Biosensors and Bioelectronics 167 (2020) 112456

Contents lists available at ScienceDirect

Biosensors and Bioelectronics

journal homepage: http://www.elsevier.com/locate/bios

Biosensor-based monitoring of the central metabolic pathway metabolites

Dongqin Ding [a,b,c,1], Jinlong Li [a,1], Danyang Bai [a,b], Huan Fang [a,b], Jianping Lin [a,d,*], Dawei Zhang [a,b,c,**]

[a] Tianjin Institutes of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, 300308, PR China
[b] Key Laboratories of Systems Microbial Biotechnology, Chinese Academy of Sciences, Tianjin, 300308, PR China
[c] University of Chinese Academy of Sciences, Beijing, 100049, PR China
[d] College of Pharmacy, Nankai University, Haihe Education Park, 38 Tongyan Road, Tianjin, 300353, PR China

ARTICLE INFO

Keywords:
Protein redesign
Screening
Biosensors
Metabolic engineering

ABSTRACT

In vivo biosensors have a wide range of applications, from the detection of metabolites to the regulation of metabolic networks, providing a versatile tool for cell biology and metabolic engineering. However, compared with the vast array of small molecules present in nature, the existing range of biosensors is far from sufficient. Here we describe the use of human hypoxanthine guanine phosphoribosyltransferase (HGPRT) as a ligand binding domain (LBD) protein, that acts by ligand-dependent stabilization, to build a biosensor for detection of the pentose phosphate pathway metabolite 5-phospho-α-D-ribose 1-diphosphate (PRPP). Using this protein as a template, we computationally redesigned a new pocket *de novo* according to the pose of the ligand, creating a binding mode exclusive to recognize another pentose phosphate metabolite, D-erythrose 4-phosphate (E4P), and glycerate-3-phosphate (3PG), from the glycolysis pathway. Furthermore, E4P biosensor was developed by fluorescence-activated cell sorting (FACS) and application of it enabled successful screening for the highest phenylalanine-producing strain reported to date. This work provides a strategy for computational design and development of biosensors for a broad range of molecules.

1. Introduction

Sensing and responding to a small-molecule signal require both recognition of the target and linking target recognition to an output response. Biosensors capable of sensing and responding to small molecules *in vivo* are widely applicable in biotechnology (Raman et al., 2014; Zhao et al., 2015) and synthetic biology (Liu et al., 2017b; Ng et al., 2015). Moreover, a wide range of new technologies have enabled rapid engineering and discovery of new biosensors, which are paving the way for a new era of biotechnological progress (Shi et al., 2018). However, the variety of available biosensors is limited in comparison with the vast number of small molecules of interest, thus limiting their potential application. Most *in vivo* biosensors have been designed based on a few naturally occurring transcriptional and translational regulatory systems such as transcription factors (Dietrich et al., 2013; Li et al., 2017) and riboswitches (Zhou and Zeng, 2015). Novel, *in vivo,* protein-based biosensors have also been developed by inserting a ligand binding protein (LBD) into a reporter protein whose signal can be readily detected. This strategy employs ligand-dependent stability, which has been demonstrated in eukaryotic hosts (Feng et al., 2015; Tucker and Fields, 2001) and bacteria (Brandsen et al., 2018). The LBD is conditionally stable, meaning that in the absence of the cognate ligand, these proteins are degraded inside living cells by the ubiquitin proteasome system (Banaszynski et al., 2006; Egeler et al., 2011). Binding of a target ligand stabilizes the reporter protein, preventing its degradation, which allows the signal to be detected (Fig. 1a).

The field of computational protein design has facilitated several biotechnological engineering innovations such as the stabilization of naturally occurring proteins (Capriotti et al., 2005), manipulation of substrate-protein binding specificity (Netzer et al., 2018), development of novel biosensors (Bick et al., 2017; Glasgow et al., 2019) and enzymes (Marshall et al., 2019; Nanda and Koder, 2010; Tinberg et al., 2013), and the creation of novel protein structures (Huang et al., 2016; Kuhlman and Bradley, 2019; Ovchinnikov et al., 2015). Rosetta is a

* Corresponding author. Tianjin Institutes of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, 300308, PR China.
** Corresponding author. Tianjin Institutes of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, 300308, PR China.
E-mail addresses: lin_jp@tib.cas.cn (J. Lin), zhang_dw@tib.cas.cn (D. Zhang).
[1] These authors contributed equally to the work.

https://doi.org/10.1016/j.bios.2020.112456
Received 2 June 2020; Received in revised form 8 July 2020; Accepted 14 July 2020
Available online 2 August 2020

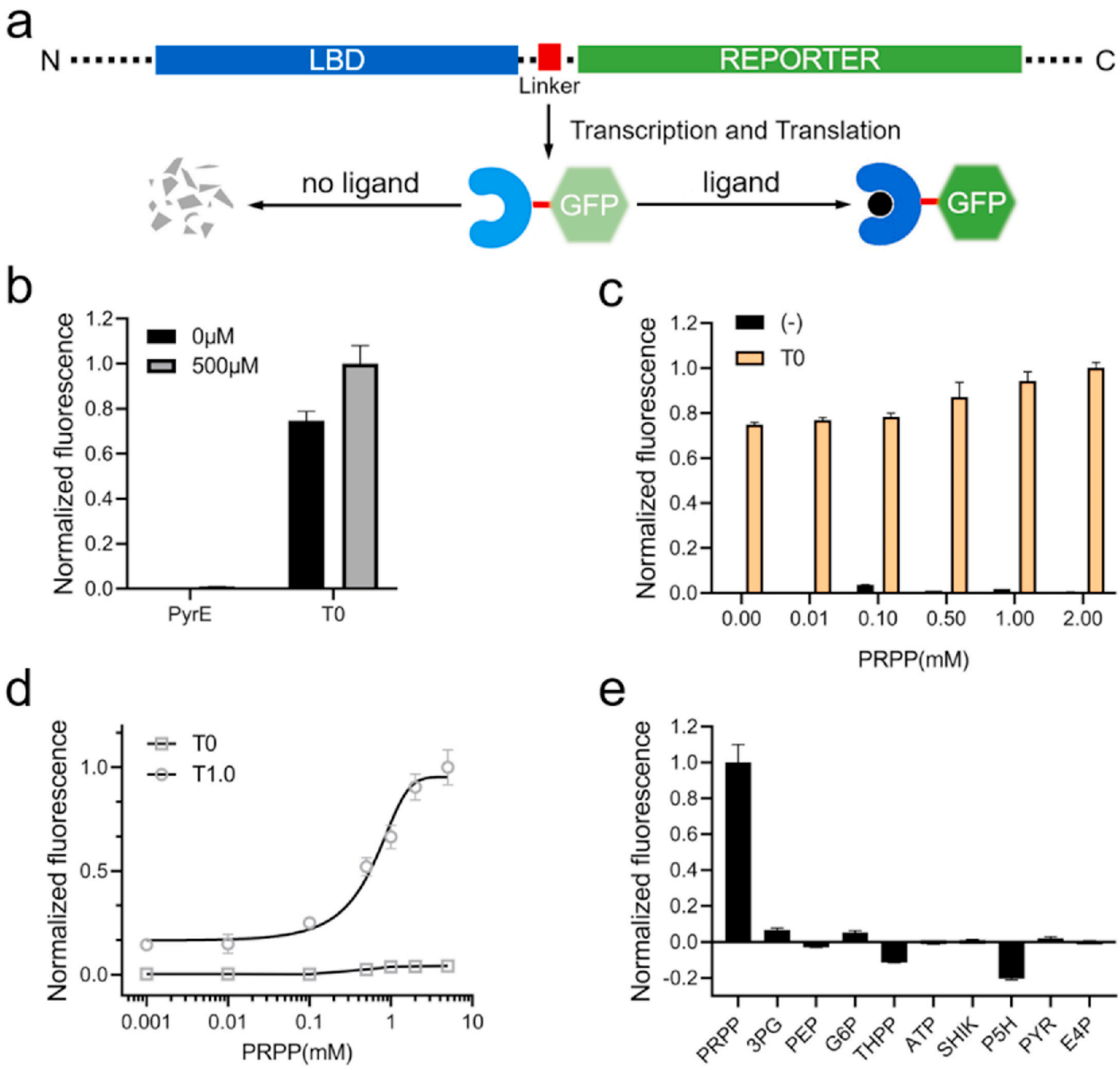

Fig. 1. Design and construction of a PRPP biosensor. **a** Modular biosensor construction from a conditionally destabilized LBD fused to a reporter gene will result in degradation of the reporter unless the small molecule target is present in cells. **b** GFPmut2 fluorescence of PyrE-GFP and T0-GFP biosensors upon addition of 500 μM PRPP. Fluorescence was measured and normalized after incubation at 37 °C for 5 h. **c** Concentration-dependent response to PRPP by the T0-GFP biosensor driving GFPmut2 expression. **d** GFPmut2 fluorescence in *E. coli* strain DH5α expressing the T0 variant T1.0 as a function of PRPP concentration. The corresponding dose-response curve of T0 is included for reference. **e** Ligand specificity of T1.0. The corresponding response of T1.0 in the absence of any molecules is included for reference. PEP (phosphoenolpyruvic acid), G6P (glucose-6-phosphate), THPP (thiamine pyrophosphate chloride), ATP (adenosine triphosphate), SHIK (shikimic acid), P5H (Pyridoxal 5′-phosphate hydrate), PYR (pyruvic acid). Experiments were conducted in triplicate and measurements are represented as the mean ± s.d.

comprehensive software suite that includes algorithms for analysis and design of protein structures and computational modeling. Rosetta has enabled several significant scientific advances in computational biology, including *de novo* protein design (Jiang et al., 2008; Rothlisberger et al., 2008; Siegel et al., 2010), enzyme design (Bick et al., 2017; Li et al., 2018; Tinberg et al., 2013), and structural prediction of biological macromolecules and macromolecular complexes (Chen et al., 2019; Lapidoth et al., 2018; Nagarajan et al., 2015), among other functions. Among its applications, Enzyme Design is used to analyze structural input consisting of a ligand/substrate docked to a protein and provide alternative binding/interacting residues for that substrate. Alternatively, specific catalytic/binding contacts between ligand and protein can be selected beforehand, which are then optimized by the application. By designing specific substrate-protein binding sites, enzyme design can be used to engineer proteins that recognize new molecules (Bick et al., 2017; Feng et al., 2015; Loshbaugh and Kortemme, 2020).

Due to limited knowledge, it is difficult to build biosensors without existing regulatory or binding proteins, and one of our goals is to solve these problems. For small molecules without transcription factors, with unknown or unsuitable binding proteins, we can take an existing protein as a protein template and use computational protein design to *de novo* design the binding pocket so that it can respond to the target molecules, so as to make it possible to build biosensors of the small molecules. Here, for the first time, we built three biosensors for PRPP, E4P and 3PG from the central metabolic pathway from one target protein HGPRT mutant F36L (hereafter T0) based on increased biosensor stability due to ligand binding, which then induces expression of a reporter gene. We also completed a proof of concept validation of our design process by using an optimized E4P biosensor to develop an L-PHE-hyperproducing strain by fluorescence screening over multiple rounds of mutagenesis. This work thus provides an effective pipeline for development of biosensors for biosynthetic engineering. To make it easier for readers or more biologists to understand and use this method, a complete description of the biosensor design workflow is included on the application home page. Biosensor design workflow is freely available at https://github.com/Gerald-li/BioSensor_design-workflow/.

2. Materials and methods

2.1. Bacterial strains, plasmids, cultures

The *E. coli* strains and plasmids used in the study are described in Table S2. The primers used in this study are listed in Table S3.

Biological replicates were defined as samples inoculated from distinct colonies. Growth media consisted of lysogeny broth (LB) and M9 minimal medium.

2.2. Measurement of fluorescence

A single colony was inoculated into LB and cultured overnight at 37 °C. Subsequently, the fresh culture containing M9 medium or LB medium were inoculated respectively with 1:50 or 1:100 LB precultures and incubated at 37 °C with shaking. After incubation at 37 °C, when the culture OD_{600} was ~0.2-0.4, IPTG (final concentration, 1 mM) was added to induce expression of the biosensor and various amounts of ligands were then added into individual culture samples to control the intracellular ligand concentration. Finally, the fluorescence (483 nm excitation, 525 nm emission for GFPmut2) of 200 μl of the diluted cultures was monitored via multimode microplate reader. All fluorescence was normalized using formula $\hat{Y} = (Fx\text{-}a)/(b\text{-}a)$, where $\hat{Y}$ is the normalized fluorescence at small molecule concentration x, a is the minimum response or specific control response when control molecules are designated (a is the minimum response by default, except where specified), b is the maximum response, and Fx is the actual response measured at molecule concentration x. It is important to note that all the Fx, a and b values are fluorescence normalized with cell density by measuring the absorbance at 600 nm.

2.3. LBD-GFP library construction

The LBD sequence was cloned by Gibson assembly (Bordat et al., 2015) into a p15ASI plasmid containing GFPmut2 (Cormack et al., 1996). The LBD sequence was randomized by error-prone PCR using an Instant Error-prone PCR kit V 1.3 from TIANDZ.

2.4. LBD-GFP library selections

Libraries of T0-GFP, G8-GFP and E30-GFP integrated into *E. coli* strain DH5α were subject to two rounds of fluorescence activated cell soring in a BD FACSAria IIu. For the first round, cells were grown overnight to an OD_{600} of ~1.0 in LB containing ligands (300 μM PRPP or 300 μM E4P), and cells showing the top 0.03% of fluorescence activation were collected and expanded overnight to an OD_{600} of ~1.0 in LB lacking ligands. In the second sort, cells displaying the lowest ~3.81% of fluorescence activation were collected. The sorted libraries were expanded in LB liquid culture and plated on solid LB media with resistance to chloramphenicol. Colonies from each library were clonally isolated and grown overnight in deep well plates containing 500 μl of LB with a final concentration of 30 μg/mL chloramphenicol. Candidates were diluted 1:50 into two cell culture plates with M9 media: one plate was supplemented with IPTG and ligand, the other just with IPTG. Cells were grown for another 7–9 h, and the fluorescence (483 nm excitation, 525 nm emission) of 200 μl of the cultures were monitored via a multimode microplate reader. Highest induction candidates were subject to Sanger sequencing with primers flanking the LBD sequence.

2.5. Protein and ligand preparation

The crystal structure (PDB ID: 1D6N (Balendiran et al., 1999)) of *Homo sapiens* HGPRT was extracted from the Protein Data Bank (PDB) (Burley et al., 2019). Then, Protein Preparation Wizard (PrepWizard) (Schrödinger Suite, 2018-1 Protein Preparation Wizard; Epik, 2018) was employed to prepare the protein structure using Maestro in Schrödinger, 2018 (Maestro, 2018) as follows: water molecules and inorganic molecules were removed, all hydrogen atoms were added, the H-bond assignment was optimized with exhaustive sampling, and finally, the protein structure was energy minimized using the OPLS3 force field (Harder et al., 2016). The 3PG, E4P and PEP used for protein design were treated with LigPrep, (2018). For each ligand, the lowest energy conformation was kept after the LigPrep step, and then ConfGen (Watts et al., 2010) was applied to each structure to generate no more than 100 conformers with force field minimization. All conformations were used for the RosettaMatch (Leaver-Fay et al., 2011) process.

2.6. Computational biosensor protein redesign

Computational redesign was conducted in RosettaMatch and Rosetta Enzyme Design applications. The RosettaMatch application searches a protein scaffold to find positions where a binding site for a desired small molecule target can be introduced. The target consists of a molecule and a set of desired side-chain interactions with that molecule. That interaction is specified via a constraint file (cstfile); this file can be used both for matching, and later, for enzyme design. The cstfile consists of blocks, and for each interaction between two residues (in this study, e.g. the hydrogen bond interaction between D135 on chain A and a substrate), there needs to be one block. The information in each block defines constraints between three atoms on D135 and three atoms on the substrate. Then, six parameters are specified, representing the ligand's six rigid body degrees of freedom. These parameters were given as one distance, two angles, and three dihedrals in each conformation generated by ConfGen. Based on each cstfile, RosettaMatch was employed to create one protein-conformer complex. Then, Optimize design was

implemented with the Rosetta Enzyme Design approach (Rajagopalan et al., 2014) and the following command line parameters were used: -enzdes -detect_design_interface -cut1 6.0 -cut2 10.0 -cut3 15.0 -cut4 20.0 -cst_opt -chi_min -bb_min -cst_min -cst_design -design_min_cycles 2 -lig_packer_weight 1.8 -packing:use_input_sc -packing:soft_rep_design -packing:linmen_li 10 -nstruct 5 or 100; all parameters were written in a "flagfile". In addition to these tuning parameters, to keep the substrate and enzyme within the steady interaction state, the cstfile used in RosettaMatch was also used in the process of Rosetta Enzyme Design.

3. Results and discussion

3.1. Design and construction of a PRPP biosensor

The PPP metabolite PRPP is ubiquitously found in living organisms and is used in substitution reactions with the formation of glycosidic bonds and links carbon and nitrogen metabolism. PRPP is also a useful precursor compound in the synthesis of many valuable products, such as tryptophan, vitamin B6 and purines (Fig. S1) (Lin et al., 2014; Trondle et al., 2020). In order to build a PRPP biosensor to assess its production *in vivo,* we began with HGPRT mutant F36L (Raman et al., 2004) (T0) cloned from *Homo sapiens*, which binds PRPP with nanomolar affinities, and PyrE (orotate phosphoribosyltransferase) cloned from *Mycobacterium tuberculosis*, which binds PRPP with micromolar affinities, as the LBD candidates. We then fused the T0 and PyrE genes to a GFPmut2 (*i.e.,* T0-GFP and PyrE-GFP) to construct the LBD candidate biosensors and subsequently induced their expression in *E. coli* DH5α. The PyrE-GFP fusion construct showed no change in fluorescence in response to PRPP, while T0-GFP showed a minute response (Fig. 1b), confirming that high binding affinity for selection of likely LBD is apparently quite impactful for success. Next, to determine the sensitivity of T0-GFP to different concentrations of PRPP, 0–2 mM PRPP was added to the culture medium and cells were examined. The results showed that the fluorescence intensity increased with the increase of PRPP concentration, indicating that T0 exhibited a dose-dependent response to PRPP (Fig. 1c).

We then sought to increase the sensitivity of the T0 reporter to PRPP. To this end, we randomly mutagenized T0 by error-prone PCR and subjected libraries of 10^5 transformants to two rounds of FACS, sorting alternately for high fluorescence in the presence of 300 μM PRPP and low fluorescence in their absence. After two rounds of FACS sorting, two colonies T1.0 (G16D) and T1.1 (N26H) were identified with full-length biosensors that showed improved activation by PRPP (Fig. S2). By adding ligands *in vitro*, we found that T1.0 had a linear response to PRPP (Fig. 1d) with the widest dynamic range (3-fold) and selectivity for PRPP over other metabolites (Fig. 1e). The effect of T1.0 on PRPP response was also evaluated by Constraint Network Analysis (Gohlke et al., 2017; Krüger et al., 2013) (CNAnalysis), as listed in Fig. S3.

3.2. Design of T0 template to generate biosensors for E4P

To investigate whether the T0 protein can be modified to respond to other analytes, we investigated a smaller, phosphate-containing molecule, E4P, that is also an essential intermediate in the PPP (Fig. S4) but does not induce a response by T0 (Fig. 2a). The workflow is divided into three stages (Fig. 2b). First, based on the binding mode between T0 and PRPP, the parameters of the protein template were defined for the new target molecule. This process begins with preparation of the constraint files from the initial protein template and ligand conformation files. We then specified the constraint file (cstfile) parameters for the D135 carboxyl group and the phosphate group (Fig. 2c) that were applied in one block, and included one distance (2.59 Å), two angles (87.3° and 111.8°), and three torsions (159.3°, −160.0° and −131.3°). At the same time, E4P was processed using the ligprep (LigPrep, 2018) and confgen (Watts et al., 2010) functions of schrödinger software.

The second stage is to output the qualified mutants after two rounds of computational evaluation. All conformations of E4P (64 in this case) and the cstfile were submitted to the RosettaMatch calculation process, generating 104 match results. All results were sent to Rosetta Enzyme Design for preliminary evaluation. Given the limitations on computing resources and calculation time, the pose of each match was set to quickly output the results of 5 designs during the first round of screening. From the crystal structure of the template, we knew that the binding pocket was divided into two binding regions, one with PRPP and another with PPO (3h-pyrazolo [4,3-d] pyrimidin-7-ol) (Fig. 2d, left). Taking this finding into consideration, two combined modes, one with the smallest all_cst value (total constraint score of the defined residues) and another with a binding energy value < -10.0 with the minimum all_cst value, were selected for the next round of mutation design (Fig. 2d). This rapid output screen returned two T0-E4P binding patterns, one with E4P docking extended into the PPO binding region (Fig. 2d, middle), and one with E4P located in the PRPP binding region (Fig. 2d, right). In the second round of design, each pose output was set to 100 mutants, while all other parameters remained the same.

The third stage is to analyze and evaluate the output proteins according to various criteria, so as to select the optimal mutant for subsequent experimental verification. All of the mutation models were classified according to binding energy and all_cst value. Some informative criteria for evaluation of each design included total_score, all_cst value, and ligand_binding energy value, among others (Li et al., 2018; Rajagopalan et al., 2014). In this case, the ratios of all_cst values < 1.0 in the combination modes designated as E30 (23 mutated AAs) and E89 (19 mutated AAs) were 4% and 0%, respectively, and the ratios of binding energy values < -10.0 were 0% and 18%, respectively. However, it may be possible to generate better results by increasing the proportions of various criteria for some designs by adjusting the output parameters of Rosetta Enzyme Design process. In order to verify the effectiveness of the parameters when outputs = 100, we thus selected the E30 and E89 biosensors for experimental validation because of their similar hydrogen bond networks and binding mode (see Fig. S5), in addition to their meeting the following criteria: total_score <0.0; and all_cst <1.0 with a lowest binding energy value or ligand_binding energy < -10.0 with a minimum all_cst value (Table S1).

3.3. Evaluation of the redesigned biosensors

To test whether the proteins designed by the computer could respond to E4P, we individually introduced candidate biosensors E30-GFP and E89-GFP into *E. coli* DH5α. First, *E. coli* DH5α containing E30-GFP was cultured in M9 or LB media, and the growth was uninhibited. However, when different concentrations of E4P molecules were added, we found that in M9 medium, at 1mM E4P, growth of the culture was significantly decreased, and when the concentration of E4P reached 2 mM, the strain failed to grow. In LB medium, growth was not inhibited until the E4P concentration reached 3 mM (Fig. S6). Thus, the strain showed higher tolerance to E4P in LB medium. For subsequent experiments testing the fluorescence of strains carrying E30-GFP and E89-GFP, 1 mM E4P was added to LB medium when the OD reached 0.2–0.4, and cultures were then incubated for 5–7 h prior to evaluation. The E30-GFP fluorescence expression increased in the presence of E4P compared to its fluorescence in the absence of E4P, while E89-GFP fluorescence was not increased, indicating that the E30 protein responds to E4P (Fig. 3a).

An important feature of biosensors intended for pathway engineering is their ability to detect a product with minimal activation by other related chemicals. To explore the specificity of E4P recognition by the E30 mutant, we tested a panel of molecules from the central carbon metabolism (G6P, 3PG, PEP, and PYR) and other primary metabolites (Thpp, P5H, ATP, and SHIK) for their ability to induce fluorescence expression in addition to the natural inducer of T0, PRPP, as a control. We found that the induction of E30-controlled GFP expression was much higher in response to E4P than to the other compounds tested (Fig. 3b), indicating that the E30 mutant exhibited good specificity for E4P

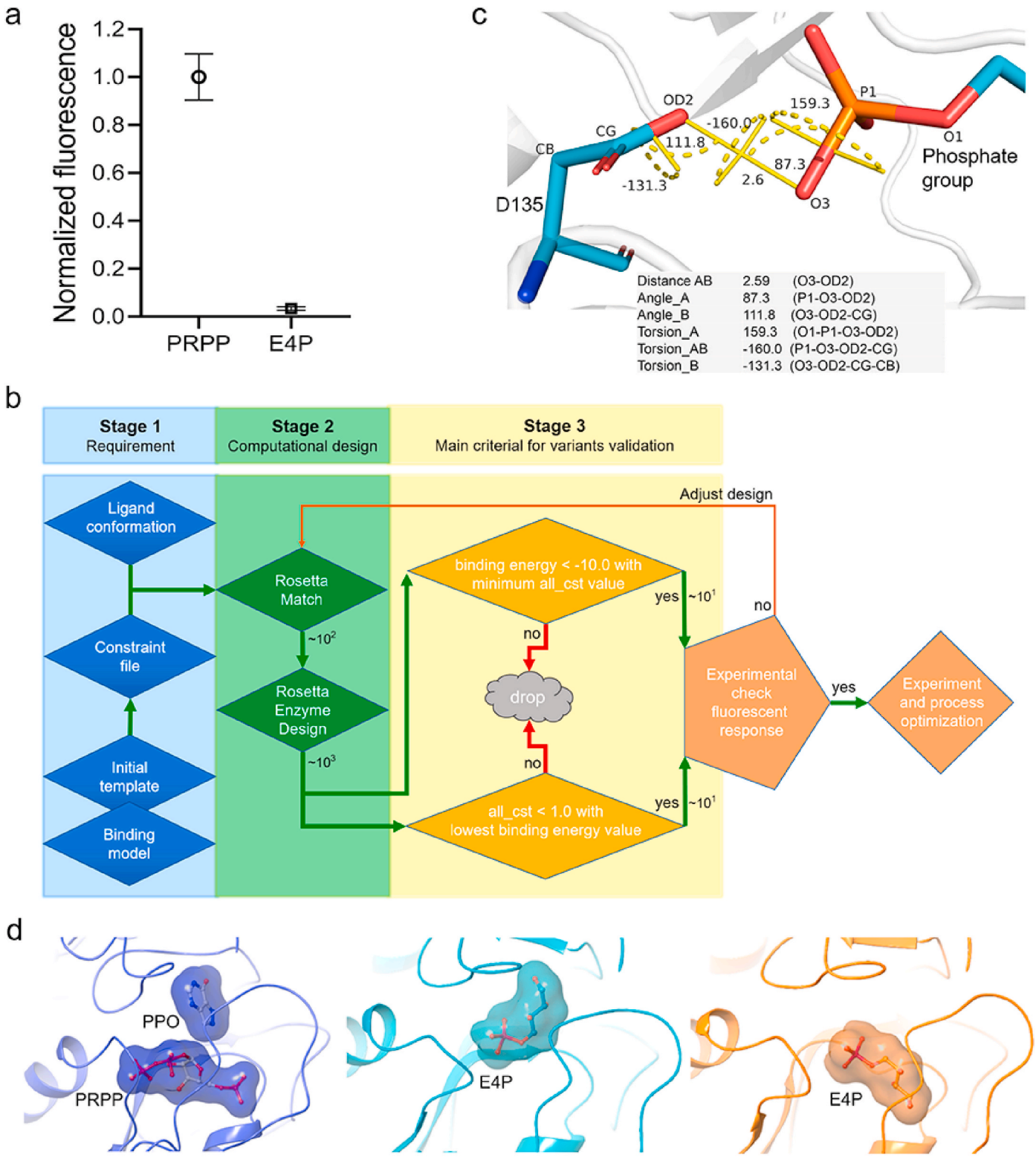

Fig. 2. Computational redesign of T0 for biosensor detection of E4P. **a** T0-GFP fluorescence response upon addition of 1 mM E4P and 2mM PRPP. **b** Schematic representation of a three-stage strategy for the computational redesign of an LBD for recognition of alternative molecules. **c** Geometric constraints of enzyme-substrate interactions for binding with E4P phosphate group (orange). The yellow lines represent distances, angles, and dihedral angles. **d** The binding patterns of T0 mutants with E4P and the binding models of PRPP and PPO in the crystal structure of 1D6N (left). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

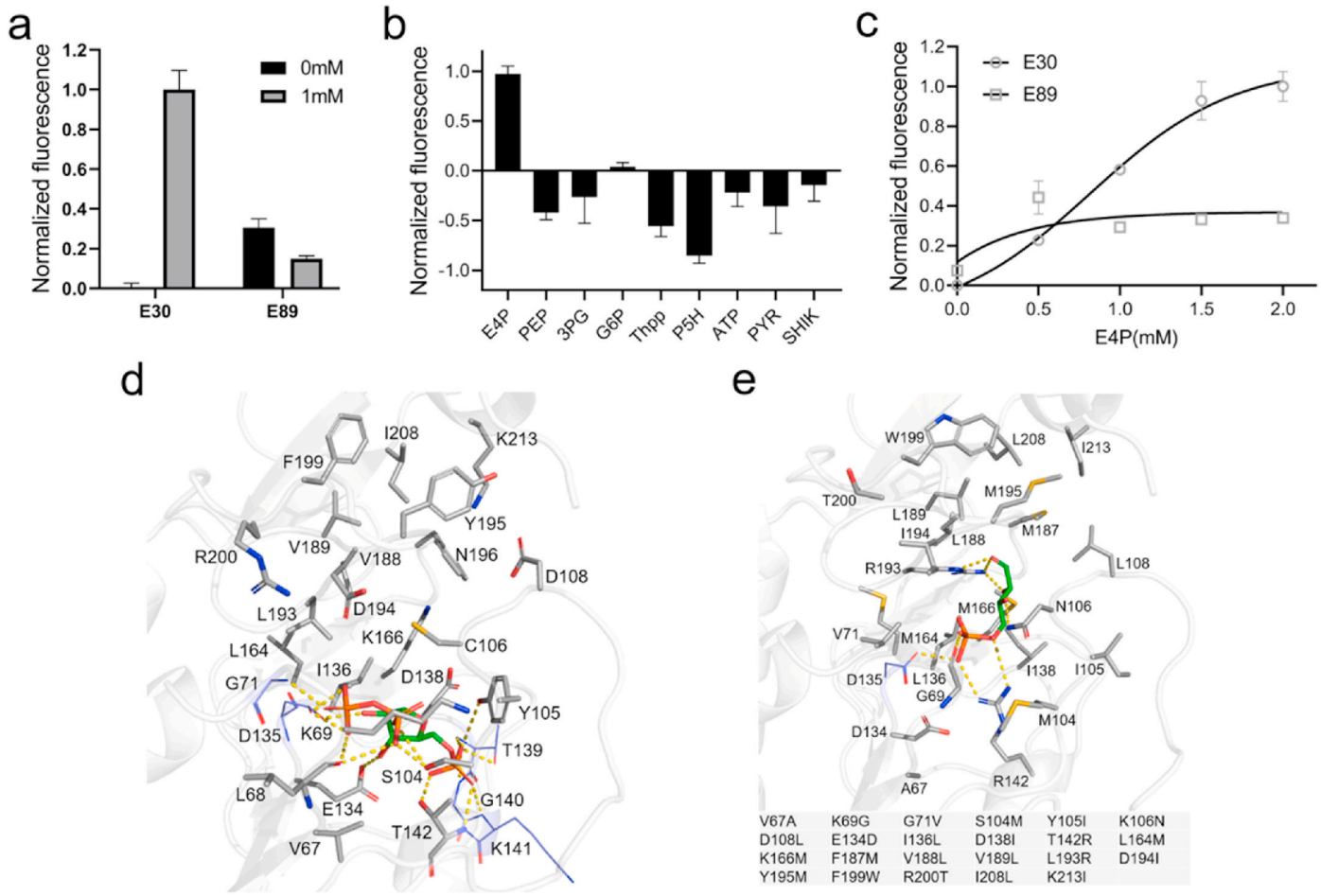

Fig. 3. Testing the redesigned protein as an *in vivo* biosensor for E4P. **a** The response of the two candidate biosensors E30-GFP and E89-GFP to E4P. **b** Ligand specificity of the E30 mutant. The corresponding response to PRPP by E30 is included for reference. **c** The concentration-dependent response by the E30 designed enzyme for the E4P target substrate. The corresponding response to E4P by E89 is included for reference. **d** The T0 binding site for PRPP, showing the native hydrogen-bonding network. **e** The Rosetta-predicted E30 binding site for E4P, highlighting disruption of the native hydrogen-bonding network and the predicted formation of a new binding pocket. Experiments were conducted in triplicate and measurements are represented as the means ± s.d.

induction.

In light of our data showing that the E30 mutant designed with the assistance of Rosetta was specific to E4P recognition, we then investigated if the response was dose dependent by exposing the *E. coli* DH5α cells carrying the E30 variant to a gradient of E4P concentrations. This experiment subsequently revealed that a range of 0–1.5 mM E4P induced a gradient increase in GFP expression (Fig. 3c), indicating that the T0 LBD template was successful for production of an E4P-specific biosensor. The E30 design contains 23 mutations, which were shown through modeling to result in E4P binding to a position in the PPO binding pocket, partially overlapping with the PRPP binding site, and which disrupted the native hydrogen-bonding network while promoting new hydrogen bonds between the E4P phosphate and carbonyl group and the surrounding residues (Fig. 3d and e). Having developed the second biosensor using the T0 template protein, we then sought to further optimize the dynamic range for applicability.

3.4. Design, evaluation, and improvement of the T0 template for building a 3PG biosensor

Encouraged by the successful design of E4P, we then explored the potential for the T0 bioreporter template to respond to analytes beyond the initial pathway. To this end, we applied the T0 template and our combined approach to develop biosensors for the 3PG and PEP metabolites of the glycolysis pathway. Compared to the E4P target analyte, both 3PG and PEP are smaller, with PEP having the smallest volume (Fig. S4). Since phosphoric acid radicals are present in these molecules, the cstfile generated for phosphoric acid with E4P was used again to perform the conformation analysis for 3PG and PEP. The RosettaMatch function generated 194 matches for 3PG and 109 matches for PEP, with five design outputs for each match. Through the quick screening process, we found that almost all total_score values for PEP were greater than zero, suggesting that these designs would be still unstable in the presence of PEP, and therefore failed the preliminary screening. We then evaluated the parameters for 3PG and determined that the ratio of total_score values < 0 was about 53.8%, and thus conducted further screening.

Finally, three combination modes were selected. These designs all interacted with the ligand in the PRPP catalytic region (Fig. 4b), although one design was biased toward the PPO binding site (Fig. 4b, left). However, unlike E30, in this model 3PG was not fully bound to the PPO binding site. One protein design was biased toward the PRPP binding site (Fig. 4b, right). However, one of the protein designs differed from the other two in that the carboxyl group oxygen bound to the same position as the phosphate group in the template protein. These structures were then selected for the next round of design, with parameters and criteria consistent with the E4P design process and screening. A total of three designs were selected from the subsequent screen for a 3PG biosensor, including G8 (18 mutated AAs), G25 (20 mutated AAs), and G74 (14 mutated AAs) (Table S1). To verify that the three redesigned proteins were responsive to 3PG, we introduced these candidate biosensors into *E. coli* DH5α, separately. Both G8 and G74 showed fluorescence expression in response to 3PG (Fig. 4c), thereby demonstrating the success of the computational design of T0 template for development of a 3PG biosensor, regardless of G25 did not exhibit any obvious GFP fluorescence. As with the previous biosensors, we found that 3PG induced the strongest response from the G8-GFP bioreporter (Fig. 4d).

In addition to substrate specificity, bioreporter sensitivity and concentration dependence for 3PG were also tested for G8 and G74. A concentration range of 0–10 mM 3PG produced a gradient increase in GFP expression by the G8 and G74 bioreporter, respectively (Fig. 4e). Modeling showed that the mutation of 18 AAs in G8 disrupted the native hydrogen-bonding network while enhancing the interactions between the 3PG carboxyl and hydroxy groups and their surrounding residues (Fig. 4f, see G74 in Fig. S7). These experiments did support the success of the computationally designed 3PG biosensors, and performance improvements were needed to adjust parameters. Then we changed output = 2500 and selected a relatively large set of 10 computational designs for experimental validation, of which seven mutants responded to 3PG (Fig. 4g). Among these seven reporter constructs, G1127 had an improved performance, particularly in dynamic range. This design carried 23 mutations, and compared with I136M mutation in G8, the I136Q transition in G1127 added a specific hydrogen bond to its interaction with 3PG (Fig. 4h, Table S1). Thus, by adjusting the parameters, we were able to further enhance the performance of designed biosensors.

3.5. Optimization and application of biosensors for screening bacterial L-PHE production

Although E30 showed a positive response to E4P, the sensitivity of the biosensor needed improvement for biological applications. To enhance the sensitivity and magnitude of fluorescent response to E4P, the E30 GFP-fusion biosensor was randomly mutagenized using error-prone PCR, cloned into the vector, and then transformed into *E. coli* DH5α. After two rounds of FACS sorting in positive/negative screens, one strain, designated E1.0, was isolated that carried a full-length biosensor which showed improved activation by E4P (Fig. S8). Assays of specific fluorescence by *E. coli* DH5α expressing the mutant E1.0 as a function of E4P concentration showed a ~3.9-fold induction of fluorescence by E4P over a concentration range of 0–2.5 mM (Fig. 5a).

Previous works showed that supply of E4P was the bottleneck for L-PHE production (Note S1). To investigate the feasibility of using the E1.0 E4P biosensor to screen for E4P hyperproducing strains which could further increase the production of L-PHE, we introduced the E1.0 construct into strain DP01, which repressed 3-deoxy-7-phosphoheptulonate synthase (*aroG* and *aroF*), and then subsequently removed plasmid p15A-3 in strain HDH6-D12 (Ding et al., 2016; Liu et al., 2017a). In order to obtain a genome-wide perturbation mutant library, DP01 strains were treated with atmospheric and room-temperature plasma (ARTP) for 14 s, according to the manufacturer's instructions (Fig. S9). In a parallel control, colonies without ARTP treatment were cultured under the same conditions for comparison.

After three rounds of mutagenesis by ARTP and FACS screening, a total of 2000 cells were finally selected by FACS, and of which 782 single colonies successfully grew on solid medium. Approximately 63.68% of the mutants showed higher fluorescence than the DP01 strain (Fig. S10), with DP01-H4 showing the highest fluorescence, supported by spectrograms of DP01 and DP01-H4 showing effective separation of the strains by screening with the E4P biosensor (Fig. 5c). To determine if screening for E4P was also successful at identifying hyperproduction of L-PHE, we also needed to confirm that L-PHE was increased in these strains. To this end, we reconstructed strain DP02 by introducing back the *aroG* and *aroF* genes, along with the co-expression plasmid p15A-3. The resulting biosynthesis of L-PHE in strain DP02 was 10.72 g/L, approximately 15.4% higher than the initial strain HDH6-D12. Furthermore, Fig. 5d shows that strain DP02 was able to efficiently produce 106 g/L of L-PHE in 5 L fed-batch fermentations, which to our knowledge is the highest yield of L-PHE reported to date.

4. Conclusion

In summary, we successfully built biosensors of three crucial metabolites from the central metabolic pathway and the three biosensors are all designed from one target protein. In addition, the concentration of the detected target molecules covers the intracellular concentration range and the magnitude of reporter response observed with the biosensors developed here is comparable with some engineered biosensors based on ligand-dependent stabilization (Brandsen et al., 2018; Feng et al., 2015; Tucker and Fields, 2001). The method presented here thus introduces a generalizable way to create sensing systems that can be used in diverse biological contexts to respond to different molecules. Combined with computational protein design, which is a rational design

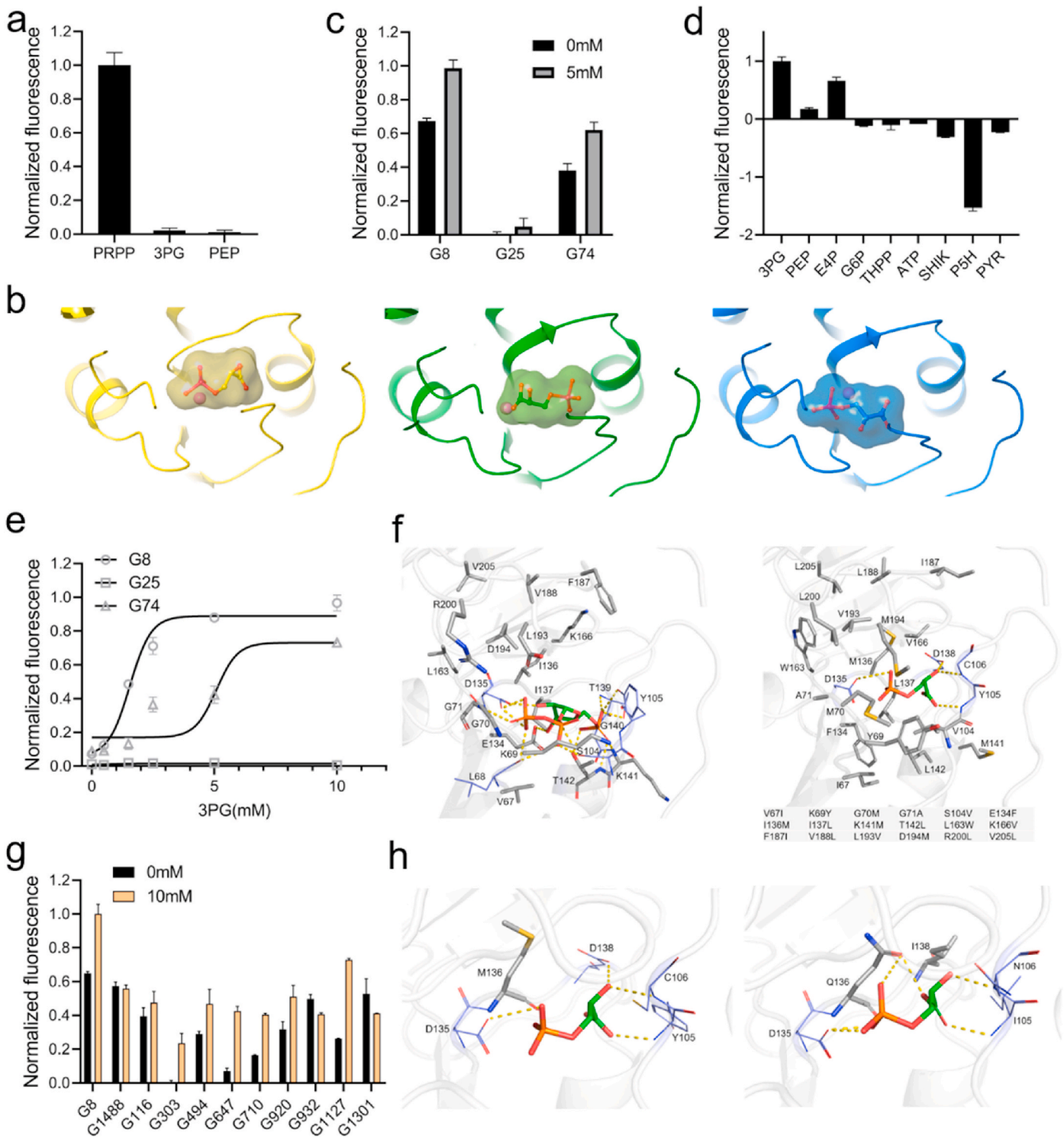

Fig. 4. Computational redesign of T0 as a 3PG responsive biosensor. **a** Fluorescence response of T0-GFP biosensors to the addition of 5 mM 3PG and 5 mM PEP. The corresponding response of T0 to PRPP (2 mM) is included for reference. **b** Three binding patterns of T0 mutants with 3PG were selected in the first round of rapid screening of the computationally redesigned protein. **c** Fluorescent response to 3PG by the three candidate biosensors G8-GFP, G25-GFP, and G74-GFP. **d** Ligand specificity of mutants G8. The corresponding response to PRPP by G8 is included for reference. **e** The concentration-dependent response to 3PG by the designed enzymes G8 and G74. The corresponding response of G25 to 3PG is included for reference. **f** Comparison of the native hydrogen-bonding network of T0 with PRPP (left) and the Rosetta-predicted binding site of G8 with 3PG (right). **g** Response to the target 3PG for 10 enzymes following the second round of computationally aided design. **h** Comparison of the first Rosetta-predicted 3PG binding site of G8 (left) with the second Rosetta-predicted 3PG binding site of G1127 (right). Experiments were conducted in triplicate and measurements are represented as the mean ± s.d.

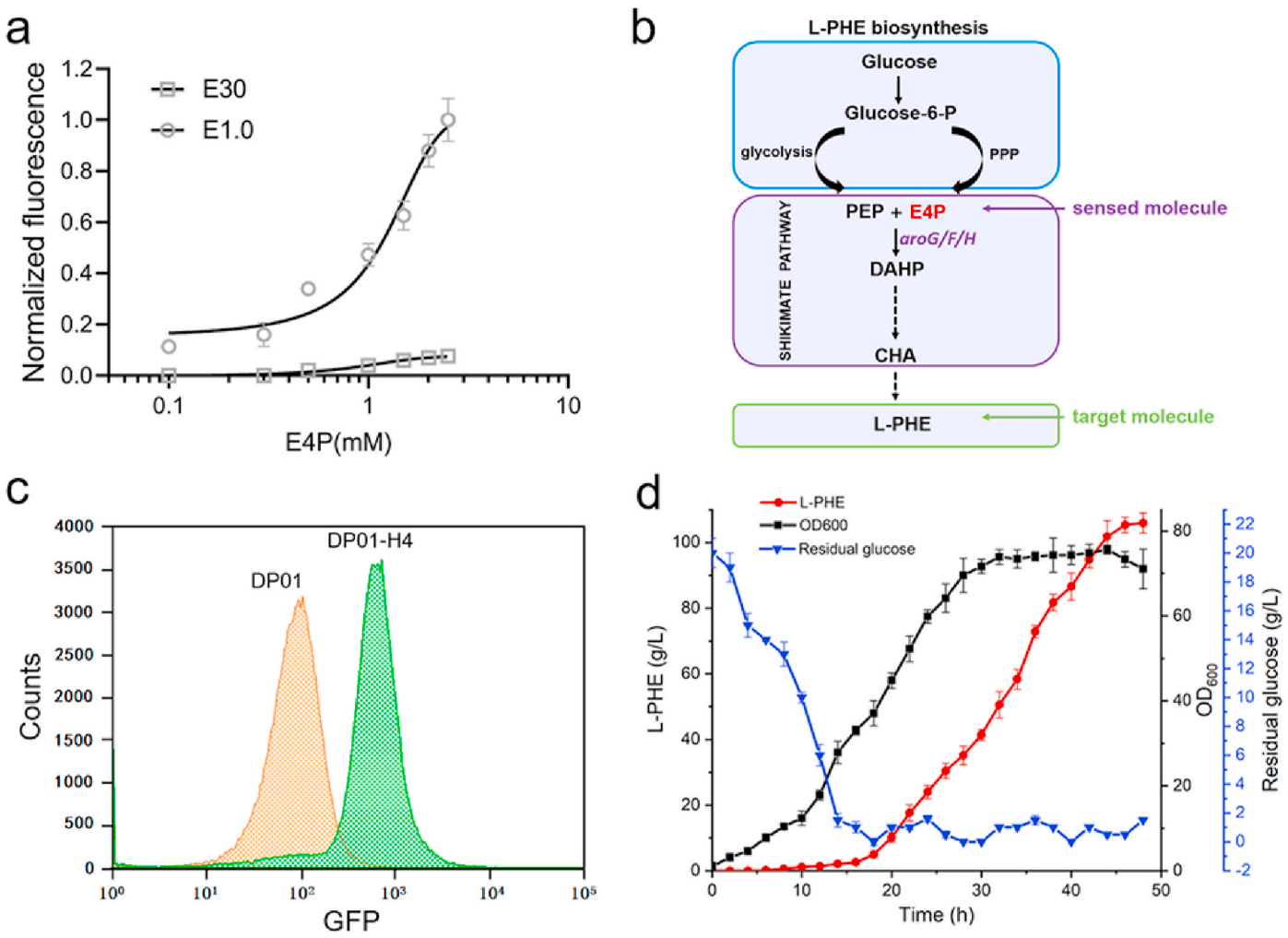

Fig. 5. E4P biosensor screening for bacterial production of L-PHE. **a** *E. coli* strain DH5α expression of GFPmut2 fluorescence through the E30 variant E1.0, as a function of E4P concentration. The corresponding dose-response curve of E30 is included for reference. **b** L-PHE is produced from glucose by converting PEP and E4P into DAHP and then on to L-PHE. DAHP (3-deoxy-D-arabino-heptulosonate 7-phosphate); CHA (chorismic acid). **c** FACS discrimination between two E4P producers (DP01 and DP01-H4) using the E4P biosensor as a screening tool. **d** Time course of L-PHE production by strain DP02 in a 5 L fermenter. Values represent means ± SD of three independent experiments.

method with large capacity, fast speed, and has been programmed and streamlined, this strategy can be widely used in the design of biosensors for molecules derived from complex metabolic pathways or genetic circuits, and will be of particular significance for improved yield of target products governed by unclear genetic mechanisms. In addition, *de novo*-designed binders opens the possibility of generating biosensors to detect molecules with unsuitable or unknown binding proteins. Furthermore, this method is also a good choice for building biosensors with multiple target molecules at the same time. After the analysis of multiple target molecules, a protein template or multiple protein templates can be selected and the computational protein design can be carried out at the same time, so as to design different mutants responding to different small molecules in a short time. Based on the existing ligand binding for protein stability biosensor construction, the method here improves the biosensor construction process and provides a new choice for building biosensor. As the performance of computational algorithms continues to improve, we foresee in silico design and screening can be combined to create binding proteins that are more responsive to target molecules, providing an increasingly powerful approach to creating a new generation of small molecule receptors for biosensor construction.

CRediT authorship contribution statement

Dongqin Ding: Conceptualization, Investigation, Methodology, Formal analysis, Data curation, Visualization, Writing - review & editing. **Jinlong Li:** Conceptualization, Software, Formal analysis, Data curation, Writing - review & editing. **Danyang Bai:** Methodology, Data curation, Validation. **Huan Fang:** Writing - review & editing, Supervision, Formal analysis. **Jianping Lin:** Writing - review & editing, Supervision, Funding acquisition. **Dawei Zhang:** Conceptualization, Methodology, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Key R&D Program of China (2018YFA0901600, 2018YFA0900300, 2018YFD0901001), the National Natural Science Foundation of China (NSFC 31800086, 31900052), the Tianjin Science Fund for Distinguished Young Scholars (17JCJQJC45300), the Science and Technology Service Network (STS) Initiative of the Chinese Academy of Sciences (CAS) (KFJ-STS-ZDTP-065), and the National Key R&D Plan Special Project for "Synthetic biology" (2018YFA0901402), Tianjin Synthetic Biotechnology Innovation Capacity Improvement Project (TSBICIP-KJGG-004-03).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bios.2020.112456.

References

Balendiran, G.K., Molina, J.A., Xu, Y., Torres-Martinez, J., Stevens, R., Focia, P.J., Eakin, A.E., Sacchettini, J.C., Craig 3rd, S.P., 1999. Protein Sci. 8 (5), 1023–1031.
Banaszynski, L.A., Chen, L.C., Maynard-Smith, L.A., Ooi, A.G., Wandless, T.J., 2006. Cell 126 (5), 995–1004.
Bick, M.J., Greisen, P.J., Morey, K.J., Antunes, M.S., La, D., Sankaran, B., Reymond, L., Johnsson, K., Medford, J.I., Baker, D., 2017. elife 6.
Bordat, A., Houvenaghel, M.C., German-Retana, S., 2015. Virol. J. 12, 89.
Brandsen, B.M., Mattheisen, J.M., Noel, T., Fields, S., 2018. ACS Synth. Biol. 7 (9), 1990–1999.
Burley, S.K., Berman, H.M., Bhikadiya, C., Bi, C., Chen, L., Di Costanzo, L., Christie, C., Dalenberg, K., Duarte, J.M., Dutta, S., Feng, Z., Ghosh, S., Goodsell, D.S., Green, R. K., Guranovic, V., Guzenko, D., Hudson, B.P., Kalro, T., Liang, Y., Lowe, R., Namkoong, H., Peisach, E., Periskova, I., Prlic, A., Randle, C., Rose, A., Rose, P., Sala, R., Sekharan, M., Shao, C., Tan, L., Tao, Y.P., Valasatava, Y., Voigt, M., Westbrook, J., Woo, J., Yang, H., Young, J., Zhuravleva, M., Zardecki, C., 2019. Nucleic Acids Res. 47 (D1), D464–D474.
Capriotti, E., Fariselli, P., Casadio, R., 2005. Nucleic Acids Res. 33, W306–W310 (Web Server issue).
Chen, Z., Boyken, S.E., Jia, M., Busch, F., Flores-Solis, D., Bick, M.J., Lu, P., VanAernum, Z.L., Sahasrabuddhe, A., Langan, R.A., Bermeo, S., Brunette, T.J., Mulligan, V.K., Carter, L.P., DiMaio, F., Sgourakis, N.G., Wysocki, V.H., Baker, D., 2019. Nature 565 (7737), 106–111.
Cormack, B.P., Valdivia, R.H., Falkow, S., 1996. Gene 173 (1 Spec No), 33–38.
Dietrich, J.A., Shis, D.L., Alikhani, A., Keasling, J.D., 2013. ACS Synth. Biol. 2 (1), 47–58.
Ding, D., Liu, Y., Xu, Y., Zheng, P., Li, H., Zhang, D., Sun, J., 2016. Sci. Rep. 6, 32208.
Egeler, E.L., Urner, L.M., Rakhit, R., Liu, C.W., Wandless, T.J., 2011. J. Biol. Chem. 286 (36), 31328–31336.
Feng, J., Jester, B.W., Tinberg, C.E., Mandell, D.J., Antunes, M.S., Chari, R., Morey, K.J., Rios, X., Medford, J.I., Church, G.M., Fields, S., Baker, D., 2015. eLife 4.
Glasgow, A.A., Huang, Y.M., Mandell, D.J., Thompson, M., Ritterson, R., Loshbaugh, A. L., Pellegrino, J., Krivacic, C., Pache, R.A., Barlow, K.A., Ollikainen, N., Jeon, D., Kelly, M.J.S., Fraser, J.S., Kortemme, T., 2019. Science (80-.) 366, 1024–1028, 6468.
Gohlke, H., Ben-Shalom, I.Y., Kopitz, H., Pfeiffer-Marek, S., Baringhaus, K.H., 2017. J. Chem. Theor. Comput. 13 (4), 1495–1502.
Harder, E., Damm, W., Maple, J., Wu, C., Reboul, M., Xiang, J.Y., Wang, L., Lupyan, D., Dahlgren, M.K., Knight, J.L., Kaus, J.W., Cerutti, D.S., Krilov, G., Jorgensen, W.L., Abel, R., Friesner, R.A., 2016. J. Chem. Theor. Comput. 12 (1), 281–296.
Huang, P.S., Boyken, S.E., Baker, D., 2016. Nature 537 (7620), 320–327.
Jiang, L., Althoff, E.A., Clemente, F.R., Doyle, L., Rothlisberger, D., Zanghellini, A., Gallaher, J.L., Betker, J.L., Tanaka, F., Barbas 3rd, C.F., Hilvert, D., Houk, K.N., Stoddard, B.L., Baker, D., 2008. Science (80-.) 319, 1387–1391, 5868.
Krüger, D.M., Rathi, P.C., Pfleger, C., Gohlke, H., 2013. Nucleic Acids Res. 41 (W1), W340–W348.
Kuhlman, B., Bradley, P., 2019. Nat. Rev. Mol. Cell Biol. 20 (11), 681–697.
Lapidoth, G., Khersonsky, O., Lipsh, R., Dym, O., Albeck, S., Rogotner, S., Fleishman, S. J., 2018. Nat. Commun. 9 (1), 2780.
Leaver-Fay, A., Tyka, M., Lewis, S.M., Lange, O.F., Thompson, J., Jacak, R., Kaufman, K., Renfrew, P.D., Smith, C.A., Sheffler, W., Davis, I.W., Cooper, S., Treuille, A., Mandell, D.J., Richter, F., Ban, Y.E., Fleishman, S.J., Corn, J.E., Kim, D.E., Lyskov, S., Berrondo, M., Mentzer, S., Popovic, Z., Havranek, J.J., Karanicolas, J., Das, R., Meiler, J., Kortemme, T., Gray, J.J., Kuhlman, B., Baker, D., Bradley, P., 2011. Methods Enzymol. 487, 545–574.
Li, H., Liang, C., Chen, W., Jin, J.M., Tang, S.Y., Tao, Y., 2017. Biosens. Bioelectron. 98, 457–465.
Li, R., Wijma, H.J.A.-O.h.o.o., Song, L., Cui, Y.A.-O.h.o.o., Otzen, M., Tian, Y., Du, J., Li, T., Niu, D., Chen, Y., Feng, J., Han, J., Chen, H., Tao, Y., Janssen, D.B., Wu, B.A.-O.h.o.o., 2018. Nat. Chem. Biol. 1552-4469 (Electronic).
LigPrep, S., 2018. LLC, New York, NY.
Lin, Z., Xu, Z., Li, Y., Wang, Z., Chen, T., Zhao, X., 2014. Microb. Cell Factories 13, 104.
Liu, Y., Zhuang, Y., Ding, D., Xu, Y., Sun, J., Zhang, D., 2017a. ACS Synth. Biol. 6 (5), 837–848.
Liu, Z., Zhang, Y.C., Jia, X.G., Hu, M.Z., Deng, Z.X., Xu, Y.C., Liu, T., 2017b. ACS Synth. Biol. 6 (4), 701–709.
Loshbaugh, A.L., Kortemme, T., 2020. Proteins 88 (1), 206–226.
Maestro, S., 2018. LLC, New York, NY.
Marshall, L.R., Zozulia, O., Lengyel-Zhand, Z., Korendovych, I.V., 2019. ACS Catal. 9 (10), 9265–9275.
Nagarajan, D., Deka, G., Rao, M., 2015. BMC Biochem. 16, 18.
Nanda, V., Koder, R.L., 2010. Nat. Chem. 2 (1), 15–24.
Netzer, R., Listov, D., Lipsh, R., Dym, O., Albeck, S., Knop, O., Kleanthous, C., Fleishman, S.J., 2018. Nat. Commun. 9 (1), 5286.
Ng, C.Y., Khodayari, A., Chowdhury, A., Maranas, C.D., 2015. Curr. Opin. Chem. Biol. 28, 105–114.
Ovchinnikov, S., Kinch, L., Park, H., Liao, Y.X., Pei, J.M., Kim, D.E., Kamisetty, H., Grishin, N.V., Baker, D., 2015. eLife 4.
Rajagopalan, S., Wang, C., Yu, K., Kuzin, A.P., Richter, F., Lew, S., Miklos, A.E., Matthews, M.L., Seetharaman, J., Su, M., Hunt, J.F., Cravatt, B.F., Baker, D., 2014. Nat. Chem. Biol. 10 (5), 386-U100.
Raman, J., Sumathy, K., Anand, R.P., Balaram, H., 2004. Arch. Biochem. Biophys. 427 (1), 116–122.
Raman, S., Rogers, J.K., Taylor, N.D., Church, G.M., 2014. Proc. Natl. Acad. Sci. U.S.A. 111 (50), 17803–17808.
Rothlisberger, D., Khersonsky, O., Wollacott, A.M., Jiang, L., DeChancie, J., Betker, J., Gallaher, J.L., Althoff, E.A., Zanghellini, A., Dym, O., Albeck, S., Houk, K.N., Tawfik, D.S., Baker, D., 2008. Nature 453 (7192), 190–195.
Schrödinger Suite, 2018 -1. Protein Preparation Wizard. Epik, S., LLC, New York, NY, 2016; Impact, Schrödinger, LLC, New York, NY, 2016; Prime, Schrödinger, LLC, New York, NY, 2018.
Shi, S., Ang, E.L., Zhao, H., 2018. J. Ind. Microbiol. Biotechnol. 45 (7), 491–516.
Siegel, J.B., Zanghellini, A., Lovick, H.M., Kiss, G., Lambert, A.R., St Clair, J.L., Gallaher, J.L., Hilvert, D., Gelb, M.H., Stoddard, B.L., Houk, K.N., Michael, F.E., Baker, D., 2010. Science (80-.) 329, 309–313, 5989.
Tinberg, C.E., Khare, S.D., Dou, J., Doyle, L., Nelson, J.W., Schena, A., Jankowski, W., Kalodimos, C.G., Johnsson, K., Stoddard, B.L., Baker, D., 2013. Nature 501 (7466), 212–216.

Trondle, J., Schoppel, K., Bleidt, A., Trachtmann, N., Sprenger, G.A., Weuster-Botz, D., 2020. J. Biotechnol. 307, 15–28.
Tucker, C.L., Fields, S., 2001. Nat. Biotechnol. 19 (11), 1042–1046.
Watts, K.S., Dalal, P., Murphy, R.B., Sherman, W., Friesner, R.A., Shelley, J.C., 2010. J. Chem. Inf. Model. 50 (4), 534–546.
Zhao, Y., Hu, Q., Cheng, F., Su, N., Wang, A., Zou, Y., Hu, H., Chen, X., Zhou, H.M., Huang, X., Yang, K., Zhu, Q., Wang, X., Yi, J., Zhu, L., Qian, X., Chen, L., Tang, Y., Loscalzo, J., Yang, Y., 2015. Cell Metabol. 21 (5), 777–789.
Zhou, L.B., Zeng, A.P., 2015. ACS Synth. Biol. 4 (6), 729–734.