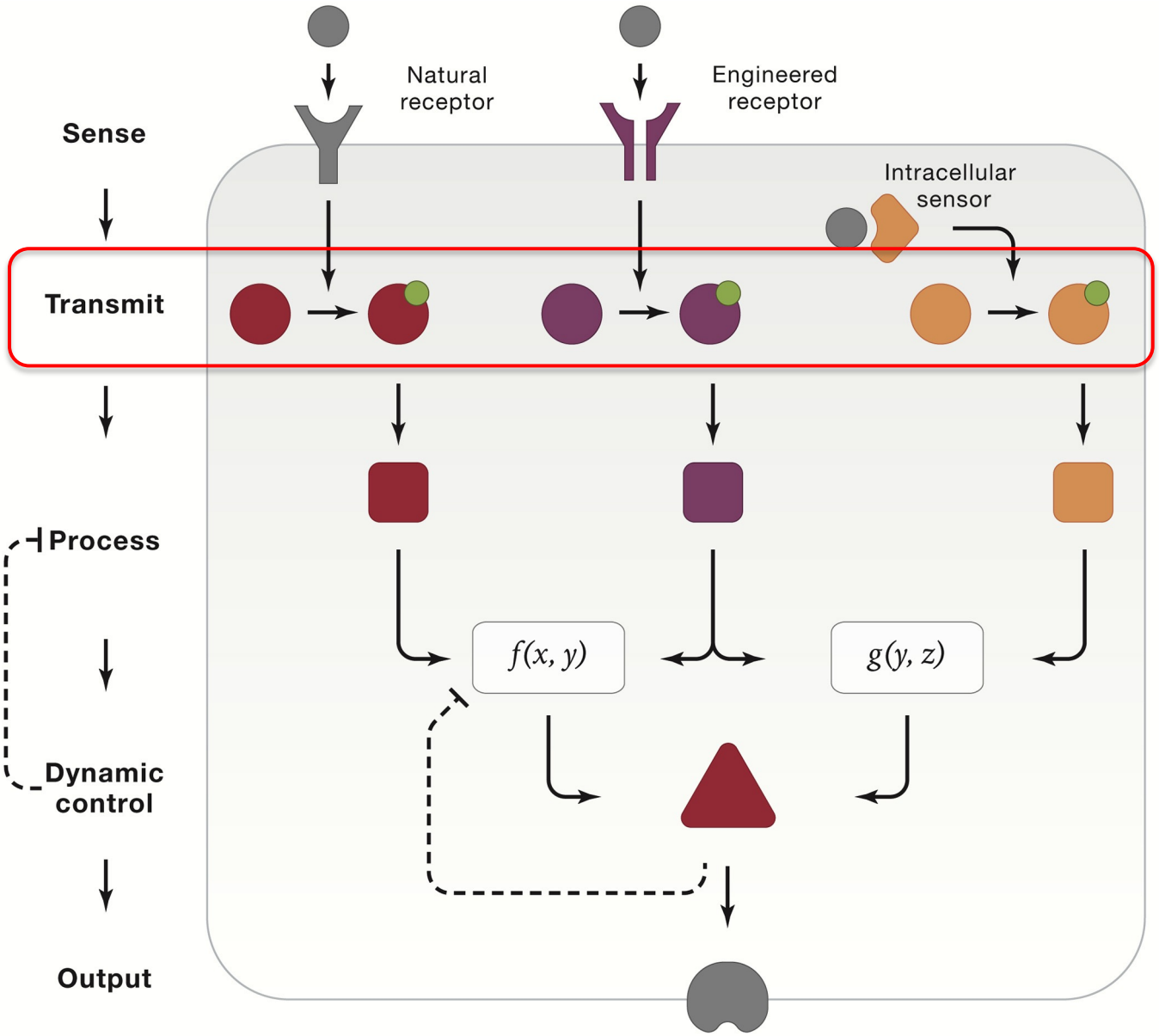
Limitations of current sensing schemes

1. Sensors rely on highly specific target-binding domains.
Hard to efficiently and reliably design binders selective in cellular environment

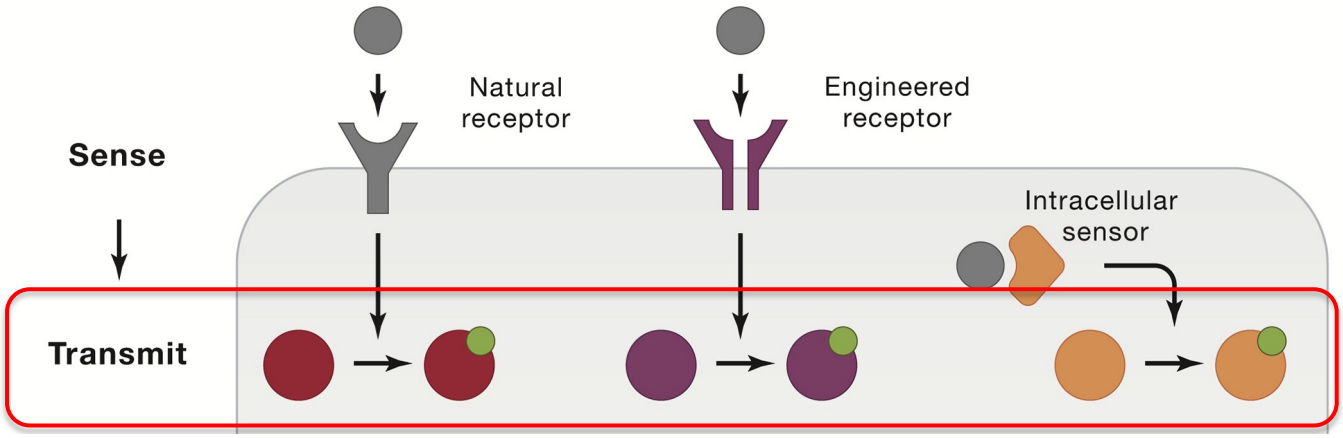
2. Sensors rely on binding. How can binding be converted into an arbitrary downstream protein signal (degradation, post-translational modifications) ? Current sensors are not very modular and produce a single type of output. Need more flexibility in signal conversion.

3. Current technology relies on genetically fusing effector domains onto the proteins of interest or having nanobodies that directly bind to surface hotspots of target proteins
potential interference with the folding or function of the detected proteins

Signal transmission



Signal transmission



Once information is sensed, protein circuits must:

1. **propagate** it.
2. **convert** it among distinct molecular carriers (e.g., from protein concentration to phosphorylation)
3. physically **transport** it from one cellular compartment to another.

Two types of implementation of these information transmission steps:

1. **directly**, through *specific protein-protein interactions*
2. **indirectly**, through *scaffolds that recruit source and destination molecules to the same site*, facilitating their interactions.

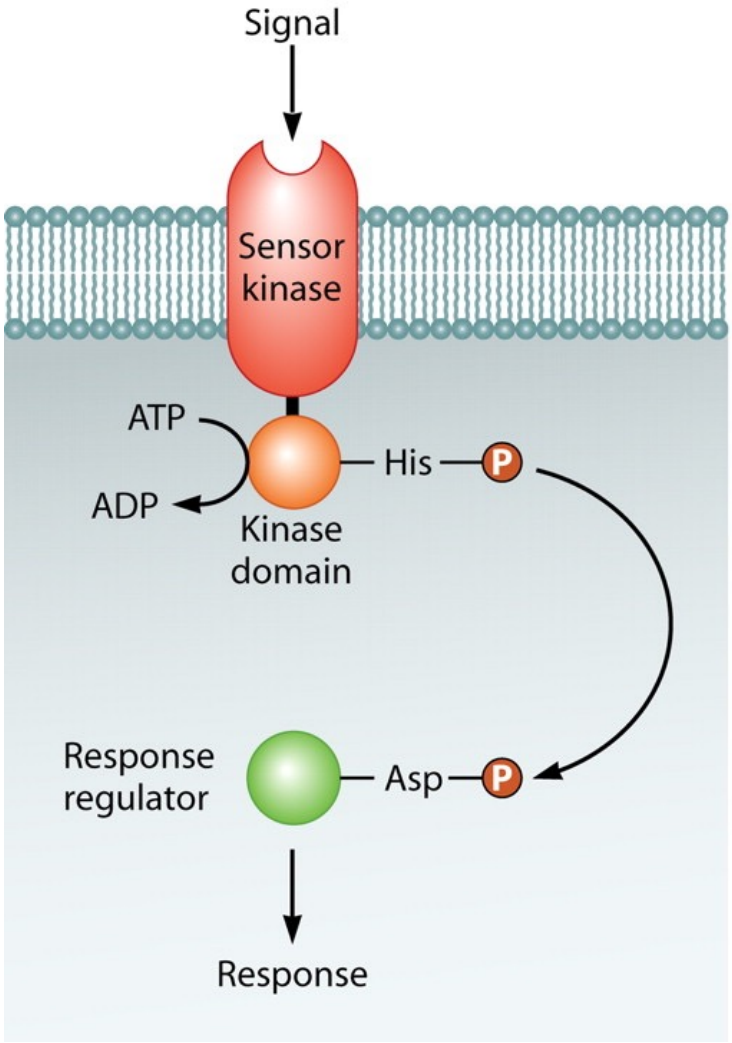
Direct signal transmission

The most direct way to control transmission is through a protein-protein interaction specificity "code" in which different amino acid sequence variants on one protein specifically interact with corresponding sequences on another protein.

Binding orthogonality is key

DO YOU KNOW A NATURAL PROTEIN SYSTEM THAT HAS BEEN EVOLVED TO DO SO ?

The bacterial 2-component signal transduction system

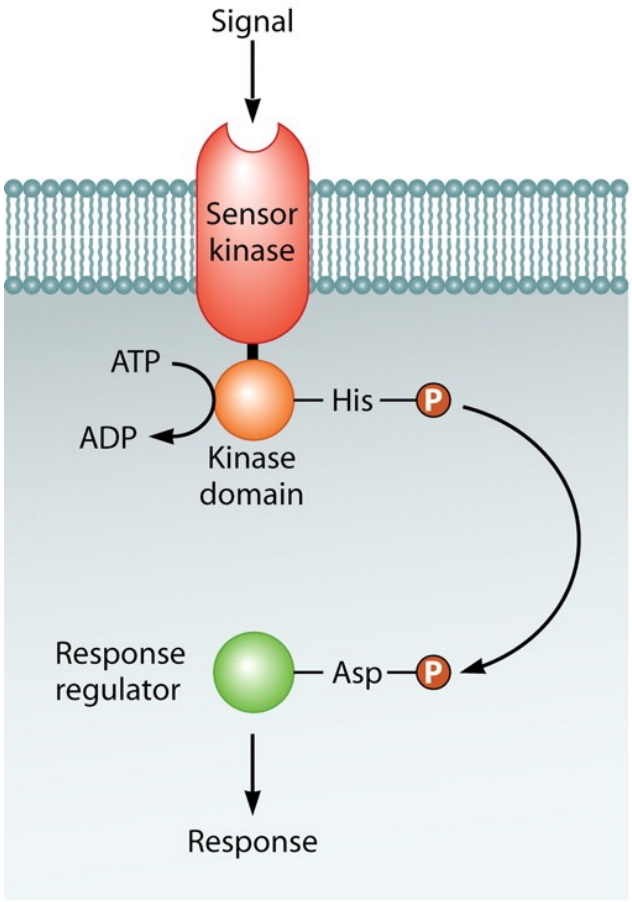


Histidine kinases transfer phosphates specifically to cognate response regulators.

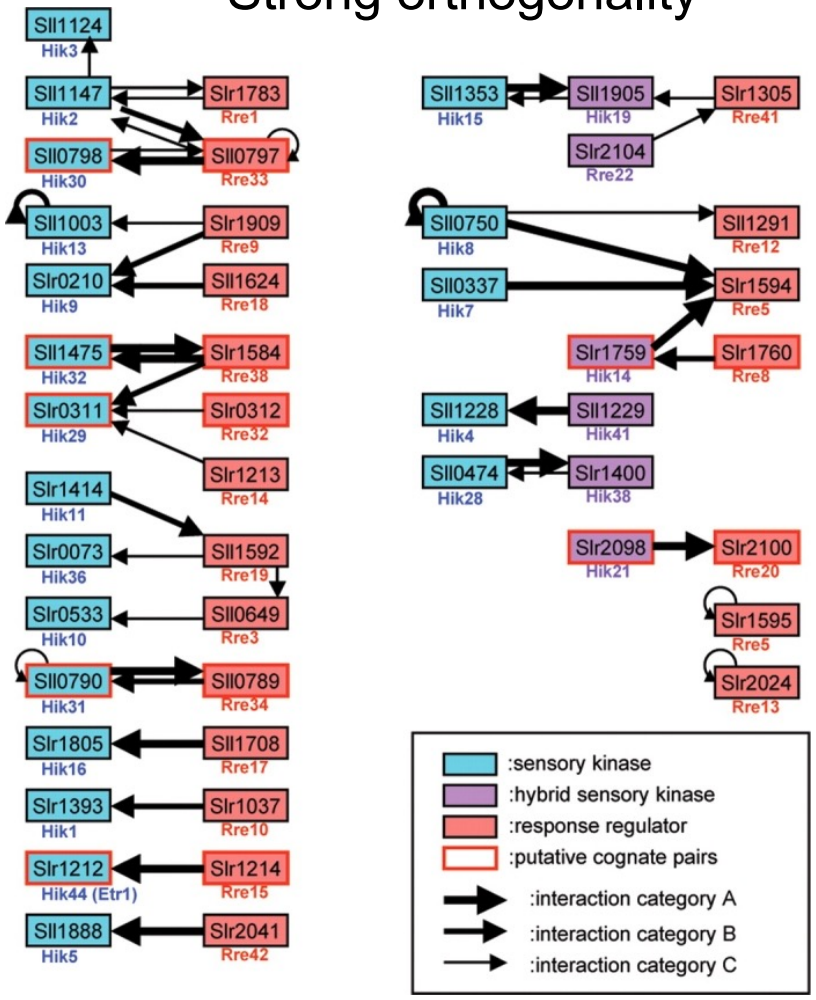
A handful of specificity- determining residues control which histidine kinase interacts with which response regulator, usually in an orthogonal, or one-to-one, manner.

Altering the specificity-determining residues is sufficient to rewire kinase connections.

The bacterial 2-component signal transduction system

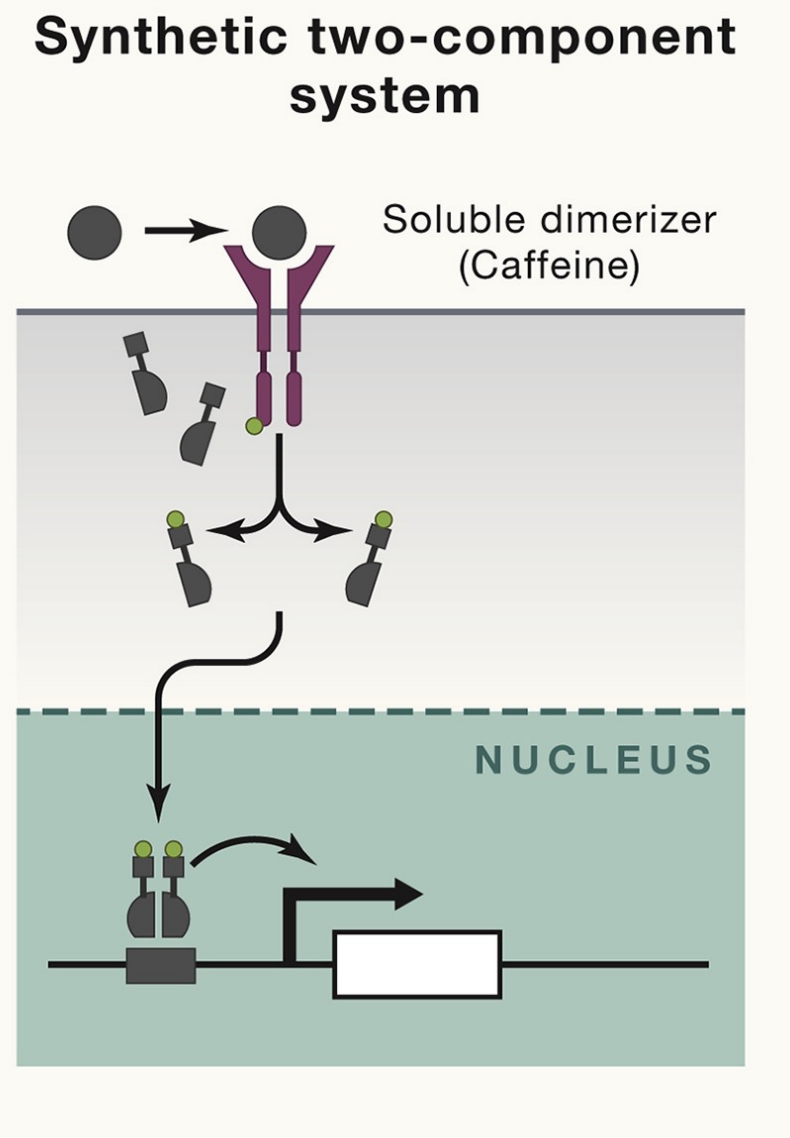


Strong orthogonality



The well-studied specificity code,
in combination with the modularity of two-component systems and their absence in higher eukaryotes makes it ideal as a synthetic mammalian signal transmission system

Repurposing the bacterial 2-component signal transduction system



Transplanted in mammalian cells

Sense [small molecules]

Non responsive to endogenous molecules

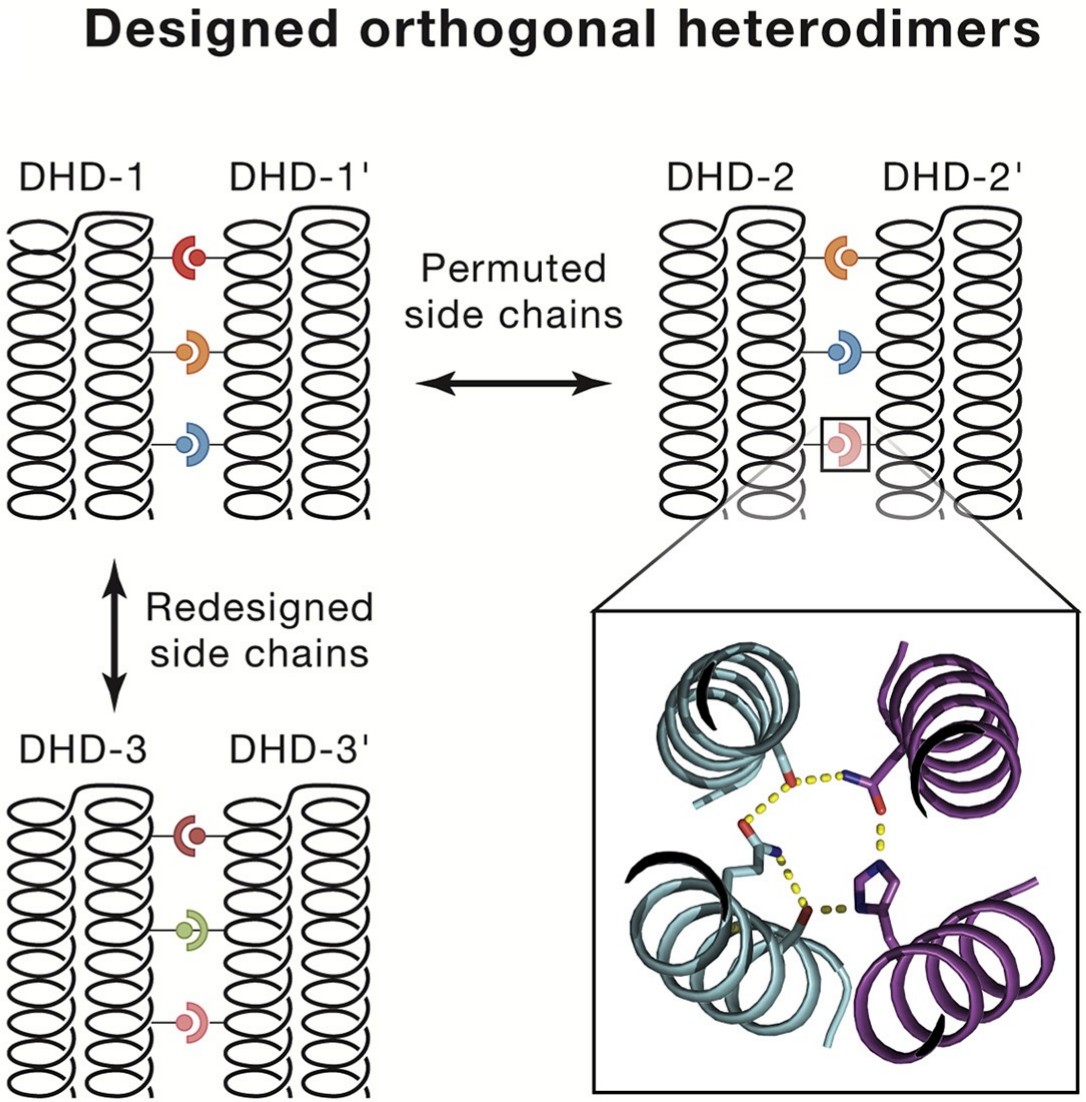
De novo protein-mediated information transmission

Could one engineer a similar specificity-determining code from scratch?

One solution: de novo design of a protein-protein interaction code based on accurate placement of buried hydrogen bond networks at the binding interfaces between structurally repeating protein helical bundles

=> high level of orthogonality. Why?

De novo protein-mediated information transmission



Indirect protein-mediated signal transmission

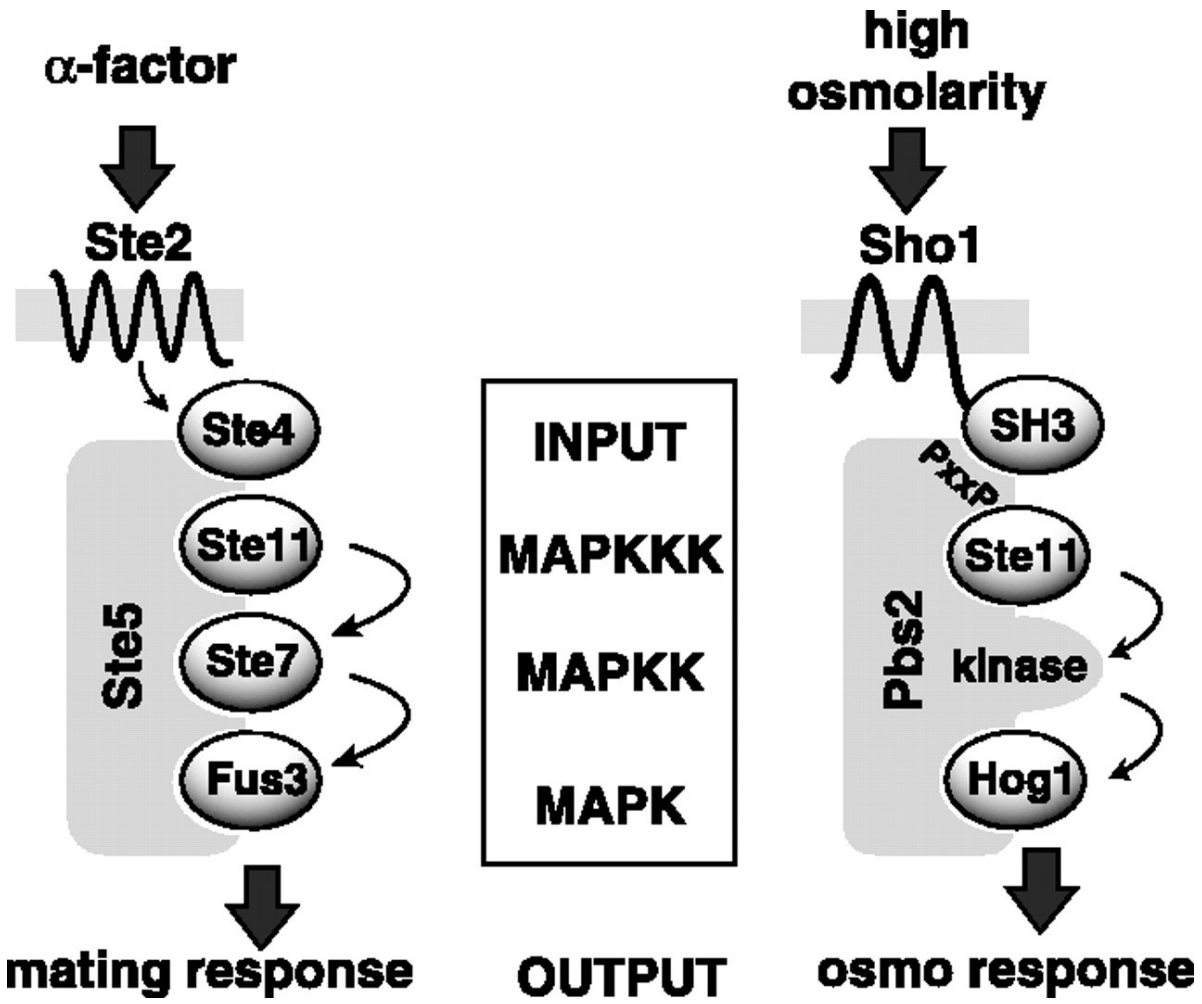
How is this achieved in native cells?

In natural circuits, scaffold proteins route signals by recruiting otherwise weakly interacting proteins into close proximity

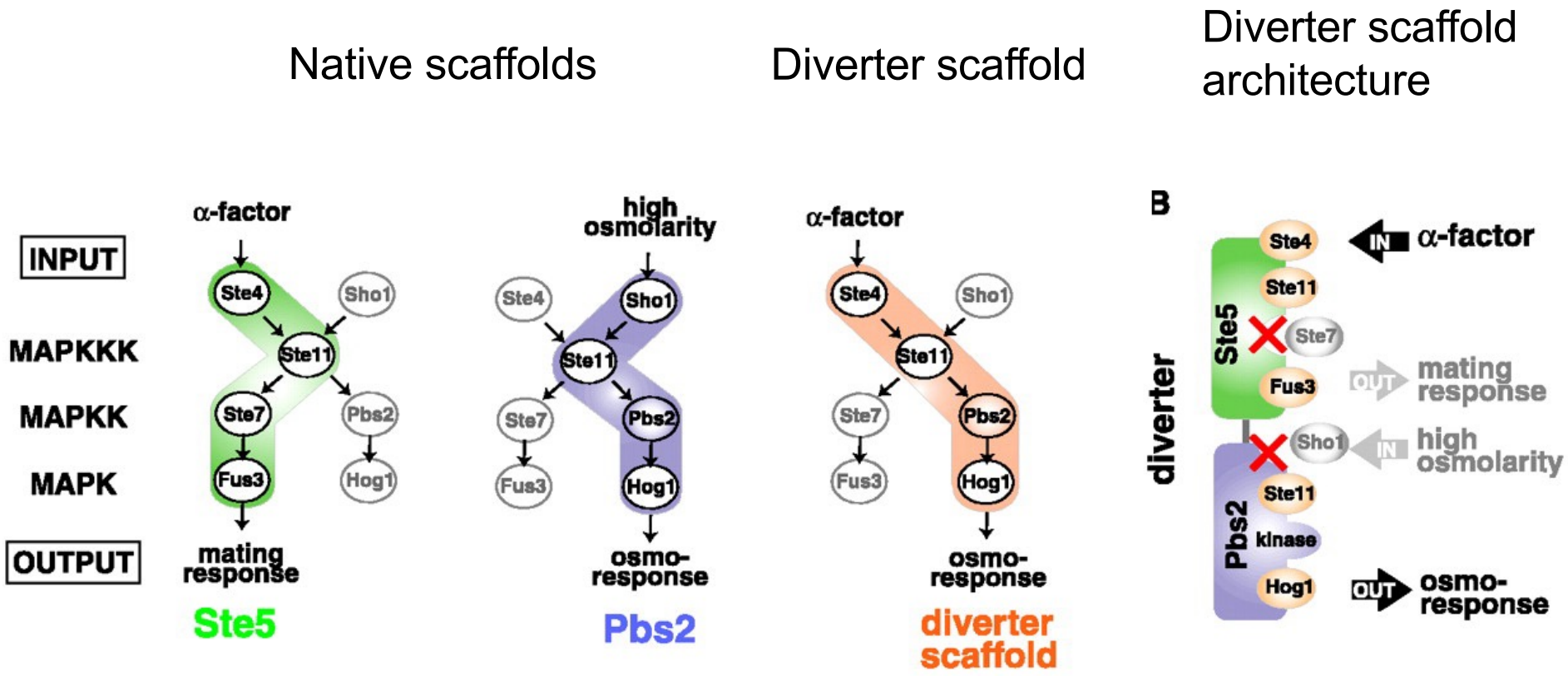
Engineered scaffold systems can similarly allow controlled routing through swappable recruitment domains.

Scaffold-mediated information transmission

Yeast mating and high-osmolarity MAPK pathways require scaffold proteins Ste5 and Pbs2. Both pathways require the shared MAPKKK Ste11 but exhibit no cross-signaling under normal conditions

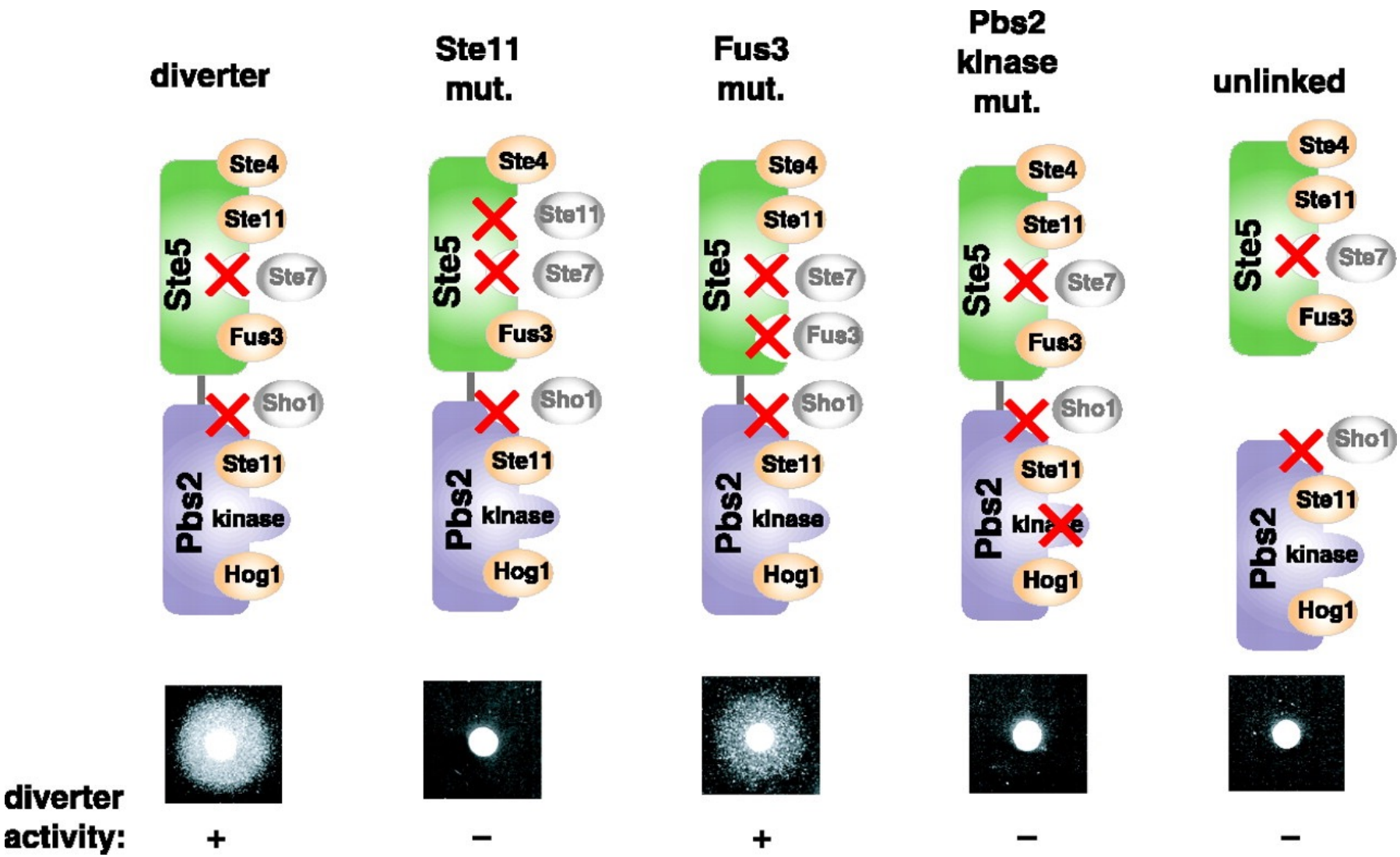


A synthetic diverter scaffold mediates an artificial MAPK pathway

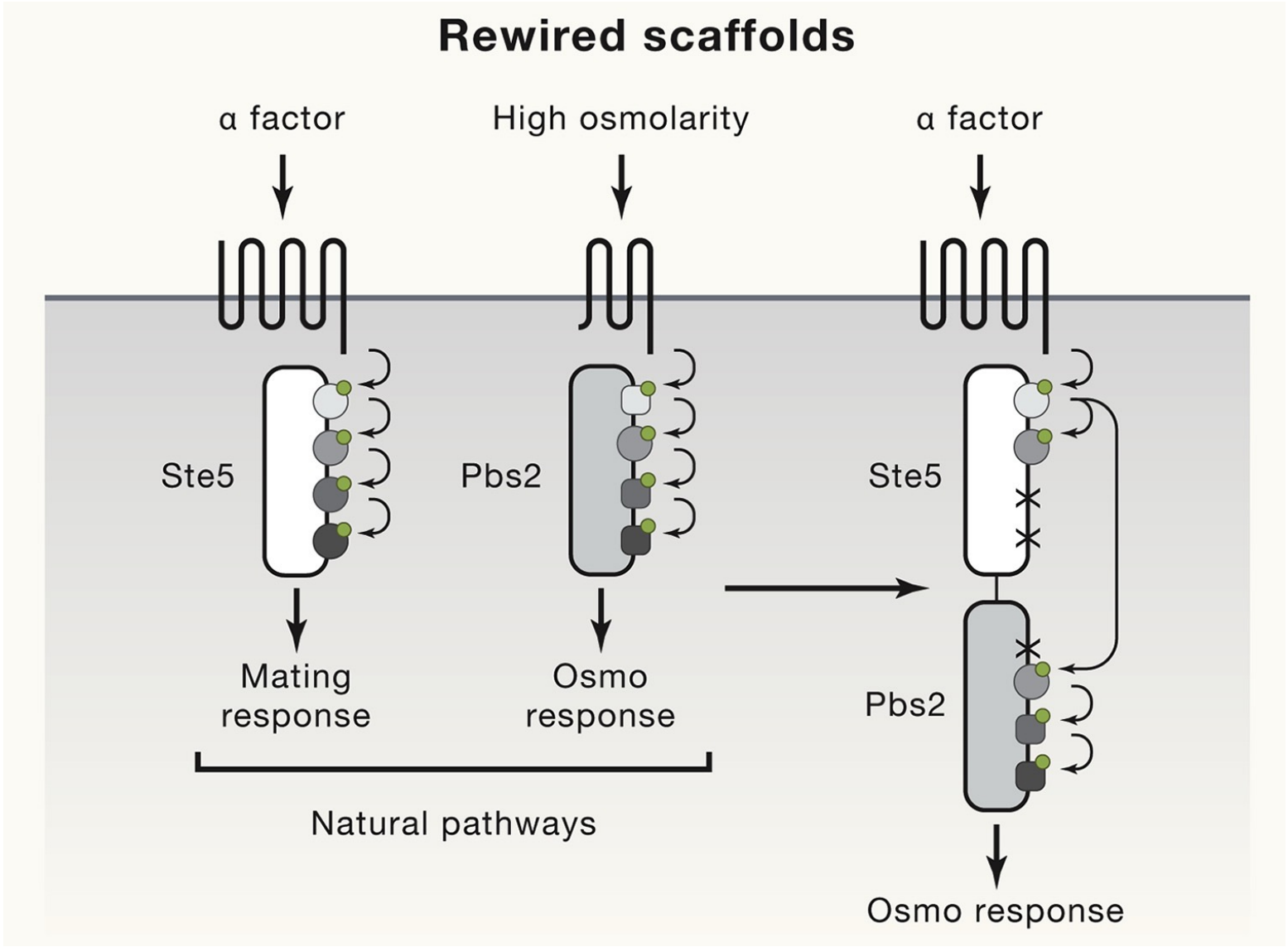


A diverter scaffold through genetic fusion of Ste5 and Pbs2 would mediate a novel input-output linkage via Ste11 (α-factor stimulation selectively activates the osmoresponse)

Analysis of diverter scaffold requirements



Rewired scaffold-mediated information transmission



Rewired scaffold-mediated information transmission. conclusions

1. Scaffolds are conceptually similar to promoters: **modular** and **flexible organizing centers** for controlling the **flow of information** in signaling or transcription, respectively.

2. Similarly, the regulation of a **transcriptional response** can be **modulated** by simple alterations in the presence or arrangement of diverse **transcription factor docking sites**.

3. These organizing structures are **optimized for evolvability**, a property that may provide **increased fitness** in the face of constantly changing environmental challenges and signaling needs.

4. Conversely, just as **promoter engineering** can be used to control cellular behavior and to create useful tools (e.g., yeast two-hybrid systems), **scaffold engineering** may allow for systematic **manipulation of cytoplasmic signaling pathways**.

Take home messages

Protein circuits harder to design than genetic circuits

Protein circuits have more potential for encoding diverse functions especially in eukaryotic cells

Engineered sensors and signal transmitter are mostly based on designed binding specificity

Questions?