

What is claimed is:

CLAIMS

1. A method for the discovery and evolution of metabolic pathways that produce
5 molecules that modulate protein function comprising
contacting a population of host cells that comprise a protein of interest with a
population of expression vectors comprising different metabolic pathways, wherein the host
cells are amenable to transfer of the population of expression vectors;
expressing the metabolic pathways in the population of host cells, wherein a cell or
10 subset of the population of host cells produce a detectable output when the metabolic
pathway produces a product that modulates the protein of interest;
screening the population of host cells under conditions that enable measurement of the
detectable output in the cell or the subset of the population of host cells;
isolating the cell or the subset of the population of host cells that produce a detectable
15 output;
isolating the expression vectors that yield detectable outputs higher than the output of
a reference vector that harbors a reference pathway, for example, a vector that encodes a
pathway that does not produce molecules with concentrations and/or potencies sufficient to
modulate the activity of a protein of interest, in the cell or the subset of the population of host
20 cells; and
characterizing the products of the metabolic pathways encoded by the expression
vectors that yield detectable outputs that are higher than the output of said reference vector in
the cell or the subset of the population of host cells.
- 25 2. The method of claim 1, wherein the host cells comprise a genetically encoded system
in which the activity of a protein of interest controls the assembly of a protein complex with
an activity that is not possessed by either of two or more components of the complex and,
thus, yields a detectable output in proportion to the amount of complex formed.
- 30 3. The method of claim 1 or claim 2, wherein the protein of interest is an enzyme that
adds a post-translational modification that causes two proteins, which are initially
dissociated, to be covalently linked or to form a noncovalent complex.

4. The method of any one of claims 1-3, wherein the complex is formed by two proteins with a dissociation constant (K_d) less than or equal to the K_d of the complexes formed between SH2 domains and their phosphorylated substrates.
- 5 5. The method of any one of claims 1-4, wherein the metabolic pathways produce phenylpropanoids or nonribosomal peptides.
6. The method of any one of claims 1-5, wherein the expression vectors comprising different metabolic pathways comprise a library of pathways generated by mutating one or
10 more genes within a starting metabolic pathway.
7. The method of any one of claims 1-6, wherein one or more of the metabolic pathways comprises a set of genes of unknown biosynthetic capability.
- 15 8. The method of any one of claims 1-7, wherein one or more of the metabolic pathways that produces a detectable output higher than the output of the reference pathway produces a product that differs from the products of other metabolic pathways.
9. The method of any one of claims 1-8, wherein one or more of the metabolic pathways
20 that produces a detectable output higher than the output of the reference pathway produces a larger quantity of a product than the quantity of product generated by other metabolic pathways.
10. The method of any one of claims 1-9, wherein one or more of the metabolic pathways
25 that produces a detectable output higher than the output of the reference pathway exhibits a lower cellular toxicity than other metabolic pathways.
11. The method of any one of claims 1-10, wherein the products of the metabolic pathways are characterized by standard analytical methods, preferably by gas
30 chromatography-mass spectrometry (GC/MS), liquid chromatography-mass spectrometry (LC/MS), and/or nuclear magnetic resonance (NMR) spectroscopy.
12. The method of any one of claims 1-11, further comprising isolating the products.

13. The method of claim 12, further comprising concentrating the products, preferably using a rotary evaporator.
- 5 14. The method of claim 12 or claim 13, further comprising testing the effects of the products on the protein of interest.
- 15 15. The method of any one of claims 1-14, wherein the protein of interest is a ubiquitin ligase, a SUMO transferase, a methyltransferase, a demethylase, an acetyltransferase, a glycosyltransferase, a palmitoyltransferase, or a related hydrolase.
- 10 16. A composition or system comprising a population of host cells that comprise a protein of interest and a population of expression vectors comprising different metabolic pathways, wherein a cell or subset of the population of host cells produce a detectable output when the metabolic pathway produces a product that modulates the protein of interest, optionally wherein the expression vectors yield detectable outputs higher than the output of a reference vector that harbors a reference pathway, for example, a vector that encodes a pathway that does not produce molecules with concentrations and/or potencies sufficient to modulate the activity of a protein of interest, in the cell or the subset of the population of host cells.
- 15 17. The composition or system of claim 16, wherein the host cells comprise a genetically encoded system in which the activity of a protein of interest controls the assembly of a protein complex with an activity that is not possessed by either of two or more components of the complex and, thus, yields a detectable output in proportion to the amount of complex formed.
- 20 18. The composition or system of claim 16 or claim 17, wherein the protein of interest is an enzyme that adds a post-translational modification that causes two proteins, which are initially dissociated, to be covalently linked or to form a noncovalent complex.
- 25 30

19. The composition or system of any one of claims 16-18, wherein the complex is formed by two proteins with a dissociation constant (K_d) less than or equal to the K_d of the complexes formed between SH2 domains and their phosphorylated substrates.
- 5 20. The composition or system of any one of claims 16-19, wherein the metabolic pathways produce phenylpropanoids or nonribosomal peptides.
21. The composition or system of any one of claims 16-20, wherein the expression vectors comprising different metabolic pathways comprise a library of pathways generated by
10 mutating one or more genes within a starting metabolic pathway.
22. The composition or system of any one of claims 16-21, wherein one or more of the metabolic pathways comprises a set of genes of unknown biosynthetic capability.
- 15 23. The composition or system of any one of claims 16-22, wherein one or more of the metabolic pathways that produces a detectable output higher than the output of the reference pathway produces a product that differs from the products of other metabolic pathways.
24. The composition or system of any one of claims 16-23, wherein one or more of the
20 metabolic pathways that produces a detectable output higher than the output of the reference pathway produces a larger quantity of a product than the quantity of product generated by other metabolic pathways.
25. The composition or system of any one of claims 16-24, wherein one or more of the
25 metabolic pathways that produces a detectable output higher than the output of the reference pathway exhibits a lower cellular toxicity than other metabolic pathways.
26. The composition or system of any one of claims 16-25, wherein the protein of interest is a ubiquitin ligase, a SUMO transferase, a methyltransferase, a demethylase, an
30 acetyltransferase, a glycosyltransferase, a palmitoyltransferase, or a related hydrolase.
27. A kit comprising a population of expression vectors of any one of claims 16-26.

28. The kit of claim 27, further comprising the population of host cells of any one of claims 16-26.