A Review of the Optical Control of Protein Function Through Unnatural Amino Acid Mutagenesis and Other Optogenetic Approaches

by

Austin S. Baker

B.S., University of Kentucky, 2011

Submitted to the Graduate Faculty of the

Kenneth P. Dietrich Graduate School of Arts & Sciences in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2014

UNIVERSITY OF PITTSBURGH

Dietrich School of Arts and Sciences

This thesis was presented

by

Austin S. Baker

It was defended on

April 11th, 2014

and approved by

W. Seth Horne, Assistant Professor, Department of Chemistry

Peter Wipf, Distinguished Professor, Department of Chemistry

Committee Chair: Alexander Deiters, Professor, Department of Chemistry

Copyright © by Austin S. Baker

2014

A Review of the Optical Control of Protein Function via Unnatural Amino Acid Mutagenesis and Other Optogenetic Approaches

Austin S. Baker, M.S.

University of Pittsburgh, 2014

Biological processes are regulated with high spatial and temporal resolution at the molecular, cellular, and systems level. To control and study processes with the same resolution, researchers have employed light-sensitive moieties to optically activate and deactivate protein function. Optical control is a non-invasive technique in which the amplitude, wavelength, spatial location, and timing of the light illumination can be easily controlled. This review focuses on applications of genetically encoded unnatural amino acids containing light-removable protecting groups to optically trigger protein function, while also discussing select optogenetic approaches using natural light-sensing domains to engineer optical control of biological processes.

TABLE OF CONTENTS

1.0 INTRODUCTION TO THE OPTICAL CONTROL OF PROTEIN FUNCTION 1

2.0 OPTOGENETIC CONTROL WITH NATURALLY OCCURRING LIGHT-SENSITIVE DOMAINS..4

2.1 OPTOCHEMICAL CONTROL OF MEMBRABE LOCALIZATION THROUGH PHY-PIF DIMERIZATION..4

2.2 OPTOCHEMICAL CONTROL OF TRANSCRIPTION THROUGH CRY-CIB DIMERIZATION..6

2.3 OPTOCHEMICAL CONTROL OF TRANSCRIPTION USING LOV DOMAINS 8

2.4 OPTOCHEMICAL CONTROL OF MOTOR FUNCTION USING A CHANNELRHODOPSIN ION CHANNEL..9

3.0 EXPANSION OF THE GENETIC CODE FOR THE INCORPORATION OF UNNATURAL AMINO ACIDS...11

3.1 EVOLUTION OF THE BIOSYNTHETIC MACHINERY...................................12

3.2 UNNATURAL AMINO ACID MUTAGENESIS 14

4.0 PHOTOCAGED AMINO ACID MUTAGENESIS FOR THE OPTOCHEMICAL CONTROL OF PROTEIN FUNCTION...17

4.1 EARLY APPROACHES TO THE OPTOCHEMICAL CONTROL OF PROTEIN FUNCTION: CAGED ION CHANNEL, CAGED CASPASE AND CAGED GALACTOSIDASE 17

4.2 OPTOCHEMICAL CONTROL OF SIGNAL TRANSDUCTION THROUGH CAGED KINASES AND THEIR SUBSTRATES 20

4.3 OPTOCHEMICAL CONTROL OF PROTEIN TRANSLOCATION THROUGH CAGED TRANSLOCATION SEQUENCES 23

4.4 OPTOCHEMICAL CONTROL OF GENE EXPRESSION: CAGED RECOMBINASES, CAGED NUCLEASES AND CAGED POLYMERASES 24

4.5 OPTOCHEMICAL CONTROL OF CAGED GFP AND LUCIFERASE PROTEIN REPORTERS 27

5.0 CONCLUDING REMARKS 28

BIBLIOGRAPHY 30

LIST OF FIGURES

Figure 1: Different approaches to optochemically control protein function with light. 3

Figure 2: Optogenetic approaches to control protein function with Phytochromes and Cryptochromes..6

Figure 3: Optogenetic approaches to control protein function with LOV domains and Opsin Ion Channels 8

Figure 4: Expansion of the genetic code for UAA incorporation..14

Figure 5: Chemical structures of some genetically encoded unnatural amino acids 16

Figure 6: Optochemical control of Ion Channels and Signaling Pathways 19

Figure 7: Optochemical control of kinase substrate activity and nuclear translocation 22

Figure 8: Optochemical control over T7RNAp and firefly luciferase...26

ABBREVIATIONS

ATP – adenosine triphosphate

aaRS – Amino acyl tRNA synthetase

ChR2 – Channelrhodopsin 2

CIB – Cryptochrome Interacting Basic helix-loop-helix

C2P – Cryptochrome 2 Photolyase related domain

Cry2 – Cryprochrome 2

D1 / D2 – Dopamine 1 or Dopamine 2 receptor

EGFP – Enhanced Green Fluorescent Protein

ERK1/2 – Extracellular signal-Related Kinases 1 or 2

E. coli – Escherichia Coli

FAD – Flavin Adenine Dinucleotide

FRET – Fluorescence Resonance Energy Transfer

JAK – Janus Kinase

LOV – Light-Oxygen-Voltage Sensing domains

MEK1 – Mitogen Activated Protein Kinase Kinase 1

MKP – Mitogen Activated Protein Kinase Phosphatase

nAChR – Nicotinic Acetylcholine Receptor

NLS – nuclear localization sequence

ONB – *ortho*-nitrobenzyl

PhyB – Phytochrome B

PIF – Phytochrome Interacting Factor

qRT-PCR – quantitative reverse transcriptase–polymerase chain reaction

STAT – Signal Transducer and Activator of Transcription

TALE – Transcriptional Activator-Like Effector

T7-RNAp – T7 bacteriophage RNA polymerase

UAAs – Unnatural Amino Acids

UV – Ultraviolet light

ZFN – Zinc Finger Nuclease

1.0 INTRODUCTION TO THE OPTICAL CONTROL OF PROTEIN FUNCTION

Key to the study of biological processes is the ability to selectively control protein function in order to link it to cellular response. Traditional methods to regulate the function at the genome level, such as knock-in, knockdown, or small molecule approaches typically suffer from long induction times, off-target effects, stability and inefficient delivery. Pharmacological manipulation often lacks cell-type specificity and temporal precision; typical genetic approaches often don't have the temporal precision necessary to drive specific perturbations in response to external behavioral and environmental triggers. However, by directly regulating processes at the genomic level, these strategies are often more robust in their control; direct gene regulation can induce more specific and global changes in the gamut of protein function. The use of light to control protein activity, however, provides precise spatial and temporal control to biological processes using a non-invasive control element, and allows more precise control than other biological perturbations.[1–10] Furthermore, optically controlled systems have been developed within the realm of "classic" gene manipulation techniques, to allow more spatiotemporal control.[11–15]

The genetic encoding of light-responsive proteins, termed "optogenetics", has taken advantage of naturally occurring photosensitive protein domains. Optogenetic approaches have been used to control biological processes such as transcription, protein localization, protein phosphorylation, and ion channel activity.[16–22] Commonly employed, naturally occurring

photoreceptors include the rhodopsin family of ion channels, Light-Oxygen-Voltage (LOV) sensitive domains, as well as phytochrome and cryptochrome proteins. These receptors respond with reversible conformational changes upon exposure to light.

The genetic incorporation of unnatural structures relies on suppression of naturally occurring stop codons, through the use of tRNA that naturally suppress these truncation sites. Stop codon suppressors were noted with point mutations in the DNA that didn't correspond to the resulting amino acid, where either the natural or structurally similar amino acid was incorporated. Reports for the suppression of a naturally occurring stop codon emerged in the 70s and early 80s, and design of *in vitro* suppression systems were developed using acylated tRNAs.[23–25] An *in vitro* stop codon suppression system was reported, where several unnatural phenylalanine derivatives were incorporated through chemically misacylated suppressor tRNAs,[26] and set the foundations for the incorporation of unnatural amino acids (UAAs). Although the use of the *ortho*-nitrobenzyl (ONB) group as a light-sensitive protecting group dates back to the 1960s,[27] it's use as a light-sensitive protecting group for amino acids by R. B. Woodward, gave it an impetus for further applications.[28] It was reported much later as a general method for masking nascent amino acid functionality; through stop codon suppression, an ONB-protected Aspartate was incorporated into the phage T4 lysozyme.[29] It wasn't until mutation of the synthetase active site generated a fully genetically encoded system, as developed by the Schultz lab.[30,31]

Using this methodology, light-responsive proteins have also been engineered based on the expansion of the genetic code with unnatural, photocaged amino acids. UAAs can be genetically encoded through the addition of orthogonal protein biosynthetic machinery to pro- and eukaryotic cells.[32,33] The use of light-cleavable protecting groups – so called "caging groups"

– allows for modulation of protein function with a single, small, covalently-linked chromophore that can be removed through exposure to non-damaging UV light, thereby restoring the native protein function.[8–10] Here, recent developments in the engineering of light-sensitive proteins and their application to the optical control of enzymatic and cellular processes will be discussed, with an emphasis on the use of genetically encoded, caged amino acids (**Figure 1**).

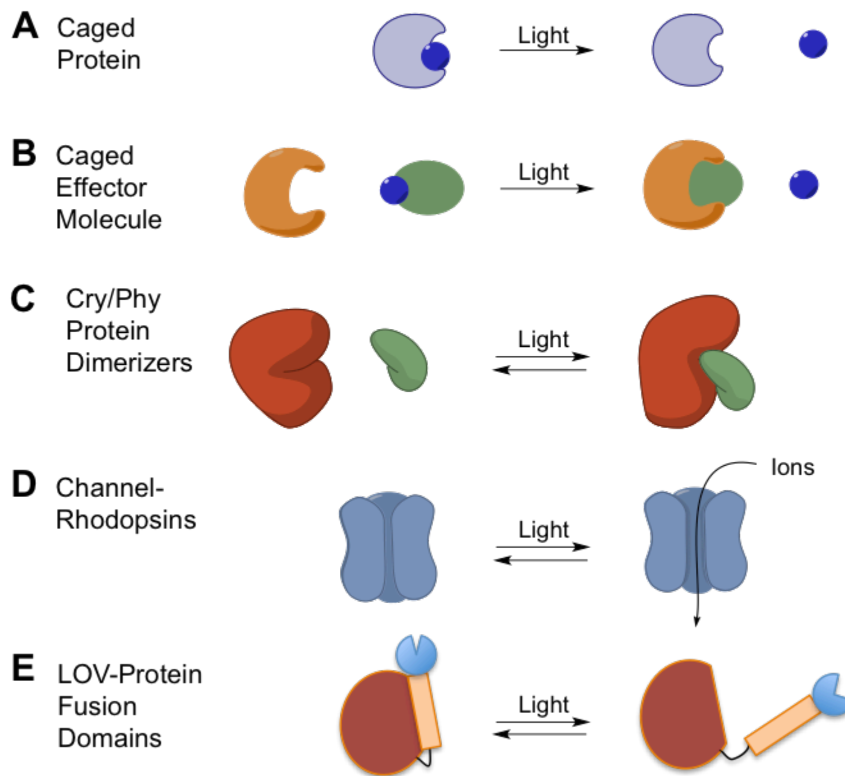

Figure 1: Different approaches to optochemically control protein function with light.

A) Caged amino acids can be genetically incorporated into proteins of interest, thereby rendering them inactive, until UV irradiation removes the caging group and restores protein activity. B) An effector molecule can be caged and UV irradiation can regenerate the active molecule. C) Protein dimerization can be induced upon light exposure, e.g., of cryptochrome (Cry) and the Cry Interacting Basic helix-loop-helix (CIB). D) Light-sensitive ion channels, such as channelrhodopsin, can be activated upon light exposure. E) Protein activity can be controlled by fusion of a LOV domain that generates an active enzyme upon irradiation.

2.0 OPTOGENETIC CONTROL WITH NATURALLY OCCURRING LIGHT-SENSITIVE DOMAINS

Recent developments in the understanding and engineering of naturally light-sensitive proteins have greatly advanced the application of optogenetic tools to regulate complex biological processes. Photon absorption induces conformational changes in these proteins through isomerization of a double bond (e.g., in opsin ion channels or in phytochromes) or covalent adduct formation between the protein and the chromophore (e.g., in cryptochromes or in LOV domains). Regeneration of the chromophore can occur via absorption of another photon of different energy, or as a slower thermal process in the dark. Most often, these optogenetic techniques require significant protein engineering and optimization in order to obtain excellent light-switching behavior, if the desired process to be regulated differs from optochemical ion channel activation. However, these genetically encoded fusion proteins are capable of a high degree of spatiotemporal control over diverse processes.

2.1 OPTOCHEMICAL CONTROL OF MEMBRABE LOCALIZATION THROUGH PHY-PIF DIMERIZATION

Phytochromes (Phy) are proteins that switch between two stable states through isomerization of the bilin chromophore by absorbing red (Pr state) and far-red light (Pfr).[34]

Absorption of ~650 nm light converts the Pr form into the Pfr conformation, to allow binding with the Phytochrome Interacting Factor (PIF); irradiation with ~750 nm far-red light switches it back to the Pr state. These processes are rapid and repeatable with no observable chromophore decomposition. Thus, any proteins of interest could be brought in close proximity to one another through an optically triggered dimerization, if fused to Phy or PIF. The optical control of protein-protein interactions in cells was first demonstrated with the PhyB-PIF3 pair, showing that PIF3 dimerized with PhyB in the Pfr, but not the Pr state.[35] This reversible, PhyB-PIF3 dimerizing system was used to optically control transcription in yeast through fusion of Gal4 binding (Gal4BD) and activation domains (Gal4AD) to PhyB and PIF3, respectively.

The Phy-PIF system was also employed to optically regulate protein localization to the membrane, enabling spatial activation.[36] Using PhyB-PIF constructs, rates of membrane association and dissociation were determined to be on the second timescale, faster than previous methods that used small molecules to induce membrane localization.[37–39] The guanine exchange factor (Tiam-DH-PH) activates the GTPase Rac1, which is in turn responsible for dynamic regulation of the actin cytoskeleton. Using a Tiam-DH-PH–PIF and a membrane bound PhyB fusion constructs, reorganization of the cytoskeleton was demonstrated in live NIH3T3 cells, showing that spatially localized, activating red light could induce lamellapodial protrusion formation through localized activation of Rac1, via light-induced PhyB-PIF dimerization, and recruitment to the membrane by the Rac-activating Tiam (**Figure 2A**).

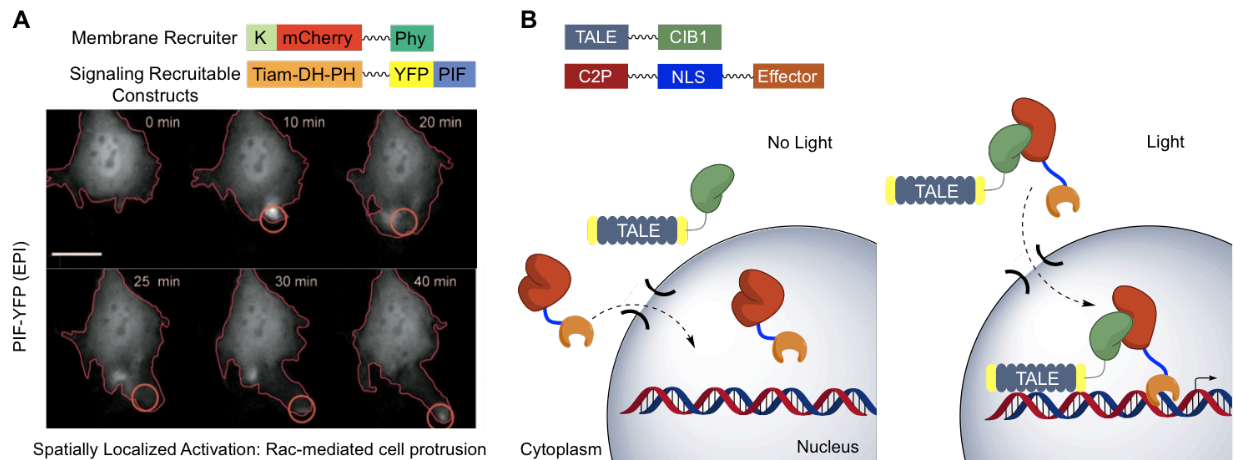

Figure 2: Optogenetic approaches to control protein function with Phytochromes and Cryptochromes.

A) Fused constructs for membrane localization and recruitment of signaling proteins (top). Light-induced dimerization of the fusion constructs, leads to protrusion formation upon localization of red light irradiation away from the cell body (bottom). Scale bar is 20 μm. B) Nuclear translocation upon blue light irradiation and CRY2-CIB dimerization to control transcriptional activation. TALE: Transcriptional Activator-Like Effector, C2P: CRY2-PHR (PHotolyase Related domain), Effector: transcriptional activator (VP64) or repressor (SID4X). Figures taken with permission from references [36] (A) and [40] (B).

2.2 OPTOCHEMICAL CONTROL OF TRANSCRIPTION THROUGH CRY-CIB DIMERIZATION

Cryptochromes (Cry) are blue-light sensing flavin binding proteins that undergo covalent adduct formation with the flavin adenine dinucleotide (FAD) chromophore and a corresponding conformational shift when irradiated.[41–43] These conformational changes allow dimerization of Cry and its binding partner CIB. Activity of fused proteins can be controlled in a similar manner

to the Phy-PIF system, and dimerization is thermally reversible in the dark. Dimerizing Cry-CIB constructs have been used to induce membrane localization through blue light illumination using cytosolic Cry2-mCherry and membrane-associated CIB-EGFP components. Near complete membrane localization occurs 10 seconds after irradiation in HEK293T cells.[16] Additionally, a split Cre recombinase system, fused to Cry2 and a truncated CIB, was developed for the optochemical control of DNA recombination through light-induced protein dimerization. Cre recombinase catalyzes strand exchange between defined target sequences, recognizing palindromic *loxP* sites to insert, delete or invert DNA sequences.[44] A low level of recombination was observed in the dark, with a 158-fold increase in EGFP expression through excision of a stop codon after irradiation with 450 nm blue light.

A similar system was designed, using a customizable Transcription Activator-Like Effector (TALE) DNA-binding domain and TALE activator domains, fused to either Cry2 or CIB1, to photochemically modulate transcription of endogenous mammalian DNA upon Cry-CIB dimerization.[40] Using a TALE domain specific for the *Grm2* gene sequence and a transcriptional activator (VP64), TALE(*Grm2*)-CIB1 and CRY2-NLS-VP64 constructs were generated, which reduced background transcriptional activity of TALE in the absence of light by incorporating a nuclear localization signal (NLS) in the effector fusion domain. Irradiation allowed for nuclear translocation of the Cry-CIB dimer and activation of *Grm2* transcription (**Figure 2B**). Customization of the TALE binding domain allows for localization to any gene sequence, and switching from a transcriptional activator to a repressor allows for positive and negative regulation of transcriptional activity, thus creating a highly modular optogenetic gene-control system.

2.3 OPTOCHEMICAL CONTROL OF TRANSCRIPTION USING LOV DOMAINS

Similar to Cryptochromes, LOV domains also use a covalent flavin adduct to modulate conformational changes.[45] Earlier studies used structural shifts in LOV fusion proteins to modulate the activity of a tryptophan-activated protein,[46] dihydrofolate reductase,[47] and zinc finger transcription factors.[48] More recently, a LOV fusion was used to optically control peptide affinity for their binding partners by reversibly blocking binding-critical residues (**Figure 3A**).[49] For example, invasive plasmin antigen A (ipaA), a vinculin binding protein (VinD1), shows a sequence overlap with the LOV Jα helix, allowing for optochemical control over ipaA-VinD1 dimerization. Gal4AD-LOV-ipaA and VinD1-Gal4BD fusion proteins were generated and blue-light-activated transcription was achieved in yeast. Irradiation increased Gal4-driven *LacZ* expression 10-fold compared to the dark state, achieving 30% of wild-type ipaA activity. A similar system was also developed for mammalian cells using small stable RNA A (SsrA) protein and its adapter protein SspB, illustrating that photoswitchable, steric blocking of peptide-protein interactions can be designed based on LOV fusions and may be generally applicable.

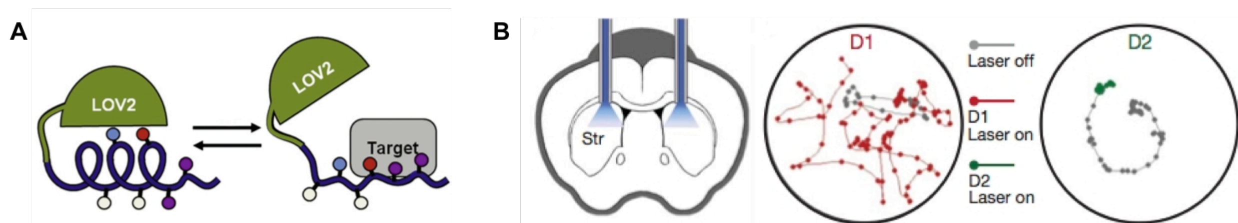

Figure 3: Optogenetic approaches to control protein function with LOV domains and Opsin Ion Channels

A) LOV domain fused to a peptide sequence. Irradiation induces helix deformation and allows the target to bind. Blue circles: critical for LOV association. Purple circles: critical for target-peptide recognition. Red circles: critical for both. White circles: unimportant for binding. B) Fiber optic probes were inserted into the dorsomedial

striatum of mice to deliver blue light and optochemically control motor function (left). Movement path (lines) for D1-ChR2 mice and D2-ChR2 mice were measured every 300 ms for 20 sec total (right), to activate and inhibit movement, respectively. Figures taken with permission from references [49] (A) and [50] (B).

2.4 OPTOCHEMICAL CONTROL OF MOTOR FUNCTION USING A CHANNELRHODOPSIN ION CHANNEL

Opsin ion channels bind retinal as a blue light-sensitive chromophore and photon absorption converts the all-*trans* isomer to the 13-*cis* isomer, inducing conformational changes in the channel that lead to pore opening and ion influx. The channelrhodopsin (ChR) subfamily has been widely applied as they provide benefits over other optogenetic ion channels. They show ion flux comparable to natural ion channels[51] and better efficacy over previous photostimulatory methods.[52] ChR2 allows Na^+, Ca^{2+}, H^+ and K^+ influx in a pH dependent manner, and controls whole-cell electrochemical activation with rapid switching between lit and dark states within milliseconds.[19] ChR2 has been used to optochemically control pathway-specific motor function, involved in Parkinson's disease, within the dorsomedial striatum region of the brain of mice.[50] Two dopamine receptor-based pathways, D1 and D2, have been proposed to activate and inhibit motor function, respectively.[53,54] A Cre recombinase-dependent viral expression system was used, in which ChR2-YFP is only expressed in neurons containing either D1- or D2-receptors.[55] Light irradiation through fiber optic cannulas inserted into mice striata selectively activated ChR2-expressing neurons and confirmed that D1-pathway activation increased movement, while D2-pathway activation decreased movement (**Figure 3B**). The mouse's gait was found to be unaffected in D1-ChR2 and D2-ChR2 mice after illumination, indicating that the dorsomedial

striatum mediates motor control, but not motor coordination. Taken together, these optogenetic applications emphasize the diversity of biological processes that are optochemically controlled by light using naturally occurring photosensory domains.

3.0 EXPANSION OF THE GENETIC CODE FOR THE INCORPORATION OF UNNATURAL AMINO ACIDS

With proteins being the primary workforce of the cell, they are able to carry out a diverse set of functions through a limited set of structures; the chemistries available to the canonical 20 amino acids are sufficient to sustain life. However, exogenous cofactors such as flavin, heme groups, vitamins, and metal ions, are often needed to assist in the desired transformations. Moreover, post-translational modifications further the diversity of the natural 20 amino acids through modifications such as acetylation, methylation, phosphorylation and glycosylation. With the human genome having roughly 25,000 genes and 500,000 proteins, added modifications are required to drive a biological system, notably through these post-translational modifications to regulate function. Each post-translational modification, and requisite cofactor, is highly and dynamically regulated. Several archaea and eubacteria have added to their chemical diversity, and been shown to encode noncanonical amino acids, such as selenocysteine[56] and pyrrolysine.[57] Thus the creation of an expanded genetic code, beyond the common 20 amino acids, allows for the incorporation of enhanced or novel chemistries to probe protein function within a biological context.

3.1 EVOLUTION OF THE BIOSYNTHETIC MACHINERY

To encode an unnatural structure, site-specifically, into a defined peptide sequence, one requires a host-orthogonal method to acylate a tRNA with the UAA and a codon that can recognize the amino-acylated tRNA to allow specific incorporation into the growing peptide chain. Specifically, tRNA synthetases have been engineered to acylate their cognate tRNAs exclusively with the UAA and not any of the canonical amino acids, and retain selective recognition of the orthogonal tRNA over any endogenous tRNAs.[58–62] The UAA-tRNA pair can then deliver the UAA site-specifically into the growing peptide chain; to accomplish this, a less common stop codon or frameshift codon is used to allow recognition of the orthogonal tRNA by the host biosynthetic machinery, over any endogenous tRNAs.[63,64] The host-orthogonal aaRS binds the UAA and facilitates acylation of the UAA onto the orthogonal tRNA; the UAA is then delivered to the peptide chain in response to a stop or frameshift codon, e.g., in response to the amber stop codon, TAG, (**Figure 4**). Any endogenous aaRS or tRNA will not cross-react with the unnatural aaRS or amino acid. The amino-acylated tRNA delivers the UAA site-specifically to any desired position within a protein of interest in both pro- and eukaryotic live cells.

In order for the aminoacyl tRNA synthetase (aaRS) to discern and install the UAA onto the tRNA, the active site of the aaRS can be randomly mutated to generate a library of mutant synthetases, which are then put through rounds of positive and negative selection to ensure selective UAA incorporation.[31,33,65] Evolution of the synthetase active site is accomplished by identifying residues within the binding pocket, and randomly mutating each residue to any of the common 20 amino acids, this typically generates about 10^6 to 10^8 mutants. The positive selection round is based upon the conference of resistance to chloramphenicol by suppression of an amber mutation at a permissive site in chloramphenicol acetyl transferase. Resistance to

chloramphenicol is conferred if either the natural or unnatural amino acid is charged onto the cognate tRNA. The surviving mutants are then put through a negative selection process using the barnase gene, and in the absence of the unnatural amino acid. If the natural amino acid is incorporated into the barnase gene, the cells will die. Sequential rounds of positive and negative selection are done to generate a manageable library of mutants that are selective for the UAA in response to the stop codon.

Initial efforts to reprogram the genetic code focused on *E. coli,* because of the wealth of knowledge of its translational machinery, and the ease of manipulation of the genetic code.[26] The amber codon is the least used stop codon in *E. coli* and mammalian cells, therefore using it to incorporate additional amino acids is less detrimental than replacing a more commonly used stop codon. In addition, recognition of the tRNA by the ribosome can be weakly dependent on the anticodon arm,[66] and the length and binding recognition of the variable arm is species dependent;[67] thus, non-native tRNA sequences can allow for the tRNA to still be accepted by the host ribosomal machinery. This approach of engineering an orthogonal aaRS/tRNA pair has been used with other host species, such as mammalian and yeast cells, and tRNA allows for selectivity such that it incorporates the UAA and not any native amino acids, and still functions within the host machinery.

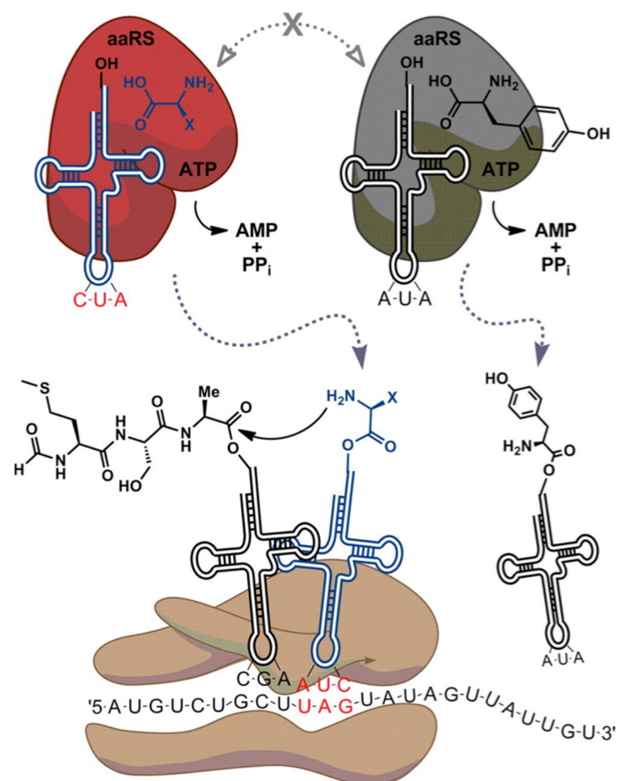

Figure 4: Expansion of the genetic code for UAA incorporation.

The orthogonal aaRS (*red*) acylates the orthogonal tRNA (*blue with red anticodon*) with the unnatural amino acid (*blue X*), which then transfers the UAA to the growing peptide chain in response to the amber codon (*red UAG*). The endogenous aaRS/tRNA pair is shown (*grey*), where it acylates the endogenous tRNA with the natural amino acid (*black Tyr*). The ribosome is shown in *brown*. Figure taken from reference [68] with permission.

3.2 UNNATURAL AMINO ACID MUTAGENESIS

Since the discovery of the incorporation of selenocysteine[56] and pyrrolysine[57] encoded by archaea and eubacteria, there have been nearly 100 unnatural amino acids that have been genetically encoded for various chemistries, a sample of which are shown in **Figure 5**. To this end there have been several UAAs that have been incorporated as bioorthogonal handles (**1-4**). *p*-benzoylphenylalanine (**1**) and the diazarine (**2**) have been incorporated as photocrosslinking

agents for investigating biomolecular interactions; irradiation at ~350 nm UV light induces non-discriminate insertion within C-H bonds.[69,70] Similarly *p*-azidophenylalanine (**3**) and proprargyloxyphenylalanine (**4**) have been used as species for "click chemistry" as bioorthogonal conjugates.[59,71,72] The tyrosine derivatives 3-hydroxytyrosine (**5**) and 3-nitrotyrosine (**6**) have been incorporated into myoglobin[73] and MnSOD[74] proteins, respectively, as the unnatural amino acids are commonly involved in an altered redox state. In addition, aminophenylalanine (**7**) was incorporated as a probe for a redox-active tyrosine,[75] used for the generation of a bacterium with an expanded genetic code, which could biosynthesize the unnatural amino acid,[76] and as a synthetase active site probe.[65] Hydroxycoumarin[77] (**8**) and dansyl[78] (**9**) derivatives have been incorporated as fluorescent amino acids, which have been used as an indicator of protein conformation and intermolecular interactions. They are more advantageous than GFP, as they can replace single residues within a protein versus a full sequence peptide.

Furthermore, other light-sensitive amino acid derivatives have been incorporated into proteins of interest. Diazophenylalanine (**10**) was incorporated into the BamHI endonuclease,[79] and catabolite activator protein (CAP)[80], both at their homodimeric interfaces, and is being pursued for other applications (unpublished results, Deiters Lab). As well, caged amino acids have been genetically incorporated, in which irradiation induces photolytic cleavage of a protecting group and regenerates the native amino acid (and thus native protein). Various caged amino acids (**11-14**) have been used to control processes such as translocation, translation and signaling, and their applications will be presented herein.

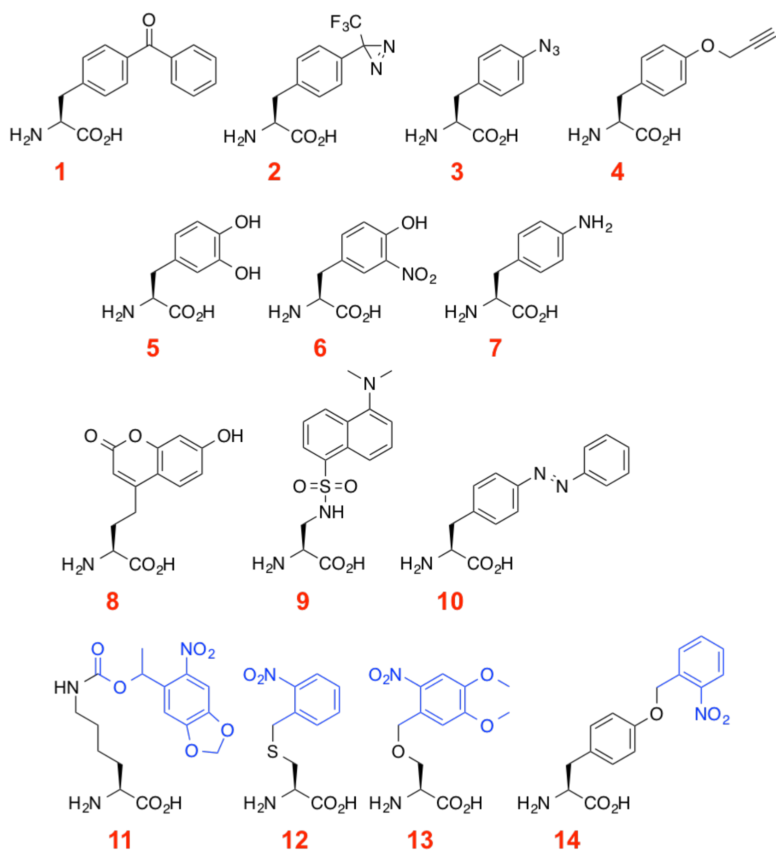

Figure 5: Chemical structures of some genetically encoded unnatural amino acids

Structures of some genetically incorporated amino acids. The *ortho*-nitrobenzyl (ONB) group of the photocaged amino acids (**11-14**) is shown in blue.

4.0 PHOTOCAGED AMINO ACID MUTAGENESIS FOR THE OPTOCHEMICAL CONTROL OF PROTEIN FUNCTION

Caging groups have commonly been introduced into proteins and peptides through chemical modification.[81,82] Issues frequently faced using these approaches are size limitations of synthetic peptides, requirements for chemical ligation approaches,[83,84] non-specific caging group installation onto expressed proteins, and the necessity to deliver the caged peptides and proteins into cells and organisms for biological studies. These problems have been addressed through the genetic encoding of caged amino acids in cells with an expanded genetic code.[32,33] Using this amber codon suppression approach, ONB derivatives (**Figure 5**) of Lys (**11**),[62,85–87] Cys (**12**),[58,88] Ser (**13**),[61]and Tyr (**14**),[60,89–91] have been genetically encoded in pro- and eukaryotic cells, and have been applied to the optochemical control of a wide range of biological processes.

4.1 EARLY APPROACHES TO THE OPTOCHEMICAL CONTROL OF PROTEIN FUNCTION: CAGED ION CHANNEL, CAGED CASPASE AND CAGED GALACTOSIDASE

Preceding opsin-mediated control, nicotinic acetylcholine receptor (nAChR) ion channels have been light-activated through site-specific incorporation of a caged tyrosine.[92] This was achieved through chemical acylation of an amber-suppressor tRNA and injection into *Xenopus*

oocytes. The caged tyrosine was placed at three specific, highly conserved Tyr residues of a nAChR, two of which participate in acetylcholine binding. Voltage clamping was used to measure membrane potential during flash decaging of the ONB caging group and ion channel activation with precise temporal control. Irradiation power and duration was tuned to decage ~5% of the caged tyrosine residues per flash, allowing for depolarizing activation of decaged ion channels, while the cells were kept in the presence of high concentrations of agonist. Changes in electrochemical current after light pulses marks an increase in cell potential, followed by a compensatory repolarization to restore the cell to resting potential (**Figure 6A**). Time-resolved kinetic analyses of changes in electrochemical potential at each residue showed that each residue has differing effects on pore opening. Two-activation phases were noted: one closely resembling physiological activation, likely due to incomplete incorporation or baseline activity, and the other indicating a conformational change critical for activation after diffusion of the caging group. The generation of light-activated ion channels allowed for the unprecedented light-mediated control over pore activation, and structure-function interpretations of agonist binding.

In order to eliminate the need for chemical tRNA acylation and injection, tRNA/tRNA synthetase pairs were engineered to genetically encode caged amino acids. Early applications included amber codon suppression in EGFP,[62] optochemical control of caspase activity through a caged cysteine,[58] and β-galactosidase function through a caged tyrosine.[60] The first incorporation of a caged cysteine into proteins in mammalian cells was also recently reported.[88] The caged cysteine was incorporated into the Tobacco Etch Virus (TEV) protease in live HEK293T cells and optochemical control of protease activity was monitored with a FRET assay.

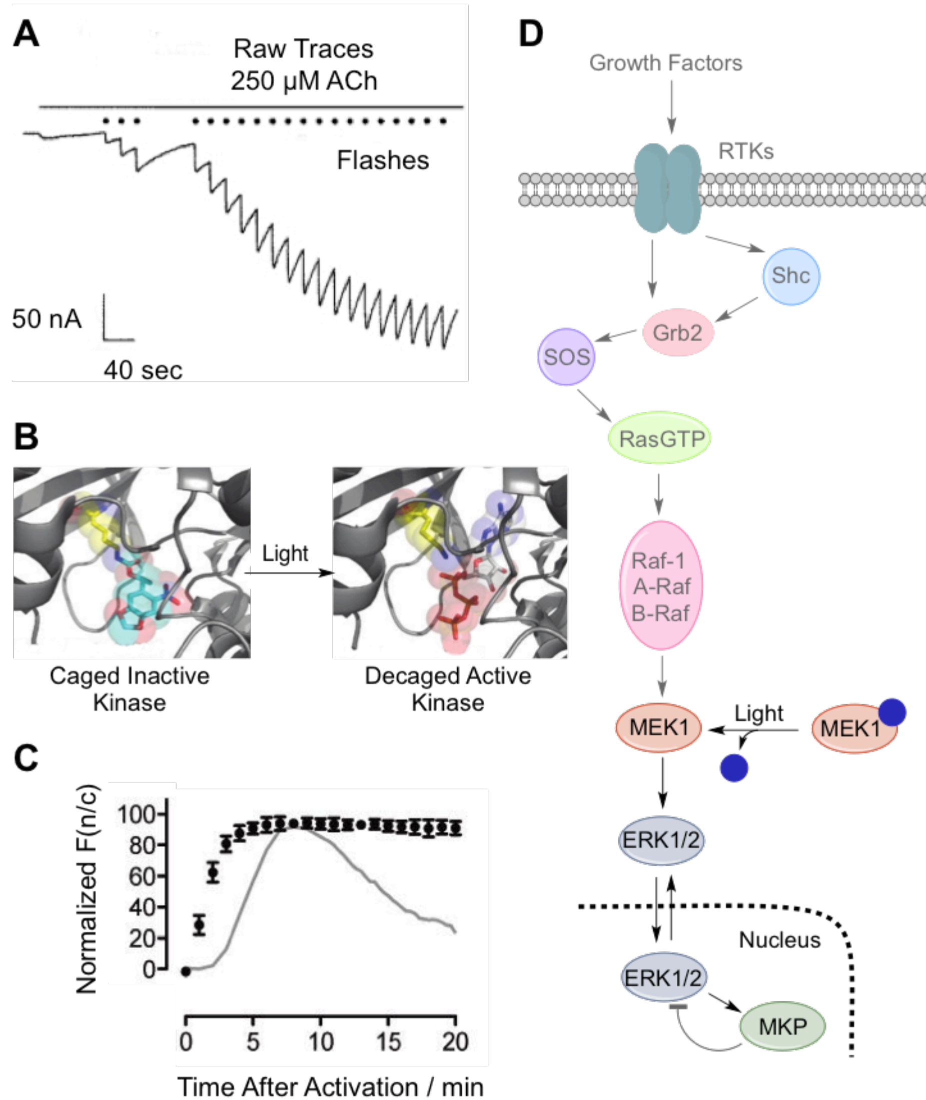

Figure 6: Optochemical control of Ion Channels and Signaling Pathways

A) Electrochemical current over time induced through flash decaging, every 20 sec with an 80 sec interval between flash 3 and 4 (as indicated by dots), of a tyrosine-caged nAChR in *Xenopus* oocytes. B) Caged MEK1 kinase is inactive due to occlusion of ATP from the active site by the caging group, through blocking of a critical ATP-Lys interaction. Irradiation removes the caging group and restores kinase activity. C) Normalized nuclear-to-cytoplasmic fluorescence, F(n/c), measuring ERK1/2 nuclear translocation after photoactivation, indicating a steady state of ERK1/2 in the nucleus (black circles). In contrast to light-activation, EGF stimulation (grey line) leads to a peak in nuclear ERK1/2. D) Stimulation of the MEK signaling cascade by growth factors activates the whole

pathway, whereas photoactivation of caged MEK1, enables partial pathway activation. Figures taken, with permission, from references [92] (A) and [87] (B-D).

4.2 OPTOCHEMICAL CONTROL OF SIGNAL TRANSDUCTION THROUGH CAGED KINASES AND THEIR SUBSTRATES

Optochemical control over kinase signaling was achieved through genetic incorporation of a caged lysine into the MAP kinase kinase MEK1, by replacing a highly conserved Lys that is critical for anchoring ATP within the catalytic domain (**Figure 6B**).[87] MEK1 is part of a signaling network that propagates signals through ERK1/2 phosphorylation. The caged kinase was completely inactive showing no catalytic activity until decaging through UV exposure, as shown by the analysis of phosphorylation levels of downstream substrates of ERK1/2. After phosphorylation by activated MEK1, ERK translocates from the cytoplasm to the nucleus. Optochemical activation, in conjunction with an EGFP-tagged ERK2, allowed for the measurement of translocation kinetics from a precisely defined starting point (**Figure 6C**). ERK2 translocation occurs faster through direct light-activation of caged MEK1 ($t_{1/2}$ = 1.5 min) and a steady state of ERK2 in the nucleus was observed. In contrast, EGF-stimulated whole-pathway activation leads to slower translocation ($t_{1/2}$ = 4.5 min) and only a temporary peak in nuclear ERK2. This type of desensitization and adaption to persistent stimuli has been documented in cellular systems and is likely to involve negative feedback. Here, once in the nucleus, ERK2 is dephosphorylated by MKP, constituting a negative feedback loop and leading to diffusion of ERK2 back into the cytoplasm. However, light-activation of caged MEK1 shows a stationary stimulus and unlike EGF stimulation of the entire pathway (**Figure 6D**); the light-activated sub-

network is not subjected to adaptation. Thus, adaptation most likely occurs upstream of MEK1, independent of MKP function.[93]

In contrast to caging the active site of a kinase, a corresponding phosphorylation substrate has been caged as well. This was demonstrated by incorporating caged tyrosine at the phosphorylation site of Signal Transducers and Activators of Transcription (STAT1).[89] The Janus Kinase (JAK) /STAT pathway is activated through cytokine binding to extracellular surface receptors and phosphorylation of STAT to induce its homodimerization, nuclear translocation, and activation of transcription.[94,95] Interferon stimulation of HEK293 cells expressing a caged STAT1 showed phosphorylation only after photolytic cleavage of the caging group from the phosphorylation site (**Figure 7A**).

To monitor the kinetics of nuclear translocation of the transcription factor Pho4, a caged serine was incorporated at several phosphorylation sites.[61] Depending on the concentration of inorganic phosphate, Pho4 is either phosphorylated at the Ser6 position and localized to the nucleus to activate gene expression, or is phosphorylated at Ser2 and Ser3 and is exported to the cytoplasm. Introduction of a caged serine at Ser3 blocked this phosphorylation site and retained GFP-tagged Pho4 in the nucleus until irradiation removed the caging group, allowing for Ser3 phosphorylation and nuclear export (**Figure 7B**). Nuclear-to-cytoplasmic GFP-Pho4 localization was quantified by a fluorescence line scan, enabling the measurement of translocation kinetics from a defined start point (**Figure 7B**). Optochemical activation of Ser3-caged Pho4 showed faster translocation (with a time constant of ~29 sec) than Ser2-caged Pho4 (with a time constant of ~47.0 sec), in agreement with previously reported kinetics.[96]

Common strategies for perturbing signaling networks include gene knock-out and knock-in approaches. Even though these approaches are suitable for epistasis experiments, they suffer

from long induction times, which limits their usefulness in dynamic studies and enables cellular compensation effects. In contrast, the examples of optochemical activation using caged amino acids summarized here can provide precise spatiotemporal control over specific nodes of signaling networks with very high temporal resolution.

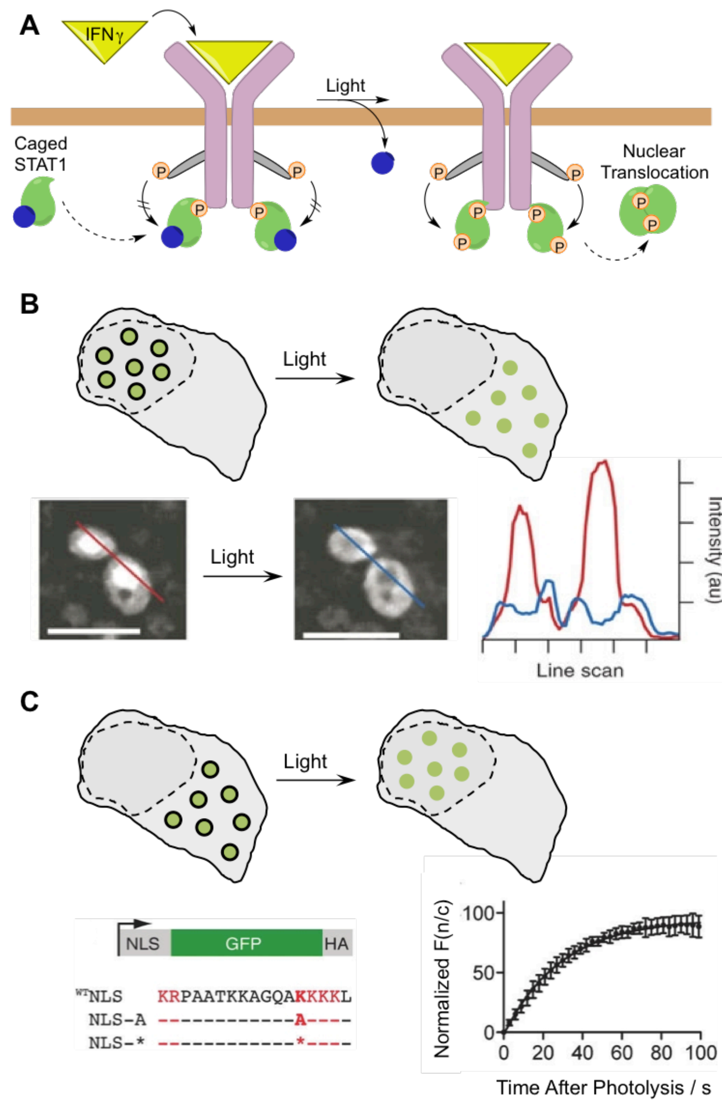

Figure 7: Optochemical control of kinase substrate activity and nuclear translocation

A) Caged STAT1 is unable to be phosphorylated by JAK and thus is localized in the cytoplasm. Irradiation removes the caging group to allow phosphorylation and STAT1 homodimerization, which induces translocation to the nucleus. B) Irradiation of the caged Pho4 transcription factor leads to its export from the nucleus. Changes in fluorescence were measured by a fluorescence line

scan of two magnified cells. Scale bars are 10 μm. C) Irradiation of a caged NLS-GFP construct induces its translocation to the nucleus. An asterisk in the sequences indicates the caged lysine residue. Analysis of the changes in F(n/c) as a function of time after irradiation reveals a time constant for translocation of ~20 sec. Figures taken with permission from references [89] (A), [61] (B) and [86] (C).

4.3 OPTOCHEMICAL CONTROL OF PROTEIN TRANSLOCATION THROUGH CAGED TRANSLOCATION SEQUENCES

Optochemical control of nuclear translocation was also achieved by site-specific incorporation of caged lysine into a classical bipartite NLS (**Figure 7C**).[86] The bipartite recognition sequence consists of two basic amino acid domains, each of which helps to form a complex with an importin receptor to cross the nuclear membrane.[97] An NLS-EGFP reporter was constructed to analyze nuclear translocation after UV-induced NLS decaging in live cells. The caged NLS-EGFP construct was found to be cytosolic, consistent with alanine mutants; irradiation increased nuclear import roughly four-fold when measuring the nuclear-to-cytosolic fluorescence, revealing a time constant of ~20 sec (**Figure 7C**). The NLS of p53[86] and SatB1[98] were also caged to monitor nuclear translocation; thus, optical protein localization may be applicable to any protein of interest.

4.4 OPTOCHEMICAL CONTROL OF GENE EXPRESSION: CAGED RECOMBINASES, CAGED NUCLEASES AND CAGED POLYMERASES

Several DNA cleaving proteins have been engineered to control sequence-specific strand breaks.[99–102] These systems have been applied to manipulate genomes, induce gene deletions, and gene additions. Optochemical control over gene editing was first demonstrated using the Cre recombinase-*loxP* system.[103] A nucleophilic tyrosine in the active site was caged, rendering the protein completely inactive until the caging group was removed through UV exposure. Spatiotemporal control of DNA recombination in mammalian cells was demonstrated through transfection of the purified protein, followed by localized irradiation and a fluorescent readout of the simultaneous deactivation of GFP and activation of DsRed.

Moreover, a zinc-finger nuclease (ZFN) system has been caged for optical control of DNA cleavage through caged tyrosine incorporation at a site crucial for DNA binding.[104] The caged chimeric proteins consist of an N-terminal domain, with three zinc finger domains for sequence-specific DNA binding and a C-terminal FokI restriction enzyme. This artificial endonuclease recognizes a 24 bp long DNA sequence and initiates a double strand break. The caged ZFN was completely inactive, thereby preventing previously reported ZFN-mediated non-specific DNA cleavage[105] and cellular toxicity.[106] A brief UV irradiation restored DNA-cleavage activity and application of the light-activated ZFN in mammalian cells was demonstrated by using a luciferase reporter. The ZFN recognized a FokI cleavage site introduced within the luciferase gene sequence. Light-activation of the ZFN induces a double-strand break, followed by homologous recombination that repairs the previously non-functional luciferase gene and leads to the expression of an intact enzyme. Transfection of the reporter construct and the caged ZFN into HEK293T cells showed no luciferase activity before irradiation, however, UV

exposure activates the ZFN to wild-type levels leading to an optically induced, sequence-specific double-strand break and luciferase expression.

A similar strategy was employed with both a caged DNA[91] and RNA[85,90] polymerase in order to obtain optical control of oligonucleotide polymerization. The T7 RNA polymerase (T7-RNAp) consists of three classical domains resembling a right hand, a structure that is commonly found in other polymerases.[107,108] The O-helix of the finger domain is critical for activity by facilitating orientation of the incoming nucleotide, and is highly conserved across many DNA and RNA polymerases. A crucial tyrosine within the O-helix plays a role in ribose vs. deoxyribose nucleotide discrimination, and moving the new RNA strand out of the active site.[109,110] Thus, caging of this tyrosine allowed for optochemical control of T7RNAp function.[90]

A different approach was taken through the site-specific caging of a lysine residue that is involved in binding the incoming nucleotide phosphate groups to enable optical control of T7RNAp activity in mammalian cells (**Figure 8A**).[85] T7-driven luciferase expression showed no activity of the caged polymerase before irradiation and a 10-fold increase in activity after a brief UV exposure of HEK293T cells. Additionally, EGFP expression under T7-driven promoter control was used to demonstrate spatial control of T7RNAp activity as only illuminated cells showed reporter gene expression (**Figure 8B**). This system was then applied to photochemically control RNA interference through T7-driven expression of a small hairpin RNA (shRNA). As proof-of-principle, an shRNA targeting Eg5 gene repression was optically triggered (**Figure 5C**), showing an expected binucleated cell phenotype only after UV illumination. qRT-PCR quantification of Eg5 mRNA showed a 55% reduction in the case of both wild-type and light-activated T7RNAp. Overall, these examples illustrate the ability to optically control DNA and

RNA editing and expression through site-specific incorporation of genetically encoded photocaged amino acids into oligonucleotide-processing enzymes.

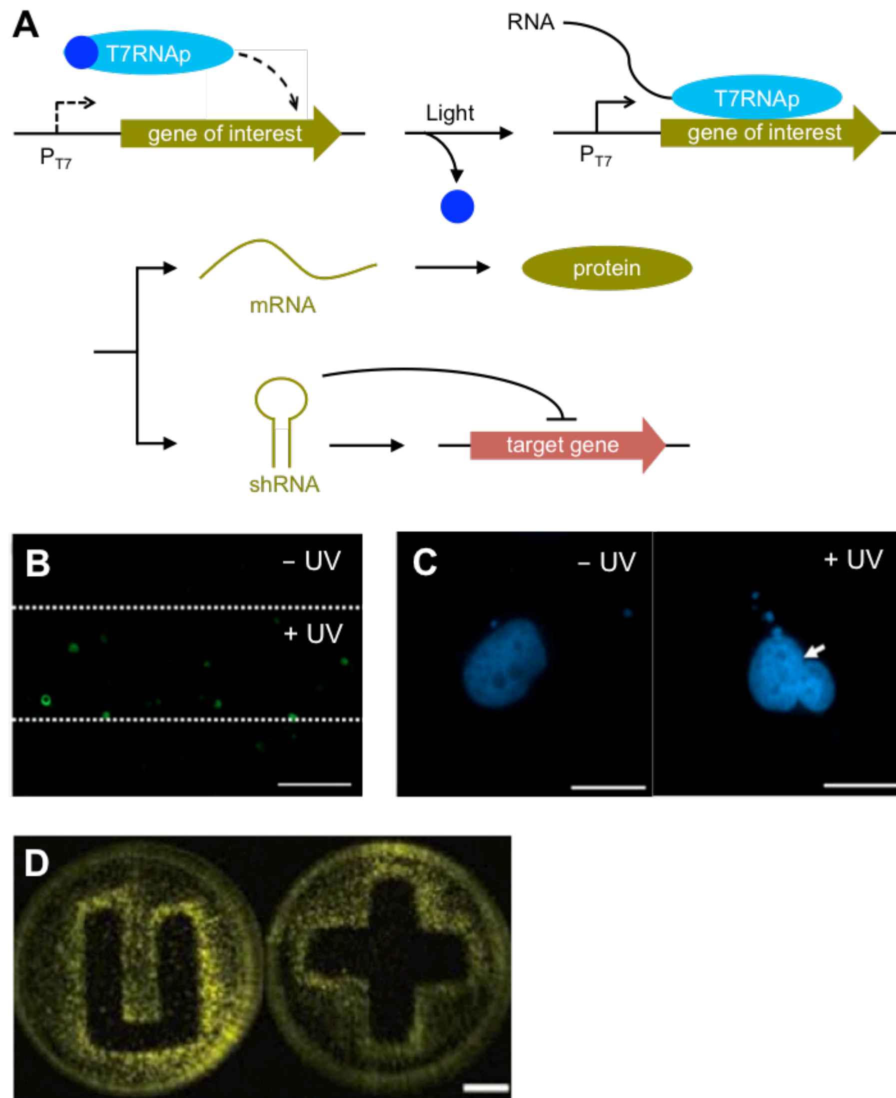

Figure 8: Optochemical control over T7RNAp and firefly luciferase

A) A caged Lys was incorporated into the active site of T7RNAp in mammalian cells rendering the polymerase inactive. Irradiation removes the caging group and allows for translation of, e.g., mRNA or shRNA sequences of interest. B) Masked irradiation of live HEK293T cells expressing caged T7RNAp to drive localized EGFP expression after irradiation. Scale bar represents 200 μm. C) In the absence of UV irradiation the T7RNAp is inactive, whereas irradiation produces active Eg5 shRNA expression leading to a binuclear phenotype, shown by DAPI-stained nuclei. Scale

bars are 20 μm. D) Masked irradiation of live HEK293T cells expressing caged firefly luciferase to generated localized luminescence after UV irradiation. Scale bar represents 10 mm. Figures taken with permission from references [85] (A-C) and [111] (D).

4.5 OPTOCHEMICAL CONTROL OF CAGED GFP AND LUCIFERASE PROTEIN REPORTERS

Light-activated GFP mutants have been generated through incorporation of photocaged tyrosine into the GFP chromophore.[112,113] Incorporation of caged fluorotyrosine analogs enabled not only light-activation but also modulation of the electronic properties and thus the fluorescence wavelengths of GFP.[112] Since the caged tyrosine-GFP mutants still seem to form the mature fluorophores,[113] photoregulation efficiency may be the result of lower quantum yields due to quenching of the fluorophore singlet-state by the nitrobenzyl group.[114]

Recently, a caged firefly luciferase (fLuc) was reported through incorporation of caged lysine at the active site residue Lys529, thereby blocking binding of ATP and formation of the luminescent oxyluciferin (**Figure 8C**).[111] Time-resolved irradiation of HEK293T cells expressing a caged fLuc enabled the measurement of intracellular ATP dynamics, in contrast to more cumbersome earlier methods.[115] In contrast to a constitutively active fLuc, the caged fLuc does not disrupt cellular ATP concentrations as it is inactive until UV-induced decaging, allowing for easier quantification. These experiments demonstrated the ability for light-activated, spatiotemporally controlled gain-of-function studies within live cells in which reporter activity can be restored through light exposure.

5.0 CONCLUDING REMARKS

The optical activation of proteins in live cells and organisms provides high spatial and temporal control over biological processes, with a resolution that is theoretically only limited by light diffraction. A wide range of protein functions has been placed under optical control, enabling precise investigations of their molecular mechanisms in their respective biological environments. Unnatural amino acid mutagenesis with caged amino acids and other optogenetic approaches using natural, light-responsive proteins and protein domains, have been developed to achieve genetically encoded light-activated protein functions. These optical tools have been used to induce light-mediated control of ion channels, signal transduction, protein localization, transcription, translation, gene silencing, and other fundamental biological processes. The incorporation of caged amino acids into protein active sites in live cells has several distinct advantages: (1) the very small caging group (150-250 Da) induces no (or only minimal) perturbation on the protein structure, enabling precise photocontrol of enzymatic function; (2) the required location of the caging group in order to inhibit protein function can be predicted based on structural and mechanistic protein data; (3) irradiation cleaves the caging group and delivers the wild-type, non-modified, and active protein. The advantages of optogenetic approaches using natural light-responsive proteins are: (1) the protein biosynthetic machinery of cells and organisms does not need to be engineered to expand the genetic code with an additional, caged amino acid; (2) the optical regulation of protein function is reversible, adding

an additional layer of control; and (3) optical regulation of protein function has been demonstrated in systems as complex as freely moving mice. The evolution of both technologies is currently expanding to the control of biological processes in additional model organisms, as unnatural amino acid mutagenesis has been demonstrated in flies[116] and worms[117,118] and optogenetic methods have been applied in non-human primates.[119–121] Full control of protein function with the flip of a light switch holds tremendous promise to help understand mechanisms at the molecular, cellular, and systems level with unprecedented spatial and temporal resolution.

BIBLIOGRAPHY

(1) Gradinaru, V., Zhang, F., Ramakrishnan, C., Mattis, J., Prakash, R., Diester, I., Goshen, I., Thompson, K. R., and Deisseroth, K. (2010) Molecular and cellular approaches for diversifying and extending optogenetics. *Cell 141*, 154–65.

(2) Deisseroth, K. (2011) Optogenetics. *Nat. Methods 8*, 26–29.

(3) Ellis-Davies, G. C. R. (2007) Caged compounds: photorelease technology for control of cellular chemistry and physiology. *Nat. Methods 4*, 619–628.

(4) Brieke, C., Rohrbach, F., Gottschalk, A., Mayer, G., and Heckel, A. (2012) Light-controlled tools. *Angew. Chem. Int. Ed. Engl. 51*, 8446–8476.

(5) Deiters, A. (2010) Principles and applications of the photochemical control of cellular processes. *Chembiochem 11*, 47–53.

(6) Gardner, L., and Deiters, A. (2012) Light-controlled synthetic gene circuits. *Curr. Opin. Chem. Biol. 16*, 292–299.

(7) Lee, H.-M., Larson, D. R., and Lawrence, D. S. (2009) Illuminating the chemistry of life: design, synthesis, and applications of "caged" and related photoresponsive compounds. *ACS Chem. Biol. 4*, 409–427.

(8) Young, D. D., and Deiters, A. (2007) Photochemical control of biological processes. *Org. Biomol. Chem. 5*, 999–1005.

(9) Riggsbee, C. W., and Deiters, A. (2010) Recent advances in the photochemical control of protein function. *Trends Biotechnol. 28*, 468–75.

(10) Deiters, A. (2009) Light activation as a method of regulating and studying gene expression. *Curr. Opin. Chem. Biol. 13*, 678–86.

(11) Lin, W., Albanese, C., Pestell, R. G., and Lawrence, D. S. (2002) Spatially discrete, light-driven protein expression. *Chem. Biol. 9*, 1347–53.

(12) Link, K. H., Shi, Y., and Koh, J. T. (2005) Light activated recombination. *J. Am. Chem. Soc. 127*, 13088–13089.

(13) Young, D. D., and Deiters, A. (2007) Photochemical activation of protein expression in bacterial cells. *Angew. Chem. Int. Ed. Engl. 46*, 4290–2.

(14) Young, D. D., Garner, R. A., Yoder, J. a, and Deiters, A. (2009) Light-activation of gene function in mammalian cells via ribozymes. *Chem. Commun. (Camb).* 568–70.

(15) Connelly, C. M., Uprety, R., Hemphill, J., and Deiters, A. (2012) Spatiotemporal control of microRNA function using light-activated antagomirs. *Mol. Biosyst. 8*, 2987–93.

(16) Kennedy, M. J., Hughes, R. M., Peteya, L. a, Schwartz, J. W., Ehlers, M. D., and Tucker, C. L. (2010) Rapid blue-light-mediated induction of protein interactions in living cells. *Nat. Methods 7*, 973–5.

(17) Airan, R. D., Thompson, K. R., Fenno, L. E., Bernstein, H., and Deisseroth, K. (2009) Temporally precise in vivo control of intracellular signalling. *Nature 458*, 1025–9.

(18) Hughes, R. M., Vrana, J. D., and Tucker, C. L. (2012) Light-dependent , Dark-promoted Interaction between Arabidopsis Cryptochrome 1 and Phytochrome B Proteins. *J. Biol. Chem. 287*, 22165–22172.

(19) Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G., and Deisseroth, K. (2005) Millisecond-timescale, genetically targeted optical control of neural activity. *Nat. Neurosci. 8*, 1263–1268.

(20) Wu, Y., Frey, D., Lungu, O., Jaehrig, A., Schlichting, I., Kuhlman, B., and Hahn, K. (2009) A Genetically Encoded Photoactivatable Rac Controls the Motility of Living Cells. *Nature 461*, 104–110.

(21) Tabor, J. J., Levskaya, A., and Voigt, C. A. (2011) Multichromatic control of gene expression in Escherichia coli. *J. Mol. Biol. 405*, 315–324.

(22) Toettcher, J. E., Weiner, O. D., and Lim, W. A. (2013) Using optogenetics to interrogate the dynamic control of signal transmission by the ras/erk module. *Cell 155*, 1422–34.

(23) Miller, J. H., Ganem, D., Lu, P., and Schmitz, A. (1977) Genetic Studies of the lac Repressor. *J. Mol. Biol.*

(24) Bossi, L., and Ruth, J. R. (1980) The influence of codon context on genetic code translation. *Nature 286*, 123–127.

(25) Bruce, A. G., Wills, N., and Gesteland, R. F. (1982) Replacement of anticodon loop nucleotides to produce functional tRNAs : Amber suppressors derived from yeast-tRNAPhe. *Proc. Natl. Acad. Sci. U. S. A. 79*, 7127–7131.

(26) Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C., and Schultz, P. G. (1989) A general method for site-specific incorporation of unnatural amino acids into proteins. *Science (80-.). 244*, 182–188.

(27) Barltrop, J. A., Plant, P. J., and Schofield, P. (1966) Photosensitive Protecting Groups. *Chem. Commun. (Camb). 22*, 822–823.

(28) Patchornik, A., Amit, B., and Woodward, R. B. (1970) Photosensitive Protecting Groups. *J. Am. Chem. Soc. 92*, 6333–6335.

(29) Mendel, D., Ellman, J. A., and Schultz, P. G. (1991) Construction of a Light-Activated Protein by Unnatural Amino Acid Mutagenesis. *J. Am. Chem. Soc.* 2758–2760.

(30) Liu, D. R., and Schultz, P. G. (1999) Progress toward the evolution of an organism with an expanded genetic code. *Proc. Natl. Acad. Sci. U. S. A. 96*, 4780–5.

(31) Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) Expanding the genetic code of Escherichia coli. *Science 292*, 498–500.

(32) Chin, J. W. (2011) Reprogramming the genetic code. *EMBO J. 30*, 2312–24.

(33) Liu, C. C., and Schultz, P. G. (2010) Adding new chemistries to the genetic code. *Annu. Rev. Biochem. 79*, 413–444.

(34) Rockwell, N. C., Su, Y.-S., and Lagarias, J. C. (2006) Phytochrome structure and signaling mechanisms. *Annu. Rev. Plant Biol. 57*, 837–858.

(35) Shimizu-Sato, S., Huq, E., Tepperman, J. M., and Quail, P. H. (2002) A light-switchable gene promoter system. *Nat. Biotechnol. 20*, 1041–4.

(36) Levskaya, A., Weiner, O. D., Lim, W. A., and Voigt, C. A. (2009) Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature 461*, 997–1001.

(37) Inoue, T., Heo, W. Do, Grimley, J. S., Wandless, T. J., and Meyer, T. (2005) An inducible translocation strategy to rapidly activate and inhibit small GTPase signaling pathways. *Nat. Methods 2*, 415–418.

(38) Suh, B.-C., Inoue, T., Meyer, T., and Hille, B. (2006) Rapid chemically induced changes of PtdIns(4,5)P2 gate KCNQ ion channels. *Science 314*, 1454–7.

(39) Castellano, F., Montcourrier, P., Guillemot, J. C., Gouin, E., Machesky, L., Cossart, P., and Chavrier, P. (1999) Inducible recruitment of Cdc42 or WASP to a cell-surface receptor triggers actin polymerization and filopodium formation. *Curr. Biol. 9*, 351–60.

(40) Konermann, S., Brigham, M. D., Trevino, A. E., Hsu, P. D., Heidenreich, M., Cong, L., Platt, R. J., Scott, D. A., Church, G. M., and Zhang, F. (2013) Optical control of mammalian endogenous transcription and epigenetic states. *Nature 500*, 472–6.

(41) Bouly, J.-P., Schleicher, E., Dionisio-Sese, M., Vandenbussche, F., Van Der Straeten, D., Bakrim, N., Meier, S., Batschauer, A., Galland, P., Bittl, R., and Ahmad, M. (2007)

Cryptochrome blue light photoreceptors are activated through interconversion of flavin redox states. *J. Biol. Chem. 282*, 9383–91.

(42) Giovani, B., Byrdin, M., Ahmad, M., and Brettel, K. (2003) Light-induced electron transfer in a cryptochrome blue-light photoreceptor. *Nat. Struct. Biol. 10*, 489–90.

(43) Chaves, I., Pokorny, R., Byrdin, M., Hoang, N., Ritz, T., Brettel, K., Essen, L.-O., van der Horst, G. T. J., Batschauer, A., and Ahmad, M. (2011) The cryptochromes: blue light photoreceptors in plants and animals. *Annu. Rev. Plant Biol. 62*, 335–64.

(44) Sternberg, N., and Hamilton, D. (1981) Bacteriophage Pl Site-specific Recombination: I: Recombination between loxP sites. *J. Mol. Biol. 150*, 467–486.

(45) Swartz, T. E., Wenzel, P. J., Corchnoy, S. B., Briggs, W. R., and Bogomolni, R. A. (2002) Accelerated Publications Vibration Spectroscopy Reveals Light-Induced Chromophore and Protein Structural Changes in the LOV2 Domain of the Plant Blue-Light Receptor Phototropin 1. *Biochemistry 41*, 7183–7189.

(46) Strickland, D., Yao, X., Gawlak, G., Rosen, M. K., Gardner, K. H., and Sosnick, T. R. (2010) Rationally improving LOV domain-based photoswitches. *Nat. Methods 7*, 623–6.

(47) Lee, J., Natarajan, M., Nashine, V., Socolich, M., Vo, T., Russ, W. P., Benkovic, S., and Ranganathan, R. (2008) Surface Sites for Engineering Allosteric Control in Proteins. *Science 322*, 438–442.

(48) Polstein, L. R., Gersbach, C. A. (2012) Light-inducible spatiotemporal control of gene activation by customizable zinc finger transcription factors. *J. Am. Chem. Soc. 134*, 16480–3.

(49) Lungu, O. I., Hallett, R. A., Choi, E. J., Aiken, M. J., Hahn, K. M., and Kuhlman, B. (2012) Article Designing Photoswitchable Peptides Using the AsLOV2 Domain. *Chem. Biol. 19*, 507–517.

(50) Kravitz, A. V, Freeze, B. S., Parker, P. R. L., Kay, K., Thwin, M. T., Deisseroth, K., and Kreitzer, A. C. (2010) Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature 466*, 622–6.

(51) Banghart, M., Borges, K., Isacoff, E., Trauner, D., and Kramer, R. H. (2004) Light-activated ion channels for remote control of neuronal firing. *Nat. Neurosci. 7*, 1381–1386.

(52) Zemelman, B. V, Lee, G. A., Ng, M., and Miesenböck, G. (2002) Selective photostimulation of genetically chARGed neurons. *Neuron 33*, 15–22.

(53) Alexander, G. E., and Crutcher, M. D. (1990) Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci. 13*, 266–271.

(54) DeLong, M. R. (1990) Primate models of movement disorders of basal ganglia origin. *Trends Neurosci. 13*, 281–285.

(55) Gong, S., Doughty, M., Harbaugh, C. R., Cummins, A., Hatten, M. E., Heintz, N., and Gerfen, C. R. (2007) Targeting Cre recombinase to specific neuron populations with bacterial artificial chromosome constructs. *J. Neurosci. 27*, 9817–23.

(56) Bock, A., Forchhammer, K., Heider, J., Leinfelder, W., Sawers, G., Veprek, B., and Zinoni, F. (1991) MicroReview Selenocysteine : the 21st amino acid. *Mol. Microbiol. 5*, 515–520.

(57) Srinivasan, G., James, C. M., and Krzycki, J. a. (2002) Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science 296*, 1459–62.

(58) Wu, N., Deiters, A., Cropp, T. A., King, D., and Schultz, P. G. (2004) A genetically encoded photocaged amino acid. *J. Am. Chem. Soc. 126*, 14306–7.

(59) Deiters, A., and Schultz, P. G. (2005) In vivo incorporation of an alkyne into proteins in Escherichia coli. *Bioorg. Med. Chem. Lett. 15*, 1521–4.

(60) Deiters, A., Groff, D., Ryu, Y., Xie, J., and Schultz, P. G. (2006) A genetically encoded photocaged tyrosine. *Angew. Chem. Int. Ed. Engl. 45*, 2728–31.

(61) Lemke, E. A., Summerer, D., Geierstanger, B. H., Brittain, S. M., and Schultz, P. G. (2007) Control of protein phosphorylation with a genetically encoded photocaged amino acid. *Nat. Chem. Biol. 3*, 769–72.

(62) Chen, P. R., Groff, D., Guo, J., Ou, W., Cellitti, S., Geierstanger, B. H., and Schultz, P. G. (2009) A facile system for encoding unnatural amino acids in mammalian cells. *Angew. Chem. Int. Ed. Engl. 48*, 4052–5.

(63) Bossi, L., and Roth, J. R. (1981) Four-base codons ACCA, ACCU and ACCC are recognized by frameshift suppressor sufJ. *Cell 25*, 489–96.

(64) Anderson, J. C., Wu, N., Santoro, S. W., Lakshman, V., King, D. S., and Schultz, P. G. (2004) An expanded genetic code with a functional quadruplet codon. *Proc. Natl. Acad. Sci. U. S. A. 101*, 7566–71.

(65) Santoro, S. W., Wang, L., Herberich, B., King, D. S., and Schultz, P. G. (2002) An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. *Nat. Biotechnol. 20*, 1044–8.

(66) Fechter, P., Rudinger-Thirion, J., Tukalo, M., and Giegé, R. (2001) Major tyrosine identity determinants in Methanococcus jannaschii and Saccharomyces cerevisiae tRNA(Tyr) are conserved but expressed differently. *Eur. J. Biochem. 268*, 761–7.

(67) Lenhard, B., Orellana, O., Ibba, M., and Weygand-Durasević, I. (1999) tRNA recognition and evolution of determinants in seryl-tRNA synthesis. *Nucleic Acids Res. 27*, 721–9.

(68) Young, T. S., and Schultz, P. G. (2010) Beyond the canonical 20 amino acids: expanding the genetic lexicon. *J. Biol. Chem. 285*, 11039–44.

(69) Chin, J. W., Martin, A. B., King, D. S., Wang, L., and Schultz, P. G. (2002) Addition of a photocrosslinking amino acid to the genetic code of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A. 99*, 11020–11024.

(70) Tippmann, E. M., Liu, W., Summerer, D., Mack, A. V, and Schultz, P. G. (2007) A genetically encoded diazirine photocrosslinker in Escherichia coli. *Chembiochem 8*, 2210–4.

(71) Deiters, A., Cropp, T. A., Mukherji, M., Chin, J. W., Anderson, J. C., and Schultz, P. G. (2003) Adding amino acids with novel reactivity to the genetic code of Saccharomyces cerevisiae. *J. Am. Chem. Soc. 125*, 11782–3.

(72) Chin, J. W., Santoro, S. W., Martin, A. B., King, D. S., Wang, L., and Schultz, P. G. (2002) Addition of p-azido-L-phenylalanine to the genetic code of Escherichia coli. *J. Am. Chem. Soc. 124*, 9026–7.

(73) Alfonta, L., Zhang, Z., Uryu, S., Loo, J. A., and Schultz, P. G. (2003) Site-Specific Incorporation of a Redox-Active Amino Acid into Proteins 14662–14663.

(74) Neumann, H., Hazen, J. L., Weinstein, J., Mehl, R. A., and Chin, J. W. (2008) Genetically encoding protein oxidative damage. *J. Am. Chem. Soc. 130*, 4028–33.

(75) Chang, M. C. Y., Yee, C. S., Nocera, D. G., and Stubbe, J. (2004) Site-specific replacement of a conserved tyrosine in ribonucleotide reductase with an aniline amino acid: a mechanistic probe for a redox-active tyrosine. *J. Am. Chem. Soc. 126*, 16702–3.

(76) Mehl, R. a, Anderson, J. C., Santoro, S. W., Wang, L., Martin, A. B., King, D. S., Horn, D. M., and Schultz, P. G. (2003) Generation of a bacterium with a 21 amino acid genetic code. *J. Am. Chem. Soc. 125*, 935–9.

(77) Wang, J., Xie, J., and Schultz, P. G. (2006) A genetically encoded fluorescent amino acid. *J. Am. Chem. Soc. 128*, 8738–9.

(78) Summerer, D., Chen, S., Deiters, A., Chin, J. W., and Schultz, P. G. (2006) A genetically encoded fluorescent amino acid. *J. Am. Chem. Soc. 103*, 9785–9789.

(79) Nakayama, K., Endo, M., and Majima, T. (2004) Photochemical regulation of the activity of an endonuclease BamHI using an azobenzene moiety incorporated site-selectively into the dimer interface. *Chem. Commun. (Camb).* 2386–7.

(80) Bose, M., Groff, D., Xie, J., Brustad, E., and Schultz, P. G. (2006) The incorporation of a photoisomerizable amino acid into proteins in E. coli. *J. Am. Chem. Soc. 128*, 388–9.

(81) Curley, K., and Lawrence, D. S. (1999) Caged regulators of signaling pathways. *Pharmacol. Ther. 82*, 347–54.

(82) Lawrence, D. S. (2005) The preparation and in vivo applications of caged peptides and proteins. *Curr. Opin. Chem. Biol. 9*, 570–5.

(83) Pellois, J.-P., and Muir, T. W. (2005) A ligation and photorelease strategy for the temporal and spatial control of protein function in living cells. *Angew. Chem. Int. Ed. Engl. 44*, 5713–7.

(84) Hahn, M. E., and Muir, T. W. (2004) Photocontrol of Smad2, a multiphosphorylated cell-signaling protein, through caging of activating phosphoserines. *Angew. Chem. Int. Ed. Engl. 43*, 5800–3.

(85) Hemphill, J., Chou, C., Chin, J. W., and Deiters, A. (2013) Genetically encoded light-activated transcription for spatiotemporal control of gene expression and gene silencing in mammalian cells. *J. Am. Chem. Soc. 135*, 13433–9.

(86) Gautier, A., Nguyen, D. P., Lusic, H., An, W., Deiters, A., and Chin, J. W. (2010) Genetically encoded photocontrol of protein localization in mammalian cells. *J. Am. Chem. Soc. 132*, 4086–8.

(87) Gautier, A., Deiters, A., and Chin, J. W. (2011) Light-Activated Kinases Enable Temporal Dissection of Signaling Networks in Living Cells. *J. Am. Chem. Soc. 133*, 2124–2127.

(88) Nguyen, D. P., Mahesh, M., Elsasser, S., Hancock, S. M., Uttampinant, C., and Chin, J. W. (2014) Genetic Encoding of Photocaged Cysteine Allows Photoactivation of TEV Protease in Live Mammalian Cells. *J. Am. Chem. Soc. Article AS*.

(89) Arbely, E., Torres-Kolbus, J., Deiters, A., and Chin, J. W. (2012) Photocontrol of tyrosine phosphorylation in mammalian cells via genetic encoding of photocaged tyrosine. *J. Am. Chem. Soc. 134*, 11912–5.

(90) Chou, C., Young, D. D., and Deiters, A. (2010) Photocaged t7 RNA polymerase for the light activation of transcription and gene function in pro- and eukaryotic cells. *Chembiochem 11*, 972–7.

(91) Chou, C., Young, D. D., and Deiters, A. (2009) A light-activated DNA polymerase. *Angew. Chem. Int. Ed. Engl. 48*, 5950–3.

(92) Miller, J. C., Silverman, S. K., England, P. M., Dougherty, D. A., and Lester, H. A. (1998) Flash decaging of tyrosine sidechains in an ion channel. *Neuron 20*, 619–24.

(93) Cirit, M., Wang, C.-C., and Haugh, J. M. (2010) Systematic quantification of negative feedback mechanisms in the extracellular signal-regulated kinase (ERK) signaling network. *J. Biol. Chem. 285*, 36736–44.

(94) Levy, D. E., and Darnell, J. E. (2002) Stats: transcriptional control and biological impact. *Nat. Rev. Mol. Cell Biol. 3*, 651–62.

(95) Mohr, A., Chatain, N., Domoszlai, T., Rinis, N., Sommerauer, M., Vogt, M., and Müller-Newen, G. (2012) Dynamics and non-canonical aspects of JAK/STAT signalling. *Eur. J. Cell Biol. 91*, 524–32.

(96) Kaffman, A., Rank, N. M., O'Neill, E. M., Huang, L. S., and O'Shea, E. K. (1998) The receptor Msn5 exports the phosphorylated transcription factor Pho4 out of the nucleus. *Nature 396*, 482–486.

(97) Lange, A., Mills, R. E., Lange, C. J., Stewart, M., Devine, S. E., and Corbett, A. H. (2007) Classical nuclear localization signals: definition, function, and interaction with importin alpha. *J. Biol. Chem. 282*, 5101–5105.

(98) Engelke, H., Chou, C., Uprety, R., Jess, P., and Deiters, A. (2014) Control of Protein Function through Optochemical Translocation. *ACS Synth. Biol. In Press*.

(99) Kim, Y. G., Cha, J., and Chandrasegaran, S. (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. U. S. A. 93*, 1156–1160.

(100) Urnov, F. D., Miller, J. C., Lee, Y.-L., Beausejour, C. M., Rock, J. M., Augustus, S., Jamieson, A. C., Porteus, M. H., Gregory, P. D., and Holmes, M. C. (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature 435*, 646–651.

(101) Miller, J. C., Tan, S., Qiao, G., Barlow, K. A., Wang, J., Xia, D. F., Meng, X., Paschon, D. E., Leung, E., Hinkley, S. J., Dulay, G. P., Hua, K. L., Ankoudinova, I., Cost, G. J., Urnov, F. D., Zhang, H. S., Holmes, M. C., Zhang, L., Gregory, P. D., and Rebar, E. J. (2011) A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol. 29*, 143–148.

(102) Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., and Church, G. M. (2013) RNA-guided human genome engineering via Cas9. *Science 339*, 823–6.

(103) Edwards, W. F., Young, D. D., and Deiters, A. (2009) Light-activated Cre recombinase as a tool for the spatial and temporal control of gene function in mammalian cells. *ACS Chem. Biol. 4*, 441–5.

(104) Chou, C., and Deiters, A. (2011) Light-activated gene editing with a photocaged zinc-finger nuclease. *Angew. Chem. Int. Ed. Engl. 50*, 6839–42.

(105) Radecke, S., Radecke, F., Cathomen, T., and Schwarz, K. (2010) Zinc-finger nuclease-induced gene repair with oligodeoxynucleotides: wanted and unwanted target locus modifications. *Mol. Ther. 18*, 743–753.

(106) Bohne, J., and Cathomen, T. (2008) Genotoxicity in gene therapy: an account of vector integration and designer nucleases. *Curr. Opin. Mol. Ther. 10*, 214–223.

(107) Steitz, T. A. (2006) Visualizing polynucleotide polymerase machines at work. *EMBO J. 25*, 3458–68.

(108) Cramer, P. (2002) Common structural features of nucleic acid polymerases. *Bioessays 24*, 724–9.

(109) Temiakov, D., Patlan, V., Anikin, M., McAllister, W. T., Yokoyama, S., and Vassylyev, D. G. (2004) Structural basis for substrate selection by t7 RNA polymerase. *Cell 116*, 381–91.

(110) Achenbach, J. C., Chiuman, W., Cruz, R. P. G., and Li, Y. (2004) DNAzymes: from creation in vitro to application in vivo. *Curr. Pharm. Biotechnol. 5*, 321–336.

(111) Zhao, J., Lin, S., Huang, Y., Zhao, J., and Chen, P. R. (2013) Mechanism-Based Design of a Photoactivatable Firefly Luciferase. *J. Am. Chem. Soc. 135*, 7410–7413.

(112) Wilkins, B. J., Marionni, S., Young, D. D., Liu, J., Wang, Y., Di Salvo, M. L., Deiters, A., and Cropp, T. A. (2010) Site-specific incorporation of fluorotyrosines into proteins in Escherichia coli by photochemical disguise. *Biochemistry 49*, 1557–1559.

(113) Groff, D., Wang, F., Jockusch, S., Turro, N. J., and Schultz, P. G. (2010) A new strategy to photoactivate green fluorescent protein. *Angew. Chem. Int. Ed. Engl. 49*, 7677–9.

(114) Kobayashi, T., Urano, Y., Kamiya, M., Ueno, T., Kojima, H., and Nagano, T. (2007) Highly activatable and rapidly releasable caged fluorescein derivatives. *J. Am. Chem. Soc. 129*, 6696–7.

(115) Gajewski, C. D., Yang, L., Schon, E. A., and Manfredi, G. (2003) New insights into the bioenergetics of mitochondrial disorders using intracellular ATP reporters. *Mol. Biol. Cell 14*, 3628–3635.

(116) Bianco, A., Townsley, F. M., Greiss, S., Lang, K., and Chin, J. W. (2012) Expanding the genetic code of Drosophila melanogaster. *Nat. Chem. Biol. 8*, 748–50.

(117) Greiss, S., and Chin, J. W. (2011) Expanding the genetic code of an animal. *J. Am. Chem. Soc. 133*, 14196–9.

(118) Parrish, A. R., She, X., Xiang, Z., Coin, I., Shen, Z., Briggs, S. P., Dillin, A., and Wang, L. (2012) Expanding the genetic code of Caenorhabditis elegans using bacterial aminoacyl-tRNA synthetase/tRNA pairs. *ACS Chem. Biol. 7*, 1292–302.

(119) Gerits, A., Farivar, R., Rosen, B. R., Wald, L. L., Boyden, E. S., and Vanduffel, W. (2012) Optogenetically induced behavioral and functional network changes in primates. *Curr. Biol. 22*, 1722–1726.

(120) Cavanaugh, J., Monosov, I. E., McAlonan, K., Berman, R., Smith, M. K., Cao, V., Wang, K. H., Boyden, E. S., and Wurtz, R. H. (2012) Optogenetic Inactivation Modifies Monkey Visuomotor Behavior. *Neuron 76*, 901–907.

(121) Diester, I., Kaufman, M. T., Mogri, M., Pashaie, R., Goo, W., Yizhar, O., Ramakrishnan, C., Deisseroth, K., and Shenoy, K. V. (2011) An optogenetic toolbox designed for primates. *Nat. Neurosci. 14*, 387–397.