Accepted Manuscript
Version of record at: https://doi.org/10.1038/s41589-022-01005-7

Dynamic modulation of enzyme activity by synthetic CRISPR–Cas6 endonucleases

Alexander A. Mitkas, Mauricio Valverde and Wilfred Chen

Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, DE, USA.
e-mail: wilfred@udel.edu

In nature, dynamic interactions between enzymes play a crucial role in defining cellular metabolism. By controlling the spatial and temporal organization of these supramolecular complexes called metabolons, natural metabolism can be tuned in a highly dynamic manner. Here, we repurpose the CRISPR–Cas6 family proteins as a synthetic strategy to create dynamic metabo-lons by combining the ease of RNA processing and the predictability of RNA hybridization for protein assembly. By disturb-ing RNA–RNA networks using toehold-mediated strand displacement reactions, on-demand assembly and disassembly are achieved using both synthetic RNA triggers and *mCherry* messenger RNA. Both direct and 'Turn-On' assembly of the pathway enzymes tryptophan-2-monooxygenase and indoleacetamide hydrolase can enhance indole-3-acetic acid production by up to ninefold. Even multimeric enzymes can be assembled to improve malate production by threefold. By interfacing with endog-enous mRNAs, more complex metabolons may be constructed, resulting in a self-responsive metabolic machinery capable of adapting to changing cellular demand.

Over the past two decades, microorganisms have been engineered to produce high-value chemicals ranging from biofuels to pharmaceuticals[1–4]. The ability to fine-tune pathway fluxes is crucial to maximizing cell productivity and product titer[5–7]. However, tedious, and often unrewarding optimization of the turnover and expression rates is required to minimize the stress inflicted by the new pathway on endogenous metabolism[8,9]. Traditional approaches based on overexpression of rate-limiting enzymes and deletion of competing pathways are mostly static in nature and do not provide dynamic regulation of pathway fluxes based on cellular resource availability[7]. A wide range of genetic circuit designs have been implemented to provide dynamic control of gene expression and pathway fluxes, resulting in substantial improvements in product titers[10–12]. Although these dynamic strategies exhibit excellent and rapid 'upregulation' control, downregulation of specific pathway flux is much slower because it requires degradation of the associated biological components such as proteins and RNAs. Ideally, post-translational modulation without altering enzyme levels can be exploited to provide on-demand flux control in a highly dynamic manner[13].

Nature has evolved an alternative strategy to regulate complex metabolic reactions by the formation of metabolons to elicit proximity control over enzyme organization[14]. Metabolons are multienzyme complexes that allow the product of one enzyme to be quickly processed by the downstream enzymes to increase overall metabolic flux[15,16]. Substrate channeling of intermediates limits their diffusion into the cytoplasm, maintains separate pools of intermediates, facilitates fast turnover of labile or toxic intermediates and prevents undesired cross-talk between different metabolic pathways. In some eukaryotes, metabolon formation can be dynamically modulated to enable rapid switching of biosynthesis profiles to prioritize cellular resources[17,18]. Considering their natural importance in redirecting metabolic resources, dynamic metabolons offer an untapped synthetic biology strategy to rapidly up- and downregulate metabolic pathways beyond altering the expression level or turnover rate of the enzymes involved. Synthetic metabolons have been created using artificial protein or DNA scaffolds to colocalize enzymes in a controlled manner for enhanced product synthesis[19–22]. Despite their success in improving product yields, current synthetic metabolons are based on static assembly and lack the ability to dynamically dis/assemble based on endogenous cellular cues[19,23].

RNA scaffolds have emerged as an attractive strategy to guide the assembly of synthetic metabolons because highly organized structures can be achieved simply by the expression of different RNA transcripts[24]. Target proteins are recruited onto these scaffolds by insertion of protein-binding RNA moieties into defined locations[25,26]. Moreover, dynamic RNA–RNA assembly can be induced using toehold-mediated strand displacement (TMSD)[27–29]. A strategy that combines the predictability of RNA hybridization and the ease of RNA processing while offering reversible metabolon assembly on demand will greatly accelerate the use of RNA scaffolds for modulating a wide range of cellular functions.

The Cas6 family of proteins are small (~25 kDa) endoribonucleases from the Type I Cascade system that bind in a sequence-specific manner to short RNA hairpin structures (20–30 bp) and cleave at the 3′ end to generate CRISPR RNAs containing a unique 5′ RNA handle[30]. After cleavage, the Cas6 protein stays bound to the hairpin with pM affinity[31–34]. By taking advantage of the hairpin specificity of two well-characterized Cas6 proteins, Csy4 from *Pseudomonas aeruginosa*[35,36] and Cse3 from *Escherichia coli*[37], we set out to generate orthogonal protein–RNA hybrids suitable for dynamic assembly (Fig. 1a). A single RNA scaffold is first used to recruit two orthogonal Cas6–enzyme fusions onto their respective hairpin. After cleavage of the RNA scaffold by the first Cas6 protein at the 3′ end, two unique enzyme–RNA complexes are generated. Spontaneous formation of an enzyme complex is achieved based on hybridization between the two complementary 5′ RNA handles. A small toehold sequence is added to the 5′ end of the hybridization sequence to drive the spontaneous TMSD reaction in the presence of an RNA trigger.

Here, we report a modular toolkit for dynamic metabolon assembly by combining the RNA processing capability of CRISPR–Cas6 proteins with TMSD. By exploiting the native function of CRISPR–Cas6 to cleave and form stable Cas6–RNA complexes[38], we successfully repurposed the Cas6 family of

Accepted Manuscript
Version of record at: https://doi.org/10.1038/s41589-022-01005-7

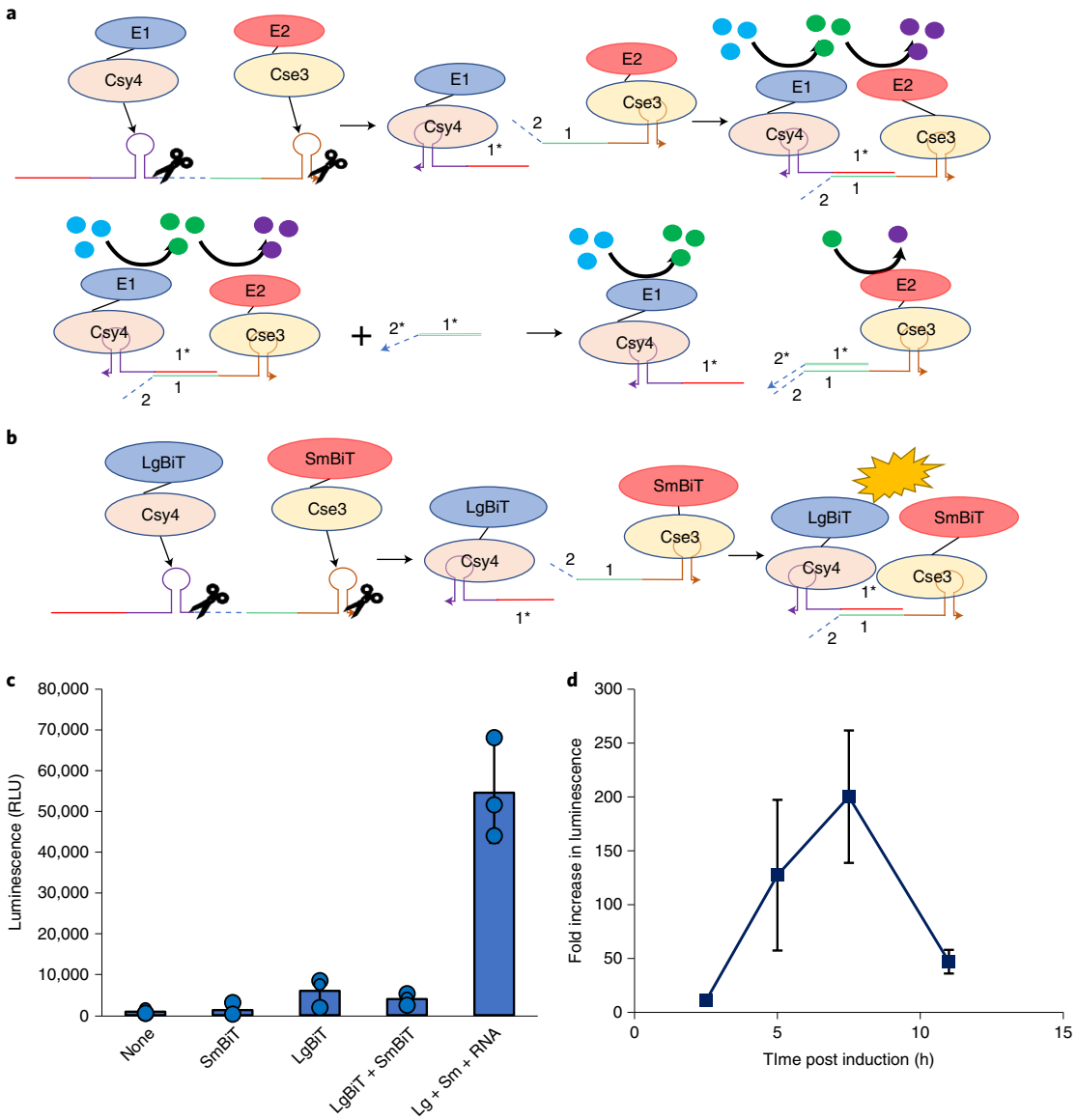


Fig. 1 | Dynamic assembly of metabolon by Cas6-mediated RNA display. a, Different enzyme–RNA hybrids are generated by Cas6-mediated binding and cleavage to drive metabolon assembly, which increases metabolic flux through the pathway by substrate channeling. Expression of a specific RNA trigger disrupts hybridization by TMSD, resulting in disassembly of metabolon. **b**, Schematic of in vivo assembly of split Nluc. **c**, Luminescence levels of different induced samples at 2.5 h post induction. All samples had the same plasmids, but different components were induced. A higher level of luminescence was detected only when both Cas6 fusions and RNA scaffold were induced. **d**, Stability of the assembled split Nluc. The fold increase in luminescence continued to increase with time. Results are presented as mean ± s.d. of three biological replicates.

proteins for site-specific RNA binding/processing and the subsequent assembly of metabolons based on flanking RNA hybridization. Dynamic assembling was made possible by RNA-driven strand displacement to enable on or off control of enzyme function and metabolite synthesis.

Results

Orthogonal Cas6 proteins for RNA-guided protein assembly. To demonstrate the use of Cas6 fusions for RNA-guided protein assembly, the split nanoluciferase (Nluc) system was first selected because of its low noise-to-signal ratio and the ability to provide real-time monitoring of protein assembly and disassembly[39]. The split Nluc is comprised of two components, the Large BiT (LgBiT) of ~18 kDa and the Small BiT (SmBiT) of ~1.3 kDa. The two components on their own demonstrate limited association and activity. When brought together by a strong interacting pair, they reconstitute into a fully functional Nluc.

The Cas6 strategy was first tested for in vitro split Nluc assembly based on hybridization between the two complementary 5′ RNA handles (Extended Data Fig. 1a). We created two Csy4 fusions, one with LgBiT (Lg) and one with SmBiT (Sm) (Extended Data Fig. 1b). Two synthetic RNA transcripts (named A or B) containing the Csy4 hairpin sequence and a flanking sequence (13 nucleotides) complementary to each other were synthesized by in vitro transcription (Extended Data Fig. 1c). Different combinations were mixed and incubated for 30 min to generate protein–RNA complexes (Extended Data Fig. 1d). Lg-A alone exhibited only a low level of luminescence, whereas Sm-B was completely inactive. Mixing Lg and Sm together resulted in a twofold increase in luminescence, indicative of low background affinity. Only mixing

Accepted Manuscript
Version of record at: https://doi.org/10.1038/s41589-022-01005-7

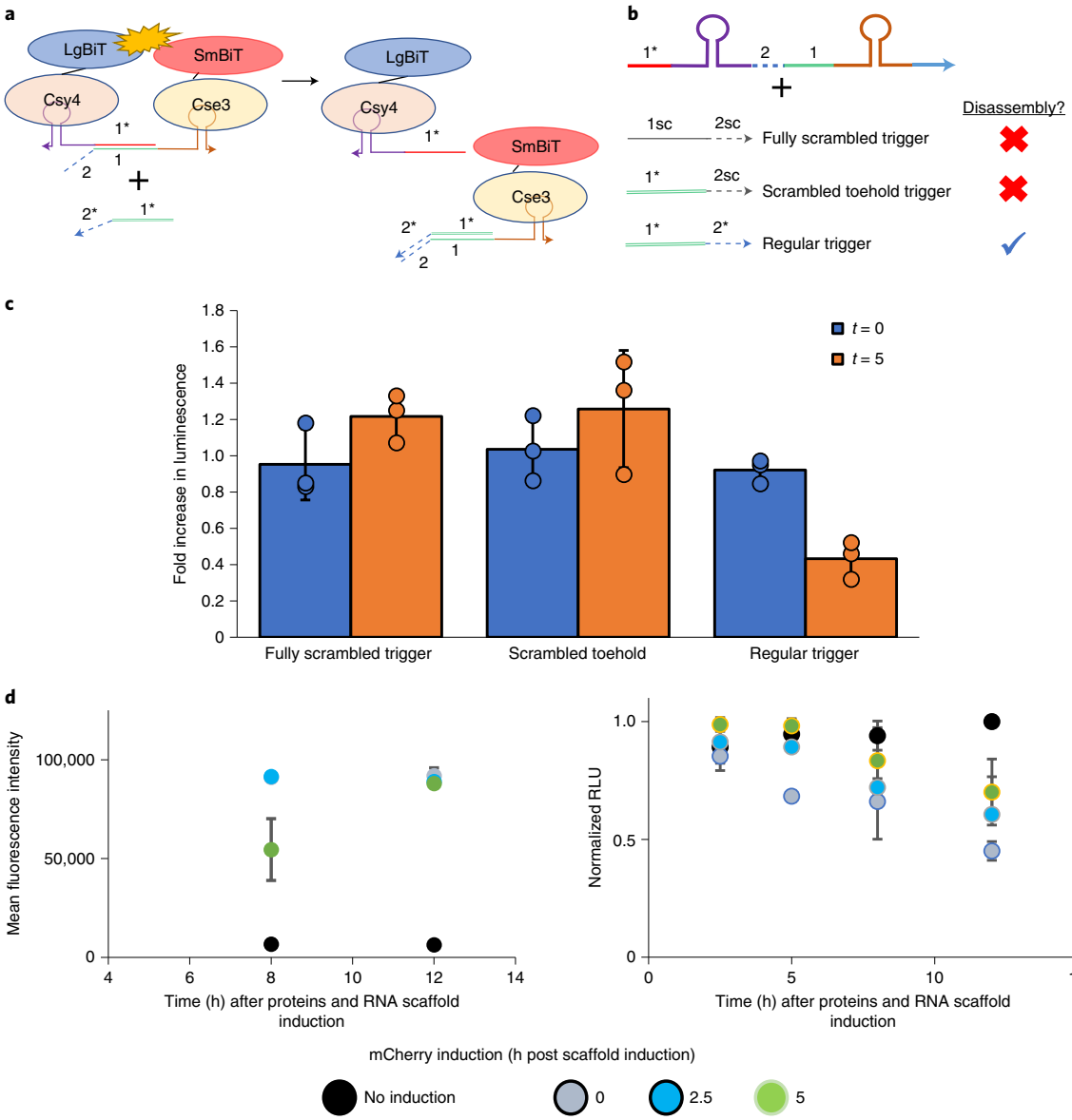


Fig. 2 | In vivo protein disassembly by TMSD. a, Mechanism of disassembly of the 'Turn-Off' design. Following assembly as shown in Fig. 1b, split Nluc can be disassembled upon induction of a trigger RNA strand. Strand displacement is initiated when the trigger strand binds to its complementary sequence (2) and branch migration allows for the formation of a new duplex. **b**, Three different trigger strands were used to demonstrate specific disassembly by TMSD. **c**, In vivo assay demonstrating split Nluc disassembly after trigger induction. Trigger was induced 2.5 h after proteins and scaffold RNA induction. Sample luminescence was measured upon trigger induction ($t = 0$) and 5 h post trigger induction ($t = 5$). The fold increase in luminescence was calculated by taking the luminescence ratio of samples with trigger RNA, scaffold RNA and proteins induced to samples with only scaffold RNA and proteins induced. **d**, Full-length *mCherry* mRNA was used as the trigger for split Nluc disassembly. Arabinose was added to induce mCherry expression at 0 h (gray), 2.5 h (blue) or 5 h (green) after proteins and RNA scaffold expression. A control (black) without mCherry induction is included for comparison. Whole-cell mCherry fluorescent and luminescent levels were measured over time after inoculation. The normalized luminescent levels were calculated as [RLU(sample) − RLU(blank with no cells)]/[RLU(maximum for the control at 12 h) − RLU(blank with no cells)]. Results are presented as mean ± s.d. of three biological replicates.

of Lg-A and Sm-B that are capable of hybridization increased luminescence by threefold over the background signal of the Lg-Sm sample, consistent with the proximity-based activity reconstitution. By contrast, mixing of Lg-A and Sm-C (an RNA transcript containing a scrambled binding sequence) only resulted in background signal. Collectively, these results clearly demonstrated the ability of Csy4-mediated protein colocalization and activity reconstitution based on specific RNA-guided assembly, a key feature necessary for our Cas6-based enzyme assembly strategy.

To adapt the strategy for in vivo assembly, Csy4–SmBiT was replaced with Cse3–SmBiT. Cse3 is an orthogonal Cas6 protein with a distinct hairpin-binding sequence, which ensured the selective colocalization of SmBiT and LgBiT after hybridization (Fig. 1b). A long scaffold RNA containing two distinct Cas6 RNA-binding hairpins and complementary 5′ hybridization sequences was used such that Csy4 and Cse3 can bind and cleave the RNA scaffold into two smaller RNA transcripts that can hybridize together to guide split Nluc assembly. A 5′ toehold is added to the Cse3 sequence to

Accepted Manuscript
Version of record at: https://doi.org/10.1038/s41589-022-01005-7

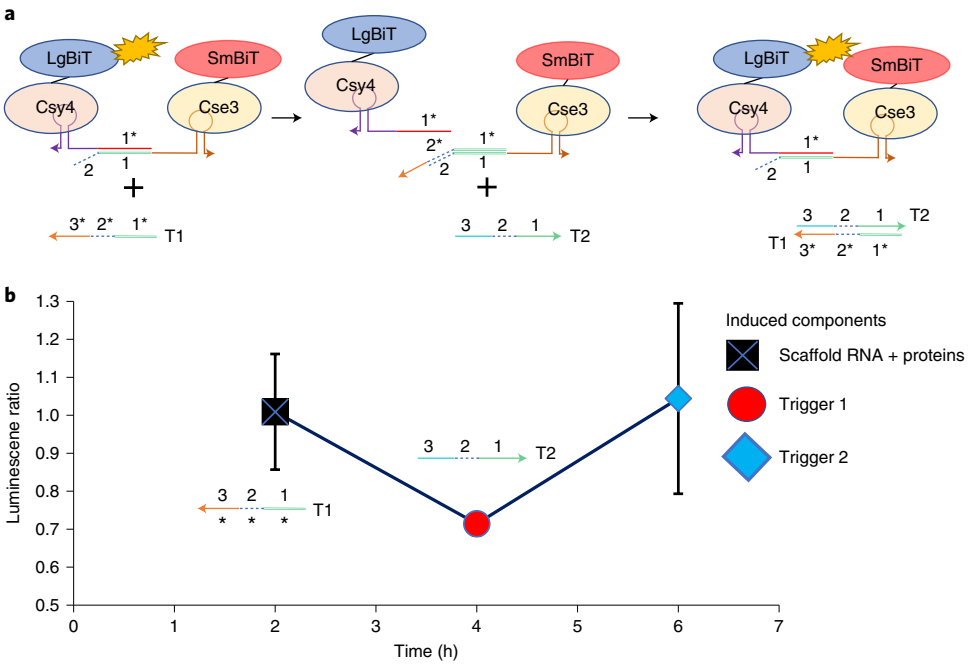


Fig. 3 | Proposed dynamic protein assembly system. a, Two different RNA triggers (T1 and T2) are used to drive cycling between the assembled and disassembled states. Induction of T1 drives protein disassembly, whereas induction of T2 drives protein assembly. **b**, Cycling of split Nluc between assembled and disassembled states. The luminescence ratio is that of a sample with cycling to a sample with only RNA scaffold and protein induction. Proteins and RNA scaffold were induced at 0 h. T1 was induced at 2 h and T2 was induced at 4 h. Results are presented as mean ± s.d. of three biological replicates.

allow potential strand displacement. Expression of individual components was achieved using a three-promoter, two-plasmid system (Extended Data Fig. 2a). Expression of either one or both fusion proteins alone (Extended Data Fig. 2b) without the RNA scaffold resulted in only background luminescence (Fig. 1c). Coexpression of the RNA scaffold increased luminescence by ninefold, indicating that all three components are required for in vivo protein assembly (Fig. 1c). Because an increase in luminescence was detected only 2.5 h after induction, this result demonstrates that the RNA scaffold was processed correctly and rapidly in vivo by the two Cas6 proteins, and sequences capable of hybridization were successful in bringing LgBiT and SmBiT together for reconstitution even in a crowded cellular environment.

To further test the stability of the assembly, the two Cas6 fusions were combined into one anhydrotetracycline (aTc)-inducible operon to allow better balance in expression (Extended Data Fig. 2c). Again, a ninefold increase in luminescence was observed rapidly at 2.5 h post induction, a level comparable with that of the three-promoter system (Fig. 1d). The luminescence level increased drastically over time, peaking at ~205-fold at 7.5 h. The sharp increase in luminescence supports continuous accumulation of all Cas6 components and the highly stable nature of the intracellular-assembled protein–RNA complex during exponential growth. The drop in luminescence at 11 h is mainly due to the transition to stationary growth, which resulted in both protein and RNA turnover. The physical tethering of the two Cas6 proteins by RNA hybridization was further confirmed by coimmunoprecipitation using an anti-his mouse antibody (Extended Data Fig. 2d). This was accomplished by replacing the his tag on Cse3–SmBiT with a FLAG tag. Both Csy4–LgBiT-his and Cse3–SmBiT-FLAG were recovered when the RNA scaffold was coexpressed. This result confirms that the increase in luminescence observed upon scaffold induction is indeed due to the colocalization of the LgBiT and the SmBiT on the RNA scaffold.

In addition to the 13-bp hybridization region, we further tested the use of longer 5′ handles for split Nluc assembly (Extended Data Fig. 3a). The two additional hybridization lengths would correspond to a greater distance between the scaffolded proteins, as well as different orientations (*cis or trans*). Although the longer hybridization lengths were less efficient, the 26-bp design is slightly better, suggesting that the proper orientations may be more important than the hybridization length (Extended Data Fig. 3b).

Dynamic protein disassembly via TMSD (Turn-Off system). Having demonstrated the in vivo Cas6-mediated protein assembly with high stability and specificity, we next tested the ability for dynamic disassembly via TMSD (Fig. 2a). Three separate trigger RNAs—a regular trigger, a fully scrambled trigger and a scrambled toehold trigger—were designed and placed under the control of a pBAD arabinose-inducible promoter (Fig. 2b). If dissociation is exclusively triggered via TMSD, we would expect a drop in luminescence only using the regular trigger sequence. After inducing production of Cas6 fusions and the scaffold RNA, the different trigger RNAs were induced 2.5 h later. In the presence of a regular trigger, a 55% drop in luminescence was observed 5 h post trigger induction (Fig. 2c). In comparison, when the fully scrambled trigger or the scrambled toehold trigger was induced, no change in luminescence was detected. The results highlight that dynamic disassembly occurs only in the presence of a correct trigger RNA and that the driving force of the disassembly is TMSD.

We next tested to see whether a full-length mRNA can act as a trigger for split Nluc disassembly. The *mCherry* mRNA was chosen as the target because we can easily follow expression on a real-time basis. We designed a new scaffold RNA by replacing the toehold and hybridization region to be responsive to the *mCherry* mRNA. After inducing expression of the Cas6 fusions and RNA scaffold, mCherry expression was turned on by arabinose addition. Although the luminescence levels continued to decrease with time independent of the timing of mCherry induction, the greatest level of reduction was observed when mCherry was coinduced with the Cas6 fusions and RNA scaffold. (Fig. 2d). More importantly, no notable decrease

Accepted Manuscript
Version of record at: https://doi.org/10.1038/s41589-022-01005-7

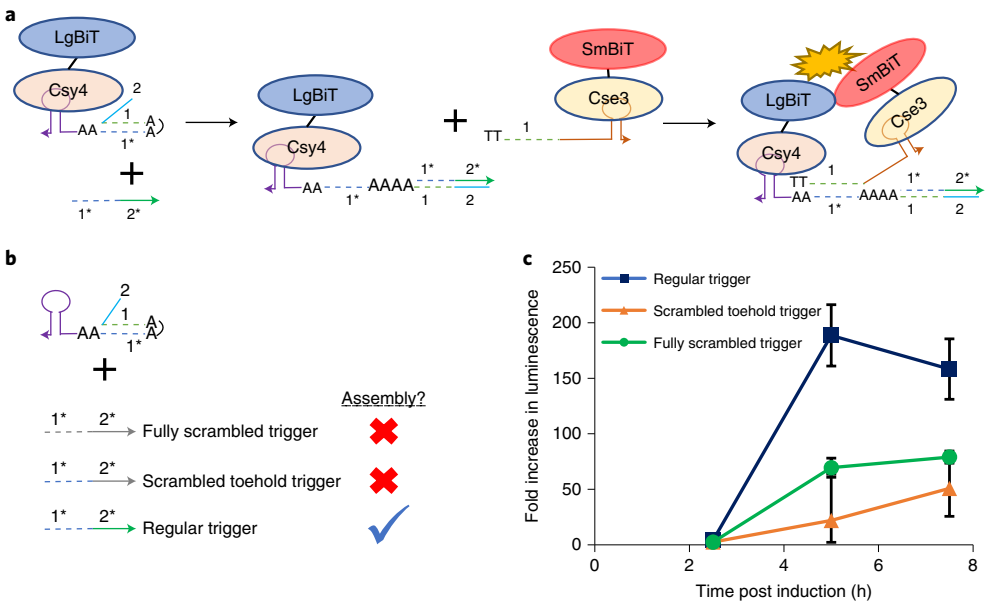


Fig. 4 | Development of a 'Turn-On' system. a, Design principle of the 'Turn-On' system. In this case, the scaffold RNA is expressed as two separate pieces, which initially cannot assemble due to blocking by a hairpin in the Csy4–RNA component. Following induction of a correct trigger strand, the blocking hairpin is unwound, thus exposing the complementary sequence 2*, upon which the Cse3 scaffold RNA can bind with its 2 region. The Cse–SmBiT scaffold RNA contains two extra nucleotides on the hybridization region to favor binding with the Csy4–LgBiT scaffold RNA rather than with the trigger. **b**, Three different trigger strands were used to determine the specificity of the 'Turn-On' system by TMSD. **c**, In vivo assay to determine split Nluc assembly for the 'Turn-On' system. All components were induced at $t = 0$ and luminescence of different samples was measured at 2.5, 5 and 7.5 h post induction. The fold increase in luminescence was calculated by taking the ratio of luminescence of a sample with trigger induction to a sample with no trigger induction. Results are presented as mean ± s.d. of three biological replicates.

in mCherry fluorescence was observed, indicating that translation is not inhibited even after the mRNAs were used as the trigger for TMSD (Fig. 2d). These results highlight the feasibility of creating a 'Turn-Off' system that is responsive to native mRNAs without compromising their endogenous cellular functions.

Although the displacement rate is slightly slower than the smaller synthetic trigger, this is expected because of the presence of more secondary structures on the longer mRNA[27]. Another potential reason for the slower displacement is probably the continuous synthesis of new RNA scaffolds, which must be displaced in addition to the already hybridized ones. This was confirmed by performing a medium exchange experiment by removing both aTc and IPTG before inducing mCherry expression with arabinose. Consistent with other in vivo TMSD results, a rapid decline in luminescence was detected within 1 h with almost complete disassembly to the background level observed within 3 h (Extended Data Fig. 4). These results demonstrated that protein disassembly based on in vivo TMSD is fast and efficient.

Dynamic protein assembly. To provide fully dynamic assembly, cycling between the assembled and disassembled states is needed. The proposed cycling system, shown in Fig. 3a, builds on the direct assembly system developed above. The trigger sequence used above for disassembly was extended by 18 nucleotides. This extension serves as a secondary toehold, upon which a second trigger (T2) can bind and sequester the primary trigger (T1). Upon this binding, the scaffold RNAs are free to reassemble. The individual triggers can be expressed sequentially to drive disassembly followed by assembly (T1 first, followed by T2). The secondary toehold sequence (3*) was designed to be longer with higher GC content than the primary toehold sequence (2) to promote rapid and strong binding of the T1 and T2 RNA sequences. The new sequences were designed to have minimum secondary structures using NUPACK[40]. The dual plasmid system is shown in Extended Data Fig. 5.

The cycling experiment was performed by cotransforming *E. coli* with the trigger plasmid and the RNA scaffold/protein plasmid. Two hours after RNA scaffold and protein induction with aTc, IPTG was added to induce T1 expression. Two hours following T1 induction, arabinose was added to induce T2 expression. A control in which only the RNA scaffold and proteins were produced without T1 and T2 induction was also included to measure the luminescence of the assembled Nluc without cycling. Data shown in Fig. 3b represent the luminescence ratio of cultures with cycling (that is, T1 and T2 induction) to that without cycling. The luminescence levels of both cultures were similar before T1 induction, suggesting little variability in the split Nluc assembly. Two hours after T1 induction, the luminescence ratio decreased to 0.7, indicating that 30% of the reconstituted split Nluc was disassembled by strand displacement. The level of scaffold disassembly aligns with the original disassembly experiments in which a 55% drop in luminescence was observed 5 h after induction of the trigger strand. More importantly, a luminescence ratio of 1 was again detected 2 h after T2 induction, indicating that the second trigger can successfully reassemble split Nluc to a similar level as the control. These results demonstrate that protein assembly can dynamically cycle between the ON and OFF states with simple toehold design using cascading trigger strands.

Dynamic protein assembly via TMSD (Turn-On system). To achieve fully dynamic assembly, the Cas6 system was redesigned to activate protein assembly upon trigger induction. The system architecture was modified by splitting the scaffold RNA into two distinct sequences (Fig. 4a). The hybridization region on the Csy4 RNA scaffold is initially sequestered by a kinetically trapped hairpin followed by a toehold sequence. The Cse3 RNA scaffold is designed to contain two extra nucleotides in the 5′ hybridization region (2) to favor binding with the Csy4 RNA scaffold rather than with the trigger. Expression of the correct trigger sequence unlocks the blocked hybridization region of the Csy4 scaffold RNA to drive protein

Accepted Manuscript
Version of record at: https://doi.org/10.1038/s41589-022-01005-7

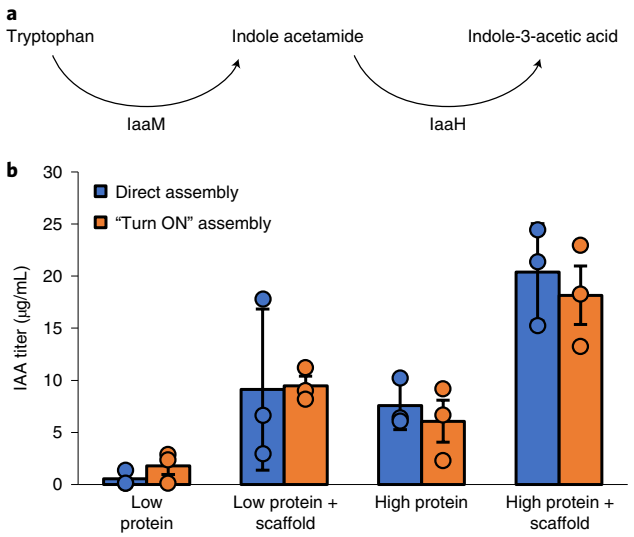


Fig. 5 | Enhancing IAA pathway titer by Cas6-mediated enzyme assembly. **a**, Production of IAA by a two-step pathway with tryptophan as the initial substrate. **b**, Production of IAA by either direct (blue) or 'Turn-On' (orange) assembly of the IaaM/IaaH metabolon. Samples had either only the Cas6 fusions expressed (Csy4–IaaH and Cse3–IaaM) or had the fusion proteins assembled into a metabolon. Two different protein expression levels were employed to evaluate their impact on pathway enhancement. Samples were harvested 16 h post induction. HPLC analysis was performed on the sample supernatant to determine the IAA titer. Results are presented as mean ± s.d. of three biological replicates.

assembly. To achieve this new design, the trigger RNA was placed under IPTG induction, whereas the two scaffold RNA molecules were expressed separately under the constitutive promoter J23100[41]. Both Cas6 fusions were still expressed via aTc induction. Three separate constructs, each expressing a different trigger RNA sequence (Fig. 4b) were created to assess the Turn-On assembly via TMSD.

Expression of the regular trigger strand resulted in a 200-fold increase in luminescence at the peak (Fig. 4c), which is comparable with the luminescence levels observed in the original scaffold configuration (Extended Data Fig. 6). However, a much lower fold increase in luminescence was observed for cells expressing either a fully scrambled trigger or a scramble toehold trigger. This indicates that the 'Turn-On' design can successfully activate protein assembly upon expression of a correct trigger sequence. The 'Turn-On' architecture can potentially overlay with the original 'Turn-Off' system to afford dynamic switching from one assembly state to the other using two orthogonal triggers. It should be noted that is some increase in luminescence even with the incorrect trigger sequences indicating that the blocking might not be perfect. This can be minimized by extending the stem region of the hairpin[27,28].

Enhanced indole-3-acetic acid synthesis by metabolon assembly. Our Cas6-mediated, RNA-driven approach for protein assembly is further exploited to create synthetic metabolons for enhanced product synthesis. The ability to dynamically modulate metabolon assembly using orthogonal RNA triggers has the potential to provide on-demand control of metabolic flux. As a model system, we designed a simple two-enzyme metabolon to control the flux through the indole-3-acetic acid (IAA) production pathway.

IAA is produced from tryptophan in a two-step enzymatic reaction (Fig. 5a). Tryptophan-2-monooxygenase (IaaM) first converts tryptophan to indole-3-acetamide, which is subsequently converted to IAA by indoleacetamide hydrolase (IaaH). The pathway was selected because it has two steps, with relatively small proteins, and use of synthetic metabolon has been shown to enhance IAA synthesis[42]. To adapt our designs to the IAA pathway, we replaced the LgBiT and SmBiT portion of the fusion proteins to generate Csy4–IaaM and Cse3–IaaH. The protein operon was again placed under aTc induction, whereas the scaffold RNA was either under IPTG induction or controlled by the constitutive promoter J23100 depending on the design architecture (Extended Data Fig. 7).

To guide formation of the IaaM/IaaH metabolon, IPTG was added to induce expression of the RNA scaffold or the trigger in the 'Turn-On' system. Two different aTc levels, low (1 ng μl^{-1}) and high (100 ng μl^{-1}), were used to ensure that protein expression has the minimum impact of enhancing IAA synthesis. Cells were cultured for 16 h, after which the supernatant was harvested and the product titer was measured via HPLC. Independent of enzyme expression levels, coexpression of the RNA scaffold greatly increased IAA titers demonstrating successful assembly of the metabolon (Fig. 5b). Although up to a ninefold increase in IAA titer was observed in the low protein expression regime, only a 2.7-fold enhancement was obtained at higher protein expression. This is consistent with other reports suggesting that enzyme concentration directly influences the substrate channeling efficiency with lower enhancement expected at higher concentrations because of enzyme crowding[43]. It should be noted that a similar IAA titer was achieved for the low protein induction culture assembling metabolon and the high induction culture without metabolon, indicating that the metabolon can effectively boost the IAA flux even at a 100-fold lower protein induction level.

More importantly, a similar level of enhancement in IAA production was observed using the 'Turn-On' design (Fig. 5b). In this case, higher IAA titers were detected only when the trigger was induced with 300 μM IPTG. The level of enhancement was 4.5- and 3-fold for the low and high protein induction level, respectively. Whereas the 'Turn-On' system had a slightly lower fold increase probably due to higher background metabolon assembly, this design when combined with the 'Turn-Off' architecture offers the potential to provide autonomous and dynamic switching of metabolic flux based on intracellular mRNA levels as indicators.

Metabolon assembly using multimeric enzymes. To demonstrate that this strategy could be used to assemble metabolons using larger, multimeric enzymes, we created a new enzyme assembly to improve malate production by rerouting flux through the glyoxylate shunt. The pyruvate dehydrogenase complex (PDH), which converts pyruvate to acetyl-CoA, and isocitrate lyase (ICL), which converts isocitrate to glyoxylate and succinate, were chosen. Because malate synthase (MS) requires both acetyl-CoA and glyoxylate to form malate (Fig. 6a), assembling PDH and ICL together would help increase the throughput of MS by creating an environment with a localized increase in the concentrations of both substrates required for the reaction to proceed forward (Fig. 6b). Because PDH is a large complex composed of three subunits—E1, E2 and E3 (ref. [44])—and ICL is a tetramer[45], successful assembly of these two enzymes allows us to demonstrate the ability to generalize the Cas6-mediated strategy for more complex enzyme cascades. More importantly, effective product titer increase with this method of scaffolding would be a promising approach to metabolic engineering via substrate clustering.

To test the design, *E. coli* cells were cotransformed with the plasmids shown in Extended Data Fig. 8. Cells were cultured in Terrific Broth (TB) supplemented with 20 g L^{-1} of glucose. The high level of glucose was added to ensure that reactions in the tricarboxylic acid cycle are driven to the forward reaction. Cultures were induced with 100 ng ml^{-1} aTc to induce protein expression. Different levels of IPTG were used to assess the impact of RNA scaffold expression on malate production. Cultures were grown for 24 h, after which the supernatant was harvested for HPLC analysis. In addition to direct

Accepted Manuscript
Version of record at: https://doi.org/10.1038/s41589-022-01005-7

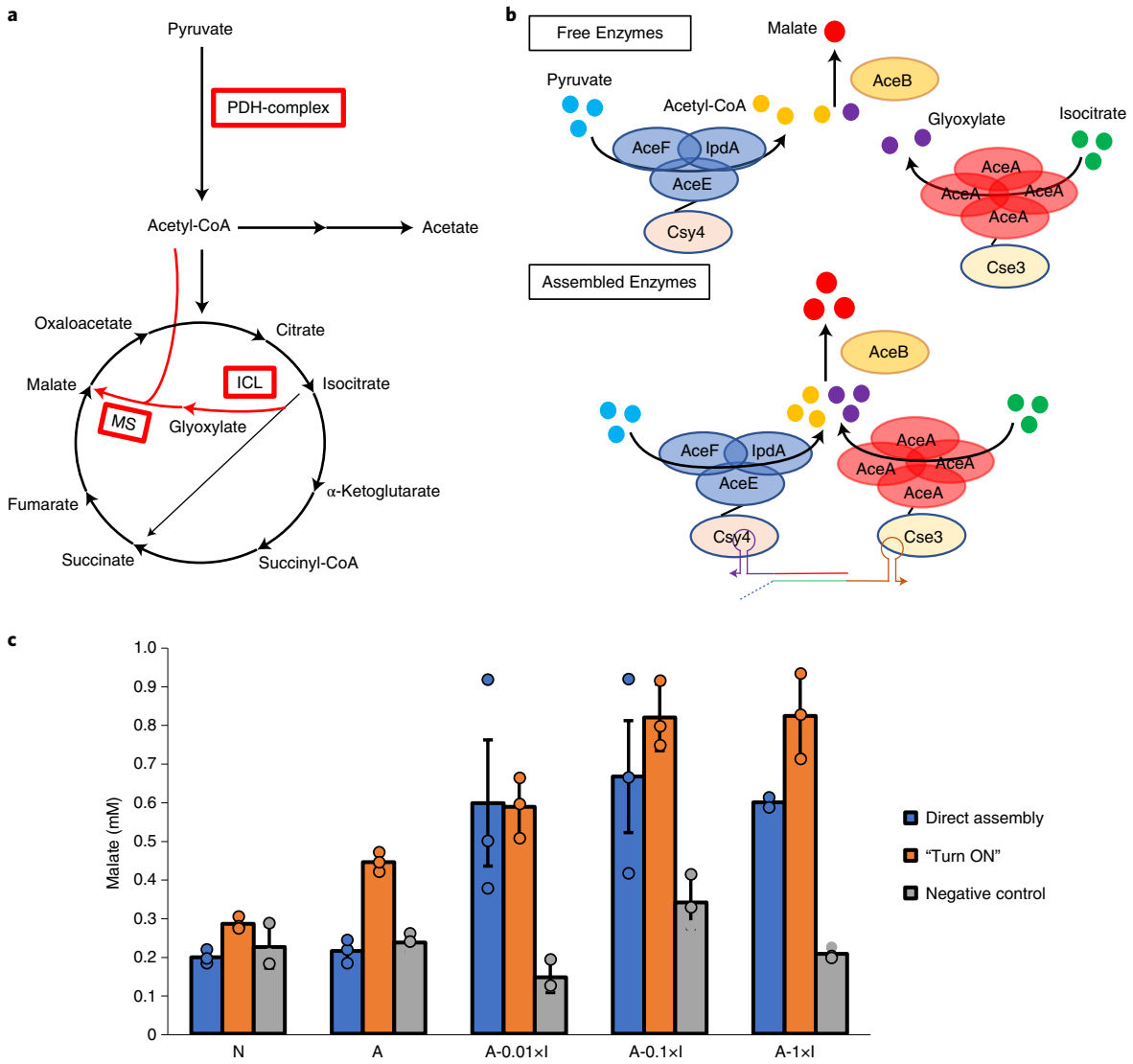


Fig. 6 | Cas6-mediated assembly of multimeric enzymes for malate production. a, The basic network of central carbon metabolism and the tricarboxylic acid cycle. The pathway of interest and the enzymes of interest are highlighted in red. **b**, The proposed mechanism for increased malate production by assembling PDH and ICL to cluster the two substrates for MS. ICL is a homotetramer, whereas PDH is a heterotrimer of AceE (E1), AceF (E2), and IpdA (E3). MS is also known as AceB. Without enzyme assembly, the local substrate concentrations of acetyl-CoA and glyoxylate are low, leading to low malate production. By contrast, enzyme assembly creates a microenvironment with higher substrate availability for MS. **c**, The malate concentration in the supernatant was measured 24 h post induction. Blue bars are for the direct assembly system, orange bars are for the "Turn ON" system, and gray bars are for the negative control of direct split Nluc assembly. Different levels of IPTG were used to vary the amount of either RNA scaffold or RNA trigger produced. N, no induction; A, 100 ng ml^{-1} aTc only for proteins induction; A-0.01×I, aTc + 3 μM IPTG for low RNA induction; A-0.1×I, aTc + 30 μM IPTG for intermediate RNA induction; A-1×I, aTc + 300 μM IPTG for high RNA induction. Results are presented as mean ± s.d. of three biological replicates except for the direct assembly at 0.1× IPTG which is from two independent experiments.

enzyme assembly, trigger-mediated assembly and a control expressing split Nluc were used for comparison.

For the direct assembly strategy, a background level of 0.2 mM of malate was observed (N) (Fig. 6c) when no fusion enzymes were expressed. When just the fusion enzymes were induced with aTc (A), the level of malate in the system was unchanged, indicating that fusion enzymes alone are not enough to drive flux through the glyoxylate shunt and increase malate production. However, when the scaffold was induced alongside the fusion enzymes, a maximum threefold increase in malate production was observed. The highest fold change was observed with the intermediate induction (0.1× IPTG or 30 μM) because excess RNA scaffold probably results in the formation of RNA hybrids containing only one of the two enzymes. More importantly, induction of RNA scaffold by itself without the fusion enzymes (control) has no impact on malate production. Collectively, these results provide strong evidence that the direct assembly system can increase malate production in vivo as expected via the proposed substrate clustering approach.

Enhancements in malate production were similarly observed using the 'Turn-On' design. A slightly higher level of malate was detected when only the enzymes were expressed. The higher level of background assembly could be a result of the unstable hairpin structure used in the 'Turn-On' design leading to a small level of background assembly. However, inducing trigger expression by IPTG still resulted in increased malate production, peaking at around 2.9-fold with the strongest induction condition (1× IPTG

or 300 μM). Even with the higher background assembly, the overall fold increase in malate production is comparable with that of the direct assembly system.

Discussion

In nature, many enzymes are frequently clustered together to form a functional metabolon to enhance the efficiency of metabolism via a process called substrate channeling[15]. In some organisms, switchable biosynthesis can be achieved in a highly dynamic manner by controlling the spatial and temporal organization of these supramolecular complexes[16,17]. Here, we demonstrate a new framework based on orthogonal Cas6 proteins to create dynamic metabolons by combining the predictability of RNA hybridization and the ease of RNA processing, while offering reversible protein assembly on demand, for metabolic pathway regulation. We successfully repurpose the unique ability of Cas6 family proteins to specifically bind and cleave long RNAs into shorter RNA scaffolds containing the desired 5′ flanking sequence[38] to generate orthogonal protein–RNA hybrids suitable for colocalization based on RNA hybridization. Selective disassembly or assembly by mRNA-driven TMSD is used to elicit dynamic modulation. Our strategy offers a simple 'plug and play' design and can be easily manipulated by changing RNA sequences or by replacing the fused proteins of interest. Because TMSD is elicited in a sequence-specific manner, the high degree of specificity for both hybridization and strand displacement results in minimal cross-talk between our synthetic designs and endogenous information.

Using split Nluc assembly as a model system, our results demonstrate that assembly is dependent on having the correct RNA 5′ hybridization regions. The simplicity of the design enables us to create RNA scaffolds to guide protein assembly with precise control. Although the effect on split Nluc assembly is minimal, the flexibility to tune protein spacing and orientation will likely have a bigger impact on the catalysis efficiency of the assembled metabolon. Addition of a flanking toehold sequence allows selective disassembly only upon induction of a correct trigger strand, highlighting the highly specific nature of the design. With only minor refinements, we successfully extend the 'Turn-Off' strategy by using the full-length *mCherry* mRNA as a trigger, suggesting the feasibility to interface native signals to synthetic cellular outputs. Although the initial disassembly using mCherry achieves a reduction in luminescence of more than 50% at 5 h post trigger induction, the slower than expected displacement kinetics is due to competition of newly synthesized RNA scaffolds for the mCherry trigger, as demonstrated by our medium exchange experiment. It may be possible to take advantage our recent toehold-gated dCas9 system[27] to concomitantly downregulate protein and RNA scaffold synthesis using the same RNA trigger to improve the disassembly kinetics. The kinetics of the displacement reaction can be further enhanced by optimizing the toehold sequence through either increasing the length and/or redesigning the sequence to confer stronger binding characteristics[46].

The flexibility of the design allows us to insert a hairpin into the Csy4 portion of the RNA scaffold to block the initial hybridization until an appropriate RNA trigger is expressed to release the blocking. By developing both a 'Turn-On' and a 'Turn-Off' system that are controlled by two orthogonal RNA triggers, we are expanding the capabilities of the Cas6 toolkit for dynamic assembly. The Turn-Off module can be used to downregulate metabolic flux in the case of toxic product accumulation to keep the concentration of the said product within acceptable levels. By contrast, the Turn-On system can be used to increase metabolic flux through protein localization only when substrate availability is at an optimum. By combining these two features, the synthetic scaffold strategy developed here has the ability to transition from the assembled to the disassembled state and vice versa.

Because the main benefit of the spatiotemporal control of protein localization is to increase metabolic flux in non-native pathways by imitating substrate channeling seen in natural metabolons, we demonstrate that both scaffolding architectures are successful in creating synthetic metabolons to substantially increase the IAA titer. Although a modest level of background split Nluc assembly is observed for the 'Turn-On' strategy, this is not the case for IAA synthesis. This is mostly likely due to the inherent background affinity between LgBiT and SmBiT in the micromolar range. Even more interesting is the fact that in the presence of the RNA scaffold, 100-fold less induction can be used to achieve a similar product titer to that in the condition with only protein expression. This indicates that Cas6-mediated assembly is highly effective in colocalizing enzymes for enhanced pathway efficiency.

Although we only demonstrated the assembly of a two-enzyme metabolon, the current framework can be expanded to larger metabolic pathways either by using longer RNA handles to recruit more enzymes using addition Cas6 proteins or by incorporating other Type I CRISPR–Cas system proteins for more complex assemblies. Ideally, this assembly approach can be used for other proximity-based cellular controls such as phosphorylation and protein degradation. These applications are currently under investigation.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41589-022-01005-7.

Received: 23 June 2021; Accepted: 25 February 2022;
Published online: 25 April 2022

References

1. Kim, B., Du, J., Eriksen, D. T. & Zhao, H. Combinatorial design of a highly efficient xylose-utilizing pathway in *Saccharomyces cerevisiae* for the production of cellulosic biofuels. *Appl. Environ. Microbiol.* **79**, 931–941 (2013).
2. Da Silva, N. A. & Srikrishnan, S. Introduction and expression of genes for metabolic engineering applications in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **12**, 197–214 (2012).
3. Westfall, P. J. et al. Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. *Proc. Natl Acad. Sci. USA* **109**, E111–E118 (2012).
4. Pitera, D. J., Paddon, C. J., Newman, J. D. & Keasling, J. D. Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*. *Metab. Eng.* **9**, 193–207 (2007).
5. Li, N., Zeng, W., Xu, S. & Zhou, J. Toward fine-tuned metabolic networks in industrial microorganisms. *Synth. Syst. Biotechnol.* **5**, 81–91 (2020).
6. Ferreira, R. et al. Model-assisted fine-tuning of central carbon metabolism in yeast through dCas9-based regulation. *ACS Synth. Biol.* **8**, 2457–2463 (2019).
7. Toya, Y. & Shimizu, H. Flux controlling technology for central carbon metabolism for efficient microbial bio-production. *Curr. Opin. Biotechnol.* **64**, 169–174 (2020).
8. Tai, M. & Stephanopoulos, G. Engineering the push and pull of lipid biosynthesis in oleaginous yeast *Yarrowia lipolytica* for biofuel production. *Metab. Eng.* **15**, 1–9 (2013).
9. Lan, E. I. & Liao, J. C. ATP drives direct photosynthetic production of 1-butanol in cyanobacteria. *Proc. Natl Acad. Sci. USA* **109**, 6018–6023 (2012).
10. Xu, P., Li, L., Zhang, F., Stephanopoulos, G. & Koffas, M. Improving fatty acids production by engineering dynamic pathway regulation and metabolic control. *Proc. Natl Acad. Sci. USA* **111**, 11299–11304 (2014).
11. Zhang, F., Carothers, J. M. & Keasling, J. D. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat. Biotech.* **30**, 354–359 (2012).
12. Win, M. N. & Smolke, C. D. A modular and extensible RNA-based gene-regulatory platform for engineering cellular function. *Proc. Natl Acad. Sci. USA* **104**, 14283–14288 (2007).
13. Venayak, N., Anesiadis, N., Cluett, W. R. & Mahadevan, R. Engineering metabolism through dynamic control. *Curr. Opin. Biotechnol.* **34**, 142–152 (2015).

14. Agapakis, C. M., Boyle, P. M. & Silver, P. A. Natural strategies for the spatial optimization of metabolism in synthetic biology. *Nat. Chem. Biol.* **8**, 527–535 (2012).
15. Menard, L., Maughan, D. & Vigoreaux, J. The structural and functional coordination of glycolytic enzymes in muscle: evidence of a metabolon? *Biology* **3**, 623–644 (2014).
16. Jørgensen, K. et al. Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. *Curr. Opin. Plant Biol.* **8**, 280–291 (2005).
17. Møller, B. L. Dynamic metabolons. *Science* **330**, 1328–1329 (2010).
18. Laursen, T., Møller, B. L. & Bassard, J.-E. Plasticity of specialized metabolism as mediated by dynamic metabolons. *Trends Plant Sci.* **20**, 20–32 (2015).
19. Dueber, J. E. et al. Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotech.* **27**, 753–759 (2009).
20. Conrado, R. J., Varner, J. D. & DeLisa, M. P. Engineering the spatial organization of metabolic enzymes: mimicking nature's synergy. *Curr. Opin. Biotechnol.* **19**, 492–499 (2008).
21. Chen, R. et al. Biomolecular scaffolds for enhanced signaling and catalytic efficiency. *Curr. Opin. Biotechnol.* **28**, 59–68 (2014).
22. Berckman, E. A. & Chen, W. Exploiting dCas9 fusion proteins for dynamic assembly of synthetic metabolons. *Chem. Commun.* **55**, 8219–8222 (2019).
23. Zhao, E. M. et al. Light-based control of metabolic flux through assembly of synthetic organelles. *Nat. Chem. Biol.* **15**, 589–597 (2019).
24. Chen, R. P., Hunt, V. M., Mitkas, A. A., Siu, K.-H. & Chen, W. Controlling metabolic flux by toehold-mediated strand displacement. *Curr. Opin. Biotechnol.* **66**, 150–157 (2020).
25. Sachdeva, G., Garg, A., Godding, D., Way, J. C. & Silver, P. A. In vivo co-localization of enzymes on RNA scaffolds increases metabolic production in a geometrically dependent manner. *Nucleic Acids Res.* **42**, 9493–9503 (2014).
26. Delebecque, C. J., Lindner, A. B., Silver, P. A. & Aldaye, F. A. Organization of intracellular reactions with rationally designed RNA assemblies. *Science* **333**, 470–474 (2011).
27. Siu, K. H. & Chen, W. Riboregulated toehold-gated gRNA for programmable CRISPR–Cas9 function. *Nat. Chem. Biol.* **15**, 217–220 (2019).
28. Green, A. A., Silver, P. A., Collins, J. J. & Yin, P. Toehold switches: de-novo-designed regulators of gene expression. *Cell* **159**, 925–939 (2014).
29. Chen, R. P., Blackstock, D., Sun, Q. & Chen, W. Dynamic protein assembly by programmable DNA strand displacement. *Nat. Chem.* **10**, 474–481 (2018).
30. Carte, J., Wang, R., Li, H., Terns, R. M. & Terns, M. P. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.* **22**, 3489–3496 (2008).
31. Niewoehner, O., Jinek, M. & Doudna, J. A. Evolution of CRISPR RNA recognition and processing by Cas6 endonucleases. *Nucleic Acids Res.* **42**, 1341–1353 (2014).
32. Plagens, A. et al. In vitro assembly and activity of an archaeal CRISPR–Cas type I-A cascade interference complex. *Nucleic Acids Res.* **42**, 5125–5138 (2014).
33. van der Oost, J., Westra, E. R., Jackson, R. N. & Wiedenheft, B. Unravelling the structural and mechanistic basis of CRISPR–Cas systems. *Nat. Rev. Microbiol.* **12**, 479–492 (2014).
34. Sokolowski, R. D., Graham, S. & White, M. F. Cas6 specificity and CRISPR RNA loading in a complex CRISPR–Cas system. *Nucleic Acids Res.* **42**, 6532–6541 (2014).
35. Sternberg, S. H., Haurwitz, R. E. & Doudna, J. A. Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. *RNA* **18**, 661–672 (2012).
36. Makarova, K. S. et al. Evolution and classification of the CRISPR–Cas systems. *Nat. Rev. Microbiol.* **9**, 467–477 (2011).
37. Jore, M. M. et al. Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat. Struct. Mol. Biol.* **18**, 529–536 (2011).
38. Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR–Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014).
39. Hall, M. P. et al. Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. *ACS Chem. Biol.* **7**, 1848–1857 (2012).
40. Zadeh, J. N. et al. NUPACK: analysis and design of nucleic acid systems. *J. Comput. Chem.* **32**, 170–173 (2011).
41. Meng, H. et al. Quantitative design of regulatory elements based on high-precision strength prediction using artificial neural network. *PLoS ONE* **8**, e60288 (2013).
42. Zhu, L.-y et al. Spatial organization of heterologous metabolic system in vivo based on TALE. *Sci. Rep.* **6**, 26065 (2016).
43. Ellis, G. A. et al. Artificial multienzyme scaffolds: pursuing in vitro substrate channeling with an overview of current progress. *ACS Catal.* **9**, 10812–10869 (2019).
44. Škerlová, J., Berndtsson, J., Nolte, H., Ott, M. & Stenmark, P. Structure of the native pyruvate dehydrogenase complex reveals the mechanism of substrate insertion. *Nat. Commun.* **12**, 5277 (2021).
45. Britton, K. L. et al. The structure and domain organization of *Escherichia coli* isocitrate lyase. *Acta Crystallogr. D* **57**, 1209–1218 (2001).
46. Srinivas, N. et al. On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Res.* **41**, 10641–10658 (2013).

Accepted Manuscript
Version of record at: https://doi.org/10.1038/s41589-022-01005-7

Methods

Strains used and plasmid construction. All strains and plasmids used for this study are listed in Supplementary Table 1.

The plasmids used for the in vitro assay containing the fusion proteins Csy4–LgBiT or Csy4–SmBiT were created using the high-copy pET29b plasmid backbone. The plasmid contained a pL*lac*O-1 promoter. Plasmids were transformed into BL21-Gold(DE3) cells for expression. The Csy4 gene and the LgBiT were amplified from pre-existing stock in the laboratory. The SmBiT sequence was created by designing overlapping oligonucleotides, which were synthesized and purified by a commercial vendor (Integrated DNA Technologies). The overlapping oligonucleotides were treated with polynucleotide kinase (PNK) and annealed. The pieces (backbone, Csy4, LgBiT/SmBiT) were assembled into a fully functional plasmid with traditional subcloning techniques using unique enzyme restriction sites.

The plasmids used in the in vivo assay containing the pL*tet*O-1 and pBAD promoters were created using a low-copy p15a plasmid backbone. Two separate constructs were used, one containing the pL*tet*O-1 scaffold RNA sequence and one containing the pBAD–Csy4–LgBiT fusion protein. The two pieces were then ligated together along with the backbone plasmid to form the full plasmid. The Cse3–SmBiT containing plasmid was constructed from the in vitro assay Csy4–SmBiT plasmid by replacing the Csy4 gene sequence with the Cse3 one using traditional subcloning techniques. The Cse3 gene was amplified from a gene sequence synthesized and purified by a commercial vendor (Integrated DNA Technologies). To create the pL*tet*O-1 Cse3–SmBiT–Csy4–LgBiT plasmid the two fusion proteins genes were amplified from previously constructed plasmids and ligated with the p15a backbone in a three-piece ligation using standard molecular cloning techniques. All the scaffold RNA plasmids were created using the high-copy pET29b plasmid backbone with the pL*lac*O-1 promoter. The scaffold RNA sequences were created using overlapping oligonucleotides which were synthesized and purified by a commercial vendor (Integrated DNA Technologies), PNK treated, annealed then added into the backbone via traditional molecular cloning techniques. The remainder of the in vivo assay plasmids were created using the same approaches illustrated above.

To generate the cycling system constructs, the plasmid architecture was modified. The system consisted of the J23113 constitutive promoter with the protein operon downstream of it and the aTc promoter with the scaffold RNA. To create this plasmid, a five-piece Gibson assembly approach was used. The five pieces used were the backbone containing pL*Tet*O promoter, single-cell RNA (scRNA) which was assembled using oligonucleotides treated with PNK (New England Biolabs (NEB)), the constitutive promoter J23113 flanked by two DNA linker sequences and the protein operon (Cse3–SmBiT–Csy4–LgBiT). All the fragments were mixed at equimolar ratios of 0.25 pmol and incubated with the NEBuilder HiFi DNA Assembly Master Mix (NEB). Unique restriction cut sites flanking both the scRNA sequences and the protein sequences were retained to allow for easy substitution of different components. To generate a plasmid containing the trigger sequences the ampicillin-resistant plasmid containing pL*lac*O-1 and pBAD was used as the template. Using simple restriction digests, the sequences downstream of the two promoters were replaced by the T1–T2 trigger pair. The trigger sequences were created using overlapping oligonucleotides which were synthesized and purified by a commercial vendor (Integrated DNA Technologies), PNK treated, annealed then added into the backbone via traditional molecular cloning techniques.

To generate the direct assembly RNA plasmid, a three-piece ligation was performed. The backbone used was plasmid CR564 with the arabinose promoter sequence excised. The two scRNA hairpins with their constitutive promoters were created using annealed oligonucleotides treated with PNK (NEB). The two inserts were added at equimolar ratios at a 1:8 ratio to the vector and the reaction was incubated at 25 °C for 1 h, 20 °C for 1 h and 16 °C for 24 h. All remaining plasmids used were created using conventional cloning techniques.

For the construction of the IaaM/IaaH plasmid, IaaM and IaaH genetic sequences were obtained from the KEGG online gene database and ordered as gBlocks (Integrated DNA Technologies). The gBlocks were then cloned into the protein plasmid backbone using standard cloning techniques. For the construction of the AceE/AceA plasmid, the gene sequences were amplified from *E. coli* chromosomal DNA and once again cloned into the protein plasmid backbone using standard cloning techniques (restriction digest and ligation). Cloning had to be performed in two steps. First, the SmBiT sequence was replaced with the IaaH or AceA sequence. Following confirmation of those plasmid constructs, the LgBiT sequence was replaced with the IaaM/AceE sequence. All plasmid constructs were confirmed via either colony polymerase chain reaction or restriction digest confirmation (or both).

In vitro scaffold hybridization assay. The Csy4–LgBiT and Csy4–SmBiT fusion proteins were expressed from the corresponding plasmids in *E. coli*. The cultures were grown in TB (12 g l^{-1} tryptone, 24 g l^{-1} yeast extract, 0.4% v/v glycerol, 0.017 M KH_2PO_4 and 0.072 M K_2HPO_4). Expression was induced with 500 µM IPTG at an optical density (OD) of ~0.9. Following induction, the cultures were grown at 20 °C for 16 h. The cells were harvested by centrifugation (3,000 r.p.m. for 20 min) and resuspended to a final OD of 15 in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4). The cells were lysed with a standard sonication protocol and then centrifuged (30,000 r.p.m. for 15 min) to separate the soluble and insoluble phases. The fusion proteins were purified using a standard histidine tag purification protocol with Ni-NTA His•Bind Resin (MilliporeSigma). Protein concentration was measured using a standard Bradford assay and protein purity was assessed by running samples using PAGE with a 7.5% SDS acrylamide gel.

The scaffold RNAs (A, B, C) were synthesized in vitro using oligonucleotides that were synthesized and purified by a commercial vendor (Integrated DNA Technologies). Transcription was performed using the HiScribe T7 Quick High Yield RNA Synthesis Kit (NEB) and purified by standard phenol–chloroform extraction and ethanol precipitation. RNA purity and quality were analyzed by spectrophotometry on a NanoDrop 2000 UV–Vis spectrophotometer (ThermoFisher Scientific) and PAGE using a 8 M urea 10% polyacrylamide gel.

For the in vitro assay, the purified proteins and scaffold RNAs were mixed in equimolar ratios (10 nm) in a Costar black 96-well plate and incubated for 15 min at 37 °C while shaken in binding buffer (100 mM NaCl, 50 mM Tris–HCl, 10 mM $MgCl_2$, 100 µg ml^{-1} BSA, pH 7.9). Following incubation, luminescence was measured according to NanoGlo vendor's instructions (Promega) using a Synergy H4 Hybrid microplate reader. The data was analyzed using Gen5 software (BioTek Instruments).

In vivo scaffold hybridization assay. Depending on the experiment performed, the corresponding plasmids were transformed into BL21-Gold(DE3) *E. coli* cells via heat shock and plated on LB medium (10 g l^{-1} tryptone, 5 g l^{-1} yeast extract, 10 g l^{-1} NaCl, 15 g l^{-1} agar) with ampicillin (100 µg ml^{-1}) and chloramphenicol (25 µg ml^{-1}) antibiotics for selection. Colonies were picked from the agar plates and cultured overnight in TB supplemented with 100 µg ml^{-1} carbenicillin and 25 µg ml^{-1} chloramphenicol at 37 °C. The following day these cultures were used to inoculate subcultures at an OD of ~0.05. The subcultures were grown to an OD of ~1.2 at which time they were induced with the appropriate inducer depending on the promoter system. For the pL*lac*O-1 promoter, 300 µM IPTG was used, for the pL*tet*O-1 promoter, 1 ng ml^{-1} of anhydrotetracycline was used, and for the pBAD promoter, 1% w/v arabinose was used. The pL*lac*O-1 promoter controlled scaffold RNA expression, the pL*tet*O-1 promoter controlled protein expression and the pBAD promoter controlled trigger expression in the original system. For the Turn-On system, the scaffold RNA was expressed constitutively using the j23100 promoter and the trigger was expressed using the pL*lac*O-1 promoter. Depending on the experiment, time points were taken at 2.5, 5 and 7.5 h.

For each time point, the OD of the cultures was measured first after samples were taken, and live cells were harvested by centrifuging at 2,500 r.p.m. for 5 min. The samples were resuspended at an OD of 2.0 in 200 µl of PBS. Then 100 µl of each sample was combined with 100 µl of the NanoGlo buffer/substrate mix in a 96-well Costar black plate. Luminescence was measured according to NanoGlo vendor's instructions for 45 min (Promega) using a Synergy H4 Hybrid microplate reader. Data were analyzed using Gen5 software (BioTek Instruments).

Scaffold dissociation and media exchange assay. For the scaffold dissociation experiments, the method used up to induction of the proteins and scaffold RNA was identical to that outlined for scaffold assembly. The plasmids used for the scaffold dissociation assays are listed in Supplementary Table 1. These were cotransformed with CR424. Two and a half hours after the protein (IPTG) and scaffold RNA (aTc) were induced, the trigger RNA was induced using arabinose (1% w/v). Luminescence was measured 5 h after trigger induction, following the method outlined in the in vivo scaffold hybridization assay section.

For the scaffold dissociation experiments using mRNA as the trigger sequence, plasmid CR652 was used. The same process was followed for induction, except in this case the mCherry sequence was induced at 0, 2.5 or 5 h following protein and scaffold RNA induction (again with 1% arabinose w/v). Luminescence was measured using previously outlined methods. For the fluorescence measurements, 50 µl of the samples (at an OD of 2.0) was combined with 150 µl of PBS in different wells of a Costar black 96-well plate to a final OD of 0.5. Fluorescence was measured using a Synergy H4 Hybrid microplate reader with excitation at 580 nm. The reported fluorescence values represent the peak emission signal observed at 610 nm. The data were also analyzed using Gen5 software (BioTek Instruments).

For the media exchange experiments, the same initial method was followed as outlined above. At 2.5 h after induction of the proteins and scaffold RNA, the first time point luminescence measurement was taken. Samples were pelleted by centrifugation at 2,500 r.p.m. for 5 min, washed once with PBS and pelleted again. After that, one sample was resuspended in fresh TB media with 100 µg ml^{-1} ampicillin and 25 µg ml^{-1} chloramphenicol (no trigger induction) and one sample was resuspended in fresh TB media with 100 µg ml^{-1} ampicillin, 25 µg ml^{-1} chloramphenicol and 1% arabinose w/v (trigger induced). Luminescence was measured again 1 and 3 h post media exchange using the method outlined earlier.

Dynamic cycling of enzyme assembly. The corresponding plasmids were transformed into *E. coli* BL21-Gold(DE3) cells via heat shock and plated on LB medium (10 g l^{-1} tryptone, 5 g l^{-1} yeast extract, 10 g l^{-1} NaCl, 15 g l^{-1} agar) with ampicillin (100 µg ml^{-1}) and chloramphenicol (25 µg ml^{-1}) antibiotics for selection. Colonies were picked from the agar plates and cultured overnight in TB

supplemented with 100 μg ml^{-1} carbenicillin and 25 μg ml^{-1} chloramphenicol at 37 °C. The following day these cultures were used to inoculate subcultures at an OD of ~0.05. The subcultures were grown to an OD of ~1.2 at which time they were induced with 10 ng μl^{-1} of aTc. Two hours post aTc induction, samples were taken for luminescence measurements according to the standard protocol. The cultures were then spun down at 2,200 r.c.f. for 5 min, the supernatant was removed and the bacterial pellet was washed with fresh TB supplemented with ampicillin (100 μg ml^{-1}) and chloramphenicol (25 μg ml^{-1}), and spun down once more at 2,200 r.c.f. for 5 min. Following the second spin, the supernatant was again removed, and the cells were resuspended in fresh TB with the same antibiotic concentrations. For the sample with no cycling, no inducer was added. For the sample with cycling, the samples were resuspended to a final OD_{600} of 1.6 (to return cell density to appropriate levels for induction) and the medium was supplemented with 300 μM IPTG to induce T1. Two hours post IPTG induction, samples for luminescence measurements were taken again. Following that, a second media exchange was performed on the cultures with scaffold cycling using the procedure outlined above. The fresh TB added following the media exchange was supplemented with 2% w/v arabinose to drive T2 expression. Finally, 2 h post arabinose addition, the final samples for luminescence measurements were harvested.

Coimmunoprecipitation assay. For the coimmunoprecipitation assay, BL21-Gold(DE3) cells were transformed with plasmids CR605 and CR435. A culture was grown following the protocol outlined for the in vivo assay. The culture was induced at an OD ~1.2 with 100 ng ml^{-1} anhydrotetracycline and 300 μM IPTG. Following induction, the sample was grown for 7.5 h and then harvested via centrifugation at 2,500 r.p.m. for 5 min. The bacterial pellet was resuspended in 1× Monarch DNA/RNA Protection Reagent (NEB) to a final OD of 5. Some 200 μl of the sample was taken and the pH was adjusted to ~7 using 5 M NaOH (~0.25 μl). Coimmunoprecipitation was performed using protein G magnetic beads (NEB) according to the manufacturer's protocol . For the immunoprecipitation, 2.5 μg of mouse monoclonal anti-his antibody was used (GenScript Biotech, Piscataway, NJ, USA). Following coimmunoprecipitation, the samples were analyzed using PAGE on two 7.5% SDS acrylamide gels. One of the gels was transferred using a standard western blot transfer protocol on a nitrocellulose membrane. The membrane was then cut in half (each half containing the same samples); one half was probed using mouse anti-his antibody (1:10,000 dilution) and the other half was probed using mouse anti-FLAG antibody (1:1,000 dilution). The blots were then both probed with horseradish peroxidase-conjugated goat anti-mouse antibodies (1:8,000) and were imaged using Amersham ECL Prime western blotting detection reagent (GE Healthcare).

IAA production assay. For the IAA production assay, BL21-Gold(DE3) cells were transformed with the plasmids CR549 and CR435 or CR618. A culture was grown following the protocol outlined in the in vivo assay method with the only difference being the culture medium was 2XYT (16 g l^{-1} tryptone, 10 g l^{-1} yeast extract, 5 g l^{-1} NaCl). The culture was induced at an OD of ~1.2 (1 ng ml^{-1} (low dose) or 100 ng ml^{-1} (high dose) anhydrotetracycline, 300 μM IPTG or both). Following induction, the cultures were grown for 16 h after which the supernatant was harvested by centrifugation of the cultures at 5,000 r.p.m. for 5 min. Following that, the supernatants were analyzed on Agilent 1260 HPLC (Agilent Technologies) using a Zorbax Eclipse XDB C18 (Agilent Technologies) column, 0.1% trifluoroacetic acid buffer A and acetonitrile, 0.1% trifluoroacetic acid buffer B. IAA eluted at around 21 min. The standards were run in 2XYT media. Analysis of HPLC data was performed using OpenLab software (Agilent Technologies).

Malate production assay. For the malate assay, BL21-Gold(DE3) cells were transformed with the appropriate plasmids. A culture was grown following the protocol outlined in the in vivo assay method, the only difference being that the culture medium was TB supplemented with 20 g l^{-1} of glucose. The cultures were induced at an OD of ~1.2 with 100 ng μl^{-1} aTc, and 3, 30 and 300 μM IPTG depending on the induction condition. Post induction, the cultures were grown for 24 h, after which the OD_{600} was recorded, and the samples were harvested. The samples were spun at 5,000 r.p.m. for 5 min, after which the supernatant was stored at −80 °C until HPLC analysis.

The supernatants were analyzed on Agilent 1200 HPLC (Agilent) using a Bio-Rad organic acid analysis column, Aminex HPX-87H ion-exclusion column (Bio-Rad) with 5 mM sulfuric acid buffer A and acetonitrile, and water as buffer B. A method previously developed for bacterial metabolites was used to elute the compounds of interest[47]. The standards were TB media. Analysis of HPLC data was performed using OpenLab software (Agilent).

Statistical analysis. All experiments were performed in triplicate (three separate biological replicates) unless otherwise stated. The error bars in the graphs represent the standard deviation unless stated otherwise. All plotting and statistical analysis was performed using Microsoft Excel.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Sequences of all constructs studied are included in the Supplementary Information. Source data are provided with this paper. Additional data are available from the corresponding author upon reasonable request.

References

47. Charubin, K. & Papoutsakis, E. T. Direct cell-to-cell exchange of matter in a synthetic *Clostridium* syntrophy enables CO_2 fixation, superior metabolite yields, and an expanded metabolic space. *Metab. Eng.* **52**, 9–19 (2019).

Acknowledgements

This work was supported by grants from NSF (CBET1803008, MCB1817675 and MCB2013991).

Author contributions

A.A.M. and W.C. conceived the project. A.A.M. and W.C. designed the experiments. A.A.M. and M.V. performed the experiments. A.A.M. M.V. and W.C. analyzed the data. A.A.M. and W.C. wrote the paper. All authors discussed the results and commented on the manuscript.

Competing interests

A.A.M. and W.C. have filed a patent application ('Dynamic Control of Colocalization of Proteins', US patent application 16/751,793) describing the use of orthogonal Cas6 proteins for enzyme assembly and chemical synthesis.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41589-022-01005-7.

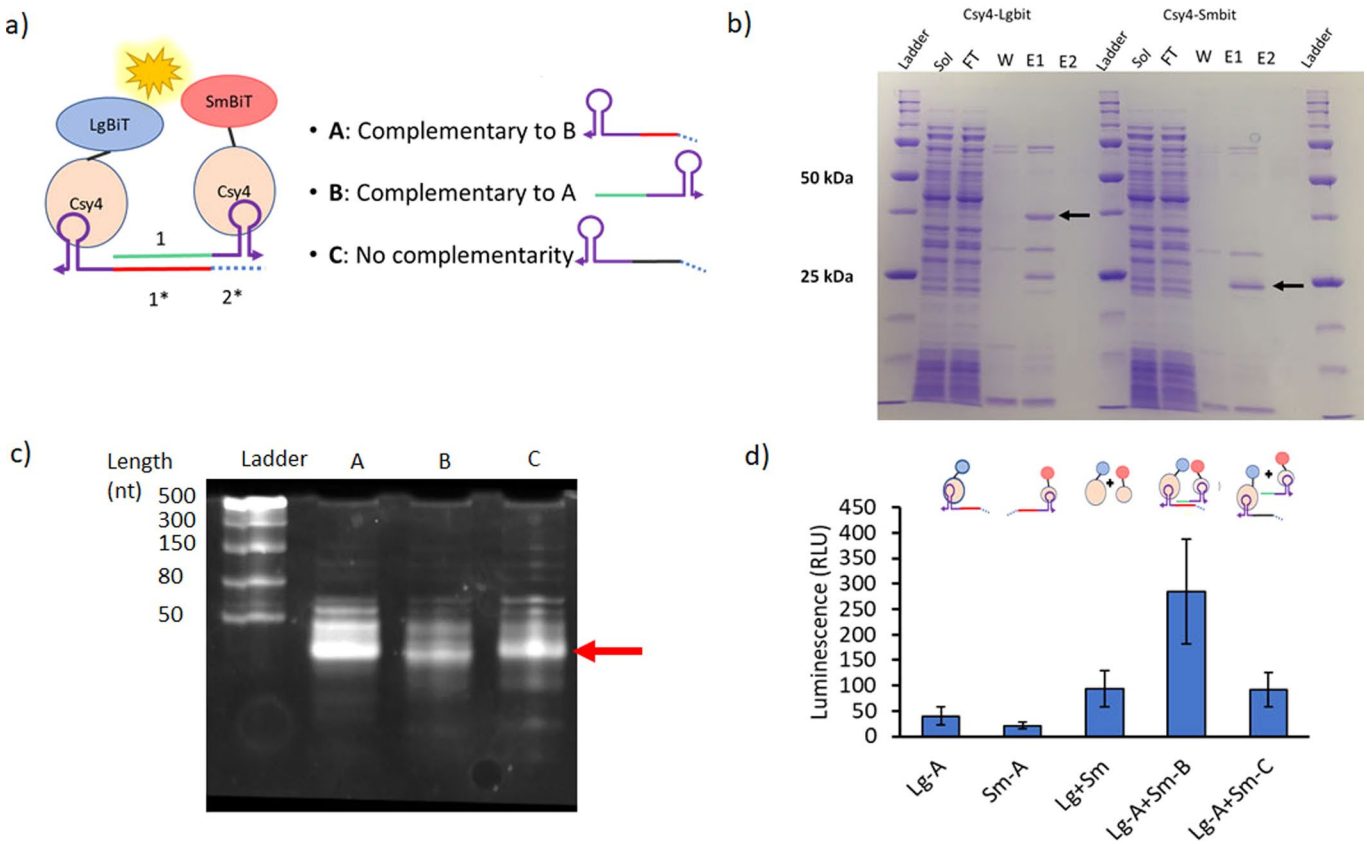
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41589-022-01005-7.

Correspondence and requests for materials should be addressed to Wilfred Chen.

Peer review information *Nature Chemical Biology* thanks Matthew Chang, Alexander Green and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

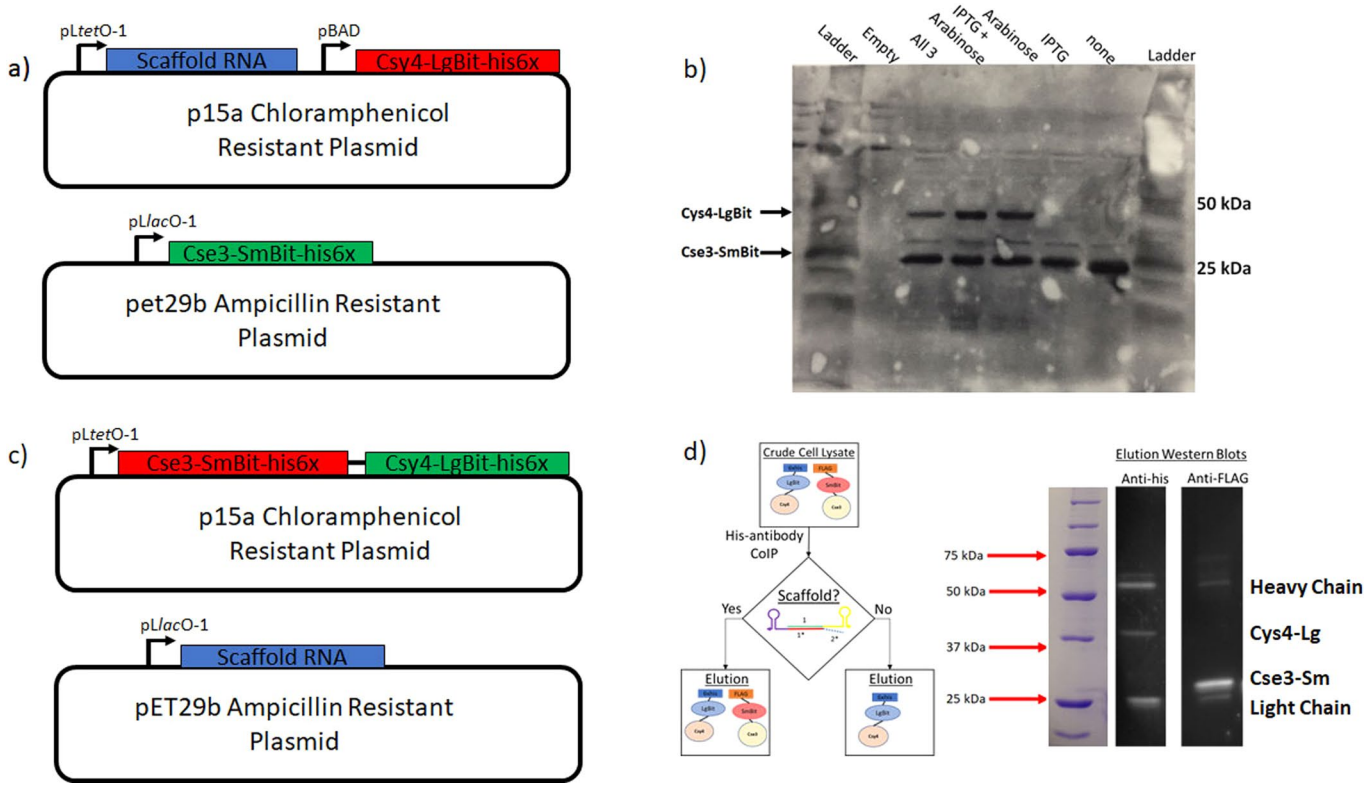
Reprints and permissions information is available at www.nature.com/reprints.

Accepted Manuscript
Version of record at: https://doi.org/10.1038/s41589-022-01005-7



Extended Data Fig. 1 | *In vitro* Cas6-mediated split Nluc assembly. a) The assembly components, Csy4-LgBit, Csy4-SmBit and small RNAs A, B, and C, were generated through protein expression and *in vitro* RNA transcription. b) Production of Cas6 fusions. 7.5% SDS Acrylamide gel indicating full-length expression and purification of the novel fusion proteins Csy4-LgBit and Csy4-SmBit. The full length of the proteins are indicated with around on the bands. Legend is as follows: Sol: soluble fraction, FT: Flow through, W: Wash, E1: Elution fraction 1, E2, Elution fraction 2. c) Different *in vitro* components were mixed together and luminescence was assayed after 15 minutes of incubation. The two fusion proteins together exhibited minimal background activity as compared to the blank (RLU = 6). A 3-fold increase in luminescence was observed (Lg-A + Sm-B) when the RNA scaffolds added are complementary, while no increase was observed when the two RNA scaffolds do not have complementarity (Lg-A + Sm-C). Results are presented as mean ± s.d. of three biological replicates. d) 10% Urea-PAGE acrylamide gel for confirmation of *in vitro* transcript products. Products A, B, and C, all should have a full length of around 60 nt. 60 nt is identified on the gel with the red arrow on the gel.

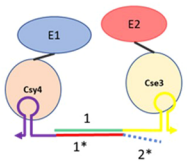
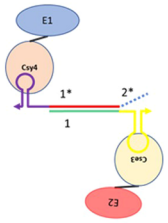
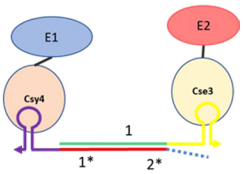
Accepted Manuscript
Version of record at: https://doi.org/10.1038/s41589-022-01005-7



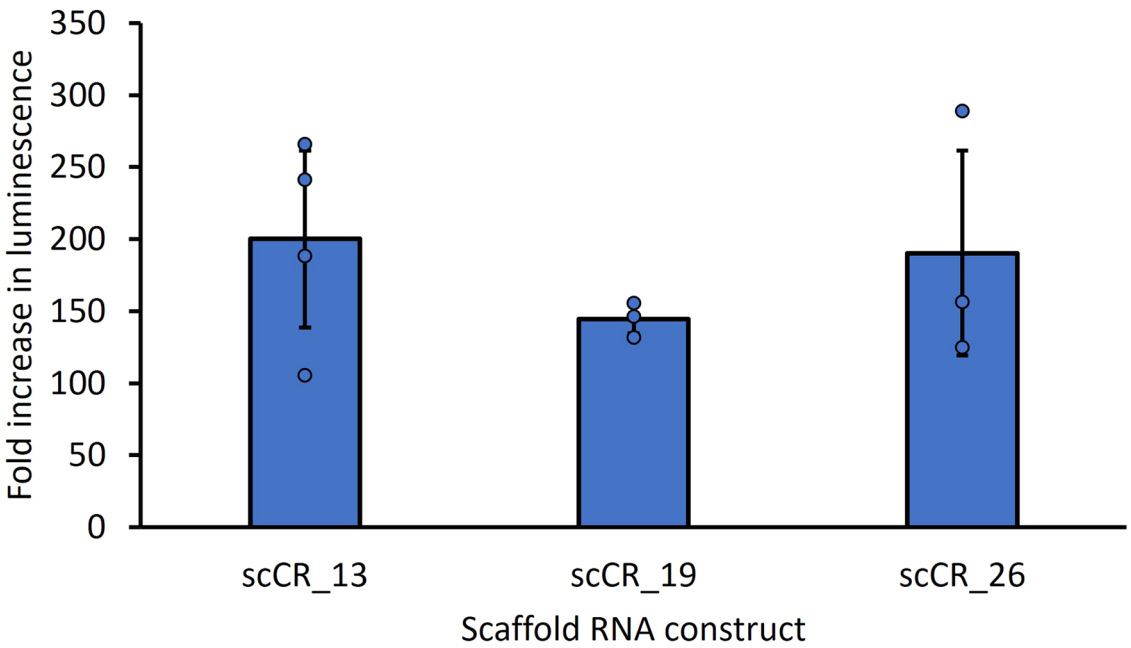
Extended Data Fig. 2 | *In vivo* Cas6-mediated split Nluc assembly. (a) Original plasmid configuration for *in vivo* expression of the scaffold system. The p15a plasmid contains a tetracycline inducible promoter, under which the scaffold RNA is placed as well as an arabinose inducible promoter under which the protein Csy4-LgBit-his6x was placed. The pET29b plasmid contains an IPTG inducible promoter under which the Cse3-SmBit-his6x protein was placed. The plasmids were cotransformed in *E. coli* to assess initial scaffold assembly. This scaffold configuration was used for the experiment in Fig. 1e. (b) Western Blot probing for histidine tags of samples in the experiment in Fig. 1e. The location of each protein is denoted with an arrow on the left of the Western Blot. Each lane represents a different induction condition. (c) Optimized modular expression system. The two fusion proteins were combined onto one plasmid which was placed on the p15a plasmid under the pLtetO-1 promoter. The pET29b plasmid contains the scaffold RNA sequence under control of the pLlacO-1 promoter. Both plasmids were designed to be modular, having unique cut sites around each DNA sequence of interest, which can allow users to easily introduce new RNA scaffold sequences and/or fusion proteins. This plasmid combination was used for the experiment in Fig. 1f. (d) Co-immunoprecipitation schematic illustrating that upon expression of the scaffold RNA, the two fusion proteins are tethered together via RNA interactions and can both be pulled down via an immunoprecipitation using an anti-his antibody. Results of the Western Blot performed on the coimmunoprecipitation elution show successful pulldown of the Cse3-SmBit upon probing with an anti-his mouse antibody. All visible bands on the Western Blot are labelled accordingly.

Accepted Manuscript
Version of record at: https://doi.org/10.1038/s41589-022-01005-7

a)

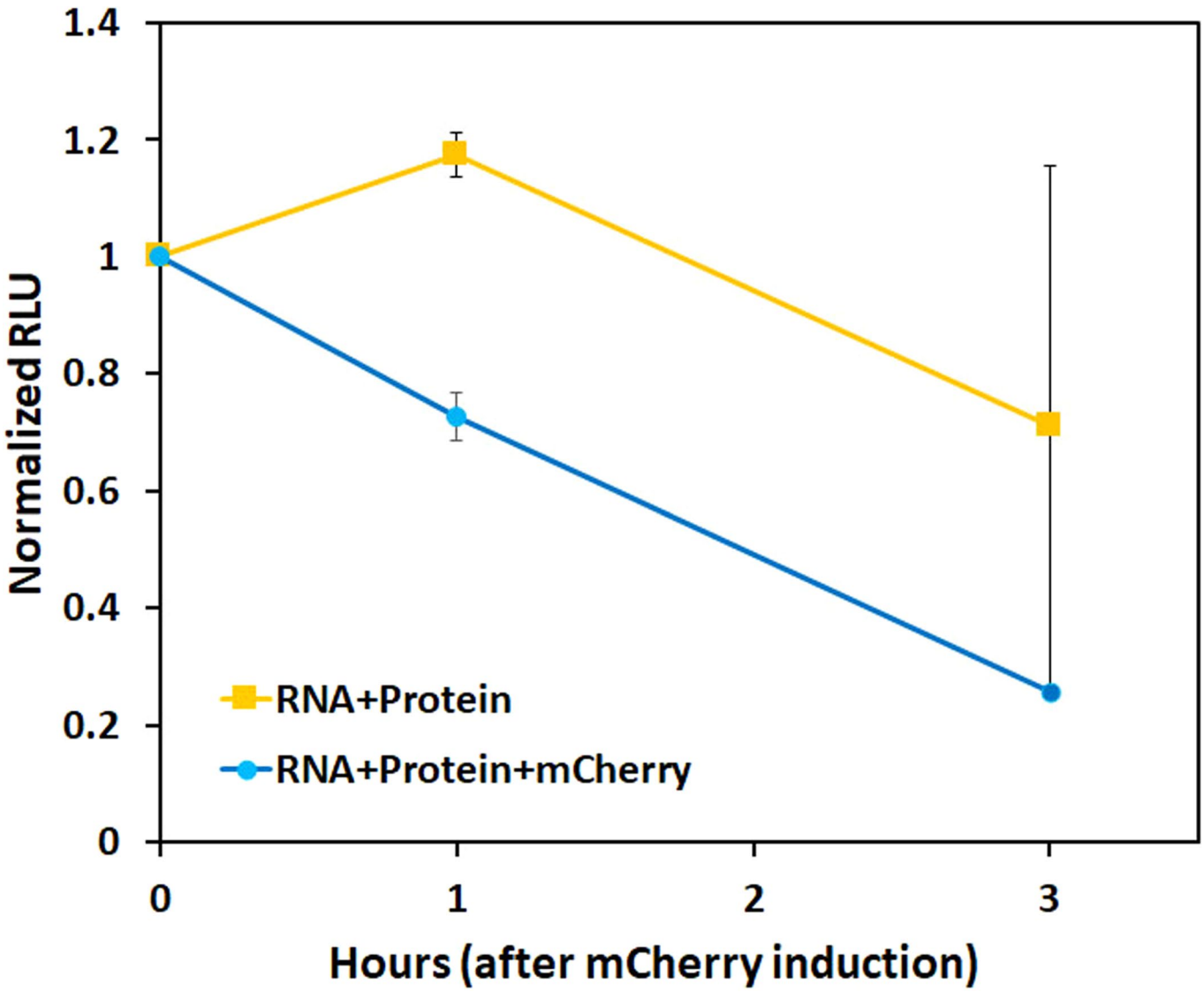
Construct name	Hybridization length (nt)	# of turns of RNA	Orientation of proteins	
scCR_13	13	1	Cis	
scCR_19	19	1.5	Trans	
scCR_26	26	2	Cis	

b)



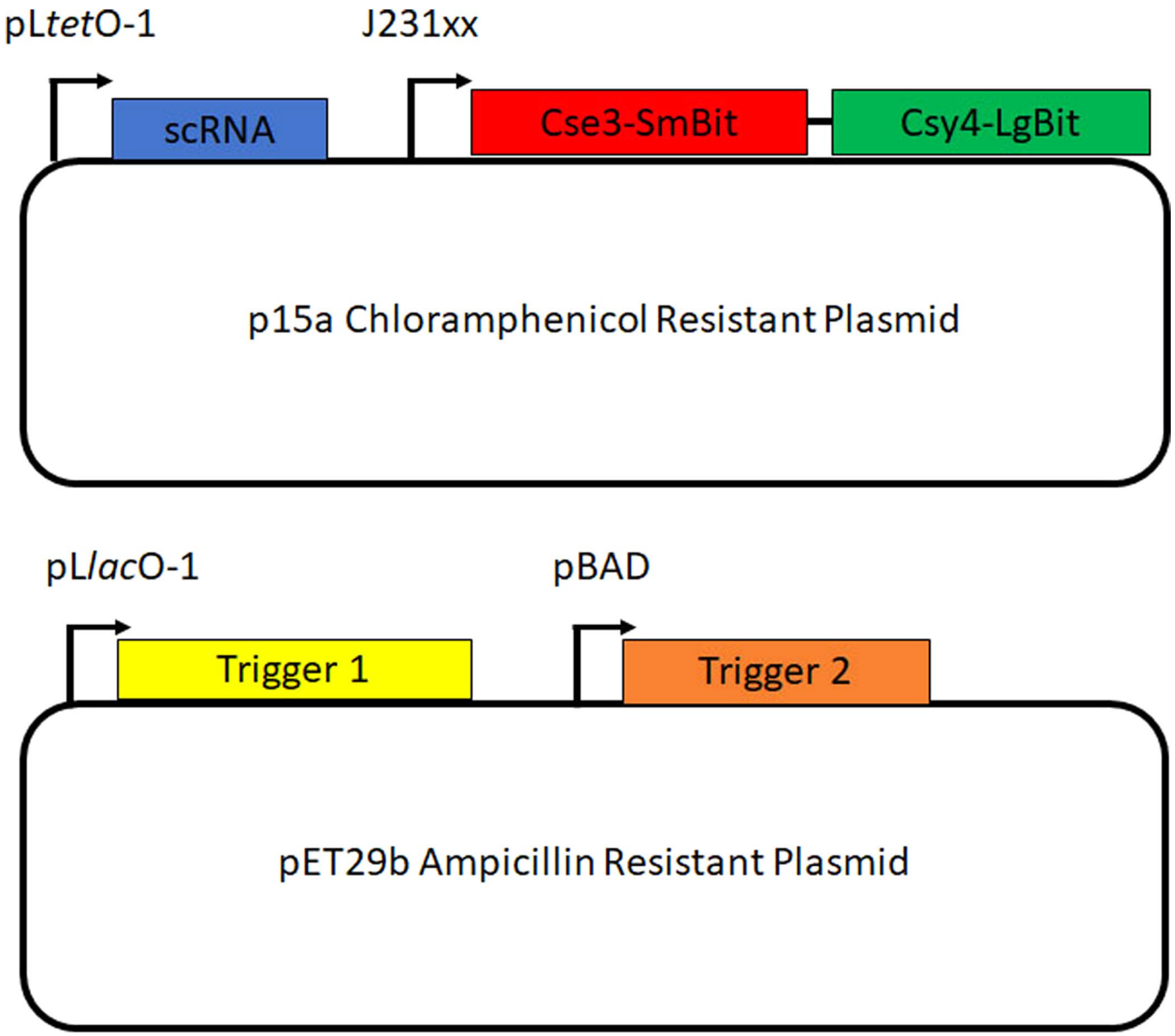
Extended Data Fig. 3 | Effect of longer 5′ RNA handles on split Nluc assembly. a) Table depicting the characteristics of the three scaffold RNA sequences tested. They are the same overall, with the only difference being the length of the hybridization region. The lengths tested were 13, 18, and 26 nt which correspond respectively to a *cis, trans*, and *cis* orientation of the two proteins when they are scaffolded. The number of RNA turns are also highlighted on the table. b) Fold increase in luminescence 8 hours post protein and RNA induction for each of the three different RNA scaffold constructs. Fold increase was calculated as the luminescence ratio of: (proteins + scaffold RNA induction)/ (proteins only induction). Induction conditions were kept the same in all three cases and experiments were performed in parallel. Luminescence results indicate that both proximity of proteins as well as orientation play a role on the assembly of the functional Nluc. Results are presented as mean ± s.d. of three biological replicates.

Accepted Manuscript
Version of record at: https://doi.org/10.1038/s41589-022-01005-7



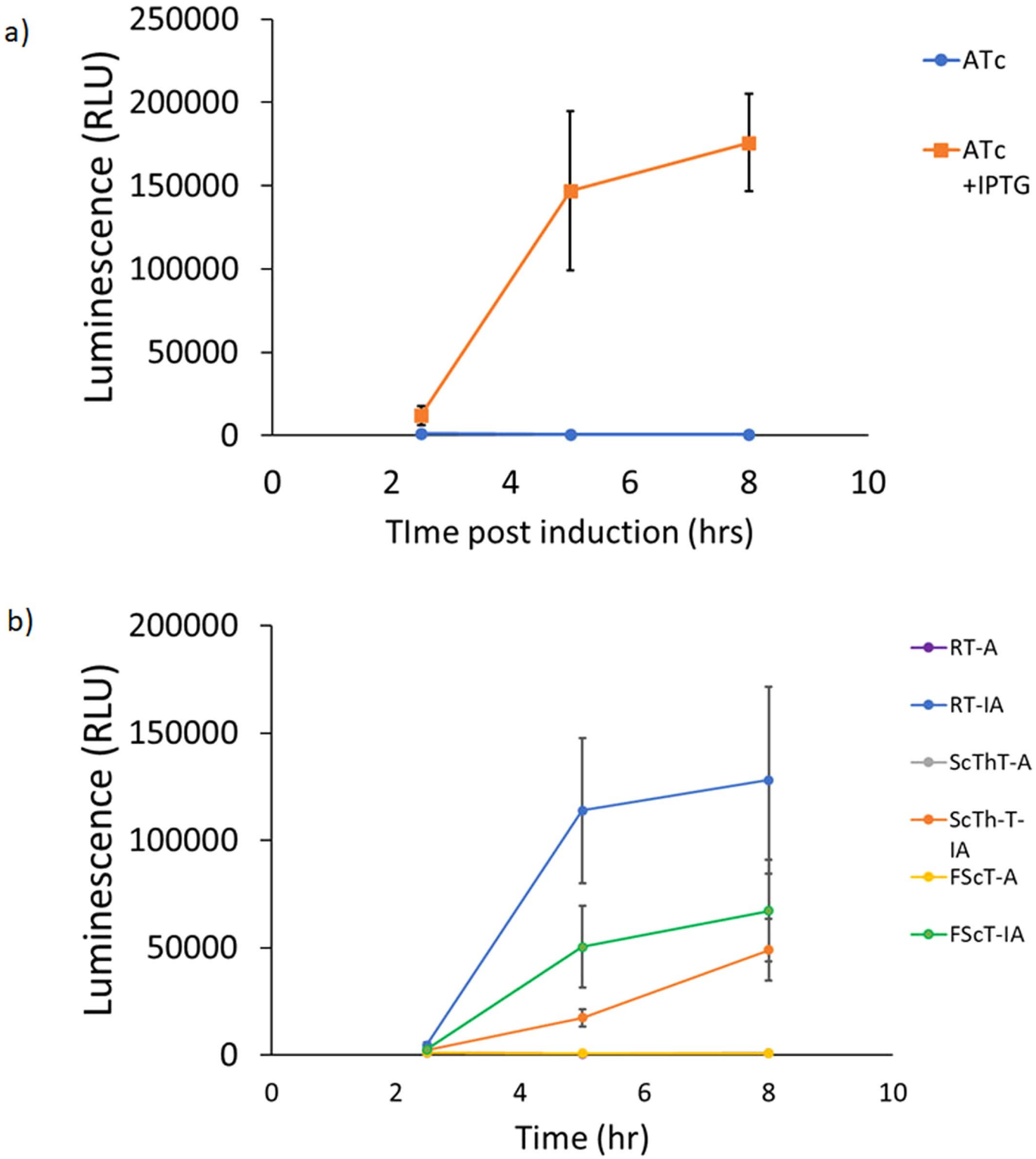
Extended Data Fig. 4 | Improved protein disassembly by *mCherry* mRNA with media exchange. Before *mCherry* mRNA induction via arabinose addition (blue line), the media of the cell culture was exchanged to fresh media without any IPTG or aTc (only arabinose). The luminescence of the sample with trigger expression (blue) drops to 60% of that of the sample with no trigger expression 1 hour post trigger induction and 3 hours post trigger induction it is down to 30%. This result illustrates that the slow reaction observed in Fig. 2d is due to the continued expression of the scaffold RNA due to the presence of IPTG in the medium. Results are presented as mean ± the range of two biological replicates.

Accepted Manuscript
Version of record at: https://doi.org/10.1038/s41589-022-01005-7



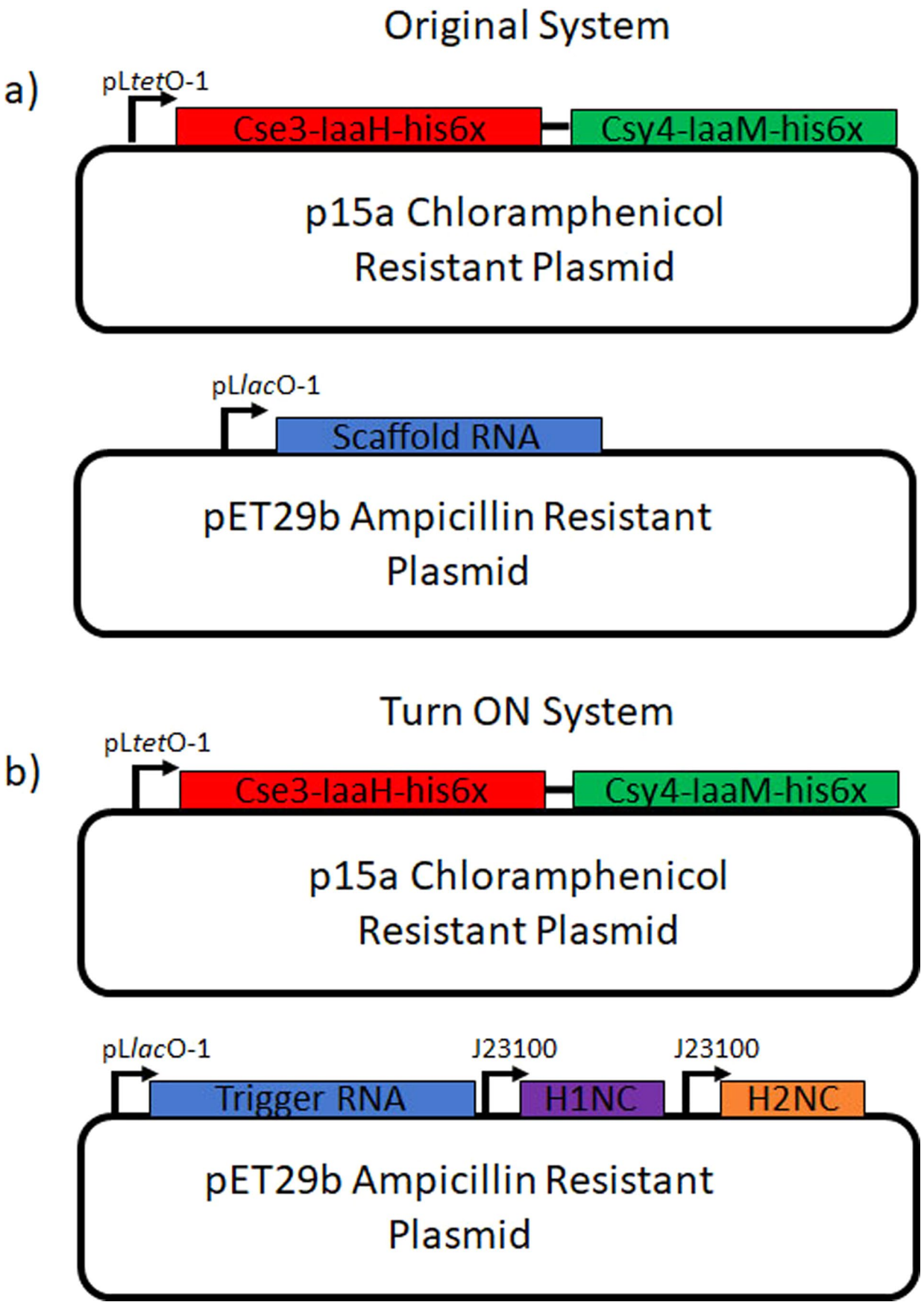
Extended Data Fig. 5 | Dual plasmid system for dynamic protein assembly. Expression system for the cycling scaffold architecture. The chloramphenicol resistance plasmid contains the scaffold RNA under pLtetO promoter and the protein operon under the J23113 constitutive promoter system. The two triggers are on the ampicillin resistance plasmid, with trigger 1 under control of the pLlacO-1 promoter and trigger 2 under the pBAD promoter.

Accepted Manuscript
Version of record at: https://doi.org/10.1038/s41589-022-01005-7



Extended Data Fig. 6 | Raw luminescence data for the Turn-ON system. (a) Raw Luminescence data for the direct assembly indicating that there only when the scaffold and proteins are induced (orange) is there a high amount of luminescent signal. (b) Raw luminescence data for the turn ON assembly system. RT: Regular trigger, ScThT: Scrambled toehold trigger, FSc: Fully scrambled trigger. In this case the only conditions upon which any observable increase in luminescence is observed is upon induction of one of the three trigger alongside the proteins (orange, green, yellow).

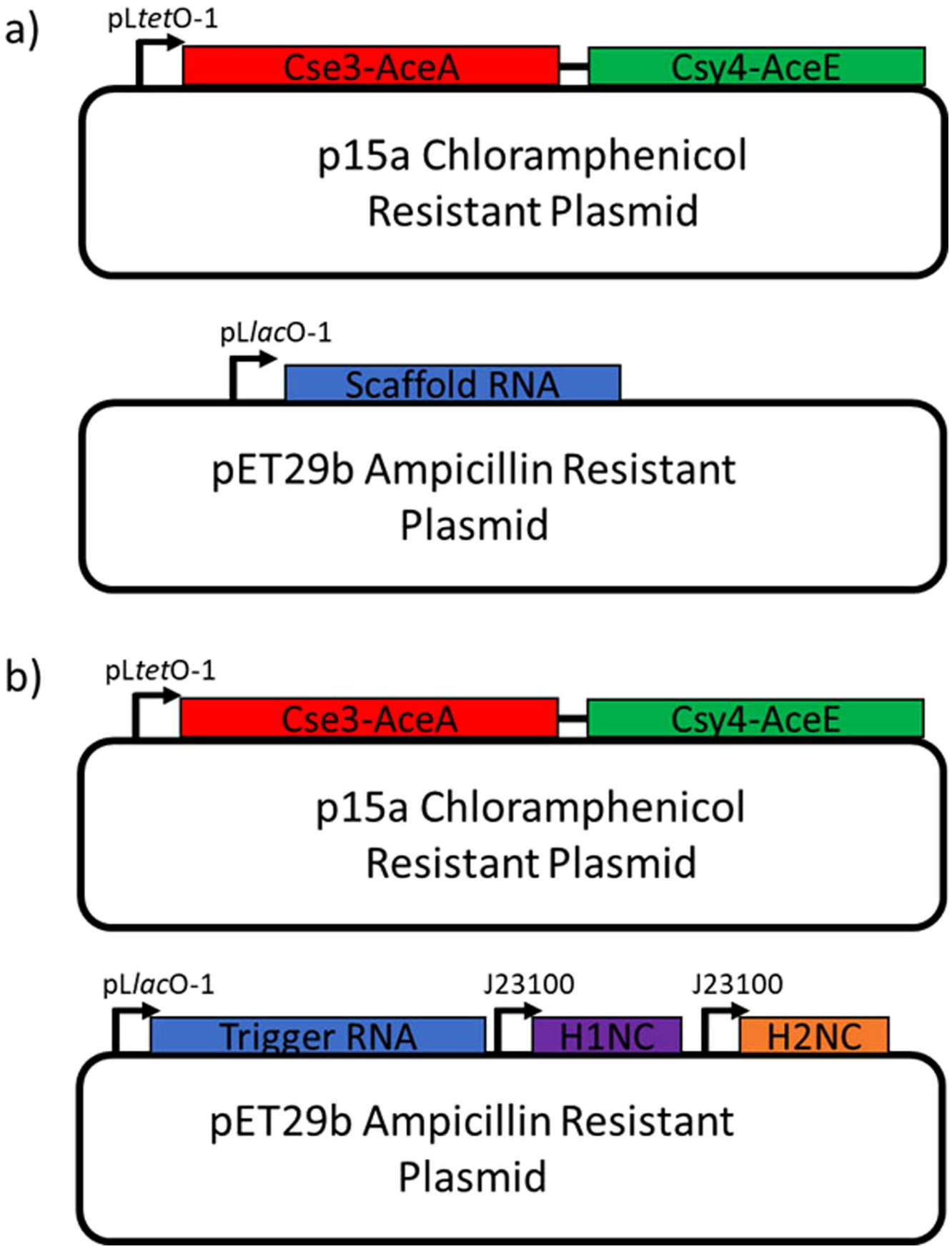
Accepted Manuscript
Version of record at: https://doi.org/10.1038/s41589-022-01005-7



Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Expression system for indole-3-acetic acid synthesis. (a) Plasmid structure for indole-3-acetic acid using direct scaffold assembly. The system is the same as the on in Supplementary Fig. 2(a) except the SmBit and LgBit have now been replaced with the enzymes IaaH and IaaM. The unique restriction sites allowed for easy substitution of the new fusion partners with traditional subcloning techniques. (b) Plasmid structure for indol-3-acetic acid using turn ON scaffold assembly. In this system, the scaffold RNA plasmid has been modified, such that the two separate hairpins of the turn ON system are constitutively expressed via the J23100 synthetic promoter, and the trigger sequence is under control of the pLlacO-1 promoter.

Accepted Manuscript
Version of record at: https://doi.org/10.1038/s41589-022-01005-7



Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Expression system for enhanced malate synthesis. (a) Plasmid structure for malate enhancement using direct scaffold assembly. The system is the same as the on in Supplementary Fig. 7(a) except the IaaH and IaaM have now been replaced with the enzymes AceA and AceE. The unique restriction sites allowed for easy substitution of the new fusion partners with traditional subcloning techniques. (b) Plasmid structure for malate enhancement using turn ON scaffold assembly. In this system, the scaffold RNA plasmid has been modified, such that the two separate hairpins of the turn ON system are constitutively expressed via the J23100 synthetic promoter, and the trigger sequence is under control of the pLlacO-1 promoter.