ACS SyntheticBiology

Research Article

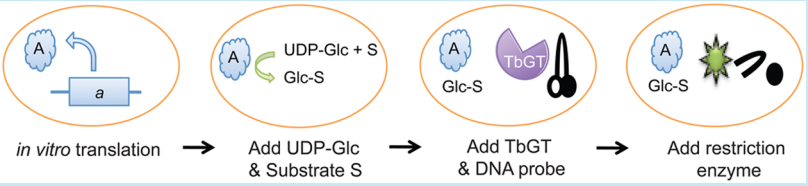
pubs.acs.org/synthbio

DNA-Linked Enzyme-Coupled Assay for Probing Glucosyltransferase Specificity

David J. Sukovich, Cyrus Modavi, Markus de Raad, Robin N. Prince, and J. Christopher Anderson*

Department of Biological Engineering, Synthetic Biology Institute, University of California, Berkeley, Berkeley, California 94704, United States

Supporting Information



ABSTRACT: Traditional enzyme characterization methods are low-throughput and therefore limit engineering efforts in synthetic biology and biotechnology. Here, we propose a DNA-linked enzyme-coupled assay (DLEnCA) to monitor enzyme reactions in a high-throughput manner. Throughput is improved by removing the need for protein purification and by limiting the need for liquid chromatography mass spectrometry (LCMS) product detection by linking enzymatic function to DNA modification. We demonstrate the DLEnCA methodology using glucosyltransferases as an illustration. The assay utilizes cell free transcription/translation systems to produce enzymes of interest, while UDP-glucose and T4-β-glucosyltransferase are used to modify DNA, which is detected postreaction using qPCR or a similar means of DNA analysis. OleD and two glucosyltransferases from *Arabidopsis* were used to verify the assay's generality toward glucosyltransferases. We further show DLEnCA's utility by mapping out the substrate specificity for these enzymes.

KEYWORDS: *synthetic biology, glycosyltransferase, enzyme specificity, biochemical assays*

Enzymatic biochemical reactions have been studied for decades, as they are responsible for the thousands of metabolic processes required for life. With the rise of biotechnology, synthetic biology, and metabolic engineering, enzymes are increasingly viewed as tools to catalyze a desired chemical reaction,[1−3] even those that are non-natural.[4,5] As enzymes often act on an array of structurally related compounds, they are often probed for activity on substrates other than their natural one to form a structure activity relationship (SAR).[6] Additionally, when searching for the best performing enzyme, it is often desirable either to apply protein engineering to create new enzymes with a desired characteristic[7] or to screen several homologues.[8,9]

The space to mine for the best performing enzyme is quite large. Large protein libraries can be generated in *E. coli* when applying protein engineering[7,10−13] and over 85 000 enzymatic reaction observations are cataloged in the largest public database (BRENDA).[14] Even this is dwarfed by estimates of total gene diversity in the environment.[15−17] Although there are many assays for monitoring enzymatic reactions, many reactions can only be observed using liquid or gas chromatography mass spectrometry (LCMS or GCMS) applied to purified protein. These are laborious and low-throughput and therefore severely limit the space that can be experimentally mined. As gene synthesis costs continue to decrease, the ability to perform such assays becomes the dominant bottleneck in mining efforts.

One generalized method to circumvent LCMS and GCMS analysis in the preliminary screening phase is to indirectly monitor enzymatic activity on a given substrate through turnover of a required cofactor. For example, numerous assays have been developed to follow the production or depletion of NADH or NADPH utilizing the absorbance spectra of these cofactors or alternative reaction products.[18−20] In terms of glycosyltransferases, cofactor-monitoring assays may follow the production of nucleotide formation during glycosyltransferase reactions. For example, BellBrooks Lab has developed a glycosyltransferase assay that utilizes an UDP-antibody and fluorescence polarization readout.[21] When UDP is released through a glycosyltransferase activity, the free UDP displaces an UDP-fluorophore from an immobilized antibody, resulting in a fluorescence-detectable readout. Similarly, Kumagai *et al.* described a fluorescent assay that follows glycosyltransferase activities by detecting nucleotide production after the glycosyltransfer reaction.[22] While these particular assays are beneficial in that one can identify enzymatic activity in a multiwell setting, one remains unable to link this activity to a particular protein in a mixed multienzyme library.

We sought to develop methodology to improve throughput and cost of enzyme characterization that matched the scalability of new DNA synthesis methods. Here, we describe a DNA-

Received: October 21, 2014
Published: January 26, 2015

DOI: 10.1021/sb500341a
ACS Synth. Biol. 2015, 4, 833−841

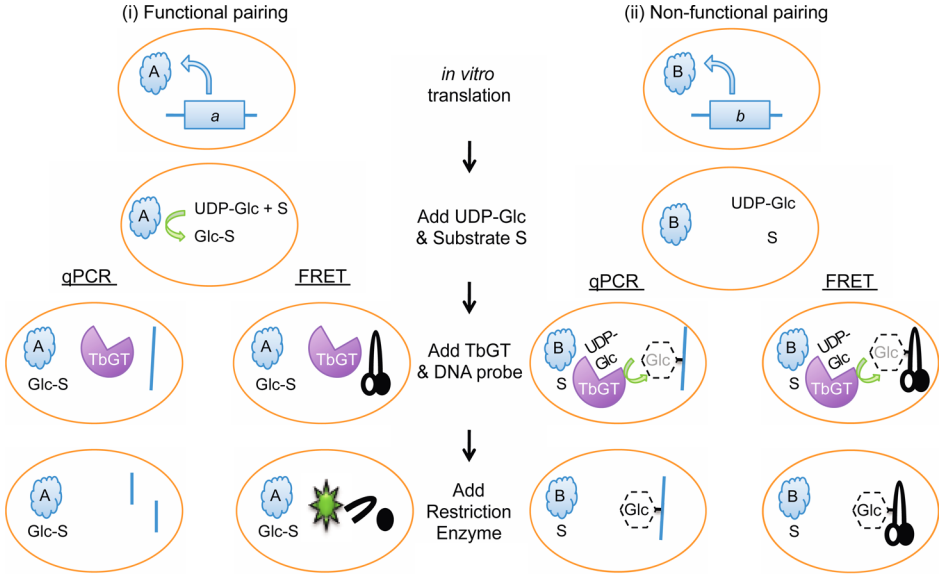


Figure 1. DNA-linked enzyme-coupled assay (DLEnCA). An overview of DLEnCA where the end readout is either qPCR or fluorescence using FRET. (i) A functional pairing of enzyme A and substrate S, leading to depletion of UDP-Glc and no modification of either qPCR or FRET probe by TbGT. The end result is the digestion of probes upon restriction enzyme addition. (ii) A nonfunctional pairing of enzyme B and chemical S, leading to no depletion of UDP-Glc and no modification of probes by TbGT. The end result is the protection of the probes from restriction enzyme digestion. Key: TbGT = T4-β-glucosyltransferase; Glc = glucose.

linked enzyme-coupled assay (DLEnCA) to monitor enzyme reactions in a manner amenable to multiplex gene synthesis[23−25] linked to multiplex readout in droplets and deep sequencing instrumentation.[26,27] The sample preparation steps of cloning plasmid DNA, transformation, and protein purification are eliminated, as is LCMS for analytics. We achieve this by uniting PCR and cell-free transcription/translation systems to generate protein and link the enzymatic output to modifications in DNA for easy analysis with agarose gels, quantitative PCR (qPCR), fluorescence resonance energy transfer (FRET) probes, or potentially next generation sequencing.[28−30]

DLEnCA is based on a competition between two enzymatic reactions for a common cofactor. The reactions are separated temporally, such that depletion of the cofactor in the first reaction results in no modification in the second reaction. In DLEnCA, the second reaction results in a DNA modification. In the case of a glucosyltransferase version of DLEnCA, the assay employs the enzyme T4-β-glucosyltransferase (TbGT). TbGT modifies 5-hydroxymethyl-cytosine (5-hm-cytosine) residues in a DNA with the glucose moiety from uridine diphosphate glucose (UDP-Glc).[28,30−33] This atypical base can be incorporated into DNA using PCR containing 5-hydroxymethyl-dCTP[34] or by phosphoramidite synthesis.[35,36] Therefore, the 5-hm-cytosine-modified DNA, referred to as the "probe" DNA within this publication, can be a small linear DNA fragment, modified FRET probe, or even the original DNA template that initiates the assay. The modification of the 5-hm-cytosines of probe can be detected because glucosylation of double-stranded DNA can block recognition by other DNA modification enzymes, including restriction endonucleases.[28,30] Therefore, depletion of the cofactor in the first reaction is recorded as an absence of a chemical modification to the DNA probe, and this difference can be translated into an easily observed signal by DNA digestion, followed by agarose gel analysis, qPCR, fluorescence, or deep sequencing.

A schematic overview of the DLEnCA assay is shown in Figure 1. Two workflows were developed to follow an enzymatic reaction using a cell-free transcription/translation system: (1) a qPCR-amenable workflow and (2) a fluorescence-amenable workflow. Both assay schemes are initiated upon the addition of linear DNA encoding a promoter and an enzyme of interest to a cell-free transcription/translation system. After an initial incubation period, UDP-Glc and chemical are added to the reaction. If the enzyme of interest is able to glucosylate the substrate, UPD-Glc concentrations within the reaction are depleted (Figure 1, i). The opposite is true if the enzyme of interest is not able to glucosylate the substrate (Figure 1, ii). After a second incubation period, TbGT and DNA probe are added to the reaction. The probe is a linear DNA fragment where all cytosines have been modified to 5-hm-cytosine (qPCR readouts) or a hairpin DNA oligonucleotide modified to contain two 5-hm-cytosines within an MfeI recognition site, a 5′ fluorophore, and a 3′ quencher (fluorescence readouts). After a third incubation period is completed, MfeI restriction endonuclease is added to the assay and the results are read using qPCR or fluorometry, respectively.

RESULTS AND DISCUSSION

In order to verify that DNA protection using TbGT could be used to follow an enzymatic reaction under conditions amenable to downstream DNA analyses, the ability of recombinant OleD to reduce concentrations of UDP-Glc while in the presence of the substrate kaempferol was first tested. OleD is a well-characterized enzyme of interest to biotechnology due to its ability to glucosylate pharmaceuticals and, as a result, modify their physical properties.[2,37−40] In particular, OleD has been shown to glucosylate kaempferol. The schematic of this reaction is shown in Figure 2A. To identify if this reaction could be monitored by modification of DNA, recombinant OleD was purified (Supporting Information Figure S1) and incubated with 5-hm-cytosine-modified probe

DOI: 10.1021/sb500341a
ACS Synth. Biol. 2015, 4, 833−841

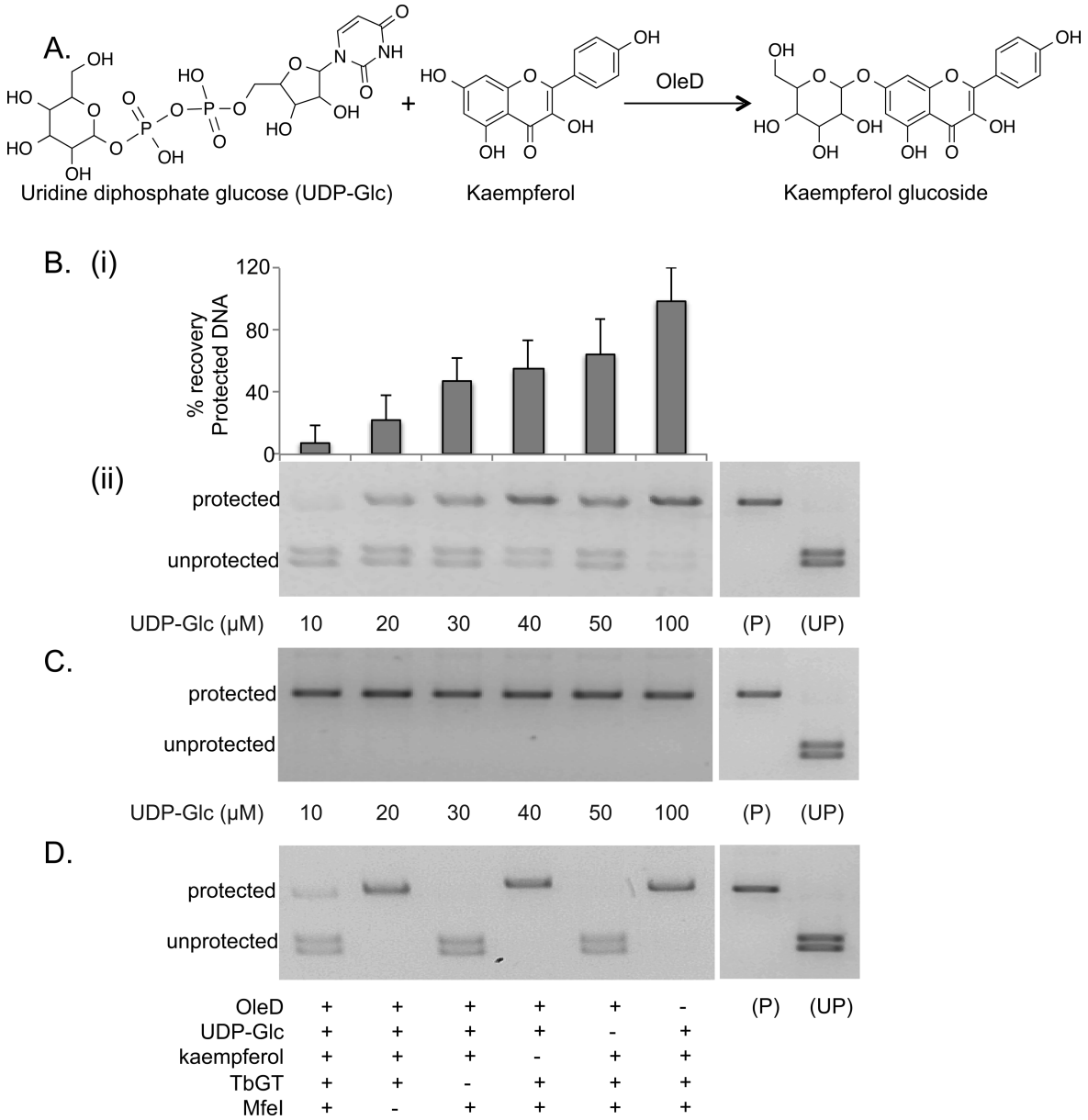


Figure 2. DNA-linked enzyme-coupled assay using purified model enzyme (OleD). (A) The OleD-catalized glucosylation of kaempferol using UDP-Glc as a donor-molecule. (B) Purified recombinant OleD (0.5 μM) was incubated with probe DNA (4.5 nM) in the presence of kaempferol (0.5 mM) and various concentrations of UDP-Glc at 37 °C for 1 h. Following incubation with TbGT, DNA was digested with MfeI. Digests were analyzed on a 2% agarose gel. (i) % recovery of intact probe DNA was identified using band intensities quantified by ImageJ and compared to completely protected probe intensities (n = 4, error bars = standard error). (ii) A representative gel of the UDP-Glc titration reaction, as-well-as a "ladder" depicting expected DNA digestion sizes, is shown. (C) Purified recombinant OleD (0.5 μM) was incubated with probe DNA (4.5 nM) in the presence of 4-hydroxybenzoic acid (0.5 mM) as acceptor and various concentrations of UDP-Glc at 37 °C for 1 h. Following incubation with TbGT, DNA was digested with MfeI. Digests were analyzed on a 2% agarose gel (n = 4). A representative gel of the titration, along with a "ladder" depicting expected DNA digestion sizes, is shown. (D) Purified recombinant OleD (0.5 μM) was incubated with probe DNA (4.5 nM) in the presence of kaempferol (0.5 mM) and UDP-Glc (20 μM) at 37 °C for 1 h. Following incubation with TbGT, DNA was incubated with MfeI. Digests were analyzed on a 2% agarose gel (n = 4). A representative gel, including component knockdown reactions, along with a "ladder" depicting expected DNA digestion sizes is shown. Key: (P) = protected DNA after digestion with MfeI; (UP) = unprotected DNA after digestion with MfeI; (+) = presence of component in reaction; (−) = absence of component in reaction.

DNA (4.5 nM probe DNA representing 2.5 μM modified 5-hm-cytosines, with the expectation that all cytosines in the probe were modified to 5-hm-cytosine, as described in Methods), kaempferol (0.5 mM), and UDP-Glc concentrations varying from 10 μM to 1 mM in a buffered solution. As controls, probe DNA was also incubated under identical conditions with no UDP-Glc or 10 mM UDP-Glc. Following TbGT incubation and MfeI digestion, DNA integrity was identified using standard agarose gels (Figure 2B). As expected, when less UDP-Glc was present at the start of the assay, more probe DNA was digested using the restriction enzyme MfeI. To determine whether an enzymatic reaction—and not spontaneous UDP-Glc hydrolysis—was being followed, the assay was performed under identical conditions using a different UDP-Glc acceptor (4-hydroxybenzoic acid) (Figure 2C). Under these conditions, no DNA was digested after the UDP-Glc

DOI: 10.1021/sb500341a
ACS Synth. Biol. 2015, 4, 833−841

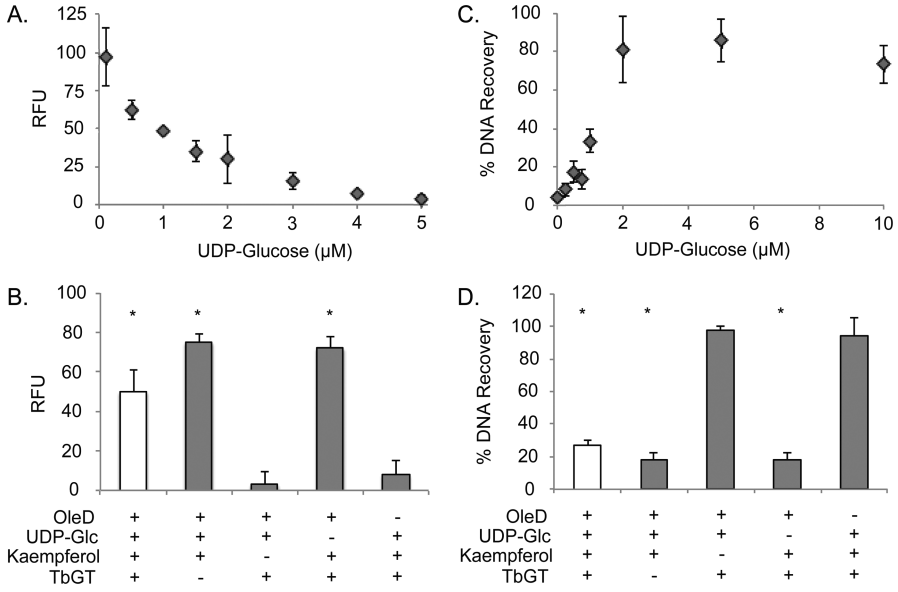


Figure 3. DNA-linked enzyme-coupled assay using cell-free transcription/translation system. (A) UDP-Glc was titrated into reactions containing 250 nM FRET probe, TbGT, and a cell-free transcription/translation system. After 3 h incubation at 37 °C, probe was digested with MfeI, and resulting digests were analyzed using fluorescent spectrometry (n = 6, error bars = standard error). (B) DNA encoding OleD (5 nM) was incubated in a cell-free transcription/translation system for 3 h at 37 °C, followed by the addition of UDP-Glucose (5 μM) and kaempferol (1 mM). After a subsequent incubation of 3 h, TbGT and FRET probe (250 nM) were added. Following MfeI addition, fluorescence formation was monitored using fluorometry. Experimental results, as-well-as component knockdown reactions, are plotted as Relative Fluorescent Units (RFU) (n = 8; error bars = standard error). (C) UDP-Glc was titrated into assays containing probe DNA (0.2 nM), TbGT, and a cell-free transcription/translation system. After 3 h incubation at 37 °C, probe was purified and digested with MfeI. Digests were analyzed using qPCR. Results are plotted as % recovery of probe DNA (recovery of experimental DNA/recovery of completely protected control DNA) (n = 10, error bars = standard error). (D) DNA encoding OleD (5 nM) was incubated for 3 h at 37 °C in a cell-free transcription/translation system, followed by the addition of UDP-Glc (2 μM) and kaempferol (1 mM). After a subsequent incubation of 3 h, TbGT and probe DNA (0.2 nM) were added. Following probe purification and MfeI digestion, intact DNA probe was identified using qPCR. Experimental results, as-well-as component knockdown reactions, are plotted as % recovery of probe DNA (n = 11; error bars = standard error). Key: (+) = presence of component in reaction; (−) = absence of component in reaction. * = p < 0.05 vs no chemical control.

titration enzyme assay was completed. As a final verification, the assay was performed with various component dropouts using 20 μM as starting UDP-Glc concentration (Figure 2D). As expected, when all components of the assay were present, significant DNA digestion occurred. Under conditions where negligible protection of probe DNA was expected (i.e., no UDP-Glc or TbGT), no protected DNA was identified. In contrast, when protection was expected (no OleD or kaempferol), significant concentrations of protected DNA was found. From this, it was concluded that an enzymatic reaction could be followed using DNA protected from MfeI digestion.

In order to further develop the assay toward a mid-throughput workflow, the protocol was modified to utilize nonpurified protein derived from a cell-free transcription/translation system as starting material. This assay was first examined using a fluorescent readout (FRET DLEnCA, Figure 1, FRET workflow). A 38-mer hairpin oligonucleotide containing a Cy3-fluorophore at the 5′-termini and a Black Hole Quencher at the 3′-termini, as-well-as two 5-hm-cytosine residues overlapping an MfeI-recognition site, was purchased; the FRET probe was quenched in its initial state but became fluorescent upon release of the quencher by digestion with MfeI. To determine a threshold for probe protection under conditions amenable to the FRET DLEnCA workflow, UDP-Glc was titrated into mixtures of the cell-free transcription/translation system containing FRET probe (250 nM) and TbGT. After digestion with the restriction enzyme MfeI, the amount of relative fluorescence was monitored by fluorometry (Figure 3A). A threshold of protection at 5 μM UDP-Glc was identified. This corresponded to a 10:1 molar ratio of UDP-Glc needed to protect the 5-hm-cytosines in the probe. To verify that an enzymatic reaction could be monitored using cell-free transcription/translation-derived protein and the FRET probe as readout, the OleD-catalyzed glucosylation reaction of kaempferol was followed using DNA encoding a T7 promoter and OleD, kaempferol (1 mM), and 5 μM UDP-Glc as starting donor material (Figure 3B). As expected, when all components of the assay were present, a fluorescent reading of 50 RFU was detected. In contrast, omission of the OleD gene or kaempferol resulted in significant protection of the FRET probe (3 RFU and 8 RFUs, respectively). When no protection occurred, as when UDP-Glc or TbGT were omitted from the assay, 75 RFU and 72 RFUs were detected (Figure 3B). From this, it was determined that the FRET-DLEnCA workflow could be used to monitor an enzymatic reaction.

The second DLEnCA workflow developed involved qPCR as a final readout (qPCR DLEnCA, Figure 1, qPCR workflow). A linear fragment of DNA was modified to contain 5-hm-cytosines throughout. To determine the protection threshold for the probe required for qPCR DLEnCA, UDP-Glc was titrated into mixtures of the cell-free transcription/translation system containing 5-hm-cytosine-modified DNA probe (0.2 nM probe DNA containing 0.1 μM modified 5-hm-cytosines, with the expectation that all cytosines in the probe were modified to 5-hm-cytosine, as described in Supporting Information Methods) and TbGT. After digestion with the

DOI: 10.1021/sb500341a
ACS Synth. Biol. 2015, 4, 833−841

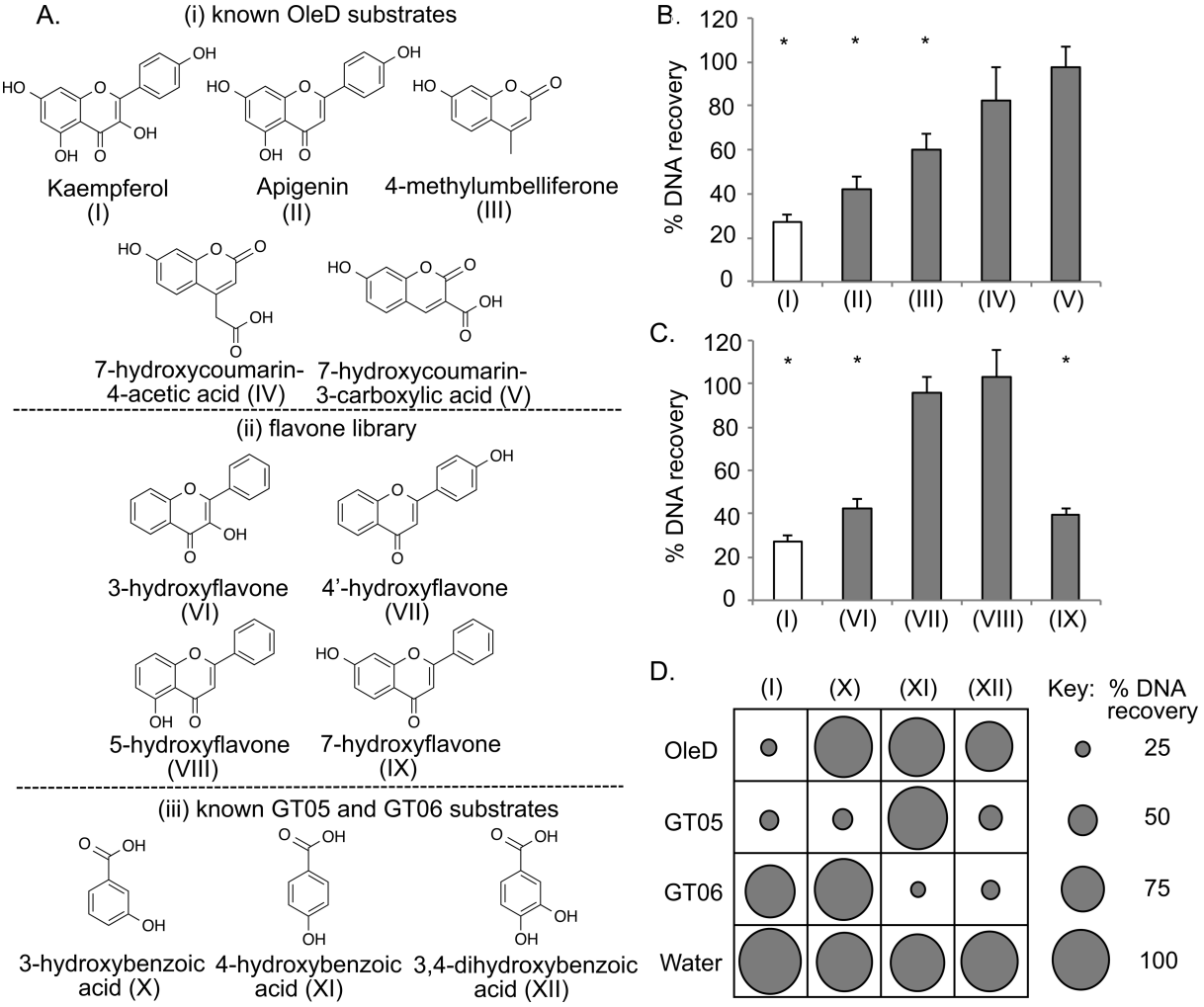


Figure 4. qPCR DLEnCA and substrate specificity studies. (A) Chemicals used as substrates in subsequent assays: (i) known substrates of OleD; (ii) flavone library; (iii) known substrates of GT05 (UGT72B1 (Q9M156)) and GT06 (UGT89B1 (AT1G73880)). (B) DNA encoding OleD (5 nM) was incubated in a cell-free transcription/translation system for 3 h at 37 °C, followed by the addition of chemical (1 mM) and UDP-Glc (2 μM). After a subsequent incubation of 3 h, TbGT and probe (0.2 nM) were added. After probe purification and MfeI digestion, probe integrity was identified using qPCR. Percent recovery of probe is reported (n = 8; error bars = standard error). (C) DNA encoding OleD (5 nM) was incubated for 3 h at 37 °C in a cell-free transcription/translation system, followed by the addition of flavone (1 mM) and UDP-Glc (2 μM). After a subsequent incubation of 3 h, TbGT and probe (0.2 nM) were added. After probe purification and MfeI addition, probe integrity was identified using qPCR. Percent recovery of probe is reported (n = 8; error bars = standard error). (D) DNA encoding OleD, GT05, or GT06 (5 nM) were incubated for 3 h at 37 °C in a cell-free transcription/translation system, followed by the addition of chemical (1 mM) and UDP-Glc (2 μM). After a subsequent incubation of 3 h, TbGT and probe (0.2 nM) were added. After probe purification and MfeI digestion, probe integrity was identified using qPCR. Percent recovery of probe is reported as a heat map (n = 8) (specific recovery figures and percent error can be found in Supporting Information Figure S12). * = $p < 0.05$ vs no chemical control.

restriction enzyme MfeI, the amount of intact DNA probe was determined using qPCR (Figure 3C). As expected, greater starting concentrations of UDP-Glc resulted in greater protection of the probe DNA, with the threshold for protection identified as a 20:1 molar ratio of UDP-Glc to total 5-hmcytosine in probe DNA. This UDP-Glc amount was used as the starting concentration of UDP-Glc for all qPCR experiments. To test the feasibility of qPCR DLEnCA, OleD was again tested upon the substrate kaempferol (Figure 3D), using DNA encoding a T7 promoter and OleD as starting material. As expected, when all components of the assay were present, 27% of intact probe DNA was recovered. Under conditions where negligible protection of the probe DNA was expected (i.e., no UDP-Glc or TbGT conditions), only 17% and 18% of the probe was recovered, as compared to 97% and 94% recovery under conditions where substantial protection was expected (no OleD or kaempferol, respectively). LCMS was used to verify that glucosylated products were produced under expected conditions (Supporting Information Figure S2).

Further verification of qPCR DLEnCA capability was shown when OleD was tested against other known OleD substrates (Figure 4A, i).[2,37,40] Using the previously described protocol, DNA encoding OleD was incubated with kaempferol (**I**), apigenin (**II**), 4-methylumbelliferone (**III**), 7-hydroxycoumarin-4-acetic acid (**IV**), and 7-hydroxycoumarin-3-carboxylic acid (**V**). Results to this assay are shown in Figure 4B as percent DNA recovery. A threshold of activity can be seen when one takes into consideration the lower activity rate of OleD against substrates **II** and **III**. In previous studies, OleD was reported to glucosylate these substrates at a reduced rate or with less efficiency than with **I**.[2,37,40] Consistent with earlier observations, while LCMS analysis identified the chemicals as

DOI: 10.1021/sb500341a
ACS Synth. Biol. 2015, 4, 833−841

substrates (Supporting Information Figure S3−S4), the assay was only able to recover 42% and 60% of the probe DNA, respectively. Other substrates tested include **IV** and **V**. Both were found in previous publications to be very weak substrates of OleD.[2,37,40] This observation was verified using both qPCR DLEnCA (82% and 97% recovery of probe respectively) and LCMS (Supporting Information Figure S5−S6). From this, it was concluded qPCR DLEnCA could be used to follow an enzymatic reaction.

During the initial verification of the assay, it was observed by LCMS that two predominant glucosides were produced when kaempferol was incubated with OleD (Supporting Information Figure S2). Kaempferol has four hydroxyl-motifs amenable for glucosylation in its structure. To help identify which hydroxyl groups were being preferentially glucosylated, qPCR DLEnCA was employed using four monohydroxylated flavones as substrates for the enzyme (Figure 4A, ii). As shown in Figure 4C, 3-hydroxyflavone (**VI**) and 7-hydroxyflavone (**IX**) resulted in a significant loss in protection of the DNA probe, at 42% and 39%, respectively. This was not the case with 4′-hydroxyflavone (**VII**) and 5-hydroxyflavone (**VIII**). Flavone glucosylation was also verified by LCMS (Supporting Information Figure S7−S10). It was concluded that the 3-hydroxy and the 7-hydroxy positions of kaempferol were the preferential glucosylation motifs of OleD. This observation was verified by use of glucoside standards (Sigma-Aldrich) on LCMS. As shown in Supporting Information Figure S11, the purchased kaempferol-3-glucoside standard eluted from LCMS at the same time as the first product peak in the experimental reaction, while the elution time of the kaempferol-7-glucoside standard was identical to the second peak.

Finally, to demonstrate the assay's generalizability to the glucosyltransferase enzyme class, qPCR DLEnCA was used to test the activity of two *Arabidopsis* glucosyltransferases against previously identified substrates (Figure 4A, iii).[41,42] GT05 (71B2) and GT06 (89B1) were cloned from cDNA obtained from *Arabidopsis thaliana* and used as starting material for the cell-free transcription/translation assays; OleD was also included for verification reasons (Figure 4D). As expected, OleD only showed activity with **I**. This was identified by qPCR DLEnCA and verified with LCMS (Figure 4D, Supporting Information Figures S12−S16). GT05 was shown to not only have activity with 3-hydroxybenzoic acid (**X**) (33% recovery) and 3,4-dihydroxybenzoic acid (**XII**) (40% recovery) but was also found to interact with **I** (32% recovery) (Figure 4D; Supporting Information Figures S12−S16). Finally, as expected, GT06 was found to interact with 4-hydroxybenzoic acid (**XI**) (25% recovery) and (**XII**) (30% recovery) (Figure 4D; Supporting Information Figures S12−S16). This highlighted that qPCR DLEnCA could follow glucosyltransferase activities from eukaryotes.

Three enzyme assays that utilize DNA modifications as a final readout were described herein with similar results. Each has its own advantages and disadvantages. For example, the first assay described is relatively easy and uses equipment most laboratories have access to, though the need to purify protein makes it more time-consuming than the others. In contrast, FRET DLEnCA and qPCR DLEnCA are relatively quicker assays, with FRET DLEnCA using less hands-on time than the others. This results in less error due to human and mechanical errors. Unfortunately, due to the 5-hm-cytosine modifications in the fluorescent probe, FRET DLEnCA is a more costly assay.

The technology described also has a number of advantages over existing methods for the biochemical analysis of glucosyltransferases. For example, DLEnCA can alleviate the need to purify proteins or use time-consuming analytics. At the same time, the speed of the assay allows for a quick screen of either substrate specificities of an enzyme of interest or testing different enzymes for activity on a given substrate. For example, in the course of a single day the glucosylation-specificity of OleD was mapped to the 3- and 7-hydroxy-motifs found on kaempferol. This mapping was later verified using LCMS and glucoside standards.

The ability to link an enzymatic reaction to a DNA modification is a significant advantage for DLEnCA. This advantage permits a researcher to link a protein's activity to its encoding DNA, allowing for the potential use of multiplexed or mixed enzyme libraries in a single assay. Additionally, the throughput of the assay could be further improved by coupling the assay with liquid handling robotics and/or emulsion technologies, and next generation sequencing techniques. Indeed, additional work beyond the scope of this manuscript has technically demonstrated the ability to connect DLEnCA with deep sequencing (manuscript in preparation).

FRET DLEnCA and qPCR DLEnCA are dependent upon the use of the PURExpress cell-free transcription/translation system. While this is beneficial, giving a researcher the ability to study the enzymatic activities of toxic or chronically insoluble proteins without the need for protein purification and/or concentration, some issues of note should be mentioned. For example, less-purified commercially available transcription/translation systems were attempted, but their background levels of enzymatic activity with the cofactor and DNase activity were too high for efficient use in this assay (data not shown). The assays are also dependent upon the ability of the transcription/translation system to produce protein efficiently and consistently; low protein yields due to transcription/translation reagent variability or starting DNA material integrity can result in problems with assay activity or assay reproducibility. Enzymatic rate determination and differentiation between a weak enzymatic reaction or the absence of an enzymatic reaction are also difficult using DLEnCA. One way to circumvent these issues is to use purified protein as a starting material. Though this diminishes FRET DLEnCA and qPCR DLEnCA's advantages of speed, it still circumvents the need for LCMS during initial screens. The assay also depends upon the efficient activity of a restriction endonuclease. While MfeI was used in this study to follow glucosylation of the probe DNAs, it was found to be inefficient in our hands. Even when no protection of the probe was expected, the assay was able to recover a fraction of intact probe DNA. While this causes difficulty if one wants to identify enzymatic rates for the protein of interest, it did not inhibit the analysis of overall activities of the glucosyltransferases of interest. It should also be noted that many chemicals are known to interact with double stranded DNA, modifying enzymatic reactions in a beneficial or potentially adverse manner.[43,44] While this is a phenomenon that could affect the enzyme assay described here, the subsequent LCMS analyses performed in the study show the presence or the lack of glucosylation of the chemical of interest. Most of the chemicals used had also been previously tested for enzyme specificity. Therefore, given the phenomenon described above, while the assay could be used for an initial screen of a large library of chemicals, subsequent analysis is important for confirmation of activity. Finally, poor substrate purity could

DOI: 10.1021/sb500341a
ACS Synth. Biol. 2015, 4, 833−841

result in false positive signals caused by reactions of contaminating compounds. However, because the assay is stoichiometric, the contaminant would need to be present in excess of ~10 μM to see such false signals.

In summary, DNA-linked enzyme-coupled assays can monitor the glucosyltransferase reactions of known or unknown enzymes with diverse substrates. A major challenge in screening large numbers of genes for function has been the ability to connect genotype with phenotype. By linking enzymatic function to chemical modifications of DNA, DLEnCA provides the solution toward screening and characterizing a large class of genes encoding enzymes relevant to constructing microbial chemical factories.

METHODS

Enzymes and Reagents. T4 Phage β-glucosyltransferase (M0357L), MfeI (R0589L), UDP-Glucose (S2200S), dNTP mix (N0447L), and the PURExpress *in vitro* Protein Synthesis Kit (E6800L) were purchased from NEB. The glucosyltransferase was dialyzed against PBS for 2 h prior to use. All other reagents were employed without further purification. 5-Hydroxymethylcytosine dNTP mix (D1040) was purchased from Zymo Research. Phusion High-Fidelity DNA polymerase (F-530L) was purchased from Thermo Scientific. One Shot TOP10 Chemically Competent *E. coli* (404003) was purchased from Invitrogen. L-Arabinose (A3256), kaempferol (60010-25MG), kaempferol-3-glucoside (68437-5MG), kaempferol-7-glucoside (18854-1MG), 3-hydroxybenzoic acid (H20008-100G), 3,4-dihydroxybenzoic acid (37580-25G-F), and 5-hydroxyflavone (H4405-250MG) were purchased from Sigma-Aldrich; 4′-hydroxyflavone (H-411) was purchased from Indofine Chemical Company Inc.; 7-hydroxyflavone (H0852) was purchased from TCI America; 3-hydroxyflavone (A18110) and 4-hydroxybenzoic acid (10170920) were purchased from AlfaAesar; apigenin (sc3529A), 4-methylumbelliferone (sc206910), and 7-hydroxycoumarin-4-acetic acid (sc210628) were purchased from Santa Cruz Biotechnology; 7-hydroxycoumarin-3-carboxylic acid (81215) was purchased from AnaSpec. Plates for qPCR were purchased from Bio-Rad (2239441). Water Optima-LCMS (W6-4), Formic Acid Optima-LCMS (A117-50), and Acetonitrile Optima-LCMS (A955-4) were purchased from Fisher Scientific. All oligonucleotides were purchased from Integrated DNA Technologies.

Cloning of *oleD*. Genetic constructs for expression of glucosyltransferases were fabricated as clonal plasmid samples from preexisting DNAs. Each construct contained the glucosyltransferase under the transcriptional control of a T7 promoter and a T_{rrnB} terminator. The sequences of primers used in this study are listed in Supporting Information Table S1. For the OleD construct, genomic DNA was extracted from *Streptomyces antibioticus* using the UltraClean Microbial DNA Isolation Kit (MO BIO Laboratories). The gene encoding OleD (accession #DQ195536.2) was PCR amplified with primers P01 and P02 using Phusion polymerase. The pET15B backbone containing the T7 promoter, origin of replication, and ampicillin-resistance gene was amplified using primers P03 and P04. The PCRs were carried out using a PTC-200 Peltier Thermo Cycler at the following temperatures for OleD; 98 °C 2 min followed by 35 cycles at 98 °C 30 s, 60 °C 30 s, and 72 °C for 1.5 min, with a final single extended elongation phase at 72 °C for 10 min. For the backbone, the elongation time was extended to 5 min. The amplified DNA was gel purified and added to Gibson Assembly Master Mix (NEB E2611S) and assembled using the recommended protocol. Plasmid was transformed into chemically competent *E. coli* strain DH10B.

Cloning of Other GTs. *Arabidopsis thaliana* cDNA was kindly donated by the Feldman laboratory of UC Berkeley. GT05 (accession #UGT72B1) was amplified using primers P05 and P06, while GT06 (accession #UGT89B1) was amplified using primers P07 and P08. The PCRs were carried out using the following temperature program; 98 °C 2 min followed by 35 cycles at 98 °C 30 s, 62 °C 30 s, and 72 °C for 1.5 min, with a final single extended elongation phase at 72 °C for 10 min. DNA was then digested using Restriction Enzymes NcoI and *Bam*HI and cloned into a predigested pET15b vector.

DNA for Assay. Linear DNA was amplified from sequence-verified plasmids with primers P09 and P10 using the following temperatures: 98 °C 2 min followed by 35 cycles at 98 °C 30 s, 62 °C 30 s, and 72 °C for 1.5 min, with a final single extended elongation phase at 72 °C for 10 min. DNA was purified using DNA Clean & Concentrator (Zymo Research). DNA was quantified using NanoDrop.

Production of Probe DNA. The *araC* gene was amplified from pNE2001, a derivative of pBAD/Myc-His A (Life Technologies) using primers 011 and 012, Phusion DNA polymerase, and the 5-hydroxymethylcytosine dNTP mix. The PCR was carried out using the following temperatures; 98 °C 2 min followed by 35 cycles at 98 °C 30 s, 60 °C 30 s, and 72 °C for 1 min, with a final single extended elongation phase at 72 °C for 10 min. DNA was gel purified using the Zymoclean Gel DNA recovery Kit (Zymo Research).

Purification of OleD. A pre-existing plasmid containing OleD under the transcriptional control of a pBAD promoter was used. Recombinant protein was expressed in TOP10 *E. coli* cultures using 0.2% arabinose for induction. Cultures were grown at 30 °C, induced during logarithmic phase, and allowed to reach saturation before cells were collected and frozen at −80 °C for future use. Protein was purified using the B-PER 6xHis Fusion Protein Spin Purification Kit as described by the user manual. Eluted protein was dialyzed 3× overnight in buffer (50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 10% glycerol, pH 7.9) using the Slide-A-Lyzer MINI Dialysis Device (Thermo Scientific 88402). Protein purity was verified using Bio-Rad Mini-Protean TGX Gels (456-1096), and the protein concentration was identified using the Bradford Protein Assay.

Purified Protein Assays. Dialyzed recombinant OleD (0.5 μM) was incubated with UDP-glucose and kaempferol (0.5 mM) in a buffered solution (50 mM NaCl, 10 mM Tris-HCl, 10 mM MgCl2, pH 7.9). Samples were incubated for 1 h at 37 °C. One unit of T4 Phage β-glucosyltransferase and probe DNA at a concentration of 4.5 nM were added, followed by a 2 h incubation at 37 °C. One unit of MfeI was then added followed by a 4 h incubation at 37 °C. DNA was visualized on a 2% agarose gel with 0.1% GelGreen Nucleic Acid Stain (Biotium; 41004) and blue light. DNA band intensities were identified using ImageJ. Intensities were compared to protected or unprotected samples.

qPCR Assay Protocol. PCR tubes were used for most reactions. All incubations were performed at 37 °C. Assays were initiated upon the addition of linear DNA (5 nM final concentration) to 4 μL PURExpress Solution A and 3 μL PURExpress Solution B. Samples were incubated at 37 °C for 3 h prior to addition of UDP-Glucose and chemical (1 mM final concentration) in DMSO. Samples were incubated for an additional 3 h prior to addition of 1 unit T4 Phage β-glucosyltransferase containing 0.2 nM probe. Samples were

DOI: 10.1021/sb500341a
ACS Synth. Biol. 2015, 4, 833−841

incubated an additional 3 h before DNA was purified using the Zymo DNA Clean and Concentrator Kit. Samples were digested overnight with MfeI prior to analysis. Concentration of intact DNA probe was identified using the iQ SYBR Green Supermix (1708880) with probe DNA as standards. qPCR was carried out on the Bio-Rad iQ5Multicolor Real-Time PCR Detection System using primers P13 and P14 with the following temperatures; 94 °C 30 s followed by 40 cycles of 94 °C 10 s and 55 °C for 30 s. All results were compared to uncut probe DNA that had been carried through the process as negative controls (considered full recovery of probe).

Fluorescence Assay Protocol. Assays were performed using 250 nM fluorescent probe (Sequence provided in Supporting Information Table S1) in place of the linear probe. Fluorescence was detected using a Tecan Safire^2 using the following settings: excitation wavelength 550 nm, emission wavelength 564 nm, excitation bandwidth 5 nm, emission bandwidth 5 nm, gain (manual) 120, number of reads 10, FlashMode High Sensitivity, integration time 100 μs, lag time 0 μs, Z-position 12 000 μm, temperature 37 °C. Development of fluorescence after MfeI addition was monitored over an 8 h period.

LCMS. Determination of glucosylation was accomplished by means of an LCMS system consisting of an Agilent Technologies 1200 series HPLC with an Agilent Technologies 6520 Accurate Mass qTOF LC/MS. An Eclipse Plus C18 (4.6 mm × 100 mm inner diameter, 3.5-μm packing, Agilent Technologies) reverse-phase column with a guard Zorbax Eclipse Plus C18 column (4.6 cm × 12.5 cm, 5 μm packing, Agilent Technologies) was used for separating the samples. Water + 0.1% formic acid and acetonitrile + 0.1% formic acid were used as mobile phases at a flow rate of 500 μL/min. The elution gradient water/acetonitrile ratio) was ramped as follows: 98:2 (v/v) (0−2 min), 98:2−5:95 (v/v) linearly (2−17 min), 5:95 (v/v) (17−27 min), and 5:95−98−2 (v/v) (27−28 min). Full scanning mode (50−750 m/z) was used for data acquisition in a positive-ion mode, and the operation parameters were as follows: ESI probe capillary voltage, +3.5 kV with a scan rate of 1.01 scans/second. The nebulizer gas flow rate was 7 L/min. During the analysis two references (121.0509 m/z ($C_5H_4N_4$) and 922.0098 m/z ($C_{18}H_{18}O_6N_3P_3F_{24}$)) were continuously measured to allow constant mass correction.

Samples were prepared directly from cell-free reactions. Equal-molar amounts of chemical and UDP-glucose were incubated with DNA encoding the glucosyltransferase of interest or water in PURExpress for 3 h at 37 °C. Then, 10 μL reactions were combined with 20 μL 100% ethanol, followed by centrifugation at 13.4 krpm for 5 min. Supernatant was added to 20 μL water in LCMS vials. The reaction/water mixture (10 μL) was injected onto the LCMS per run. Compound presence was identified using MS parent ions.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Experimental characterization and assay results. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel: 510-666-3611. Fax: 510-666-3611. E-mail: jcanderson@berkeley.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Professor L. Feldmen and Dr. Jiang for their donation of *Arabidopsis* cDNA. The Synthetic Biology Institute is supported by a grant from Agilent Technologies as an industry partner.

■ REFERENCES

(1) Kim, D. Y., Rha, E., Choi, S.-L., Song, J. J., Hong, S.-P., Sung, M.-H., and Lee, S.-G. (2007) Development of bioreactor system for L-tyrosine synthesis using thermostable tyrosine phenol-lyase. *J. Microb. Biotechnol. 17*, 116−122.

(2) Williams, G. J., Zhang, C., and Thorson, J. S. (2007) Expanding the promiscuity of a natural-product glycosyltransferase by directed evolution. *Nat. Chem. Biol. 3*, 657−662.

(3) Volpato, G., Rodrigues, R. C., and Fernandez-Lafuente, R. (2010) Use of enzymes in the production of semi-synthetic penicillins and cephalosporins: Drawbacks and perspectives. *Curr. Med. Chem. 17*, 3855−3873.

(4) Meyer, D., Walter, L., Kolter, G., Pohl, M., Muller, M., and Tittmann, K. (2011) Conversion of pyruvate decarboxylase into an enantioselective carboligase with biosynthetic potential. *J. Am. Chem. Soc. 133*, 3609−3616.

(5) Seelig, B., and Szostak, J. W. (2007) Selection and evolution of enzymes from a partially randomized non-catalytic scaffold. *Nature 448*, 828−831.

(6) Cisneros, J. A., Bjorklund, E., Gonzalez-Gil, I., Hu, Y., Canales, A., Medrano, F. J., Romero, A., Ortega-Gutierrez, S., Fowler, C. J., and Lopez-Rodriguez, M. L. (2012) Structure−activity relationship of a new series of reversible dual monoacylglycerol lipase/fatty acid amide hydrolase inhibitors. *J. Med. Chem. 55*, 824−836.

(7) Hibbert, E. G., Senussi, T., Smith, M. E., Costelloe, S. J., Ward, J. M., Hailes, H. C., and Dalby, P. A. (2008) Directed evolution of transketolase substrate specificity towards an aliphatic aldehyde. *J. Biotechnol. 134*, 240−245.

(8) Kubota, T., Tanaka, Y., Hiraga, K., Inui, M., and Yukawa, H. (2013) Characterization of shikimate dehydrogenase homologues of *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol. 97*, 8139−8149.

(9) Canganella, F., Andrade, C. M., and Antranikian, G. (1994) Characterization of amylolytic and pullulytic enzymes from thermophilic archaea and from a new *Fervidobacterium* species. *Appl. Microbiol. Biotechnol. 42*, 239−245.

(10) Brissos, V., Eggert, T., Babral, J. M. S., and Jaeger, K. E. (2008) Improving activity and stability of cutinase towards the anionic detergent AOT by complete saturation mutagenesis. *Protein Eng. Des. Sel. 21*, 387−393.

(11) Miyazaki, K., and Arnold, F. H. (1999) Exploring non-natural evolutionary pathways by saturated mutagenesis: Rapid improvement of protein function. *J. Mol. Evol. 49*, 716−720.

(12) Reetz, M. T., Prasad, S., Carballeira, J. D., Gumulya, Y., and Bocola, M. (2010) Iterative saturation mutagenesis accelerates laboratory evolution of enzyme stereoselectivity: Rigorous comparison with traditional methods. *J. Am. Chem. Soc. 132*, 9144−9152.

(13) Zaccolo, M., and Gherardi, E. (1999) The effect of high-frequency random mutagenesis on *in vitro* protein evolution: A study on TEM-1 β-lactamase. *J. Mol. Biol. 285*, 775−783.

(14) Schomburg, I., Chang, A., Placzek, S., Söhngen, C., Rother, M., Lang, M., Munaretto, C., Ulas, S., Stelzer, M., Grote, A., Scheer, M., and Schomburg, D. (2013) BRENDA in 2013: Integrated reactions, kinetic data, enzyme function data, improved disease classification: New options and contents in BRENDA. *Nucleic Acids Res. 41*, D764−D772.

(15) Henne, A., Daniel, R., Schmitz, R. A., and Gottschalk, G. (1999) Construction of environmental DNA libraries in *Escherichia coli* and

DOI: 10.1021/sb500341a
ACS Synth. Biol. 2015, 4, 833−841

screening for the presence of genes conferring utilization of 4-hydroxybutyrate. *Appl. Environ. Microbiol. 65*, 3901−3907.
(16) Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., Nielsen, R., Pons, N., Levenez, F., Yamada, T., Mende, D. R., Li, J., Xu, J., Li, S., Li, D., Cao, J., Wang, B., Liang, H., Zheng, H., Wie, Y., Tap, J., Lepage, P., Bertalan, M., Batto, J.-M., Hansen, T., Le Paslier, D., Linneberg, A., Nielsen, H. B., Pelletier, E., Renault, P., Sicheritz-Ponten, T., Turner, K., Zhu, H., Yu, C., Li, S., Jian, M., Zhou, Y., Li, Y., Zhang, X., Li, S., Qin, N., Yang, H., Wang, H., Brunak, S., Dore, J., Guarner, F., Kristiansen, K., Pedersen, O., Parkhill, J., and Weissenbach, J. (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature 464*, 59−65.
(17) Venter, J. C., Remington, K., Heidelverg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., Wu, D., Paulsen, I., Nelson, K., Nelson, W., Fouts, D. E., Levy, S., Knap, A. H., Lomas, M. W., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., Baden-Tillson, H., Pfannkoch, C., Rogers, Y.-H., and Smith, H. O. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science 304*, 66−74.
(18) Desquiret-Dumas, V., Gueguen, N., Leman, G., Baron, S., Nivet-Antoine, V., Chupin, S., Chevrollier, A., Vessieres, E., Ayer, A., Ferre, M., Bonneau, D., Reynier, P., and Provaccio, V. (2013) Resveratrol induces a mitochondrial complex I-dependent increase in NADH oxidation responsible for sirtuin activation in liver cells. *J. Biol. Chem. 288*, 36662−36675.
(19) Gray, J. P., Alavian, K. N., Jonas, E. A., and Heart, E. A. (2012) NAD kinase regulates the size of the NADPH pool and insulin secretion in pancreatic β-cells. *Am. J. Physiol. Endocrinol. Metab. 303*, E191−E199.
(20) Smith, I., Vierheller, T., and Thorne, C. (1988) Assay of glutathione reductase in crude tissue homogenates uwing 5,5′-dithiobis(2-nitrobenzoic acid). *Anal. Biochem. 175*, 408−413.
(21) Lowery, R.; Kleman-Leyer, K.; Staeben, M.; Westermeyer, T. Detecting the activity of an screening acceptor substrates, inhibitors, or activators of enzymes catalyzing group transfer reactions to facilitate the development of more selective and therapeutic drugs. US patent No. 20080233592 (2008).
(22) Kumagai, K., Kojima, H., Okabe, T., and Nagano, T. (2014) Development of a highly sensitive, high-throughput assay for glycosyltransferases using enzyme-coupled fluorescence detection. *Anal. Biochem. 447*, 146−155.
(23) Borovkov, A. Y., Loskutov, A. V., Robida, M. D., Day, K. M., Cano, J. A., Olson, T. L., Patel, H., Brown, K., Hunter, P. D., and Sykes, K. F. (2010) High-quality gene assembly directly from unpurified mixtures of microarray-synthesized oligonucleotides. *Nucleic Acids Res. 38*, e180−e180.
(24) Kosuri, S., Eroshenko, N., LeProust, E. M., Super, M., Way, J., Li, J. B., and Church, G. M. (2010) Scalable gene synthesis by selective amplification of DNA pools from high-fidelity microchips. *Nat. Biotechnol. 28*, 1295−1299.
(25) Quan, J., Saaem, I., Tang, N., Ma, S., Negre, N., Gong, H., White, K. P., and Tian, J. (2011) Parallel on-chip gene synthesis and application to optimization of protein expression. *Nat. Biotechnol. 29*, 449−452.
(26) Mardis, E. R. (2008) Next-generation DNA sequencing methods. *Annu. Rev. Genomics Hum. Genet. 9*, 387−402.
(27) Shendure, J., and Ji, H. (2008) Next-generation DNA sequencing. *Nat. Biotechnol. 26*, 1135−1145.
(28) Kinney, S. M., Chin, H. G., Vaisvilla, R., Bitinaite, J., Zheng, Y., Esteve, P. O., Feng, S., Stroud, H., Jacobson, S. E., and Pradhan, S. (2011) Tissue-specific distribution and dynamic changes of 5-hydroxymethylcytosine in mammalian genomes. *J. Biol. Chem. 286*, 24685−24693.
(29) Laird, P. W. (2010) Principles and challenges of genome-wide DNA methylation analysis. *Nat. Rev. Genet. 11*, 191−203.
(30) Terragni, J., Bitinaite, J., Zheng, Y., and Pradhan, S. (2012) Biochemical characterization of recombinant β-glucosyltransferase and analysis of global 5-hydroxymethylcytosine in unique genomes. *Biochemistry 51*, 1009−1019.
(31) Morera, S., Imberty, A., Aschke-Sonnenborn, U., Ruger, W., and Freemont, P. S. (1999) T4 phage β-glucosyltransferase: Substrate binding and proposed catalytic mechanism. *J. Mol. Biol. 292*, 717−730.
(32) Robertson, A. B., Dahl, J. A., Vagbo, C. B., Tripathi, P., Krokan, H. E., and Klungland, A. (2011) A novel method for the efficient and selective identification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res. 39*, e55.
(33) Szwagierczak, A., Bultmann, S., Schmidt, C. S., Spada, F., and Leonhardt, H. (2010) Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res. 38*, e181.
(34) Huang, Y., Pastor, W. A., Shen, Y., Tahiliani, M., Liu, D. R., and Rao, A. (2010) The behavior of 5-hydroxymethylcytosine *n*-bisulfite sequencing. *PLoS One 5*, e8888.
(35) Jin, S.-G., Kadam, S., and Pfeifer, G. P. (2010) Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic Acids Res. 38*, e125.
(36) Xu, Y., et al. (2011) Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol. Cell 42*, 451−464.
(37) Choi, S. H., Ryu, M., Yoon, Y. J., Kim, D.-M., and Lee, E. Y. (2012) Glycosylation of various flavonoids by recombinant oleandomycin glycosyltransferase from *Streptomyces antibioticus* in batch and repeated batch modes. *Biotechnol. Lett. 34*, 499−505.
(38) Gantt, R. W., Peltier-Pain, P., and Thorson, J. S. (2011) Enzymatic methods for glycol (diversification/randomization) of drugs and small molecules. *Nat. Prod. Rep. 28*, 1811−1853.
(39) Gantt, R. W., Goff, R. D., Williams, G. J., and Thorson, J. S. (2008) Probing the aglycon promiscuity of an engineered glycosyltransferase. *Angew. Chem., Int. Ed. 47*, 8889−8892.
(40) Yang, M., Proctor, M., Bolam, D., Errey, J., Field, R., Gilbert, H., and Davis, B. (2005) Probing the breadth of macrolide glycosyltransferases: *In vitro* remodeling of a polyketide antibiotic creates active bacterial uptake and enhances potency. *J. Am. Chem. Soc. Commun. 127*, 9336−9337.
(41) Lim, E. K., Doucet, C., Li, Y., Elias, L., Worrall, D., Spencer, S., Ross, J., and Bowles, D. (2002) The activity of *Arabidopsis* glycosyltransferases toward salicylic acid, 4-hydroxybenzoic acid, and other benzoates. *J. Biol. Chem. 277*, 586−592.
(42) Messner, B., Thylke, O., and Shaffner, A. R. (2003) *Arabidopsis* glucosyltransferase with activities toward both endogenous and zenobiotic substrates. *Planta 217*, 138−146.
(43) Datta, S., Dou, Z., Shibley, A., and Datta, B. (2012) DNA template-assisted modulation of horseradish peroxidase activity. *Int. J. Biol. Macromol. 50*, 552−557.
(44) Lin, J.-L., and Wheeldon, I. (2013) Kinetic enhancements in DNA-enzyme nanostructures mimic the Sabatier principle. *ACS Catal. 3*, 560−564.

DOI: 10.1021/sb500341a
ACS Synth. Biol. 2015, 4, 833−841