

DISCOVERY AND EVOLUTION OF BIOLOGICALLY ACTIVE METABOLITES**RELATED APPLICATION**

5 This application claims the benefit under 35 U.S.C. § 119(e) of U.S. provisional application No. 62/958,368, filed January 8, 2020, which is incorporated by reference herein in its entirety.

GOVERNMENT SUPPORT

10 This invention was made with U.S. Government support under grant 1750244 awarded by the National Science Foundation. The Government has certain rights to this invention.

FIELD

15 Disclosed herein are systems, methods, reagents, apparatuses, vectors, and host cells for the discovery and evolution of metabolic pathways that produce small molecules that modulate enzyme function.

BACKGROUND

20 Natural products and their derivatives represent a longstanding source of pharmaceuticals and medicinal preparations¹⁻³. These molecules—perhaps, as a result of their biological origin—tend to exhibit favorable pharmacological properties (e.g., bioavailability and “metabolite-likeness”)^{1,4} and can exert a striking variety of therapeutic effects (e.g., analgesic, antiviral, antineoplastic, anti-inflammatory, cytotoxic, immunosuppressive, and immunostimulatory)⁵⁻¹⁰. Recent advances in synthetic biology and metabolic engineering
25 have supplied new approaches for the efficient biosynthesis and functionalization of known, pharmaceutically relevant natural products¹¹⁻¹³; complementary methods for the discovery and optimization of new products with specific, therapeutically relevant activities, however, remain underdeveloped¹⁴.

30 Existing strategies for natural product discovery are largely undirected and/or limited in scope. For example, screens of large natural product libraries—augmented, on occasion, with combinatorial (bio)chemistry¹⁵⁻¹⁷—have uncovered molecules with important medicinal properties¹⁸, but these screens are resource-intensive and largely subject to serendipity¹⁹. Bioinformatic tools, by contrast, permit the identification of biosynthetic gene clusters^{20,21},

where co-localized resistance genes, if present, can reveal the biochemical function of their products²². The therapeutic activities of many pharmaceutically relevant metabolites, however, differ from their native functions²³, and most biosynthetic pathways can, when appropriately reconfigured, yield entirely new—and, perhaps, more effective—therapeutic molecules^{12,24}.

Microbial systems have emerged as powerful platforms for the biosynthesis of natural products from unculturable or low-yielding organisms.^{25,26} Recent work showed that such systems can also permit the discovery and evolution of metabolic pathways with specific, therapeutically relevant activities (PCT/US2019/40896).

SUMMARY

Disclosed herein are systems, methods, reagents, apparatuses, vectors, and host cells for the discovery and evolution of metabolic pathways that produce small molecules that modulate enzyme function. For example, a microorganism is provided in which a first genetically encoded system links cell growth to the activity of a target enzyme and in which a second genetically encoded system—to be discovered or evolved—produces a metabolite that modulates the activity of the target enzyme. This disclosure applies this approach to a subset of target enzymes that post-translationally modify proteins, to metabolic pathways that produce phenylpropanoids or nonribosomal peptides, and to the discovery of cryptic metabolic pathways. Some aspects of this disclosure provide specific reconfigured or evolved pathways that produce specific modulators of enzyme activity, that yield improved titers of such modulators (relative to a starting pathway), and/or that exhibit reduced host toxicity (relative to a starting pathway). Metabolic products with specific inhibitory effects are also disclosed.

According to one aspect, methods for the discovery and evolution of metabolic pathways that produce molecules that modulate protein function are provided. The methods include contacting a population of host cells that comprise a protein of interest, such as an enzyme 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 subset of the population of host cells produce a detectable output when the metabolic pathway within said cell or population of host cells produces a product that modulates the protein of interest, such as the enzyme 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 output; isolating the expression vectors that yield detectable outputs higher than ($p < 0.05$) 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, such as an enzyme of interest, in the cell or the subset of the population of host 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.

In some embodiments, the host cells comprise a genetically encoded system in which the activity of a protein of interest, such as an enzyme 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.

In some embodiments, 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.

In some embodiments, 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.

In some embodiments, the enzyme of interest is an enzyme that adds a post-translational modification other than the addition or removal of a phosphate, and that modification causes two proteins, which are initially dissociated inside of the cell, to be covalently linked or to form a complex with a dissociation constant (K_d) less than or equal to the K_d of the complex formed between a SH2 domain and a phosphorylated SH2-substrate domain (e.g., as shown in FIG. 1a).

In some embodiments, the metabolic pathways produce phenylpropanoids or nonribosomal peptides.

In some embodiments, the expression vectors comprising different metabolic pathways comprise a library of pathways generated by mutating one or more genes within a starting metabolic pathway.

In some embodiments, one or more of the metabolic pathways comprises a set of genes of unknown biosynthetic capability.

In some embodiments, one or more of the metabolic pathways that produces a detectable output higher than the output of the reference pathway produces a product that
5 differs from the products of other metabolic pathways.

In some embodiments, one or more of the 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.

In some embodiments, one or more of the metabolic pathways that produces a
10 detectable output higher than the output of the reference pathway exhibits a lower cellular toxicity than other metabolic pathways.

In some embodiments, the products of the metabolic pathways are characterized by standard analytical methods, preferably by gas chromatography–mass spectrometry (GC/MS), liquid chromatography-mass spectrometry (LC/MS), and/or nuclear magnetic
15 resonance (NMR) spectroscopy.

In some embodiments, the methods further include isolating the products.

In some embodiments, the methods further include concentrating the products, preferably using a rotary evaporator.

In some embodiments, the methods further include testing the effects of the products
20 on the protein of interest, such as the enzyme of interest.

In some embodiments, the protein of interest, such as the enzyme of interest, is a ubiquitin ligase, a SUMO transferase, a methyltransferase, a demethylase, an acetyltransferase, a glycosyltransferase, a palmitoyltransferase, or a related hydrolase.

In some embodiments, the products or molecules identified (e.g., amorphadiene and
25 derivatives, taxadiene and derivatives, β -bisabolene and derivatives, α -bisabolene and derivatives, and α -longipinene and derivatives) are provided as drugs or drug leads for the treatment of diseases to which PTPs contribute, for example, type 2 diabetes, HER2-positive breast cancer, or Rett syndrome, as are methods of treatment of such diseases by administering an effective amount of the molecule(s) to a subject in need of such treatment.

30 According to another aspect, compositions or systems are provided that include 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, and 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
5 the cell or the subset of the population of host cells.

In some embodiments, 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.

10 In some embodiments, 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.

In some embodiments, 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
15 their phosphorylated substrates.

In some embodiments, the metabolic pathways produce phenylpropanoids or nonribosomal peptides.

In some embodiments, the expression vectors comprising different metabolic pathways comprise a library of pathways generated by mutating one or more genes within a
20 starting metabolic pathway.

In some embodiments, one or more of the metabolic pathways comprises a set of genes of unknown biosynthetic capability.

In some embodiments, one or more of the metabolic pathways that produces a detectable output higher than the output of the reference pathway produces a product that
25 differs from the products of other metabolic pathways.

In some embodiments, one or more of the 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.

In some embodiments, one or more of the metabolic pathways that produces a
30 detectable output higher than the output of the reference pathway exhibits a lower cellular toxicity than other metabolic pathways.

In some embodiments, 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.

According to another aspect, kits are provided that include a population of expression
5 vectors as described herein. In some embodiments, the kits also include the population of host cells that comprise a protein of interest as described herein.

Each of the limitations of the invention can encompass various embodiments of the invention. It is therefore anticipated that each of the limitations of the invention involving any one element or combinations of elements can be included in each aspect of the invention.
10 This invention is not limited in its application to the details of construction and the arrangement of components set forth in the following description or illustrated in the drawings. The invention is capable of other embodiments and of being practiced or of being carried out in various ways.

15 BRIEF DESCRIPTION OF DRAWINGS

FIGs. 1a-1e. Development of a bacterial-two hybrid system that links the inhibition of PTP1B to antibiotic resistance. FIG. 1a, A bacterial two-hybrid (B2H) system that detects phosphorylation-dependent protein-protein interactions. Major components include (i) a substrate domain fused to the omega subunit of RNA polymerase (yellow), (ii) an SH2
20 domain fused to the 434 phage cI repressor (light blue), (iii) an operator for 434cI (dark green), (iv) a binding site for RNA polymerase (purple), (v) Src kinase, and (vi) PTP1B. Src-catalyzed phosphorylation of the substrate domain enables a substrate-SH2 interaction that activates transcription of a gene of interest (GOI, black). PTP1B-catalyzed dephosphorylation of the substrate domain prevents that interaction; inhibition of PTP1B re-enables it. **FIG. 1b,**
25 A version of the B2H system that both (i) lacks PTP1B and (ii) contains p130cas as the substrate domain and luxAB as the GOI. Inducible plasmids were used to increase expression of specific components in *E. coli*; secondary induction of Src from one such plasmid enhanced luminescence. **FIG. 1c,** A version of the B2H system that both (i) lacks PTP1B and Src and (ii) includes an SH2 domain (SH2*) with an enhanced affinity for phosphopeptides, a
30 variable substrate domain, and LuxAB as the GOI. An inducible plasmid was used to increase expression of Src in *E. coli*. Sequences for substrates p130cas (SEQ ID NO: 24), MidT (SEQ ID NO: 25), EGFR (SEQ ID NO: 27), and ShcA (SEQ ID NO: 26) are shown. **FIG. 1d,** The B2H system from c with either p130cas or MidT as substrates. A second plasmid was used to

overexpress either (i) Src and PTP1B or (ii) Src and an inactive variant of PTP1B (C215S) in *E. coli*. Right: Two single-plasmid B2H systems. **FIG. 1e**, The optimized system includes SH2*, the midT substrate, optimized promoters and ribosome binding sites (bb034 from **FIG. 1d**), and SpecR as the GOI. Inactivation of PTP1B enabled a strain of *E. coli* harboring this plasmid-borne system to survive at high concentrations of spectinomycin (> 250 µg/ml). Error bars in **FIG. 1b- FIG. 1d** denote standard error with n = 3 replicates.

FIGs. 2a-2c. Biosynthesis of PTP1B-inhibiting terpenoids enables cell survival. FIG. 2a,

A plasmid-borne pathway for terpenoid biosynthesis: (i) pMBIS, which harbors the mevalonate-dependent isoprenoid pathway of *S. cerevisiae*, converts mevalonate to isopentyl pyrophosphate (IPP) and farnesyl pyrophosphate (FPP). (ii) pTS, which encodes a terpene synthase (TS) and, when necessary, a geranylgeranyl diphosphate synthase (GGPPS), converts IPP and FPP to sesquiterpenes or diterpenes. **FIG. 2b**, Four terpene synthases: amorphaadiene synthase (ADS), γ-humulene synthase (GHS), abietadiene synthase (ABS), and taxadiene synthase (TXS). **FIG. 2c**, The spectinomycin resistance of strains of *E. coli* that harbor both (i) the bacterial two-hybrid (B2H) system (ii) a TS-specific terpenoid pathway (pTS includes GGPPS only when ABS or TXS are present). ADS enabled survival in the presence of high concentrations of spectinomycin. Note: ABS_{D404A/D621A} is catalytically inactive. B2H* contains PTP1B_{C215S}, which is inactive.

FIGs. 3a-3g. Strategy for microbially assisted directed evolution (MADE). FIG. 3a,

Error-prone PCR and/or site-saturation mutagenesis of a subset of genes within a metabolic pathway yield a library of metabolic pathways. **FIG. 3b**, Microbes, each of which harbors both (i) the B2H system and (ii) a member of the pathway library, are grown in liquid culture.

Note: The system shown is an *E. coli* host that harbors both (i) the B2H system and (ii) mutated terpenoid pathways (i.e., pMBIS + pTS with mutations; see Fig. 2a). **FIG. 3c**, After liquid culture, the transformants are plated on solid media with different concentrations of antibiotic; hits comprise colonies that grow at antibiotic concentrations at which the wild-type pathway does not permit growth. **FIG. 3d**, The pathways of the hits are sequenced; their mutations are reintroduced into the wild-type pathway; and these reconstructed pathway variants are rescreened with drop-based plating (10 µL) on solid media with different concentrations of antibiotic. This step removes false positives (e.g., colonies that survived because of mutations located outside of the target genes). **FIG. 3e**, The confirmed hits are

grown in liquid culture; their products are extracted with a hexane overlay, as needed, and concentrated in a rotary evaporator. **FIG. 3f**, GC/MS enables the identification and quantification of mutant products; NMR can assist with identification. **FIG. 3g**, Interesting metabolites (purchased or purified from culture extract) are characterized with *in vitro* kinetic measurements or cell studies of target modulation and/or ITC analyses of target-metabolite binding.

FIGs. 4a-4d. Genetically encoded systems that detect metabolite-mediated modulation of post-translational modification (PTM) enzymes.

FIG. 4a, A genetically encoded system that detects metabolite-mediated activation of enzymes E1 and/or E2. E1 adds a PTM to protein P1, allowing it to bind to P2; the newly formed P1-P2 complex activates transcription of a gene of interest (GOI, black). E2 removes the PTM from P1 and, thus, prevents complex formation. When the GOI confers a fitness advantage, inhibitors of E2 or activators of E1 enhance cell survival. When the GOI is toxic, inhibitors of E1 or activators of E2 enhance cell survival. **FIG. 4b**, An alternative detection system. E1 adds a PTM to protein P1, allowing it to bind to P2; the newly formed P1-P2 complex assembles a split protein (e.g., a fluorescent protein, a luciferase, or an enzyme that confers antibiotic resistance). E2 removes the PTM from P1 and, thus, prevents complex formation. When the reconstituted split protein confers a fitness advantage, inhibitors of E2 or activators of E1 enhance cell survival. When, by contrast, the reconstituted protein is toxic, inhibitors of E1 or activators of E2 enhance cell survival. **FIG. 4c**, A genetically encoded system that detects metabolite-mediated activation of PTM enzymes that control protein ligation (e.g., a SUMO transferase, a ubiquitin ligase, or associated peptidases). E1 attaches P1 to a lysine residue (K) of P2, and the newly formed P1-P2 complex activates transcription of a GOI. E2 breaks this complex apart. **FIG. 4d**, An alternative system. E1 attaches P1 to P2, and the newly formed P1-P2 complex permits the assembly of a split protein. E2-mediated proteolysis breaks this complex apart.

FIGs. 5a-5c. Alternative metabolic pathways.

FIG. 5a, Phenylpropanoid pathways developed by Young-Soo Hong and colleagues⁴⁵. Abbreviations: TAL, ammonia-lyase from *S. espanaensis*; Sam5, 4-coumarate 3-hydroxylase from *S. espanaensis*; COM, O-methyltransferase from *A. thaliana*; ScCCL, cinnamate/4-coumarate:CoA ligase from *Streptomyces coelicolor*; CHS, chalcone synthase from *A. thaliana*; STS, stilbene synthase from *Arachis hypogaea*. **FIG. 5b**, The pathways encoded by the plasmids from **FIG. 5a**.

FIG. 5c, A genetically encodable yersiniabactin (Ybt) synthetase, as described by Khosla and colleagues⁴⁶. Ybt is a polyketide-nonribosomal peptide. The substrates necessary for Ybt production appear in blue. Abbreviations: ArCP, aryl carrier protein; A, adenylation; PCP, peptidyl carrier proteins; Cy, cyclization; KS, ketosynthase; ACP, acyl carrier protein; AT, acyltransferase; KR, NADPH-dependent ketoreductase; MT, methyltransferase; SAM, S-adenosylmethionine; TE, thioesterase. See the text for details on biosynthesis.

FIGs. 6a-6b. An approach for the discovery of cryptic metabolic pathways. FIG. 6a, Mutagenesis and/or reorganization of a multi-step pathway inactivates a biosynthetic gene and, thus, permits the accumulation of a metabolic intermediate. **FIG. 6b**, Mutagenesis and/or reorganization of a multi-step pathway inactivates a repressor gene and, thus, permits the expression of pathway genes.

FIGs. 7a-7i. Microbial evolution of terpenoid inhibitors. FIG. 7a- FIG. 7b, Homology models for (**FIG. 7a**) ADS and (**FIG. 7b**) GHS show the locations of residues targeted for site-saturation mutagenesis (SSM). A substrate analogue from an aligned structure of 5-epi-aristolochene synthase (pdb entry 5eat) appears in blue. **FIG. 7c- FIG. 7d**, Measurements of the spectinomycin resistance conferred by mutants of (**c**) ADS (LB plates) and (**FIG. 7d**) GHS (TB plates). ALP corresponds to a quintuple mutant of GHS (A336C/T445C/S484C/I562L/M565L) that generates α -longipinene as a major product. Shades denote colony densities: diffuse (≥ 10 colonies, light gray), circular diffuse (gray), and circular lawn (black). **FIG. 7e**, The product profiles of mutants of ADS that enable growth at higher antibiotic concentrations than the wild-type enzyme. **FIG. 7f**, ADS_{G43S/K51N} and ADS yield similar amorphadiene titers in liquid cultures. **FIG. 7g**, ADS_{G43S/K51N} yields higher colony densities than the wild-type enzyme in the presence of an inactive B2H system (B2Hx); these densities suggest that ADS_{G43S/K51N} is less toxic than ADS. **FIG. 7h**, The product profiles of wild-type GHS and several GHS mutants that yield enhanced antibiotic resistance; discrepancies between profiles of these mutants suggest differences in the composition of intracellular terpenoids that might give rise to enhanced antibiotic resistance. **FIG. 7i**, GHS_{A319Q} yields a higher terpenoid titer than GHS. Error bars in **FIG. 7f** and **FIG. 7i** denote standard deviation with n = 3 biological replicates.

FIGs. 8a-8d. Analysis of evolved mutants. FIG. 8a, Analysis of the antibiotic resistance conferred by mutants of ADS. Images show the growth of *E. coli* on LB plates seeded from drops of liquid culture (10 μ L). Each mutant was prepared by using site-directed mutagenesis to introduce mutations identified in the selection experiment (i.e., hits) into the starting ADS plasmid. Shades denote colony densities: diffuse (≥ 10 colonies, light gray), circular diffuse (gray), and circular lawn (black) **FIG. 8b,** A replicate of the experiment described in **FIG. 8a.** **FIG. 8c,** Analysis of the antibiotic resistance conferred by mutants of GHS. Images show the growth of *E. coli* on TB plates seeded from drops of liquid culture (10 μ L). **FIG. 8d,** A replicate of the experiment described in **FIG. 8c.** In **FIG. 8a- FIG. 8d,** blue highlights denote mutants that enabled growth at higher concentrations of spectinomycin than the wild-type enzymes in two biological replicates (i.e., these mutants appear in FIGS. 3c and 3d).

FIGs. 9a-9c. Analysis of the products of different terpene synthases. FIG. 9a, Titers of the dominant terpenoids (i.e., amorphadiene, γ -humulene, taxadiene, or abietadiene) generated by each TS-specific strain in the absence (top) and presence (bottom) of the B2H system. Similar titers indicate that the B2H system does not interfere with terpeneoid biosynthesis. **FIG. 9b,** GC/MS chromatograms of the terpenoids generated by each strain in the absence (top) and presence (bottom) of the B2H system ($m/z = 204$). Similar profiles indicate that the B2H system does not alter product distributions. **FIG. 9c,** Analysis of the contributions of either (i) TS activity or (ii) B2H function to the death and survival of various strains. Inactivation of GHS does not enhance the survival of the GHS strain, an indication that this enzyme does not produce growth-inhibiting terpenoids. Inactivation of either ADS or the B2H system, by contrast, weakens the antibiotic resistance of the ADS strain, an indication that maximal resistance requires both terpeneoid production and B2H activation. Labels denote the following controls: GHS_{D/A}, an inactive GHS; ADS_{D/A}, an inactive ADS; B2H*, a constitutively active B2H; B2H_x, an inactive B2H. Note: The left and right images show LB plates seeded with drops of liquid culture (10 μ L) from two biological replicates. Error bars in **FIG. 9a** denote standard error for $n \geq 3$ biological replicates.

FIGs. 10a-10e. Analysis of the products of various terpenoids. **FIG. 10a,** Chromatograms show expected dominant products (*) for each TS-specific strain from Fig. 2c (the B2H system is present). **FIG. 10b,** Titers of major products generated by ADS and TXS. **FIG. 10c,** Initial rates of PTP1B-catalyzed hydrolysis of pNPP in the presence of increasing

concentrations of amorphadiene and taxadiene. Lines show fits to a Michaelis-Menten model, which provides evidence of noncompetitive inhibition (amorphadiene) and mixed inhibition (taxadiene). **FIG. 10d**, A depiction of a HEK293T/17 cell. Insulin stimulates phosphorylation of the membrane-bound insulin receptor (IR); PTP1B dephosphorylates IR, and the inhibition of PTP1B restores phosphorylation. **FIG. 10e**, ELISA-based measurements of IR phosphorylation in starved wild-type HEK293T/17 cells exposed to 3% dimethyl sulfoxide (DMSO, n = 2), 930 μ M amorphadiene (AD, in 3% DMSO, n = 3), and 405 μ M α -bisabolene (Abis, 3% DMSO, n = 1) for 10 minutes. The results indicate that both amorphadiene and α -bisabolene can cross the cell membrane, inhibit intracellular PTP1B, and, thus, increase IR phosphorylation. Error bars in **FIG. 10b** denote standard error with n = 3 biological replicates. Error bars in **FIG. 10c** denote standard error with n $\geq$ 3 measurements. Error bars in **FIG. 10e** denote standard error with n values indicated (we note: for these measurements, we subtracted a reference signal produced by lysis buffer alone, n = 3).

FIGs. 11a-11d. Analysis of alternative terpene synthases. **FIG. 11a- FIG. 11b**, The spectinomycin resistance of strains of *E. coli* that harbor (i) an active or inactive bacterial two-hybrid system (B2H and B2Hx, respectively, as in FIGS. 1, 2, and 7-9) and (ii) the terpenoid pathway from FIG. 2 with each of the following terpene synthases: γ -humulene synthase from *Abies grandis* (GHS), β -bisabolene synthase from *Zingiber officinale* (ZoBBA), β -bisabolene synthase from *Santalum album* (SaBBA), and α -bisabolene synthase (ABB) from *Abies grandis* (ABS). SaBBA and, most prominently, ABB enable survival at high concentrations of spectinomycin. **FIG. 11c**, chemical structures of β -bisabolene and α -bisabolene. **FIG. 11d**, analysis of PTP1B activity on p-nitrophenyl phosphate (pNPP) in the presence of increasing concentrations of α -bisabolene (measured as amorphadiene equivalents) purified from culture extract. Lines show fits to a Michaelis-Menten Model.

FIGs. 12a-12g. Analysis of selective inhibitors of PTP1B. **FIG. 12a**, Initial rates of pNPP hydrolysis by PTP1B₃₂₁, TCPTP₂₉₂, and PTP1B₂₈₂ in the presence of increasing concentrations of amorphadiene. Lines show fits to models of inhibition. A comparison of the first and second plots (or, more specifically, the IC₅₀'s derived from the plotted data) indicates that amorphadiene is a ~ five-fold more potent inhibitor of PTP1B₃₂₁ than TCPTP₂₉₂, the most closely related PTP in the human genome (by sequence identity); this

selectivity suggests that amorphadiene binds outside of the active site of PTP1B. A comparison of the second and third plots, in turn, indicate that amorphadiene inhibits PTP1B₂₈₂ ~four-fold less potently than PTP1B₃₂₁; this discrepancy suggests that the $\alpha 7$ helix, which is present in PTP1B₃₂₁ but missing in PTP1B₂₈₂ (and which is proximal to a known allosteric binding site of PTP1B), is involved in the PTP1B₃₂₁-amorphadiene interaction. **FIG. 12b**, the chemical structure of amorphadiene. **FIG. 12c**, a preliminary crystal structure of PTP1B bound to amorphadiene. **FIG. 12d**, Data used to solve the structure in **FIG. 12c** shows electron density near the allosteric site of PTP1B (F280 appears on the left of this image); this density is consistent with the structure of amorphadiene. **FIG. 12e**, the chemical structure of α -bisabolol, a structural analogue of α -bisabolene. **FIG. 12f**, a preliminary crystal structure of PTP1B bound to α -bisabolol. **FIG. 12g**, Data used to solve the structure in **FIG. 12f** shows electron density near the allosteric site of PTP1B (F280 appears in the upper left of this image); this density is consistent with the structure of α -bisabolol.

FIG. 13. Optimization of the bacterial-two hybrid (B2H) system. FIG. 13, We optimized the transcriptional response of the B2H system by adjusting the strength of various genetic elements. In three sequential phases, we changed (1) the promoter for Src/CDC37, (2) the ribosome binding site (RBS) for Src/CDC37, and (3) and the RBS for PTP1B. In phases 1 and 2, we used a PTP1B-deficient system with either a wild-type (WT, EPQYEEIPYL (SEQ ID NO:1)) or non-phosphorylatable (Mut, EPQFEEIPYL (SEQ ID NO:2)) substrate domain. Here, “none” indicates that absence of an additional promoter; the labeled “Pro1” controls the transcription of all five genes to its left. In phase 3, we used a complete B2H system with either a wild-type (WT) or catalytically inactive (C215S, Mut) variant of PTP1B. The remaining B2H component of each phase are detailed in **TABLE 2**. Error bars denote standard error with $n \geq 3$ biological replicates.

FIG. 14. Analysis of different selection conditions. FIG. 14, A comparison of the antibiotic resistance conferred by B2H systems with different RBSs for PTP1B (see **TABLE 2** for the remaining components of each system). Images show the growth of *E. coli* on agar plates (LB) seeded from drops of liquid culture (10 μ L) with two biological replicates for each condition. The RBS bb034 confers a greater sensitivity to spectinomycin on agar plates; concentrations of spectinomycin in the liquid culture, by contrast, do not have a strong

influence on bacterial growth. Informed by this analysis, we incorporated bb034 into our “optimized” B2H system and ceased adding spectinomycin to liquid culture.

5 **FIGs. 15a-15b. FIG. 15a**, A GC chromatogram of pure amorphadiene (purchased from Ambeed). **FIG. 15b**, The mass spectrum of the indicated peak from **FIG. 15a**.

FIGs. 16a-16b. GC/MS analysis of γ -humulene production. FIG. 16a, A GC chromatogram shows the production of γ -humulene by a strain of *E. coli* engineered to produce it (i.e., pMBIS + pGHS). **FIG. 16b**, The mass spectrum of the indicated peak from
10 **FIG. 16a**.

FIGs. 17a-17b. Supplementary Fig. 4 | GC/MS analysis of abietadiene production. FIG. 17a, A GC chromatogram shows the production of abietadiene by a strain of *E. coli* engineered to produce it (i.e., pMBIS + pABS). **FIG. 17b**, The mass spectrum of the
15 indicated peak from **FIG. 17a**.

FIGs. 18a-18b. GC/MS analysis of taxadiene production. FIG. 18a, A GC chromatogram shows the production of pure taxadiene (a kind gift from Phil Baran). **FIG. 18b**, The mass spectrum of the indicated peak from **FIG. 18a**.

20 **FIGs. 19a-19b. GC/MS analysis of β -bisabolene production. FIG. 19a**, A GC chromatogram shows the production of β -bisabolene by a strain of *E. coli* engineered to produce it (i.e., pMBIS + pGHS_{L450G}). **FIG. 19b**, The mass spectrum of the indicated peak from **FIG. 19a**.

25 **FIG. 20. Standard curve for pNPP assay.** This standard curve was generated by dissolving various concentrations of p-nitrophenol (p-NP) in 100 μ L water and measuring their absorbance with a plate reader. Absorbance measurements collected in our pNPP kinetics analysis were converted to concentrations using this curve.

30 **FIGs. 21a-21e. Development of a bacterial-two hybrid system that links the inhibition of PTP1B to antibiotic resistance.** This figure elaborates on FIG. 1 by including the orientation of genes. **FIG. 21a**, A bacterial two-hybrid (B2H) system in which a phosphorylation-

dependent protein-protein interaction modulates transcription of a gene of interest (GOI, black). Major components include (i) a substrate domain fused to the omega subunit of RNA polymerase (yellow), (ii) an SH2 domain fused to the 434 phage cI repressor (light blue), (iii) Src kinase and PTP1B, (iv) an operator for 434cI (dark green), (v) a binding site for RNA polymerase (purple), and (vi) a gene of interest (GOI, black). **FIG. 21b**, The luminescence generated by a B2H system with a p130cas substrate, LuxAB as the GOI, and no PTP1B. We used an inducible plasmid to increase expression of specific components. **FIG. 21c**, The luminescence generated by B2H systems with an SH2 domain that exhibits enhanced affinity for phosphopeptides (SH2*), one of four substrate domains, LuxAB as the GOI, and no Src or PTP1B. We used an inducible plasmid to control the expression of Src. Sequences for substrates p130cas (SEQ ID NO: 24), MidT (SEQ ID NO: 25), EGFR (SEQ ID NO: 27), and ShcA (SEQ ID NO: 26) are shown. **FIG. 21d**, The B2H system from **c** with either p130cas or MidT substrates. We used a second plasmid to control the expression of Src and an active or inactive (C215) variant of PTP1B. Right: Two optimized single-plasmid systems. **FIG. 21e**, The final B2H system. Inactivation of PTP1B enabled a strain of *E. coli* harboring this system to survive at high concentrations of spectinomycin (> 250 µg/ml). Error bars in **FIGS. 21b-d** denote standard error with n = 3 biological replicates.

FIGS. 22a-22g. Biosynthesis of PTP1B-inhibiting terpenoids enables cell survival. This figure elaborates on Figures 2 and 10. **FIG. 22a**, The plasmid-borne pathway for terpenoid biosynthesis: (i) pMBIS_{CmR}, which harbors the mevalonate-dependent isoprenoid pathway of *S. cerevisiae*, converts mevalonate to isopentyl pyrophosphate (IPP) and farnesyl pyrophosphate (FPP). (ii) pTS, which encodes a terpene synthase (TS) and, when necessary, a geranylgeranyl diphosphate synthase (GGPPS), converts IPP and FPP to sesquiterpenes or diterpenes. **FIG. 22b**, Five terpene synthases examined in this study: amorphadiene synthase (ADS), γ-humulene synthase (GHS), α-bisabolene synthase (ABA), abietadiene synthase (ABS), and taxadiene synthase (TXS). **FIG. 22c**, The spectinomycin resistance of strains of *E. coli* that harbor both (i) the bacterial two-hybrid (B2H) system (ii) a TS-specific terpenoid pathway. Note: ABS*, a positive control, has a constitutively active B2H (i.e., it includes PTP1B_{C215S}). **FIG. 22d**, Chromatograms show expected major products (i.e., namesake; *) for each TS-specific strain from **c** in the presence of the B2H system. Values are normalized to the largest peak within a given sample. **FIG. 22e**, Initial rates of PTP1B-catalyzed hydrolysis of pNPP in the presence of increasing concentrations of (AD) amorphadiene or

(AB) α -bisabolene. Lines show the best-fit kinetic models of inhibition (**TABLE 12**). **FIG. 22f**, Estimated IC_{50} 's. **FIG. 22g**, Titters of the major products generated by ADS and ABA. Error bars denote (**FIG. 22e**) standard error and (**FIG. 22f**) 95% confidence intervals for $n \geq 3$ independent measurements, and (**FIG. 22g**) standard deviation for $n = 3$ biological replicates.

FIGs. 23a-23h. Biophysical analysis of terpenoid-mediated inhibition. This figure builds on Figure 12 by including additional kinetic measurements. **FIG. 23a**, Aligned X-ray crystal structures of PTP1B bound to TCS401, a competitive inhibitor (yellow protein, orange highlights, and green spheres; pdb entry 5k9w), and BBR, an allosteric inhibitor (gray protein, blue highlights, and light blue spheres; pdb entry 1t4j). **FIG. 23b**, Aligned structures of PTP1B bound to BBR (white protein and light blue ligand) and amorphadiene (cyan protein and dark blue ligand, pdb entry 6W30). **FIG. 23c**, Dihydroartemisinic acid (DHA), a structural analogue of amorphadiene with a carboxyl group likely to disrupt binding to the hydrophobic cleft. **FIG. 23d**, DHA is eight-fold less potent than amorphadiene. Lines show the best-fit kinetic models of inhibition (**TABLE 12**). Error bars denote standard error for $n = 3$ independent measurements with a 95% confidence interval for the IC_{50} . **FIG. 23e**, Dixon plot showing V_o^{-1} vs. [TCS401] at various concentrations of AD (black, blue, purple markers). The parallel lines indicate that TCS401 and AD cannot bind simultaneously. **FIG. 23f**, Dixon plot showing V_o^{-1} vs. [orthovanadate] at various concentrations of AD (black, blue, purple markers). The intersecting lines indicate that orthovanadate and AD can bind simultaneously. **FIG. 23g**, Both amorphadiene and α -bisabolene inhibit PTP1B much more potently than TC-PTP; the removal of the $\alpha 7$ helix (or equivalent) from both enzymes reduces the selectivity of AD, but not AB. Error bars show propagated 95% confidence intervals estimated from $n \geq 3$ independent measurements at each condition. **FIG. 23h**, Amorphadiene (930 μ M) and α -bisabolene (405 μ M) stimulate IR phosphorylation in HEK293T/17 cells; at the same concentrations, dihydroartemisinic acid (DHA) and α -bisabolol (ABOL) exhibit reduced signals consistent with their reduced potencies (#: $p < 0.05$, compared to negative control, *: $p < 0.05$). All inhibitors are dissolved in 3% DMSO (v/v; negative control). Error bars in **FIGs. 23d-f** denote standard error for $n=3-12$ biological replicates. Error bars in **FIG. 23g** denote propagated 95% confidence intervals for $n \geq 3$

independent measurements. Error bars in **FIG. 23h** denote standard error propagated from a buffer-only control ($n = 3$ biological replicates).

FIGs. 24a-24e. Analysis of uncharacterized terpene synthase genes. FIG. 24a, A

- 5 bioinformatic analysis of terpene synthases. We assembled a cladogram of 4,464 members of the largest terpene synthase family (PF03936) and annotated it with functional data. We selected three genes from each of eight clades (curved boxes): six with no characterized genes (i.e., genes with known functions) and two with no characterized genes. **FIG. 24b**, The spectinomycin resistance conferred by the selected genes alongside pMBIS_{CmR} and pB2H_{opt}.
 10 Hits with robust growth beyond 400 $\mu\text{g/mL}$ spectinomycin appear in blue. “n.m.” indicates the condition was not measured. **FIG. 24c**, A0A0C9VSL7 produces (+)-1(10),4-cadinadiene as a dominant product ($m/z=204$). **FIG. 24d**, Structure of (+)-1(10),4-cadinadiene. **FIG. 24e**, The inhibition of PTP1B by (+)-1(10),4-cadinadiene (85% purity, 10% DMSO). Lines show the best-fit kinetic models of inhibition (**TABLE 12**).

15

FIGs. 25a-c. Extension to other disease-related PTPs. FIG. 25a, The spectinomycin resistance of strains harboring B2H systems modified to detect the inactivation of different disease-relevant PTPs. Inactivating mutations^{86–88} confer survival at high concentrations of antibiotic. **FIG. 25b**, A comparison of the resistance conferred by PTP1B- and TC-PTP-specific B2H systems in the presence of metabolic pathways for amorphadiene and α -bisabolene (i.e., pMBIS_{CmR} + ADS or ABA). The PTP1B-specific system exhibits a prominent survival advantage, a finding consistent with the selectivity of both terpenoids for this enzyme. **FIG. 25c**, The titers of AD and AB in strains harboring both the B2H systems and associated metabolic pathways are indistinguishable between strains.

25

- FIG. 26a-d. Analysis of the products of different terpene synthases.** This figure builds on Figure 9 by including additional measurements. **FIG. 26a**, Total terpene titers generated by each TS-specific strain in the absence (red) and presence (blue) of the B2H system. These results indicate that the B2H system does not disrupt terpenoid biosynthesis. **FIG. 26b**, GC/MS chromatograms of the terpenoids generated by the diterpene synthases in the absence (top) and presence (bottom) of the B2H system ($m/z=272$). **FIG. 26c**, GC/MS chromatograms of the terpenoids generated by the sesquiterpene synthases in the absence (top) and presence (bottom) of the B2H system ($m/z=204$). Similar profiles in **FIG. 26b** and

FIG. 26c indicate that the B2H system does not alter product distributions. **FIG. 26d**, Analysis of the contributions of either (i) TS activity or (ii) B2H function to the death and survival of GHS, ADS, and ABA strains. Inactivation of GHS does not enhance survival, an indication that this enzyme does not produce growth-inhibiting terpenoids. Inactivation of either ADS, ABA, or the B2H system, by contrast, weakens the antibiotic resistance of the ADS and ABA strains; maximal resistance thus requires both terpenoid production and B2H activation. Labels denote the following controls: D/A, an inactive terpene synthase (contains a D/A mutation at the catalytic aspartic acid, preventing the initial metal-binding step in terpene cyclization); *, a constitutively active B2H (contains PTP1B_{C215S}, preventing dephosphorylation); X, an inactive B2H (contains a substrate domain with a Y/F mutation, prohibiting phosphorylation and thus binding with the SH2 domain). Images show LB plates seeded with drops of liquid culture (10 μ L) from two biological replicates. **TABLE 2** details the B2H systems used for these analyses. Error bars in **FIG. 26a** denote standard deviation for $n \geq 3$ biological replicates.

15

FIG. 27. An annotated cladogram of terpene synthases. This cladogram of the PF03936 family is surrounded by a heatmap that shows the presence/absence of known EC numbers of the form 4.2.3.# (which includes terpene cyclization reactions) from the Uniprot database. We selected three genes from each of eight clades: six with no characterized genes (red) and two with characterized genes (blue). **TABLE 1** summarizes the genes.

20

FIG. 28. Analysis of selected genes. We searched for sesquiterpene inhibitors of PTP1B by screening each of the 24 uncharacterized genes alongside the FPP pathway (i.e., pMBIS). These pictures show the antibiotic resistance conferred by each gene. We selected strains with antibiotic resistance exceeding 400 μ g/ml as hits (blue). Importantly, for these genes, the reduced survival of B2Hx controls indicates that enhanced resistance requires activation of the B2H system. In the top diagrams, n.m. indicates conditions that were not measured.

25

FIG. 29. Product profiles of selected hits. The product profiles of selected hits (extracted ion chromatograms, $m/z = 204$). In brief, we grew up hits (i.e., pB2H_{opt}, pMBIS_{CmR}, and pTS) in liquid culture for 72 hours. With the exception of A0A0G2ZSL3, all hits were grown in 10 mL of 2% TB; A0A0G2ZSL3 was grown in a 4-mL culture of 2% TB. Notably, both A0A0C9VSL7 and A0A2H3DKU3 generate one dominant product: (+)-1(10),4-cadinadiene

30

and β -farnesene, respectively. We focused on A0A0C9VSL7 because (+)-1(10),4-cadinadiene is a structural analog of amorphadiene, an inhibitor identified in our initial screen.

5 **FIG. 30. Crystallographic analysis of PTP1B bound to AD.** Crystal structures of PTP1B collected in the (left) presence or (right) absence of AD. Resolutions: 2.10 Å (PTP1B-AD) and 1.94 Å (PTP1B). We refined these structures by modeling (top) the PTP1B-AD complex or (bottom) the apo form PTP1B. For PTP1B soaked with AD (left), the 1.0 σ 2Fo-Fc electron density supports the modeled position of AD but suggest multiple conformations;
10 this density appears even when AD is excluded from the model. For apo PTP1B (right), the 1.0 σ 2Fo-Fc electron does not support a bound AD molecule; small regions of unexplained density may reflect water molecules or partial occupancy of the α 7 helix¹⁵.

FIG. 31. Crystallographic analysis of PTP1B bound to ABol. Crystal structures of PTP1B
15 collected in the (left) presence or (right) absence of ABol. Resolutions: 2.11 Å (PTP1B-ABol) and 1.94 Å (PTP1B). We refined these structures by modeling (top) the PTP1B-ABol complex or (middle/bottom) the apo form PTP1B. For PTP1B soaked with ABol (left), the 0.90 σ 2Fo-Fc electron density is consistent with the modeled position of ABol, but it becomes less pronounced when ABol is excluded from the model. The apo form of PTP1B
20 (right) shows similar density for both models; small differences in the shape of the 0.90 σ 2Fo-Fc electron density between datasets suggests that this density may have a different origin (e.g., a ligand vs. partial occupancy of the α 7 helix). The unambiguous determination of a binding site for α -bisabolol requires additional data.

25 **FIGs. 32a-32c. Evidence of multiple bound conformations. FIG. 32a,** Snapshots from molecular dynamics (MD) simulations of PTP1B bound to amorphadiene (AD). Arrows indicate clusters of ligand. **FIG. 32b,** A crystal structure of PTP1B bound to AD highlights residues that undergo high-frequency contacts. Here, contacts have residue-ligand distances < 4 Å, and high frequencies exceed 10% of all snapshots in the MD simulations. **FIG. 32c,**
30 Estimates of the average root-mean-square deviation (RMSD) of the complete system (PL), the protein (P), the protein core (P_{core}; residues 1-287), the disordered region of the protein (P_{tail}; residues 288-321), and the ligand (L) over MD simulations indicate that both AD and the disordered region of the protein are mobile (the latter more so than the former), while the

protein core remains fixed. The average RMSDs of both (i) the re-centered ligand (Int), a metric for rotational and vibrational fluctuations, and (ii) the center of mass (COM) of the ligand, a metric for its positional deviation, are large, an indication that the ligand can adopt multiple bound conformations and/or positions.

5

FIGs. 33a-33m. Summary of kinetics analyses. **FIG. 33a**, Aligned crystal structures of PTP1B (gray, pdb entry 5k9w) and TC-PTP (blue, pdb entry 1l8k). Highlights on PTP1B: a competitive inhibitor (orange), the $\alpha 7$ helix (red), and truncation points used for kinetic studies (281 and 283, the 281-equivalent of TC-PTP). **FIG. 33b**, Sequence alignment of the $\alpha 6/7$ regions of PTP1B (SEQ ID NO: 140) and TC-PTP (SEQ ID NO: 141). The truncation points used in our kinetics analysis. **FIG. 33c**, aligned structures of the binding sites of BBR (gray, pdb entry 1t4j) and amorphadiene (blue). **FIG. 33d- FIG. 33m**, Initial rates of pNPP hydrolysis by various PTPs in the presence of increasing concentrations of (FIG. 33d- FIG. 33g) amorphadiene, (FIG. 33h- FIG. 33k) α -bisabolene, (FIG. 33l) dihydroartemesinic acid, and (FIG. 33m) α -bisabolol inhibition. In all figures, lines show the best-fit models of inhibition (TABLE 12). Error bars in FIG. 33d- FIG. 33m represent standard error of at least 3 measurements. Error in IC_{50} 's represent 95% confidence intervals determined from fits to models of inhibition (TABLE 12).

FIGs. 34a-34d. Expanded analysis of selectivity . **FIG. 34a**, Initial rate data for AD inhibition of SHP1. The lower panel shows the same data as % inhibition for a subset of points at two different substrate concentrations (open vs. closed circles). **FIG. 34b**, Initial rate data for AD inhibition of SHP2. The lower panel shows the same data as % inhibition for a subset of points at two different substrate concentrations (open vs. closed circles). **FIG. 34c**, Initial rate data for AB inhibition of SHP1. The lower panel shows the same data as % inhibition for a subset of points at two different substrate concentrations (open vs. closed circles). **FIG. 34d**, Initial rate data for AB inhibition of SHP2. The lower panel shows the same data as % inhibition for a subset of points at two different substrate concentrations (open vs. closed circles). In FIG. 34a, FIG. 34c, and FIG. 34d, our inability to measure inhibition >25% (lower panel) at the solubility limit of AD, in combination with the high K_m for 4-methylumbelliferyl phosphate (4-MUP), precluded accurate inhibition model fitting, K_I , and IC_{50} determination. However, the weak inhibition observed suggests AD/AB are less

potent inhibitors of these enzymes than PTP1B. In all panels, error bars denote standard error of n=3 biological replicates and lines show fit to a noncompetitive inhibition model.

FIG. 35a-35c. Analysis of PTP1B-mediated IR dephosphorylation. **FIG. 35a**, A depiction of insulin signaling in HEK293T/17 cells. Extracellular insulin binds to the transmembrane insulin receptor (IR), triggering phosphorylation of its intracellular domain. PTP1B, which localizes to the endoplasmic reticulum (ER) of mammalian cells, dephosphorylates this domain to regulate downstream signaling pathways. In starved cells, exogenously supplied inhibitors can permeate the cell membrane and inhibit PTP1B-mediated dephosphorylation of the IR. **FIG. 35b**, A screen of inhibitor concentrations for enzyme-linked immunosorbent assay (ELISAs). An enzyme-linked immunosorbent assay (ELISA) of IR phosphorylation in HEK293T/17 cells incubated with various concentrations of amorphadiene, α -bisabolene, and their structural analogues. We used this screen to identify biologically active concentrations of amorphadiene and α -bisabolene to study further. **FIG. 35c**, ELISA-based measurements of IR phosphorylation in HEK293T/17 cells incubated with amorphadiene (AD), α -bisabolene (AB), dihydroartimesnic acid (DHA), and α -bisabolol (ABOL). Curves denote fits to the four-parameter logistic equation: $y = d + (a - d) / (1 + (x/c)^b)$, where y is absorbance at 450 nm, and x is the sample dilution (e.g., 1 denotes no dilution, 0.5 denotes a 2-fold dilution, and so on). These signals indicate that amorphadiene and α -bisabolene can increase IR phosphorylation over a negative control (3% DMSO) and their less inhibitory analogs. Error bars denote standard error with $n \geq 3$ biological replicates.

FIGs. 36a-36c. Full datasets for B2H-mediated antibiotic resistance. **FIG. 36a**, Biological replicates for **FIG. 22c**. **FIG. 36b**, Biological replicates for **FIG. 25a**. **FIG. 36c**, Biological replicates for **FIG. 25b**. Orange highlights correspond to the data displayed in Figs. 2c and 5a-b.

FIGs. 37a-37b. GC/MS analysis of α -bisabolene production. **FIG. 37a**, A GC/MS chromatogram shows the production of α -bisabolene by a strain of *E. coli* engineered to produce it (i.e., pMBIS + pABA). **FIG. 37b**, The mass spectrum of the indicated peak from **FIG. 37a**.

FIGs. 38a-38b. Supplementary Fig. 20 | GC/MS analysis of (+)-1(10),4-Cadinadiene.

FIG. 38a, A GC/MS chromatogram shows the production of (+)-1(10),4-Cadinadiene by a strain of *E. coli* engineered to produce it (i.e., pMBIS + pA0A0C9VSL7). **FIG. 38b**, The mass spectrum of the indicated peak from **FIG. 38a**.

5

FIGs. 39a-39b. A standard curve for *p*-nitrophenol (*p*-NP). This figure elaborates on

Figure 20 by including additional measurements. **FIG. 39a**, We dissolved different amounts of *p*-nitrophenol (*p*-NP) in 100 μ L buffer (50 mM HEPES, pH=7.3) and measured the absorbance of the resulting solutions with a SpectraMax M2 plate reader. A linear fit to this

10 curve allowed us to convert absorbance measurements taken during kinetic assays (*p*NPP) to *p*-NP concentrations. **FIG. 39b**, We dissolved different amounts of 4-methyl umbelliferone (4-MU) in 100 μ L buffer (50 mM HEPES, pH=7.3) and measured the FLUORESCENCE of the resulting solutions with a SpectraMax M2 plate reader. A linear fit to this curve allowed us to convert absorbance measurements taken during kinetic assays (4-MUP) to 4-MU
15 concentrations.

DETAILED DESCRIPTION

E. coli is a valuable platform for the production of terpenoids²⁷⁻²⁹. The inventors hypothesized that a strain of *E. coli* programmed to detect the inactivation of a human drug
20 target might enable the rapid discovery and biosynthesis of terpenoids that inhibit that target. To program such a strain, a bacterial two-hybrid (B2H) system was assembled in which a protein tyrosine kinase (PTK) and protein tyrosine phosphatase (PTP) from *H. sapiens* control gene expression. PTKs are targets of over 30 FDA-approved drugs³⁰; PTPs lack clinically approved inhibitors but contribute to an enormous number of diseases^{31,32}. The first
25 proof-of-concept system was specifically designed to detect inhibitors of protein tyrosine phosphatase 1B (PTP1B), an elusive therapeutic target for the treatment of type 2 diabetes, obesity, and breast cancer (Fig. 1a)³¹⁻³⁵. In this system, Src kinase phosphorylates a substrate domain, enabling a protein-protein interaction that activates transcription of a gene of interest (GOI). PTP1B dephosphorylates the substrate domain, preventing that interaction, and the
30 inactivation of PTP1B re-enables it. *E. coli* is a particularly good host for this detection system because its proteome is sufficiently orthogonal to the proteome of *H. sapiens* to minimize off-target growth defects that can result from the regulatory activities of Src and PTP1B³⁶.

B2H development was carried out in several steps. To begin, a luminescent “base” system was assembled in which Src modulates the binding of a substrate domain to a substrate homology 2 (SH2) domain; this system was based on a previous design in which protein-protein association controls GOI expression³⁷. The initial system did not yield a phosphorylation-dependent transcriptional response, however, so it was complemented with inducible plasmids—each harboring a different system component—to identify proteins that might exhibit suboptimal activities. Notably, secondary induction of Src increased luminescence, an indication that insufficient substrate phosphorylation depressed GOI expression in the base system (Fig. 1b). Accordingly, this system was modified by swapping in different substrate domains, by adding mutations to the SH2 domain that enhance its affinity for phosphopeptides³⁸, and by removing the gene for Src. With this configuration, induction of Src from a second plasmid increased luminescence most prominently for the MidT substrate (Fig. 1c); simultaneous induction of both Src and PTP1B, in turn, prevented that increase (Fig. 1d). The MidT system was finalized by integrating genes for Src and PTP1B, by adjusting promoters and ribosome binding sites to amplify its transcriptional response further (Figs. 1d, 13, and 14), and by adding a gene for spectinomycin resistance (SpecR) as the GOI. The final plasmid-borne detection system required the inactivation of PTP1B to permit growth at high antibiotic concentrations (Fig. 1e).

The B2H system was used to identify new inhibitors of PTP1B by coupling it with metabolic pathways that might generate such molecules in *E. coli*. Previous screens of plant extracts have identified structurally complex terpenoids that inhibit PTP1B³⁹; pathways were, thus, constructed for several simpler terpenoid scaffolds that lack established inhibitory effects: amorphadiene, γ -humulene, abietadiene, and taxadiene. Abietadiene is a metabolic precursor to a weak inhibitor of PTP1B⁴⁰; the other three terpenoids represent a structurally diverse set of molecules. Each pathway consisted of two plasmid-borne modules (Fig. 2a): (i) the mevalonate-dependent isoprenoid pathway from *S. cerevisiae*⁴¹ and (ii) a terpene synthase supplemented—when necessary for diterpenoid production—with a geranylgeranyl diphosphate synthase. These modules enabled terpenoid titers of 0.5-100 μ M in *E. coli* (Fig. 9).

Each pathway was screened for its ability to produce inhibitors of PTP1B by transforming *E. coli* with plasmids harboring both the pathway of interest and the B2H system. GC-MS traces confirmed that all pathways generated terpenoids in the presence of the B2H system (Fig. 2d). Surprisingly, the amorphadiene pathway permitted survival at high

concentrations of antibiotic; importantly, maximal resistance required a functional B2H system (Fig. 9c). This result suggests that the amorphadiene pathway produces an inhibitor of PTP1B.

Microbially-assisted directed evolution (MADE) refers to the approach described
5 herein for using microbial systems to discover and evolve metabolic pathways that produce inhibitors or activators of a therapeutically relevant enzyme target, wherein both the metabolic pathway and the target enzyme exist within a host cell, for example, an *E. coli* cell (Fig. 3). Some aspects of this approach provide a method for building a genetically encoded system that detects the activity of a target enzyme within a host cell, for example a system
10 that links changes in the activity of a target enzyme to changes in the antibiotic resistance of the host cell (Fig. 1).

Previous work demonstrated (i) the assembly of a detection system that links the activities of a protein kinase and a protein phosphatase to antibiotic resistance (Fig. 1) and (ii) the use of that system, in combination with MADE, to discover inhibitors of a protein
15 phosphatase (Fig. 2). These results are detailed in PCT/US2019/40896.

Described herein are strategies, systems, methods, and reagents to expand the scope of capabilities of MADE and to address the needs of previously described evolution experiments. The MADE methods herein utilize one or more of the following: 1) target enzymes that post-translationally modify proteins (PTM enzymes) in a manner other than
20 adding or removing a phosphate group; 2) a metabolic pathway that generates phenylpropanoids or nonribosomal peptides; 3) a cryptic gene cluster that encodes putative natural products; and 4) natural products with specific inhibitory effects.

In some embodiments, provided are methods for using MADE to discover and evolve metabolic pathways that produce inhibitors or activators of PTM enzymes (Fig. 3), wherein
25 said PTM enzymes modulate a protein-protein interaction that controls a detectable output, wherein both the PTM enzymes and the detectable output are encoded by at least one plasmid or one genome, wherein a metabolic pathway that produces natural products is encoded by at least one plasmid or one genome, and wherein said plasmids and genomes exist within the same host cell. In some embodiments, a pool of said host cells, each of which contains a
30 different metabolic pathway, is screened for a detectable output, and the cells that yield the highest detectable output are selected as hits. These hits are analyzed with the following steps: 1) their metabolic pathways are reassembled from a starting pathway; 2) the re-assembled pathways are re-screened in host cells (a confirmation step); 3) the cells that yield

the highest detectable outputs are, once again, selected as hits; 4) these selected cells are grown in liquid culture; 5) the products generated in said liquid culture are identified and quantified with standard analytical methods, for example, gas chromatography–mass spectrometry (GC/MS); 6) the products generated in liquid culture are concentrated with a rotary evaporator; and 7) the modulatory effects of the concentrated products are tested on purified PTM enzymes (Fig. 3).

In some embodiments, the target PTM enzyme naturally inhibits the growth of a host cell, for example, an *S. cerevisiae* cell in which a heterologously expressed kinase slows cell growth .

10 In some embodiments, the PTM enzymes are ubiquitin ligases, SUMO transferases, methyltransferases, demethylases, acetyltransferases, glycosyltransferases, palmitoyltransferases, and/or related hydrolases. In some embodiments, a bacterial two-hybrid (B2H) system links the activity of one or more PTM enzymes to the transcription of a gene of interest (GOI; Fig. 4a). In some embodiments, the PTM enzymes modulate the
15 assembly of a split protein, for example, a fluorescent protein, a luciferase, or an enzyme that confers antibiotic resistance (Fig. 4b). In some embodiments, the target enzymes covalently link or proteolyze two proteins, wherein the assembly of these proteins activates the transcription of a gene of interest (Fig. 4c) or reassembles a split protein (Fig. 4d).

In some embodiments, provided are methods for the discovery and evolution of
20 phenylpropanoids or nonribosomal peptides that inhibit or activate a target enzyme, wherein a metabolic pathway that produces phenylpropanoids or nonribosomal peptides is encoded by at least one plasmid or one genome (Fig. 5), wherein said plasmid and said genome exist within a host cell, wherein mutagenesis and/or modulation of said metabolic pathways permit the production of an inhibitor or activator of the target enzyme, and wherein MADE enables
25 the identification of pathways thus mutated and/or reconfigured.

In some embodiments, provided are methods for the discovery and evolution of cryptic metabolic pathways that generate inhibitors or activators of a target enzyme, wherein said cryptic metabolic pathways comprise a set of genes with unknown or poorly characterized products, or wherein said cryptic metabolic pathways comprise a set of genes in
30 which one gene hinders the biosynthesis of an important product, wherein subsequent mutagenesis and/or reconfiguration of said pathway causes it to generate more of that product, and wherein MADE enables the discovery of a pathway thus mutated and/or reconfigured. For example, the removal of a biosynthetic gene may enable the accumulation

of a metabolic intermediate that modulates the activity of a target enzyme (Fig. 6a); alternatively, the removal of a gene for a transcriptional repressor may permit the activation of the entire metabolic pathway (Fig. 6b).

In some embodiments, provided are methods for the discovery and evolution of metabolic pathways with higher titers and/or lower toxicities, wherein starting pathways are mutated and/or reconfigured to create a library of pathways, and said library of pathways is screened using MADE to identify pathways that (i) produce higher quantities of inhibitor or activator than the starting pathway and/or (ii) exhibit a lower toxicity than the starting pathway (Fig. 7). For example, mutagenized and/or reconfigured pathways may contain genes for a mutant enzyme, for example, a terpene synthase, that exhibits a higher activity than the wild-type enzyme; alternatively, mutagenized and/or reconfigured pathways may contain genes for a mutant terpene synthase that is more soluble or otherwise less toxic than a wild-type enzyme.

Some aspects of this disclosure provide molecules that inhibit protein tyrosine phosphatases (PTPs), for example, protein tyrosine phosphatase 1B (PTP1B; Figs. 9 and 10). Examples include amorphadiene and derivatives, taxadiene and derivatives, β -bisabolene and derivatives, α -bisabolene and derivatives, and α -longipinene and derivatives. In some embodiments, these molecules are provided as drugs or drug leads for the treatment of diseases to which PTPs contribute, for example, type 2 diabetes⁴², HER2-positive breast cancer⁴³, or Rett syndrome⁴⁴, as are methods of treatment of such diseases by administering an effective amount of the molecule(s) to a subject in need of such treatment.

Also provided are compositions or systems that include 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, and 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.

In some embodiments, 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. In some embodiments, 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. In some embodiments, the complex is formed by two proteins with a dissociation
5 constant (K_d) less than or equal to the K_d of the complexes formed between SH2 domains and their phosphorylated substrates.

In some embodiments, the metabolic pathways encoded by the expression vectors produce phenylpropanoids or nonribosomal peptides. In some embodiments, 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. In some embodiments, one or more of the metabolic pathways comprises a set of genes of unknown biosynthetic capability.

In some embodiments, one or more of the metabolic pathways that produces a detectable output higher than the output of the reference pathway produces a product that
15 differs from the products of other metabolic pathways. In some embodiments, one or more of the 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. In some embodiments, one or more of the metabolic pathways that produces a detectable output higher than the output of the reference pathway
20 exhibits a lower cellular toxicity than other metabolic pathways.

In some embodiments, 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.

Also provided herein are kits that include a population of expression vectors as
25 described herein. In some embodiments, the kits also include the population of host cells that comprise a protein of interest as described herein.

The summary above is meant to illustrate, in a non-limiting manner, some of the embodiments, advantages, features, and uses of the technology described herein. Other embodiments, advantages, features, and uses of the technology disclosed herein will be
30 apparent from the Detailed Description, Drawings, Examples, and Claims.

Definitions

The term “metabolic pathway,” as used herein, refers to a collection of genes that enable the synthesis of metabolite.

5 The term “metabolite,” as used herein, refers to an organic molecule assembled within a living system.

The term “small molecule,” as used herein, refers to a molecule with a molecular weight less than 900 daltons.

The term “phenylpropanoids,” as used herein, refers to an organic compound synthesized from the amino acids phenylalanine and/or tyrosine.

10 The term “nonribosomal peptide,” as used herein, refers to peptides synthesized without messenger RNA. For example, peptides synthesized from nonribosomal peptide synthases.

The term “modulator,” as used herein, refers to a molecule, peptide, protein, polynucleotide, or entity that changes the activity of another molecule, peptide, protein, polynucleotide, or entity.

15 The term “inhibitor,” as used herein, refers to a small molecule that reduces the activity of an enzyme.

The term “activator,” as used herein, refers to a small molecule that increases the activity of an enzyme.

20 The term “natural product,” as used herein, refers to a chemical compound or substance produced by a living organism.

The term “detection system,” as used herein, refers to a system that links the activity of a target enzyme to a detectable output.

25 The term “bacterial two-hybrid (B2H) system,” as used herein, refers to a genetically encoded system that links a protein-protein interaction to a detectable output.

The term “detectable output,” as used herein, refers to an output that can be detected with standard analytical instrumentation. Examples include fluorescence, luminescence, antibiotic resistance, or microbial growth.

30 The term “split protein,” as used herein, refers to a protein that exists as two separate halves, which, upon reassembly, restore the function of the protein.

The term “substrate domain,” as used herein, refers to a protein that includes a peptide fragment or protein component acted upon by a protein of interest. For example, a substrate

domain may include the peptide fragment of a receptor protein targeted by a kinase or phosphatase of interest.

The term “vector,” as used herein, refers to a deoxyribonucleic acid (DNA) molecule used as a vehicle to artificially carry foreign genetic material into a cell.

5 The term “host cell,” as used herein, refers to a cell that can host the genetically encoded systems, on vectors or genomes, necessary for MADE. For example, as host cell may contain plasmids that encode both (i) a genetically encoded detection system that links the activity of a target enzyme to a detectable output and (ii) a metabolic pathway capable of synthesizing molecules that might or might not inhibit said target enzyme.

10

EXAMPLES

Example 1

In previous work, a strain of *E. coli* was generated with two genetically encoded
15 modules—a B2H system that links the inhibition of PTP1B to the expression of a gene for antibiotic resistance, and a metabolic pathway for the production of amorphadiene—
exhibited greater antibiotic resistance than similar strains with different metabolic pathways (Fig. 2). In recent work, this result was explored further. First, it was shown that maximal
resistance required both an active amorphadiene synthase (ADS) and a functional B2H
20 system (Fig. 9). Second, the inhibitory effect of amorphadiene, the dominant product of ADS, was confirmed by measuring its influence on PTP1B-catalyzed hydrolysis of p-nitrophenyl
phosphate (pNPP; Fig. 10c). Initial rates exhibited a saturation behavior characteristic of noncompetitive or uncompetitive inhibition; most importantly, the IC₅₀ for amorphadiene was
~53 μM, a concentration lower than the 72 μM generated in liquid culture. For comparison,
25 the IC₅₀ for taxadiene was 119 μM, a concentration far lower than its titer in liquid culture. Results of the *in vitro* studies thus indicate that amorphadiene confers antibiotic resistance by
inhibiting PTP1B. Finally, an enzyme-linked immunosorbent assay (ELISA) was used to demonstrate the ability of amorphadiene to inhibit PTP1B inside of a HEK293T/17 cell (Fig.
10d-10e).

30 The microbial system provides an interesting opportunity to explore how metabolic pathways evolve to generate functional molecules. To look for evolutionarily accessible changes in the activities ADS and GHS that improve their ability to generate inhibitors of PTP1B, mutants of both enzymes were prepared. For ADS, error-prone PCR and site-

saturation mutagenesis of poorly conserved residues was used; for GHS, site-saturation mutagenesis of the wild-type enzyme was paired with a screen of several previously developed mutants with distinct product profiles⁴⁷ (Figs. 7a, 7b). At least one mutant from each library consistently conferred survival at higher antibiotic concentrations than the wild-type enzyme (Fig. 7c, 7d).

The G34S/K51N mutant of ADS, which improved antibiotic resistance more than other mutants, is particularly intriguing because its mutated residues are located outside of the active site and alter neither product profile nor titer (Fig. 7e, f). It was hypothesized that these mutations might reduce a minor growth deficiency caused by heterologous ADS expression (e.g., they might reduce the formation of inclusion bodies). To test this hypothesis, the survival conferred by wild-type and mutant strains in the presence of an inactive B2H system was compared; the mutant strain showed more robust growth at high concentrations of antibiotic (Fig. 7g). These results suggest that the engineered strain can select for less toxic enzyme mutants which, in the presence of other stresses, might improve production of inhibitory metabolites.

Intriguingly, the mutants of GHS that conferred enhanced antibiotic resistance (relative to the wild-type enzyme) altered product profile and/or titer (Figs. 7h and 7i). Two examples include GHS_{A336C/T445C/S484C/I562L/M565L} (or ALP), which primarily generates α -longipinene, and GHS_{A319Q}, which enhances terpenoid titer by ~ tenfold. The GHS mutants thus indicate that the engineered strain can select for enzyme mutants that generate different products and/or higher titers than a starting wild-type enzyme.

To expand the study, the survival conferred by terpene synthases that primarily generate β -bisabolene and α -bisabolene was also examined. Both of these enzymes enhanced antibiotic resistance; strikingly, kinetic studies of α -bisabolene purified from culture supernatant indicate that this molecule is particularly potent (i.e., IC_{50} ~20 μ M in 10% DMSO; Fig. 11).

The results of the analyses of terpene synthases suggest that amorphadiene and derivatives, taxadiene and derivatives, α -longipinene and derivatives, β -bisabolene and derivatives, and α -bisabolene and derivatives, and may provide an important source of pharmaceutically relevant PTP inhibitors.

Methods

Bacterial strains. *E. coli* DH10B, chemically competent NEB Turbo, or electrocompetent One Shot Top10 (Invitrogen) were used to carry out molecular cloning and to perform preliminary analyses of terpenoid production; *E. coli* BL2-DE31 were used to express proteins for *in vitro* studies; and *E. coli* s1030⁴⁸ were used for luminescence studies and for all experiments involving terpenoid-mediated growth (i.e., evolution studies).

For all strains, chemically competent cells were generated by carrying out the following steps: (i) each strain was plated on LB agar plates with the required antibiotics. (ii) One colony of each strain was used to inoculate 1 mL of LB media (25 g/L LB with appropriate antibiotics listed in **TABLE 2**) in a glass culture tube, and this culture was grown overnight (37°C, 225 RPM). (iii) The 1-mL culture was used to inoculate 100-300 mL of LB media (as above) in a glass shake flask, and this culture was grown for several hours (37°C, 225 RPM). (iv) When the culture reached an OD of 0.3-0.6, the cells were centrifuged (4,000 x g for 10 minutes at 4°C), the supernatant was removed, and the cells were resuspended in 30 mL of ice cold TFB1 buffer (30 mM potassium acetate, 10 mM CaCl₂, 50 mM MnCl₂, 100 mM RbCl, 15% v/v glycerol, water to 200 mL, pH=5.8, sterile filtered), and the suspension was incubated at 4°C for 90 min. (v) Step iv was repeated, but resuspended in 4 mL of ice cold TFB2 buffer (10 mM MOPS, 75 mM CaCl₂, 10 mM RbCl₂, 15% glycerol, water to 50 mL, pH=6.5, sterile filtered). (iv) The final suspension as split into 100 µL aliquots and frozen at -80°C until further use.

Electrocompetent cells were generated by following an approach similar to the one above. In step iv, however, the cells were resuspended in 50 mL of ice cold MilliQ water and repeated this step twice—first with 50 mL of 20% sterile glycerol (ice cold) and, then, with 1 mL of 20% sterile glycerol (ice cold). The pellets were frozen as before.

Materials. Methyl abietate was purchased from Santa Cruz Biotechnology; trans-caryophyllene, farnesol, tris(2-carboxyethyl)phosphine (TCEP), bovine serum albumin (BSA), M9 minimal salts, phenylmethylsulfonyl fluoride (PMSF), and DMSO (dimethyl sulfoxide) were purchased from Millipore Sigma; glycerol, bacterial protein extraction reagent II (B-PERII), and lysozyme from were purchased VWR; cloning reagents were purchased from New England Biolabs; amorphadiene was purchased from Ambeed, Inc.; and all other reagents (e.g., antibiotics and media components) were purchased from Thermo Fisher. Taxadiene was a kind gift from Phil Baran of the The Scripps Research Institute.

Mevalonate was prepared by mixing 1 volume of 2 M DL-mevalanolactone with 1.05 volumes of 2 M KOH and incubating this mixture at 37°C for 30 minutes.

Cloning and molecular biology. All plasmids were constructed by using standard methods (i.e., restriction digest and ligation, Golden Gate and Gibson assembly, Quikchange mutagenesis, and circular polymerase extension cloning). **TABLE 1** describes the source of each gene; **TABLES 2** and **3** describe the composition of all final plasmids.

Construction of the B2H system was begun by integrating the gene for HA4-rpoZ from pAB094a into pAB078d and by replacing the ampicillin resistance marker of pAB078d with a kanamycin resistance marker (Gibson Assembly). The resulting “combined” plasmid was modified, in turn, by replacing the HA4 and SH2 domains with kinase substrate and substrate recognition (i.e., SH2) domains, respectively (Gibson assembly), and by integrating genes for Src kinase, CDC37, and PTP1B in various combinations (Gibson assembly). The functional B2H system was finalized by modifying the SH2 domain with several mutations known to enhance its affinity for phosphopeptides (K15L, T8V, and C10A, numbered as in Kaneko et. al.⁴⁰), by exchanging the GOI for luminescence (LuxAB) with one for spectinomycin resistance (SpecR), and by toggling promoters and ribosome binding sites to enhance the transcriptional response (Gibson assembly and Quickchange Mutagenesis, Agilent Inc.). Note: For the last step, Pro1 to ProD was also converted by using the Quickchange protocol. When necessary, plasmids with arabinose-inducible components were constructed by cloning a single component from the B2H system into pBAD (Golden Gate assembly). **TABLES 4** and **5** list the primers and DNA fragments used to construct each plasmid.

Pathways for terpenoid biosynthesis were assembled by purchasing plasmids encoding the first module (pMBIS) and sesquiterpene synthases (ADS or GHS in pTRC99a) from Addgene, and by building the remaining plasmids. Genes for ABS, TXS, and GGPPS were integrated into pTRC99t (i.e., pTRC99a without BsaI sites), and a version of pADS was modified by adding a gene for P450_{BM3} with three mutations that enable the epoxidation of amorphadiene (F87A, R47L, and Y51F; P450_{G3}; Gibson Assembly and Quickchange Mutagenesis)⁴⁹. **TABLE 6** lists the primers and DNA fragments used to construct each plasmid.

Luminescence assays. Preliminary B2H systems (which contained LuxAB as the GOI) were characterized with luminescence assays. In brief, necessary plasmids were transformed into *E. coli* s1030 (**TABLE 2**), the transformed cells were plated onto LB agar plates (20 g/L agar, 10 g/L tryptone, 10 g/L sodium chloride, and 5 g/L yeast extract with antibiotics described in **TABLE 2**), and all plates were incubated overnight at 37°C. Individual colonies were used to inoculate 1 ml of terrific both (TB at 2%, or 12 g/L tryptone, 24 g/L yeast extract, 12 mL/L 100% glycerol, 2.28 g/L KH₂PO₄, 12.53 g/L K₂HPO₄, pH = 7.0, and antibiotics described in **TABLE 2**), and we incubated these cultures overnight (37°C and 225 RPM). The following morning, each culture was diluted by 100-fold into 1 ml of TB media (above), and these cultures were incubated in individual wells of a deep 96-well plate for 5.5 hours (37°C, 225 RPM). (Note: When pBAD was present, the TB media was supplemented with 0-0.02 w/v % arabinose). An amount of 100µL of each culture was transferred into a single well of a standard 96-well plate and measured both OD₆₀₀ and luminescence (gain: 135, integration time: 1 second, read height: 1 mm) on a Biotek Synergy plate reader. Analogous measurements of cell-free media were performed to measure background signals, which were subtracted from each measurement prior to calculating OD-normalized luminescence (i.e., Lum / OD₆₀₀).

Analysis of antibiotic resistance. The spectinomycin resistance conferred by various B2H systems in the absence of terpenoid pathways was evaluated by carrying out the following steps: (i) *E. coli* were transformed with the necessary plasmids (**TABLE 2**) and the transformed cells were plated onto LB agar plates (20 g/L agar, 10 g/L tryptone, 10 g/L sodium chloride, 5 g/L yeast extract, 50 µg/ml kanamycin, 10 µg/ml tetracycline). (ii) Individual colonies were used to inoculate 1-2 ml of TB media (12 g/L tryptone, 24 g/L yeast extract, 12 mL/L 100% glycerol, 2.28 g/L KH₂PO₄, 12.53 g/L K₂HPO₄, 50 µg/ml kanamycin, 10 µg/ml tetracycline, pH = 7.0), and these cultures were incubated overnight (37°C, 225 RPM). In the morning, each culture was diluted by 100-fold into 4 ml of TB media (as above) with 0-500 µg/ml spectinomycin (spectinomycin was used only for the results depicted in **FIG. 14**), and these cultures were incubated in deep 24-well plates until wells containing 0 µg/ml spectinomycin reached an OD₆₀₀ of 0.9-1.1. (iv) Each 4-ml culture was diluted by 10-fold into TB media with no antibiotics and plated 10-µL drops of the diluent onto agar plates with various concentrations of spectinomycin. (v) Plates were incubated overnight (37°C) and photographed the following day.

To examine terpenoid-mediated resistance, steps i and ii were performed as described above with the addition of 34 µg/ml chloramphenicol and 50 µg/ml carbenicillin in all liquid/solid media. The experiment then proceeded with the following steps: (iii) Samples were diluted from 1-ml cultures to an OD₆₀₀ of 0.05 in 4.5 ml of TB media (supplemented with 12 g/L tryptone, 24 g/L yeast extract, 12 mL/L 100% glycerol, 2.28 g/L KH₂PO₄, 12.53 g/L K₂HPO₄, 50 µg/ml kanamycin, 10 µg/ml tetracycline, 34 µg/ml chloramphenicol, and 50 µg/ml carbenicillin), which were incubated in deep 24-well plates (37°C, 225 RPM). (iv) At an OD₆₀₀ of 0.3-0.6, 4 ml of each culture was transferred to a new well of a deep 24-well plate, 500 µM isopropyl β-D-1-thiogalactopyranoside (IPTG) and 20 mM of mevalonate was added, and incubated for 20 hours (22°C, 225 RPM). (v) Each 4-ml culture was diluted to an OD₆₀₀ of 0.1 with TB media and plated 10 µL of the diluent onto either LB or TB plates supplemented with 500 µM IPTG, 20 mM mevalonate, 50 µg/ml kanamycin, 10 µg/ml tetracycline, 34 µg/ml chloramphenicol, 50 µg/ml carbenicillin, and 0-1200 µg/ml spectinomycin (for both plates, 20 g/L agar was used with media and buffer components described above). Note: to control the range of antibiotic resistance, LB plates were used for ADS and its mutants, and TB plates, which improve terpenoid titers, were used for GHS and its mutants. (iv) All plates were incubated at 30°C and photographed after 2 days.

Terpenoid biosynthesis. *E. coli* were prepared for terpenoid production by transforming cells with plasmids harboring requisite pathway components (**TABLE 2**) and plating them onto LB agar plates (20 g/L agar, 10 g/L tryptone, 10 g/L sodium chloride, and 5 g/L yeast extract with antibiotics described in **TABLE 2**). One colony from each strain was used to inoculate 2 ml TB (12 g/L tryptone, 24 g/L yeast extract, 12 mL/L 100% glycerol, 2.28 g/L KH₂PO₄, 12.53 g/L K₂HPO₄, pH = 7.0, and antibiotics described in **TABLE 2**) in a glass culture tube for ~16 hours (37°C and 225 RPM). These cultures were diluted by 75-fold into 10 ml of TB media and the new cultures were incubated in 125 mL glass shake flasks (37°C and 225 RPM). At an OD₆₀₀ of 0.3-0.6, 500 µM IPTG and 20 mM mevalonate were added. After 72-88 hours of growth (22°C and 225 RPM), terpenoids were extracted from each culture.

To measure terpenoid production over time, the approach described above was used with the following modifications: (i) Overnight cultures were diluted with 1:75 mL in 4.5 mL TB supplemented with antibiotics in a glass culture tube. (ii) When cultures reached an OD₆₀₀ of 0.3-0.6, 4 mL of each culture were moved to a new culture tube and 500 µM IPTG, 20 mM

mevalonate, 0-800 µg/mL spectinomycin, and 1 mL dodecane were added (to extract terpenoids). Every 4 hours, 100 µL of the dodecane sample was removed for GC/MS analysis.

- 5 **Protein expression and purification.** PTPs were expressed and purified as described previously⁴². Briefly, *E. coli* BL21(DE3) cells were transformed with pET21b vectors, and induced with 500 µM IPTG at 22 °C for 20 hours. PTPs were purified from cell lysate by using desalting, nickel affinity, and anion exchange chromatography (HiPrep 26/10, HisTrap HP, and HiPrep Q HP, respectively; GE Healthcare). The final protein (30–50 µM) was
10 stored in HEPES buffer (50 mM, pH 7.5, 0.5 mM TCEP) in 20% glycerol at –80°C.

- Extraction and purification of terpenoids.** Hexane was used to extract terpenoids generated in liquid culture. For 10-mL cultures, 14 mL of hexane was added to 10 ml of culture broth in 125-mL glass shake flasks, the mixture (100 RPM) shaken for 30 minutes, centrifuged (4000
15 x g), and 10 mL of the hexane layer was withdrawn for further analysis. For 4-mL cultures, 600 µL hexane were added to 1 mL of culture broth in a microcentrifuge tube, the tubes were vortexed for 3 minutes, the tubes were centrifuged for 1 minute (17000 x g), and 300-400 µL of the hexane layer was saved for further analysis.

- To purify amorphadiene, 500-1000 mL culture broth was supplemented with hexane
20 (16.7% v/v), the mixture was shaken for 30 minutes (100 RPM), the hexane layer was isolated with a separatory funnel, the isolated organic phase was centrifuged (4000 x g), and the hexane layer withdrawn. To concentrate the terpenoid products, excess hexane was evaporated in a rotary evaporator to bring the final volume to 500 µL, and the resulting mixture was passed over a silica gel one or two times (Sigma-Aldrich; high purity grade, 60
25 Å pore size, 230-400 mesh particle size)). Elution fractions (100% hexane) were analyzed on the GC/MS and pooled fractions with the compound of interest (amorphadiene). Once purified, pooled fractions were dried under a gentle stream of air, the terpenoid solids were resuspended in DMSO, and the final samples were quantified as outlined below.

- 30 **GC-MS analysis of terpenoids.** Terpenoids generated in liquid culture were measured with a gas chromatograph / mass spectrometer (GC-MS; a Trace 1310 GC fitted with a TG5-SilMS column and an ISQ 7000 MS; Thermo Fisher Scientific). All samples were prepared in hexane (directly or through a 1:100 dilution of DMSO) with 20 µg/ml of caryophyllene or

5 methyl abietate as an internal standard. When the peak area of an internal standard exceeded $\pm 30\%$ of the average area in hexane samples containing only standard, the corresponding samples were re-analyzed. For all runs, the following GC method was used: hold at 80°C (3 min), increase to 250°C ($15^{\circ}\text{C}/\text{min}$), hold at 250°C (6 min), increase to 280°C ($30^{\circ}\text{C}/\text{min}$), and hold at 280°C (3 min). To identify various analytes, m/z ratios were scanned from 50 to 550.

Sesquiterpenes generated by variants of ADS were examined by using select ion mode (SIM) to scan for the molecular ion ($m/z = 204$). For quantification, we used Eq. 1:

$$C_i = C_{std} * \frac{A_i}{A_{std}} * R \quad (\text{Eq. 1})$$

$$10 \quad R = \frac{A_{std,o}/C_{std,o}}{A_{ref,o}/C_{ref,o}} \quad (\text{Eq. 2})$$

where A_i is the area of the peak produced by analyte i , A_{std} is the area of the peak produced by C_{std} of caryophyllene in the sample, and R is the ratio of response factors for caryophyllene and amorphadiene in a reference sample.

15 Sesquiterpenes generated by variants of GHS were quantified by using the aforementioned procedure with several modifications: Methyl abietate was used as an internal standard (several mutants of GHS generate caryophyllene as a product); both $m/z=204$ and $m/z=121$, a common ion between sesquiterpenes and methyl abietate were scanned for; a ratio of response factors for amorphadiene and methyl abietate at $m/z = 121$ for R was used; and peak areas were calculated at $m/z = 121$. For all analyses, the analysis was
20 focused on peaks with areas that exceeded 1% of the total area of all peaks at $m/z=204$.

Diterpenoids were quantified by, once again, accompanying the general procedure with several modifications: A different molecular ion ($m/z = 272$) and an ion common to both diterpenoids and caryophyllene ($m/z=93$) was scanned for; a ratio of response factors for pure taxadiene (a kind gift from Phil Baran) and caryophyllene at $m/z = 93$ was used; and peak
25 areas $m/z = 93$ were calculated. For all analyses, only peaks with areas that exceeded 1% of the total area of all peaks at $m/z=272$ were examined.

Molecules were identified by using the NIST MS library and, when necessary, this identification was confirmed with analytical standards or mass spectra reported in the literature. Note: The assumption of a constant response factor for different terpenoids (e.g.,
30 all sesquiterpenes and diterpenes ionize like amorphadiene and taxadiene, respectively) can certainly yield error in estimates of their concentrations; the analyses described herein, which are consistent with those of other studies of terpenoid production in microbial systems^{50,51},

thus supply rough estimates of concentrations for all compounds except amorphadiene and taxadiene (which had analytical standards).

Homology modeling of ADS and GHS. Homology models of ADS and GHS were

- 5 constructed by using SWISS-MODEL with structures for α -bisabolol synthase (pdb entry 4gax) and α -bisabolene synthase (pdb entry 3sae) as templates, respectively⁵². This software package uses ProMod3 to build models from a target-template alignment, which preserves the structures of conserved regions and remodels insertions and deletions with a fragment library^{53,54}.

10

- Preparation of mutant libraries.** Libraries of enzyme mutants were prepared by using site-saturation mutagenesis (SSM) and error-prone PCR (ePCR). For SSM, the following steps were performed: (i) Genes were amplified with NNK primers that targeted select sites. (ii) The amplified genes were digested with DpnI, purified with gel electrophoresis, and either
- 15 Gibson Assembly or circular polymerase extension cloning (CPEC)⁵⁵ was used to integrate them into plasmids (*pTS_{xx}*). (iii) Heat shock was used to transform the fully assembled plasmids into chemically competent NEB Turbo cells. (iv) Library size was determined by plating dilutions of the transformation reactions on several LB agar plates (20 g/L agar, 10 g/L tryptone, 10 g/L sodium chloride, 5 g/L yeast extract, 50 μ g/ml carbenicillin), and all
- 20 remaining cells were plated over 9-10 plates for subsequent analysis. (v) Colonies were sequenced to verify that at least 5 of 6 transformants contained mutated genes. (vi) Plates were scraped into LB media (25 g/L LB broth mix, no antibiotics) and the final transformants were miniprepmed to recover the DNA Library. (vii) All final libraries were frozen in MilliQ water at -20°C.

- 25 For ePCR, the Genemorph II kit (Agilent) was used with ~0.5-2.5 mutations/kb. The final plasmids were dialyzed and electroporated into One Shot electrocompetent Top 10 cells, and the final plasmids were sequenced, extracted, and stored as described above.

- Analysis of mutant libraries.** Each mutant library was screened by carrying out the
- 30 following steps: (i) 100 ng of each site-specific SSM library for a given terpene synthase was pooled. (ii) Each complete library (i.e., ePCR or pooled SSM) was dialyzed for 2 hours. (iii) Up to 10 μ L (< 1 μ g) of each library was electroporated into a strain of *E. coli* harboring both the pMBIS pathway and the B2H system. (iv) 1 mL of SOC was added to the transformed

cells and incubated for 1 hour (37°C and 225 RPM). (v) 100 µL of the SOC outgrowth was serial diluted and plated onto LB agar plates (20 g/L agar, 10 g/L tryptone, 10 g/L sodium chloride, 5 g/L yeast extract, 50 µg/ml carbenicillin, 10 µg/ml tetracycline, 50 µg/ml kanamycin, and 34 µg/ml chloramphenicol) and the plates were incubated overnight (37°C).

5 This step allowed for quantification of the number of transformants screened (i.e., a number determined by counting colonies). (vi) The remaining 900 µL of transformed cells was added to 100 mL of TB (12 g/L tryptone, 24 g/L yeast extract, 12 mL/L 100% glycerol, 2.28 g/L KH₂PO₄, 12.53 g/L K₂HPO₄, 50 µg/ml carbenicillin, 10 µg/ml tetracycline, 34 µg/ml chloramphenicol, 50 µg/ml kanamycin, pH = 7.0) in 500-mL Erlenmeyer flasks, and these

10 flasks were incubated overnight (37°C and 225 RPM). (vii) In the morning, an aliquot of each culture was diluted to an OD₆₀₀ of 0.05 in 4 mL of TB and incubated in glass culture tubes (37°C and 225 RPM). (viii) At an OD₆₀₀ of 0.3-0.6, terpenoid production was induced by adding 5-20 mM mevalonate and 500 µM IPTG, and the resulting cultures were incubated for 20 hours (22°C and 225 RPM). (ix) Each culture was diluted to an OD₆₀₀ of 0.001 and 100 µL

15 of diluent was plated onto agar plates containing 500 µM IPTG, 5-20 mM mevalonate, 50 µg/ml kanamycin, 10 µg/ml tetracycline, 34 µg/ml chloramphenicol, 50 µg/ml carbenicillin, and 0-1000 µg/ml spectinomycin. (x) Colonies that survived high concentrations of spectinomycin were used to inoculate 4 mL of LB media (25 g/L LB broth mix, 50 µg/ml carbenicillin, 10 µg/ml tetracycline, 34 µg/ml chloramphenicol, 50 µg/ml kanamycin, which

20 was incubated overnight (37°C, 225 RPM). (xi) Plasmid DNA was extracted from the overnight culture for Sanger sequencing.

The influence of interesting mutations—and a check for false positive— were confirmed by rescreening them in freshly prepared mutants. Site directed mutagenesis was used to introduce mutations found in the hits and then their antibiotic resistance was analyzed

25 using the drop-based plating method described above.

Enzyme kinetics. To examine terpenoid-mediated inhibition, PTP1B-catalyzed hydrolysis of p-nitrophenyl phosphate (pNPP) was measured in the presence of various concentrations of terpenoids. Each reaction included PTP1B (0.05 µM), pNPP (0.33, 0.67, 2, 5, 10, and 15

30 mM), inhibitors (110 µM, 50 µM, and 15 µM for amorpha-4,11-diene; 100 µM, 50 µM, and 16.7 µM for taxadiene), and buffer (50 mM HEPES pH=7.5, 0.5 mM TCEP, 50 µg/ml BSA, 10% DMSO). The formation of p-nitrophenol was monitored by measuring absorbance at 405 nm every 10 seconds for 5 minutes on a Spectramax M2 plate reader.

Kinetic models were evaluated in three steps: (i) Initial-rate measurements collected in the absence and presence of inhibitors were fitted to Michaelis-Menten and inhibition models, respectively (here, the *nlinfit* and *fminsearch* functions from MATLAB were used). (ii) An F-test was used to compare the mixed model to the single-parameter model with the least sum squared error (here, the *fcd* function from MATLAB was used to assign p-values), and the mixed model was accepted when $p < 0.05$. (iii) The Akaike's Information Criterion (AIC) was used to compare the best-fit single parameter model to each alternative single parameter model, and the “best-fit” model was accepted when the difference in AIC (Δ_i) exceed 10 for all comparisons.⁵⁶ Note: For amorphadiene, this criterion was not met; both noncompetitive and uncompetitive models, however, yielded indistinguishable IC_{50} 's.

The half maximal inhibitory concentration (IC_{50}) of inhibitors were estimated by using the best-fit kinetic models to determine the concentration of inhibitor required to reduce initial rates of PTP-catalyzed hydrolysis of 15 mM of pNPP by 50%. The MATLAB function “nlparci” was used to determine the confidence intervals of kinetic parameters, and those intervals were propagated to estimate corresponding confidence on IC_{50} 's.

References for Example 1

1. Newman, D. J. & Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014. *Journal of Natural Products* **79**, 629–661 (2016).
2. Koehn, F. E. & Carter, G. T. The evolving role of natural products in drug discovery. *Nature Reviews Drug Discovery* **4**, 206–220 (2005).
3. Harvey, A. L., Edrada-Ebel, R. & Quinn, R. J. The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug Discov.* **14**, 111–129 (2015).
4. Rodrigues, T., Reker, D., Schneider, P. & Schneider, G. Counting on natural products for drug design. *Nature Chemistry* **8**, 531–541 (2016).
5. Pathan, H. & Williams, J. Basic opioid pharmacology: an update. *Br. J. Pain* **6**, 11–16 (2012).
6. Vidal, V. *et al.* Library-Based Discovery and Characterization of Daphnane Diterpenes as Potent and Selective HIV Inhibitors in *Daphne gnidium*. (2011).
doi:10.1021/np200855d
7. Weaver, B. A. How Taxol/paclitaxel kills cancer cells. **25**, (2014).
8. Camuesco, D. *et al.* The intestinal anti-inflammatory effect of quercitrin is associated with an inhibition in iNOS expression. *Br. J. Pharmacol.* **143**, 908–918 (2004).

9. Ling, T., Lang, W. H., Maier, J., Quintana Centurion, M. & Rivas, F. Cytostatic and Cytotoxic Natural Products against Cancer Cell Models. *Molecules* **24**, 2012 (2019).
10. Jantan, I., Ahmad, W. & Bukhari, S. N. A. Plant-derived immunomodulators: An insight on their preclinical evaluation and clinical trials. *Frontiers in Plant Science* **6**, 5 (2015).
11. Galanie, S., Thodey, K., Trenchard, I. J., Filsinger Interrante, M. & Smolke, C. D. Complete biosynthesis of opioids in yeast. *Science*. **349**, 1095–1100 (2015).
12. Luo, X. *et al.* Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature* (2019). doi:10.1038/s41586-019-0978-9
- 10 13. Zhang, R. K. *et al.* Enzymatic assembly of carbon–carbon bonds via iron-catalysed sp³ C–H functionalization. *Nature* (2019). doi:10.1038/s41586-018-0808-5
14. Davis, A. M., Plowright, A. T. & Valeur, E. Directing evolution: The next revolution in drug discovery? *Nature Reviews Drug Discovery* **16**, 681–698 (2017).
15. Maier, M. E. Design and synthesis of analogues of natural products. *Organic and Biomolecular Chemistry* **13**, 5302–5343 (2015).
16. Chen, M. S. & White, M. C. A predictably selective aliphatic C–H oxidation reaction for complex molecule synthesis. *Science*. **318**, 783–787 (2007).
17. Cho, I., Jia, Z. J. & Arnold, F. H. Site-selective enzymatic C–H amidation for synthesis of diverse lactams. *Science*. **364**, 575–578 (2019).
- 20 18. Harvey, A. L. Natural products in drug discovery. *Drug Discovery Today* **13**, 894–901 (2008).
19. Henrich, C. J. & Beutler, J. A. Matching the power of high throughput screening to the chemical diversity of natural products. *Nat. Prod. Rep.* **30**, 1284–1298 (2013).
20. Medema, M. H. *et al.* AntiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* **39**, (2011).
- 25 21. Jensen, P. R. Natural Products and the Gene Cluster Revolution. *Trends Microbiol.* **24**, 968–977 (2016).
22. Yan, Y. *et al.* Resistance-gene-directed discovery of a natural-product herbicide with a new mode of action. (2018). doi:10.1038/s41586-018-0319-4
- 30 23. Zhabinskii, V. N., Khripach, N. B. & Khripach, V. A. Steroid plant hormones: Effects outside plant kingdom. *Steroids* **97**, 87–97 (2015).
24. Li, Y. *et al.* Complete biosynthesis of noscapine and halogenated alkaloids in yeast.

- Proc. Natl. Acad. Sci. U. S. A.* **115**, E3922–E3931 (2018).
25. Zhang, H., Wang, Y., Wu, J., Skalina, K. & Pfeifer, B. A. Complete biosynthesis of erythromycin A and designed analogs using *E. coli* as a heterologous host. *Chem. Biol.* **17**, 1232–1240 (2010).
 - 5 26. Antosch, J., Schaefer, F. & Gulder, T. A. M. Heterologous Reconstitution of Ikarugamycin Biosynthesis in *E. coli*. *Angew. Chemie Int. Ed.* **53**, 3011–3014 (2014).
 27. Choi, O. *et al.* Biosynthesis of plant-specific phenylpropanoids by construction of an artificial biosynthetic pathway in *Escherichia coli*. *J. Ind. Microbiol. Biotechnol.* (2011). doi:10.1007/s10295-011-0954-3
 - 10 28. Pfeifer, B. A., Wang, C. C. C., Walsh, C. T. & Khosla, C. Biosynthesis of Yersiniabactin, a Complex Polyketide-Nonribosomal Peptide, Using *Escherichia coli* as a Heterologous Host. *Appl. Environ. Microbiol.* (2003). doi:10.1128/AEM.69.11.6698-6702.2003
 29. Ajikumar, P. K. *et al.* Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science* **330**, 70–74 (2010).
 - 15 30. Chang, M. C. Y., Eachus, R. A., Trieu, W., Ro, D.-K. & Keasling, J. D. Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s. *Nat. Chem. Biol.* **3**, 274–277 (2007).
 31. Morrone, D. *et al.* Increasing diterpene yield with a modular metabolic engineering system in *E. coli*: Comparison of MEV and MEP isoprenoid precursor pathway engineering. *Appl. Microbiol. Biotechnol.* **85**, 1893–1906 (2010).
 - 20 32. Ferguson, F. M. & Gray, N. S. Kinase inhibitors: The road ahead. *Nature Reviews Drug Discovery* **17**, 353–376 (2018).
 33. Stanford, S. M. & Bottini, N. Targeting Tyrosine Phosphatases: Time to End the Stigma. *Trends in Pharmacological Sciences* (2017). doi:10.1016/j.tips.2017.03.004
 - 25 34. Tonks, N. K. Protein tyrosine phosphatases: from genes, to function, to disease. *Nat. Rev. Mol. Cell Biol.* **7**, 833–846 (2006).
 35. Tautz, L., Pellicchia, M. & Mustelin, T. Targeting the PTPome in human disease. *Expert Opin. Ther. Targets* **10**, 157–77 (2006).
 - 30 36. Tonks, N. K. Protein tyrosine phosphatases - From housekeeping enzymes to master regulators of signal transduction. *FEBS Journal* **280**, 346–378 (2013).
 37. Scott, L. M., Lawrence, H. R., Sebt, S. M., Lawrence, N. J. & Wu, J. Targeting protein tyrosine phosphatases for anticancer drug discovery. *Curr. Pharm. Des.* **16**,

- 1843–62 (2010).
38. Montalibet, J. & Kennedy, B. P. Using yeast to screen for inhibitors of protein tyrosine phosphatase 1B. *Biochem. Pharmacol.* **68**, 1807–1814 (2004).
39. Badran, A. H. *et al.* Continuous evolution of *Bacillus thuringiensis* toxins overcomes
5 insect resistance. *Nature* **533**, 58–63 (2016).
40. Kaneko, T. *et al.* Superbinder SH2 domains act as antagonists of cell signaling. *Sci. Signal.* **5**, (2012).
41. Jiang, C.-S., Liang, L.-F. & Guo, Y.-W. Natural products possessing protein tyrosine phosphatase 1B (PTP1B) inhibitory activity found in the last decades. *Acta
10 Pharmacol. Sin.* **33**, 1217–1245 (2012).
42. Hjortness, M. K. *et al.* Abietane-Type Diterpenoids Inhibit Protein Tyrosine Phosphatases by Stabilizing an Inactive Enzyme Conformation. *Biochemistry* **57**, 5886–5896 (2018).
43. Martin, V. J. J., Pitera, D. J., Withers, S. T., Newman, J. D. & Keasling, J. D.
15 Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat. Biotechnol.* **21**, 796–802 (2003).
44. He, R., Yu, Z., Zhang, R. & Zhang, Z. Protein tyrosine phosphatases as potential therapeutic targets. *Acta Pharmacol. Sin.* **35**, 1227–1246 (2014).
45. Bentires-Alj, M. & Neel, B. G. Protein-tyrosine phosphatase 1B is required for
20 HER2/Neu-induced breast cancer. *Cancer Res.* (2007). doi:10.1158/0008-5472.CAN-06-4610
46. Krishnan, N. *et al.* PTP1B inhibition suggests a therapeutic strategy for Rett syndrome. *J. Clin. Invest.* (2015). doi:10.1172/JCI80323
47. Yoshikuni, Y., Ferrin, T. E. & Keasling, J. D. Designed divergent evolution of enzyme
25 function. *Nature* **440**, 1078–1082 (2006).
48. Carlson, J. C., Badran, A. H., Guggiana-Nilo, D. A. & Liu, D. R. Negative selection and stringency modulation in phage-assisted continuous evolution. *Nat. Chem. Biol.* **10**, 216–222 (2014).
49. Dietrich, J. A. *et al.* A novel semi-biosynthetic route for artemisinin production using
30 engineered substrate-promiscuous P450BM3. *ACS Chem. Biol.* **4**, 261–267 (2009).
50. Edgar, S. *et al.* Mechanistic Insights into Taxadiene Epoxidation by Taxadiene-5 α -Hydroxylase. *ACS Chem. Biol.* **11**, 460–469 (2016).
51. Chen, X. *et al.* Statistical experimental design guided optimization of a one-pot

- biphasic multienzyme total synthesis of amorpha-4,11-diene. *PLoS One* **8**, e79650 (2013).
52. Waterhouse, A. *et al.* SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* (2018). doi:10.1093/nar/gky427
- 5 53. Guex, N., Peitsch, M. C. & Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis* (2009). doi:10.1002/elps.200900140
54. Benkert, P., Biasini, M. & Schwede, T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* (2011).
10 doi:10.1093/bioinformatics/btq662
55. Tian, J. Q. and J. & Quan, J. Circular Polymerase Extension Cloning of Complex Gene Libraries and Pathways. *PLoS One* **4**, e6441 (2009).
56. Burnham, K. P. & Anderson, D. R. *Model Selection and Multimodel Inference: a Practical Information-theoretic Approach*, 2nd edn. Springer-Verlag, New York. New York Springer **60**, (2002).
15
57. Davis, J. H., Rubin, A. J. & Sauer, R. T. Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Res.* **39**, 1131–1141 (2011).
58. Salis, H. M. The ribosome binding site calculator. *Methods Enzymol.* **498**, 19–42 (2011).
- 20 59. Sato, M., Ozawa, T., Inukai, K., Asano, T. & Umezawa, Y. Fluorescent indicators for imaging protein phosphorylation in single living cells. *Nat Biotechnol* **20**, 287–294 (2002).

Example 2

The design of small molecules that inhibit disease-relevant proteins represents a longstanding challenge of medicinal chemistry. Here, we describe an approach for encoding this challenge—the inhibition of a human drug target—into a microbial host and using it to guide the discovery and biosynthesis of targeted, biologically active natural products. This approach identified two previously unknown terpenoid inhibitors of protein tyrosine phosphatase 1B (PTP1B), an elusive therapeutic target for the treatment of diabetes and cancer. At least one inhibitor targets an allosteric site, which confers unusual selectivity; both can inhibit PTP1B in living cells. A screen of 24 uncharacterized terpene synthases from a pool of 4,464 genes uncovered additional hits, demonstrating a scalable discovery approach, and the incorporation of different PTPs into the microbial host yielded PTP-specific detection systems. Findings illustrate the potential for using microbes to discover and build natural products that exhibit precisely defined biochemical activities yet possess unanticipated structures and/or binding sites.

Despite advances in structural biology and computational chemistry, the design of small molecules that bind tightly and selectively to disease-relevant proteins remains exceptionally difficult¹. The free energetic contributions of rearrangements in the molecules of water that solvate binding partners and structural changes in the binding partners themselves are particularly challenging to predict and, thus, to incorporate into molecular design^{2,3}. Drug development, as a result, often begins with screens of large compound libraries⁴.

Nature has endowed living systems with the catalytic machinery to build an enormous variety of biologically active molecules—a diverse natural library⁵. These molecules evolved to carry out important metabolic and ecological functions (e.g., the phytochemical recruitment of predators of herbivorous insects⁶) but often also exhibit useful medicinal properties. Over the years, screens of environmental extracts and natural product libraries—augmented, on occasion, with combinatorial (bio)chemistry^{7–9}—have uncovered a diverse set of therapeutics, from aspirin to paclitaxel¹⁰. Unfortunately, these screens tend to be resource intensive¹¹, limited by low natural titers¹², and largely subject to serendipity¹³. Bioinformatic tools, in turn, have permitted the identification of biosynthetic gene clusters^{14,15}, where co-localized resistance genes can reveal the biochemical function of their products^{16,17}. The

therapeutic applications of many natural products, however, differ from their native functions¹⁸, and many biosynthetic pathways can, when appropriately reconfigured, produce entirely new and, perhaps, more effective therapeutic molecules^{19,20}. Methods for efficiently identifying and building natural products that inhibit specific disease-relevant proteins remain largely undeveloped.

Protein tyrosine phosphatases (PTPs) are an important class of drug targets that could benefit from new approaches to inhibitor discovery. These enzymes catalyze the hydrolytic dephosphorylation of tyrosine residues and, together with protein tyrosine kinases (PTKs), contribute to an enormous number of diseases (e.g., cancer, autoimmune disorders, and heart disease, to name a few)^{21,22}. The last several decades have witnessed the construction of many potent inhibitors of PTKs, which are targets for over 30 approved drugs²³. Therapeutic inhibitors of PTPs, by contrast, have proven difficult to develop. These enzymes possess well conserved, positively charged active sites that make them difficult to inhibit with selective, membrane-permeable molecules²⁴; they lack targeted therapeutics of any kind.

In this study, we describe an approach for using microbial systems to find natural products that inhibit difficult-to-drug proteins. We focused on protein tyrosine phosphatase 1B (PTP1B), a therapeutic target for the treatment of type 2 diabetes, obesity, and HER2-positive breast cancer²⁵. PTP1B possesses structural characteristics that are generally representative of the PTP family²⁶ and regulates a diverse set of physiological processes (e.g., energy expenditure²⁷, inflammation²⁸, and neural specification in embryonic stem cells²⁹). In brief, we assembled a strain of *Escherichia coli* with two genetic modules—(i) one that links cell survival to the inhibition of PTP1B and (ii) one that enables the biosynthesis of structurally varied terpenoids. In a study of five well-characterized terpene synthases, this strain identified two previously unknown terpenoid inhibitors of PTP1B. Both inhibitors were selective for PTP1B, exhibited distinct binding mechanisms, and increased insulin receptor phosphorylation in mammalian cells. A screen of 24 uncharacterized terpene synthases from eight phylogenetically diverse clades uncovered additional hits, demonstrating a scalable approach for finding inhibitor-synthesizing genes. A simple exchange of PTP genes, in turn, permitted the facile extension of our genetically encoded detection system to new targets. Our findings illustrate a versatile approach for using microbial systems to find targeted, readily synthesizable inhibitors of disease-relevant enzymes.

Development of a genetically encoded objective

E. coli is a versatile platform for building natural products from unculturable or low-yielding organisms^{30,31}. We hypothesized that a strain of *E. coli* programmed to detect the inactivation of PTP1B (i.e., a genetically encoded objective) might enable the discovery of natural products that inhibit it (i.e., molecular solutions to the objective). To program such a strain, we assembled a bacterial two-hybrid (B2H) system in which PTP1B and Src kinase control gene expression (**FIG. 21a**). In this system, Src phosphorylates a substrate domain, enabling a protein-protein interaction that activates transcription of a gene of interest (GOI). PTP1B dephosphorylates the substrate domain, preventing that interaction, and the inactivation of PTP1B re-enables it. *E. coli* is a particularly good host for this detection system because its proteome is sufficiently orthogonal to the proteome of *H. sapiens* to minimize off-target growth defects that can result from the regulatory activities of Src and PTP1B (Note 1)³².

We carried out B2H development in several steps. To begin, we assembled a luminescent “base” system in which Src modulates the binding of a substrate domain to an Src homology 2 (SH2) domain (**FIG. 21b**); this system, which includes a chaperone that helps Src to fold (Cdc37)³³, is similar to other B2H designs that detect protein-protein binding³⁴. Unfortunately, our initial system did not yield a phosphorylation-dependent transcriptional response, so we complemented it with inducible plasmids—each harboring a different system component—to identify proteins with suboptimal expression levels (**FIG. 21b**). Interestingly, secondary induction of Src increased luminescence, an indication that insufficient substrate phosphorylation and/or weak substrate-SH2 binding depressed GOI expression in our base system. We modified this system by swapping in different substrate domains, by adding mutations to the SH2 domain that enhance its affinity for phosphopeptides³⁵, and by removing the gene for Src—a modification that allowed us to control expression exclusively from a second plasmid. With this configuration, induction of Src increased luminescence most prominently for the MidT substrate (**FIG. 1c**), and simultaneous induction of both Src and PTP1B prevented that increase—an indication of intracellular PTP1B activity (**FIG. 21d**). We finalized the MidT system by incorporating genes for PTP1B and Src, by adjusting promoters and ribosome binding sites to amplify its transcriptional response further (**FIG. 21d**, **FIG. 13**, and **FIG. 14**), and by adding a gene for spectinomycin resistance (SpecR) as the GOI. The final plasmid-borne detection system

required the inactivation of PTP1B to permit growth at high concentrations of antibiotic (FIG. 21e).

Biosynthesis of PTP1B inhibitors

5 To search for inhibitors of PTP1B that bind outside of its active site, we coupled the B2H system with metabolic pathways for terpenoids, a structurally diverse class of secondary metabolites with largely nonpolar structures (FIG. 22a), some of which are known to inhibit PTP1B^{36,37}. Terpenoids include over 80,000 known compounds and represent nearly one-third of all characterized natural products³⁸ (the basis of approximately 50% of clinically
10 approved drugs³⁹). To begin, we focused on a handful of structurally diverse terpenoids without established inhibitory effects (FIG. 22b): Amorphadiene (AD), γ -humulene, α -bisabolene (AB), abietadiene, and taxadiene. Each terpenoid pathway consisted of two plasmid-borne modules: (i) the mevalonate-dependent isoprenoid pathway from *S. cerevisiae* (optimized for expression in *E. coli*⁴⁰) and (ii) a terpene synthase previously demonstrated to
15 express and produce one of the five selected terpenoids in *E. coli*⁴⁰⁻⁴⁴. The terpene synthase was supplemented, when necessary for diterpenoid production, with a geranylgeranyl diphosphate synthase. These modules generated terpenoids at titers of 0.3-18 mg/L in *E. coli* (FIG. 26).

We screened each pathway for its ability to produce inhibitors of PTP1B by
20 transforming *E. coli* with plasmids harboring both the pathway of interest and the B2H system (FIG. 22c). To our surprise, pathways for AD and AB permitted survival at high concentrations of antibiotic. Critically, GC-MS traces confirmed that all pathways generated terpenoids in the presence of the B2H system (FIG. 22d, FIG. 26), and maximal resistance of the AD- and AB-producing strains required both an active terpene synthase and a
25 functional B2H system (FIG. 26d).

We confirmed the inhibitory effects of purified terpenoids by examining their influence on PTP1B-catalyzed hydrolysis of *p*-nitrophenyl phosphate (*p*NPP; FIG. 22e, TABLE 12). The IC₅₀s for AD and AB were 53 $\pm$ 8 μ M and 13 $\pm$ 2 μ M, respectively, in 10% DMSO (FIG. 22f). These IC₅₀s are surprisingly strong for small, unfunctionalized
30 hydrocarbons; the ligand efficiencies of both inhibitors are high (TABLE 15), and their potencies are similar to those of larger molecules that form hydrogen bonds and other stabilizing interactions with PTP1B^{21,45}. Both IC₅₀s are also similar to the respective terpenoid concentrations in liquid culture (FIG. 22g), a finding consistent with *in vivo*

inhibition (terpenoids tend to accumulate intracellularly⁴⁶, so *in vivo* concentrations may be even higher). Our growth-coupled assays, kinetic assays, and production measurements, taken together, indicate that AD and AB activate the B2H system by inhibiting PTP1B inside the cell.

5

Biophysical analysis of PTP1B inhibitors

Allosteric inhibitors of PTPs are valuable starting points for drug development. These molecules bind outside of the well conserved, positively charged active sites of PTPs and tend to have improved selectivities and membrane permeabilities over substrate analogs²¹.

10 Motivated by these considerations, an early screen identified a benzbromarone derivative that inhibited PTP1B weakly ($IC_{50} = 350 \mu M$) without competing with substrates; subsequent optimization of this compound led to two improved inhibitors (IC_{50} 's = 8 and 22 μM) that bind to an allosteric site⁴⁵ (**FIG. 23a**). Over the next 15 years, efforts to find new inhibitors that bind to this or other allosteric regions on the catalytic domain have been largely
15 unsuccessful⁴⁷. Benzbromarone derivatives are the only allosteric inhibitors with crystallographically verified binding sites. (Although, an allosteric inhibitor that binds to a disordered region of the full-length protein has been characterized with NMR²⁵). New approaches for finding allosteric inhibitors are clearly needed.

Our microbial system could grant access to new compounds that bind in unexpected
20 ways. AD and AB provide examples. They are highly nonpolar and, thus, incapable of engaging in the hydrogen bonds and electrostatic interactions on which most other PTP inhibitors rely^{21,45}. To examine their binding mechanisms in detail, we sought to collect X-ray crystal structures of PTP1B bound to AD and α -bisabolol, a soluble analogue of AB (a ligand for which poor solubility precluded soaking experiments). Unfortunately, only the structure
25 of PTP1B bound to AD was sufficient for unambiguous determination of a binding site (**FIG. 30 and FIG. 31**). This inhibitor binds to the same allosteric site targeted by benzbromarone derivatives. Its binding mode, however, is distinct: (i) AD causes the $\alpha 7$ helix of PTP1B to reorganize to create a hydrophobic cleft (**FIG. 23b**); this type of reorganization is interesting because it is typically slow (micro- to millisecond)⁴⁸ and difficult to incorporate into
30 computational ligand design⁴⁹. (ii) It likely adopts multiple bound conformations (i.e., the electron density indicates regions of disorder; **FIG. 30**). This behavior, which is supported by molecular dynamics simulations, is consistent with prior work on the binding of proteins to hydrocarbon moieties, which tend to be "mobile" in their binding pockets.

We probed the binding of AD and AB further with several additional analyses. First, we examined the inhibition of PTP1B by dihydroartemisinic acid. This structural analogue of AD has a carboxyl group that, according to our crystal structure, should interfere with binding to the hydrophobic cleft created by the $\alpha 7$ helix (**FIG. 23c**). The IC_{50} of this molecule was eight-fold higher than that of AD, a reduction in potency consistent with its crystallographic pose (**FIG. 23d** and **FIG. 33**). Second, we studied the competition between AD and two inhibitors that bind to the active site: (i) TCS401, which causes the WPD loop to adopt a closed conformation, and (ii) orthovanadate, which does not. For background, benzobromarones, upon binding to the C-terminal allosteric site, stabilize the WPD loop in an open conformation that is incompatible with the binding of TCS401, but not orthovanadate. Our kinetic data suggest that AD behaves similarly (**FIG. 23e** and **FIG. 23f**), a finding consistent with a shared binding site and mechanism of modulation. Finally, we assessed the inhibitory effects of AD and AB against TC-PTP, the closest homolog of PTP1B. Intriguingly, both molecules inhibited TC-PTP five- to six-fold less potently than PTP1B (**FIG. 23g** and **FIG. 33**). This finding is consistent with binding to the poorly conserved allosteric site. Importantly, this selectivity may seem modest, but it matches or exceeds the selectivities of most pre-optimized inhibitors (including benzobromarone derivatives) and is exceedingly rare for unfunctionalized hydrocarbons⁵⁰. We assessed the contribution of the $\alpha 7$ helix to selectivity, in turn, by removing the equivalent region from PTP1B and TC-PTP (**FIG. 23g**). This modification caused a four-fold reduction in the selectivity of AD, an effect consistent with the involvement of the $\alpha 7$ helix in its binding. Intriguingly, the selectivity of AB was insensitive to this modification; the unambiguous determination of the binding site of this ligand requires additional data.

AD and AB are lipophilic molecules that could be valuable for their ability to pass through the membranes of mammalian cells. To examine the biological activity of these molecules, we incubated them with HEK293T/17 cells and used an enzyme-linked immunosorbent assay to measure shifts in insulin receptor (IR) phosphorylation. IR is a receptor tyrosine kinase that undergoes PTP1B-mediated dephosphorylation from the cytosolic side of the plasma membrane (PTP1B, in turn, localizes to the endoplasmic reticulum of the cell). Both molecules increased IR phosphorylation over a negative control (**FIG. 23h** and **FIG. 35**). We checked for off-target contributions to this signal, in turn, by repeating the ELISA with equivalent concentrations of dihydroartemisinic acid and α -

bisabolol. To our satisfaction, both molecules led to a reduction in signal consistent with their reduced potencies.

Other PTPs can promote IR dephosphorylation; SHP1 and SHP2 provide two examples^{51–53}. To examine the potential contribution of these enzymes to the increase in IR phosphorylation observed in our ELISA, we measured their inhibition by AD and AB. Briefly, AD inhibited SHP2 three-fold less potently than PTP1B, and its inhibition of SHP1 was too weak to measure (**FIGs. 34a–34b**). The low potency of AB against SHP1 and SHP2 also precluded experimental measurement (**FIGs. 34c–34d**). These potencies, together with the aforementioned analysis of weakly inhibitory structural analogs, suggest that the inhibition of PTP1B by AD and AB is the primary cause of the increase in IR phosphorylation observed in our ELISA experiments.

A scalable approach to molecular discovery

Our microbial strain provides a powerful tool for screening genes for their ability to generate novel PTP1B inhibitors. Most terpenoids, as a case study, are not commercially available, and even when their metabolic pathways are known, their biosynthesis, purification, and *in vitro* analysis is a resource-intensive process that is difficult to parallelize with existing methods⁵⁴. Our B2H system offers a potential solution: It can identify inhibitor-synthesizing genes with a simple growth-coupled assay. We explored its application to discovery efforts by using it to screen a diverse set of uncharacterized biosynthetic genes. In brief, we carried out a bioinformatic analysis of the largest terpene synthase family (PF03936) by building and annotating a cladogram of its 4,464 constituent members (**FIG. 27**); from here, we synthesized three uncharacterized genes from each of eight clades: six with no characterized genes and two with some characterized genes (**FIG. 24a**). We reasoned that these 24 phylogenetically diverse genes (8 from fungi, 13 from plants, and 3 from bacteria) might encode enzymes with distinct product profiles and potentially, through the inclusion of uncharacterized clades, novel sesquiterpene scaffolds.

Guided by our initial screen, we searched for sesquiterpene inhibitors by pairing each of the uncharacterized genes with the FPP pathway. To our surprise, six genes conferred a significant survival advantage (**FIG. 24b**), and maximal resistance required an active B2H system (**FIG. 28**). Each hit generated distinct product profiles (**FIG. 29**); we focused our analysis on A0A0C9VSL7, which produced mostly (+)-1(10),4-cadinadiene as a major product (**FIGs. 24c–24d**). This terpenoid is a structural analog of AD but has a weaker

potency ($IC_{50} = 165 \pm 33 \mu M$; **FIG. 24e**); a titer of $33 \pm 18 \mu M$ suggests that intracellular accumulation may allow it to inhibit PTP1B inside the cell. Our ability to detect a weak inhibitor suggests that the B2H system can capture a broad set of scaffolds in molecular discovery efforts. The purification and analysis of additional hits, the incorporation of isoprenoid substrates of different sizes (through the use of geranyl diphosphate synthase or geranyl geranyl diphosphate synthase), and the inclusion of more uncharacterized genes could expand the scope of such efforts.

Design of alternative PTP-specific objectives

We explored the versatility of our B2H system by assessing its ability to detect the inactivation of several other diseases-relevant PTPs. In short, we swapped out the gene for PTP1B with genes for PTPN2, PTPN6, or PTPN12; these enzymes are targets for immunotherapeutic enhancement⁵⁵, the treatment of ovarian cancer⁵⁶, and acute myocardial infarction⁵⁷, respectively. Their catalytic domains share 31-65% sequence identity with the catalytic domain of PTP1B. Interestingly, the new B2H systems were immediately functional; PTP inactivation permitted growth at high concentrations of spectinomycin (**FIG. 25a**). This finding suggests that our detection system can be easily extended to other members of the PTP family.

PTP-specific B2H systems could facilitate the identification of natural products that selectively inhibit one PTP over another. We explored this application by comparing the antibiotic resistance conferred by PTP1B- and TC-PTP-specific systems in response to metabolic pathways for AD and α -bisabolene (**FIG. 25b**). As expected, the PTP1B-specific system permitted growth at higher concentrations of antibiotic, a result consistent with the selectivity of both terpenoids for PTP1B. Indistinguishable terpenoid titers between the two strains suggest that this survival advantage does not result from difference in intracellular concentration (**FIG. 25c**). Findings thus indicate that a simple comparison of B2H systems—a potential secondary screen—offers a simple approach for evaluating the selectivity PTP-inhibiting gene products. Notably, high concentrations of inhibitors in two strains could swamp out selective effects; in such cases, terpenoid levels could be reduced with lower mevalonate concentrations.

This study addresses an important challenge of medicinal chemistry—the design of molecular structures that inhibit disease-relevant enzymes—by using a desired biochemical

activity (i.e., an objective) as a genetically encoded constraint to guide molecular biosynthesis. This approach enabled the identification of two selective, biologically active inhibitors of PTP1B, an elusive drug target⁵⁸. These molecules are not drugs, but they are promising scaffolds for lead development. Their mechanisms of modulation—which elicit
5 allosteric conformational changes yet appear to rely on loose, conformationally flexible binding—are unusual (and computationally elusive⁵⁹), and demonstrate the ability of microbial systems to find new solutions to difficult challenges in molecular design. Our identification of unusual inhibitors in relatively small libraries, in turn, suggests that microbial systems can access a rich molecular landscape that is not efficiently explored by
10 existing approaches to molecular discovery.

The B2H system at the core of our approach is a valuable tool for identifying biologically active natural products, which are structurally complex, difficult to synthesize, and often hidden in cryptic gene clusters⁶⁰. It has several key advantages over contemporary approaches to inhibitor discovery: (i) It incorporates synthesizability as a search criterion—an
15 important attribute of drug leads⁶¹. (ii) It is scalable. We used a growth-coupled assay to screen 24 uncharacterized terpene synthases; this type of assay is also compatible with very large mutagenesis libraries (e.g., 10^{10})⁶². (iii) It can use cellular machinery to stabilize proteins (e.g., CDC37 for Src); this capability could facilitate the integration of unstable and/or disordered targets. Future efforts to exploit these advantages by incorporating large
20 libraries of mutated and/or reconfigured pathways, alternative biosynthetic enzymes (e.g., cytochromes P450, halogenases, and methyltransferases), or new classes of disease-relevant enzymes would be informative.

The B2H system also has important limits. When used alongside metabolic pathways, it links survival not only to the potency of metabolites, but also to their titers, off-target
25 effects, and pathway toxicities. These limitations can be beneficial; they bias the discovery process toward potent, readily synthesizable inhibitors and could, thus, facilitate post-discovery efforts to improve the titers of interesting molecules⁶³. Nonetheless, they will exclude some types of structurally complex molecules that are difficult to synthesize in *E. coli*. The use of similar activity-based screens in other organisms (e.g., *Streptomyces*) could
30 be interesting.

The compatibility of our discovery approach with different PTPs is valuable in light of their increasingly well validated potential as a rich—and essentially untapped—source of new therapeutic targets⁶⁴. We anticipate that some PTPs will require the use of chaperones

and/or transcriptional adjustments to be incorporated into B2H systems. Our systematic optimization of the PTP1B-based system provides an experimental framework for exploring these modifications. Side-by-side comparisons of B2H systems, in turn, offer a promising strategy for evaluating inhibitor selectivity in secondary screens. In future work, new varieties of objectives (e.g., B2H systems or genetic circuits that detect the selective inhibition—or, perhaps, activation—of one PTP over another) could facilitate the discovery of molecules with sophisticated mechanisms of modulation in primary screens. The versatility of genetically encoded objectives highlights the power of using microbial systems to find targeted, biologically active molecules.

Note 1: The orthogonality of proteomes. *E. coli* and *S. cerevisiae* are both well-developed platforms for the production of pharmaceutically relevant natural products^{20,65,66}. We chose to use *E. coli* for this study because its machinery for phosphorylating proteins is dissimilar from that of eukaryotic cells and thus less likely to interfere with the function of genetically encoded systems that link the inhibition of PTP1B to cellular growth⁶⁷. By contrast, the overexpression of Src kinase in *S. cerevisiae* is lethal and is mitigated by PTP1B⁶⁸; these effects are inconsistent with our biochemical objective. More broadly, *S. cerevisiae* and humans, despite having evolved from a common ancestor approximately 1 billion years ago⁶⁹, share many functionally equivalent proteins; orthologous genes, in fact, account for more than one-third of the yeast genome⁷⁰. Most strikingly, a recent study found that nearly half (47%) of 414 essential genes from *S. cerevisiae* could be replaced with human orthologs without growth defects⁷¹. This finding suggests that yeast is a particularly restrictive host for genetically encoded systems that link arbitrary changes in the activities of human regulatory enzymes to fitness advantage.

METHODS

Bacterial strains. We used *E. coli* DH10B, chemically competent NEB Turbo, or electrocompetent One Shot Top10 (Invitrogen) to carry out molecular cloning and to perform preliminary analyses of terpenoid production; we used *E. coli* BL2-DE31 to express proteins for *in vitro* studies; and we used *E. coli* s1030⁷² for our luminescence studies and for all experiments involving terpenoid-mediated growth (i.e., evolution studies).

For all strains, we generated chemically competent cells by carrying out the following steps: (i) We plated each strain on LB agar plates with the required antibiotics. (ii) We used

one colony of each strain to inoculate 1 mL of LB media (25 g/L LB with appropriate antibiotics listed in **TABLE 8**) in a glass culture tube, and we grew this culture overnight (37°C, 225 RPM). (iii) We used the 1-mL culture to inoculate 100-300 mL of LB media (as above) in a glass shake flask, and we grew this culture for several hours (37°C, 225 RPM).

5 (iv) When the culture reached an OD of 0.3-0.6, we centrifuged the cells (4,000 x g for 10 minutes at 4°C), removed the supernatant, resuspended them in 30 mL of ice cold TFB1 buffer (30 mM potassium acetate, 10 mM CaCl₂, 50 mM MnCl₂, 100 mM RbCl, 15% v/v glycerol, water to 200 mL, pH=5.8, sterile filtered), and incubated the suspension at 4°C for 90 min. (v) We repeated step iv, but resuspended in 4 mL of ice cold TFB2 buffer (10 mM

10 MOPS, 75 mM CaCl₂, 10 mM RbCl₂, 15% glycerol, water to 50 mL, pH=6.5, sterile filtered). (iv) We split the final suspension into 100 µL aliquots and froze them at -80°C until further use.

We generated electrocompetent cells by following an approach similar to the one above. In step iv, however, we resuspended the cells in 50 mL of ice cold MilliQ water and

15 repeated this step twice—first with 50 mL of 20% sterile glycerol (ice cold) and, then, with 1 mL of 20% sterile glycerol (ice cold). We froze the pellets as before.

Materials. We purchased methyl abietate from Santa Cruz Biotechnology; trans-caryophyllene, tris(2-carboxyethyl)phosphine (TCEP), bovine serum albumin (BSA), M9

20 minimal salts, phenylmethylsulfonyl fluoride (PMSF), and DMSO (dimethyl sulfoxide) from Millipore Sigma; glycerol, bacterial protein extraction reagent II (B-PERII), and lysozyme from VWR; cloning reagents from New England Biolabs; AD from Ambeed, Inc.; and all other reagents (e.g., antibiotics and media components) from Thermo Fisher. Taxadiene was a kind gift from Phil Baran of the The Scripps Research Institute. We prepared mevalonate by

25 mixing 1 volume of 2 M DL-mevalanolactone with 1.05 volumes of 2 M KOH and incubating this mixture at 37°C for 30 minutes.

Cloning and molecular biology. We constructed all plasmids by using standard methods (i.e., restriction digest and ligation, Golden Gate and Gibson assembly, Quikchange

30 mutagenesis, and circular polymerase extension cloning). **TABLE 7** describes the source of each gene; **TABLE 8** and **TABLE 3** describe the composition of all final plasmids.

We began construction of the B2H system by integrating the gene for HA4-RpoZ from pAB094a into pAB078d and by replacing the ampicillin resistance marker of pAB078d

with a kanamycin resistance marker (Gibson Assembly). We modified the resulting “combined” plasmid, in turn, by replacing the HA4 and SH2 domains with kinase substrate and substrate recognition (i.e., SH2) domains, respectively (Gibson assembly), and by integrating genes for Src kinase, CDC37, and PTP1B in various combinations (Gibson assembly). We finalized the functional B2H system by modifying the SH2 domain with several mutations known to enhance its affinity for phosphopeptides (K15L, T8V, and C10A, numbered as in Kaneko et. al.³⁵), by exchanging the GOI for luminescence (LuxAB) with one for spectinomycin resistance (SpecR), and by toggling promoters and ribosome binding sites to enhance the transcriptional response (Gibson assembly and Quickchange Mutagenesis, Agilent Inc.). We note: For the last step, we also converted ProI to ProD by using the Quickchange protocol. When necessary, we constructed plasmids with arabinose-inducible components by cloning a single component from the B2H system into pBAD (Golden Gate assembly). **TABLE 4, TABLE 9, and TABLE 10** list the primers and DNA fragments used to construct each plasmid.

We assembled pathways for terpenoid biosynthesis by purchasing plasmids encoding the first module (pMBIS) and various sesquiterpene synthases (ADS or GHS in pTRC99a) from Addgene, and by building the remaining plasmids. We replaced the tetracycline resistance in pMBIS with a gene for chloramphenicol resistance to create pMBIS_{CmR}. We integrated genes for ABS, TXS, ABA, and GGPPS into pTRC99t (i.e., pTRC99a without BsaI sites). **TABLE 4, TABLE 9, and TABLE 10** list the primers and DNA fragments used to construct each plasmid.

Luminescence assays. We characterized preliminary B2H systems (which contained LuxAB as the GOI) with luminescence assays. In brief, we transformed necessary plasmids into *E. coli* s1030 (**TABLE 8**), plated the transformed cells onto LB agar plates (20 g/L agar, 10 g/L tryptone, 10 g/L sodium chloride, and 5 g/L yeast extract with antibiotics described in **TABLE 8**), and incubated all plates overnight at 37°C. We used individual colonies to inoculate 1 ml of terrific both (TB at 2%, or 12 g/L tryptone, 24 g/L yeast extract, 12 mL/L 100% glycerol, 2.28 g/L KH₂PO₄, 12.53 g/L K₂HPO₄, pH = 7.3, and antibiotics described in **TABLE 8**), and we incubated these cultures overnight (37°C and 225 RPM). The following morning, we diluted each culture by 100-fold into 1 ml of TB media (above), and we incubated these cultures in individual wells of a deep 96-well plate for 5.5 hours (37°C, 225 RPM). (We note: When pBAD was present, we supplemented the TB media with 0-0.02 w/v

% arabinose). We transferred 100 μ L of each culture into a single well of a standard 96-well clear plate and measured both OD₆₀₀ and luminescence on a Biotek Synergy plate reader (gain: 135, integration time: 1 second, read height: 1 mm). Analogous measurements of cell-free media allowed us to measure background signals, which we subtracted from each measurement prior to calculating OD-normalized luminescence (i.e., Lum / OD₆₀₀).

Analysis of antibiotic resistance. We evaluated the spectinomycin resistance conferred by various B2H systems in the absence of terpenoid pathways by carrying out the following steps: (i) We transformed *E. coli* with the necessary plasmids (**TABLE 8**) and plated the transformed cells onto LB agar plates (20 g/L agar, 10 g/L tryptone, 10 g/L sodium chloride, 5 g/L yeast extract, 50 μ g/ml kanamycin, 10 μ g/ml tetracycline). (ii) We used individual colonies to inoculate 1-2 ml of TB media (12 g/L tryptone, 24 g/L yeast extract, 12 mL/L 100% glycerol, 2.28 g/L KH₂PO₄, 12.53 g/L K₂HPO₄, 50 μ g/ml kanamycin, 10 μ g/ml tetracycline, pH = 7.3), and we incubated these cultures overnight (37°C, 225 RPM). In the morning, we diluted each culture by 100-fold into 4 ml of TB media (as above) with 0-500 μ g/ml spectinomycin (we used spectinomycin in the liquid culture only for **FIG. 14**), and we incubated these cultures in deep 24-well plates until wells containing 0 μ g/ml spectinomycin reached an OD₆₀₀ of 0.9-1.1. (iv) We diluted each 4-ml culture by 10-fold into TB media with no antibiotics and plated 10- μ L drops of the diluent onto agar plates with various concentrations of spectinomycin. (v) We incubated plates overnight (37°C) and photographed them the following day.

To examine terpenoid-mediated resistance, we began with steps i and ii as described above with the addition of 34 μ g/ml chloramphenicol and 50 μ g/ml carbenicillin in all liquid/solid media. We then proceeded with the following steps: (iii) We diluted samples from 1-ml cultures to an OD₆₀₀ of 0.05 in 4.5 ml of TB media (supplemented with 12 g/L tryptone, 24 g/L yeast extract, 12 mL/L 100% glycerol, 2.28 g/L KH₂PO₄, 12.53 g/L K₂HPO₄, 50 μ g/ml kanamycin, 10 μ g/ml tetracycline, 34 μ g/ml chloramphenicol, and 50 μ g/ml carbenicillin), which we incubated in deep 24-well plates (37°C, 225 RPM). (iv) At an OD₆₀₀ of 0.3-0.6, we transferred 4 ml of each culture to a new well of a deep 24-well plate, added 500 μ M isopropyl β -D-1-thiogalactopyranoside (IPTG) and 20 mM of mevalonate, and incubated for 20 hours (22°C, 225 RPM). (v) We diluted each 4-ml culture to an OD₆₀₀ of 0.1 with TB media and plated 10 μ L of the diluent onto either LB or TB plates supplemented with 500 μ M IPTG, 20 mM mevalonate, 50 μ g/ml kanamycin, 10 μ g/ml tetracycline, 34

µg/ml chloramphenicol, 50 µg/ml carbenicillin, and 0-1200 µg/ml spectinomycin (for both plates, we used 20 g/L agar with media and buffer components described above).

- Terpenoid biosynthesis.** We prepared *E. coli* for terpenoid production by transforming cells with plasmids harboring requisite pathway components (**TABLE 8**) and plating them onto LB agar plates (20 g/L agar, 10 g/L tryptone, 10 g/L sodium chloride, and 5 g/L yeast extract with antibiotics described in **TABLE 8**). We used one colony from each strain to inoculate 2 ml TB (12 g/L tryptone, 24 g/L yeast extract, 12 mL/L 100% glycerol, 2.28 g/L KH₂PO₄, 12.53 g/L K₂HPO₄, pH = 7.0, and antibiotics described in **TABLE 8**) in a glass culture tube for ~16 hours (37°C and 225 RPM). We diluted these cultures by 75-fold into 10 ml of TB media and incubated the new cultures in 125 mL glass shake flasks (37°C and 225 RPM). At an OD₆₀₀ of 0.3-0.6, we added 500 µM IPTG and 20 mM mevalonate. After 72-88 hours of growth (22°C and 225 RPM), we extracted terpenoids from each culture as outlined below.
- Protein expression and purification.** We expressed and purified PTPs as described previously⁷³. Briefly, we transformed *E. coli* BL21(DE3) cells with pET16b or pET21b vectors (see **TABLE 8** for details), and we induced with 500 µM IPTG at 22°C for 20 hours. We purified PTPs from cell lysate by using desalting, nickel affinity, and anion exchange chromatography (HiPrep 26/10, HisTrap HP, and HiPrep Q HP, respectively; GE Healthcare). We stored the final protein (30–50 µM) in HEPES buffer (50 mM, pH 7.5, 0.5 mM TCEP) in 20% glycerol at –80°C.

- Extraction and purification of terpenoids.** We used hexane to extract terpenoids generated in liquid culture. For 10-mL cultures, we added 14 mL of hexane to 10 ml of culture broth in 125-mL glass shake flasks, shook the mixture (100 RPM) for 30 minutes, centrifuged it (4000 x g), and withdrew 10 mL of the hexane layer for further analysis. For 4-mL cultures, we added 600 µL hexane to 1 mL of culture broth in a microcentrifuge tube, vortexed the tubes for 3 minutes, centrifuged the tubes for 1 minute (17000 x g), and saved 300-400 µL of the hexane layer for further analysis.

- To purify AD, AB, and (+)-1(10),4-cadinadiene, we supplemented 500-1000 mL culture broth with hexane (16.7% v/v), shook the mixture for 30 minutes (100 RPM), isolated the hexane layer with a separatory funnel, centrifuged the isolated organic phase (4000 x g), and withdrew the hexane layer. To concentrate the terpenoid products, we evaporated excess

hexane in a rotary evaporator to bring the final volume to 500 μ L, and we passed the resulting mixture over a silica gel 1-3 times (Sigma-Aldrich; high purity grade, 60 Å pore size, 230-400 mesh particle size). We analyzed elution fractions (100% hexane) on the GC/MS and pooled fractions with the compound of interest (AD). Once purified, we dried pooled fractions under a gentle stream of air, resuspended the concentrated terpenoids in DMSO, and quantified the final samples as outlined below. We repeated the purification process until samples (in DMSO) were >95% pure by GC/MS unless otherwise noted.

GC-MS analysis of terpenoids. We measured terpenoids generated in liquid culture with a gas chromatograph / mass spectrometer (GC-MS; a Trace 1310 GC fitted with a TG5-SilMS column and an ISQ 7000 MS; Thermo Fisher Scientific). We prepared all samples in hexane (directly or through a 1:100 dilution of DMSO) with 20 μ g/ml of caryophyllene as an internal standard. Highly concentrated samples were diluted 10-20x prior to preparation to bring concentrations within the MS detection limit. When the peak area of an internal standard exceeded $\pm$ 40% of the average area of all samples containing that standard, we re-analyzed the corresponding samples. For all runs, we used the following GC method: hold at 80°C (3 min), increase to 250°C (15°C/min), hold at 250°C (6 min), increase to 280°C (30°C/min), and hold at 280°C (3 min). To identify various analytes, we scanned m/z ratios from 50 to 550.

We examined sesquiterpenes generated by variants of ADS by using select ion mode (SIM) to scan for the molecular ion (m/z = 204). For quantification, we used Eq. 1: where A_i

$$C_i = C_{std} * \frac{A_i}{A_{std}} * R \quad (\text{Eq. 1})$$

$$R = \frac{A_{std,o}/C_{std,o}}{A_{ref,o}/C_{ref,o}} \quad (\text{Eq. 2})$$

is the area of the peak produced by analyte i, A_{std} is the area of the peak produced by C_{std} of caryophyllene in the sample, and R is the ratio of response factors for caryophyllene and AD in a reference sample. TABLE 11 provides the concentrations of all standards and reference compounds used in this analysis.

We quantified diterpenoids by, once again, accompanying our general procedure with several modifications: We scanned for a different molecular ion (m/z = 272) and an ion common to both diterpenoids and caryophyllene (m/z=93); we used a ratio of response factors for pure taxadiene (a kind gift from Phil Baran) and caryophyllene at m/z = 93; and

we calculated peak areas $m/z = 93$. For all analyses, we examined only peaks with areas that exceeded 1% of the total area of all peaks at $m/z=272$.

We identified molecules by using the NIST MS library and, when necessary, confirmed this identification with analytical standards or mass spectra reported in the literature. We note: The assumption of a constant response factor for different terpenoids (that is, the assumption that all sesquiterpenes and diterpenes ionize like AD and taxadiene, respectively) can certainly yield error in estimates of their concentrations; our analyses, which are consistent with those of other studies of terpenoid production in microbial systems^{74,75}, supply rough estimates of concentrations for all compounds except AD and taxadiene (which had analytical standards).

Bioinformatics. We used a bioinformatic analysis to identify a phylogenetically diverse set of terpene synthases. Briefly, we downloaded (i) all constituent genes of PF03936 (the largest terpene synthase family grouped by a C-terminal domain) from the PFAM Database and (ii) all enzymes with Enzyme Commission (EC) number of 4.2.3.# from the Uniprot Database; this string, which defines carbon oxygen lyases that act on phosphates, includes terpene synthases. We cleaned both datasets in Excel (i.e., we ensured that every identifier had only one row), and we used a custom R script to designate each PF03936 member as characterized (i.e., in possession of a Uniprot-based EC number) or uncharacterized. Finally, we used FastTree⁷⁶ with default settings to create a phylogenetic tree of the PF03936 family and the R-package ggtree⁷⁷ to visualize the resulting tree and function data as a cladogram and heatmap.

After annotating the cladogram by hand, we selected three genes from each of six clades: six with no characterized genes and two with some characterized genes. We avoided clades proximal to known monoterpene synthases or diterpene synthases known to act on GGPP isomers absent in our system (e.g., ent-copalyl diphosphate); these enzymes are unlikely to act on FPP, the primary product of pMBIS_{CmR}. When selecting enzymes within clades, we biased our choice towards bacterial/fungal species and selected genes with a minimal number of common ancestors within the clade. The selected genes were synthesized and cloned into the pTrc99a vector by Twist Biosciences and assayed for antibiotic resistance as described above.

Enzyme kinetics. To examine terpenoid-mediated inhibition, we measured PTP-catalyzed hydrolysis of *p*-nitrophenyl phosphate (*p*NPP) or 4-methylumbelliferyl phosphate (4-MUP, used when K_M for *p*NPP was large) in the presence of various concentrations of terpenoids. Each reaction included PTP (0.05 μ M PTP1B/TCPTP or 0.1 μ M SHP1/SHP2 in 50 mM

5 HEPES, 0.5 mM TCEP, 50 μ g/ml BSA), *p*NPP (0.33, 0.67, 2, 5, 10, and 15 mM) or 4-MUP (0.13, 0.27, 0.8, 2.27, 2.93, 4.53, 7.07, and 8 mM), inhibitor (with concentrations listed in the figures), buffer (50 mM HEPES pH=7.3, 50 μ g/ml BSA), and DMSO at 10% v/v. We monitored the formation of *p*-nitrophenol by measuring absorbance at 405 nm every 10

10 seconds for 5 minutes on a SpectraMax M2 plate reader and the formation of 4-methylumbelliferyl by measuring fluorescence at 450 nm (370 nm ex, 435 nm cutoff, medium gain).

We used a custom MATLAB script to process all raw kinetic data. This script removed all concentration values that fell outside of either (i) the range of our standard curve (absorbance/fluorescence vs. μ M; **FIG. 39**) or (ii) the initial rate regime (>10% of the *p*NPP

15 or 4-MUP concentration used in the assay). When this step reduced kinetic dataset to fewer than ten points, we re-measured those datasets to collect at least ten. We fit final datasets, in turn, with a linear regression model (using Matlab's backslash operator).

We evaluated kinetic models in three steps: (i) We fit initial-rate measurements collected in the absence and presence of inhibitors to Michaelis-Menten and inhibition

20 models, respectively (here, we used the *nlinfit* and *fminsearch* functions from MATLAB; **TABLE 12**). (ii) We used an F-test to compare the mixed model to the single-parameter model with the least sum squared error (here, we used the *fcd* function from MATLAB to assign p-values), and we accepted the mixed model when $p < 0.05$. (iii) We used the Akaike's Information Criterion (AIC) to compare the best-fit single parameter model to each

25 alternative single parameter model, and we accepted the "best-fit" model when the difference in AIC (Δ_i) exceed 5 for all comparisons.⁷⁸ We note: For AD, AB, and (+)-1-(10),4-cadinadiene this criterion was not met; both noncompetitive and uncompetitive models, however, yielded indistinguishable IC_{50} 's.

We estimated the half maximal inhibitory concentration (IC_{50}) of inhibitors by using

30 the best-fit kinetic models to determine the concentration of inhibitor required to reduce initial rates of PTP-catalyzed hydrolysis of 15 mM of *p*NPP by 50%. We used the MATLAB function "nlparci" to determine the confidence intervals of kinetic parameters, and we propagated those intervals to estimate corresponding confidence intervals for each IC_{50} .

X-ray crystallography. We prepared crystals of PTP1B by using hanging drop vapor diffusion. In brief, we added 2 μ L of PTP1B (~600 μ M PTP1B, 50 mM HEPES, pH 7.3) to 6 μ L of crystallization solution (100 mM HEPES, 200 mM magnesium acetate, and 14% polyethylene glycol 8000, pH 7.5) and incubated the resulting droplets over crystallization solution for one week at 4°C (EasyXtal CrystalSupport, Qiagen). We soaked crystals with ligand by transferring them to droplets formed with 6 μ L of crystallization solution and 1 μ L of ligand solution (10 mM in DMSO), which we incubated for 2-5 days at 4°C. We prepared all ligands for freezing by soaking them in cryoprotectant formed from a 70/30 (v/v) mixture of buffer (100 mM HEPES, 200 mM magnesium acetate, and 25% polyethylene glycol 8000, pH 7.5) and glycerol.

We collected X-ray diffraction data through the Collaborative Crystallography Program at Lawrence Berkeley National Lab (ALS ENABLE, beamline 8.2.1, 100 K, 1.00003 Å). We performed integration, scaling, and merging of X-ray diffraction data using the xia2 software package⁷⁹, and we carried out molecular replacement and structure refinement with the PHENIX graphical interface,⁸⁰ supplemented with manual model adjustment in COOT⁸¹ and one round of PDB-REDO⁸² (the latter, only for the PTP1B-AD complex).

Molecular dynamics (MD) simulations. Full-length PTP1B contains a disordered region that extends beyond the α 7 helix (i.e., 299-435). In this study, we used a well-studied truncation variant (i.e., PTP1B₁₋₃₂₁) that includes residues from the disordered region. To model PTP1B, we used CAMPARI v.2⁸³ to generate structures of the disordered region of each complex (i.e., residues 288-321 for PTP1B-AD) from a crystal structure without a disordered tail. To quickly thermalize the tail structures, we ran short Monte Carlo (MC) simulations using the ABSINTH implicit-solvent force field^{84,85}, fixing the coordinates of the atoms in the ligand and the protein core.

We performed MD simulations using GROMACS 2020⁸⁶. Briefly, we used the CHARMM36m protein force field⁸⁷, a CHARMM-modified TIP3P water model⁸⁸, and ligand parameters generated by CGenFF^{89,90}. We solvated each PTP1B-ligand complex (initialized from the corresponding crystal structure) in a dodecahedral box with edges positioned ≥ 10 Å from the surface of the complex, and we added six sodium ions to neutralize each system. We used the LINCS algorithm⁹¹ to constrain all bonds involving hydrogen atoms, the Verlet

leapfrog algorithm to numerically integrate equations of motion with a 2-fs time step, and the particle-mesh Ewald summation⁹² (cubic interpolation with a grid spacing of 0.16 nm) to calculate long-range electrostatic interactions; we used a cutoff of 1.2 nm, in turn, for short-range electrostatic and Lennard-Jones interactions. We independently coupled the protein-ligand complex and solvent molecules to a temperature bath (300K) using a modified Berendsen thermostat⁹³ with a relaxation time of 0.1 ps, and we fixed pressure coupling to 1 bar using the Parrinello–Rahman algorithm⁹⁴ with a relaxation time of 2 ps and isothermal compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$.

For each system, we carried out 30 independent MD simulations to reduce sampling bias. For each MD trajectory, we minimized energy using the steepest decent method followed by 100-ps solvent relaxation in the NVT ensemble and 100-ps solvent relaxation in the NPT ensemble. After an additional 5-ns NPT equilibration, we carried out production runs for 5 ns in the NPT ensemble and registered coordinate data every 10 ps.

Analysis of PTP1B inhibition in HEK293T Cells . We prepared HEK293T/17 cells for an enzyme-linked immunosorbent assay (ELISA) by growing them in 75 cm² culture flasks (Corning) with DMEM media supplemented with 10% FBS, 100 units/ml penicillin, and 100 units/ml streptomycin. We replaced the media every day for 3-5 days until the cells reached 80-100% confluency.

We measured the influence of inhibitors on insulin receptor (IR) phosphorylation by using an IR-specific ELISA (**FIG. 35**). Briefly, we starved cells for 48 hours in FBS-free media and incubated the with inhibitors (all at 3% DMSO) for 10 minutes. After incubation, we lysed cells with lysis buffer (9803, Cell Signaling Technology) supplemented with 1X halt phosphatase inhibitor cocktail and 1X halt protease inhibitor cocktail (Thermo Fisher Scientific) for 10 min, pelleted the cell debris, and used the lysis buffer to dilute each sample to 60 mg/ml total protein. We measured IR phosphorylation in subsequent dilutions of the 60 mg/ml samples with the PathScan® Phospho-Insulin Receptor β (panTyr) Sandwich ELISA Kit (Cell Signaling Technology; #7082). We note: To identify biologically active concentrations of AB and AD, we screened several concentrations and chose those that gave the highest signal (405 μM for AB and 930 μM for AD); similar concentrations of weak inhibitors did not yield a detectable signal (**FIGs. 35b and 35c**).

Statistical analysis and reproducibility. We determined statistical significance (**FIG. 23h**) with a two-tailed Student's t-test (details in **TABLE 14**), and we used an F-test to compare one- and two-parameter models of inhibition (**TABLE 12**).

5 References for Example 2

1. Olsson, T. S. G., Williams, M. a., Pitt, W. R. & Ladbury, J. E. The Thermodynamics of Protein-Ligand Interaction and Solvation: Insights for Ligand Design. *J. Mol. Biol.* **384**, 1002–1017 (2008).
2. Fox, J. M., Zhao, M., Fink, M. J., Kang, K. & Whitesides, G. M. The Molecular Origin of Enthalpy/Entropy Compensation in Biomolecular Recognition. *Annu. Rev. Biophys.* **47**, (2018).
3. Mobley, D. L. & Gilson, M. K. Predicting Binding Free Energies: Frontiers and Benchmarks. *Annu. Rev. Biophys.* **46**, 531–558 (2017).
4. Hert, J., Irwin, J. J., Laggner, C., Keiser, M. J. & Shoichet, B. K. Quantifying biogenic bias in screening libraries. *Nat. Chem. Biol.* **5**, pages479–483 (2009).
5. Smanski, M. J. *et al.* Synthetic biology to access and expand nature's chemical diversity. *Nature Reviews Microbiology* **14**, 135–149 (2016).
6. Fürstenberg-Hägg, J., Zagrobelny, M. & Bak, S. Plant defense against insect herbivores. *Int. J. Mol. Sci.* **14**, 10242–10297 (2013).
7. Maier, M. E. Design and synthesis of analogues of natural products. *Organic and Biomolecular Chemistry* **13**, 5302–5343 (2015).
8. Chen, M. S. & White, M. C. A predictably selective aliphatic C-H oxidation reaction for complex molecule synthesis. *Science (80-.)*. **318**, 783–787 (2007).
9. Cho, I., Jia, Z. J. & Arnold, F. H. Site-selective enzymatic C-H amidation for synthesis of diverse lactams. *Science (80-.)*. **364**, 575–578 (2019).
10. Atanasov, A. G. *et al.* A Historical overview of natural products in drug discovery. *Metabolites* **33**, 1582–1614 (2012).
11. Paul, S. M. *et al.* How to improve RD productivity: The pharmaceutical industry's grand challenge. *Nature Reviews Drug Discovery* **9**, 203–214 (2010).
12. Li, J. W. H. & Vederas, J. C. Drug discovery and natural products: End of era or an endless frontier? *Biomeditsinskaya Khimiya* **57**, 148–160 (2011).
13. Jensen, P. R., Chavarria, K. L., Fenical, W., Moore, B. S. & Ziemert, N. Challenges and triumphs to genomics-based natural product discovery. *J. Ind. Microbiol.*

- Biotechnol.* **41**, 203–209 (2014).
14. Medema, M. H. *et al.* AntiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* **39**, W339–W346 (2011).
 - 5 15. Jensen, P. R. Natural Products and the Gene Cluster Revolution. *Trends Microbiol.* **24**, 968–977 (2016).
 16. Yan, Y. *et al.* Resistance-gene-directed discovery of a natural-product herbicide with a new mode of action. *Nature* **559**, 415–418 (2018).
 17. Culp, E. J. *et al.* Evolution-guided discovery of antibiotics that inhibit peptidoglycan
10 remodelling. *Nature* **578**, 582–587 (2020).
 18. Zhabinskii, V. N., Khripach, N. B. & Khripach, V. A. Steroid plant hormones: Effects outside plant kingdom. *Steroids* **97**, 87–97 (2015).
 19. Li, Y. *et al.* Complete biosynthesis of noscapine and halogenated alkaloids in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E3922–E3931 (2018).
 - 15 20. Luo, X. *et al.* Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature* **567**, 123–126 (2019).
 21. He, R., Yu, Z., Zhang, R. & Zhang, Z. Protein tyrosine phosphatases as potential therapeutic targets. *Acta Pharmacol. Sin.* **35**, 1227–1246 (2014).
 22. Paul, M. K. & Mukhopadhyay, A. K. Tyrosine kinase – Role and significance in
20 Cancer. *Int. J. Med. Sci.* **1**, 101–115 (2012).
 23. Ferguson, F. M. & Gray, N. S. Kinase inhibitors: The road ahead. *Nature Reviews Drug Discovery* **17**, 353–376 (2018).
 24. Stanford, S. M. & Bottini, N. Targeting Tyrosine Phosphatases: Time to End the Stigma. *Trends in Pharmacological Sciences* **38**, 524–540 (2017).
 - 25 25. Krishnan, N. *et al.* Targeting the disordered C terminus of PTP1B with an allosteric inhibitor. *Nat. Chem. Biol.* **10**, 558–566 (2014).
 26. Barr, A. J. *et al.* Large-Scale Structural Analysis of the Classical Human Protein Tyrosine Phosphatome. *Cell* **136**, 352–363 (2009).
 27. Banno, R. *et al.* PTP1B and SHP2 in POMC neurons reciprocally regulate energy
30 balance in mice. *J. Clin. Invest.* **120**, 720–734 (2010).
 28. Zabolotny, J. M. *et al.* Protein-tyrosine phosphatase 1B expression is induced by inflammation in vivo. *J. Biol. Chem.* **283**, 14230–14241 (2008).
 29. Matulka, K. *et al.* PTP1B is an effector of activin signaling and regulates neural

- specification of embryonic stem cells. *Cell Stem Cell* **13**, 706–719 (2013).
30. Zhang, H., Wang, Y., Wu, J., Skalina, K. & Pfeifer, B. A. Complete biosynthesis of erythromycin A and designed analogs using *E. coli* as a heterologous host. *Chem. Biol.* **17**, 1232–1240 (2010).
- 5 31. Antosch, J., Schaefer, F. & Gulder, T. A. M. Heterologous Reconstitution of Ikarugamycin Biosynthesis in *E. coli*. *Angew. Chemie Int. Ed.* **53**, 3011–3014 (2014).
32. Montalibet, J. & Kennedy, B. P. Using yeast to screen for inhibitors of protein tyrosine phosphatase 1B. *Biochem. Pharmacol.* **68**, 1807–1814 (2004).
33. Piserchio, A., Cowburn, D. & Ghose, R. Expression and purification of Src-family
10 kinases for solution NMR studies. *Methods Mol. Biol.* **831**, 111–131 (2012).
34. Badran, A. H. *et al.* Continuous evolution of *Bacillus thuringiensis* toxins overcomes insect resistance. *Nature* **533**, 58–63 (2016).
35. Kaneko, T. *et al.* Superbinder SH2 domains act as antagonists of cell signaling. *Sci. Signal.* **5**, ra68–ra68 (2012).
- 15 36. Jiang, C. S., Liang, L. F. & Guo, Y. W. Natural products possessing protein tyrosine phosphatase 1B (PTP1B) inhibitory activity found in the last decades. *Acta Pharmacologica Sinica* (2012). doi:10.1038/aps.2012.90
37. Hjortness, M. K. *et al.* Abietane-Type Diterpenoids Inhibit Protein Tyrosine Phosphatases by Stabilizing an Inactive Enzyme Conformation. *Biochemistry* **57**,
20 5886–5896 (2018).
38. Christianson, D. W. Structural and Chemical Biology of Terpenoid Cyclases. *Chem. Rev.* **106**, 3412–3442 (2006).
39. Newman, D. J. & Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014. *Journal of Natural Products* **79**, 629–661 (2016).
- 25 40. Martin, V. J. J., Pitera, D. J., Withers, S. T., Newman, J. D. & Keasling, J. D. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat. Biotechnol.* **21**, 796–802 (2003).
41. Morrone, D. *et al.* Increasing diterpene yield with a modular metabolic engineering system in *E. coli*: Comparison of MEV and MEP isoprenoid precursor pathway
30 engineering. *Appl. Microbiol. Biotechnol.* **85**, 1893–1906 (2010).
42. Williams, D. C. *et al.* Heterologous expression and characterization of a ‘pseudomature’ form of taxadiene synthase involved in paclitaxel (Taxol) biosynthesis and evaluation of a potential intermediate and inhibitors of the multistep diterpene

- cyclization reaction. *Arch. Biochem. Biophys.* (2000). doi:10.1006/abbi.2000.1865
43. Yoshikuni, Y., Ferrin, T. E. & Keasling, J. D. Designed divergent evolution of enzyme function. *Nature* **440**, 1078–1082 (2006).
 44. Peralta-Yahya, P. P. *et al.* Identification and microbial production of a terpene-based
5 advanced biofuel. *Nat. Commun.* (2011). doi:10.1038/ncomms1494
 45. Wiesmann, C. *et al.* Allosteric inhibition of protein tyrosine phosphatase 1B. *Nat. Struct. Mol. Biol.* **11**, 730–737 (2004).
 46. Zhang, C., Chen, X., Stephanopoulos, G. & Too, H. P. Efflux transporter engineering markedly improves AD production in *Escherichia coli*. *Biotechnol. Bioeng.* (2016).
10 doi:10.1002/bit.25943
 47. Keedy, D. A. *et al.* An expanded allosteric network in PTP1B by multitemperature crystallography, fragment screening, and covalent tethering. *Elife* **7**, doi: 10.7554/eLife.36307 (2018).
 48. Vallurupalli, P., Bouvignies, G. & Kay, L. E. Studying ‘invisible’ excited protein
15 states in slow exchange with a major state conformation. *J. Am. Chem. Soc.* **134**, 8148–8161 (2012).
 49. Amamuddy, O. S. *et al.* Integrated computational approaches and tools for allosteric drug discovery. *Int. J. Mol. Sci.* **21**, 847 (2020).
 50. Shimada, T. *et al.* Selectivity of Polycyclic Inhibitors for Human Cytochrome P450s
20 1A1, 1A2, and 1B1. *Chem. Res. Toxicol.* **11**, 1048–1056 (1998).
 51. Goldstein, B. J., Bittner-Kowalczyk, A., White, M. F. & Harbeck, M. Tyrosine dephosphorylation and deactivation of insulin receptor substrate-1 by protein-tyrosine phosphatase 1B. Possible facilitation by the formation of a ternary complex with the GRB2 adaptor protein. *J. Biol. Chem.* **275**, 4283–4289 (2000).
 - 25 52. Choi, E. *et al.* Mitotic regulators and the SHP2-MAPK pathway promote IR endocytosis and feedback regulation of insulin signaling. *Nat. Commun.* **10**, (2019).
 53. Dubois, M. J. *et al.* The SHP-1 protein tyrosine phosphatase negatively modulates glucose homeostasis. *Nat. Med.* **12**, 549–556 (2006).
 54. Hubert, J., Nuzillard, J. M. & Renault, J. H. Dereplication strategies in natural product
30 research: How many tools and methodologies behind the same concept?
Phytochemistry Reviews **16**, 55–95 (2017).
 55. Manguso, R. T. *et al.* In vivo CRISPR screening identifies Ptpn2 as a cancer immunotherapy target. *Nature* **547**, 413–418 (2017).

56. Varone, A., Spano, D. & Corda, D. Shp1 in Solid Cancers and Their Therapy. *Frontiers in Oncology* **10**, 935 (2020).
57. Yang, C. F. *et al.* Targeting protein tyrosine phosphatase PTP-PEST (PTPN12) for therapeutic intervention in acute myocardial infarction. *Cardiovasc. Res.* **116**, 1032–1046 (2020).
58. Zhang, S. & Zhang, Z. Y. PTP1B as a drug target: recent developments in PTP1B inhibitor discovery. *Drug Discov. Today* **12**, 373–381 (2007).
59. Oleinikovas, V., Saladino, G., Cossins, B. P. & Gervasio, F. L. Understanding Cryptic Pocket Formation in Protein Targets by Enhanced Sampling Simulations. *J. Am. Chem. Soc.* **138**, 14257–14263 (2016).
60. Rutledge, P. J. & Challis, G. L. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nature Reviews Microbiology* **13**, 509–523 (2015).
61. Hartenfeller, M. & Schneider, G. De Novo Drug Design. in *Chemoinformatics and Computational Chemical Biology* (ed. Bajorath, J.) 299–323 (Humana Press, 2011). doi:10.1007/978-1-60761-839-3_12
62. Packer, M. S. & Liu, D. R. Methods for the directed evolution of proteins. *Nat. Rev. Genet.* **16**, 379–394 (2015).
63. Johnston, C. W., Badran, A. H. & Collins, J. J. Continuous bioactivity-dependent evolution of an antibiotic biosynthetic pathway. *Nat. Commun.* **11**, 4202 (2020).
64. Chen, M. J., Dixon, J. E. & Manning, G. Genomics and evolution of protein phosphatases. *Sci. Signal.* **10**, 1–17 (2017).
65. Galanie, S. *et al.* Complete biosynthesis of opioids in yeast. *Science* (80-.). **349**, 1095–1100 (2015).
66. Paddon, C. J. & Keasling, J. D. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nat. Rev. Microbiol.* **12**, 355–367 (2014).
67. Grangeasse, C., Nessler, S. & Mijakovic, I. Bacterial tyrosine kinases: Evolution, biological function and structural insights. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 2640–2655 (2012).
68. Montalibet, J. *et al.* Residues distant from the active site influence protein-tyrosine phosphatase 1B inhibitor binding. *J. Biol. Chem.* **281**, 5258–5266 (2006).
69. Douzery, E. J. P., Snell, E. A., Baptiste, E., Delsuc, F. & Philippe, H. The timing of eukaryotic evolution: Does a relaxed molecular clock reconcile proteins and fossils?

- Proc. Natl. Acad. Sci. U. S. A.* **101**, 15386–15391 (2004).
70. O'Brien, K. P., Remm, M. & Sonnhammer, E. L. L. Inparanoid: A comprehensive database of eukaryotic orthologs. *Nucleic Acids Res.* **33**, D476–D480 (2005).
71. Kachroo, A. H. *et al.* Systematic humanization of yeast genes reveals conserved
5 functions and genetic modularity. *Science* (80-.). **348**, 921–925 (2015).
72. Carlson, J. C., Badran, A. H., Guggiana-Nilo, D. A. & Liu, D. R. Negative selection and stringency modulation in phage-assisted continuous evolution. *Nat. Chem. Biol.* **10**, 216–222 (2014).
73. Hjortness, M. K. *et al.* Evolutionarily Conserved Allosteric Communication in Protein
10 Tyrosine Phosphatases. *Biochemistry* **57**, 6443–6451 (2018).
74. Chen, X. *et al.* Statistical experimental design guided optimization of a one-pot biphasic multienzyme total synthesis of amorpho-4,11-diene. *PLoS One* **8**, e79650 (2013).
75. Edgar, S. *et al.* Mechanistic Insights into Taxadiene Epoxidation by Taxadiene-5 α -
15 Hydroxylase. *ACS Chem. Biol.* **11**, 460–469 (2016).
76. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490 (2010).
77. Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T. Y. ggtree: an r package for
20 visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* **8**, 28–36 (2017).
78. Burnham, K. P. & Anderson, D. R. *Model Selection and Multimodel Inference: a Practical Information-theoretic Approach*, 2nd edn. Springer-Verlag, New York. New York Springer **60**, (2002).
79. Winter, G. Xia2: An expert system for macromolecular crystallography data reduction.
25 *J. Appl. Crystallogr.* **43**, 186–190 (2010).
80. Afonine, P. V *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D. Biol. Crystallogr.* **68**, 352–67 (2012).
81. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
- 30 82. Joosten, R. P., Long, F., Murshudov, G. N. & Perrakis, A. The *PDB_REDO* server for macromolecular structure model optimization. *IUCrJ* **1**, 213–220 (2014).
83. Vitalis, A. & Pappu, R. V. Chapter 3 Methods for Monte Carlo Simulations of Biomacromolecules. in (ed. Wheeler, R. A. B. T.-A. R. in C. C.) **5**, 49–76 (Elsevier,

- 2009).
84. Vitalis, A. & Pappu, R. V. ABSINTH: A new continuum solvation model for simulations of polypeptides in aqueous solutions. *J. Comput. Chem.* **30**, 673–699 (2009).
- 5 85. Choi, J.-M. & Pappu, R. V. Improvements to the ABSINTH Force Field for Proteins Based on Experimentally Derived Amino Acid Specific Backbone Conformational Statistics. *J. Chem. Theory Comput.* **15**, 1367–1382 (2019).
86. Abraham, M. J. *et al.* Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19–25 (2015).
- 10 87. Huang, J. *et al.* CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2016).
88. MacKerell, A. D. *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).
89. Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671–690 (2010).
- 15 90. Yu, W., He, X., Vanommeslaeghe, K. & MacKerell, A. D. Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations. *J. Comput. Chem.* **33**, 2451–2468 (2012).
- 20 91. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
92. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089 (1993).
93. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
- 25 94. Parrinello, M. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182 (1981).
95. Li, H. *et al.* Crystal Structure and Substrate Specificity of PTPN12. *Cell Rep.* (2016). doi:10.1016/j.celrep.2016.04.016
- 30 96. Paling, N. R. D. & Welham, M. J. Role of the protein tyrosine phosphatase SHP-1 (Src homology phosphatase-1) in the regulation of interleukin-3-induced survival, proliferation and signalling. *Biochem. J.* **368**, 885–894 (2002).
97. Van Vliet, C. *et al.* Selective regulation of tumor necrosis factor-induced Erk signaling

by Src family kinases and the T cell protein tyrosine phosphatase. *Nat. Immunol.* **6**, 253–260 (2005).

TABLES

TABLE 1. Gene Sources

Component Src	Organism	Plasmid	Source
	<i>H. sapiens</i>	pDONR223_SR C_WT	Addgene: 82165
CDC37	<i>H. sapiens</i>	pBACgus4x/cdc 37/RocCOR LRRK2 1867- 2176	Addgene: 40398
PTP1B	<i>H. sapiens</i>	pGEX-2T PTP- 1B	Addgene: 8602
SHP2	<i>H. sapiens</i>	PTPN11	Addgene: 38965
TC-PTP	<i>H. sapiens</i>	pBG100-TCPTP	Addgene: 33365
LuxAB		pAB078d8	Addgene: 79206
RpoZ	<i>Escherichia coli</i>	pAB094a	Addgene: 79241
cl434	<i>Escherichia virus Lambda</i>	pAB078d8	Addgene: 79206
SH2	<i>Rous sarcoma virus</i>		Addgene: 78302
p130cas	<i>H. sapiens</i>	Synthetic	Integrated DNA Technologies, Inc.
midT	<i>H. sapiens</i>	Synthetic	Integrated DNA Technologies, Inc.
EGFR	<i>H. sapiens</i>	Synthetic	Integrated DNA Technologies, Inc.
ShcA	<i>H. sapiens</i>	Synthetic	Integrated DNA Technologies, Inc.
MBIS	<i>S. cerevisiae</i>	pMBIS	Addgene: 17817
ADS	<i>Artemisia annua</i>	pADS	Addgene: 19040
GHS	<i>Abies grandis</i>	pTrcHUM	Addgene: 19003
ABS	<i>Abies grandis</i>	pSBET/AgAs	Ruben Peters, Iowa State University
TXS	<i>Taxus brevifolia</i>	M60	David W. Christianson, University of Pennsylvania
GGPPS	<i>Taxus Canadensis</i>	gBlock	Integrated DNA Technologies, Inc.

TABLE 2. Plasmids

Plasmid	Description	Antibiotic*	Addgene
F-plasmid	The F-plasmid from the S1030 strain of <i>E. coli</i> .	T	105063
pB2H _{1b}	An early version of B2H that lacks PTP1B and contains LuxAB as the GOI.	K	TBD
pBAD _{1b.Src}	Enables inducible expression of Src and CDC37	P	TBD
pBAD _{1b.SH2}	Enables inducible expression of the SH2 domain.	P	TBD
pBAD _{1b.S}	Enables inducible expression of the substrate domain.	P	TBD
pBAD _{1b.All}	Enables inducible expression of Src, CDC37, the SH2 domain, and the substrate domain.	P	TBD
pB2H _{1c.p130cas}	An early version of B2H that (i) lacks PTP1B and Src, (ii) contains LuxAB, and (iii) includes a substrate from p130cas.	K	TBD
pB2H _{1c.midT}	An early version of B2H that (i) lacks PTP1B and Src, (ii) contains LuxAB, and (iii) includes a substrate from midT.	K	TBD
pB2H _{1c.ShcA}	An early version of B2H that (i) lacks PTP1B and Src, (ii) contains LuxAB, and (iii) includes a substrate from ShcA.	K	TBD
pB2H _{1c.EGFR}	An early version of B2H that (i) lacks PTP1B and Src, (ii) contains LuxAB, and (iii) includes a substrate from EGFR.	K	TBD
pBAD _{1c}	Enables inducible expression of Src and CDC37.	P	TBD
pBAD _{1d}	Enables inducible expression of Src and PTP1B.	P	TBD
pBAD _{1d.mut}	Enables inducible expression of Src and catalytically inactive PTP1B (C215S).	P	TBD
pB2H _{S1.1Pro1}	An early version of B2H that (i) lacks PTP1B, (ii) contains LuxAB, (iii) places expression of Src, CDC37, the SH2 domain, and the substrate domain under control of the same Pro1 promoter, and (iv) uses the BB034 RBS for Src.	K	TBD
pB2H _{S1.1Pro1.mut}	Identical to pB2H _{S1.1Pro1} except for a mutation in the substrate (Y4F)	K	TBD
pB2H _{S1.1ProD}	An early version of B2H that (i) lacks PTP1B, (ii) contains LuxAB, and (iii) includes the ProD promoter and pro RBS for Src.	K	TBD
pB2H _{S1.1ProD.mut}	Identical to pB2H _{S1.1ProD} except for a mutation in the substrate (Y4F)	K	TBD
pB2H _{S1.2pro}	An early version of B2H that (i) lacks PTP1B, (ii) contains LuxAB, and (iii) includes the pro RBS for Src.	K	TBD
pB2H _{S1.2pro.mut}	Identical to pB2H _{S1.2pro} except for a mutation in the substrate (Y4F)	K	TBD
pB2H _{S1.2Sal28}	An early version of B2H that (i) lacks PTP1B, (ii) contains LuxAB, and (iii) includes the Sal28 RBS for Src.	K	TBD
pB2H _{S1.2Sal28.mu}	Identical to pB2H _{S1.2Sal28} except for a mutation in the substrate (Y4F)	K	TBD
pB2H _{S1.3RBS30}	An early version of B2H that (i) contains LuxAB and (ii) includes the bb030 RBS for PTP1B.	K	TBD
pB2H _{S1.3RBS30}	Identical to pB2H _{S1.3RBS30} except for a mutation in the substrate (Y4F)	K	TBD
pB2H _{S1.3RBS34}	An early version of B2H that (i) contains LuxAB and (ii) includes the bb034 RBS for PTP1B.	K	TBD
pB2H _{S1.3RBS34}	Identical to pB2H _{S1.3RBS34} except for a mutation in the substrate (Y4F)	K	TBD
pB2H _{S2RBS30}	An early version of B2H that (i) contains SpecR and (ii) includes the bb030 RBS for PTP1B.	K	TBD
pB2H _{S2RBS30.mu}	Identical to pB2H _{S2RBS30} except for an inactivating mutation in PTP1B (C215S)	K	TBD
pB2H _{opt}	Final, optimized B2H that (i) contains SpecR and (ii) includes the bb034 RBS for PTP1B.	K	TBD
pB2H _{opt*}	Identical to pB2H _{opt} except for an inactivating mutation in PTP1B (C215S)	K	TBD

pB2H _{optX}		K	TBD
pMBIS	A plasmid that harbors genes for the mevalonate-dependent isoprenoid pathway from <i>S. cerevisiae</i> and harbors a tetracycline resistance marker.	T	17817
pMBIS _{CmR}	A plasmid that harbors genes for the mevalonate-dependent isoprenoid pathway from <i>S. cerevisiae</i> and harbors a chloramphenicol resistance marker.	P	TBD
pTrc99t	A pTrc99a variant with BsaI removed for use in Golden Gate cloning	C	TBD
pTS _{ADS}	A plasmid that harbors ADS.	C	TBD
pTS _{ADS(G349A)}	A plasmid that harbors ADS (G349A).	C	TBD
pTS _{ADS(G400C)}	A plasmid that harbors ADS (G400C).	C	TBD
pTS _{ADS(D299A)}	A plasmid that harbors ADS (D299A, inactivating).	C	TBD
pTS _{ADS(F514E)}	A plasmid that harbors ADS (F514E).	C	TBD
pTS _{ADS(G400L)}	A plasmid that harbors ADS (G400L).	C	TBD
pTS _{ADS(F514S)}	A plasmid that harbors ADS (F514S).	C	TBD
pTS _{ADS(F514V)}	A plasmid that harbors ADS (F514V).	C	TBD
pTS _{ADS(V292I)}	A plasmid that harbors ADS (V292I).	C	TBD
pTS _{ADS(I90S/F340S)}	A plasmid that harbors ADS (I90S/F340S).	C	TBD
pTS _{ADS(I490V/M528K)}	A plasmid that harbors ADS (I490V/M528K).	C	TBD
pTS _{ADS(G34S/K51N)}	A plasmid that harbors ADS (G34S/K51N).	C	TBD
pTS _{ADSF370Y}	A plasmid that harbors ADS (F370Y).	C	TBD
pTS _{ADSR527L}	A plasmid that harbors ADS (R527L).	C	TBD
pTS _{GHS}	A plasmid that harbors GHS.	C	TBD
pTS _{GHS(BFN)}	A plasmid that harbors GHS (W315P).	C	TBD
pTS _{GHS(SIB)}	A plasmid that harbors GHS (F312Q/M339A/M447F).	C	TBD
pTS _{GHS(HUM)}	A plasmid that harbors GHS (M339N/S484C/M565I).	C	TBD
pTS _{GHS(BBA)}	A plasmid that harbors GHS (A336V/M447H/I562T).	C	TBD
pTS _{GHS(ALP)}	A plasmid that harbors GHS (A336C/T445C/S484C/I562L/M565L).	C	TBD
pTS _{GHS(LFN)}	A plasmid that harbors GHS (A317N/A337S/S484C/I562V).	C	TBD
pTS _{GHS(A319Q)}	A plasmid that harbors GHS (A319Q).	C	TBD
pTS _{GHS(S561C)}	A plasmid that harbors GHS (S561C).	C	TBD
pTS _{GHS(Y415C)}	A plasmid that harbors GHS (Y415C).	C	TBD
pTS _{GHS(S484L)}	A plasmid that harbors GHS (S484L).	C	TBD
pTS _{GHS(L450Y)}	A plasmid that harbors GHS (L450Y).	C	TBD
pTS _{GHS(L450G)}	A plasmid that harbors GHS (L450G).	C	TBD
pTS _{GHS(L450K)}	A plasmid that harbors GHS (L450K).	C	TBD
pTS _{GHS(L450T)}	A plasmid that harbors GHS (L450T).	C	TBD
pTS _{GHS(T455I)}	A plasmid that harbors GHS (T455I).	C	TBD
pTS _{ABS}	A plasmid that harbors ABS and GGPPS.	C	TBD
pTS _{TXS}	A plasmid that harbors TXS and GGPPS.	C	TBD

*Antibiotic resistance: carbenicillin (C, 50 µg/ml), kanamycin (K, 50 µg/ml), tetracycline (T, 10 µg/ml), chloramphenicol (P, 34 µg/ml), and spectinomycin (S, conditional).

TABLE 3. Components of various B2H systems.

Component	Name	DNA SEQ ID NO:	DNA	Amino Acid SEQ ID NO:	Amino Acid
Kinase	c-Src	3	ATGGGGCTCCAAGCCGCAGACTCAGG GCCTGGCCAAGGATGCCTGGGAGAT CCCTCGGGAGTCGCTGCGGCTGGAG GTCAAGCTGGGCCAGGGCTGCTTTG GCGAGGTGTGGATGGGGACCTGGAA CGGTACCACCAGGGTGGCCATCAAA ACCCTGAAGCCTGGCACGATGTCTC CAGAGGCCTTCCTGCAGGAGGCCCA GGTCATGAAGAAGCTGAGGCATGAG AAGCTGGTGCAGTTGTATGCTGTGG TTTCAGAGGAGCCCATTTACATCGT CACGGAGTACATGAGCAAGGGGAG TTTGCTGGACTTTCTCAAGGGGGAG ACAGGCAAGTACCTGCGGCTGCCTC AGCTGGTGGACATGGCTGCTCAGAT CGCCTCAGGCATGGCGTACGTGGAG CGGATGAACTACGTCCACCGGGACC TTCGTGCAGCCAACATCCTGGTGGG AGAGAACCTGGTGTGCAAAGTGGCC GACTTTGGGCTGGCTCGGCTCATTG AAGACAATGAGTACACGGCGCGGC AAGGTGCCAAATTCCCCATCAAGTG GACGGCTCCAGAAGCTGCCCTCTAT GGCCGCTTCACCATCAAGTCGGACG TGTGGTCCTTCGGGATCCTGCTGACT GAGCTCACCACAAAGGGACGGGTGC CCTACCCTGGGATGGTGAACCGCGA GGTGCTGGACCAGGTGGAGCGGGGC TACCGGATGCCCTGCCGCCGAGT GTCCCGAGTCCCTGCACGACCTCAT GTGCCAGTGCTGGCGGAAGGAGCCT GAGGAGCGGCCACCTTCGAGTACC TGCAGGCCCTCCTGGAGGACTACTT CACGTCCACCGAGCCCCAGTACCAG CCCGGGGAGAACCTCTAA	21	MGSKPQTQGLAKDAWEIP RESLRLEVKLGGQCFGEV WMGTWNGTTRVAIKTLKP GTMSPEAFMQEAQVMKKL RHEKLVQLYAVVSEEPYIV TEYMSKGSLLDFLKGETGK YLRPLQLVDMAAQIASGM AYVERMNYVHRDLRAANI LVGENLVCKVADFGLARLI EDNEYTARQGAQFPIKWT PEAAALYGRFTIKSDVWSFGI LLELTTKGRVPYPGMVNR EVLDQVERGYRMPCEPCP ESLHDLMCQWRKEPEERP TFEYLAFLQEDYFTSTEPQY QPGENL*
Chaperone	CDC37	4	ATGGTGGACTACAGCGTGTGGGACC ACATTGAGGTGTCTGATGATGAAGA CGAGACGCACCCCAACATCGACACG GCCAGTCTCTCCGCTGGCGGCATC AGGCCCCGGTGGAAACGCATGGAGC AGTTCCAGAAGGAGAAGGAGGAAC TGGACAGGGGTGCCGCGAGTGCAA GCGCAAGGTGGCCGAGTGCCAGAG GAACTGAAGGAGCTGGAGGTGGC CGAGGGCGGCAAGGCAGAGCTGGA GCGCCTGCAGGCCGAGGCACAGCAG CTGCGCAAGGAGGAGCGGAGCTGG GAGCAGAAGCTGGAGGAGATGCGC AAGAAGGAGAAGAGCATGCCCTGG AACGTGGACACGCTCAGCAAAGACG GCTTCAGCAAGAGCATGGTAAATAC CAAGCCCAGAGAAGACGGAGGAGGA	22	MVDYSVWDHIEVSDDDEDE THPNIDTASLFRWRHQARV ERMEQFQKEKEELDRGCRE CKRKVAECQRKLKELEVA EGGKAELERLQAEAAQQLR KEERSWEQKLEEMRKKEK SMPWNVDTLSDGFSKSM VNTKPEKTEEDSEEVREQK HKTFVEKYEQIKHFGMLR RWDDSQKYLSDNVHLVCE ETANYLVIWCIDLEVEEK ALMEQVAHQITVMQFILEL AKSLKVDPRACFRQFFTKI KTADRQYMEGFNDELEAF KERVGRAKLRIEKAMKE YEEEEERKKRLGPGGLDVE VYESLPEELQKCFDVKDVQ

			<p>CTCAGAGGAGGTGAGGGAGCAGAA ACACAAGACCTTCGTGGAAAAATAC GAGAAACAGATCAAGCACTTTGGCA TGCTTCGCCGCTGGGATGACAGCCA AAAGTACCTGTCAGACAACGTCCAC CTGGTGTGCGAGGAGACAGCCAATT ACCTGGTCATTTGGTGCATTGACCTA GAGGTGGAGGAGAAATGTGCACTCA TGGAGCAGGTGGCCCACCAGACAAT CGTCATGCAATTTATCCTGGAGCTG GCCAAGAGCCTAAAGGTGGACCCCC GGGCCTGCTTCCGGCAGTTCTTCACT AAGATTAAAGACAGCCGATCGCCAGT ACATGGAGGGCTTCAACGACGAGCT GGAAGCCTTCAAGGAGCGTGTGCGG GGCCGTGCCAAGCTGCGCATCGAGA AGGCCATGAAGGAGTACGAGGAGG AGGAGCGCAAGAAGCGGCTCGGCC CCGGCGGCCTGGACCCCGTCGAGGT CTACGAGTCCCTCCCTGAGGAACTC CAGAAGTGCTTCGATGTGAAGGACG TGCAGATGCTGCAGGACGCCATCAG CAAGATGGACCCCACCGACGCAAAG TACCACATGCAGCGCTGCATTGACT CTGGCCTCTGGGTCCCCAACTCTAA GGCCAGCGAGGCCAAGGAGGGAGA GGAGGCAGGTCTTGGGGACCCATTA CTGGAAGCTGTTCCCAAGACGGGCG ATGAGAAGGATGTCAGTGTGTAA</p>	<p>MLQDAISKMDPTDAKYHM QRCIDSLWVPNSKASEAK EGEEAGPGDPLLEAVPKTG DEKDVS*</p>
Phospha- tase	PTP1B	5	<p>ATGGAGATGGAAAAGGAGTTCCGAG CAGATCGACAAGTCCGGGAGCTGGG CGGCCATTTACCAGGATATCCGACA TGAAGCCAGTGACTTCCCATGTAGA GTGGCCAAGCTTCCTAAGAACAAAA ACCGAAATAGGTACAGAGACGTCAG TCCCTTTGACCATAGTCGGATTA CTACATCAAGAAGATAATGACTATA TCAACGCTAGTTTGATAAAAAATGGA AGAAGCCCAAAGGAGTTACATTCTT ACCCAGGGCCCTTTGCCTAACACAT GCGGTCACTTTTGGGAGATGGTGTG GGAGCAGAAAAGCAGGGGTGTCGT CATGCTCAACAGAGTGATGGAGAAA GGTTTCGTTAAAAATGCGCACAACT GGCCACAAAAAGAAGAAAAAGAGA TGATCTTTGAAGACACAAATTTGAA ATTAACATTGATCTCTGAAGATATC AAGTCATATTATACAGTGCAGACAGC TAGAATTGGAAAACCTTACAACCCA AGAACTCGAGAGATCTTACATTTC CACTATAACCATGGCCCTGACTTTG GAGTCCCTGAATCACCAGCCTCATT CTTGAACCTTCTTTTCAAAGTCCGAG AGTCAGGGTCACTCAGCCCGGAGCA CGGGCCCGTTGTGGTGCAGTGCAGT GCAGGCATCGGCAGGTCTGGAACCT TCTGTCTGGCTGATACCTGCCTCTTG CTGATGGACAAGAGGAAAGACCCCTT</p>	<p>23 MEMEKEFEQIDKSGSWAAI YQDIRHEASDFPCRVAKLP KNKNRNRNRYRDVSPFDHSRI KLHQEDNDYINASLIKME AQRSYILTQGPLPNTCGHF WEMVWEQKSRGVVMLNR VMEKGSCLKAQYWPQKEE KEMIFEDTNLKLTLISED SYTTRQLELENLTQETR EILHFHYTTWPDFGVPE SFLNFLFKVRESGSLSP PVVVHCSAGIGRSGTF DTCLLLMDKRKDPSSVD KVLLEMRKFRMGLIQT QLRFSYLAVIEGAKFIM SSVQDQWKELSHEDLE EHIPPPRPPKRILEPHN*</p>

			CTTCCGTTGATATCAAGAAAGTGCT GTTAGAAATGAGGAAGTTTCGGATG GGGCTGATCCAGACAGCCGACCAGC TGCGCTTCTCCTACCTGGCTGTGATC GAAGGTGCCAAATTCATCATGGGGG ACTCTTCCGTGCAGGATCAGTGGAA GGAGCTTTCCACGAGGACCTGGAG CCCCACCCGAGCATATCCCCCAC CTCCCCGGCCACCCAAACGAATCCT GGAGCCACACAATTGA		
Substrate	p130cas	6	TGGATGGAGGACTATGACTACGTCC ACCTACAGGGG	24	WMEDYDYVHLQG
Substrate	midT	7	GAACCGCAGTATGAAGAAATTCCGA TTTATCTG	25	EPQYEEIPIYL
Substrate	ShcA	8	GATCATCAGTATTATAACGATTTTCC GGGC	26	DHQYYNDFPG
Substrate	EGFR	9	CCGCAGCGCTATCTGGTGATTACAGG GCGAT	27	PQRYLVIQGD
Substrate	p130cas Y/F	10	TGGATGGAGGACTTTGACTTCGTCC ACCTACAGGGG	28	WMEDFDFVHLQG
Substrate	midT Y/F	11	GAACCGCAGTTTGAAGAAATTCCGA TTTATCTG	29	EPQFEEIPIYL
Promoter	pBAD	12	AGAAACCAATTGTCCATATTGCATC AGACATTGCCGTCACTGCGTCTTTTA CTGGCTCTTCTCGCTAACCAAACCG GTAACCCCGCTTATTAAGCATTTC TGTAACAAAGCGGGACCAAAGCCAT GACAAAAACGCGTAACAAAAGTGTC TATAATCACGGCAGAAAAGTCCACA TTGATTATTTGCACGGCGTCACACTT TGCTATGCCATAGCATTTTATCCAT AAGATTAGCG	-	N/A
Promoter	Pro1 ⁵⁷	13	TTCTAGAGCACAGCTAACACCACGT CGTCCCTATCTGCTGCCCTAGGTCTA TGAGTGGTTGCTGGATAACTTTACG GGCATGCATAAGGCTCGGTATCTAT ATTCAGGGAGACCACAACGGTTTCC CTCTACAAATAATTTGTTTAACTTT TACTAGAG	-	N/A
Promoter	placZopt ³⁹	14	CATTAGGCACCCCGGGCTTTACTCG TAAAGCTTCCGGCGCGTATGTTGTG TCGACCG	-	N/A
Promoter	ProD ⁵⁷	13	TTCTAGAGCACAGCTAACACCACGT CGTCCCTATCTGCTGCCCTAGGTCTA TGAGTGGTTGCTGGATAACTTTACG GGCATGCATAAGGCTCGGTATCTAT ATTCAGGGAGACCACAACGGTTTCC CTCTACAAATAATTTGTTTAACTTT TACTAGAG	-	N/A
RBS	Pro	15	GTGCAGTTAAAGAGGAGAAAGGTC	-	N/A
RBS	Sal28 [‡]	16	CGAAAAAAGTAAGGCGGTAATCC	-	N/A
RBS	BB030	17	TCTAGAGATTAAAGAGGAGAAATAC TAG	-	N/A
RBS	BB034	18	TCTAGAAAAGAGGAGAAATACTAG	-	N/A
GOI	LuxAB	19	ATGAAATTTGGAACTTTTGTCTTAC ATACCAACCTCCCCAATTTCCCAA ACAGAGGTAATGAAACGTTTGTTA AATTAGGTCGCATCTCTGAGGAGTG	30	MKFGNFLTYPQFSQTE VMKRLVKLGRISEECGFDT VWLEHHFTEFGLGNPYV AAAYLLGATKKLVGTAA

TGGTTTTGATACCGTATGGTTACTGG
 AGCATCATTTACGGAGTTTGGTTTG
 CTGGTAACCCCTATGTCGCTGCTGC
 ATATTTACTTGGCGCGACTAAAAAA
 TTGAATGTAGGAAGTCCCGCTATTG
 TTCTTCCACAGCCCATCCAGTACGC
 CAACTTGAAGATGTGAATTTATTGG
 ATCAAATGTCAAAAGGACGATTTTCG
 GTTGGTATTTGCCGAGGGCTTTACA
 ACAAGGACTTTCGCGTATTCGGCAC
 AGATATGAATAACAGTCGCGCCTTA
 GCGGAATGCTGGTACGGGCTGATAA
 AGAATGGCATGACAGAGGGATATAT
 GGAAGCTGATAATGAACATATCAAG
 TTCCATAAGGTAAAAGTAAACCCCG
 CGGCGTATAGCAGAGGTGGCGCACC
 GGTTTATGTGGTGGCTGAATCAGCT
 TCGACGACTGAGTGGGCTGCTCAAT
 TTGGCCTACCGATGATATTAAGTTG
 GATTATAAACTAACGAAAAGAAA
 GCACAACTTGAGCTTTATAATGAAG
 TGGCTCAAGAATATGGGCACGATAT
 TCATAATATCGACCATTGCTTATCAT
 ATATAACATCTGTAGATCATGACTC
 AATTAAAGCGAAAGAGATTGCGCG
 AAATTTCTGGGGCATTGGTATGATT
 CTTATGTGAATGCTACGACTATTTTT
 GATGATTCAAGACCAAACAAGAGGTT
 ATGATTTCAATAAAGGGCAGTGGCG
 TGACTTTGTATTAAGGACATAAA
 GATACTAATCGCCGTATTGATTACA
 GTTACGAAATCAATCCCGTGGGAAC
 GCCGCAGGAATGTATTGACATAATT
 CAAAAAGACATTGATGCTACAGGAA
 TATCAAATATTTGTTGTGGATTTGAA
 GCTAATGGAACAGTAGACGAAATTA
 TTGCTTCCATGAAGCTCTTCCAGTCT
 GATGTCATGCCATTTCTTAAAGAAA
 AACAACGTTGCTATTATATTATGG
 CGGTGGCGGTAGCGGCGGTGGCGGT
 AGCGGCGGTGGCGGTAGCGGCGGTG
 GCGGTAGCAAATTTGGATTGTTCTTC
 CTTAACTTCATCAATTCAACAAGTGT
 TCAAGAACAGAGTATAGTTTCGCATG
 CAGGAAATAACGGAGTATGTTGATA
 AGTTGAATTTTGAACAGATTTTAGT
 GTATGAAAATCATTTTTCAGATAAT
 GGTGTGTGTCGGCGCTCCTCTGACTGT
 TTCTGGTTTTCTGCTCGGTTTAAACAG
 AGAAAATTAAAATTGGTTCATTAAA
 TCACATCATTACAACATCATCCTG
 TCCGCATAGCGGAGGAAGCTTGCTT
 ATTGGATCAGTTAAGTGAAGGGAGA
 TTTATTTTAGGGTTTAGTGATTGCGA
 AAAAAAGATGAAATGCATTTTTTT
 AATCGCCCGGTTGAATATCAACAGC
 AACTATTTGAAGAGTGTATGAAAT
 CATTAACGATGCTTTAACAACAGGC

IVLPTAHPVRQLEDVNLLD
 QMSKGRFRFGICRGLYNKD
 FRVFGTDMNNSRALAECW
 YGLIKNGMTEGYMEADNE
 HIKFHKVKVNPAAYSRGG
 APVYVVAESASTTEWAAQ
 FGLPMILSWIINTNEKKAQL
 ELYNEVAQEYGHDIHNIDH
 CLSYITSVDHDSIKAKEICR
 KFLGHWYDSYVNATTIFDD
 SDQTRGYDFNKGQWRDFV
 LKGHKDTNRRIDYSYEINP
 VGTPQECIDIIQKDIDATGIS
 NICCGFEANGTVDEIIASMK
 LFQSDVMPFLKEKQRSLLY
 YGGGGSGGGSGGGSGGG
 GGSKFGLFFLNFINSTTVQE
 QSIVRMQEITEYVDKLNFE
 QILVYENHFSDNGVVGAPL
 TVSGFLLGLTEKIKIGSLNHI
 ITTHHPVRIAEAECLLDQLS
 EGRFILGFSDCEKKDEMHEF
 FNRPVEYQQQLFEECYEIN
 DALTTGYCNPNDNFYSFPK
 ISVNPHAYTPGGPRKYVTA
 TSHHIVEWAACKGIPLIFK
 WDDSDNDVRYEYAERYKAV
 ADKYDVDLSEIDHQLMILV
 NYNEDSNKAKQETRAFISD
 YVLEMHPNENFENKLEEIIA
 ENAVGNYTECITAAKLAIE
 KCGAKSVLLSFEPMNDLMS
 QKNVINIVDDNIKKYHTEY
 T*

			TATTGTAATCCAGATAACGATTTTTTA TAGCTTCCCTAAAATATCTGTAAATC CCCATGCTTATACGCCAGGCGGACC TCGGAAATATGTAACAGCAACCAGT CATCATATTGTTGAGTGGGCGGCCA AAAAAGGTATTCCTCTCATCTTTAA GTGGGATGATTCTAATGATGTTAGA TATGAATATGCTGAAAGATATAAAG CCGTTGCGGATAAATATGACGTTGA CCTATCAGAGATAGACCATCAGTTA ATGATATTAGTTAACTATAACGAAG ATAGTAATAAAGCTAAACAAGAGAC GCGTGCAATTATTAGTGATTATGTTT TTGAAATGCACCCTAATGAAAATTT CGAAAATAAACTTGAAGAAATAATT GCAGAAAACGCTGTGCGAAATTATA CGGAGTGTATAACTGCGGCTAAGTT GGCAATTGAAAAGTGTGGTGCGAAA AGTGTATTGCTGTCTTTGAACCAAT GAATGATTTGATGAGCCAAAAAAT GTAATCAATATTGTTGATGATAATA TTAAGAAGTACCACACGGAATATAC CTAA	
GOI	SpecR	20	ATGAGGGAAGCGGTGATCGCCGAA GTATCGACTCAACTATCAGAGGTAG TTGGCGTCATCGAGCGCCATCTCGA ACCGACGTTGCTGGCCGTACATTG TACGGCTCCGCAGTGGATGGCGGCC TGAAGCCACACAGTGATATTGATTT GCTGGTTACGGTGACCGTAAGGCTT GATGAAACAACGCGCGAGCTTTGA TCAACGACCTTTTGAAAACCTTCGGC TTCCCCTGGAGAGAGCGAGATTCTC CGCGCTGTAGAAGTCACCATTGTTG TGCACGACGACATCATTCCGTGGCG TTATCCAGCTAAGCGCGAACTGCAA TTTGGAGAATGGCAGCGCAATGACA TTCTTGCAAGGTATCTTCGAGCCAGCC ACGATCGACATTGATCTGGCTATCTT GCTGACAAAAGCAAGAGAACATAG CGTTGCCCTTGGTAGGTCCAGCGCG GAGGAACTCTTTGATCCGGTTCCTG AACAGGATCTATTTGAGGCGCTAAA TGAAACCTTAACGCTATGGAACCTG CCGCCCGACTGGGCTGGCGATGAGC GAAATGTAGTGCTTACGTTGTCCCG CATTGGTACAGCGCAGTAACCGGC AAAATCGCGCCGAAGGATGTCGCTG CCGACTGGGCAATGGAGCGCCTGCC GGCCAGTATCAGCCCGTCATACTT GAAGCTAGACAGGCTTATCTTGGAC AAGAAGAAGATCGCTTGGCCTCGCG CGCAGATCAGTTGGAAGAATTTGTC CACTACGTGAAAGGCGAGATCACCA AGGTAGTCGGCAAATGA	31 MREAVIAEVSTQLSEVVGV IERHLEPTLLAVHLYGS DGLKPHSDIDLLVTVTR LDETTRRALINDLLET GESEILRAVEVTIVVHDDIIP WRYPKRELQFGEWQRND ILAGIFEPATIDIDLAILTK AREHSVALVGPAEELFDP VPEQDLFEALNETLTLWNS PPDWAGDERNVVLTLSRIW YSAVTGKIAPKDVAADWA MERLPAQYQPVILEARQAY LGQEEDRLASRADQLEEFV HYVKGEITKVVGK*

[‡]RBS designed computationally using the Ribosome Binding Site Calculator.⁵⁸

TABLE 4. Primers used to assemble the bacterial two-hybrid system.

Component	F Primer SEQ ID NO:	F Primer	R Primer SEQ ID NO:	R Primer
RpoZ/HA4 with pAB078d8 overhangs	32	GTGCAGTAAGGAGGAAAA AA	54	GTCAGGGGCGGGGTTTTTTTTT TAGGGCCCTACTGACTGTTA GCAGGTGCGGTAATTGA
pAB078d8 with RpoZ/HA4 overhang piece 1	33	CAGTCAGTAGGGCCCTAA AA	55	CACAGTTCTCGTCATCAGCTC TCTGGTTGCTTTAGCTAATAC ACCATAAGCATTTTCC
pAB078d8 with RpoZ/HA4 overhang piece 2	34	TAGCTAAAGCAACCAGAG AG	56	CAGTTACGCGTGCCATTTTTT TTCCCTCCTTACTGCACTTAG CGTTTCGGCGCCGGAT
Src/CDC37 into pAB078d8	35	CAATTCCCCTCTAGAAAT AATTTTG	57	GTCAGGGGCGGGGTTTTTTTTT TAGGGCCCTACTGACTG TTACACACTGACATCCTTCTC ATCG
Insulin Receptor Substrate_ RpoZ fusion into pAB078d8*	36	CGCTGTAGAGAAAATTGG TA	58	CAGGGGCGGGGTTTTTTTTTTA GGGCCCTACTGACTGTTATT AGCCAAGATCCATCTTCA
Insulin Receptor SH2_cI fusion into pAB078d8*	37	GACGCGGAATGGTACTGG GG	59	GTTACGCGTGCCATTTTTTTTT TCCTCCTTACTGCACTTATTA CGAAACCGGATACAACA
Src/CDC37 into pBAD33t	38	ATATGGTCTCACATGTCCA AGCCGCAGACTCAG	60	ATATGGTCTCATTTACACACT GACATCCTTCTCATCG
RpoZ/p130cas substrate into pBAD33t	39	ATATGGTCTCACATGGCA CGCGTAACTGTTC	61	ATATGGTCTCATTTACCCCTG TAGGTGGACG
cI/SH2 into pBAD33t	40	ATATGGTCTCACATGAGT ATCAGCAGCAGGGTAAAA AG	62	ATATGGTCTCATTTAGCAGA CGTTGGTCAGGC
pB2H _{1b} Gibson piece 1	41	ATGACTACGTCCACCTAC AGGGGTAATAACAATTCC CCTCTAGAAATAATTTTGT TTAAC	63	AAGATAAAAAGAATAGATCC CAGCCCTGTGTATAACTCAC TACTTTAGTCAGTTCCGCA
pB2H _{1b} Gibson piece 2	42	TGAGTTATACACAGGGCT GG	64	CCCCTGTAGGTGGACGTAGT CATAGTCCTCCATCCACGCA GCTGCACGACGA
pB2H _{1b} Gibson piece 3	43	GTGCAGTAAGGAGGAAAA AAAAATGGC	65	GCCCATGGTATATCTCCTTCT TAAAGT
pB2H _{1b} Gibson piece 4	44	TAAAATTCGTAGACTACA AGGACGACGATGACAAGT GGTATTTTGGGAAGATCA CTCGT	66	ACAGTTACGCGTGCCATTTTT TTTCCTCCTTACTGCACTTA GCAGACGTTGGTCAGGC

B2H ShcA substrate	45	TAATAACAATTCCCCTCTA GAAATAATTTTGTTTAACT TTAAG	67	GGGAATTGTTATTAGCCCGG AAAATCGTTATAATACTGAT GATCCGCAGCTGCACGACG
B2H EGFR Substrate	45	TAATAACAATTCCCCTCTA GAAATAATTTTGTTTAACT TTAAG	68	GGGAATTGTTATTAATCGCC CTGAATCACCAGATAGCGCT GCGGCGCAGCTGCACGACG
B2H MidT Substrate	45	TAATAACAATTCCCCTCTA GAAATAATTTTGTTTAACT TTAAG	69	GAATTGTTATTACAGATAAA TCGGAATTTCTTCATACTGCG GTTCCGCAGCTGCACGACG
BB034 PTP1B ₁₋₃₂₁ into pBAD _{1c}	46	GTCAGTGTGTAAGTGCAG AAAGAGGAGAAATACTAG ATGGAGATGGAAAAGGAG TTCGAG	70	CTCATCCGCCAAAACAGCCT CAATTGTGTGGCTCCAGGAT TCG
BB034 Src/CDC37	47	TAATCTAGAGAAAGAGGA GAAATACTAGATGTCCAA GCCGCAGACTC	71	TTACACACTGACATCCTTCTC ATCG
ProD into B2H	48	CTCTAGTAAAAGTTAAAC AAAATTATTTGTAGAGGG	72	TTCTAGAGCACAGCTAACAC CAC
ProD Overhang ProRBS Src/CDC37	49	AACTTTTACTAGAGGAAT TCGAGCTCTTAAAGAGGA GAAAGGTCATGGGCTCCA AGCCGC	63	AAGATAAAAAGAATAGATCC CAGCCCTGTGTATAACTCAC TACTTTAGTCAGTTCGCA
Sal28 RBS Src/CDC37	50	AACTTTTACTAGAG CGAAAAAAGTAAGGCGG TAATCCATGGGCTCCAAG CCGC	73	GAACCAATGAATGATTTGAT GAGC
BB030 PTP1B into pB2H _{S1.2Sal28}	51	AGTGTGTAAGTGCAGATT AAAGAGGAGAAATACTAG ATGGAGATGGAAAAGGAG TTCGAG	74	GTTTTTTTTTAGGGCCCTACT GACTGTCAATTGTGTGGCTC CAGGATTC
BB034 PTP1B into pB2H _{1.2Sal28}	52	TCAGTGTGTAAGTGCAGT CACACAGGAAAGTACTAG ATGGAGATGGAAAAGGAG TTCGAG	74	GTTTTTTTTTAGGGCCCTACT GACTGTCAATTGTGTGGCTC CAGGATTC
B2H Swap LuxAB/SpecR	53	GCGTACATTGGCTCCGTTC ATTTGCCGACTACCTTGGT GATC	75	GACCTGCAGATTAAAGAGGA GAAAATGAGGGAAGCGGTG ATCG

*Insulin receptor substrate/SH2 domains⁵⁹ were used initially, but failed to activate the operon (data not shown)

TABLE 5. Primers used to assemble pathways for terpenoid biosynthesis.

Component	F Primer SEQ ID NO:	F Primer	R Primer SEQ ID NO:	R Primer
GGPPS into pTrec99t	76	TATTGAGCTCCACCGCGG AGGAGGAATG	80	TATTGTCGACTTATTTATTAC GCTGGATGATGTAGTC
TXS into pTrec99t	77	TATTGGTCTCCCATGAGCA GCAGCACTGGCAC	81	TATTGGTCTCCGTCCTTCCAA CGCATTCAACATGTTG
ABS into pTrec99t	78	ATAAAGGTCTCCCATGGT GAAACGAGAATTCCTCC AG	82	TATTAGGTCTCGAGCTCTTA GGCAACTGGTTGGAAGAGGC
pMBIS TetR- >CmR	79	AGATCACTACCGGGCGTA TTTTTTGAGTTATCGAGAT TTTCAGGAGCTAAGGAAG CTAAAATGGAGAAAAAAA TCACTGGATATACCAC	83	GCCGCCGGCTTCCATTTATTA CGCCCCGCCCTG

Table 6. Primers used for site-directed mutagenesis.

Mutant	F Primer SEQ ID NO:	F Primer	R Primer SEQ ID NO:	R Primer
PTP1B (C215S)	84	GTCCAGTACTTTATTGGGGTT CAGGCGGATGGAAGTGAAGCA TGTCCGAGAT	107	ATCTCGGACATGCTCAGTTCCA TCCGCCTGAACCCCAATAAAGT ACTGGAC
ABS (D404A)	85	GAGAGAGAATCCTGTTCTCTG ATATTGGGATACAGCCATG GGCCTTC	108	GAAGGCCCATGGCTGTATCCG CAATATCAGGAACAGGATTCT CTCTC
ABS (D621A)	86	ACAAAAACTTCCAATTTTCAC TGTTATTTTATGGGATCTTTA TGACGCCCATGG	109	CCATGGGCGTCATAAAGATCC GCTAAAATAACAGTGAAATTG GAAGTTTTTGT
ADS (D299A)	87	CGTAAGCATCGTAAGTGTCC GCGATCAGGGTGATAACAGC	110	GCTGTTATCACCCTGATCGCGG ACACTTACGATGCTTACG
GHS (D343A)	88	CCCATGCGTGTCTGATAAGT CCGCTAACATTGTCATCAAG ATCG	111	CGATCTTGATGACAATGTTAGC GGACTTATACGACACGCATGG G
GHS (T455X)	89	CAATGGCACCCCCAACNNKG GTATGTGTGTACTTAATCTGA TCCCG	112	GTTGGGGGTGCCATTGTTC
GHS (L450X)	90	CAACACCGGTATGTGTGTAN NKAATCTGATCCCGTTGCTG CTTATG	113	TACACACATACCGGTGTTGGG
GHS (Y415X)	91	AAACGCTTGGGAACGCNNKC TGGAAGCGTATTTGCAGGAT G	114	GCGTTCCCAAGCGTTTTTG
GHS (A319X)	92	CTTCTGGATGGCCGCGNNKA TTTCAGAACCAAGATTTAGT GGCTC	115	CGCGGCCATCCAGAAGT
GHS (S484X)	93	ACCATCTGATTGAACTGGCT NNKCGACTGGTTCGATGATGC GAG	116	AGCCAGTTCAATCAGATGGTG G
GHS (S561X)	94	CGTCCTGGCGCGGNNKATTC AGTTTATGTATAACCAGGGG GAC	117	CCGCGCCAGGACGTG
ADS (F370X)	95	CAACTGCGGTAAAGAGTTTG TTAAAGAANNKGTACGTAAC CTGATGGTTGAAGC	118	TTCTTTAACAACCTCTTTACCG CAGTTG
ADS(G400 X)	96	CATGACCCGGTTGTTATCATC ACCNNKGGTGCAAACCTGCT GACCAC	119	GGTGATGATAACAACCGGGTC ATG
ADS (L405X)	97	CCGGCGGTGCAAACCTGNNK ACCACCACTTGCTATCTGGG	120	CAGGTTTGCACCGCCGG

ADS (G439X)	98	CTGTTCCGTTACTCCGGTATT CTGNNKCGTCGTCTGAACGA CCTGATG	121	CAGAATACCGGAGTAACGGAA CAG
ADS (F514X)	99	GGCAGTAATCTACCTGTGCC AGNNKCTGGAAGTACAGTAC GCTGGTAAAG	122	CTGGCACAGGTAGATTACTGCC
MidT Substrate (Y/F)	100	CAGCTGCGGAACCGCAGTTT GAAGAAATTCCGAT	123	ATCGGAATTTCTTCAAACCTGCG GTTCCGCAGCTG
p130Cas Substrate (Y/F)	101	TGGATGGAGGACTTTGACTT CGTCCACCTACAGGGGTAAT AACAATTC	124	GTCAAAGTCCTCCATCCACGCA GCTGCACGACG
SH2 (Superbind er mutations)	102	CTCTCCGTTTCTGACTTTGAC AACGCCAAGGGGCTCAATGT GCTGCACTACAAGATCCGCA AGCTG	125	AAGTCAGAAACGGAGAGGGCA TAGGCACCTTTTACCGTCTCGC TCTCCCG
SH2 (L13K K15L)*	103	AAACACTACCTGATCCGCAA GCTGGACAGC	126	GCTGTCCAGCTTGCGGATCAGG TAGTGTTTCACATTGAGCCCCT TGGC*
pTrc99a (remove BsaI sites) piece 1	104	TATTGGTCTCTCGCGGTATCA TTGCAGCAC	127	TATTGGTCTCAGTGACCCACACA CTACCATCGG
pTrc99a (remove BsaI sites) piece 2	105	TATTGGTCTCATCACCCCATG CGAGAGTAGG	128	TATTGGTCTCACGCGTGACCCA CGCTCACCG
ADS ep PCR	106	AACAATTTACACAGGAAAC AGACC	129	GCCTGCAGGTCGACTCTAGA

*The original superbinder primer mutated the incorrect lysine residue (13 vs. 15). This primer corrects that error. The residue numbering system used for this protein matches that of Kaneko et. al.⁴⁰

TABLE 7. Gene sources.

Component	Organism	Plasmid	Source
Src	<i>H. sapiens</i>	pDONR223_SRC_WT	Addgene: 82165
CDC37	<i>H. sapiens</i>	pBACgus4x/cdc37/RocCOR LRRK2 1867-2176	Addgene: 40398
PTP1B	<i>H. sapiens</i>	pET21B_PTP1B	Nicholas Tonks, Cold Spring Harbor
TC-PTP	<i>H. sapiens</i>	pBG100-TCPTP	Addgene: 33365
PTPN6	<i>H. sapiens</i>	pGEX-2T SHP1 WT	Addgene: 8594
PTPN12	<i>H. sapiens</i>	pDONR223_PTPN12_p.E57D	Addgene: 81528
LuxAB		pAB078d8	Addgene: 79206
RpoZ	<i>Escherichia coli</i>	pAB094a	Addgene: 79241
cI434	<i>Escherichia virus Lambda</i>	pAB078d8	Addgene: 79206
SH2	<i>Rous sarcoma virus</i>	Kras-SRC FRET Biosensor	Addgene: 78302
p130cas	<i>H. sapiens</i>	Synthetic	Integrated DNA Technologies, Inc.
midT	<i>H. sapiens</i>	Synthetic	Integrated DNA Technologies, Inc.
EGFR	<i>H. sapiens</i>	Synthetic	Integrated DNA Technologies, Inc.
ShcA	<i>H. sapiens</i>	Synthetic	Integrated DNA Technologies, Inc.
MBIS	<i>S. cerevisiae</i>	pMBIS	Addgene: 17817
ADS	<i>Artemisia annua</i>	pADS	Addgene: 19040
GHS	<i>Abies grandis</i>	pTrcHUM	Addgene: 19003
ABS	<i>Abies grandis</i>	pSBET/AgAs	Reuben Peters, Iowa State University
TXS	<i>Taxus brevifolia</i>	M60	David W. Christianson, University of Pennsylvania
ABA	<i>Abies grandis</i>	pTrc99a	Addgene: 35153
GGPPS	<i>Taxus canadensis</i>	gBlock	Integrated DNA Technologies, Inc.
A0A166A5J3	<i>S. Suecicum HHB10207 ss-3</i>	Synthetic	Twist Bioscience
A0A0D9X487	<i>L. perrieri</i>	Synthetic	Twist Bioscience
F2DRF1	<i>H. vulgare</i>	Synthetic	Twist Bioscience
A2XI80	<i>O. sativa</i>	Synthetic	Twist Bioscience
A0A0D9ZGD1	<i>O. glumipatula</i>	Synthetic	Twist Bioscience
A0A0K9RZT8	<i>S. oleracea</i>	Synthetic	Twist Bioscience
A0A1I1AC30	<i>A. aquimarinus</i>	Synthetic	Twist Bioscience
A0A1S3XW43	<i>N. tabacum</i>	Synthetic	Twist Bioscience
A0A0D3D8G7	<i>B. oleracea</i>	Synthetic	Twist Bioscience
B9IF04	<i>P. trichocarpa</i>	Synthetic	Twist Bioscience
A0A067L3D3	<i>J. curcas</i>	Synthetic	Twist Bioscience
A0A0C2TFL3	<i>A. Muscaria Koide BX008</i>	Synthetic	Twist Bioscience
A0A022S1C8	<i>E. guttata</i>	Synthetic	Twist Bioscience
G4TNA6	<i>S. indica</i>	Synthetic	Twist Bioscience
A0A1L7WMZ 8	<i>P. subalpine</i>	Synthetic	Twist Bioscience
A0A078IZJ5	<i>B. napus</i>	Synthetic	Twist Bioscience

A0A0C9VSL7	<i>S. stellatus SS14</i>	Synthetic	Twist Bioscience
G2QRS0	<i>T. terrestris ATCC 38088</i>	Synthetic	Twist Bioscience
A0A2H3DKU3	<i>A.gallica</i>	Synthetic	Twist Bioscience
A0A0D2L718	<i>H. sublateritium FD-334</i> <i>SS-4</i>	Synthetic	Twist Bioscience
S9Q922	<i>C. Fuscus DSM 2262</i>	Synthetic	Twist Bioscience
T1LTV1	<i>T. urartu</i>	Synthetic	Twist Bioscience
A0A287XU99	<i>H. vulgare</i>	Synthetic	Twist Bioscience
A0A0G2ZSL3	<i>A.gephyra</i>	Synthetic	Twist Bioscience

TABLE 8. Plasmids

Plasmid	Description	Antibiotic*	Availability
F-plasmid	The F-plasmid from the S1030 strain of <i>E. coli</i> .	T	AG: 105063 [‡]
pB2H_{1b}	An early version of B2H that lacks PTP1B and contains LuxAB as the GOI.	K	Fox Lab
pBAD_{1b,Src}	Enables inducible expression of Src and CDC37	P	Fox Lab
pBAD_{1b,SH2}	Enables inducible expression of the SH2 domain.	P	Fox Lab
pBAD_{1b,S}	Enables inducible expression of the substrate domain.	P	Fox Lab
pBAD_{1b,All}	Enables inducible expression of Src, CDC37, the SH2 domain, and the substrate domain.	P	Fox Lab
pB2H_{1c,p130cas}	An early version of B2H that (i) lacks PTP1B and Src, (ii) contains LuxAB, and (iii) includes a substrate from p130cas.	K	Fox Lab
pB2H_{1c,midT}	An early version of B2H that (i) lacks PTP1B and Src, (ii) contains LuxAB, and (iii) includes a substrate from midT.	K	Fox Lab
pB2H_{1c,ShcA}	An early version of B2H that (i) lacks PTP1B and Src, (ii) contains LuxAB, and (iii) includes a substrate from ShcA.	K	Fox Lab
pB2H_{1c,EGFR}	An early version of B2H that (i) lacks PTP1B and Src, (ii) contains LuxAB, and (iii) includes a substrate from EGFR.	K	Fox Lab
pBAD_{1d}	Enables inducible expression of Src and PTP1B.	P	Fox Lab
pBAD_{1d,mut}	Enables inducible expression of Src and catalytically inactive PTP1B (C215S).	P	Fox Lab
pB2H_{S1.1Pro1}	An early version of B2H that (i) lacks PTP1B, (ii) contains LuxAB, (iii) places expression of Src, CDC37, the SH2 domain, and the substrate domain under control of the same Pro1 promoter, and (iv) uses the BB034 RBS for Src.	K	Fox Lab
pB2H_{S1.1Pro1,mut}	Identical to pB2H _{S1.1Pro1} except for a mutation in the substrate (Y4F)	K	Fox Lab
pB2H_{S1.1ProD}	An early version of B2H that (i) lacks PTP1B, (ii) contains LuxAB, and (iii) includes the ProD promoter and pro RBS for Src.	K	Fox Lab
pB2H_{S1.1ProD,mut}	Identical to pB2H _{S1.1ProD} except for a mutation in the substrate (Y4F)	K	Fox Lab
pB2H_{S1.2pro}	An early version of B2H that (i) lacks PTP1B, (ii) contains LuxAB, and (iii) includes the pro RBS for Src.	K	Fox Lab
pB2H_{S1.2pro,mut}	Identical to pB2H _{S1.2pro} except for a mutation in the substrate (Y4F)	K	Fox Lab
pB2H_{S1.2Sal28}	An early version of B2H that (i) lacks PTP1B, (ii) contains LuxAB, and (iii) includes the Sal28 RBS for Src.	K	Fox Lab
pB2H_{S1.2Sal28,mut}	Identical to pB2H _{S1.2Sal28} except for a mutation in the substrate (Y4F)	K	Fox Lab
pB2H_{S1.3RBS30}	An early version of B2H that (i) contains LuxAB and (ii) includes the bb030 RBS for PTP1B.	K	Fox Lab
pB2H_{S1.3RBS30,mut}	Identical to pB2H _{S1.3RBS30} except for a mutation in the substrate (Y4F)	K	Fox Lab
pB2H_{S1.3RBS34}	An early version of B2H that (i) contains LuxAB and (ii) includes the bb034 RBS for PTP1B.	K	Fox Lab
pB2H_{S1.3RBS34,mut}	Identical to pB2H _{S1.3RBS34} except for a mutation in the substrate (Y4F)	K	Fox Lab
pB2H_{S2RBS30}	An early version of B2H that (i) contains SpecR and (ii) includes the bb030 RBS for PTP1B.	K	Fox Lab
pB2H_{S2RBS30,mut}	Identical to pB2H _{S2RBS30} except for an inactivating mutation in PTP1B (C215S)	K	Fox Lab
pB2H_{opt}	Final, optimized B2H that (i) contains SpecR and (ii) includes the bb034 RBS for PTP1B.	K	AG: 163830
pB2H_{opt*}	Identical to pB2H _{opt} except for an inactivating mutation in PTP1B (C215S)	K	AG: 163831

pB2H_{opt}X	Identical to pB2H _{opt} except for a mutation in the substrate domain (Y4F)	K	AG: 163832
pB2H₂	Identical to pB2H _{opt} with TC-PTP in place of PTP1B	K	AG: 163833
pB2H₂*	Identical to pB2H ₂ except for an inactivating mutation in TC-PTP (R222M)	K	AG: 163834
pB2H₆	Identical to pB2H _{opt} with SHP1 (catalytic domain) in place of PTP1B	K	AG: 163835
pB2H₆*	Identical to pB2H ₆ except for an inactivating mutation in SHP1 (R459M)	K	AG: 163836
pB2H₁₂	Identical to pB2H _{opt} with PTPN12 in place of PTP1B	K	AG: 163837
pB2H₁₂*	Identical to pB2H ₁₂ except for an inactivating mutation in PTPN12 (Y64A)	K	AG: 163838
pMBIS	A plasmid that harbors genes for the mevalonate-dependent isoprenoid pathway from <i>S. cerevisiae</i> and harbors a tetracycline resistance marker.	T	AG: 17817
pMBIS_{CmR}	A plasmid that harbors genes for the mevalonate-dependent isoprenoid pathway from <i>S. cerevisiae</i> and harbors a chloramphenicol resistance marker.	P	Fox Lab
pTrec99t	A pTrec99a variant with BsaI removed for use in Golden Gate cloning	C	Fox Lab
pTS_{ADS}	A plasmid that harbors ADS.	C	AG: 19040
pTS_{ADS(D299A)}	A plasmid that harbors ADS (D299A, inactivating).	C	Fox Lab
pTS_{GHS}	A plasmid that harbors GHS.	C	AG: 19003
pTS_{GHS(D343A)}	A plasmid that harbors GHS (D343A, inactivating).	C	Fox Lab
pTS_{ABA}	A plasmid that harbors ABA.	C	Fox Lab
pTS_{ABA(D566A)}	A plasmid that harbors ABA (D566A, inactivating).	C	Fox Lab
pTS_{ABS}	A plasmid that harbors ABS and GGPPS.	C	AG: 163840
pTS_{ABS(D404A/D621A)}	A plasmid that harbors ABS (D404A/D621A, inactivating) and GGPPS.	C	Fox Lab
pTS_{TXS}	A plasmid that harbors TXS and GGPPS.	C	AG: 163839
pTS_{A0A166A5J3}	A plasmid that harbors A0A166A5J3 (Clade 1)	C	Fox Lab
pTS_{A0A0D9X487}	A plasmid that harbors A0A0D9X487 (Clade 1)	C	Fox Lab
pTS_{F2DRF1}	A plasmid that harbors F2DRF1 (Clade 1)	C	Fox Lab
pTS_{A2XI80}	A plasmid that harbors A2XI80 (Clade 2)	C	Fox Lab
pTS_{A0A0D9ZGD1}	A plasmid that harbors A0A0D9ZGD1 (Clade 2)	C	Fox Lab
pTS_{A0A0K9RZT8}	A plasmid that harbors A0A0K9RZT8 (Clade 2)	C	Fox Lab
pTS_{A0A11IAC30}	A plasmid that harbors A0A11IAC30 (Clade 3)	C	Fox Lab
pTS_{A0A1S3XW43}	A plasmid that harbors A0A1S3XW43 (Clade 3)	C	Fox Lab
pTS_{A0A0D3D8G7}	A plasmid that harbors A0A0D3D8G7 (Clade 3)	C	Fox Lab
pTS_{B9IF04}	A plasmid that harbors B9IF04 (Clade 4)	C	Fox Lab
pTS_{A0A067L3D3}	A plasmid that harbors A0A067L3D3 (Clade 4)	C	Fox Lab
pTS_{A0A0C2TFL3}	A plasmid that harbors A0A0C2TFL3 (Clade 4)	C	Fox Lab
pTS_{A0A022S1C8}	A plasmid that harbors A0A022S1C8 (Clade 5)	C	Fox Lab
pTS_{G4TNA6}	A plasmid that harbors G4TNA6 (Clade 5)	C	Fox Lab
pTS_{A0A1L7WMZ8}	A plasmid that harbors A0A1L7WMZ8 (Clade 5)	C	Fox Lab
pTS_{A0A078IZJ5}	A plasmid that harbors A0A078IZJ5 (Clade 6)	C	Fox Lab
pTS_{A0A0C9VSL7}	A plasmid that harbors A0A0C9VSL7 (Clade 6)	C	AG: 163841
pTS_{G2QRS0}	A plasmid that harbors G2QRS0 (Clade 6)	C	Fox Lab
pTS_{A0A2H3DKU3}	A plasmid that harbors A0A2H3DKU3 (Clade 7)	C	Fox Lab
pTS_{A0A0D2L718}	A plasmid that harbors A0A0D2L718 (Clade 7)	C	Fox Lab

pTS_{S9Q922}	A plasmid that harbors S9Q922 (Clade 7)	C	Fox Lab
pTS_{T1LTV1}	A plasmid that harbors T1LTV1 (Clade 8)	C	Fox Lab
pTS_{A0A287XU99}	A plasmid that harbors A0A287XU99 (Clade 8)	C	Fox Lab
pTS_{A0A0G2ZSL3}	A plasmid that harbors A0A0G2ZSL3 (Clade 8)	C	Fox Lab
pET21b_{ptp1b}	A plasmid that encodes a His-tagged catalytic domain of PTP1B (for protein expression)	C	N/A ⁺
pET16B_{TCPT}	A plasmid that encodes a His-tagged catalytic domain of TCPTP (for protein expression)	C	Fox Lab

* Antibiotic resistance: carbenicillin (C, 50 µg/ml), kanamycin (K, 50 µg/ml), tetracycline (T, 10 µg/ml), chloramphenicol (P, 34 µg/ml), and spectinomycin (S, conditional).

⁺This plasmid was a kind gift from Nicholas Tonks of Cold Spring Harbor Laboratory.

*AG=Addgene accession # (Addgene.com).

TABLE 9. Primers used to assemble pathways for terpenoid biosynthesis.

Component	F Primer SEQ ID NO:	F Primer	R Primer SEQ ID NO:	R Primer
GGPPS into pTrc99t	76	TATTGAGCTCCACCGCGG AGGAGGAATG	80	TATTGTCGACTTATTTATTAC GCTGGATGATGTAGTC
TXS into pTrc99t	77	TATTGGTCTCCCATGAGCA GCAGCACTGGCAC	81	TATTGGTCTCCGTCCTTCCAA CGCATTC AACATGTTG
ABS into pTrc99t	78	ATAAAGGTCTCCCATGGT GAAACGAGAATTTCTCCTC AG	82	TATTAGGTCTCGAGCTCTTA GGCAACTGGTTGGAAGAGGC
pMBIS TetR- >CmR	79	AGATCACTACCGGGCGTA TTTTTTGAGTTATCGAGAT TTTCAGGAGCTAAGGAAG CTAAAATGGAGAAAAAAA TCACTGGATATACCAC	83	GCCGCCGGCTTCCATTTATTA CGCCCCGCCCTG
ABA into pTrc99	130	AACAATTTACACAGGAA ACAGACCATGGCGGGTGT TTCTGCG	131	GCCTGCAGGTCGACTCTAGA TTACAGCGGCAGCGGTTTC

TABLE 10. Primers used for site-directed mutagenesis.

Mutant	F Primer SEQ ID NO:	F Primer	R Primer SEQ ID NO:	R Primer
PTP1B (C215S)	84	GTCCAGTACTTTATTGGGGTT CAGGCGGATGGAAGTGAAGCA TGTCCGAGAT	107	ATCTCGGACATGCTCAGTTCCA TCCGCCTGAACCCCAATAAAGT ACTGGAC
TCPTP (R222M)	132	CAGAGAGAAGGTGCCAGACA TCCCAATGCCTGCACTACA	136	TGTAGTGCAGGCATTGGGATGT CTGGCACCTTCTCTCTG
SHP1 (R459M)	133	CAATGATGGTGCCTGTCAATG CCGATGCCGGCGCTG	137	CAGCGCCGGCATCGGCATGAC AGGCACCATCATTTG
PTPN12 (Y64A)	134	GCTGTGATCAAATGGCAGTA TGTCCTTCGCTCTGTTCTTTT AACATTTTCTTCTTTTC	138	GAAAAAGAAGAAAATGTTAAA AAGAACAGAGCGAAGGACATA CTGCCATTGTATCACAGC
ABS (D404A)	85	GAGAGAGAATCCTGTTTCCTG ATATTGCGGATACAGCCATG GGCCTTC	108	GAAGGCCCATGGCTGTATCCG CAATATCAGGAACAGGATTCT CTCTC
ABS (D621A)	86	ACAAAAACTTCCAATTTTAC TGTTATTTTAGCGGATCTTTA TGACGCCCATGG	109	CCATGGGCGTCATAAAGATCC GCTAAAATAACAGTGAAATTG GAAGTTTTTGT
ADS (D299A)	87	CGTAAGCATCGTAAGTGTCC GCGATCAGGGTGATAACAGC	110	GCTGTTATCACCTGTATCGCGG ACACTTACGATGCTTACG
GHS (D343A)	88	CCCATGCGTGTCTGATAAGT CCGCTAACATTGTCATCAAG ATCG	111	CGATCTTGATGACAATGTTAGC GGACTTATACGACACGCATGG G
MidT Substrate (Y/F)	100	CAGCTGCGGAACCGCAGTTT GAAGAAATTCGGAT	123	ATCGGAATTTCTTCAAATGCG GTTCCGCAGCTG
p130Cas Substrate (Y/F)	101	TGGATGGAGGACTTTGACTT CGTCCACCTACAGGGGTAAT AACAATTC	124	GTCAAAGTCTCCATCCACGCA GCTGCACGACG
SH2 (Superbinder mutations)	102	CTCTCCGTTTCTGACTTTGAC AACGCCAAGGGGCTCAATGT GCTGCACTACAAGATCCGCA AGCTG	125	AAGTCAGAAACGGAGAGGGCA TAGGCACCTTTTACCGTCTCGC TCTCCCG
SH2 (L13K K15L)*	103	AAACACTACCTGATCCGCAA GCTGGACAGC	126	GCTGTCCAGCTTGCGGATCAGG TAGTGTTCACATTGAGCCCCCT TGGC*
pTrc99a (remove BsaI sites) piece 1	104	TATTGGTCTCTCGGGTATCA TTGCAGCAC	127	TATTGGTCTCAGTGACCCCA CTACCATCGG
pTrc99a (remove BsaI sites) piece 2	105	TATTGGTCTCATCACCCCATG CGAGAGTAGG	128	TATTGGTCTCACGCGTGACCCA CGCTCACCG
ABA D/A	135	AGGTGTCTGATCATGTCCGCC AGAACGGTCTGCAG	139	CTGCAGACCGTTCTGGCGGAC ATGTACGACACCT

*The original superbinder primer mutated the incorrect lysine residue (13 vs. 15). This primer corrects that error.
 5 The residue numbering system used for this protein matches that of Kaneko et. al.²⁰

TABLE 11a. Scaling factor for amorphadiene/caryophyllene (m/z=204)

Technical Replicate	A _{std} (counts*min)	A _{ref} (counts*min)	C _{std} (µg/mL)	C _{ref} (µg/mL)	R
1	74520	88358	20	0.4	0.017
2	71037	142415	20	0.4	0.010
3	75761	49011	20	0.4	0.031
				Avg R	0.019 (0.006)

* R was computed using eq. 2. Standard error is shown in parentheses.

5

TABLE 11b. Scaling factor for taxadiene/caryophyllene (m/z=93)

Technical Replicate	A _{std} (counts*min)	A _{ref} (counts*min)	C _{std} (µg/mL)	C _{ref} (µg/mL)	R
1	1399872	847009	20	10	0.83
2	1247250	605265	20	10	1.0
3	1291028	547740	20	10	1.2
				Avg R	1.0 (0.10)

10

TABLE 11c. Scaling factor for amorphadiene/methyl abietate (m/z=121)

Technical Replicate	A _{std} (counts*min)	A _{ref} (counts*min)	C _{std} (µg/mL)	C _{ref} (µg/mL)	R
1	949492	868168	20	3.162	0.17
2	920694	908257	20	3.162	0.16
3	898594	1106474	20	3.162	0.13
				Avg R	0.15 (0.01)

TABLE 12a. Analysis of the inhibition of PTP1B₁₋₃₂₁ by amorphadiene.

<i>Model</i>	<i>SSE ($\mu M^2/s^2$)</i>	<i>DF</i>	<i>Criteria</i>	<i>Reference</i>	<i>Fit par. (μM)</i>
Competitive	0.14	27	$\Delta_i = 51.2$	noncompetitive	$K_i = 2.85$
Uncompetitive**	0.023	27	$\Delta_i = 1.16$	noncompetitive	$K_i = 46.3$
Noncompetitive**	0.023	27			$K_i = 52.6^*$
Mixed	0.022	26	$F = 0.47$ $p = 0.972$	noncompetitive	$K_{i,c} = 86.2$ $K_{i,u} = 50.1$

*The SSEs of the uncompetitive and noncompetitive models are indistinguishable from one another.

**Indicate models of best fit.

5

TABLE 12b. Analysis of the inhibition of PTP1B₁₋₃₂₁ by α -bisabolene.

<i>Model</i>	<i>SSE ($\mu M^2/s^2$)</i>	<i>DF</i>	<i>Criteria</i>	<i>Reference</i>	<i>Fit par. (μM)</i>
Competitive	0.082	27	$\Delta_i = 39.1$	noncompetitive	$K_i = 1.05$
Uncompetitive**	0.023	27	$\Delta_i = 3.81$	noncompetitive	$K_i = 11.7$
Noncompetitive**	0.021	27			$K_i = 13.1$
Mixed	0.020	26	$F = 0.24$ $p = 1.0$	noncompetitive	$K_{i,c} = 9.51$ $K_{i,u} = 13.7$

10

TABLE 12c. Analysis of the inhibition of PTP1B₁₋₃₂₁ by alpha bisabolol.

<i>Model</i>	<i>SSE ($\mu M^2/s^2$)</i>	<i>DF</i>	<i>Criteria</i>	<i>Reference</i>	<i>Fit par. (μM)</i>
Competitive	0.039	27	$\Delta_i = 34.4$	uncompetitive	$K_i = 178$
Uncompetitive**	0.011	27			$K_i = 469$
Noncompetitive**	0.013	27	$\Delta_i = 4.65$	uncompetitive	$K_i = 541$
Mixed	0.011	26	$F = 0$ $p = 1.0$	uncompetitive	$K_{i,c} = 3.5e^{16}$ $K_{i,u} = 469$

TABLE 12d. Analysis of the inhibition of PTP1B₁₋₃₂₁ by dihydroartemesnic acid.

<i>Model</i>	<i>SSE ($\mu M^2/s^2$)</i>	<i>DF</i>	<i>Criteria</i>	<i>Reference</i>	<i>Fit par. (μM)</i>
Competitive	0.129	27	$\Delta_i = 60.7$	noncompetitive	$K_i = 178$
Uncompetitive	0.025	27	$\Delta_i = 15.2$	noncompetitive	$K_i = 469$
Noncompetitive	0.015	27			$K_i = 541$
Mixed**	0.013	26	$F = 2.69$ $p = 6.9e^{-3}$	noncompetitive	$K_{i,c} = 3.5e^{16}$ $K_{i,u} = 469$

TABLE 12e. Analysis of the inhibition of TCPTP₁₋₃₁₇ by amorphadiene.

<i>Model</i>	<i>SSE ($\mu M^2/s^2$)</i>	<i>DF</i>	<i>Criteria</i>	<i>Reference</i>	<i>Fit par. (μM)</i>
Competitive	0.053	27	$\Delta_i = 41.1$	uncompetitive	$K_i = 87.2$
Uncompetitive**	0.012	27			$K_i = 356$
Noncompetitive**	0.013	27	$\Delta_i = 2.22$	uncompetitive	$K_i = 400$
Mixed	0.012	26	$F = 0$ $p = 1.0$	uncompetitive	$K_{i,c} = 3.7e^{15}$ $K_{i,u} = 356$

*The SSEs of the uncompetitive and noncompetitive models are indistinguishable from one another.

**Indicate models of best fit.

20

TABLE 12f. Analysis of the inhibition of TCPTP₁₋₃₁₇ by α -bisabolene.

<i>Model</i>	<i>SSE ($\mu\text{M}^2/\text{s}^2$)</i>	<i>DF</i>	<i>Criteria</i>	<i>Reference</i>	<i>Fit par. (μM)</i>
Competitive	0.046	27	$\Delta_i = 37.6$	uncompetitive	$K_i = 13.7$
Uncompetitive**	0.012	27			$K_i = 69.2$
Noncompetitive**	0.012	27	$\Delta_i = 1.12$	uncompetitive	$K_i = 76.2$
Mixed	0.012	26	F=0 p = 1.0	uncompetitive	$K_{i,c} = 3610$ $K_{i,u} = 69.3$

5 **TABLE 12g.** Analysis of the inhibition of PTP1B₁₋₂₈₁ by amorphadiene.

<i>Model</i>	<i>SSE ($\mu\text{M}^2/\text{s}^2$)</i>	<i>DF</i>	<i>Criteria</i>	<i>Reference</i>	<i>Fit par. (μM)</i>
Competitive	0.010	27	$\Delta_i = 16.3$	noncompetitive	$K_i = 37.9$
Uncompetitive**	0.006	27	$\Delta_i = 3.51$	noncompetitive	$K_i = 210$
Noncompetitive**	0.006	27			$K_i = 244$
Mixed	0.006	26	F=0.41 p = 0.99	noncompetitive	$K_{i,c} = 157$ $K_{i,u} = 271$

TABLE 12h. Analysis of the inhibition of PTP1B₁₋₂₈₁ by α -bisabolene.

<i>Model</i>	<i>SSE ($\mu\text{M}^2/\text{s}^2$)</i>	<i>DF</i>	<i>Criteria</i>	<i>Reference</i>	<i>Fit par. (μM)</i>
Competitive	0.012	27	$\Delta_i = 14.4$	noncompetitive	$K_i = 6.51$
Uncompetitive**	0.008	27	$\Delta_i = 1.41$	noncompetitive	$K_i = 40.0$
Noncompetitive**	0.007	27			$K_i = 46.3$
Mixed	0.007	26	F=0 p = 1.0	noncompetitive	$K_{i,c} = 39.0$ $K_{i,u} = 47.7$

10 **TABLE 12i.** Analysis of the inhibition of TCPTP₁₋₂₈₁ by amorphadiene.

<i>Model</i>	<i>SSE ($\mu\text{M}^2/\text{s}^2$)</i>	<i>DF</i>	<i>Criteria</i>	<i>Reference</i>	<i>Fit par. (μM)</i>
Competitive	0.005	27	$\Delta_i = 22.9$	uncompetitive	$K_i = 87.2$
Uncompetitive**	0.002	27			$K_i = 356$
Noncompetitive**	0.002	27	$\Delta_i = 0.83$	uncompetitive	$K_i = 400$
Mixed	0.002	26	F=0.03 p = 1.0	uncompetitive	$K_{i,c} = 3.7e^{15}$ $K_{i,u} = 356$

TABLE 12j. Analysis of the inhibition of TCPTP₁₋₂₈₁ by α -bisabolene.

<i>Model</i>	<i>SSE ($\mu\text{M}^2/\text{s}^2$)</i>	<i>DF</i>	<i>Criteria</i>	<i>Reference</i>	<i>Fit par. (μM)</i>
Competitive	0.083	27	$\Delta_i = 39.1$	noncompetitive	$K_i = 13.7$
Uncompetitive**	0.023	27	$\Delta_i = 3.81$	noncompetitive	$K_i = 69.2$
Noncompetitive**	0.021	27			$K_i = 76.2$
Mixed	0.020	26	F=0 p = 1.0	noncompetitive	$K_{i,c} = 3610$ $K_{i,u} = 69.3$

15

TABLE 12k. Analysis of the inhibition of PTP1B₁₋₃₂₁ by (+)-1-(10),4-cadinadiene

<i>Model</i>	<i>SSE ($\mu\text{M}^2/\text{s}^2$)</i>	<i>DF</i>	<i>Criteria</i>	<i>Reference</i>	<i>Fit par. (μM)</i>
Competitive	0.115	27	$\Delta_i = 48.3$	uncompetitive	$K_i = 14.75$
Uncompetitive**	0.020	27			$K_i = 168.09$
Noncompetitive**	0.022	27	$\Delta_i = 2.5$	uncompetitive	$K_i = 190.44$
Mixed	0.020	26	F=0 p = 1.0	uncompetitive	$K_{i,c} = 5689.38$ $K_{i,u} = 168.78$

TABLE 12I. Analysis of the inhibition of SHP2₂₂₃₋₅₆₅ by Amorphadiene

<i>Model</i>	<i>SSE ($\mu\text{M}^2/\text{s}^2$)</i>	<i>DF</i>	<i>Criteria</i>	<i>Reference</i>	<i>Fit par. (μM)</i>
Competitive	.0024	27	$\Delta_i = 10.6$	noncompetitive	$K_i = 25.1$
Uncompetitive**	.0017	27	$\Delta_i = 0.5$	noncompetitive	$K_i = 116.51$
Noncompetitive**	.0017	27			$K_i = 145.69$
Mixed	0.0017	26	F=0.15 p = 1.0	noncompetitive	$K_{i,c} = 236.21$ $K_{i,u} = 132.37$

TABLE 13. Data collection and refinement statistics (molecular replacement)

	PTP1B:amorphadiene (6W30)	PTP1B:α-bisabolol (N/A***)
Data collection		
Space group		
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	89.03, 89.03, 105.56	89.28, 89.28, 105.51
α, β, γ (°)	90.00, 90.00, 120.00	90.00, 90.00, 120.00
Resolution (Å)	62.26-2.10 (2.13-2.10)*	77.32-2.11 (2.15-2.11)
<i>R</i> _{sym} or <i>R</i> _{merge}	0.130 (0.442)	0.086 (0.331)
<i>I</i> / σ <i>I</i>	5.4 (1.0)	6.7 (1.1)
Completeness (%)	99.8 (93.3)	100.0 (98.5)
Redundancy	10.7 (10.8)	12.1 (12.3)
Refinement		
Resolution (Å)	44.52-2.10 (2.17-2.10)	62.37-2.11 (2.18-2.11)
No. reflections	28,654	28,479
<i>R</i> _{work} / <i>R</i> _{free}	0.20 / 0.24	0.19 / 0.24
No. atoms		
Protein	2355	2320
Ligand/ion	22	17
Water	170	270
<i>B</i> -factors		
Protein	37	30
Ligand/ion	90/61	66/37
Water	47	43
R.m.s. deviations		
Bond lengths (Å)	0.42	0.42
Bond angles (°)	0.56	0.54

*Values in parentheses correspond to the highest-resolution shell.

**Number of crystals used for each structure: 1

***In light of the results detailed in Figure 31, we elected not to deposit this structure into the protein data bank.

TABLE 14. Details of hypothesis testing

<i>Figure</i>	<i>Null hypothesis</i>	$\Delta\mu$	<i>Test</i>	<i>DF</i>	<i>t</i>	<i>95% confidence intervals</i>	<i>P-value</i>
3h	AD - (-) = 0	0.212	t-test, unequal variance	2	6.61	(0.092,0.332)	0.02
3h	AB - (-) = 0	0.310	t-test, unequal variance	2	13.5	(0.138,0.482)	0.005
3h	AD - DHA = 0	0.124	t-test, unequal variance	3	3.59	(0.069,0.179)	0.04
3h	AB - ABOL = 0	0.309	t-test, unequal variance	3	12.6	(0.170,0.447)	0.001

TABLE 15. Ligand efficiency.

Ligand	IC₅₀ (μM)	# Heavy Atoms	Ligand Efficiency (kcal/mol-atom)*
<i>Amorphadiene</i>	50	15	0.39
<i>α-bisabolene</i>	13	15	0.44
<i>BBR</i>	8	41	0.17
<i>MSI-1436</i>	0.6	47	0.17

* Ligand efficiency = $(-2.303RT)/HAC \cdot \log(IC_{50})$, where R is the gas constant, T is the temperature in K, and HAC is the number of heavy atoms.

|

OTHER EMBODIMENTS

All of the features disclosed in this specification may be combined in any combination. Each feature disclosed in this specification may be replaced by an alternative feature serving the same, equivalent, or similar purpose. Thus, unless expressly stated
5 otherwise, each feature disclosed is only an example of a generic series of equivalent or similar features.

From the above description, one skilled in the art can easily ascertain the essential characteristics of the present disclosure, and without departing from the spirit and scope thereof, can make various changes and modifications of the disclosure to adapt it to various
10 usages and conditions. Thus, other embodiments are also within the claims.

EQUIVALENTS AND SCOPE

While several inventive embodiments have been described and illustrated herein, those of ordinary skill in the art will readily envision a variety of other means and/or
15 structures for performing the function and/or obtaining the results and/or one or more of the advantages described herein, and each of such variations and/or modifications is deemed to be within the scope of the inventive embodiments described herein. More generally, those skilled in the art will readily appreciate that all parameters, dimensions, materials, and configurations described herein are meant to be exemplary and that the actual parameters,
20 dimensions, materials, and/or configurations will depend upon the specific application or applications for which the inventive teachings is/are used. Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific inventive embodiments described herein. It is, therefore, to be understood that the foregoing embodiments are presented by way of example only and that,
25 within the scope of the appended claims and equivalents thereto, inventive embodiments may be practiced otherwise than as specifically described and claimed. Inventive embodiments of the present disclosure are directed to each individual feature, system, article, material, kit, and/or method described herein. In addition, any combination of two or more such features, systems, articles, materials, kits, and/or methods, if such features, systems, articles, materials,
30 kits, and/or methods are not mutually inconsistent, is included within the inventive scope of the present disclosure.

All definitions, as defined and used herein, should be understood to control over dictionary definitions, definitions in documents incorporated by reference, and/or ordinary meanings of the defined terms.

5 All references, patents and patent applications disclosed herein are incorporated by reference with respect to the subject matter for which each is cited, which in some cases may encompass the entirety of the document.

The indefinite articles “a” and “an,” as used herein in the specification and in the claims, unless clearly indicated to the contrary, should be understood to mean “at least one.”

10 The phrase “and/or,” as used herein in the specification and in the claims, should be understood to mean “either or both” of the elements so conjoined, i.e., elements that are conjunctively present in some cases and disjunctively present in other cases. Multiple elements listed with “and/or” should be construed in the same fashion, i.e., “one or more” of the elements so conjoined. Other elements may optionally be present other than the elements specifically identified by the “and/or” clause, whether related or unrelated to those elements specifically identified. Thus, as a non-limiting example, a reference to “A and/or B”, when
15 used in conjunction with open-ended language such as “comprising” can refer, in one embodiment, to A only (optionally including elements other than B); in another embodiment, to B only (optionally including elements other than A); in yet another embodiment, to both A and B (optionally including other elements); etc.

20 As used herein in the specification and in the claims, the phrase “at least one,” in reference to a list of one or more elements, should be understood to mean at least one element selected from any one or more of the elements in the list of elements, but not necessarily including at least one of each and every element specifically listed within the list of elements and not excluding any combinations of elements in the list of elements. This definition also
25 allows that elements may optionally be present other than the elements specifically identified within the list of elements to which the phrase “at least one” refers, whether related or unrelated to those elements specifically identified. Thus, as a non-limiting example, “at least one of A and B” (or, equivalently, “at least one of A or B,” or, equivalently “at least one of A and/or B”) can refer, in one embodiment, to at least one, optionally including more than one,
30 A, with no B present (and optionally including elements other than B); in another embodiment, to at least one, optionally including more than one, B, with no A present (and optionally including elements other than A); in yet another embodiment, to at least one,

optionally including more than one, A, and at least one, optionally including more than one, B (and optionally including other elements); etc.

It should also be understood that, unless clearly indicated to the contrary, in any methods claimed herein that include more than one step or act, the order of the steps or acts
5 of the method is not necessarily limited to the order in which the steps or acts of the method are recited.

In the claims, as well as in the specification above, all transitional phrases such as “comprising,” “including,” “carrying,” “having,” “containing,” “involving,” “holding,” “composed of,” and the like are to be understood to be open-ended, i.e., to mean including
10 but not limited to. Only the transitional phrases “consisting of” and “consisting essentially of” shall be closed or semi-closed transitional phrases, respectively, as set forth in the United States Patent Office Manual of Patent Examining Procedures, Section 2111.03.