

Cite this: *Chem. Sci.*, 2023, 14, 12693 All publication charges for this article have been paid for by the Royal Society of Chemistry

# Sequence-independent, site-specific incorporation of chemical modifications to generate light-activated plasmids†

Khoa Chung <sup>a</sup> and Michael J. Booth <sup>\*ab</sup>

Plasmids are ubiquitous in biology, where they are used to study gene-function relationships and intricate molecular networks, and hold potential as therapeutic devices. Developing methods to control their function will advance their application in research and may also expedite their translation to clinical settings. Light is an attractive stimulus to conditionally regulate plasmid expression as it is non-invasive, and its properties such as wavelength, intensity, and duration can be adjusted to minimise cellular toxicity and increase penetration. Herein, we have developed a method to site-specifically introduce photocages into plasmids, by resynthesising one strand in a manner similar to Kunkel mutagenesis. Unlike alternative approaches to chemically modify plasmids, this method is sequence-independent at the site of modification and uses commercially available phosphoramidites. To generate our light-activated (LA) plasmids, photocleavable biotinylated nucleobases were introduced at specific sites across the T7 and CMV promoters on plasmids and bound to streptavidin to sterically block access. These LA-plasmids were then successfully used to control expression in both cell-free systems (T7 promoter) and mammalian cells (CMV promoter). These light-activated plasmids might be used to remotely control cellular activity and reduce off-target toxicity for future medical use. Our simple approach to plasmid modification might also be used to introduce novel chemical moieties for advanced function.

Received 31st May 2023

Accepted 11th October 2023

DOI: 10.1039/d3sc02761a

rsc.li/chemical-science

## Introduction

Biological processes are complex and occur in a tightly regulated manner. To fully understand the sophisticated interplay between one gene within a whole network, its downstream effects within a signalling cascade, or its wider physiological impacts on the cellular/tissue level, precise control over its expression is desired. Light represents an ideal means for transcriptional activation as its application is localised and its properties can be finely tuned. Photoregulation of gene expression is a burgeoning field and a multitude of light-dependent technologies are available, encompassing the entire UV-visible spectrum, to mediate photoresponses.<sup>1</sup> Although natural light-responsive proteins are abundant and have shown excellent photomodulation,<sup>2,3</sup> their use is limited by multi-component systems<sup>2-9</sup> and lengthy optimisation.<sup>4,10</sup> Alternatively, photocaging involves the installation of light-sensitive blocking groups, photocages, directly on the nucleic acids to disrupt transcription or translation. The activity is

minimised/abolished until the appropriate light stimulus is applied to remove the caging groups and reactivate expression. Whilst examples of caged linear DNA<sup>11-24</sup> and RNA<sup>25-28</sup> are extensive, photocaged plasmids are limited in the literature.<sup>25,29-34</sup> Owing to their covalently-closed nature, plasmids are resistant to exonucleases abundant within the cytoplasm and therefore they have a longer lifetime and make better templates than linear nucleic acids.<sup>35-37</sup>

Early examples of generating photocaged plasmids incorporated multiple caging groups onto the phosphate backbones in a random manner.<sup>25,29,33,34</sup> Although Nagamune *et al.* were the first to control the site of photocleavable group attachment, only one such group was present, and their method created unsealed nicks in both plasmid strands.<sup>30</sup> The presence of numerous caging groups and irregularity in their positionings across the plasmid sequence likely resulted in the leaky expression and poor activation in these initial attempts. Subsequent efforts to photocage the plasmid installed the photosensitive groups onto the nucleobase in a site-specific manner.<sup>31,32</sup> These strategies exhibited a tight OFF-state, and a good ON-state upon photolysis of the caging groups. However, both methods suffered from the reliance on custom phosphoramidites, which are not commercially available, to introduce the photoresponsive groups, restricting their adoption by the wider research community. Furthermore, these methods necessitated precise placements of specific enzyme-recognition or G-quadruplex forming sequences to enable their caging effect. The

<sup>a</sup>Department of Chemistry, University of Oxford, Mansfield Road, OX1 3TA, Oxford, UK. E-mail: m.j.booth@ucl.ac.uk

<sup>b</sup>Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, UK

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d3sc02761a>



incorporation of these sequences at or near the promoter will not be possible in all plasmids and may result in unintended changes to the promoter's regulatory role. To circumvent these issues, we engineered a novel approach to site-specifically install photoresponsive modifications into a plasmid, by resynthesising one strand, to govern gene expression using UV light (Fig. 1A). By electing to block the non-Watson-Crick-Franklin (WCF) face of the nucleobase, all components necessary to affect plasmid resynthesis and optical control were commercially available, which will permit its routine use by any laboratory. As the short nickase recognition sites required to create these light-activated (LA) plasmids could be placed anywhere on the plasmid, this approach is general and compatible with most plasmid sequences, thus extending the approach's accessibility compared to the previously generated methods. This strategy was demonstrated for two nickases with plasmids containing either the T7 or CMV promoter to control cell-free and mammalian gene expression (Fig. 1B).

## Results and discussion

### Development of NEEL

To incorporate photoblocking modifications, we were inspired by Kunkel site-directed mutagenesis (SDM), which was traditionally used to introduce mutations site-specifically into DNA

sequences.<sup>38,39</sup> Briefly, in Kunkel SDM a plasmid, containing the sequence to be mutated, is transformed into an *E. coli* strain lacking the enzymes uracil deglycosidase ( $udg^-$ ) and dUTPase ( $dut^-$ ). These enzymes are involved in DNA repair by preventing dUTP insertion into bacterial chromosomes. Ultimately, bacteria produce a uracil-containing circular single-stranded DNA (cssDNA). This cssDNA serves as a template to anneal a primer bearing the desired mutations. The primer is then extended around the plasmid and ligated *in vitro*. The resultant heteroduplex comprises a mutated deoxythymine-containing strand and the original uracil-containing strand. When introduced into a  $udg^+$ ,  $dut^+$  *E. coli* strain, the latter strand is replaced with a thymine-containing strand.

We incorporated the *in vitro* plasmid resynthesis aspect of Kunkel SDM into our plasmid modification strategy (Fig. 2A). Initially, the sense strand bearing the promoter and gene of interest (GOI) was reacted with a nickase to induce a single-stranded cut. The nicked strand could then be removed with an exonuclease to give a cssDNA. In a similar manner to Kunkel mutagenesis, sequential primer annealing, extension, and ligation regenerated the plasmid. This approach was termed NEEL after the four enzymatic steps: Nicking, Exonuclease digestion, Extension, and Ligation.

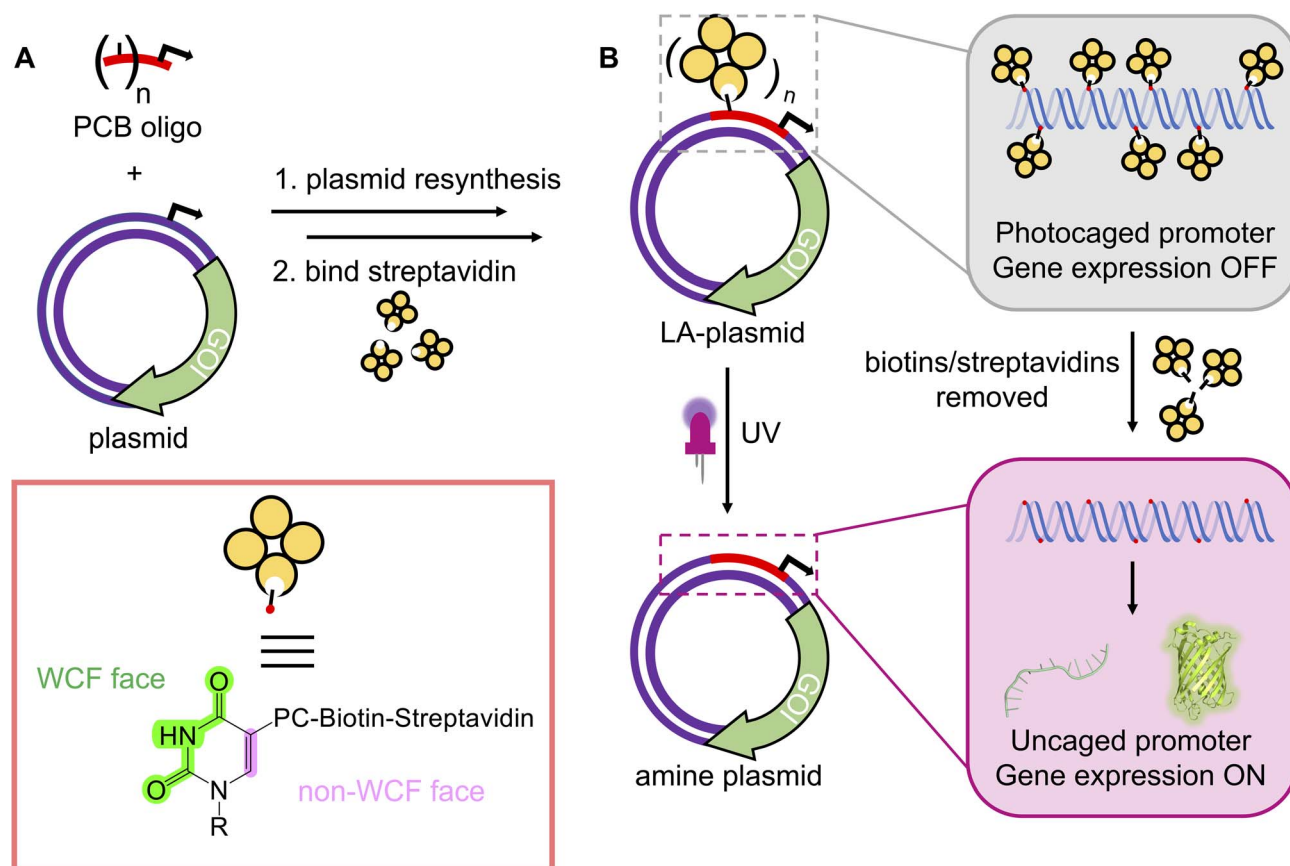


Fig. 1 The plasmid resynthesis strategy to photocage plasmids. (A) We have developed a method to introduce photocleavable (PC) biotin modifications site-specifically onto the non-WCF face of the nucleobases in a plasmid. (B) Photocaging groups are incorporated across the promoter upstream of a gene of interest (GOI). Gene expression is repressed until UV illumination removes biotins/streptavidins to reactivate the plasmids.



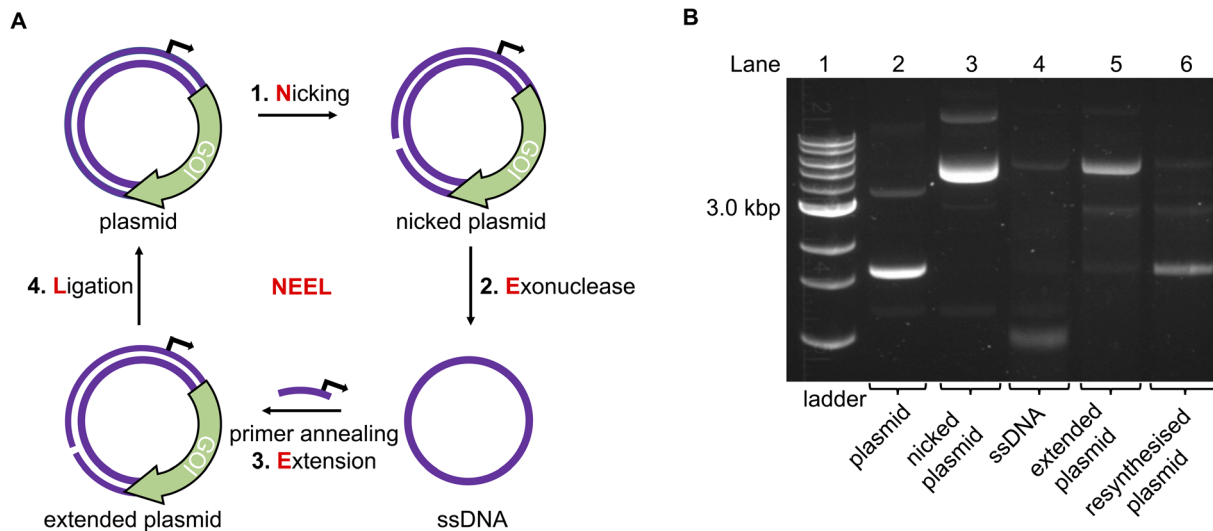


Fig. 2 The NEEL method to resynthesise a plasmid. (A) Overview of the enzymatic reactions in NEEL. (B) Using NEEL, our plasmid encoding mVenus was successfully resynthesised.

Initially, a plasmid encoding for the mVenus fluorescent protein under the control of a T7 promoter was used to develop NEEL. The recognition site for the nickase Nt·BspQI was cloned downstream of the mVenus open reading frame. Theoretically, the nicking site could be positioned anywhere on the plasmid, which is an advantage of this method over previous attempts to photocage plasmids.<sup>31,32</sup> Following nicking and exonuclease treatment, multiple Kunkel protocols from the literature were employed to perform NEEL (ESI Fig. S1†).<sup>40–43</sup> It was paramount that the plasmids obtained from NEEL were ligated, as it was well known that linear DNA templates are degraded by abundant exonucleases within the cellular milieu, which hampers their transfection efficiency.<sup>35,36</sup> Thus, to confirm ligation had occurred, the plasmids were assessed by agarose gel electrophoresis, with and without gyrase. Gyrase supercoils DNA and can only accept covalently-closed substrates, demonstrating the resynthesis of intact double-stranded plasmids. Without gyrase, the ligated plasmids appeared as ladder bands. However, with gyrase, these ladder bands were converted into a single band that co-migrated with the native supercoiled plasmid, confirming that ligation was successful. The non-ligated plasmid co-migrated with the nicked plasmid and remained at the same position even after reaction with gyrase.

The Kunkel protocol with the highest degree of ligation and supercoiling, and requiring a short reaction time (Kay *et al.*) was chosen for NEEL.<sup>43</sup> We determined that the crucial determinant for NEEL success was the use of T7 DNA Polymerase (DNAP), as ligated plasmids could not be generated with other DNAPs. To verify the generality of NEEL, the dinucleotide at the ligation junction was iteratively assessed. Sixteen primers with different  $T_m$ , GC content, and lengths, annealing at different positions across the plasmid were used for NEEL (ESI Table S1 and ESI Fig. S2A and B†). These primers were also chosen to include all the possible dinucleotide combinations at the ligation junction. Ligation was successful regardless of primer features, annealing location, and the dinucleotide identities at the ligation

junction. Having confirmed ligation was robust, gels were pre-stained hereafter as the resynthesised plasmid bands were more intense and easier to distinguish as they co-migrated with the native plasmid due to GelRed–DNA interactions (Fig. 2B).

To optimise the plasmid resynthesis, we wanted to remove any native plasmids that were not nicked or removed by exonuclease, which would contribute to an increase in the OFF-state when using photocages. To do this we used DpnI, a nuclease that preferentially cleaves dam-methylated DNA over non-methylated substrates.<sup>44</sup> NEEL-resynthesised plasmids were hemi-methylated, as they contained a parental methylated strand and a synthetic unmethylated strand. Although DpnI can cleave hemi-methylated DNA, its cleavage activity is significantly lower.<sup>45</sup> Therefore, mild DpnI conditions were optimised on the native plasmid, to minimise degradation of the resynthesised plasmid (ESI Fig. S3A†). Compared to the supplier's recommended conditions, plasmid degradation could be achieved using a much lower enzyme concentration and shorter reaction time. When tested with the resynthesised plasmid, faint bands denoting degradation of the methylated (major) and hemi-methylated plasmids (minor) were detected, but the resynthesised plasmid remained the most intense band (ESI Fig. S3B†).

### Photocaging the T7 promoter

With a robust method to resynthesise the plasmid in place, photocaging groups were then introduced to the sense strand of the promoter to control the expression of a downstream gene. To this end, the light-activated DNA (LA-DNA) platform previously developed in our group was used as a starting point.<sup>14</sup> In this approach, an oligo containing seven primary amines across the T7 promoter sequence is reacted with a UV-photocleavable 2-nitrobenzyl biotin *N*-hydroxysuccinimide (NHS) ester. The photocleavable biotin (PCB) oligo is then used as a primer in PCR to create a PCB DNA template, with the PCB T7 promoter upstream of a GOI. Binding streptavidin blocks T7 RNA Polymerase (RNAP) from accessing the promoter, thereby inhibiting



transcription. Conditional gene expression is restored upon UV-triggered removal of the biotins/streptavidins.

Analogous to the above LA-DNA templates, LA-plasmids were prepared by resynthesising the plasmid using a PCB T7 promoter oligo, followed by binding of streptavidin (Fig. 3A). The T7 amine primer was reacted with PCB-NHS ester, purified by HPLC, and analysed by LC-MS and denaturing PAGE (Fig. 3B, and ESI Fig. S4A and B†). Using mild UV conditions ( $\sim 1.06 \text{ mW cm}^{-2}$  for 4 minutes), multiple bands, representing successive removal of PCB groups, were observed. We then demonstrated that the PCB remained intact when subjected to annealing conditions, and its association with streptavidin was not disrupted (ESI Fig. S5†).

Excitingly, plasmid resynthesis using the amine or PCB T7 oligo gave amine and PCB plasmids, respectively, which illustrated T7 DNAP's high tolerance for base modifications (Fig. 3C). The latter plasmid was bound to either monovalent or wildtype tetravalent streptavidin (MVS and WTS, respectively) to give LA-plasmids. These could be observed in the gel by a decrease in electrophoretic mobility. Interestingly, little interplasmid crosslinking was observed using the WTS, which may be due to the concentrations involved and/or the favourability of intrastrand crosslinks (ESI Fig. S6†). As expected, UV illumination removed biotins/streptavidins to restore the amine

plasmids. Plasmids resynthesised using a regular primer without amino modifications did not show any changes using similar conditions (ESI Fig. S7†). This confirmed the requirement of both the amine and PCB groups to generate LA-plasmids.

These T7 promoter LA-plasmids were then applied in cell-free gene expression systems to determine their photo-modulation efficiency. *In vitro* transcription (IVT) with the T7 RNAP showed that the cssDNA was not transcriptionally active (Fig. 4A). Whereas the resynthesised plasmid without modifications or containing primary amine and PCB modifications all produced transcripts at the expected size, indicating that plasmids produced using the NEEL method were functionally active, and the T7 RNAP was tolerant to minor base modifications at the T7 promoter. However, when bulky streptavidin was appended, the T7 RNAP was unable to associate with its cognate promoter, and transcription was inhibited. Following UV illumination, the steric block was released and transcription from the plasmid was restored. Cell-free protein synthesis (CFPS), using the commercially available NEBExpress system, revealed this control was also manifested at the translational level (Fig. 4B and ESI Fig. S8†). Expression from caged plasmids was abolished (96% for both the MVS and WTS, compared to the 'amine plasmid -UV' condition). Upon UV-induced cleavage of

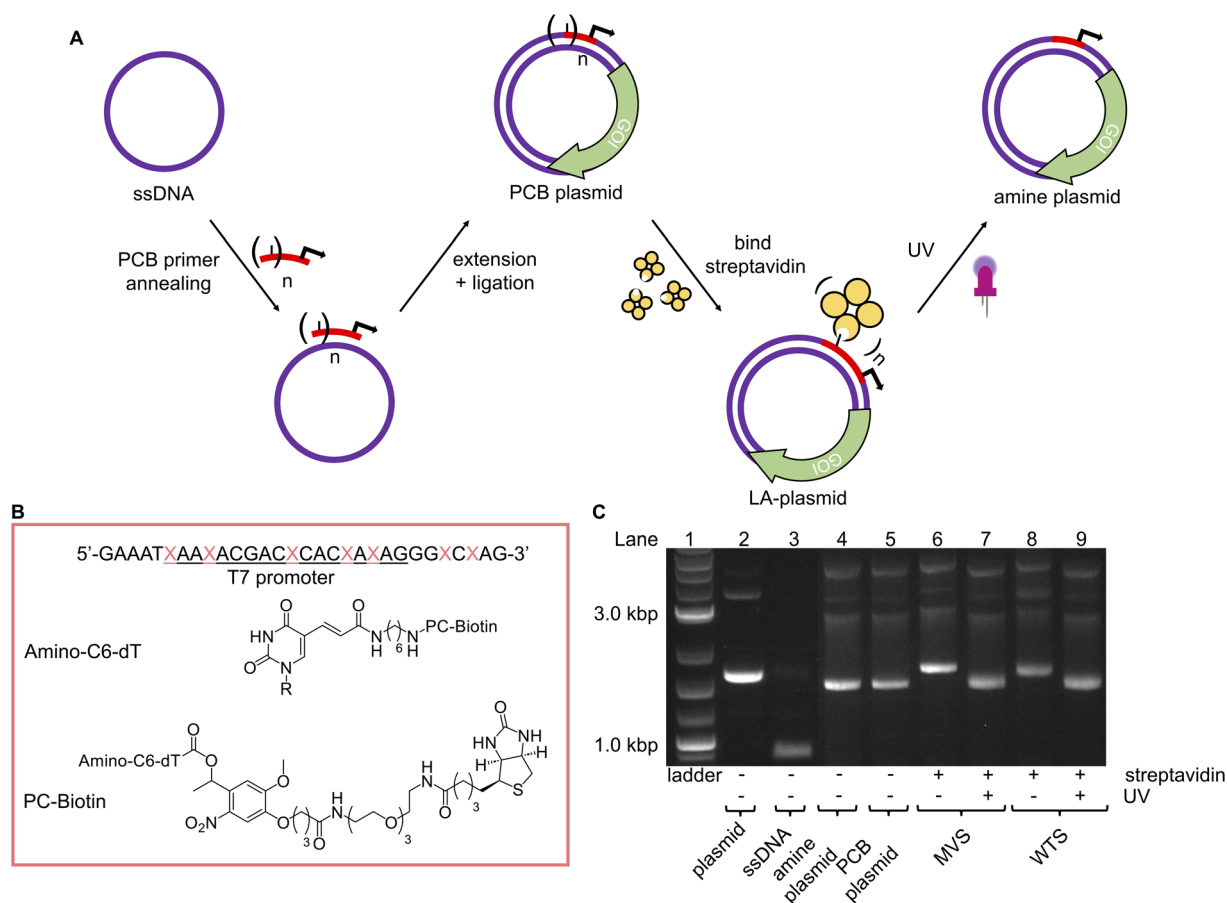


Fig. 3 LA-plasmids through photocaging the T7 promoter. (A) Overview of strategy to prepare LA-plasmids. (B) The T7 amine primer where the T7 promoter sequence is underlined, and X is the modified base. Structures of the amine nucleobase and PCB modifications are also shown. (C) T7 LA-plasmids and their photocaging.



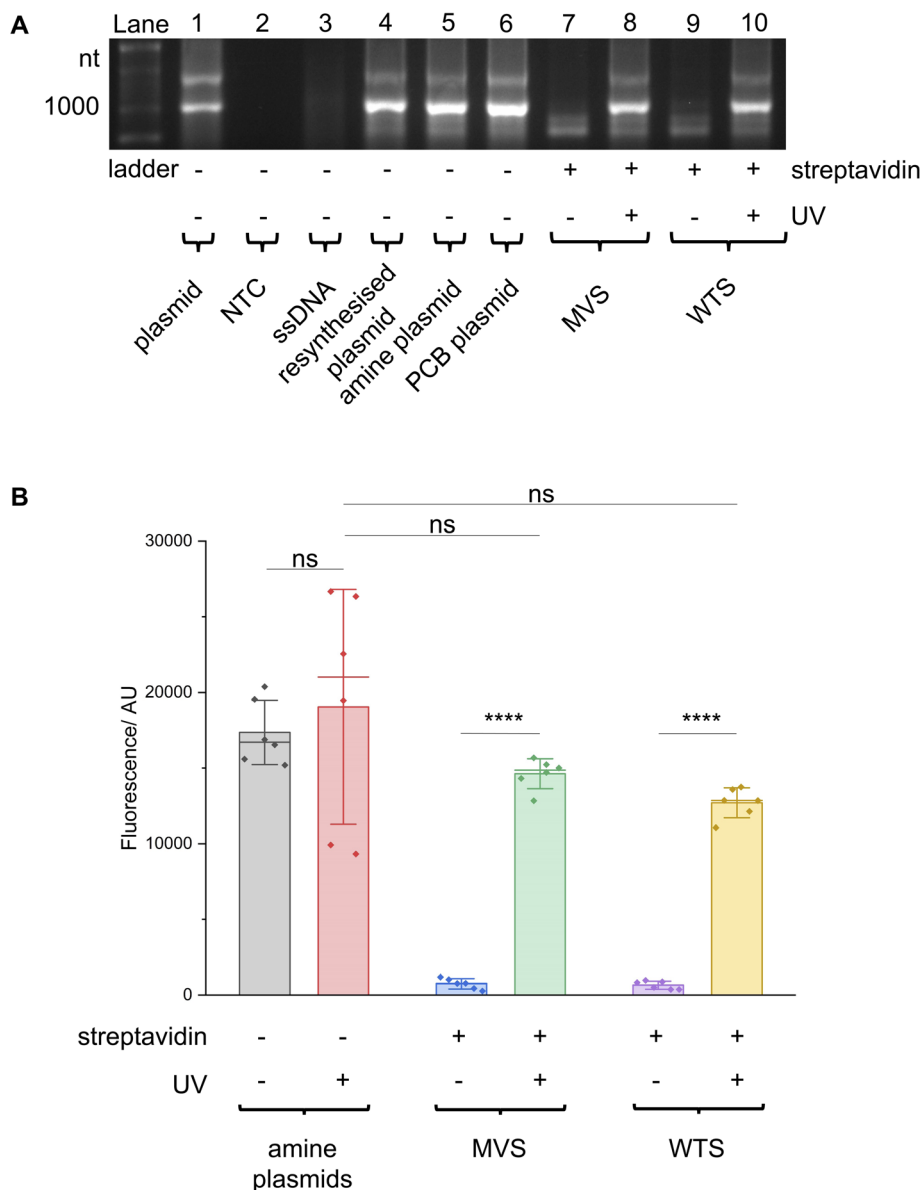


Fig. 4 T7 promoter photocaged LA-plasmid activity in cell-free gene expression. Plasmids resynthesised using NEEL were active in (A) IVT and (B) CFPS. When these plasmids were photocaged, the resultant LA-plasmids showed UV-dependent gene expression. In (B) fluorescence was background-subtracted from NTC ( $n = 3$ ). All other conditions were biological duplicates of technical triplicates ( $n = 6$ ). Error bars showed standard deviations. Two-tailed, unpaired Student's  $t$ -test was performed. ns = non-significant ( $P > 0.05$ ); \*\*\*\* $P \leq 0.0001$ .

the PCB moieties, the fluorescence increased by 19-fold for plasmids caged with either MVS or WTS. The uncaged MVS and WTS LA-plasmids obtained 77% and 67% relative to the 'amine

plasmid + UV' condition (although non-significantly different), respectively. These data clearly showed that NEEL produced T7 promoter LA-plasmids that tightly controlled gene expression.

Table 1 Comparison between plasmid core promoter elements and the corresponding consensus sequences in the CMV promoter<sup>a</sup>

Element	Consensus sequence	Plasmid sequence	Matching sequence
TATA box	TATAWAWR <sup>55</sup>	<u>TATATAAG</u>	8/8
Inr	YYANWYY <sup>56,57</sup>	<u>CCATTCT</u>	7/7
BRE <sup>u</sup>	SSRCGCC <sup>58</sup>	<u>GGA GG TC</u>	5/7
BRE <sup>d</sup>	RTDKKKK <sup>59</sup>	<u>CAATG CT</u>	4/7

<sup>a</sup> (W) A or T; (R) A or G; (S) G or C; (D) A, G, or T; (K) G or T; (Y) C or T; (N) A, C, G, or T. Plasmid sequence with underlined bases matched those of the consensus sequence.



### Photocaging the CMV promoter

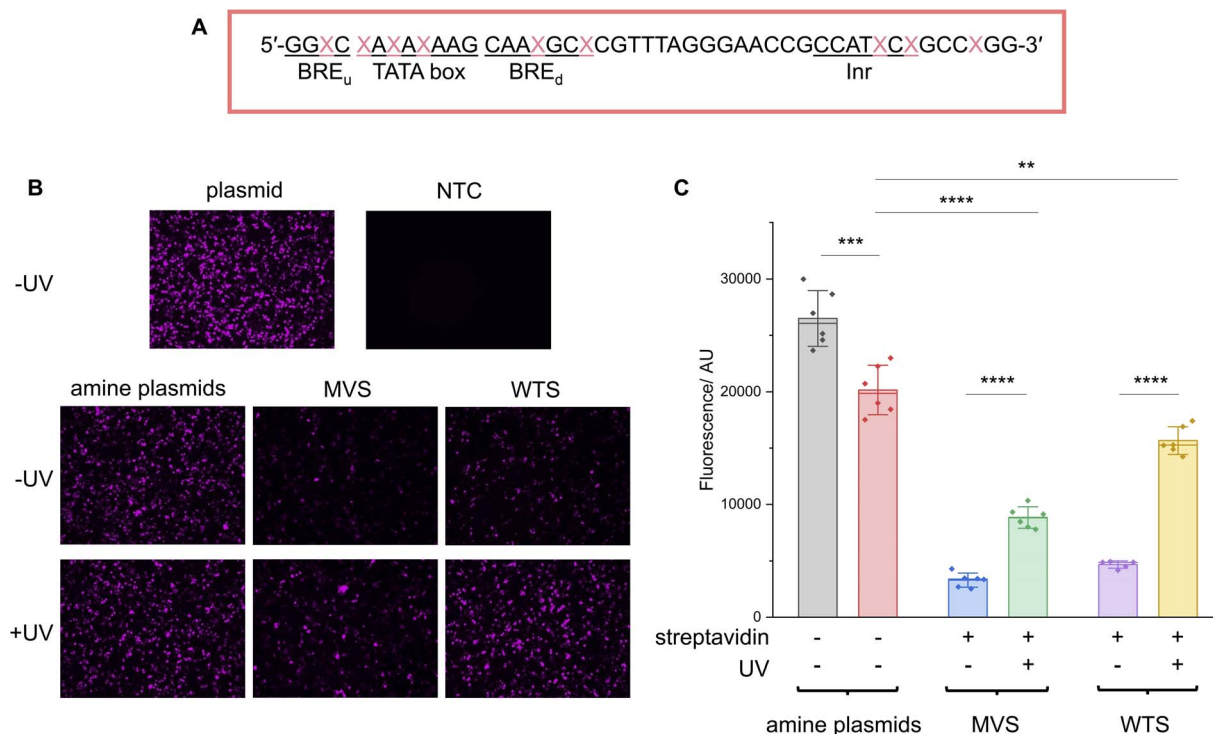
To demonstrate the generality of the NEEL protocol and LA-plasmid technology, we applied these approaches to exert conditional control of the cytomegalovirus (CMV) promoter, a widely used promoter sequence in eukaryotic expression systems. This promoter has been extensively studied and its core promoter elements are well-characterised, thus providing a good platform to illustrate our caging strategy in mammalian cells. Transcription initiation commences with the formation of the preinitiation complex (PIC) at the promoter.<sup>46–50</sup> This process begins with the recognition of the TATA-box sequence ~30 bp upstream of the transcription start site (TSS) by its cognate partner, the TATA-box binding protein (TBP), a subunit of the general transcription factor (GTF) TFIID.<sup>51</sup> The initiator element (Inr) is a 6–7 bp motif that covers the TSS and directs where transcription begins, and is also recognised by TFIID.<sup>52</sup> After TFIID binding, TBP recruits TFIIB whereby the latter can bind to the B-recognition elements immediately upstream and downstream of the TATA-box (BRE<sup>u</sup> and BRE<sup>d</sup>, respectively).<sup>53</sup> These early interactions engage other components of the PIC such as RNAP II and other GTFs.<sup>54</sup>

It has been shown *in vivo* that RNAP II transcription requires all the GTF components.<sup>46</sup> Their proximal positions relative to one another, as well as the availability of the full and partial consensus sequences (Table 1) permitted us to examine these

sites to exert light-mediated gene expression. The Nt-BbvCI recognition site was cloned into a plasmid containing the CMV promoter upstream of the green fluorescent reporter protein mNeonGreen (mNG), and this plasmid was subjected to the LA-plasmid caging strategy.

A CMV promoter oligo containing nine primary amine modifications in place of thymines across the regulatory regions was used to prepare the PCB oligo (Fig. 5A). Once again, LC-MS and denaturing PAGE were used to confirm PCB attachment (ESI Fig. S9A and B†). The UV parameters were also increased (9.88 mW cm<sup>-2</sup> for 10 minutes) to ensure the full removal of PCB groups and match the UV conditions previously used in the literature.<sup>21</sup> As has also been demonstrated by others,<sup>31</sup> these UV settings were not detrimental to cell viability (ESI Fig. S10†).

LA-plasmids containing either 9 MVS or WTS were prepared through plasmid resynthesis using the PCB CMV promoter oligo and analysed by agarose gel electrophoresis, demonstrating protein binding and release with UV illumination (ESI Fig. S11†). These plasmids were transfected into HEK293T cells and expression was assessed with and without illumination. In the absence of UV, transcription of mNG, measured by RT-qPCR, from the LA-plasmids was reduced by 82 and 80% relative to the 'amine plasmid -UV' condition for MVS and WTS LA-plasmids, respectively (ESI Fig. S12A†). Upon UV uncaging, transcription was increased by ~2 and 4.5 times for the MVS



**Fig. 5** CMV promoter photocaged LA-plasmid activity in mammalian cells. (A) The CMV amine primer, where the CMV core promoter elements are underlined and X is the modified base. The NEEL plasmids were transfected into HEK293T cells and assessed for their activity with and without UV irradiation. Expression was measured through (B). Fluorescence microscopy; images were representative of the respective conditions. (C) Fluorescence plate reader; the fluorescence was background-subtracted from NTC. Error bars showed standard deviations. Two-tailed, unpaired Student's *t*-test was performed. ns = non-significant ( $P > 0.05$ ); \*\* $0.001 \leq P \leq 0.01$ ; \*\*\* $0.0001 \leq P \leq 0.001$  and \*\*\*\* $P \leq 0.0001$ . For (B), 'plasmid' and 'NTC' conditions were technical triplicates ( $n = 3$ ), and all other conditions in (B) and (C) were biological duplicates of technical triplicates ( $n = 6$ ) except for the 'WTS-UV' condition ( $n = 5$  due to sample loss).



and WTS-caged plasmids, respectively. The latter represented a 100% recovery when compared to the 'amine + UV' control. Similar results were recapitulated at the translational level. Notably, epifluorescence microscopy confirmed that both LA-plasmid types were repressed (Fig. 5B). UV application had no discernible effect on the amine plasmid but gave a significant increase in fluorescence for the LA-plasmids. The modified plasmids and mild UV conditions employed did not affect the expression of the GAPDH housekeeping gene (ESI Fig. S13†). To quantify the change in mNG protein levels a fluorescence plate reader was used (Fig. 5C and ESI S12B†). Protein expression was diminished by 88% and 82% relative to the 'amine plasmid -UV' condition, and the increase in expression following illumination was 2.7 and 3.3-fold for the MVS and WTS LA-plasmids. Relative to the amine plasmid control, these numbers translated to a 44% and 78% recovery. This demonstrated light-activated control of expression in mammalian cells using commercially available materials.

## Conclusion

We have developed a simple and accessible approach to photoregulate plasmid expression in cell-free systems and mammalian cells. Inspired by the classical Kunkel mutagenesis method, NEEL was devised to resynthesise plasmids through four concerted enzymatic reactions using the nickase, exonuclease, polymerase, and ligase enzymes. Unlike previous approaches that relied on random, uncontrolled, and excessive caging of the phosphate backbones,<sup>25,29,33,34</sup> our method site-specifically installed the caging modifications by using a primer that contained photocleavable biotins (PCB). Additionally, our method allows for the placement of a nickase site anywhere on the plasmid, unlike previous approaches that required the incorporation of sequence-specific motifs around the promoter site.<sup>31,32</sup> We used our LA-plasmid technology to regulate expression from the T7 promoter in cell-free protein synthesis and the CMV promoter in mammalian cells. The incorporation of streptavidin proteins across the promoters provided a tight caging of activity that could be efficiently reconstituted following UV illumination. We believe these controllable plasmids might be especially relevant for studies in developmental biology<sup>26,27</sup> and population dynamics.<sup>60</sup>

Owing to the availability of nickases with diverse recognition sequences, we expect the LA-plasmid caging approach to be generalisable to most plasmid sequences. If plasmids do not contain suitable nickase sites to apply our method, these sites could be introduced through mutagenesis anywhere on the plasmid. Additionally, to allow full accessibility of our method, in the future, we envisage the use of a Cas9 nickase that, when combined with a designer gRNA, will allow our method to be applied to any plasmid sequence. With the flexibility of placement of the nickase site, the commercial availability of all reagents, and the simplicity of the workflow, we envisage this strategy could be easily adopted by non-chemistry-based laboratories to photoregulate plasmid expression. These advantages outweigh the extra steps of our method, compared to those from the literature.

The application of LA-plasmids in cell-free systems would allow the use of cheaper lysate systems that cannot accommodate linear DNA templates due to the presence of nucleases. Finally, the multivalency of streptavidin provides opportunities to incorporate other functionalities onto the plasmid. An interesting avenue is cell-penetrating moieties that would permit light-activated, self-delivered plasmid constructs. This is made possible by the availability of diverse biotinylated reagents such as transferrin and albumin that have been shown to enable gene delivery.<sup>61,62</sup> The recent approvals of plasmid-based gene therapies<sup>63–65</sup> and vaccines<sup>66,67</sup> in humans signalled a renewed interest in plasmids as therapeutic agents. Thus, the ability to precisely control their expression with a non-intrusive external stimulus such as light is timely, and could limit off-target expression and unexpected outcomes.

## Materials and methods

Unless otherwise stated, all enzymes were purchased from New England Biolabs (NEB). Seakem® LE Agarose was purchased from Lonza. XL-10 Gold competent *Escherichia coli* (*E. coli*) cells and SOC Outgrowth Medium were purchased from NEB. ATP, SuperScript IV VILO Master Mix with ezDNase, Fast SYBR Green Master Mix, Lipofectamine 3000, Fluorobrite DMEM, Opti-MEM I Reduced Serum Medium, GeneJet PCR Purification Kit were purchased from Thermo Fischer Scientific. Non-modified primers were purchased from Integrated DNA Technologies. Amine primers were purchased from ATD-Bio or Merck/Sigma. PCB-NHS ester was purchased from Click Chemistry Tools or Merck/Sigma and stored in dry dimethylformamide (DMF) at  $-80\text{ }^{\circ}\text{C}$ . HPLC solvents, trypsin-EDTA, FBS, Amicon 3 K columns,  $3\times$  GelRed, and TAE Buffer were purchased from Merck/Sigma. TEMED was purchased from Bio-Rad. Gel extraction, maxiprep, miniprep, and RNeasy Mini Kits were purchased from Qiagen. CellTiter-Glo Luminescent Cell Viability Assay was purchased from Promega. Pico488 dsDNA Quantification Reagent and the dsDNA quantitative standard ( $100\text{ ng }\mu\text{L}^{-1}$ ) in TE Buffer were purchased from Lumiprobe. HEK293T cell line was a gift from Dr Haiping Tang (Prof. Dame Carol Robinson's group). The mVenus plasmid (containing the T7 promoter) was available and previously prepared by Booth *et al.*<sup>14</sup> The pCS2-mNG-C (the CMV mNG plasmid) was purchased from Addgene (plasmid #128144).

### Agarose gel electrophoresis

100–150 ng DNA was analysed on a 1.0% pre-stained ( $1\times$  GelRed) agarose gel. The gel was run in  $1\times$  TAE buffer at 130 V for 55 min. Bands were imaged using a BioRad Molecular Imager® Geldoc™ XR + gel imager and analysed using ImageLab 6.1.

### Preparing the T7 mV and CMV mNG plasmids

Homologous recombination was used to insert the BbvCI and BspQI recognition sites into the T7 mVenus plasmid to create mV\_nicks, and the BbvCI recognition site into the CMV mNG plasmid to create mNG\_BbvCI (ESI Table S2†). For the T7 mVenus plasmid, the plasmid was digested to form a linear



fragment with BamHI before PCR amplification. For the CMV mNG plasmid, the circular plasmid was used directly for PCR. Separate PCR reactions were carried out to amplify the backbone and insert. The reactions were performed in 0.2 mL PCR tubes and contained the native or digested plasmid template (1–3 ng), the forward and reverse primers (1  $\mu\text{L}$ , 10  $\mu\text{M}$ , ESI Tables S3–S5†), Phusion High-Fidelity PCR Master Mix with HF Buffer (10  $\mu\text{L}$ , 2 $\times$ ), and made up to 20  $\mu\text{L}$  with MilliQ water.

Reactions were cycled on a Peqlab peqSTAR 96X Universal thermal cycler using the following conditions: initial denaturation at 98  $^{\circ}\text{C}$  for 30 seconds; 30 or 35 cycles [denaturation at 98  $^{\circ}\text{C}$  for 10 seconds, annealing at either 55  $^{\circ}\text{C}$  (mV\_nicks) or 68  $^{\circ}\text{C}$  (mNG\_BbvCI) for 30 seconds, and extension at 72  $^{\circ}\text{C}$  for either 1.5 minutes (backbone and insert of mV\_nicks) or 30 seconds (mNG\_BbvCI insert) and 2.5 minutes (mNG\_BbvCI backbone)]. A final extension at 72  $^{\circ}\text{C}$  for 10 minutes was performed. DpnI (0.5  $\mu\text{L}$ , 20 U  $\mu\text{L}^{-1}$ ) was added and reactions were incubated at 37  $^{\circ}\text{C}$  for 30 minutes and heat-deactivated at 80  $^{\circ}\text{C}$  for 20 minutes to remove the templates. The reactions were purified with a Genejet PCR column according to the manufacturer's protocol. Reactions were analysed with gel electrophoresis to confirm product bands.

A 7:1 insert:backbone ratio was used for homologous recombination and calculated according to NEBioCalculator. The backbone (100 ng) and insert were combined in a 1.5 mL Eppendorf tube. XL-10 Gold competent *E. coli* cells (50  $\mu\text{L}$ ) were added to the reaction. The tube was kept on ice for 30 minutes and heat-shocked at 42  $^{\circ}\text{C}$  for 30 seconds. SOC outgrowth medium (500  $\mu\text{L}$ ) was added and the tube was incubated at 37  $^{\circ}\text{C}$  with shaking at 225 rpm for 30 minutes. The tube was kept on ice for 2 minutes and 200  $\mu\text{L}$  of the transformed bacteria were plated on LB agar plates containing ampicillin (100  $\mu\text{g mL}^{-1}$ ) at 37  $^{\circ}\text{C}$  for 14–16 hours. Colonies were then picked and grown in 50 mL Falcon tubes containing 5 mL LB media with ampicillin (100  $\mu\text{g mL}^{-1}$ ) at 37  $^{\circ}\text{C}$  and shaking at 225 rpm for 14–16 hours. 300  $\mu\text{L}$  of the overnight cultures were combined with glycerol (300  $\mu\text{L}$ , 50% v/v, filtered). The glycerol stocks were kept at  $-80^{\circ}\text{C}$ . The rest of the cultures were miniprep using Qiagen's miniprep kit according to the manufacturer's protocol and eluted in their elution buffer (50  $\mu\text{L}$ ) and stored at  $-20^{\circ}\text{C}$ . Sanger sequencing was performed by SourceBioScience. After the sequence was confirmed, bulk plasmid was obtained by inoculating 2  $\times$  100 mL LB cultures with the desired glycerol stock. The cultures were grown as above and purified using Qiagen's Maxiprep kit and performed according to the manufacturer's protocol. The plasmids were resuspended in Tris (1 mL, 10 mM, pH 8.0) and stored at  $-20^{\circ}\text{C}$ .

### Nicking the T7 mV and CMV mNG plasmids

To a 1.5 mL Eppendorf tube, the mV\_nicks plasmid (75  $\mu\text{g}$ ), Nt-BspQI (62.6  $\mu\text{L}$ , 10 U  $\mu\text{L}^{-1}$ ), and NEBuffer 3.1 (62.6  $\mu\text{L}$ , 10 $\times$ ) were added and made up to 625  $\mu\text{L}$  with MilliQ water. The reactions were incubated on the 1.5 mL Thermomixer C (Eppendorf) with shaking at 600 rpm at 50  $^{\circ}\text{C}$  for 2 hours and heat-deactivated at 80  $^{\circ}\text{C}$  for 25 minutes.

The nicking reactions for the mNG\_BbvCI plasmid were performed analogously but instead used the Nt-BbvCI nickase (62.6

$\mu\text{L}$ , 10 U  $\mu\text{L}^{-1}$ ) and CutSmart buffer (62.6  $\mu\text{L}$ , 10 $\times$ ). The reactions were incubated on the Thermomixer C with shaking at 600 rpm at 37  $^{\circ}\text{C}$  for 2 hours and heat-deactivated as above. These reactions were then purified using Genejet PCR columns according to the manufacturer's protocols and eluted in their elution buffer (50  $\mu\text{L}$ ). The concentrations were measured using the Nanodrop ND-1000 Spectrophotometer and stored at  $-20^{\circ}\text{C}$ .

### Preparing cssDNA templates

The total reaction volumes varied and depended on the amount of nicked plasmid used. For example, when 60  $\mu\text{g}$  dsDNA template is used, the total reaction volume would be 500  $\mu\text{L}$ . The reactions were performed in 1.5 mL Eppendorf tubes and consisted of the plasmid (final concentration = 120.0 ng  $\mu\text{L}^{-1}$ ), Lambda Exonuclease (final concentration = 1.0 U  $\mu\text{L}^{-1}$ ), Lambda Exonuclease Reaction Buffer (final concentration = 1 $\times$ ), and made up to the desired volumes with MilliQ water. The reactions were incubated on a 1.5 mL Thermomixer C with shaking at 600 rpm at 37  $^{\circ}\text{C}$  for 3 hours, and heat-deactivated at 75  $^{\circ}\text{C}$  for 20 minutes. The reactions were purified using Genejet PCR columns according to the manufacturer's protocol and eluted in their elution buffer (50  $\mu\text{L}$ ). The concentrations were measured using the Nanodrop and stored at  $-20^{\circ}\text{C}$ .

### PCB-NHS ester reaction with the amine oligos

Sodium bicarbonate (1 M) was prepared fresh and used within 1 week. The reactions were performed in 1.5 mL Eppendorf tubes and consisted of the amine oligo (20  $\mu\text{L}$ , 100  $\mu\text{M}$ ), sodium bicarbonate (20  $\mu\text{L}$ , 1 M, pH  $\sim$  8.5), PCB-NHS ester (100  $\mu\text{L}$ , 10 mM), and made up to 200  $\mu\text{L}$  with MilliQ water. The reactions were incubated with shaking at 600 rpm at 20  $^{\circ}\text{C}$  overnight.

Tris (200  $\mu\text{L}$ , 10 mM, pH 8.0) was added and the solutions were purified on the Amicon 3 K 0.5 mL column by centrifugation at 14 000g for 15 minutes at room temperature. This was performed 3 times in total, each time decanting the eluate and making up the solution to 500  $\mu\text{L}$  with MilliQ water. The concentrate was eluted by inverting the column and centrifuging at 1000g for 2 minutes into a collection tube. This was repeated once more to maximise yield. The solution was made up to 300  $\mu\text{L}$  with MilliQ water.

### Purification of the PCB oligos

The PCB oligos were HPLC purified. Lines B/C/D were: MilliQ water, acetonitrile, and triethylammonium bicarbonate buffer (TEAB, 100 mM, pH 8.5), respectively. Line D was kept constant at 10%. Solvent percentages referred to line C, the organic phase. The flow rate was kept at 1.5 mL  $\text{minute}^{-1}$ . Separations were UV-monitored at wavelengths 260, 280, 310 and 360 nm.

The oligo (100  $\mu\text{L}$ ,  $\sim$ 6.67  $\mu\text{M}$ ) was injected into an Agilent Polaris 5 C18-A (150  $\times$  4.6 mm) analytical column heated to 50  $^{\circ}\text{C}$ . The column was equilibrated at 5% for 20 minutes. A gradient of 5–30% over 27 minutes was used to purify the oligo. The column was washed with 90% acetonitrile from 27.01 until 30 minutes. The organic phase was returned to 5% at 30.01 minutes and kept until 33 minutes for column equilibration.





Three injections were made, and the fractions were collected according to ESI Table S6.†

The fractions were then combined and lyophilised on the Christ Alpha 2–4 LDplus to dryness. The reactions were then resuspended in MilliQ water (500  $\mu\text{L}$ ) and purified using Amicon 3 K 0.5 mL columns as previously. The eluate was made up to 1 mL with MilliQ water and lyophilised to dryness. The PCB oligo was resuspended in Tris (50  $\mu\text{L}$ , 10 mM, pH 8.0) and the concentration was measured using the Nanodrop.

### Denaturing PAGE of amine and PCB oligos

A 40% w/v polyacrylamide (PAA) stock solution was prepared by dissolving 12 g PAA in MilliQ water and made up to 30 mL with MilliQ water, wrapped in tin foil and stored at 4 °C. A 10% w/v ammonium persulphate (APS) stock solution was prepared by dissolving APS (1 g) in MilliQ water and the solution was made up to 10 mL with MilliQ water and 1 mL aliquots were kept at –20 °C. A 1 mL 2 $\times$  loading dye was prepared and consisted of formamide (95% v/v), sodium dodecyl sulphate (0.02% w/v), xylene-cyanol (0.01% w/v), bromophenol blue (0.02% w/v), and EDTA (10  $\mu\text{L}$ , 1 M, pH 8.0).

The amine and PCB oligos were analysed on a denaturing PAGE. The 16% PAA gel solution was prepared in a 15 mL Falcon tube and consisted of TBE buffer (0.5 mL, 10 $\times$ ), PAA (2.0 mL, 40% w/v stock), GelRed (1 mL, 3 $\times$ ), urea (2.10 g). The solution was vigorously vortexed and sonicated and degassed at 35 °C for 10 minutes. The solution was vortexed again to ensure all the urea dissolved. APS (40  $\mu\text{L}$ , 10% w/v) and tetramethylethylenediamine (2.5  $\mu\text{L}$ ) were added to the PAA solution, inverted several times to mix, and quickly loaded onto a pre-assembled glass plate and allowed to set.

The oligo samples consisted of the nucleic acid (200 ng) and a homemade loading dye (5  $\mu\text{L}$ , 2 $\times$ ), and the solutions were made up to 10  $\mu\text{L}$  with MilliQ water. PCB oligo receiving UV was prepared without the dye, irradiated with UV and the dye was added as above. The solutions were run in 1 $\times$  TBE buffer at 200 V for 1 hour. The gel was soaked in GelRed (3 $\times$ ) for 5 minutes and imaged on a BioRad Molecular Imager® Geldoc™ XR + gel imager and analysed using ImageLab 6.1.

### LC-MS of PCB oligos

Line A was: 8.6 mM Et<sub>3</sub>N, 200 mM HFIP in 5% MeOH/H<sub>2</sub>O (v/v). Line B was: 20% buffer A in MeOH. Solvent percentages referred to line B.

Samples (10  $\mu\text{L}$ , 1  $\mu\text{M}$ ) were prepared and injected into a Waters Acquity UPLC BEH C18 column (130 Å, 1.7  $\mu\text{m}$ , 2.1 mm  $\times$  50 mm). Samples were analysed over 8 minutes with a 0–70% gradient. Masses were recorded on a Waters Xevo G2 QTOF ESI-UPLC-MS system. Analysis was carried out using MassLynx v4.1.

### Phosphorylation of oligos

The reactions were performed in 0.2 mL PCR tubes and consisted of the non-modified or amine or PCB oligo (60 pmol, ESI Table S3†), T4 Polynucleotide Kinase (2  $\mu\text{L}$ , 10 U  $\mu\text{L}^{-1}$ ), T4 Polynucleotide Kinase Reaction Buffer (4  $\mu\text{L}$ , 10 $\times$ ), ATP (0.4  $\mu\text{L}$ , 100 mM) and made up to 40  $\mu\text{L}$  with MilliQ water. The reactions

were incubated on the thermal cycler at 37 °C for 1 hour, heat-deactivated at 80 °C for 20 minutes, and stored at –20 °C.

### Annealing oligos to the cssDNA template

The total reaction volumes varied and depended on the amount of cssDNA used. For example, when 10  $\mu\text{g}$  cssDNA template is used, the total reaction volume would be 779  $\mu\text{L}$ . The molar mass of the cssDNA template was calculated based on the template length and assuming each base had an average mass of 330. The molar amount could then be deduced.

The annealing reactions were performed in 1.5 mL Eppendorf tubes and consisted of the cssDNA template (final concentration = 12.84 ng  $\mu\text{L}^{-1}$ ), the phosphorylated primer (4 molar equivalents relative to the cssDNA template), magnesium chloride (final concentration = 10 mM), Tris (pH 7.4, final concentration = 50 mM), and made up to the desired volumes with MilliQ water. The lids were parafilm to minimise evaporation. The reactions were incubated on the Thermomixer C with shaking at 600 rpm at 90 °C for 2 minutes and cooled slowly to 25 °C at a rate of –1 °C  $\text{minute}^{-1}$ .

### Extension and ligation reactions to resynthesise the plasmid

The total reaction volumes varied and depended on the amount of cssDNA used. For example, when 10  $\mu\text{g}$  cssDNA template is used, the total reaction volume would be 874  $\mu\text{L}$ . To the annealing reactions containing the cssDNA template (final concentration = 11.44 ng  $\mu\text{L}^{-1}$ ), each of the dNTP (final concentration = 0.8 mM), DTT (final concentration = 5 mM), T7 DNAP (final concentration = 0.05 U  $\mu\text{L}^{-1}$ ), T4 DNA Ligase (final concentration = 0.05 Weiss units per  $\mu\text{L}$ ) were added. The reactions were made up to the desired volumes with MilliQ water. The reactions were incubated on the Thermomixer C with shaking at 600 rpm at 20 °C for 3 hours and kept at 4 °C overnight.

### Assessing dinucleotides at the ligation junction

Each primer (0.23 pmol, ESI Table S1†) was phosphorylated similarly to above except a final concentration of 0.6 U  $\mu\text{L}^{-1}$  was used for the T4 PNK in a total reaction volume of 5  $\mu\text{L}$ . The annealing reaction was performed similarly to above except 67 ng of cssDNA template was annealed to 0.23 pmol phosphorylated primer in a 62.5  $\mu\text{L}$  total reaction volume. The extension and ligation reactions were performed similarly to above for each annealed mix, except a 75  $\mu\text{L}$  total reaction volume was used. The reactions were then lyophilised to dryness, reconstituted in MilliQ water (20  $\mu\text{L}$ ) and analysed on an agarose gel.

### Gyrase reaction to confirm ligation

The reactions were performed in 0.2 mL PCR tubes and consisted of the ligated plasmid (100 ng), Gyrase (0.5  $\mu\text{L}$ ), Gyrase Reaction Buffer (4.0  $\mu\text{L}$ , 5 $\times$ ), and made up to 20  $\mu\text{L}$  with MilliQ water. The reactions were incubated on the thermal cycler at 37 °C for 1 hour and heat-deactivated at 70 °C for 30 minutes.



### DpnI reaction to remove unmodified plasmids

After extension, the amount of DNA doubled. The total reaction volumes varied and depended on the amount of dsDNA used. For example, when 1  $\mu\text{g}$  dsDNA template is used, the total reaction volume would be 150  $\mu\text{L}$ . To the extension and ligation reaction tubes containing the ligated plasmid (final concentration = 6.7  $\text{ng } \mu\text{L}^{-1}$ ), magnesium chloride (final concentration = 10 mM), Tris (pH 7.4, final concentration = 50 mM), and DpnI (final concentration = 0.06  $\text{U } \mu\text{L}^{-1}$ ) were added. The reactions were made up to the desired volumes with MilliQ water. The reactions were incubated on the Thermomixer C with shaking at 600 rpm at 37  $^{\circ}\text{C}$  for 5 minutes. EDTA (pH 8.0, final concentration = 20 mM) was quickly added to stop the DpnI reaction. The reactions were lyophilised to dryness.

The reactions were resuspended in MilliQ water and purified using a GeneJet PCR column. The plasmids were eluted twice, each time with 50  $\mu\text{L}$  MilliQ water. Ethanol precipitation was performed to remove any residual contaminants. Briefly, sodium acetate (10  $\mu\text{L}$  or tenth volumes, 3 M, pH 5.2) and ethanol (300  $\mu\text{L}$  or 3 volumes, 99.8%) were added to the tubes. These were vortexed and kept at  $-80$   $^{\circ}\text{C}$  for 1–2 hours or  $-20$   $^{\circ}\text{C}$  overnight. The tubes were centrifuged at 15 000 rpm at 4  $^{\circ}\text{C}$  for 30 minutes. The supernatant was carefully removed whilst avoiding disturbing the pellets. Ethanol (1 mL, 70%) was added and the tubes were flicked and re-centrifuged at the above settings for 10 minutes. The supernatant was removed and the pellets were dried under vacuum on a speed vac for 10 minutes. Dried DNA pellets were re-suspended in Tris (volumes depended on the reaction scale, 10 mM, pH 8.0). The reactions were Nanodropped to determine purity, whilst the concentrations were estimated using Pico488 dye.

### Measuring concentrations with Pico488 dye

As PCB plasmids contained 2-nitrobenzyl groups, which absorb at 260 nm, their concentrations were determined using Pico488 dye. To ensure consistency, concentrations of the native and amine plasmids were also quantified using Pico488 before cell-free reactions and transfection. The standards were prepared using a dsDNA quantitative standard (Lumiprobe). Standards and samples were prepared according to the manufacturer's protocol. The solutions were transferred to a clear lid 96-well flat bottom with a low evaporation lid and centrifuged at 1250 rpm for 1 minute to remove any bubbles. The fluorescence was measured on a Tecan Infinite M1000 Pro fluorescence plate reader. Plates were shaken at 654 rpm for 10 seconds and the fluorescence was measured with  $\lambda_{\text{ex}} = 503$  nm,  $\lambda_{\text{em}} = 525$  nm, and bandwidth = 5 nm. The readings were used to plot the calibration curve. The equation for the line of best fit was used to calculate the concentrations.

### Binding streptavidin

The total reaction volumes depended on the mass of dsDNA used. For example, when 1  $\mu\text{g}$  dsDNA template is used, the total reaction volume would be 20  $\mu\text{L}$ . The molar mass of dsDNA was calculated based on the length of the plasmid and assuming the mass of

each base pair was 660. The molar amount of plasmid, and hence PCB could then be calculated, based on the number of PCBs on the PCB oligo (7 for the T7 promoter and 9 for the CMV promoter). A mass of 52.8 kDa was assumed for the monovalent and wildtype tetravalent streptavidin.

Reactions were performed in Protein LoBind tubes and consisted of the PCB plasmid (final concentration = 50.0  $\text{ng } \mu\text{L}^{-1}$ ), either MVS or WTS (3 molar equivalents relative to the amount of PCB sites), Tris (pH 7.4, final concentration = 20 mM), and made up to the desired volumes with MilliQ water. The reactions were incubated on the Thermomixer C with shaking at 600 rpm at 26  $^{\circ}\text{C}$  overnight.

### UV irradiation

Cell-free reactions were assembled and irradiated immediately before incubation at 37  $^{\circ}\text{C}$ . Cells were irradiated 4 hours after transfection. The lamp (365 nm, 700 mA, M365-L2, Thorlabs), equipped with a collimation adapter (COP5-A, Thorlabs) was collimated at a distance of  $\sim 22.5$  cm (from the base of the lamp to the bottom of the tube/plate). The UV intensity was measured with a photodiode power sensor (200–1100 nm, 50 mW,  $\varnothing = 9.5$  mm, S120VC, Thorlabs) equipped with a power meter console (PM100A, Thorlabs). The sensor measured UV power in mW. These values were then converted into irradiance units by dividing the power by the area of the aperture in  $\text{cm}^2$ . The zone of irradiation was large enough to irradiate up to 4 tubes/wells simultaneously. An irradiance of 2.47 (for the T7 promoter) or 9.88  $\text{mW cm}^{-2}$  (for the CMV promoter) was applied for 10 minutes to uncage the LA-plasmids. To ensure consistency, conditions not receiving UV were kept at room temperature whilst the +UV conditions were being irradiated. Cells were re-incubated at 37  $^{\circ}\text{C}$  and 5%  $\text{CO}_2$  for 2 days.

### IVT reactions of the T7 mV\_nicks plasmid

The reactions were assembled in 0.2 mL PCR tubes and consisted of the plasmid template (200 ng), RNAPol Reaction Buffer (2  $\mu\text{L}$ , 10 $\times$ ), each of the NTP (0.1  $\mu\text{L}$ , 100 mM), RNase Inhibitor (0.50  $\mu\text{L}$ , 40  $\text{U } \mu\text{L}^{-1}$ ), T7 RNAP (2  $\mu\text{L}$ , 50  $\text{U } \mu\text{L}^{-1}$ ), and made up to 20  $\mu\text{L}$  with MilliQ water. Reactions receiving UV were irradiated as above. They were then incubated on the thermal cycler at 37  $^{\circ}\text{C}$  for 1 hour. DNase I (1  $\mu\text{L}$ ) was then added, and the reactions were incubated on the thermal cycler at 37  $^{\circ}\text{C}$  for another 10 minutes to remove the DNA template. To stop the reactions, EDTA (1.10  $\mu\text{L}$ , 100 mM) was added, and the reactions were heat-deactivated on the thermal cycler at 75  $^{\circ}\text{C}$  for 10 minutes. 15  $\mu\text{L}$  of the reactions were mixed with RNA loading dye (15  $\mu\text{L}$ , 2 $\times$ ) and denatured on the thermal cycler at 90  $^{\circ}\text{C}$  for 2 minutes, kept on ice for 2 minutes and analysed on an agarose gel. All gel equipment was cleaned with RNaseZap beforehand to remove RNase contaminants.

### CFPS reactions of the T7 mV\_nicks plasmid

Reactions were performed in 0.2 mL PCR tubes and consisted of the plasmid (100 ng), NEBexpress<sup>®</sup> S30 Synthesis Extract (2.4  $\mu\text{L}$ ), Protein Synthesis Buffer (5.0  $\mu\text{L}$ , 2 $\times$ ), GamS Nuclease Inhibitor (0.2  $\mu\text{L}$ , 1.5  $\mu\text{g } \mu\text{L}^{-1}$ ), Murine RNase Inhibitor (0.2  $\mu\text{L}$ ,



40  $\mu\text{L}^{-1}$ ), T7 RNAP (0.2  $\mu\text{L}$ , 450  $\text{U } \mu\text{L}^{-1}$ ), and made up to 10  $\mu\text{L}$  with MilliQ water. Conditions requiring UV were irradiated as above. The reactions were incubated on the thermal cycler at 37  $^{\circ}\text{C}$  for 3 hours and kept at 4  $^{\circ}\text{C}$  for 30 minutes. The reactions were made up to 40  $\mu\text{L}$  with MilliQ water and transferred to a 384-well black, flat-bottom optiplate. The plate was centrifuged at 1250 rpm for 1 minute to remove any bubbles. The fluorescence was measured on a Tecan Infinite M1000 Pro fluorescence plate reader. Plates were shaken at 654 rpm for 10 seconds and the fluorescence was measured with  $\lambda_{\text{ex}} = 515 \text{ nm}$ ,  $\lambda_{\text{em}} = 527 \text{ nm}$ , and bandwidth = 5 nm. Data were plotted on Origin.

### Maintaining HEK293T cells

The full media contained Fluorobrite DMEM, FBS (10% v/v), Non-Essential Amino Acids (1 $\times$ ), and GlutaMAX (1 $\times$ ). Cell stocks were kept in liquid nitrogen. Cells were maintained for 20–25 passages before a new stock was thawed and immediately revived. Cells were passaged every 3 to 4 days, or earlier if confluency exceeded 80%.

To passage, the media was removed. PBS buffer (10 mL) was gently added and the flask was swirled and PBS was removed. Trypsin (2 mL) was added and the flask was incubated at 37  $^{\circ}\text{C}$ , 5%  $\text{CO}_2$  for  $\sim$ 1 minute. When the majority of cells detached, trypsin was neutralised with full media (8 mL) and the solution was transferred to a 15 mL Falcon tube. The cells were centrifuged at 1000 rpm for 3 minutes. The supernatant was carefully removed without disturbing the cell pellet. The cells were resuspended in full media (10 mL) and split in a 1:5–1:10 ratio depending on the cell density and the user's need. Cells were incubated at 37  $^{\circ}\text{C}$  and 5%  $\text{CO}_2$ .

### Cell viability assay

The CellTiter-Glo<sup>®</sup> Reagent was prepared according to the manufacturer's instructions and 2 mL aliquots in 15 mL Falcon tubes were wrapped in tin foil and kept at  $-20^{\circ}\text{C}$ . When needed, the reagent was thawed at room temperature and used immediately.

After cells were resuspended in full media (see above), the cell number was determined using the Countess II FL cell counter (Thermo Fisher Scientific). 10  $\mu\text{L}$  of the suspension was mixed with Tryptan blue (10  $\mu\text{L}$ ) and two 10  $\mu\text{L}$  measurements were recorded. The average was taken and used to prepare the working solution for plating cells. The volume of cells required was transferred to a 50 mL Falcon tube and made up to the desired volume with full media to give a final concentration of  $6.0 \times 10^2$  cells per  $\mu\text{L}$ . 100  $\mu\text{L}$  ( $6.0 \times 10^4$  cells per well) was seeded onto a 96-well plate with a flat bottom and a low evaporation lid and kept at 37  $^{\circ}\text{C}$  and 5%  $\text{CO}_2$ . After growth to  $>80\%$  confluency within 24 hours, cells were irradiated with the conditions in ESI Fig. S10<sup>†</sup> using the UV set-up described previously and re-incubated at 37  $^{\circ}\text{C}$  and 5%  $\text{CO}_2$ .

The following day, the CellTiter Reagent (100  $\mu\text{L}$ ) was added to each well and the wells were incubated on the Thermomixer C with shaking at 600 rpm at 20  $^{\circ}\text{C}$  for 2 minutes to induce cell lysis. The plates were incubated under tin foil at room

temperature for 10 minutes to stabilise the luminescence signal. Chemiluminescence was measured on a Tecan Infinite M1000 Pro fluorescence plate reader. The plates were shaken at 654 rpm for 10 seconds and the luminescence was measured using the luminescence scan mode between 559–561 nm. Values at 560 nm were used for analysis and data were plotted on Origin.

### Transfection

Before seeding cells, the wells were coated with poly-D-lysine (50  $\mu\text{L}$ ) for 5 minutes and washed twice with PBS (100  $\mu\text{L}$  each time). The cells were counted and plated as above, except  $5.5 \times 10^4$  cells in 180  $\mu\text{L}$  media were used to seed each well. The cells were incubated at 37  $^{\circ}\text{C}$  and 5%  $\text{CO}_2$  overnight.

Transfection with Lipofectamine<sup>™</sup> 3000 was performed according to the manufacturer's protocol. A 2:1 Lipofectamine 3000:DNA ratio, and 100 ng DNA were used for each well. A Lipofectamine 3000 mastermix and separate DNA mastermixes were prepared and made up to the desired volumes with MilliQ water. The Lipofectamine 3000 mastermix was added to each DNA mastermix and vortexed. The reactions were incubated at room temperature for 15 minutes to form lipoplexes. 12.40  $\mu\text{L}$  of the reaction was transfected into each well and the solution was homogenised by gently rocking the plates in a circular motion. The plates were incubated at 37  $^{\circ}\text{C}$  and 5%  $\text{CO}_2$  for 4 hours, UV was then applied as above, and the plates were re-incubated for 2 days.

### RNA extraction

A working solution of buffer RLT containing beta-mercaptoethanol ( $\beta\text{ME}$ ) was prepared fresh (10  $\mu\text{L}$   $\beta\text{ME}$  per mL buffer RLT).

The media in each well was removed and the wells were washed with PBS (100  $\mu\text{L}$ ). Trypsin (100  $\mu\text{L}$ ) was added to each well and the plates were kept at 37  $^{\circ}\text{C}$  and 5%  $\text{CO}_2$  until the majority of cells detached. Trypsin was neutralised with full media (100  $\mu\text{L}$ ) and the solutions were transferred to 1.5 mL Eppendorf tubes. RNA extraction was performed using RNeasy Mini Kit according to the manufacturer's protocol. RNA was eluted with RNase-free water (50  $\mu\text{L}$ ). RNA was ethanol-precipitated as above and resuspended in RNase-free water (20  $\mu\text{L}$ ). Concentrations were measured using the NanoDrop and the 260/230 ratios were typically  $>1.90$ . RNA was stored at  $-80^{\circ}\text{C}$ .

### ezDNase and cDNA synthesis

gDNA was removed from RNA preparations with ezDNase according to the manufacturer's protocol except the reaction was scaled to 6  $\mu\text{L}$  total volume and 60 ng RNA was used.

To each tube, Superscript IV Vilo MM (2.40  $\mu\text{L}$ , 5 $\times$ ) was added and the reactions were made up to 12.0  $\mu\text{L}$  with MilliQ water. The reactions were mixed and incubated on the thermal cycler with the following program: annealed at 25  $^{\circ}\text{C}$  for 10 minutes, reverse transcribed at 50  $^{\circ}\text{C}$  for 10 minutes, and heat-deactivated at 85  $^{\circ}\text{C}$  for 5 minutes. The cDNA was stored at  $-80^{\circ}\text{C}$ .



## RT-qPCR

Reactions were performed in 0.2 mL PCR tubes and consisted of the cDNA (2  $\mu$ L), mNG\_qPCR\_FOR (0.4  $\mu$ L, 10  $\mu$ M), mNG\_qPCR\_REV (0.4  $\mu$ L, 10  $\mu$ M, ESI Table S3<sup>†</sup>), Fast SYBR Green MM (10  $\mu$ L, 2 $\times$ ), and made up to 20  $\mu$ L with MilliQ water to amplify mNG. Normalisation with GAPDH was performed analogously using GAPDH\_FOR\_Origene and GAPDH\_REV\_IDT as primers (ESI Table S3<sup>†</sup>). Reactions were cycled on the CFX96 Real-Time System (C1000 Touch™ Thermalcycler, Bio-Rad) with SYBR Green monitoring using the following parameters: denaturation at 95  $^{\circ}$ C for 20 seconds, 30 cycles [95  $^{\circ}$ C for 3 seconds, 60  $^{\circ}$ C for 30 seconds]. The  $\Delta\Delta$ Ct method was used to quantify the transcription fold change. The data were set relative to 'amine plasmid-UV' condition and plotted using Origin.

## Plate reader of CMV mNG plasmids

After 2 days, the fluorescence was measured on the plate reader. Plates were shaken at 654 rpm for 10 seconds and the fluorescence was measured with  $\lambda_{\text{ex}} = 504$  nm,  $\lambda_{\text{em}} = 517$  nm, and bandwidth = 5 nm. Data were background-subtracted and plotted on Origin.

## Epifluorescence microscopy of CMV mNG plasmids

Cells were imaged using a Leica DMi8 inverted epifluorescence microscope with a 10 $\times$  objective lens. Cells were imaged with the brightfield and GFP channels. Analysis was performed with ImageJ.

## Author contributions

K. C. and M. J. B. designed the project. K. C. designed, performed, and analysed the experiments, with contributions from M. J. B. K. C. and M. J. B. wrote the paper.

## Data availability

All the data generated in this study are available within the article, the ESI,<sup>†</sup> figures and source data. Source data are available from DOI: [10.5281/zenodo.10048623](https://doi.org/10.5281/zenodo.10048623).

## Conflicts of interest

The authors declare no conflict of interest.

## Acknowledgements

We would like to thank Prof. M. Howarth and Dr M. Fairhead for kindly providing the monovalent streptavidin, and Professor Dame Carol Robison and Dr Haiping Tang for providing the HEK293T cell line. K. C. is supported through the Synthetic Biology Centre for Doctoral Training – EPSRC (EP/L016494/1) and CRUK funding (C2195/A27450). M. J. B. is supported by a Royal Society University Research Fellowship and an EPSRC New Investigator Award (EP/V030434/2).

## References

- 1 D. Hartmann, J. M. Smith, G. Mazzotti, R. Chowdhry and M. J. Booth, Controlling gene expression with light: a multidisciplinary endeavour, *Biochem. Soc. Trans.*, 2020, **48**, 1645–1659.
- 2 A. Möglich, R. A. Ayers and K. Moffat, Design and Signaling Mechanism of Light-Regulated Histidine Kinases, *J. Mol. Biol.*, 2009, **385**, 1433–1444.
- 3 R. Ohlendorf, R. R. Vidavski, A. Eldar, K. Moffat and A. Möglich, From Dusk till Dawn: One-Plasmid Systems for Light-Regulated Gene Expression, *J. Mol. Biol.*, 2012, **416**, 534–542.
- 4 S. Konermann, *et al.*, Optical control of mammalian endogenous transcription and epigenetic states, *Nature*, 2013, **500**, 472–476.
- 5 J. J. Tabor, A. Levskaia and C. A. Voigt, Multichromatic Control of Gene Expression in Escherichia coli, *J. Mol. Biol.*, 2011, **405**, 315–324.
- 6 M. J. Kennedy, *et al.*, Rapid blue-light-mediated induction of protein interactions in living cells, *Nat. Methods*, 2010, **7**, 973–975.
- 7 L. B. Motta-Mena, *et al.*, An optogenetic gene expression system with rapid activation and deactivation kinetics, *Nat. Chem. Biol.*, 2014, **10**, 196–202.
- 8 H. Ye, M. D.-E. Baba, R.-W. Peng and M. Fussenegger, A Synthetic Optogenetic Transcription Device Enhances Blood-Glucose Homeostasis in Mice, *Science*, 2011, **332**, 1565–1568.
- 9 S. Shimizu-Sato, E. Huq, J. M. Tepperman and P. H. Quail, A light-switchable gene promoter system, *Nat. Biotechnol.*, 2002, **20**, 1041–1044.
- 10 S. R. Schmidl, R. U. Sheth, A. Wu and J. J. Tabor, Refactoring and Optimization of Light-Switchable Escherichia coli Two-Component Systems, *ACS Synth. Biol.*, 2014, **3**, 820–831.
- 11 H. Lusic, D. D. Young, M. O. Lively and A. Deiters, Photochemical DNA Activation, *Org. Lett.*, 2007, **9**, 1903–1906.
- 12 Z. Vaníková, M. Janoušková, M. Kambová, L. Krásný and M. Hocek, Switching transcription with bacterial RNA polymerase through photocaging, photorelease and phosphorylation reactions in the major groove of DNA, *Chem. Sci.*, 2019, **10**, 3937–3942.
- 13 L. Kröck and A. Heckel, Photoinduced Transcription by Using Temporarily Mismatched Caged Oligonucleotides, *Angew. Chem., Int. Ed.*, 2005, **44**, 471–473.
- 14 M. J. Booth, V. R. Schild, A. D. Graham, S. N. Olof and H. Bayley, Light-activated communication in synthetic tissues, *Sci. Adv.*, 2016, **2**, e1600056.
- 15 M. J. Booth, V. Restrepo Schild, S. J. Box and H. Bayley, Light-patterning of synthetic tissues with single droplet resolution, *Sci. Rep.*, 2017, **7**, 9315.
- 16 X. Tang and I. J. Dmochowski, Phototriggering of Caged Fluorescent Oligodeoxynucleotides, *Org. Lett.*, 2005, **7**, 279–282.
- 17 M. A. H. Fichte, *et al.*, Three-Dimensional Control of DNA Hybridization by Orthogonal Two-Color Two-Photon Uncaging, *Angew. Chem., Int. Ed.*, 2016, **55**, 8948–8952.



- 18 D. D. Young, M. O. Lively and A. Deiters, Activation and Deactivation of DNAzyme and Antisense Function with Light for the Photochemical Regulation of Gene Expression in Mammalian Cells, *J. Am. Chem. Soc.*, 2010, **132**, 6183–6193.
- 19 X. Tang, J. Swaminathan, A. M. Gewirtz and I. J. Dmochowski, Regulating gene expression in human leukemia cells using light-activated oligodeoxynucleotides, *Nucleic Acids Res.*, 2008, **36**, 559–569.
- 20 J. M. Govan, *et al.*, Cellular Delivery and Photochemical Activation of Antisense Agents through a Nucleobase Caging Strategy, *ACS Chem. Biol.*, 2013, **8**, 2272–2282.
- 21 X. Tang, S. Maegawa, E. S. Weinberg and I. J. Dmochowski, Regulating Gene Expression in Zebrafish Embryos Using Light-Activated, Negatively Charged Peptide Nucleic Acids, *J. Am. Chem. Soc.*, 2007, **129**, 11000–11001.
- 22 D. D. Young, H. Lusic, M. O. Lively, J. A. Yoder and A. Deiters, Gene Silencing in Mammalian Cells with Light-Activated Antisense Agents, *ChemBioChem*, 2008, **9**, 2937–2940.
- 23 I. A. Shestopalov, S. Sinha and J. K. Chen, Light-controlled gene silencing in zebrafish embryos, *Nat. Chem. Biol.*, 2007, **3**, 650–651.
- 24 J. M. Govan, M. O. Lively and A. Deiters, Photochemical Control of DNA Decoy Function Enables Precise Regulation of Nuclear Factor  $\kappa$ B Activity, *J. Am. Chem. Soc.*, 2011, **133**, 13176–13182.
- 25 H. Ando, T. Furuta, R. Y. Tsien and H. Okamoto, Photo-mediated gene activation using caged RNA/DNA in zebrafish embryos, *Nat. Genet.*, 2001, **28**, 317–325.
- 26 E. V. Moroz-Omori, *et al.*, Photoswitchable gRNAs for Spatiotemporally Controlled CRISPR-Cas-Based Genomic Regulation, *ACS Cent. Sci.*, 2020, **6**, 695–703.
- 27 W. Zhou, *et al.*, Spatiotemporal Control of CRISPR/Cas9 Function in Cells and Zebrafish using Light-Activated Guide RNA, *Angew. Chem., Int. Ed.*, 2020, **59**, 8998–9003.
- 28 X. Huang, *et al.*, Light-activated RNA interference in human embryonic stem cells, *Biomaterials*, 2015, **63**, 70–79.
- 29 W. T. Monroe, M. M. McQuain, M. S. Chang, J. S. Alexander and F. R. Haselton, Targeting Expression with Light Using Caged DNA, *J. Biol. Chem.*, 1999, **274**, 20895–20900.
- 30 S. Yamaguchi, Y. Chen, S. Nakajima, T. Furuta and T. Nagamune, Light-activated gene expression from site-specific caged DNA with a biotinylated photolabile protection group, *Chem. Commun.*, 2010, **46**, 2244–2246.
- 31 J. Hemphill, J. Govan, R. Uprety, M. Tsang and A. Deiters, Site-Specific Promoter Caging Enables Optochemical Gene Activation in Cells and Animals, *J. Am. Chem. Soc.*, 2014, **136**, 7152–7158.
- 32 S. Ogasawara, Transcription Driven by Reversible Photocontrol of Hyperstable G-Quadruplexes, *ACS Synth. Biol.*, 2018, **7**, 2507–2513.
- 33 M. K. G. Jayakumar, N. M. Idris and Y. Zhang, Remote activation of biomolecules in deep tissues using near-infrared-to-UV upconversion nanotransducers, *Proc. Natl. Acad. Sci.*, 2012, **109**, 8483–8488.
- 34 A. Schroeder, *et al.*, Remotely Activated Protein-Producing Nanoparticles, *Nano Lett.*, 2012, **12**, 2685–2689.
- 35 J.-Y. Cherng, *et al.*, Effect of DNA topology on the transfection efficiency of poly((2-dimethylamino)ethyl methacrylate)–plasmid complexes, *J. Controlled Release*, 1999, **60**, 343–353.
- 36 K. Remaut, N. N. Sanders, F. Fayazpour, J. Demeester and S. C. De Smedt, Influence of plasmid DNA topology on the transfection properties of DOTAP/DOPE lipoplexes, *J. Controlled Release*, 2006, **115**, 335–343.
- 37 B. J. Ledwith, *et al.*, Plasmid DNA Vaccines: Investigation of Integration into Host Cellular DNA following Intramuscular Injection in Mice, *INT*, 2000, **43**, 258–272.
- 38 T. A. Kunkel, Rapid and efficient site-specific mutagenesis without phenotypic selection, *Proc. Natl. Acad. Sci.*, 1985, **82**, 488–492.
- 39 K. Bebenek and T. A. Kunkel, The use of native T7 DNA polymerase for site-directed mutagenesis, *Nucleic Acids Res.*, 1989, **17**, 5408.
- 40 M. A. Kay, C.-Y. He and Z.-Y. Chen, A robust system for production of minicircle DNA vectors, *Nat. Biotechnol.*, 2010, **28**, 1287–1289.
- 41 R. Tonikian, Y. Zhang, C. Boone and S. S. Sidhu, Identifying specificity profiles for peptide recognition modules from phage-displayed peptide libraries, *Nat. Protoc.*, 2007, **2**, 1368–1386.
- 42 P. Handa and U. Varshney, Rapid and reliable site directed mutagenesis using Kunkel's approach, *Indian J. Biochem. Biophys.*, 1998, **35**, 63–66.
- 43 R. Huang, P. Fang and B. K. Kay, Improvements to the Kunkel mutagenesis protocol for constructing primary and secondary phage-display libraries, *Methods*, 2012, **58**, 10–17.
- 44 C. Johnston, P. Polard and J.-P. Claverys, The DpnI/DpnII pneumococcal system, defense against foreign attack without compromising genetic exchange, *Mobile Genet. Elem.*, 2013, **3**, e25582.
- 45 A. Hermann and A. Jeltsch, Methylation Sensitivity of Restriction Enzymes Interacting with GATC Sites, *BioTechniques*, 2003, **34**, 924–930.
- 46 N. Petrenko, Y. Jin, L. Dong, K. H. Wong and K. Struhl, Requirements for RNA polymerase II preinitiation complex formation in vivo, *eLife*, 2019, **8**, e43654.
- 47 R. G. Roeder, The role of general initiation factors in transcription by RNA polymerase II, *Trends Biochem. Sci.*, 1996, **21**, 327–335.
- 48 R. C. Conaway and J. W. Conaway, General Initiation Factors for Rna Polymerase Ii, *Annu. Rev. Biochem.*, 1993, **62**, 161–190.
- 49 S. Buratowski, The basics of basal transcription by RNA polymerase II, *Cell*, 1994, **77**, 1–3.
- 50 G. Orphanides, T. Lagrange and D. Reinberg, The general transcription factors of RNA polymerase II, *Genes Dev.*, 1996, **10**, 2657–2683.
- 51 S. Buratowski, S. Hahn, L. Guarente and P. A. Sharp, Five intermediate complexes in transcription initiation by RNA polymerase II, *Cell*, 1989, **56**, 549–561.
- 52 J. Kaufmann and S. T. Smale, Direct recognition of initiator elements by a component of the transcription factor IID complex, *Genes Dev.*, 1994, **8**, 821–829.



- 53 L. V. Ngoc, Y.-L. Wang, G. A. Kassavetis and J. T. Kadonaga, The punctilious RNA polymerase II core promoter, *Genes Dev.*, 2017, **31**, 1289–1301.
- 54 C. M. Alberini, Transcription Factors in Long-Term Memory and Synaptic Plasticity, *Physiol. Rev.*, 2009, **89**, 121–145.
- 55 H. S. Rhee and B. F. Pugh, Genome-wide structure and organization of eukaryotic pre-initiation complexes, *Nature*, 2012, **483**, 295–301.
- 56 P. Bucher, Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences, *J. Mol. Biol.*, 1990, **212**, 563–578.
- 57 R. Javahery, A. Khachi, K. Lo, B. Zenzie-Gregory and S. T. Smale, DNA sequence requirements for transcriptional initiator activity in mammalian cells, *Mol. Cell. Biol.*, 1994, **14**, 116–127.
- 58 R. Evans, J. A. Fairley and S. G. E. Roberts, Activator-mediated disruption of sequence-specific DNA contacts by the general transcription factor TFIIB, *Genes Dev.*, 2001, **15**, 2945–2949.
- 59 W. Deng and S. G. E. Roberts, A core promoter element downstream of the TATA box that is recognized by TFIIB, *Genes Dev.*, 2005, **19**, 2418–2423.
- 60 J. M. Smith, D. Hartmann and M. J. Booth, Engineering cellular communication between light-activated synthetic cells and bacteria, *Nat. Chem. Biol.*, 2023, **19**, 1138–1146.
- 61 Y. Sato, *et al.*, *In vivo* gene delivery to tumor cells by transferrin-streptavidin-DNA conjugate, *Faseb. J.*, 2000, **14**, 2108–2018.
- 62 G. Guan, *et al.*, An Effective Cationic Human Serum Albumin-Based Gene-Delivery Carrier Containing the Nuclear Localization Signal, *Pharmaceutics*, 2019, **11**, 608–625.
- 63 R. Deev, *et al.*, Results of 5-year follow-up study in patients with peripheral artery disease treated with PL-VEGF165 for intermittent claudication, *Ther. Adv. Cardiovasc. Dis.*, 2018, **12**, 237–246.
- 64 S. Ylä-Herttuala, Gene Therapy of Critical Limb Ischemia Enters Clinical Use, *Mol. Ther.*, 2019, **27**, 2053.
- 65 H. Suda, A. Murakami, T. Kaga, H. Tomioka and R. Morishita, Bepermingene perplasmid for the treatment of critical limb ischemia, *Expert Rev. Cardiovasc. Ther.*, 2014, **12**, 1145–1156.
- 66 S. Mallapaty, India's DNA COVID vaccine is a world first – more are coming, *Nature*, 2021, **597**, 161–162.
- 67 C. Sheridan, First COVID-19 DNA vaccine approved, others in hot pursuit, *Nat. Biotechnol.*, 2021, **39**, 1479–1482.

