ChemComm

CRISPR/Cas-directed programmable assembly of multi-enzyme complexes

Journal:	*ChemComm*
Manuscript ID	CC-COM-02-2020-001174.R1
Article Type:	Communication

COMMUNICATION

CRISPR/Cas-directed programmable assembly of multi-enzyme complexes

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Samuel Lim[a], Jiwoo Kim[a], Yujin Kim[a], Dawei Xu[a] and Douglas S. Clark*[a,b]

We describe a versatile CRISPR/Cas-based strategy to construct multi-enzyme complexes scaffolded on a DNA template in programmable patterns. Catalytically inactive dCas9 nuclease was used in combination with SpyCatcher-SpyTag chemistry to assemble enzymes in a highly modular fashion. Five enzymes comprising the violacein biosynthesis pathway were precisely organized in nanometer proximity; a notable increase in violacein production demonstrated the benefits of scaffolding.

Spatially organized multi-enzymatic complexes, or "metabolons," are often found in nature to facilitate multi-step metabolic reactions by promoting substrate transfer between the active sites of enzymes clustered in close proximity.[1] As biocatalysis has gained increasing attention as a promising strategy for the sustainable production of chemical and biomedical products, interest in engineering the artificial equivalent of metabolons has grown accordingly.[2]

While a broad range of template materials including proteins[3,4] or lipid membranes[5] may direct the formation of multi-enzyme assemblies, nucleic acid-based enzyme scaffolding is particularly promising because the highly predictable nature of Watson-Crick base pair interaction allows for precise control over key structural features such as template geometry, number of enzyme docking sites, and inter-enzyme spacing.[6] The most straightforward way of organizing enzymes on a DNA scaffold is to use chemically attached DNA tether strands; however, this approach often suffers from low binding yields, in addition to requiring extraneous reaction steps that can potentially damage the enzymes.[7,8] Alternatively, DNA-binding protein domains such as zinc fingers (ZFs)[9] or transcription activator-like effectors (TALEs)[10] have been fused to enzymes to direct them to the desired positions on the scaffold. However, these methods require case-specific design of the scaffold and employment of different binding domains in order to change the stoichiometry or arrangement of enzymes, thereby limiting their utility in achieving modular enzyme assembly.

CRISPR-associated (Cas) nucleases represent another class of DNA-binding proteins that has recently attracted enormous attention as a gene modification tool. Unlike ZFs or TALEs, which require changes in their primary amino acid sequences to recognize different sites on DNA, Cas binding specificity can be conveniently customized by employing distinct CRISPR guide RNAs that are complementary to the target DNA sequences.[11] Considering such versatility, Cas may function as a promising interfacing component that can organize a broad range of enzymes on a DNA scaffold with minimal need for enzyme-specific variations. Recently, Berckman and Chen used Cas nuclease to construct a synthetic cellulosome by placing a cellulase domain in proximity to a binding module on a DNA template.[12] However, their model system involved only a single enzymatic component, and the possibility of forming DNA nanostructures precisely patterned with diverse types of enzymes working in concert remained unexplored.

In this study, we describe a programmable assembly of the five distinct enzymes involved in the violacein biosynthesis reaction, demonstrating the full potential of using Cas nuclease to organize enzymes involved in complex multi-step metabolic pathways. We first verify the tunable positioning of target proteins along the DNA scaffold using a fluorescent protein pair, and subsequently place the five enzymes together in varying arrangements to study the effects of scaffolding enzymes in proximity as well as their specific sequences on violacein production.

The general scheme of the CRISPR/Cas-directed programmable enzyme assembly is shown in Fig. 1. Throughout this study, we used a catalytically inactive variant of Cas9 nuclease from *Streptococcus pyogenes* (dCas9) which does not digest DNA and instead remains bound to the target sequence specified by the guide RNA.[11] As a first step, dCas9 should be conjugated to the enzymes of interest in order to anchor them to the DNA template, for which we chose to use the SpyCatcher-SpyTag conjugation system.

SpyCatcher is a 12 kDa protein domain that specifically recognizes and binds a 13-amino acid SpyTag peptide by forming a covalent isopeptide bond.[13] The SpyCatcher domain was genetically fused to the C-terminus of dCas9 to create dCas9-SpyCatcher, which can subsequently attach to any

[a] *Department of Chemical and Biomolecular Engineering, University of California, Berkeley CA 94720 (USA) E-mail: dsc@berkeley.edu*
[b] *Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720 (USA)*
†Electronic Supplementary Information (ESI) available

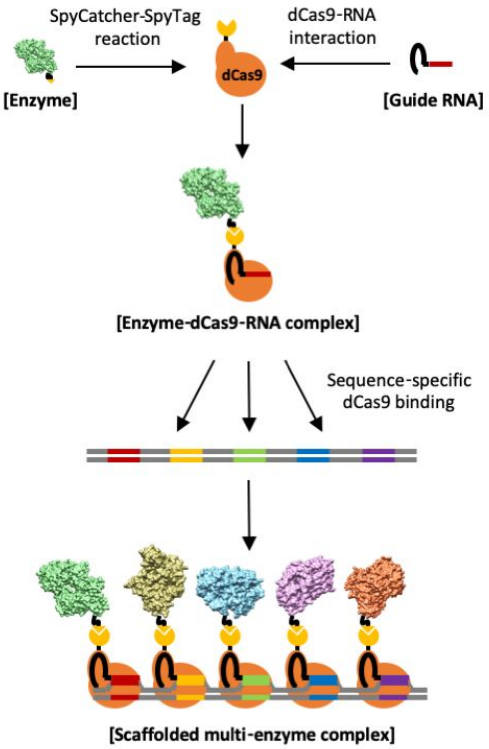


Fig. 1 General scheme of the CRISPR/Cas-directed programmable assembly of multi-enzyme complexes on a DNA scaffold.

separately expressed enzyme of interest containing SpyTag. This (enzyme-SpyTag):(dCas9-SpyCatcher) conjugation product is further loaded with the single guide RNA (sgRNA) to form the (enzyme-SpyTag):(dCas9-SpyCatcher):sgRNA complex, which functions as an individual module that delivers a particular enzyme to its target site on the DNA template. Once all components are assembled with the correct combinations of enzyme and guide RNA in separate batches, they can subsequently be mixed with the DNA template in a single pot to form the precisely patterned scaffolded multi-enzyme complex. Because this strategy is highly modular in its design, the resulting complex is easily customizable in terms of enzyme type, arrangement, and stoichiometry.

We designed a linear DNA template that has five tandem-aligned orthogonal dCas9 binding sequences designated as T1-T5, along with their relevant PAM sequences;[14] random sequences 30 base pairs (bp) in length, which corresponds to a ~10 nm distance along the linear DNA, were inserted as spacers between adjacent binding sequences (Fig. 2). To verify the successful binding, we assembled fluorescently labeled DNA templates with dCas9-SpyCatcher targeted to each site and measured their migration in an electrophoretic mobility shift assay (EMSA). Compared to the unbound DNA template, those containing a bound dCas9-SpyCatcher showed clear retardation in their mobility as demonstrated by the upward shift of the bands (Fig. 2, lanes 1, 2, 7-10). The migration patterns of the DNA with dCas9-SpyCatcher bound at five different target sequences were very similar and no unbound DNA was observed, verifying complete binding of dCas9-SpyCatcher to each target site. Next, we tested whether the DNA template could simultaneously bind dCas9-SpyCatcher at multiple locations. As the number of bound proteins increased from zero

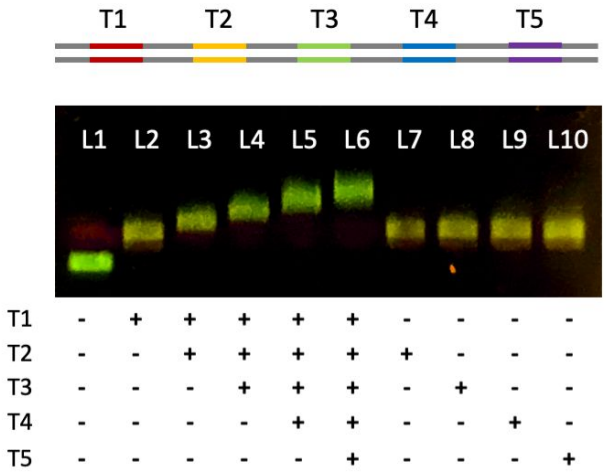


Fig. 2 Tunable binding of dCas9 to the DNA template. dCas9 binding capacity of the template was assessed using EMSA with fluorescently labeled DNA. For each lane, whether individual binding sites are occupied (+) or not (-) is indicated by the table below. As the number of bound proteins increased from zero to five, decreasing migration was observed, confirming that the multiple target sequences on the template can be conjugated with dCas9-SpyCatcher up to full occupancy.

to five, decreased migration was observed, confirming that the multiple target sequences on the template can be occupied simultaneously (Fig. 2, lanes 1-6).

Subsequently, we verified the ability to position the proteins of interest in proximity in a controllable manner using the fluorescent proteins mCerulean3 and mVenus, which are known to display shifts in their fluorescence emission when placed at nanometer-scale distances from each other (< ~10nm) through Forster Resonance Energy Transfer (FRET).[15] In particular, the excited mCerulean3 directly transfers energy to the adjacent mVenus, causing a decrease in mCerulean3 emission and an increase in mVenus emission.[16] Thus, if the spacing between the dCas9 binding sites on the DNA template is close enough, the two fluorescent proteins positioned next to each other will exhibit a FRET response (Fig. 3A). mCerulean3 and mVenus were genetically modified to contain a SpyTag peptide, and their abilities to covalently conjugate with dCas9-SpyCatcher were verified by the appearance of a new band in SDS-PAGE corresponding to the larger adduct upon mixing of the two components (Fig. S1, ESI).

We then assembled the two fluorescent proteins at varying locations on the DNA template and investigated the FRET intensity. A notable increase in the (A_{528nm}/A_{476nm}) ratio, which was caused by the increase in mVenus peak at 528 nm and a relative decrease in mCerulean3 peak at 476 nm, was observed when the two fluorescent proteins were positioned at adjacent binding sites (T1 and T2; T2 and T3; T3 and T4; T4 and T5), indicating strong FRET responses (Fig. 3B, 3C). The distance between the bound proteins calculated from spectral analysis was ~9 nm, which is consistent with the expected spacing of ~10 nm between the two binding sites assuming linear DNA. Conversely, the spectral shifts became much weaker as the fluorescent proteins were positioned farther apart (T1 and T3; T1 and T4; T1 and T5), and the (A_{528nm}/A_{476nm}) ratio decreased to the level similar to that observed from the free mCerulean3 in solution. Such results are consistent with the approximate distances of 27, 44 and 61 nm expected with increasing interval, which are too large for any FRET to occur. Moreover, a strong

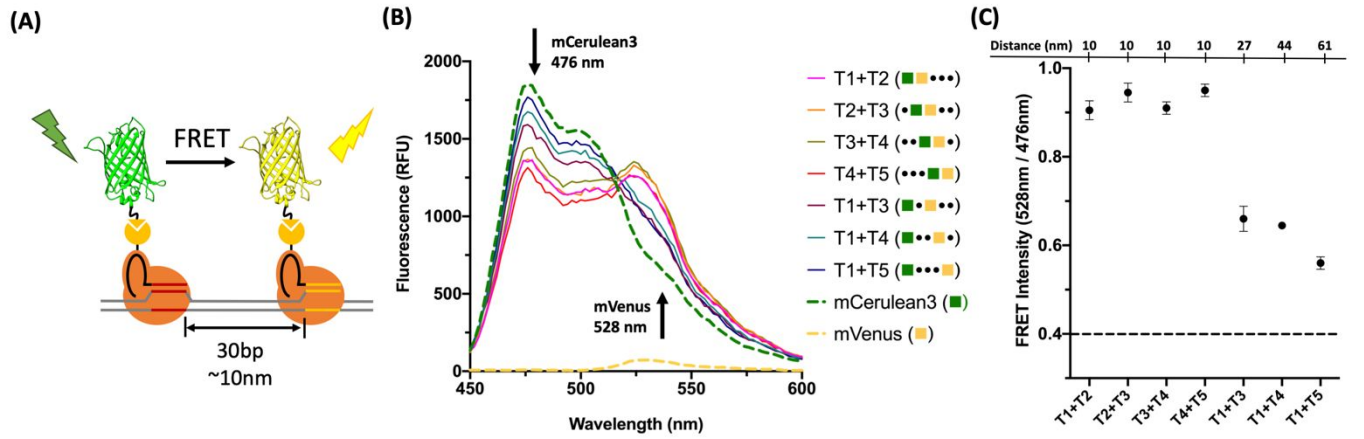


Fig. 3 Verifying the tunable positioning of fluorescent proteins using FRET. (A) Positioning two fluorescent proteins mCerulean3 and mVenus in close proximity (< 10nm) will allow for energy transfer via FRET, resulting in a shift in their emission spectra. (B) Fluorescence emission spectra of the 0.5 μM DNA-fluorescent protein complexes with mCerulean3 and mVenus bound at different locations, following excitation at 412 nm. A notably large decrease in mCerulean3 emission and a corresponding increase in mVenus emission were observed when the two fluorescent proteins were positioned at adjacent binding sites (T1 and T2; T2 and T3; T3 and T4; T4 and T5), indicating strong FRET. Conversely, the spectral shifts became much weaker as the fluorescent proteins were positioned farther apart (T1 and T3; T1 and T4; T1 and T5). (C) FRET intensity, determined by the ratio of mCerulean3 and mVenus emission peaks (A528nm/A476nm), of the assembled DNA-fluorescent protein complexes with mCerulean3 and mVenus bound at different locations. Dashed line represents the value measured from mCeruean3 alone. The expected linear distance between the binding sites for each case is indicated at the top. The error bars represent the standard deviation (SD) from at least two independent experiments.

FRET response was maintained for 24 hours after dCas9 binding to the template (Fig. S2, ESI). Overall, these results suggested that our template could be used to spatially organize the proteins of interest, and the complex formed would remain stable over extended periods of time.

Having confirmed that the dCas9 nuclease can simultaneously deliver multiple proteins to their designated locations on the DNA template, we aimed to construct a spatially organized enzyme-DNA complex, in which multiple enzymes comprising a biocatalytic cascade are assembled in a particular order along the template. As a model system, we chose the violacein biosynthesis pathway, which employs five different enzymes (VioA-E) working in concert to convert L-tryptophan into the purple-colored pigment violacein (Fig. 4A, Fig. S3, ESI).[17] Similarly to the fluorescent proteins, each enzyme was genetically modified to contain a SpyTag peptide, and successful conjugations to the dCas9-SpyCatcher protein were verified using SDS-PAGE (Fig. S4, ESI); dCas9 assembly on the DNA template was complete for each of the five enzymes (Fig. S5, ESI). Furthermore, we verified that conjugating the enzymes with dCas9-SpyCatcher had minimal effect on their combined ability to produce violacein (Fig. S6, ESI).

The five enzymes were assembled on the DNA template in two different arrangements, referred to as “optimal” and “scrambled” orders. In the “optimal” arrangement, the enzymes were positioned according to the order of the sequential reactions in violacein biosynthesis (A-B-E-D-C; from T1 to T5, respectively) to minimize the linear diffusion distance of the reaction intermediates between the enzymes along the template. Conversely, the “scrambled” arrangement represented a presumably unfavorable situation in which no enzyme was placed directly next to either the upstream or downstream enzyme in the pathway (B-D-A-E-C). The biocatalytic performance of the enzymes organized on the template in both arrangements were then compared to that of the free enzymes in solution by measuring the production of violacein over time, up to 120 min; note that in the free enzyme control, each enzyme was bound to dCas9-SpyCatcher. Notably, a ~3.2-fold increase in violacein production was observed upon scaffolding with respect to the free enzymes (Fig. 4B). Yet, the difference in violacein production from “optimal” and “scrambled” arrangements was insignificant, indicating that the effect of relative enzyme order was minimal in this system.

In general, activity enhancements in scaffolded enzyme systems have been attributed to two primary effects: 1) placing enzymes in close proximity to each other, and 2) attaching enzymes to a scaffold that may promote catalytic activation.[18] In particular, enhanced enzyme activities from attachment to DNA have been reported for a broad range of enzyme-DNA complexes,[18-20] underscoring the importance of assessing the two effects separately. Thus, we conducted a control experiment using the five different types of DNA-enzyme complexes containing only a single Vio enzyme, and determining violacein production from a mixture of all five complexes. In this design, each enzyme is bound to the DNA template individually but is completely separated from the other enzymes on different templates; thus, any catalytic enhancement may be attributed to placing the enzymes near the DNA scaffold. Notably, ~1.8-fold more violacein was produced after 120 min compared to the free enzymes, indicating the contribution from the DNA template is not sufficient to account for the ~3.2-fold improvement in violacein production observed upon assembling enzymes together on a single template (Fig. 4B). Although the exact mechanism through which the DNA template activates enzymes remains uncertain, the enzymes may benefit from a favorable local microenvironment formed near the nanostructure,[18-20] and/or affinity of the enzyme substrates to DNA.[21]

While a number of scaffolded multi-enzyme systems have been reported in the literature, whether close inter-enzyme distance plays an important role in improving the productivity remains debatable, as the evidence varied from one system to another.[2] Depending on the type of enzyme, scaffold and conjugation chemistry, the extent to which proximity contributed to the observed catalytic enhancements ranged from none[8,19] to manyfold.[7,22] Our results suggest that the system used in this study benefited both from attaching the enzymes to DNA as well as organizing them close together,

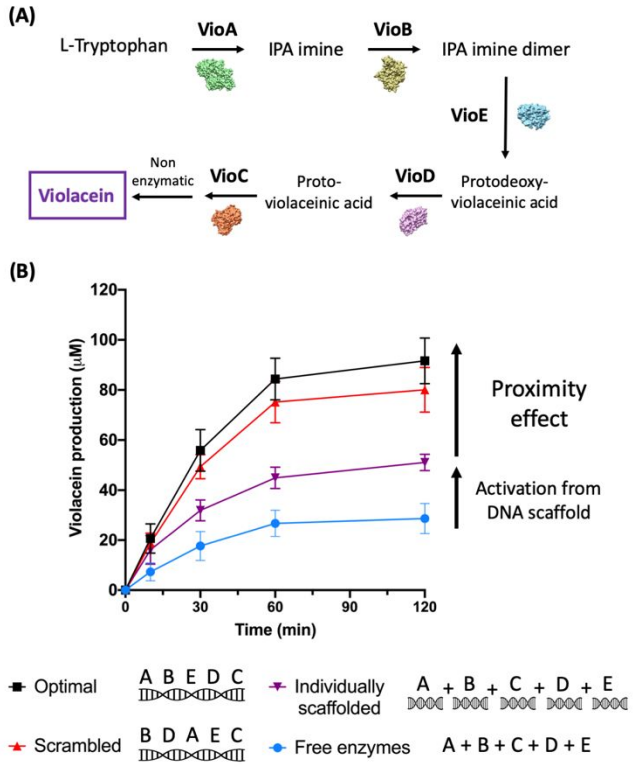


Fig. 4 Programmable assembly of the five enzymes involved in the violacein biosynthesis pathway. (A) Overall scheme of the five-step violacein biosynthesis reaction. (B) Violacein production over 120 min was compared for the different scaffolding systems. Upon scaffolding of all five enzymes together, ~3.2-fold more violacein was produced with respect to the free enzymes, whereas scaffolding each enzyme individually on separate templates showed ~1.8-fold increase. The effect of changing the enzyme order from the hypothetical optimal to a scrambled arrangement was minimal. The error bars represent the SD from at least three independent experiments.

further emphasizing the significance of nanoscale control in engineering multi-enzyme complexes. However, changing the enzyme order from "ideal" to "scrambled", which is likely to increase the distance that intermediates have to diffuse along the template to reach the next enzyme, had no significant effect. It is noteworthy that even in the "scrambled" arrangement, the five enzymes were still scaffolded in the vicinity of each other (< ~40nm along the template); varying the inter-enzyme distance within this relatively small range may have a minimal effect in this system, a phenomenon that has been previously observed from a different DNA-bound enzyme cascade.[23]

In conclusion, we have reported a versatile CRISPR/Cas-based strategy to construct precisely organized scaffolded multi-enzyme complexes. Catalytically inactive dCas9 nuclease was used as an interfacing component that could guide each enzyme to a particular location on the DNA template in a modular fashion. Programmable positioning within nanometer proximity was verified using a fluorescent protein pair, and subsequently the five enzymes comprising the violacein biosynthesis pathway were assembled together in two different arrangements. An intrinsic DNA-induced activation of the enzymes and their close proximity both contributed to a notable increase in violacein production, whereas changing the enzyme order had a minimal effect. The highly modular nature of the system described here allows not only for the scaffolding of virtually limitless types of enzymes, but also for the convenient tuning of their patterning on the template with only a minimal design change. Thus, we expect our methodology to expand the ability of DNA-based enzyme scaffolds to support complex metabolic pathways, broadening their experimental as well as practical applicability.

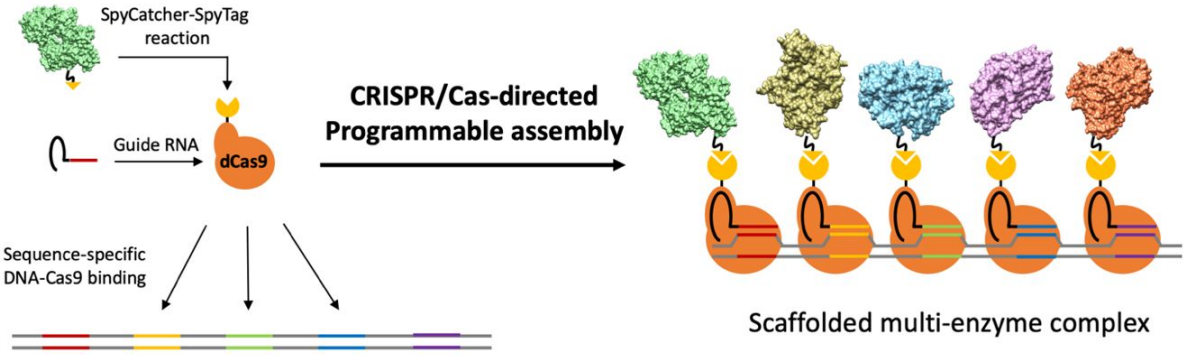
This work was supported by the Air Force Office of Scientific Research (FA9550-17-1-0451) S.L. was supported by a National Science Foundation Graduate Research Fellowship (DGE1106400, DGE1752814). The authors thank Professor John Dueber (UC Berkeley) for helpful discussions.

Conflicts of interest

There are no conflicts to declare.

Notes and references

1 I. Wheeldon, S. D. Minteer, S. Banta, S. C. Barton, P. Atanassov and M. Sigman, *Nat. Chem.*, 2016, **8**, 299.
2 G. A. Ellis, W. P. Klein, G. Lasarte-Aragones, M. Thakur, S. A. Walper and I. L. Medintz, *ACS Catal.,* 2019, **9**, 10812.
3 H. Lee, W. C. DeLoache and J. E. Dueber, *Metab. Eng.,* 2012, **14**, 242.
4 S. Lim, G. A. Jung, D. J. Glover and D. S. Clark, *Small*, 2019, **15**, 1805558.
5 J. L. Lin, J. Zhu and I. Wheeldon, *ACS Synth. Biol.,* 2017, **6**, 1534.
6 B. Sacca and C. M. Niemeyer, *Chem. Soc. Rev.*, 2011, **40**, 5910.
7 J. Fu, M. Liu, Y. Liu, N. W. Woodbury and H. Yan, *J. Am. Chem. Soc.,* 2012, **134**, 5516.
8 W. P. Klein, R. P. Thomsen, K. B. Turner, S. A. Walper, J. Vranish, J. Kjems, M. G. Ancona and I. L. Medintz, *ACS Nano,* 2019, **13**, 13677.
9 T. M. Nguyen, E. Nakata, M. Saimura, H. Dinh and T. Morii, *J. Am. Chem. Soc.,* 2017, **139**, 8487.
10 L. Y. Zhu, X. Y. Qiu, L. Y. Zhu, X. M. Wu, Y. Zhang, Q. H. Zhu, D. Y. Fan, C. S. Zhu and D. Y. Zhang, *Sci. Rep.,* 2016, **6**, 26065.
11 J. A. Doudna and E. Charpentier, *Science,* 2014, **346**, 1258096.
12 E. A. Berckman and W. Chen, *Chem. Commun.,* 2019, **55**, 8219.
13 B. Zakeri, J. O. Fierer, E. Celik, E. C. Chittock, U. Schwarz-Linek, V. T. Moy and M. Howarth, *Proc. Natl. Acad. Sci. U.S.A.* 2012, **109**, E690.
14 B. Chen, W. Zou, H. Xu, Y. Liang and B. Huang, *Nat. Commun.,* 2018, **9**, 5065.
15 S. Lim, G. A. Jung, R. J. Muckom, D. J. Glover and D. S. Clark, *Chem. Commun.,* 2019, **55**, 806.
16 D. J. Glover, S. Lim, D. Xu, N. B. Sloan, Y. Zhang and D. S. Clark, *ACS Synth. Biol.,* 2018, **7**, 2447.
17 M. E. Lee, A. Aswani, A. S. Han, C. J. Tomlin and J. E. Dueber, *Nucleic Acids Res.,* 2013, **41**, 10668.
18 Z. Zhao, J. Fu, S. Dhakal, A. Johnson-Buck, M. Liu, T. Zhang, N. W. Woodbury, Y. Liu, N. G. Walter and H. Yan, *Nat. Commun.,* 2016, **7**, 10619.
19 Y. Zhang, S. Tsitkov and H. Hess, *Nat. Commun.,* 2016, **7**, 13982.
20 S. Rudiuk, A. Venancio-Marques and D. Baigl, *Angew. Chem. Int. Ed.,* 2012, **51**, 12694.
21 J. L. Lin and I. Wheeldon, *ACS Catal.,* 2013, **3**, 560.
22 T. A. Ngo, E. Nakata, M. Saimura and T. Morii, *J. Am. Chem. Soc.,* 2016, **138**, 3012.
23 M. Liu, J. Fu, X. Qi, S. Wootten, N. W. Woodbury, Y. Liu and H. Yan, *ChemBioChem,* 2016, **17**, 1097.



We describe a versatile CRISPR/Cas-based strategy to construct precisely organized, scaffolded multi-enzyme systems with improved productivity.