Controlling gene expression with light: a multidisciplinary endeavour

Denis Hartmann* Jefferson M. Smith* Giacomo Mazzotti, Razia Chowdhry and © Michael J. Booth†

Chemistry Research Laboratory, University of Oxford, Oxford OX1 3TA, U.K.

Correspondence: Michael J. Booth (michael.booth@chem.ox.ac.uk)

The expression of a gene to a protein is one of the most vital biological processes. The use of light to control biology offers unparalleled spatiotemporal resolution from an external, orthogonal signal. A variety of methods have been developed that use light to control the steps of transcription and translation of specific genes into proteins, for cell-free to in vivo biotechnology applications. These methods employ techniques ranging from the modification of small molecules, nucleic acids and proteins with photocages, to the engineering of proteins involved in gene expression using naturally light-sensitive proteins. Although the majority of currently available technologies employ ultraviolet light, there has been a recent increase in the use of functionalities that work at longer wavelengths of light, to minimise cellular damage and increase tissue penetration. Here, we discuss the different chemical and biological methods employed to control gene expression, while also highlighting the central themes and the most exciting applications within this diverse field.

Introduction

Light is an ideal stimulus to control biological systems. It acts orthogonally to cellular signals in a large number of organisms and allows for the tightest spatiotemporal control of any input [1]. A major area of research is the use of light to control, arguably, the most fundamental biological pathway, gene expression. This includes methods to control either the transcription (DNA to mRNA) or translation (mRNA to protein) steps (Figure 1). It is possible to either activate or silence each step, or perform both reversibly. Applications of these methods include the control of cell-free systems, gene circuits, and drug/gene delivery. Specific endogenous genes can also be controlled with light by modifying existing technologies.

Controlling gene expression with light can be approached through both chemical and biological means [2]. One approach is to use chemical photocages, light-sensitive molecules that are linked to a bioactive molecule, blocking its activity. Illumination with a specific wavelength of light causes the photocage to break the linking bond, which reforms the original molecule. Photocages have been attached to small molecule, nucleic acid, or protein regulators of expression or directly to the DNA or mRNA nucleic acid templates themselves [3] (Figure 1A–C). Alternatively, several naturally light-sensitive proteins, which function through multiple different pathways, have been engineered and fused to proteins involved in expression [4] (Figure 1D). Within this review we have covered multiple different approaches that have been used to control gene expression with wavelengths of light from ultraviolet (UV) to near-infrared (NIR) (Figure 2). It is vital to cover multidisciplinary topics to highlight the collective themes and methods used. Simple reporter proteins, such as fluorescent proteins, β-galactosidase, or luciferase, are used as a common way to measure the efficiency of each method. Current applications focus on control of cell-free to in vivo gene expression. Human cell lines are widely used to demonstrate application to endogenous pathways. Light-controlled systems are also widely applied in vivo to zebrafish, as they are transparent. A few systems have also shown application in mouse models. However, when working in vivo it is important proper controls are in place due to the presence of endogenous proteins that respond to light.

*These authors contributed equally to this work.
†Michael J. Booth received the Biochemical Society’s 2019 Early Career Research Award.

Received: 13 May 2020
Revised: 19 June 2020
Accepted: 22 June 2020

Version of Record published: 13 July 2020

© 2020 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).
Figure 1. Methods of controlling transcription and translation with light discussed in this review.

Uncaging the (A) DNA or (B) mRNA template of a gene of interest with light allows for activation of transcription/translation. Whereas regulation of gene expression with light can be achieved by using (C) caged small molecules, nucleic acids, and proteins or (D) engineering naturally light-sensitive proteins.

Figure 2. An illustration of the approximate wavelength of activation of a number of photocages (A) and engineered naturally light-sensitive proteins (B, with chromophores) discussed in this review. NIR = near-infrared, NP = nanoparticle.
Chemical photocages

The most commonly used photocages are 2-nitrobenzyl derivatives [5] (Figure 2A). They are small, easy to synthesise and are often available commercially; they do, however, cleave in the UV (regularly 365 nm), which can cause cellular damage. Another commonly encountered photocage are coumarin derivatives [5]. These are also irradiated with UV (365–400 nm), but can be chemically modified to allow longer wavelengths of illumination; derivatives absorbing in the green have been prepared [6]. Other photocages that have been used within this review include quinoline (UV irradiation) [7], RuBEP (blue irradiation) [8], and cyanine. Cyanines can be activated up to the NIR [9], but are not often used due to lengthy syntheses. Visible or NIR irradiation is preferable as it causes less cellular damage than UV.

Alternative photocages are azobenzenes and styryl derivatives. They are not photocleavable, but photoswitches, as they change configuration upon irradiation. Trans-Cis isomerisation of azobenzenes and styryl derivatives occurs following UV irradiation [10,11], however modified azobenzenes can be controlled with visible or NIR irradiation [12].

Light-controlled gene expression using small molecules

Small molecules are widely used to regulate gene expression and have been extensively photocaged to control their activity with light (Figure 2A). The largest group of photocaged molecules are estradiol and tamoxifen analogues, agonists of the estrogen receptor. These have been photocaged with 2-nitrobenzyl-, coumarin- and cyanine-derivatives to control reporter gene expression in mammalian cells [13–16]. Using 2-nitrobenzyl- or thiocoumarin-caged cyclophycin, reporter gene expression and phenotype, respectively, could also be controlled in zebrafish [17,18].

A widely used inducer system in mammalian cells is the Tetracycline (Tet) system, under the control of doxycycline. Nitrobenzyl-photocaged doxycyclines have controlled reporter gene expression in mammalian cells [19], tobacco leaves [19] and developing mouse embryos and xenopus tadpoles [20]. A number of other small molecule agonists of expression in mammalian cells have been photocaged, including 2-nitrobenzyl-caged nuclear hormones [21] and β-ecdysone [22], as well as a coumarin-caged CREB inhibitor [23].

Carbohydrates are a well-known class of bacterial gene-regulatory molecules. Bacterial gene expression is often placed under control of the lac operator, which is activated using isopropyl-β-D-thiogalactopyranoside (IPTG), a non-hydrolysable allolactose mimic. 2-Nitrobenzyl-photocaged IPTG enabled the efficient regulation of reporter genes and (+)-valencene biosynthesis [24,25]. Arabinose is also widely used to control gene expression via the arabinose operator. 2-Nitrobenzyl-photocaged arabinose was shown to efficiently control expression of Violacein biosynthesis [26].

Nucleotides are vital biological signals and building blocks for transcription. Photocaged analogues are therefore able to control these functions. In vitro transcription has been controlled by photocaging adenosine triphosphate (ATP) on the terminal phosphate with a coumarin [27]. This inhibited RNA synthesis until uncaging of ATP with light. In a similar manner, synthetic cells have been activated with 2-nitrobenzyl-caged ATP [28]. Alternatively, uracil and guanine triphosphates, photocaged on the nucleobase with a 2-nitrobenzyl, prevented Watson–Crick base pairing prior to uncaging [29].

Photocaging has also been demonstrated on small molecules that directly interact with DNA or RNA. Toyocamycin is an ATP analogue that controls translation by binding to a riboyme in mRNA. Light-activated gene expression in mammalian cells was achieved through 2-nitrobenzyl-photocaging [30]. Theophylline is another small molecule used to control translation by binding to an mRNA riboswitch. 2-Nitrobenzyl-photocaged theophylline was used to control expression in bacteria [31]. G-quadruplex (G4) structures are involved in regulation of gene expression. Photocaging the G4-stabilising ligand pyridostatin with a 2-nitrobenzyl allowed the light-activated downregulation of cancer-associated genes in mammalian cells [32]. Cell-free expression can also be controlled via light-activated compaction of DNA/RNA using AzoTAB, which contains an azobenzene photo-switch and has strong nucleic acid affinity in the trans-form and weak in the cis-form [33].

Light-controlled gene expression using nucleic acids

Modification of DNA or RNA with photocages (Figure 2A) can be used to control transcription and translation with light (Figure 3). Generally, light-activated regulation of transcription and translation with nucleic acids can be broadly classified into two categories: the light-induced activation of gene expression (by using the
DNA or mRNA template) and the light-induced gene knockdown (via caged antisense oligonucleotides (ASOs) or small interfering RNAs (siRNAs)).

**Gene activation**

One approach of inhibiting transcription from a DNA template is by attaching a photocage onto the Watson–Crick face of the nucleobase, impeding base pairing. In this way, incorporation of 2-nitrobenzyl-caged thymidine into a DNA promoter has allowed control of cell-free transcription [34] and reporter gene expression in zebrafish [35]. Alternatively, a covalent inter-strand crosslink was made in a DNA promoter, using a psoralen modified DNA base-pair, to inhibit unwinding of the double helix and therefore transcription [36]. Irradiation initiated cell-free expression through decrosslinking via an adjacent pyrene.

Attachment of photocleavable molecules to other positions on the DNA or RNA can also disrupt recognition of the substrate. One approach is to modify the nucleobase with photocleavable groups at non-Watson–Crick base pairing sites. Amino-modified thymine bases were incorporated into a DNA promoter sequence and reacted with a photocleavable 2-nitrobenzyl biotin to generate a PCR primer, to allow amplification of a gene of interest. Subsequent binding of monovalent streptavidin allowed light-activated cell-free expression. Using this method, expression of protein pores created patterned conductive pathways through synthetic tissues [37, 38]. In a similar manner, integration of 2-nitrobenzyl-caged 5-hydroxymethyl(hm)-cytosine and 5-hm-uracil into a DNA template controlled cell-free transcription with light [39]. A photocage was also installed on a nucleobase of a DNA template by using a methyltransferase with a 2-nitrobenzyl-modified S-adenosyl methionine substrate, which was used to control cell-free expression [40]. Enzymatic insertion of photocleavable biotin-coumarin-modified nucleobases into mRNA also enabled controlled expression in mammalian cells [41, 42].

Photocages can also be directly attached on the oligonucleotide backbone. A plasmid with a streptavidin/coumarin backbone-modified promoter was used to control expression in mammalian cells [43]. Coumarin- [44, 45] and 2-nitrobenzyl- [46, 47] photocages have also been reacted with the phosphate backbone of plasmids and mRNA to allow control of reporter gene expression or phenotype in zebrafish and mammalian cells, and cell-free expression in synthetic cells. Study of head-specific overexpression in zebrafish was achieved using caged mRNA in this way [44].

Photoswitches have also been incorporated into DNA and RNA to reversibly control transcription and translation. By incorporation into a DNA promoter, azobenzenes controlled cell-free transcription [48] and gene expression [49], using UV or blue light as a stimulus. Incorporation of photoswitchable nucleobases into a G4 sequence of a promoter allowed for controlled expression of a reporter protein in zebrafish [50]. Reversible strategies for translation were also developed by capping the 5’ end of the mRNA with azobenzene or styryl photoswitches [50, 51]. UV light inhibited the recruitment of a translation initiation factor, while blue light activated recruitment and expression to control differentiation in mammalian cells and zebrafish.

---

Figure 3. Attachment points of photocages onto DNA (X = H) or RNA (X = OH).

(A) Photocages have been attached to various positions on DNA/RNA strands to control function, including the phosphate backbone (1), in the backbone (2), on the nucleobase (3), and on the Watson–Crick face of the nucleobase (4). These photocages can be attached to double stranded (B) or single stranded (C) nucleic acids.
Gene silencing
ASOs are commonly used to knock down gene expression in cells via targeting complementary mRNA. Base pairing-inhibited 2-nitrobenzyl-caged thymidines have been incorporated into ASOs to photoactivate gene knockdown of developmental genes in zebrafish [52] and cancer-associated genes in mammalian cells [53]. By incorporating a photoreactive nucleobase into an ASO, covalent interstrand crosslinking to its complementary mRNA was achieved in mammalian cells [54].

ASOs can also be light-activated by removal of a complementary inhibitor strand. Inclusion of a 2-nitrobenzyl in the backbone of the complementary strand enabled endogenous gene knockdown and controlled development in zebrafish [55]. The complementary strand has also been attached to the ASO via a 2-nitrobenzyl-containing hairpin, which allowed for photocontrolled knockdown of a cancer-associated gene in mammalian cells [56] and developmental genes in zebrafish [57,58].

ASOs have also been circularised with a photocleavable linker, disrupting mRNA binding until photocleavage. Circularisation, achieved with 2-nitrobenzyl- [59], coumarin- [60], quinoline- [7] and Ru-BEP linkers [8] was applied to control expression with UV or blue light in zebrafish and mammalian cells.

Synthetic siRNAs are double stranded RNA molecules that are widely used to knock down expression of target genes via RNA interference (RNAi). Modification of the phosphate backbone of siRNAs with 2-nitrobenzyl allowed for light-activated knockdown in zebrafish [61]. Similarly, by attaching 2-nitrobenzyl-moieties to the 5’ [62,63] or 5’ and 3’ phosphate termini [64] cell patterning was achieved [65], and the RNAi off-state was improved by adding larger moieties [66]. Attaching an anthracene functionality along with a porphyrin photosensitiser onto the termini of siRNA controlled its function with green or red light [67]. Hybridisation of the siRNA strands was also controlled through incorporation of an azobenzene into the backbone and controlled with red light [68]. 2-Nitrobenzyl-photocaging of the Watson–Crick face of a single nucleobase in siRNA inhibited siRNA:mRNA duplex formation, and thus knocking down a reporter protein [69] and an endogenous gene [70]. An inhibitor of microRNA (miRNA), the natural substrate of RNAi knockdown, has been developed from an RNA hairpin connected through a 2-nitrobenzyl in the backbone [71]. Cleavage with UV allowed binding of the inhibitor to the miRNA, causing altered development in nematodes.

Light-responsive nanoparticles have been extensively used to allow spatiotemporal control of siRNA function. NIR light can release siRNA bound to gold nanoparticles (AuNPs). siRNA attached to AuNPs via terminal thiol-modified siRNA [72–74], DNA bridges [75], or electrostatically [76] enabled NIR-activated knockdown in mammalian cells and therapeutic effects on tumour models in vivo. Improved delivery was also accomplished by attaching cell-penetrating peptides to the constructs [73]. NIR-activated knockdown within mammalian cells was also achieved by attachment of siRNA, via a 2-nitrobenzyl, to upconverting nanoparticles (UCNPs), which convert NIR to UV irradiation [77]. Nanoparticles of siRNA have also been formed with positively charged block-copolymers, which aided cellular delivery, containing 2-nitrobenzyls [78].

A number of alternative methods have been used to photocontrol the silencing of gene expression. ASOs have been attached to AuNPs and activated with NIR in mice [79]. Azobenzens have also been used to reversibly photocontrol binding of an inhibitory hairpin DNA attached to a DNAzyme to activate and deactivate cell-free expression of a reporter gene [80]. Gene knockdown has also been controlled by using a hairpin DNA decoy, modified with photocaged thymidines [81]. Implementing 2-nitrobenzyl-photocaged cytidines in a triplex-forming oligonucleotide also allowed for the light-controlled activation and silencing of gene expression in mammalian cells [82]. Modification of siRNA has also been used to control cell delivery with visible light [83,84].

Light-controlled gene expression using proteins
Proteins involved in gene expression have been extensively modified, both chemically and with naturally light-sensitive proteins, to control transcription and translation with light. Additionally, by combining these modified proteins with genome targeting technologies, it is possible to light-activate the expression of specific genes on a genome.

Chemical modification of proteins
Chemically modified proteins can be produced by engineering orthogonal ribosomes and tRNA/tRNA synthetase pairs to accept unnatural amino acids at a TGA codon. This allows precise placement of a photocage within a protein of interest. Using this method, a 2-nitrobenzyl was site-specifically incorporated into an RNA
polymerase to demonstrate light-activated gene expression [85] and light-activated RNAi [85] in mammalian cells. Photocages have also been installed into gene editing tools, including Cre recombinases, which enable site specific recombination between DNA sequences called LoxP sites, and Zinc Finger (ZF) nucleases, which are selective towards target DNA sequences. Essential residues in catalytic sites of Cre and ZF nucleases have been modified with 2-nitrobenzyl [86,87] and coumarin [88] to activate expression upon irradiation. Alternatively, incorporation of an azobenzene into the backbone of an epigenetic regulating peptide has allowed photocontrol over its activity in mammalian cells [89].

Naturally light-sensitive proteins

In addition to modifying proteins with photocages, naturally light-responsive proteins from plants, cyanobacteria, and algae, have also been incorporated into proteins involved in gene expression (Figure 2B). The mechanisms of these engineered proteins fall into two categories: direct modulation of gene expression using light-activated transcription factors or enzymes, and indirect modulation through light-activated signalling cascades that modulate downstream effectors (Figure 4). While many different photoreceptors have been used to control gene expression, this review focuses primarily on three of the most common classes: phytochromes, Light-Oxygen-Voltage (LOV)-domain proteins, and cryptochromes.

Phytochromes (red and NIR light-responsive)

Phytochromes are photoreceptors that utilise tetrapyrrole chromophores such as biliverdin IXα (BV) or phyco-cyanobilin (PCB) (Figure 2B) to absorb red and NIR light and induce reversible conformational changes in the protein structure. The first photoreceptor used to control gene expression was Phytochrome protein B (PhyB), and its interaction partner Phytochrome-Interacting Factor 3 (PIF3), which dimerise under red light and dissociate under far-red light. By fusing DNA-Binding Domains (DBDs) and transActivation Domains (ADs) to the distinct PhyB and PIF modules, red light inducible two-hybrid systems have been used to activate gene expression in yeast [90] and mammalian cells [91] (Figure 4A-i).

![Figure 4](https://doi.org/10.1042/BST20200014)

**Figure 4.** Controlling gene expression by using naturally light-sensitive proteins.

(A) Protein modules that dimerise in response to light, represented here as Light-Activated Domains (LADs), have been fused to effector domains to create transcription factors that activate gene expression in the presence of light, but are inactive in the absence of light. Gene expression is activated upon co-localisation of DNA-Binding Domains (DBDs) with transActivation Domains (ADs) via a light-activated two-hybrid system (i), or via the co-localisation of inactive C- and N-terminal domains of a split protein and reconstitution of the active protein (ii). (B) Light-responsive two-component systems (TCSs) are initiated when light is absorbed by the sensory domain of a histidine kinase, which stimulates/represses autophosphorylation of the kinases domains. Phosphorylated kinase domains transfer phosphate groups to downstream Transcription Factors (TF), which can then bind to consensus promoter sequences and activate gene expression. (C) Uncaging of the Jα helix in AsLOV2, when exposed to blue light, is used to reveal a shielded Nuclear Localisation Signal (NLS). The exposed NLS is recognised by the importin complex and AsLOV2 is transported into the nucleus. Gene expression can be regulated by fusing DBDs and ADs to AsLOV2, and controlling their nuclear localisation by exposing or shielding the NLS.
In bacteria, Two-Component signalling Systems (TCSs) comprising of a natural or engineered light-responsive kinase and a downstream response regulator are more prevalent. Cph8 is a light-responsive kinase that was engineered by replacing the osmosensory domain of a membrane-bound histidine kinase, with the Cph1 phytochrome. In the absence of light, Cph8 phosphorylates the transcription factor OmpR and promotes gene expression, while in the presence of red light the kinase activity and gene expression are inactivated [92] (Figure 4B). The engineered Cph8/OmpR TCS is analogous to the natural cyanobacteriochrome TCS, CcaS/CcaR, which regulates gene expression according to green and red light [93] (Figure 4B). In their most basic forms, these systems were used in bacterial edge detection algorithms [94] and for dual-colour control over gene expression [95]. More efficient versions were later developed by genetic refactoring and mutagenesis [96,97] and have been used to tightly control metabolic flux [98], cell division [99] and feedback loops in bacteria [100].

Most of the red light-responsive gene expression systems require a PCB chromophore to function. Although PCB can be added to the growth media and taken up by cells [90,91], gene cassettes encoding enzymes for PCB biosynthesis are more commonly used to enable cells to synthesise PCB from intracellular heme [92,101,102]. NIR-responsive phytochromes on the other hand utilise BV chromophores, which are produced endogenously by mammalian cells. These NIR-responsive gene expression systems are derived from BphP1 and PpsR2 proteins. PpsR2 is sequestered by BphP1 in NIR light and then released in the presence of red light or absence of light. NIR-responsive transcription factors were created by fusing BphP1 and PpsR2 to DBDs and ADs, and have been shown to activate reporter gene expression in bacteria [103], mammalian cells, and mice [104,105] (Figure 4A-i).

LOV domain proteins and Cryptochromes (blue-light responsive)

LOV domain proteins and cryptochromes are two distinct protein families that differ in their protein architecture, yet both use blue light absorbing flavin chromophores (Figure 2B) to induce conformational changes in the protein. Whereas many different LOV domain systems have been developed for blue light-activated gene expression, cryptochrome-based systems are based primarily on the interaction of Cryptochrome protein 2 (CRY2) with Cryptochrome-Interacting Basic-helix-loop-helix protein (CIB1).

The first blue light-responsive transcription system was a TCS implementing an engineered light-sensitive kinase, YF1, created by replacing an oxygen-sensing domain of a natural kinase with the YvtA LOV domain. In the absence of light, YF1 phosphorylated and activated a transcription factor, whereas in the presence of blue-light, kinase activity and gene expression were inactivated [106] (Figure 4B). Based on this system, a bacterial repression switch, pDusk, and activation switch, pDawn [107], were developed. Blue light-regulated gene expression in eukaryotic cells, however, is typically controlled with photoactivatable transcription factors. Flavin-binding, Kelch domain, F-box protein (FKF1) and its interaction partner GIGANTEA (GI) were the first photo-receptor pair to be fused to ADs and DBDs and demonstrated blue light-activated transcription in mammalian cells [108] (Figure 4A-i). Similar light-activated effector proteins based on CRY2/CIB1, and smaller LOV domain proteins also function in this way. Amongst these is Vivid (VVD), a photoreceptor that rapidly and reversibly forms homodimers. CRY2/CIB1 and VVD have been fused to a variety of DBDs and effector domains to control both transcription [109,110] and translation [111,112] in mammalian cells, bacteria [113], and yeast [110,114,115]. They have been applied in various fields including the study of oscillating gene expression [116] and in cell-based immunotherapies [117]. Blue light-induced dimerisation of CRY2/CIB1, VVD, and ‘Magnets’, improved VVD mutants [118], have also been used to control gene expression by reconstituting split enzymes, such as Cre recombinases [110,119,120] and RNA polymerases [121,122] (Figure 4A-ii).

In most cases, light-inducible transcription factors depend on the heterodimerisation of two different proteins to modulate gene expression. However, single component blue light-activated transcription factors also exist, and are typically derived from AsLOV2 or EL222. Upon exposure to blue light, both AsLOV2 and EL222 undergo conformational changes involving the release of an alpha helix from the LOV domain. By inserting nuclear localisation signals (NLS) or nuclear export signals (NES) within the 4α helix of AsLOV2, its cellular location can be switched with blue light (Figure 4C). In this way, genes have been expressed via the import of AsLOV2-based transactivators into the nucleus [123,124], or export of AsLOV2-based repressors into the cytoplasm [125,126]. Alternatively, uncaging of the 4α helix in EL222 is accompanied by the release of a DBD and exposure of a dimerization interface. EL222 fused to an AD has been shown to rapidly induce gene expression in mammalian cells [127] and zebrafish embryos following irradiation [128], and has also been used in yeast to
improve their chemical production capabilities [129]. Single component gene expression systems in bacteria [130] and cell-free expression systems [131] have used the transcription factor activity of wild type EL222.

**Genome targeting for light-activated expression**

Technologies that target specific sites on the genome allow for precise control of endogenous genes. Genome targeting has been achieved using several types of DNA binding proteins, that can be tailored to recognise specific DNA sequences [132]. These include the Clustered, Regularly Interspaced, Short Palindromic Repeat (CRISPR)-associated nuclease Cas9 system, ZF proteins, Transcription Activator-Like Effectors (TALEs), and recombinases. Photocages and naturally light-sensitive proteins have been incorporated into these technologies for spatiotemporal control of gene knockouts and transcription.

**CRISPR-Cas9**

CRISPR-Cas systems are found in bacteria and archaea as a means of adaptive immunological protection against phages. The bacterial Cas9 nuclease is targeted to the genome using a non-coding guide RNA (gRNA) and cleaves the double stranded DNA causing either a gene knockout or mutagenesis [132]. By mutating Cas9, a catalytically inactive protein (dCas9) can serve as an RNA-guided DNA-binding protein. Both the gRNA and Cas9 have been modified to create light-activated CRISPR-Cas9 systems. A protector DNA, containing 2-nitrobenzyls in the backbone has been bound to gRNA to control gene knockout in mammalian cells [133]. More efficient control was achieved by attaching 2-nitrobenzyls to the Watson–Crick face of nucleobases in the gRNA, which was used to control gene editing in zebrafish embryos [134,135]. To control the Cas9 nuclease, a 2-nitrobenzyl-modified amino acid was installed using an orthogonal tRNA/tRNA synthetase pair [136]. Cas9 has also been covalently linked to UCNPs, via a 2-nitrobenzyl photocage, allowing NIR-activated gene editing and reduction of tumour size in mice [137].

Naturally light-sensitive proteins have been used with both the Cas9 nuclease and dCas9 to achieve light-activated gene knockout and transcriptional control, respectively (Figure 4A). These rely on the dimerization of split Cas9 domains and/or dCas9 with Adams via photoreceptors, particularly CRY2/CIB1 [138–140] and magnet proteins [140,141]. Other light-activated CRISPR-Cas9 systems depend on light-activated phosphorylation [142] or cyclic diguanylate monophosphate (c-di-GMP) signalling cascades [143], as well as a dimeric green fluorescent protein, pdDronpa [144].

**Zinc finger and transcription activator-like effector proteins**

ZF proteins recognise a specific 3-base pair DNA sequence and individual TALE proteins each recognise a single base pair. Hence, effector proteins can be targeted to specific locations on a genome by fusing them to contiguous ZF and TALE domains [132]. By fusing ZFs to GI and an AD to FKF1, gene expression in mammalian cells has been controlled [145] (Figure 4A-i). TALEs and ADs using the CRY2/CIB1 pair have been used to regulate gene expression in mouse brain cortex [146] and recruitment of epigenetic modifiers in rat neuronal stem cells [147] (Figure 4A-i).

**Recombinases**

In recombinase-mediated gene editing, DNA flanked by LoxP recognition sites is commonly excised from the genome of transgenic organisms to study the response of gene knockouts [148]. By incorporating these artificial recognition sites in the genome, light-activated recombinases have been used to control the expression of targeted genes. Light-activated Cre systems have been developed using 2-nitrobenzyl-photocaged tamoxifen [149], which controlled light-dependent recombination and gene expression in mice [150]. CRY2/CIB1, magnet, or VVD pairs have also been fused to split Cre and Flp recombinase domains to control gene expression in mouse brains [151–154] and zebrafish embryos [153] (Figure 4A-ii). This has been applied to reconstruct the morphology of single neurons across an entire mouse brain [153].

**Outlook**

Controlling gene expression with light has led to a wide range of applications in cell-free to *in vivo* systems. Chemical modification of nucleic acids with photocages has been used to create conductive pathways in synthetic tissues [37,38] and control the development of zebrafish [7,55,60,134,135] and cancer therapeutics [75,76,79,137] with light. While naturally light-sensitive proteins have been used to control cell division [99],
bioreactors [129], and genome activation in zebrafish and mice [146,151,153]. Identification of new applications is vital to realising the full potential of light-controlled gene expression.

As we have focused on discussing methods and applications within this review, there has been a limited discussion of the efficiencies of each system. To allow general use of these technologies an efficient 'ON' and 'OFF' state is required, where minimal activity is observed without light, and maximal with light. It is worth noting the efficiencies vary wildly for the methods discussed. More efficient and easily accessible systems are required to allow more general use.

Photocages and naturally light-sensitive proteins each have their own advantages and limitations. Plasmid DNA encoding light-sensitive proteins and some simple photocages attached to DNA are commercially available. Most photocaged small molecules and the more advanced photocages must be synthesised in house. However, this is also an advantage as more diverse systems have been generated through the chemical synthesis of photocages with small molecules, nucleic acids, and proteins. Another advantage of photocages is their immediate generation of an active species, compared with the requirement that cells first generate the light-sensitive proteins from plasmid DNA, prior to their application. Most chemical photocages absorb in the UV, which can damage cells. Chemical photocages that absorb at longer wavelengths of light, which minimises cellular damage and increases tissue penetration [155], are becoming more popular. However, more easily accessible NIR photocages would give a major boost to this area. In contrast with this, naturally light-sensitive proteins regularly absorb UV, visible or NIR light. Hence, more orthogonal systems also exist for light-sensitive proteins, where expression of different genes can be put under the control of different wavelengths of light, even inside the same cell. Additionally, many reversible naturally light-sensitive proteins exist, compared with only a few chemical photoswitches. Regarding cell delivery, small molecules are highly advantageous as they are cell permeable. Plasmid DNA encoding naturally light-sensitive proteins and photocaged nucleic acids must be transfected. Naturally light-sensitive proteins also tend to require multiple or complex plasmids to be delivered. However, in bacteria and yeast these systems can be controlled over multiple generations as the plasmid system encoding the light-sensitive proteins is replicated. An important goal for this field is to combine the light-activation systems with cell delivery systems, where only limited examples exist [53,72,73,78,146,151,153], to produce far-reaching technologies.

**Perspective**

- **Importance:** Using light to control gene expression opens the door to myriad biotechnology applications in therapeutics, biomanufacturing, and emerging fields such as cell-free systems.
- **Summary:** Control of exogenous and endogenous gene expression has been achieved in many different ways. Chemical tools rely on attaching photocages to small molecules, nucleic acids, and proteins, while biological tools use engineered naturally light-sensitive proteins.
- **Direction:** Beyond identifying new applications, the three main areas of research that still require attention are: increased usage of longer wavelength photocages, more efficient and general systems, and combining light-activation with cell delivery.

**Competing Interests**
The authors declare that there are no competing interests associated with the manuscript.

**Funding**

D.H. is grateful to the EPSRC Centre for Doctoral Training in Synthesis for Biology and Medicine (EP/L015838/1) for a studentship, generously supported by AstraZeneca, Diamond Light Source, Defence Science and Technology Laboratory, Evotec, GlaxoSmithKline, Janssen, Novartis, Pfizer, Syngenta, Takeda, UCB and Vertex. J.M.S. is supported through the Synthetic Biology Centre for Doctoral Training — EPSRC funding (EP/L016494/1). G.M. is funded through an EPSRC studentship. R.C. is supported by funding from the Biotechnology and Biological Sciences Research Council (BBSRC) (BB/M011224/1). M.J.B. is supported by a Royal Society University Research Fellowship.
Open Access

Open access for this article was enabled by the participation of University of Oxford in an all-inclusive Read & Publish pilot with Portland Press and the Biochemical Society under a transformative agreement with JISC.

Author Contributions

All authors wrote and edited the manuscript. *D.H. and J.M.S. contributed equally.

Acknowledgements

We apologise to authors whose work we were unable to cite due to space limitations.

Abbreviations

AD, transactivation domain; ASO, antisense oligonucleotide; ATP, adenosine triphosphate; AuNP, gold nanoparticle; BV, biliverdin IXα; c-di-GMP, cyclic di guanylate monophosphate; CIB1, cryptochrome-interacting basic helix-loop-helix protein; CREB, CAMP response element-binding protein; CRISPR, clustered, regularly interspaced, short palindromic repeat; CRY2, cryptochrome protein 2; DBD, DNA binding domain; dCas9, deactivated Cas9; DNA, deoxyribonucleic acid; FKF1, Flavin-binding, Kelch domain, F-box protein; G4, G-quadruplex; GI, GIGANTEA; gRNA, guide RNA; hm, hydroxymethyl; IPTG, isopropyl-β-D-thiogalactopyranoside; LAD, light-activated domain; LOV, light-oxygen-voltage; miRNA, microRNA; mRNA, messenger RNA; NIR, near-infrared; NLS, nuclear localization signal; NP, nanoparticle; PCB, phycocyanobilin; PCR, polymerase chain reaction; PhyB, phytochrome protein B; PIF3, phytochrome interacting factor 3; RNA, ribonucleic acid; rNAi, RNA interference; siRNA, small interfering RNA; TALE, transcription activator-like effectors; TCS, two component system; Tet, tetracycline; TF, transcription factor; tRNA, transfer RNA; UCNP, upconverting nanoparticle; UV, ultraviolet; VVD, Vivid; ZF, zinc finger.

References

1 Devor, A., Sakadžić, S., Srivasan, V.J., Yaseen, M.A., Nizar, K., Saisan, P.A. et al. (2012) Frontiers in optical imaging of cerebral blood flow and metabolism. J. Cereb. Blood Flow Metab. 32, 1259–1276 https://doi.org/10.1038/jcbfm.2011.195
2 Hughes, R.M. (2018) A compendium of chemical and genetic approaches to light-regulated gene transcription. Crit. Rev. Biochem. Mol. Biol. 53, 453–474 https://doi.org/10.1080/10409238.2018.1487362
3 Deiters, A. (2009) Light activation as a method of regulating and studying gene expression. Curr. Opin. Chem. Biol. 13, 678–686 https://doi.org/10.1016/j.cpb.2009.09.026
4 de Mena, L., Rak, P. and Rincon-Limás, D.E. (2018) Bringing light to transcription: the optogenetics repertoire. Front. Genet. 9, 518 https://doi.org/10.3389/fgene.2018.00518
5 Klián, P., Soňomok, T., Bochet, C.G., Blanc, A., Šiven, R., Rubina, M. et al. (2013) Photoremovable protecting groups in chemistry and biology: reaction mechanisms and efficacy. Chem. Rev. 113, 119–191 https://doi.org/10.1021/cr300177k
6 Bojžar, M., Kormos, A., Kís-Petliát, K., Kellermayer, M. and Kele, P. (2019) Green-light activatable, water-soluble red-shifted coumarin photocages. Org. Lett. 21, 9410–9414 https://doi.org/10.1021/acs.orglett.9b03624
7 O’Connor, M.J., Beeley, L.L., Deedato, D., Ball, R.E., Page, A.T., Vanleuven, A.J. et al. (2019) Bypassing glutamic acid decarboxylase 1 (Gad1) induced craniofacial defects with a photoactivatable translation blocker morpholino. ACS Chem. Neurosci. 10, 266–278 https://doi.org/10.1021/acschemneuro.8b00231
8 Griepeenburg, J.C., Rapp, T.L., Carroll, P.J., Ebenevir, J. and Dmochowski, I.J. (2015) Ruthenium-caged antisense morpholinos for regulating gene expression in zebrafish embryos. Chem. Sci. 6, 2342–2346 https://doi.org/10.1039/C4SC0399000
9 Gorka, A.P., Nani, R.R., Zhu, J., Macken, S. and Schnermann, M.J. (2014) A near-IR uncaging strategy based on cyanine photochemistry. J. Am. Chem. Soc. 136, 14153–14159 https://doi.org/10.1021/jacs.10505203
10 Beharry, A.A. and Woolley, G.A. (2011) Azobenzene photoswitches for biomolecules. Chem. Soc. Rev. 40, 4422–4437 https://doi.org/10.1039/ c1cs15023e
11 Gömer, H. and Kuhn, H.J. (2007) Ce-trans photosoisomerization of stilbenes and stilbene-like molecules. In Advances in Photochemistry (Neckers, D.C., Volman, D.H. and van Baken, G., eds), pp. 1–117, John Wiley & Sons Ltd
12 Dong, M., Babalhavaeji, A., Collins C, V., Jarrah, K., Sadovski, O., Dai, Q. et al. (2017) Near-infrared photoswitching of azobenzenes under physiological conditions. J. Am. Chem. Soc. 139, 13483-13486 https://doi.org/10.1021/jacs.7b06671
13 Cruz, F.G., Koh, J.T. and Link, K.H. (2000) Light-activated gene expression. J. Am. Chem. Soc. 122, 8777–8778 https://doi.org/10.1021/ja001804h
14 Faal, T., Wong, P.T., Tang, S., Coulter, A., Chen, Y., Tu, C.H. et al. (2015) 4-Hydroxytamoxifen probes for light-dependent spatiotemporal control of Cre-ER mediated reporter gene expression. Mol. Biosyst. 11, 783–790 https://doi.org/10.1039/C4MB00581C
15 Wang, P.T., Roberts, E.W., Tang, S., Mukherjee, J., Cannon, J., Ni, A.J. et al. (2017) Control of an unusual photo-Claissen rearrangement in coumarin caged tamoxifen through an extended spacer. ACS Chem. Biol. 12, 1001–1010 https://doi.org/10.1021/acschembio.6b00999
16 Gorka, A.P., Yamamoto, T., Zhu, J. and Schnermann, M.J. (2018) Cyanine photocages enable spatial control of inducible Cre-mediated recombination. ChemBioChem 19, 1239–1243 https://doi.org/10.1002/cbic.201800061
17 Zhang, W., Hamouri, F., Feng, Z., Aujard, I., Ducos, B., Ye, S. et al. (2018) Control of protein activity and gene expression by cyclofen-OH uncaging. ChemBioChem 19, 1232–1238 https://doi.org/10.1002/cbic.201700630

© 2020 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).
Kamiya, Y., Arinuma, Y., Ooi, H., Kato, K., Liang, X.-G. and Asanuma, H. (2018) Development of visible-light-responsive RNA scissors based on a 10-23 DNAzyme. ChemBioChem 19, 1305–1311. https://doi.org/10.1002/cbic.201800020

Govan, J.M., Lively, M.O. and Deiters, A. (2011) Photochemical control of DNA decay function enables precise regulation of nuclear factor κB activity. J. Am. Chem. Soc. 133, 13176–13182. https://doi.org/10.1021/ja204988v

Govan, J.M., Uprety, R., Hemphill, J., Lively, M.D. and Deiters, A. (2012) Regulation of transcription through light-activation and light-deactivation of triplex-forming oligonucleotides in mammalian cells. ACS Chem. Biol. 7, 1247–1256. https://doi.org/10.1021/cb300161r

Endoh, T., Sisido, M. and Ohtsuki, T. (2009) Spatial regulation of specific gene expression through photoactivation of RNAi. J. Control. Release. 137, 241–245. https://doi.org/10.1016/j.jconrel.2009.04.015

Oliveira, S., Fretz, M.M., Hegseth, A., Storm, G. and Schifflers, R.M. (2007) Photochemical internalization enhances silencing of epidermal growth factor receptor through improved endosomal escape of siRNA. Biochem. Biophys. Acta Biomembr. 1768, 1211–1217. https://doi.org/10.1016/j.bbamem.2007.01.013

Hemphill, J., Chou, C., Chin, J.W. and Deiters, A. (2013) Genetically encoded light-activated transcription for spatiotemporal control of gene expression in mammalian cells. J. Am. Chem. Soc. 135, 13433–13439. https://doi.org/10.1021/ja4051026

Chou, C. and Deiters, A. (2011) Light-activated gene editing with a photocaged zinc-finger nuclease. Angew. Chem. Int. Ed. Engl. 50, 6839–6842. https://doi.org/10.1002/anie.201011157

Luo, J., Arbelý, E., Zhang, J., Chou, C., Uprety, R., Chin, J.W. et al. (2016) Genetically encoded optical activation of DNA recombination in human cells. Chem. Commun. 52, 8529–8532. https://doi.org/10.1039/C6CC03934K

Brown, W. and Deiters, A. (2019) Light-activation of Cre recombinase in zebrafish embryos through genetic code expansion. Methods Enzymol. 624, 265–281. https://doi.org/10.1016/bs.mie.2019.04.004

Albert, L., Xu, J., Wan, R., Srinivasan, V., Dou, Y. and Vázquez, O. (2017) Controlled inhibition of methyltransferases using photoswitchable peptidomimetics: towards an epigenetic regulation of leukemia. Chem. Sci. 8, 4612–4618. https://doi.org/10.1039/C7SC00137A

Shimizu-Sato, S., Huq, E., Tepperman, J.M. and Quail, P.H. (2002) A light-switchable gene promoter system. Nat. Biotechnol. 20, 1041–1044. https://doi.org/10.1038/nbt734

Müller, K., Engesser, R., Metzger, S., Schulz, S., Kämpf, M.M., Busacker, M. et al. (2013) A red/or red-light-responsive bi-stable toggle switch to control gene expression in mammalian cells. Nucleic Acids Res. 41, e77. https://doi.org/10.1093/nar/gkt002

Levskaya, A., Chevalier, A.A., Tabor, J.J., Simpson, Z.B., Lavery, L.A., Levy, M. et al. (2008) Engineering Escherichia coli to see light. Nature 438, 441–442. https://doi.org/10.1038/nature04405

Hirose, Y., Shinoda, T., Narikawa, R., Katayama, M. and Ikeuchi, M. (2008) Cyanobacteriochrome CcaS is the green light receptor that induces the expression of phyocobilisome linker protein. Proc. Natl Acad. Sci. U.S.A. 105, 9528–9533. https://doi.org/10.1073/pnas.0801826105

Tabor, J.J., Salis, H.M., Simpson, Z.B., Chevalier, A.A., Levskaya, A., Marcotte, E.M. et al. (2009) A synthetic genetic edge detection program. Cell 137, 1272–1281. https://doi.org/10.1016/j.cell.2009.04.048

Tabor, J.J., Levskaya, A. and Voigt, C.A. (2011) Multichromatic control of gene expression in Escherichia coli. J. Mol. Biol. 405, 315–324. https://doi.org/10.1016/j.jmb.2010.10.038

Ong, N.T. and Tabor, J.J. (2018) A miniaturized Escherichia coli green light sensor with high dynamic range. ChemBioChem 19, 1255–1258. https://doi.org/10.1002/cbic.201800007

Schmidl, S.R., Sheth, R.U., Wu, A. and Tabor, J.J. (2014) Refactoring and optimization of light-switchable Escherichia coli two-component systems. ACS Synth. Biol. 3, 820–831. https://doi.org/10.1021/acssynbio.4b00273n

Fernandez-Rodriguez, J., Moser, F., Song, M. and Voigt, C.A. (2017) Engineering RGB color vision into Escherichia coli. Nat. Chem. Biol. 13, 706–708. https://doi.org/10.1038/nchembio.2390

Mushnikov N. V., Formicheva, A., Gomelsky, M. and Bowman, G.R. (2019) Inducible asymmetric cell division and cell differentiation in a bacterium. Nat. Chem. Biol. 15, 925–931. https://doi.org/10.1038/s41588-019-0540-4

Milicevic-Argelits, A., Rullan, M., Aoki, S.K., Buchmann, P. and Krzysztofowicz, M. (2016) Automated optogenetic feedback control for precise and robust regulation of gene expression and cell growth. Nat. Commun. 7, 12546. https://doi.org/10.1038/ncomms12546

Uda, Y., Goto, Y., Oda, S., Kohchi, T., Matsuda, M. and Aoki, K. (2017) Efficient synthesis of phycocyanobilin in mammalian cells for optogenetic control of cell signaling. Proc. Natl Acad. Sci. U.S.A. 114, 11962–11967. https://doi.org/10.1073/pnas.1707190114

Kyriakakis, P., Catarina, M., Hoffner, N., Thavarajah, W., Hu, V.J., Chao, S.S. et al. (2018) Biosynthesis of orthogonal molecules using ferredoxin and ferredoxin-NAD+ reductase systems enables genetically encoded phb optogenetics. ACS Synth. Biol. 7, 606–717. https://doi.org/10.1021/acssynbio.7b00413

Ong, N.T., Olson, E.J. and Tabor, J.J. (2018) Engineering an E. coli near-Infrared light sensor. ACS Synth. Biol. 7, 240–248. https://doi.org/10.1021/acssynbio.7b00289

Kabeniuk, A., Shenmet, A.A. and Verkhusha V.V. (2016) A bacterial phytochrome-based optogenetic system controllable near-infrared light. Nat. Methods 13, 591–597. https://doi.org/10.1038/nmeth.3868

Redchuk, T.A., Omelina, E.S., Chernov, K.G. and Verkhusha V.V. (2017) Near-infrared optogenetic pair for protein regulation and spectral multiplexing. Nat. Chem. Biol. 13, 633–639. https://doi.org/10.1038/nchembio.2343

Möglicher, A., Ayers, R.A. and Moffatt, K. (2009) Design and signaling mechanism of light-Regulated histidine kinases. J. Mol. Biol. 385, 1433–1444. https://doi.org/10.1016/j.jmb.2008.12.017

Ohlendorf, R., Wisniewski, R., Eldar, A., Moffat, K. and Möglich, A. (2012) From dusk till dawn: one-plasmid systems for light-regulated gene expression. J. Mol. Biol. 416, 534–542. https://doi.org/10.1016/j.jmb.2012.01.061

Yazawa, M., Sadaghiani, A.M., Huseh, B. and Dolsenetsch, R.E. (2009) Induction of protein–protein interactions in live cells using light. Nat. Biotechnol. 27, 941–945. https://doi.org/10.1038/nbt.1569

Wang, X., Chen, X. and Yang, Y. (2012) Spatiotemporal control of gene expression by a light-switchable transgene network. Nat. Methods 9, 266–269. https://doi.org/10.1038/nmeth.1892

Taslim, A., Zoltowski, B., Miranda, J.G., Pathak, G.P., Hughes, R.M. and Tucker, C.L. (2016) Optimized second-generation CRY2-CIB dimers and photoactivatable Cre recombinase. Nat. Chem. Biol. 12, 425–430. https://doi.org/10.1038/nchembio.2063
Shao, J., Wang, M., Yu, G., Zhu, S., Yu, Y., Heng, B.C. et al. (2018) Synthetic far-red light-mediated CRISPR-dCas9 device for inducing functional neuronal differentiation. Proc. Natl Acad. Sci. U.S.A. 115, E6722–E6730 https://doi.org/10.1073/pnas.1802448115

Zhou, X.X., Zou, X., Chung, H.K., Gao, Y., Liu, Y., Qi, L.S. et al. (2018) A single-chain photoswitchable CRISPR-Cas9 architecture for light-inducible gene editing and transcription. ACS Chem. Biol. 13, 443–448 https://doi.org/10.1021/acschembio.7b00603

Polstein, L.R. and Gersbach, C.A. (2012) Light-inducible spatiotemporal control of gene activation by customizable zinc finger transcription factors. J. Am. Chem. Soc. 134, 16480–16483 https://doi.org/10.1021/ja3065667

Konermann, S., Brigham, M.D., Trevino, A.E., Hsu, P.D., Heidenreich, M., Cong, L. et al. (2013) Optical control of mammalian endogenous transcription and epigenetic states. Nature 500, 472–476 https://doi.org/10.1038/nature12466

Lo, C.L., Choudhury, S.R., Irudayaraj, J. and Zhou, F.C. (2017) Epigenetic editing of Ascl1 gene in neural stem cells by optogenetics. Sci. Rep. 7, 42027 https://doi.org/10.1038/srep42027

Nagy, A. (2000) Cre recombinase: The universal reagent for genome tailoring. Genesis 26, 99–109 https://doi.org/10.1002/(SICI)1526-968X(200002)26:2<99::AID-GENE1>3.0.CO;2-B

Inlay, M.A., Choe, V., Bharathi, S., Fernhoff, N.B., Baker, J.R., Weissman, I.L. et al. (2013) Synthesis of a photocaged tamoxifen for light-dependent activation of Cre-ER recombinase-driven gene modification. Chem. Commun. 49, 4971–4973 https://doi.org/10.1039/c3cc42179a

Lu, X., Agasti, S.S., Vinegoni, C., Waterman, P., Depinho, R.A. and Weissleder, R. (2012) Optochemogenetics (OCG) allows more precise control of genetic engineering in mice with CreER regulators. Bioconjug. Chem. 23, 1945–1951 https://doi.org/10.1021/bc300319c

Schindler, S.E., McCall, J.G., Yan, P., Hyrc, K.L., Li, M., Tucker, C.L. et al. (2015) Photo-activatable Cre recombinase regulates gene expression in vivo. Sci. Rep. 5, 13627 https://doi.org/10.1038/srep13627

Jung, H., Kim, S.W., Kim, M., Hong, J., Yu, D., Kim, J.H. et al. (2019) Noninvasive optical activation of Flp recombinase for genetic manipulation in deep mouse brain regions. Nat. Commun. 10, 314 https://doi.org/10.1038/s41467-018-08282-8

Yao, S., Yuan, P., Ouellette, B., Zhou, T., Mortrud, M., Balaram, P. et al. (2020) Recv recombinase system for in vivo targeted optogenomic modifications of single cells or cell populations. Nat. Methods 17, 422–429 https://doi.org/10.1038/s41592-020-0774-3

Morikawa, K., Furuhashi, K., de Sena-Tomas, C., Garcia-Garcia, A.L., Bekdash, R., Klein, A.D. et al. (2020) Photoactivatable Cre recombinase 3.0 for in vivo mouse applications. Nat. Commun. 11, 2141 https://doi.org/10.1038/s41467-020-16030-0

Olejniczak, J., Carlng, C.J. and Almutairi, A. (2015) Photocontrolled release using one-photon absorption of visible or NIR light. J. Control. Release 219, 19–30 https://doi.org/10.1016/j.jconrel.2015.09.030