ABSTRACT

UPRETY, RAJENDRA. Synthesis of Light-Activated Nucleotides and Unnatural Amino Acids for Biological Applications. (Under the direction of Alexander Deiters).

Nature controls biological processes, such as gene regulation and protein function, with high spatio-temporal resolution. Light, one of the most spatially and least invasive external control elements, can be used for the photo-regulation of biomolecules in order to understand cellular processes in a spatially and temporally controlled manner. Installation of photo-cleavable moieties onto biomolecules through small organic molecule synthesis is a unique and efficient technique to access light-responsive biomolecules. This dissertation explains on the syntheses and characterization of a number of light-responsive biomolecules including caged phosphoramidites, nucleotides, and amino acids.

This research work presents syntheses of new light-activated phosphoramidites and oligonucleotides to study the photochemical control of DNA and RNA functions. Applications of these caged oligonucleotides range from basic biological studies of gene expression to potential precursors for gene therapeutics. In addition, we present efficient syntheses of a variety of new unnatural caged amino acids based on lysine and tyrosine. Our newly synthesized analogues include a variety of lysine derivatives with a distinct functional moiety at the $^{\varepsilon}N$ position, such as a photocrosslinking group, one-and two-photon caging groups, fluorescent probes, a dithiolane unit, or a spin labeled probe. Selective photo-decaging of these amino acid residues in proteins upon UV irradiation to release caging groups and restore biological activities is a powerful technology to investigate protein structure, dynamics, localization, and biomolecular interactions in a real-time fashion.

Synthesis of Light-Activated Nucleotides and Unnatural Amino Acids for Biological Applications

by
Rajendra Uprety

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Chemistry

Raleigh, North Carolina

2013

APPROVED BY:

_____	_____
Alexander Deiters Committee Chair	Daniel L.Comins
_____	_____
Jonathan S. Lindsey	Christian Melander

DEDICATION

To my father Rabindra Kumar Uprety and my mother Lila Devi Uprety.

BIOGRAPHY

The author, Rajendra Uprety, was born and raised in Terathum, Nepal. After graduating from high school, Rajendra moved to capital city Kathmandu and continued his higher education at Tribhuvan University, Nepal where he accomplished his Bachelors' Degree in Science (B.Sc.) and then Masters' Degree in Science (M.Sc.) in Chemistry. In 2001, he began to work at Cosmos College of Management and Technology, Lalitpur, Nepal as an instructor for Chemistry. After six years of teaching, he arrived to the United States to pursue his further study and joined Western Carolina University, NC, USA in August 2006. At WCU, his thesis work was on the synthesis of interlocking molecules under the supervision of Professor William R. Kwochka. Rajendra graduated from WCU with Master's of Science (MS) Degree in Chemistry. In August 2008, Rajendra enrolled at North Carolina State University (NCSU) for PhD program in Chemistry. At NCSU, he has been working for his doctoral research on organic synthesis of the light-responsive nucleotides and various unnatural amino acids under the supervision of Professor Alexander Deiters. Rajendra Uprety is a member of the American Chemical Society, ACS Division of Organic Chemistry, Phi Lambda Upsilon (Honorary Chemical Society), NCSU, and Life member of Nepal Chemical Society, Nepal.

ACKNOWLEDGMENTS

First and foremost, I would like to acknowledge my advisor Professor Alexander Deiters for his guidance and financial support during my graduate study. His time and constructive suggestions during this research work, which also accounts the foundation for my future profession, have been very much appreciated. My deepest gratitude goes to the members of Dissertation Advisory Committee, Professor Daniel L. Comins, Professor Jonathan S. Lindsey, Professor Christian Melander, and Professor Donald W. Brenner for their knowledge and assistance in my doctoral studies. It is my pleasure to be educated from professors Dr. Daniel L. Comins, and Dr. Christian Melander for graduate level organic chemistry courses whose contribution to my doctoral education would always be appreciated. My grateful thanks are also extended to our collaborators Dr. Chin at the MRC Laboratory of Molecular Biology, UK; Dr. Cropp at Virginia Commonwealth University, and Dr. Sombers at NCSU.

My special thanks go to Laura Gardner and Alex Prokup for editing this dissertation at my convenience. I am also grateful to all members in the Deiters Laboratory; Colleen, Jeane, Jessica, Qingyang, Sander, James, Meryl, Kalyn, Jihe, Matt, and Ji. I would also like to take this opportunity to thank previous group members in the Deiters lab; Dr. Chou, Dr. Lockney, Dr. Lusic, Dr. McIver, Dr. Zou, Dr. Young, Dr. Georgianna, Rikki, Jie, and Chad for their help at various levels. I would like to appreciate Dr. Enamorado, Dr. Tsukanov, Dr. Rogers, and Dr. DeSousa for their friendship. I appreciate the kind cooperation from the members in the Melander group, Ghiladi group and Lindsey group. I would like to thank all of the staff and faculties that I worked with during my graduate study at NCSU. I am very

fortunate to have many friends who made my life enjoyable, memorable, and meaningful; and I am very much thankful to all of them.

I am forever indebted to my parents, other beloved family members, and generous relatives for their all kinds of supports and endless love. Finally to my beloved wife, Januka, nothing brings more happiness than your love, inspiration, and support in my life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SCHEMES

LIST OF ABBREVIATIONS

AIBN	Azobisisobutyronitrile
ATP	Adenosine triphosphate
AMP	Adenosine monophosphate
Boc	t-Butoxycarbonyl
$CHCl_3$	Chloroform
$CDCl_3$	Deuterated chloroform
^{13}C NMR	Carbon nuclear magnetic resonance
CuAAC	Copper(I) catalyzed azide-alkyne [3+2] cycloaddition
d	Doublet
DCM	Dichloromethane
dd	Doublet of doublet
D_2O	Deuterated water
DEAD	Diisopropyl azodicarboxylate
DIBAL	Diisobutylaluminium hydride
DIPEA	Diisopropylethylamine
DMSO	Dimethyl sulfoxide
DMSO-d_6	Deuterated dimethyl sulfoxide
DMF	Dimethyl formamide
DMT	Dimethoxytrityl
DSC	Disuccinimidyl carbonate
ESI	Electrospray ionization

EtOAc	Ethyl acetate
EtOH	Ethanol
Fmoc	Fluorenylmethoxycarbonyl
GFP	Green fluorescent protein
HBTU	*O*-(Benzotriazol-1-yl)-*N*,*N*,*N'*,*N'*-tetramethyluronium hexafluorophosphate
^{1}H NMR	Proton nuclear magnetic resonance
HPLC	High performance liquid chromatography
h	Hour
HRMS	High resolution mass spectrometry
IR	Infrared
J	Coupling constant
LRMS	Low resolution mass spectrometry
M	Molar
m	Multiplet
MeOH	Methanol
MHz	Megahertz
NBS	*N*-Bromosuccinamide
NPOM	Nitropiperonyloxymethyl
MOM	Methoxymethyl
mmol	Millimole
NB	Nitrobenzyl
PPi	Pyrophosphate inorganic

ppm	Parts per million
rt	Room temperature
s	Singlet
sfGFP	superfold green fluorescent protein
TBDMS	t-Butyldimethylsilyl
TES	Triethylsilane
TEA	Triethylamine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TMS	Trimethylsilyl
tRNA	Transfer ribonucleic acid
TPS	2,4,6-Triisopropylphenylsulfonyl
UAA	Unnatural amino acid
UV	Ultra violet

CHAPTER 1: UNNATURAL NUCLEOTIDES IN BIOLOGY

1. Introduction

1.1. Central dogma of molecular biology

A sequence of four nucleobases adenine (A), guanine (G), cytosine (C) and thymine/uracil (T/U) in deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) preserves the genetic information of an organism.[1-3] The specific pairing (A/T, or G/C) signatures among these four base pairs are the codes of genetic information that play a central role in life.[1, 4] The transfer of the genetic information from parent to next generation of cells or organisms occurs through the transformation of DNA to RNA to protein, known as the "Central Dogma" of molecular biology.[5, 6] The genetic integrity can only be secured when DNA polymerase, an enzyme that catalyzes DNA synthesis, replicates DNA with high fidelity and the DNA is stable under physiological conditions.[6, 7] The error rate of DNA polymerase is less than 10^{-5}.[8]

The central dogma of molecular biology proclaims that DNA is the original source of genetic information. This double-stranded molecule is transcribed into single-stranded RNA in the process of transcription as illustrated in Figure 1. An RNA molecule may be replicated or reverse-transcribed into DNA under specific circumstances.[9, 10] Following transcription, RNA then serves as the template molecule for protein synthesis during the process of translation. Both the transcription and translation processes are tightly regulated in spatiotemporal fashion.[6, 11]

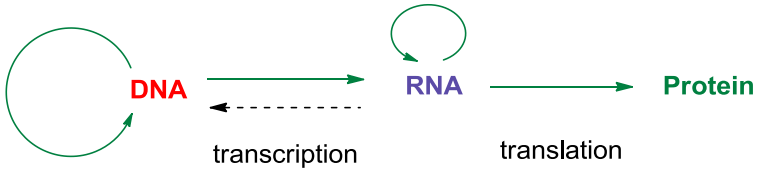


Figure 1. The central dogma of molecular biology.

Over the last sixty years after the discovery of molecular structure of DNA,[2] the study of biological processes has achieved remarkable progress in the fundamental understanding of gene regulation and protein function.[12, 13] The discoveries on the control of flow of genetic information,[5] recombinant DNA technology,[14-16] the development of the polymerase chain reaction (PCR),[17-19] DNA sequencing technology,[20, 21] DNA oligonucleotide synthesis,[22] non-coding RNAs,[23-25] G-protein-coupled receptors,[26] and signaling pathways,[27, 28] are landmark achievements in the history of biology.

Investigations on cellular function revealed that many components must work in unison to execute a specific function in a spatiotemporal fashion such as transcription or translation.[29, 30] The transcription of a particular gene requires *cis* and *trans* regulatory proteins in addition to polymerase enzymes.[31] To manipulate gene function in the laboratory, various integrated approaches using functionalized organic compounds as molecular probes have been reported.[32, 33 34-38]

1.2.Oligonucleotide tools for studying gene function

Progress on high-throughput gene synthesis technology has made it possible to integrate various probes into the gene of interest.[22, 39, 40] Installation of new probes into a

target gene alters the genetic makeup, as well as its biochemical properties. Incorporation of unnatural nucleosides modulates selectivity and efficiency of a gene by altering either hydrogen bonding topologies, hydrophobic interfaces, charge linkages or shape complementarities.[41-45] A common method to regulate gene activity is through the use of antisense oligonucleotides that bind messenger RNA (mRNA) targets by forming a stable duplex and cause target degradation or inhibition of the translation process.[46]

Antisense oligonucleotides have been classified into three categories:[47, 48] 1) oligonucleotides with a modified phosphate backbone (first generation) **1–2**, 2) oligonucleotides with alkyl substitution at 2′O in the ribose ring (second generation) **3**, and 3) oligonucleotides where the ribose ring has been replaced with a new chemical moiety such as morpholino unit or backbone modification with others functionalities (third generation) **4**–**8** (Figure 2). Each of these specific modifications provides unique advantages to the oligonucleotides despite limitations in solubility or toxicity.[49] For instance, a phosphorothionate backbone increases the stability of the molecule as well as delivery into the cell.[50, 51] Methyl phosphonate or morpholino phosphoramidate containing oligonucleotides provide a neutral backbone with an enhanced solubility and a lower toxicity.[52, 53] Furthermore, introduction of locked nucleic acids improves the stability through the change in conformation,[54] and addition of the guanidinium unit introduces a polycationic backbone, providing tight binding with DNAs through ionic interactions.[44]

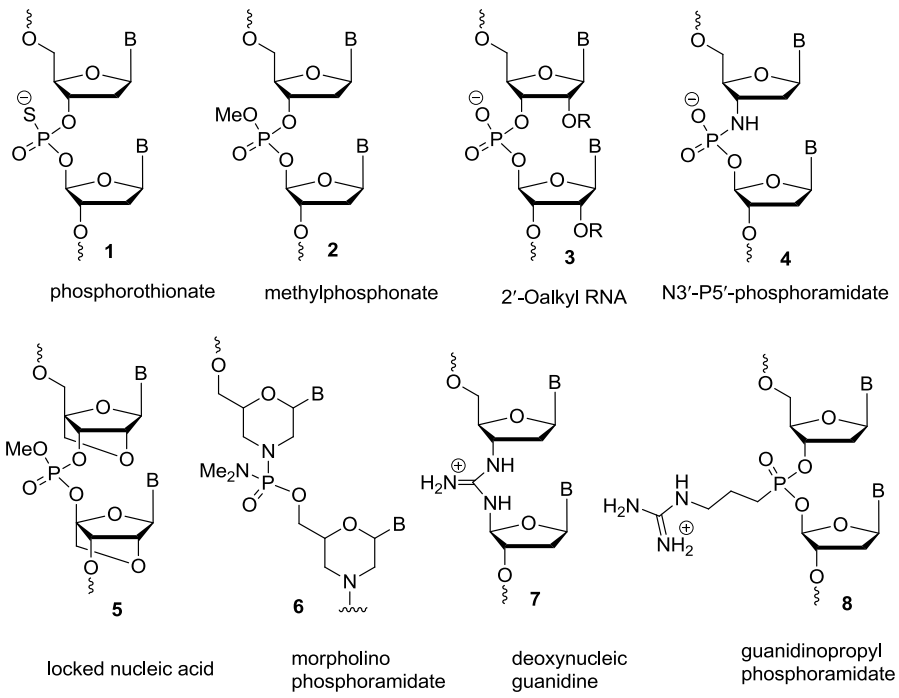


Figure 2. Unnatural oligonucleotides: a) first generation **1**–**2**, b) second generation **3**, and c) third generation **4**–**8**. B = nucleobase, e.g., G, C, T, U, or A.

Particularly, heteroatoms in either the sugar or base moiety of an oligonucleoside significantly aided in the understanding of altered functions of newly generated oligonucleosides.[55] This method provides an outstanding example of structure modulation for tunable function. Knowledge on further structural modifications for desired functions can always be borrowed from a vast repository of chemical inventories including modified nucleosides,[12, 56] fluorescent dyes,[57] modified amino acids,[58] and natural products[59] to name a few. In addition to the research on DNA nucleosides, the discovery of new RNA nucleosides helps to explore the structure-property relationships of gene regulatory tools for specific gene silencing or activation.[56, 60]

1.3. Caged nucleotides for studying cellular functions

In-depth knowledge on gene regulation and protein function is essential to understand and address complex biological processes at the cellular level.[61] Biological processes such as gene regulation and protein function take place with high spatiotemporal resolution in nature.[62] There are many aspects of cellular dynamics that may not be understood outside a cellular environment because of their context dependent nature.[63]

A number of oligonucleotide-based tools, with the advancement of nucleic acid synthesis and automated DNA synthesis technology, have been developed for therapeutic and diagnostic applications.[64] Such oligonucleotides have been used comprehensively for both *in vitro* and *in vivo* applications.[61, 65-68] Synthetic oligonucleotides have been a subject of investigation for a variety of purposes; such as catalysts (DNAzymes and ribozymes),[69] small molecule sensors (aptamers),[70] gene regulatory components (antisense DNA, siRNA, and miRNA),[68, 71, 72] and DNA computation platforms.[73, 74] Recently developed unnatural nucleosides with light-cleavable handles (photocaged oligonucleotides) not only have expanded oligonucleotides applications, but also has shown a great promise for achieving light-mediated gene modulation with high precision.[45, 62, 75] These light-cleavable oligonucleotides include caged ATP **9**,[76] caged phosphates **10**–**11**,[77-80] caged sugars **12**–**13**,[81, 82] and caged bases **14**–**15** (Figure 3).[83, 84]

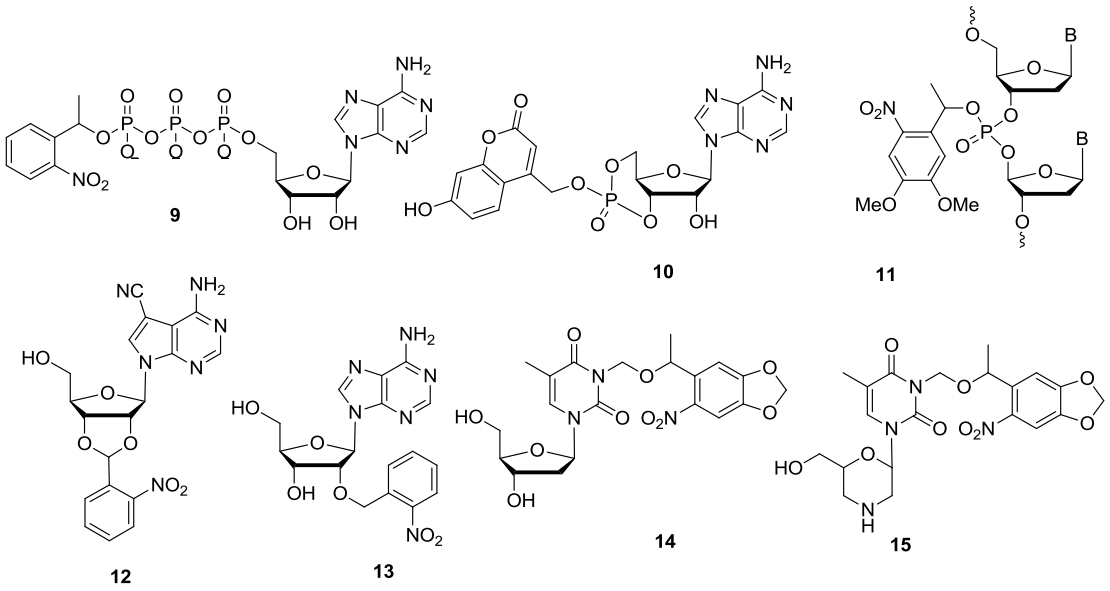


Figure 3. Caged nucleosides and nucleotides include caged ATP **9**, caged phosphates **10**–**11**, caged toyocamycin **12**, caged adenosine **13**, caged thymidine **14**, and caged morpholino **15**.

1.4. Molecular caging and decaging

Light can act as a unique modulator and has several advantages over traditional effectors for controlling biological processes (*vide infra*). Photomodulation of cellular chemistry with light is a powerful means for investigating cellular dynamics[62] through both in vitro and in vivo studies.[62, 85-88] Still, the design and synthesis of a light-cleavable (caged) unnatural nucleoside with good photo-physical properties and the ability to incorporate into biological systems is challenging.

In 1978, Hoffmann et al. reported the first synthesis of the caged ATP molecule **9**.[76] Since then, the application of photo-caged compounds for the study of cellular processes has greatly advanced. Over the past 30 years, many caged nucleotides, neurotransmitters,

secondary metabolites, metal ions, and proteins have been reported.[89, 90] Molecular caging and the subsequent activation of the caged compound through light irradiation (decaging) involves two major steps: 1) installation of a caging group to render the substrate biologically inactive and 2) removal of the caging group (activation) using light as a benign trigger. In general, upon brief irradiation with UV light (>365 nm), the caged compound undergoes a photochemical reaction to release the caging group thereby restoring the activity of the substrate (Figure 4).[62, 91]

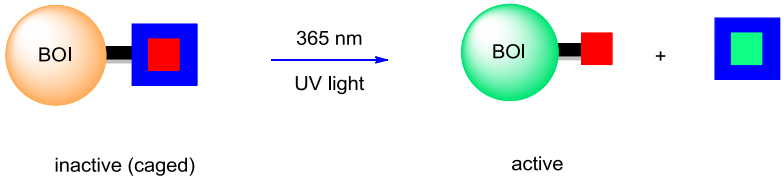


Figure 4. Activation of a caged biomolecule of interest (BOI) with light (365 nm).

Light can be used as a photomodulator for studying a variety of biological processes due to several distinct advantages:[62, 75] 1) Light is an ideal external trigger signal that acts as an orthogonal tool in many circumstances except for photosynthetic and photoreceptor cells.[85] 2) Light can easily be controlled in amplitude allowing the desired biological effects to be tuned. 3) Light can be used as a non-invasive tool causing only a minimal secondary perturbation of cellular processes. 4) Light can be controlled in both a spatial and a temporal fashion.[65]

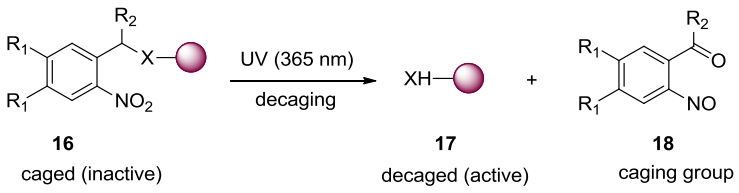
For a successful caging application, a caging group should satisfy certain criteria:[62, 75, 90] 1) It should possess selective reactivity and compatibility with the molecule of interest in high yield, 2) the installed caging group should render the substrate biologically inactive, 3) the caged substrate should be stable under physiological conditions, 4) it should have high quantum yield (*vide infra*) and a short decaging half-life ($t_{1/2}$), 5) upon photolysis of the caged substrate, the biomolecule should be able to completely restore its native form, and 6) only benign byproducts should be released upon photolysis.

The quantum yield (ϕ) gives a measure for efficiency of a photochemical process and is defined as the ratio of the number of excited molecules that undergo photolysis to the total number of photons absorbed.[92] The rate of photolysis can be estimated by measuring the half-life ($t_{1/2}$) of the chemical reaction, which is the time required for half of the amount of the starting material to be consumed.

$$Quantum\ Yield\ (\phi)\ for\ photolysis = \frac{Number\ of\ excited\ molecules\ that\ undergo\ photolysis}{Total\ number\ of\ photons\ absorbed}$$

A wide variety of caging groups and their syntheses are reported in the literature.[90, 93] However, the most popular caging groups are analogues of the *ortho*-nitrobenzyl (*o*-NB) moiety (Scheme 1).[87, 94] A few appealing reasons for the wide acceptance of *o*-NB compounds are: 1) ease of synthesizing its derivatives in high yields, 2) efficient photolysis of caged substrates, 3) compatibility with a variety of functional groups including hydroxyl, amino, thio, and carboxy groups, and 4) tunable photochemical properties of the *o*-NB group

by electronic modulation through substituent effects on the aromatic ring.[95, 96] For example, introduction of an electron–donating group (e.g., -OCH_3) in **16** results in a red shift of the absorption maximum allowing for the use of longer wavelength light, which is biologically less photodamaging.[65]



Scheme 1. Light activation of an *o*-nitrobenzyl caged substrate using UV irradiation (365 nm). Adapted with permission from *ChemBioChem* **2010**, *11*, 47-53.

Taking advantages of fine tunability of the *o*-NB caging moiety, a variety *o*-NB derivatives have been incorporated into nucleosides (Figure 5) for specific proposes in designing caged oligonucleotides (**19**–**29**).[95, 97, 98] A number of caging groups used in the past include modifications on a few core frame works such as *o*-nitrobenzyl, *o*-nitrodibenzofuryl, and 7-hydroxy- or aminocoumaryl core structure; albeit several other types of caging groups have been reported suitable for various biological applications.[77, 97-101] Overall outstanding photophysical properties, robustness for DNA synthesis conditions, minimal byproduct toxicity upon decaging, ease of synthesis on a multi-gram scale, and structural simplicity for their characterization are rationales behind the popularity of using derivatives based on the aforementioned core structures.[85, 102, 103]

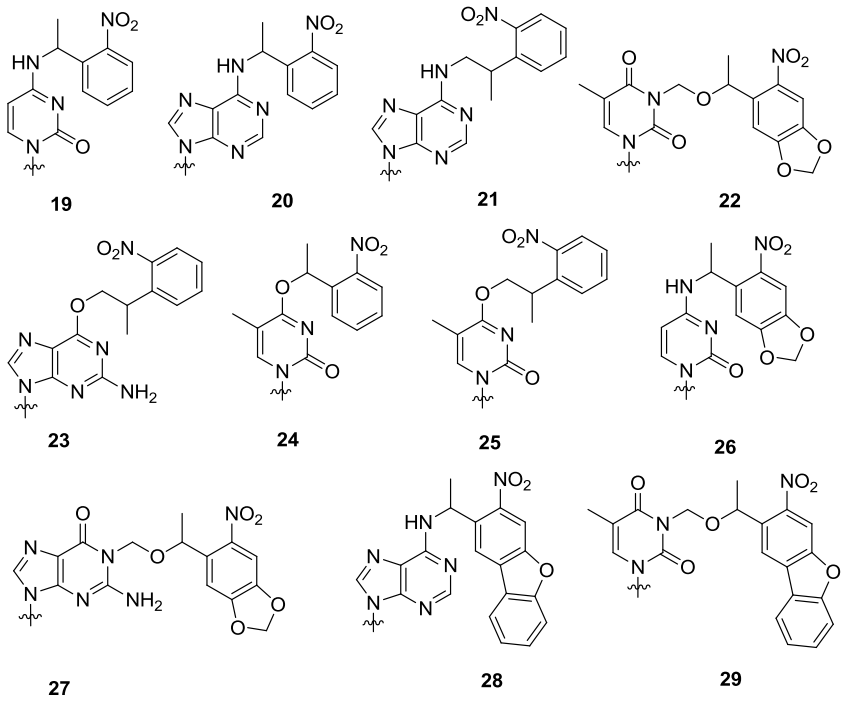
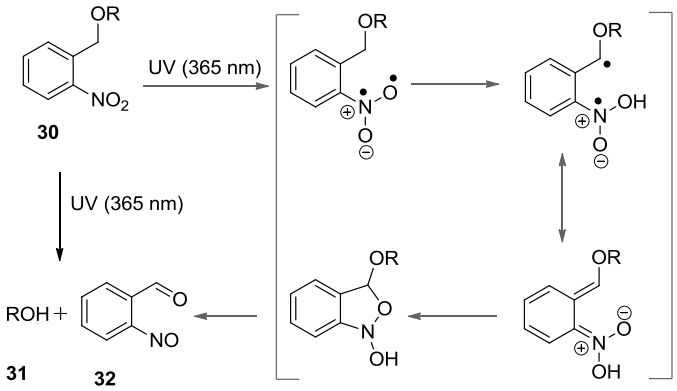


Figure 5. Caged nucleosides for DNA or RNA synthesis: *o*-NB caged **19**–**27,** and *o*-nitrodibenzofuryl caged **28**–**29** nucleotide bases.

1.4.1. Single-photon excitation and decaging mechanism

In 1970, Patchornik et al. reported that *o*-nitrobenzyl type caged substrates undergo decaging under UV (320 nm) irradiation as a result of a single photon excitation phenomenon. Single-photon excitation is a linear process in which the probability of excitation is linearly related to the power of the incident irradiation.[104] The caged substrate **30** decaging process is based on a Norrish type II mechanism (Scheme 2) resulting in an active substrate **31** and a nitroso-benzocarbonyl byproduct **32**.[105] The rate of photolysis depends upon the nature of the caged substrate, pH, and the composition of the buffer and/or

dielectric constant of the medium.[106] In addition, depending on the caging group, the use of a protic solvent such as water favors the photolysis compared to that of an aprotic solvent such as an acetonitrile. The caging group also poses a few limitations despite its use for various purposes. The *o*-nitrobenzyl caging group requires UV irradiation (~360 nm) which may result in photo-damaging of a biological system to some extent. In addition, the photochemical degradation of the *o*-nitrobenzyl caging group caused the formation of an amine reactive *o*-nitroso carbonyl compound.



Scheme 2. Mechanism of photolysis of an *o*-NB type substrate.

Generally, absorption maxima of nitrobenzyl caging groups vary between 254–375 nm resulting in the release of substrate at the order of microseconds to milliseconds.[90, 107, 108] Absorption maximum is a value of a specific wavelength of light required to achieve an optimum absorption for a particular compound. Photolysis of a nitrobenzyl caging group depends upon the nature of substituents on the caging group **33**–**38**, including the substrate, and the UV irradiation.[95] For instance, a variety of *o*-NB caging groups possess significantly

lower quantum yields upon irradiation with 365 nm UV light compared to that of 254 nm UV irradiation (Figure 6).[97] It was also observed that the quantum yield increases with a substitution at the benzylic position and decreases upon substitutions at the aromatic ring.[109]

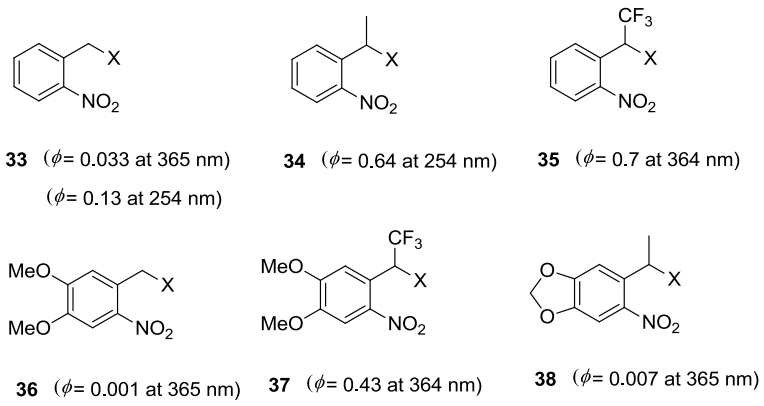


Figure 6. Substitution effect on quantum yield (ϕ) of *o*-NB caging groups (X = substrate).

1.4.2. Two-photon excitation and decaging mechanism

In 1930s German physicist Maria Goeppert-Mayer proposed that a molecule can simultaneously absorb two or more photons (within 1fs) of lower energy (non-linear excitation) such that the sum of the absorbed energy is equal to the energy required for the linear excitation (one-photon) of the molecule as depicted in Figure 7.[110] Thirty years after the discovery, two-photon excitation was experimentally observed by Kaiser and Garrett in 1961.[111] In this technique, it is possible to create an event of excitation at the very focal point of radiation by the use of a high power laser irradiation technique that allows control over the three-dimensional spatial selectivity of excitation to a volume less than 1 fL.[112, 113]

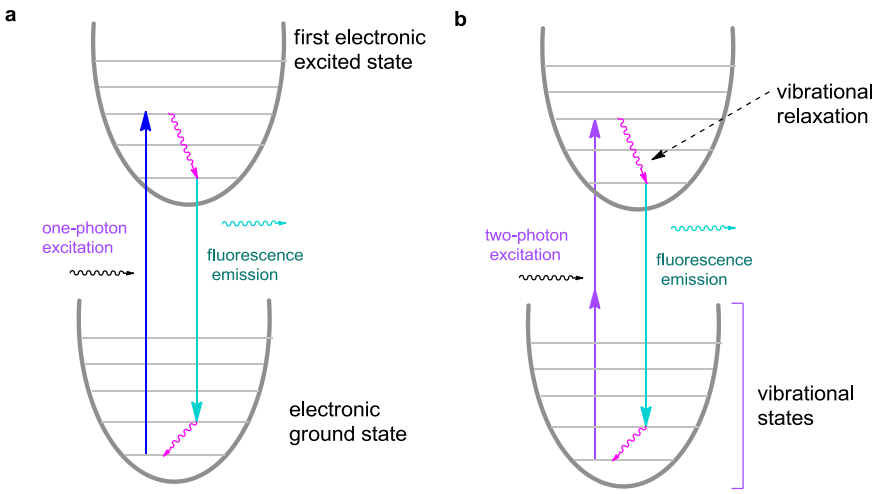
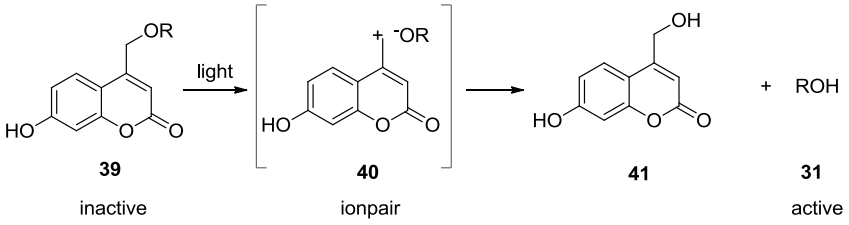


Figure 7. An illustration of a molecular excitation via Jablonski diagram: **a**) a linear excitation (one-photon), and **b**) a non-linear excitation (two-photon).

A two-photon excitation technique has several advantages for the study of cellular or biological processes, including higher tissue penetration up to a few hundred micrometers, less photobleaching, photodamaging, and increase in fluorescence collection efficiency through the collection of a portion of the scattered fluorescent light from the sample [114] [104, 115, 116] Furuta et al. reported a coumarin derivative as the first two-photon caging group.[117] They discovered that the 6-bromo-7-hydroxycoumarinmethyl (Bhc) **44** caging group readily undergoes two-photon photolysis with a δ_u value of 0.72 GM at 740 nm. In 2007 Bendig et al. reported a decaging mechanism of coumarin caged substrate **39** via an ion-pair solvent-assisted heterolysis (Scheme 3). [118, 119] Thus, the solvent polarity plays an important role on photolysis resulting in significantly higher quantum yield on using a polar protic solvent such as methanol or water compared to that of organic solvent like acetonitrile or hexane. For

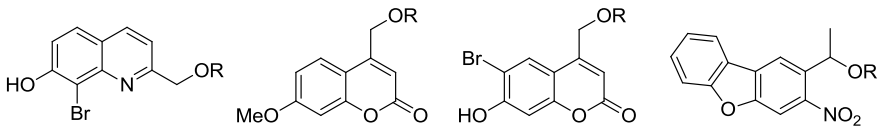
example, the fluorescence quantum yield of 7-methoxy coumarin **43** at 330 nm irradiation in methanol is 0.10 whereas it is greatly diminished to 0.012, 0.007, and 0.006 when using acetonitrile, THF, and dioxane, respectively, as solvents.[119]



Scheme 3. Mechanism of photolysis of coumarin upon a two-photon excitation (R = substrate).

Several promising two-photon caging groups based on core structures 8-bromo-7-hydroxyquinoline (BHQ) **42**,[120] 7-methoxy coumarin (MCM) **43**, bromohydroxycoumarin (Bhc) **44**,[117, 119] *o*-nitrodibenzofuran (NDBF) **45**,[121, 122] and 7-nitroindolines[123] have been reported elsewhere. Recently, the Deiters group reported the synthesis and decaging of an *o*-nitrodibenzofuran (NDBF) **45** caged thymidine as a two-photon caged nucleoside.[101] Unlike coumarin and hydroxyquinoline, the *ortho*-nitrobenzyl functionality in NDBF caging group **45** undergoes Norrish type II decaging pathways of photolysis (Scheme x). The scope of two-photon excitation is expanding with the discovery of a variety of fluorescent molecular probes and an accessibility of two-photon microscopy to study various areas of biological studies.[57, 124] However, one of the limitations of the coumarin caging group is its intrinsic

fluorescence both before and after decaging, , which may require washing steps in cases of fluorescence imaging.



42 **43** **44** **45**

Figure 8. Caging groups **42**–**45** suitable for two-photon excitation (R = substrate such as DNA or protein).

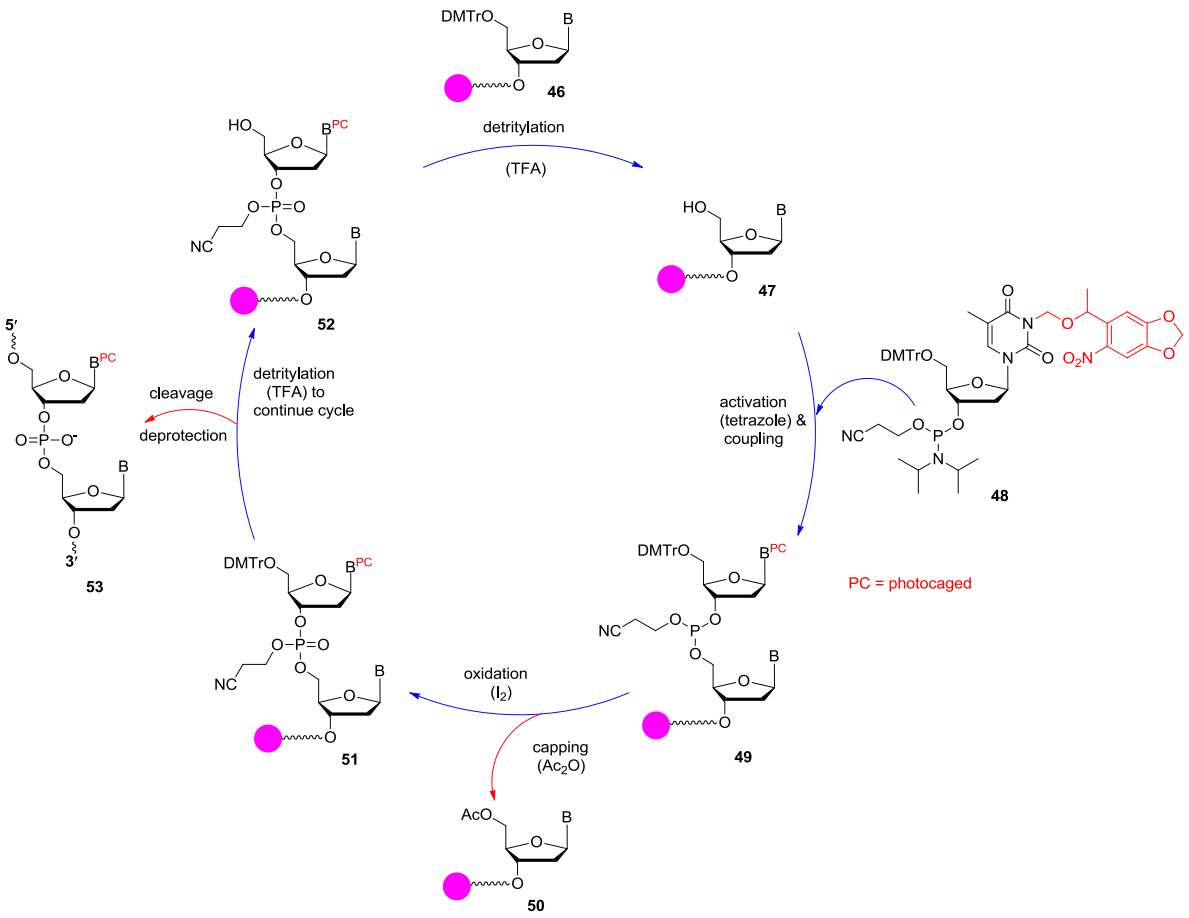
1.5. Synthesis of caged oligonucleotides

Oligonucleotides are nucleic acid sequences (DNA or RNA) that are synthesized from nucleotide precursors generating polymers of up to ~30 nucleotides (directing from 3′ to 5′). At present, oligonucleotides are synthesized using standard DNA synthesis protocols (Scheme 4). The automated DNA synthesis can be used for the incorporation of a caged phosphoramidites or any other unnatural nucleotides at a desired step during DNA or RNA synthesis.[125, 126]

The oligonucleotide synthesis is carried out in the 3′→5′ direction, after the first nucleoside is immobilized on a controlled pore glass (CPG) resin column at the 3′ sugar position **46**, subsequent detritylation, coupling, capping, and oxidation steps are used in sequence to couple each additional nucleotide. The duration for each step in an automated DNA synthesis is about 30–50 seconds but may vary in different protocols. Trichloroacetic acid (TCA) is used to remove the dimethoxytrityl group by liberating a reactive 5′ OH group

in **47** which is then reacted with a new incoming phosphoramidite **48** in the presence of tetrazole to form a dinucleotide **49**. The coupling time may vary up to several minutes (about 4-6 minutes for our caged phosphoramidite) to ensure completion of the reaction. After the coupling, a capping step is performed to acetylate unreacted 5′ OH groups in **47** by using acetic anhydride in the presence of *N*-methyl imidazole resulting in **50**. Then, the trivalent phosphite **49** is oxidized to a more stable phosphate group **51** through treatment with a mixture of iodine:water:THF:pyridine. This is the final step of a single nucleotide addition leaving the 5′ ODMT group for the successive cycles. Once all the nucleotides are added, the 5′ DMT group is removed by treating with trichloroacetic acid. Concentrated ammonium hydroxide is used to cleave the oligonucleotide from the resin, to remove the cyanoethyl protecting group as well as exocyclic amine protecting groups on certain bases (A, C, G). After the cleavage, the DNA oligonucleotide product **53** may be purified by polyacrylamide gel electrophorosis (PAGE) or high performance liquid chromatography (HPLC), and confirmed by mass spectrometry.

Although this method of oligonucleotide synthesis is common and widely accepted, limitations may arise when a specific functionality has to be inserted into the oligonucleotide. For instance, an azide group can be a universal ligation tool for bioconjugation reactions, but this particular functionality is not compatible with aforementioned oligonucleotide synthesis protocol, since the phosphoramidite monomer contains trivalent phosphorus, inducing azide reduction.[127]



Scheme 4. Automated DNA oligomer synthesis, as explained in the text.[128]

CHAPTER 2: CONTROLLING GENE FUNCTION WITH CAGED OLIGONUCLEOTIDES

2. Controlled gene expression

Protein levels, the ultimate functional units of gene expression, depend upon both the transcription and the translation efficiency of corresponding genes.[129, 130] A number of factors including transcriptional regulators, transcript localization and stability, translational regulation, protein stability and post-translational modifications play crucial roles in gene expression.[129] Understanding the functional roles of over 20,000 proteins from vertebrates is a challenge in the post genome era.[131] Continuous development of facile tools to advance the deconvolution of gene function in biology is essential to dissect the regulation of the protein coding genes.[132] It appeared that the molecular mechanism of gene expression is enormously complex and context dependent, albeit various tools and protocols have been previously developed to understand gene expression on a molecular scale.[129, 130, 133] The two main approaches to study gene regulation, either forward or reverse genetics approach, have their own advantages and limitations.[134, 135]

2.1. Caged oligonucleotides for controlled gene expression

Caged oligonucleotides can be used as an antisense agent for the deactivation of a target gene using light.[80] This methodology is particularly appealing as an alternative approach to gene knock down via gene silencing because of the noninvasive and orthogonal nature of light in cells.[77, 136] Additionally, recent investigations revealed that caged oligonucleotide technology can be extended to target gene activation.[137] Caged oligonucleotides have been used for various applications including but not limited to: 1) photochemical control of DNA-DNA interactions,[83] 2) DNA logic gate activation,[73] 3) activation and deactivation of the polymerase chain reaction,[138] 4) DNA decoy activity,[71] 5)

triplex-forming oligonucleotide (TFO) function,[137] 6) DNA-RNA interaction,[139] 7) RNA-RNA interactions,[72] and 8) DNA(RNA)-protein interactions.[140, 141]

Mechanisms of gene silencing have been studied both at the transcriptional and post-transcriptional level (Figure 9). The mode of action for silencing a target gene is dependent on its binding partners. For example, a triplex-forming oligonucleotide (TFO)[142, 143] or H-DNA[144, 145] sequence specifically interacts with a DNA duplex by forming a stable triplex, which inhibits transcription. On the other hand, antisense oligonucleotides bind with messenger RNA (mRNA) as the target gene by forming a stable duplex and then induce degradation of the target gene or block ribosome function.[46]

The inhibition of gene expression using an antisense oligomer can be accomplished by one of two pathways (Figure 9). First, an antisense DNA oligonucleotide can bind with the target mRNA and block gene expression by forming a stable duplex. Second, a DNA antisense agent binds with the target mRNA forming a duplex which is subsequently recognized by RNase H for sequence specific degradation. Third, the duplex (siRNA:RNA) incorporates into the RNA-induced silencing complex (RISC) which selectively degrades the target gene. The aforementioned approaches have been used to investigate gene regulation and protein functions with a variety of probes in the past.

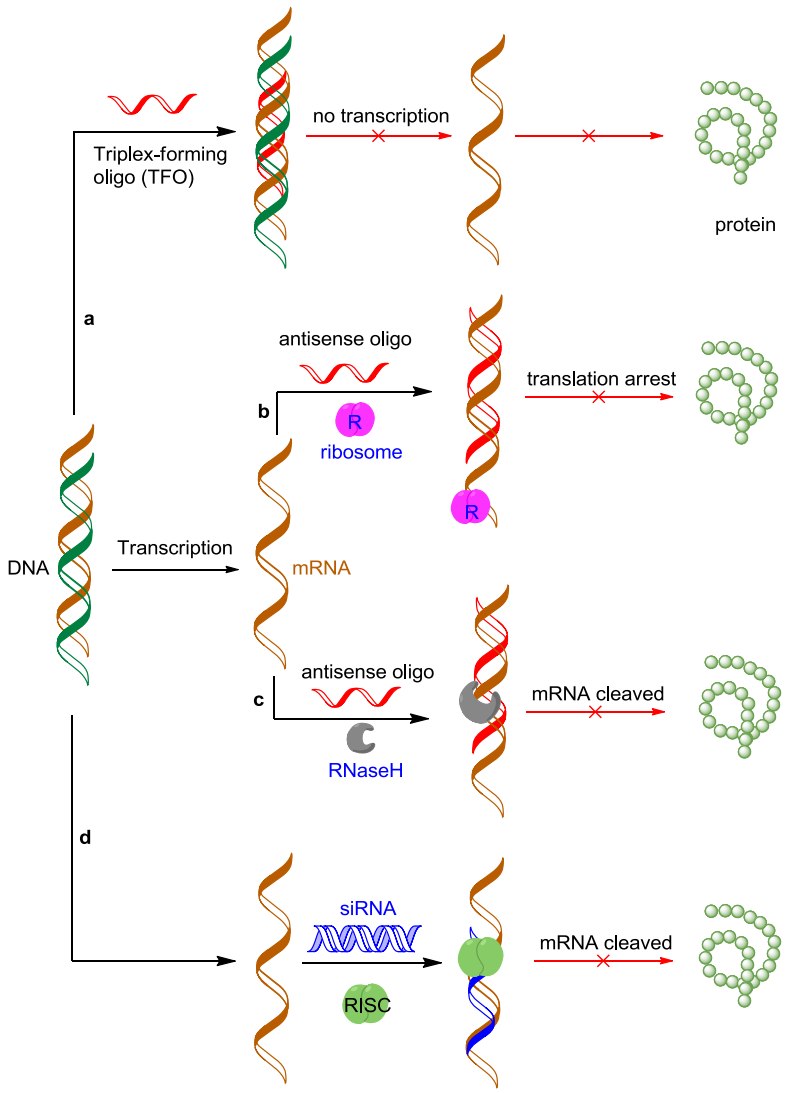


Figure 9. Mechanisms of gene silencing using different oligonucleotides.

Investigations have shown that caged oligonucleotides are sufficient enough to produce gene silencing activities in a wide range of cell lines.[146] In addition, the photo-mediated spatiotemporal activation or deactivation of target genes using caged oligonucleotides for therapeutic treatment is a possibility for the future.[12] However, a few

disadvantages of the caged oligonucleotides such as potential toxicity caused from off-target effects and the lower efficiency in cellular uptake[147] need to be addressed.

Nucleosides modifications have been carried out both to improve the photophysical properties and to optimize spatiotemporal control resolutions of the gene functions. Generally, synthetic modification strategies address the following major concerns: the stability of the modified oligonucleotide in DNA synthesis conditions, robustness in cellular environment, efficacy of DNA binding properties, and the ease of the monomer synthesis. A variety of nucleoside caged analogues (Figure 3 and 5) have been developed from a few different synthetic approaches. Literature examples show the installation of a caging group at the exocyclic oxygen **23**–**25**, nitrogen **19**–**21** or heterocyclic nitrogen **14**–**15**. Among these modifications, caging groups installed at the nitrogen atoms demonstrated better stability.

In general, the presence of a caging group interrupts either the hydrogen bonding between oligonucleotides, or sterically blocks the DNA-protein recognition. Upon exposure to light, the caging groups are removed from the oligonucleotide restoring hydrogen bonding ability (Figure 10).[86]

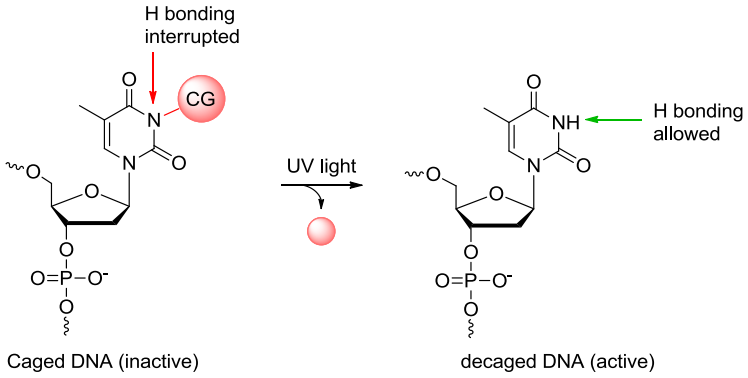


Figure 10. Activation of a caged oligonucleotide using light.

The efficacy of a caged oligonucleotide indisputably relies on the number of caging groups and their distributions. For example: 1) evenly distributed three caging groups per a 19-mer was sufficient to conduct a light-triggered PCR reaction,[148] 2) evenly distributed 3-4 caging groups per 19-mer oligonucleotide was effective for achieving gene silencing in mammalian cells,[61] 3) a caged antagomir with three caging groups was used to investigate light-mediated spatiotemporal control of microRNA function,[72] 4) a single caging group at the 5′ end of an antisense strand was used to regulate kinase activity,[149] 5) about 1.4 caging groups per duplex was used to control light-activated RNA interference,[77] 6) a single caging group was sufficient enough to accomplish photochemical regulation of peroxidase activity,[150] and 7) specifically positioned 1-2 caging groups per 15-mer DNA were sufficient to mask aptamer function.[151]

2.1.1. Transcriptional regulation using caged Triplex-Forming Oligonucleotides (TFO)

Triplex-forming oligonucleotides (TFO),[142, 152] can recognize short sequences of about 10-30 nucleotides duplex DNA. The site-specific sequences of those duplex DNAs, also known as TFO target sequence (TTS), are mainly found in the promoter region and provide invaluable information on genome manipulations.[153-155] It was estimated that 97.8% of the known human genome and 95.2% of the known mouse genome have at least one potential TTS in a promoter region or transcription region.[156]

TFOs regulate the gene expression through transcription inhibition. They recognize poly-purine or poly-pyrimidine rich regions of a double stranded DNA from the major groove of DNA by forming Hoogsteen hydrogen bonds.[157, 158] Generally, adenine hybridizes to an adenine:thymine base pair forming an T:A:T triplex (Figure 11a), while guanine binds to guanine:cystosine base pair through reverse Hoogsteen hydrogen bonds forming a G:G:C triplex structure (Figure 11b).[154, 155, 159, 160]

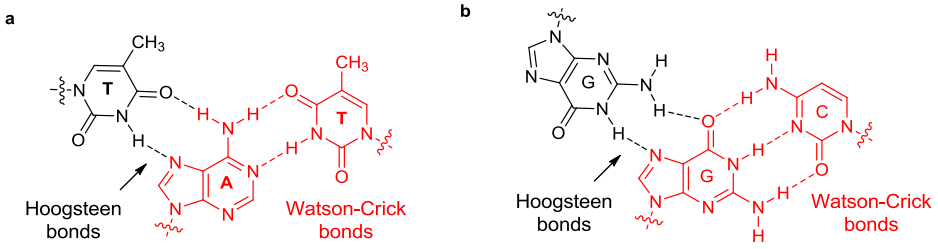


Figure 11. An interaction of a TFO with a DNA duplex through Hoogsteen bonds: a) T:A:T triplex, and b) G:G:C triplex. Adapted with permission from *ACS Chem. Biol.* **2012**, *7*, 1247-56.

TFOs bind double-stranded DNA in a sequence-specific manner for performing a variety of functions[161] (Figure 12) such as inhibition of protein-DNA interaction,[157] gene expression,[160] and DNA replication.[162] They have also been applied to induce site-specific DNA damage,[163] to enhance DNA recombination,[164] to perform DNA mutagenesis,[154, 155] and are also used as electrochemical sensors for double-stranded DNA.[165]

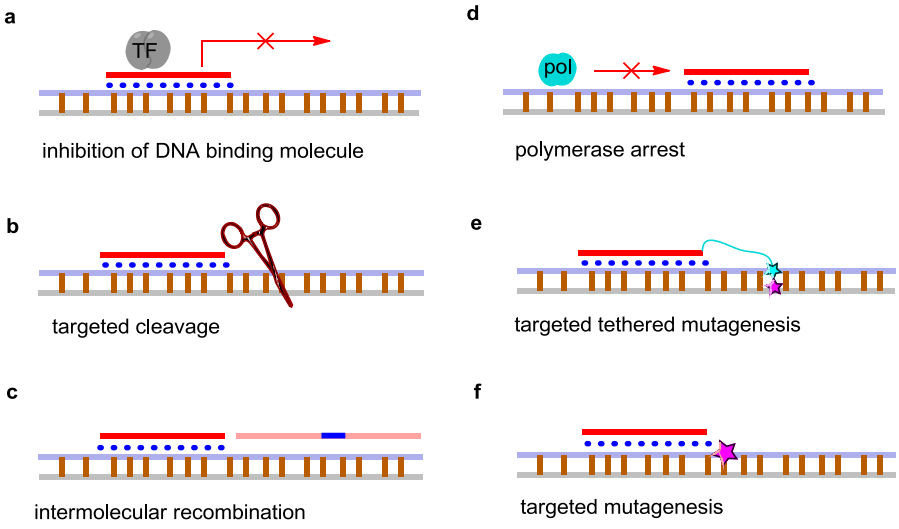


Figure 12. TFOs as potential therapeutics. a) Inhibition of transcription factor from binding to DNA, b) targeted DNA cleavage, c) molecular recombination, d) polymerase arrest, and e-f) mutagenesis.

A high level of gene suppression, the foremost benefit of the TFOs, can be accomplished by using TFOs in comparison to that of the antisense agents because cells contain relatively low copies of DNAs compared to mRNAs.[155]

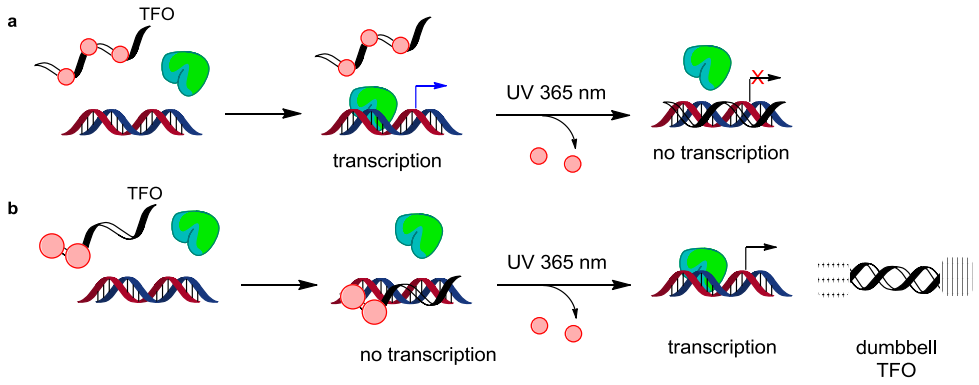


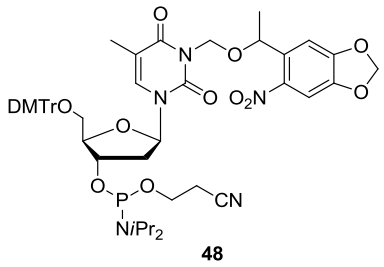
Figure 13. Transcription regulation of a target gene using a caged TFO: a) transcription inhibition through a light-activation, and b) transcription regulation through a light-activation. Adapted with permission from *ACS Chem. Biol.* **2012**, *7*, 1247-56.

Both the inhibition and the activation of the target gene (*vide infra*) can be achieved by an appropriate design of the caged TFO (Figure 13). First, the presence of caging groups in the TFO prevents interaction with the duplex DNA thereby rendering it inactive. Upon irradiation, the caging groups can be removed from the TFO resulting in an active TFO that site-specifically binds to the target sequence blocking gene expression (Figure 13a). Second, the presence of caging groups in a specially designed TFO forms a triplex with the target DNA, inhibiting the transcription factor. Once the caging groups are removed the resulting TFO forms an inactive dumbbell oligonucleotide resulting in the activation of the gene expression (Figure 13b).

In order to control the transcription inhibition and activation of the target gene using light, we prepared caged TFOs using both NPOM-caged thymidine **48** and NPE-caged 2′ deoxycytidine **61** (*Vide infra*). In the past, 1-(2-nitrophenyl) ethyl caged deoxycytidine has been reported in the activation of aptamers using light.[166] It has been shown that the efficiency of photolysis of the 2-nitrobenzyl caging group depends upon the nature of its aromatic substituents.[167] Electron-rich substituents on the benzene ring and a methyl substituent in the benzylic position enhance the photolytic properties of such caging groups enabling the use of long-wavelength UV light of 365 nm and thus are less photo-damaging for biological systems.[109, 168] Taking advantages of electron donating nature of the methylenedioxy moiety and an α-methyl group, we synthesized a nitropiperonyl ethyl (NPE)-caged deoxycytidine phosphoramidite (Scheme 5) for preparing caged DNA oligonucleotides with light-activable deoxycytidine residues. The caged 2′-deoxycytidine phosphoramidite (*vide infra*), and caged thymidine phosphoramidite[169] were synthesized according to previously reported procedures. Both the NPOM caged thymidine TFO and the NPE caged 2′-deoxycytidine TFO were then employed to perform photo-mediated target gene inhibition and activation. An endogenous gene, Cyclin D1, was selected as the target gene for the photocaged TFO due to its critical regulatory role in the cell cycle.[170-172]

Cyclin D1 is essential for G_1 to S phase transition during the cell cycle.[173, 174] The misregulation of this process has been associated with a number of diseases.[175, 176] For instance, over-expression of cyclin D1 causes multiple downstream effects, including anchorage-independent growth, vascular endothelial growth factor production,

tumorigenicity, and resistance to chemotherapeutic agents.[176] Previously, inhibition of cyclin D1 gene transcription[177] and translation[178] has been achieved using oligonucleotides.

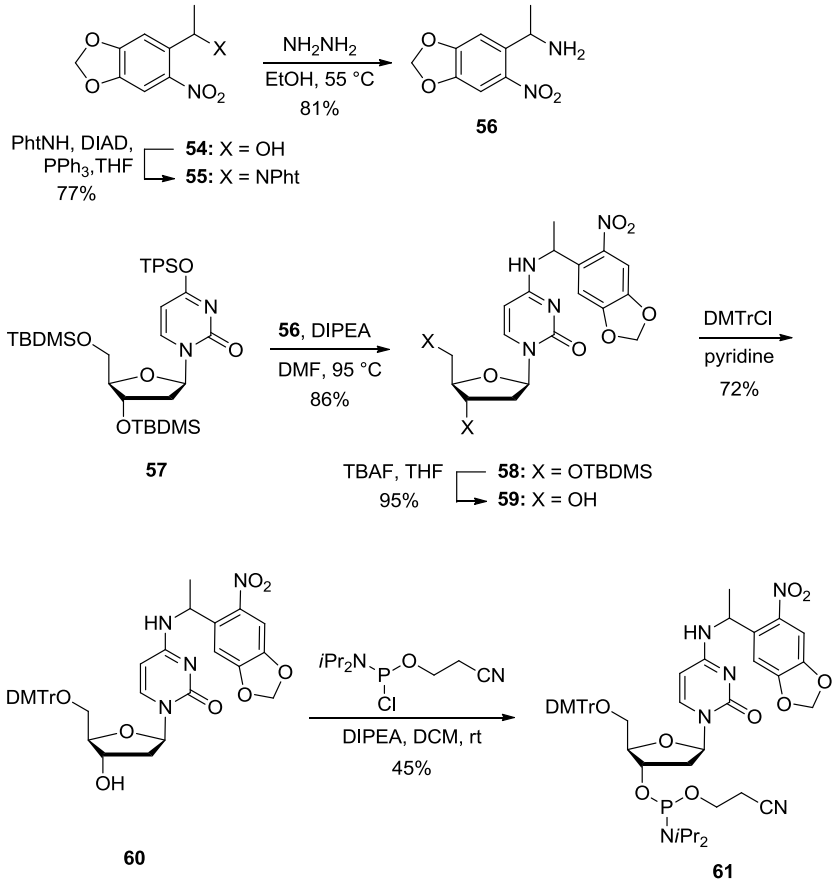


48

Figure 14. NPOM-caged thymidine phosphoramidite **48.**[83]

Our synthesis of the caged phosphoramidite **61** commenced with the coupling of phthalimide with the known alcohol[179] **54** in the presence of diisopropyl azodicarboxylate and triphenylphosphine in THF, affording the corresponding nitrodione derivative **55** in 77% yield.[180] The dione **55** was then reacted with hydrazine in refluxing ethanol delivering the amine **56** in 81% yield. The amine **56** was subsequently coupled with TBDMS-protected triisopropylbenzenesulfonyl activated deoxyuridine **57** (synthesized in 2 steps from 2′-deoxyuridine)[181] in the presence of diisopropyl ethyl amine (DIPEA) in DMF under reflux conditions to obtain the NPE-caged deoxycytidine nucleoside **58** in 86% yield. The TBDMS groups on **58** were removed through treatment with TBAF in THF at low temperature to give the corresponding alcohol **59** in 95% yield. Selective tritylation of the 5′ hydroxyl group on **59** with dimethoxytritylchloride in pyridine was carried out at room temperature and delivered the compound **60** in 72% yield. Finally, activation of the 3′ hydroxyl group on **60**

was achieved by reacting it with 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite in the presence of DIPEA in dichloromethane providing the desired caged deoxycytidine phosphoramidite **61** as a white solid in 45% yield.



Scheme 5. Synthesis of the NPE-caged deoxycytidine phosphoramidite **61**.

2.1.2. Light-activated transcription using a caged TFO

It is anticipated that the specifically designed caged dumbbell TFO binds to the DNA duplex, forming a triplex structure that blocks the transcription factors, thereby inhibiting the

gene expression. A brief UV irradiation at 365 nm removes the caging group of the dumbbell TFO thereby resulting in an inactive DNA dumbbell structure that may not bind to the DNA duplex allowing the gene expression. In order to test this hypothesis, our group designed both the regular dumbbell TFO (DB-TFO-1) and the other two caged dumbbell DNA oligonucleotides as well. Our newly synthesized NPE-caged cytidine **61** was used to prepare the first caged dumbbell TFO (CDB-TFO-1) that contains four NPE-caged cytidine **61** residues, and the second one CDB-TFO-2 with five NPE-caged cytidines by James Hemphill in the Deiters group. Light-mediated transcription activation of a target gene was then investigated, using caged dumbbell TFOs, by Jeane Govan in the Deiters group.

In this particular experiment, the TFO (DB-TFO-1) serves as a control. In addition, additional oligonucleotides were prepared such that a hairpin-protected TFO (HP-TFO-1) and a corresponding caged TFO (CHP-TFO-1) containing four NPOM-caged thymidine **48** units remained at an interval of every 3-6 nucleotides throughout the binding sites for obtaining an optimum efficiency.[182, 183] The particular design of the hairpin-protected TFO was selected based on the previous report, which suggested that a hairpin loop structures on the 5′ and 3′ termini of antisense agents help stabilize oligonucleotides rendering its antisense properties.[184, 185] Performed gel shift assays suggested that only three TFOs namely HP-TFO-1, CDB-TFO-1, and CDB-TFO-2 efficiently bind with the DNA duplex whereas the other two TFOs namely CHP-TFO-1 and DB-TFO-1 remained unbound with DNA duplex.

All of the aforementioned non-caged and caged dumbbell TFOs were tested for their ability to control gene activity in cell culture. HEK 293T cells were co-transfected with plasmids encoding firefly luciferase (pCyclin D1 Δ-944), *Renilla* luciferase (pRL-TK), and

with the TFOs. As a positive control, the CHP-TFO-1 showed no inhibition of gene expression until irradiation that induces decaging and results in nearly a complete inhibition of luciferase gene expression (Figure 15). The other control TFO (DB-TFO-1) shows no inhibition of gene expression regardless of the presence or the absence of UV irradiation. However, the dumbbell TFO (DB-TFO-1) did not inhibit gene expression. The CDB-TFO-1 remained as an active TFO, inhibiting luciferase gene expression until irradiated with a UV light at 365 nm. Caging groups are removed upon irradiation allowing the TFO to form dumbbell structure resulting in a substantial increase in luciferase activity, an indication of gene expression. Similarly, the CDB-TFO-2 inhibits gene expression and the gene expression is restored after UV exposure. These results demonstrate that the caged dumbbell TFOs can be complementary tools to the caged TFOs. Together they enable the promoter-specific light-activation and light-deactivation of transcription with excellent efficiency. This application may be useful to explore transcriptional on/off light-switches using TFOs.

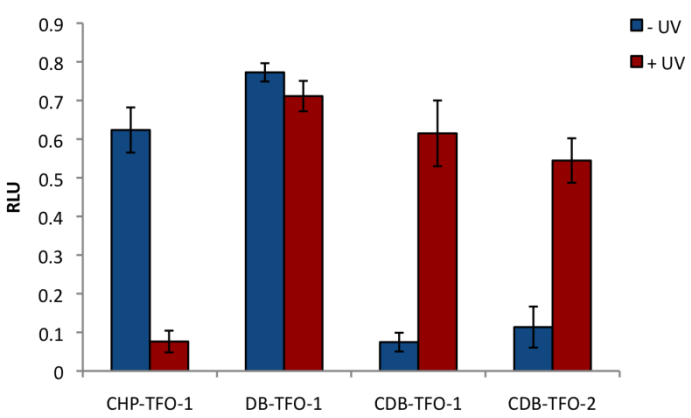


Figure 15. Photochemical activation of gene transcription in HEK 293T cells. Relative luciferase units (RLU) represent a firefly luciferase signal normalized to a *Renilla* luciferase. Error bars represent standard deviations from three independent experiments, (Experiments were performed by Jeane Govan). Adapted with permission from *ACS Chem. Biol.* **2012**, *7*, 1247-56.

2.2. Cellular delivery and photochemical activation of antisense agents

The research interest in the area of antisense agents is ever expanding. In the past, only a few antisense agents have been employed to study gene regulations or a controlled gene expression[186, 187] mainly due to their weak stability under physiological conditions, poor bioavailability and toxicity.[188-191] Literature reports show that a wide variety of antisense agents have thoroughly been investigated for their therapeutic potential. In addition, microinjection technology and use of a transfection agent can assure efficient delivery of antisense agents into cells. However, cellular toxicity of available transfection reagents is one

of the major concerns of their usage[192] Alternatively, new synthetic derivatives may be investigated.

We envisioned that a caged thymidine derivative may be used to improve the cellular delivery of antisense agents and perhaps can bypass the toxicity associated with commercial transfection reagents such as lipofectamine. Therefore, we synthesized propargyl-6-nitroveratryloxymethyl (PNVOM) caged thymidine **67** and explored it as a bioconjugation handle for the modification of caged antisense agents.

After oligonucleotide synthesis, the alkyne moiety may be coupled with azide functionalized cell penetrating peptides (CPPs), such as HIV-TAT-azide, using 'Click' chemistry to prepare a corresponding PNVOM caged-HIV TAT-oligonucleotide conjugate **69** (vide infra). The resulting conjugate may facilitate cellular delivery without need for transfection. After their delivery in cells, the photo-cleavable peptide linker could be removed from the conjugate by using UV light to restore a native antisense oligonucleotide (Figure 16b). Advantages of this approach are: 1) efficient delivery of an antisense agent into cells without transfection, and 2) the controlled release of a native antisense agent from the CPP resulting in its concomitant activation with light.

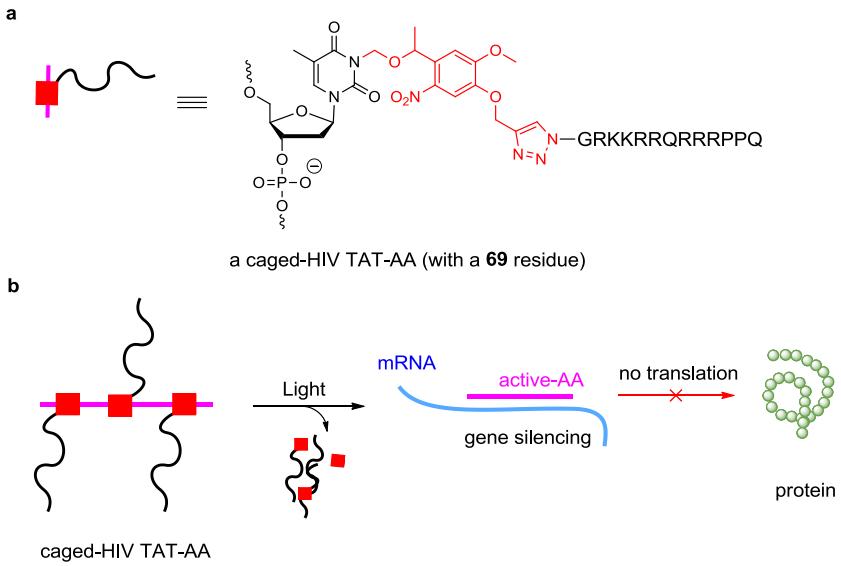
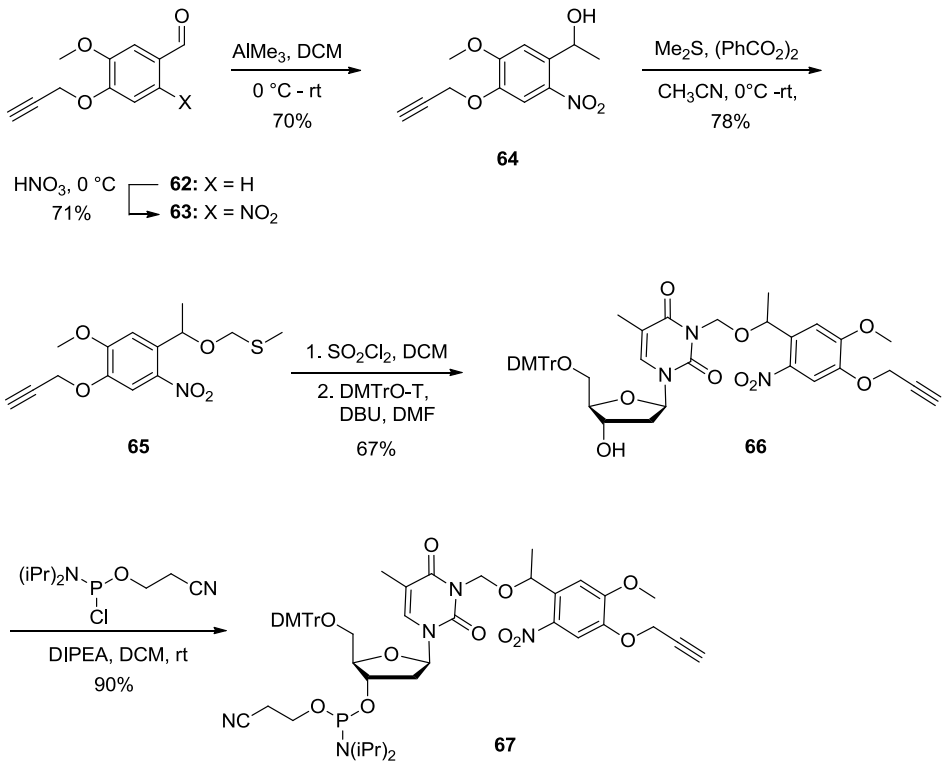


Figure 16. A caged-HIV TAT-oligonucleotide conjugate for a controlled antisense activity. a) NVOM caged-HIV TAT-AA conjugate, and b) a caged TAT-oligonucleotide conjugate on gene silencing; HIV TAT = GRKKRRQRRRPPQ.

CPPs are one of the most investigated agents for cellular delivery of cargos. The proposed modes of action for cellular delivery are based on passive diffusion and endocytosis-like pathways.[193] Studies showed that the modes of cellular uptake rely on several parameters, including CPP concentration, cargo size, CPP type, and cell type.[194, 195] Results from research studies suggest that the HIV TAT CPP may undergo internalization through endocytosis pathways.[196, 197] The efficiency of a particular CPP-conjugate for the effective cargo delivery depends on the nature of CPP, the cargo molecule, and the covalent linker.[198]

Previously CPPs, short (8–30) synthetic peptides that facilitate the cellular uptake of a cargo,[199-201] have been conjugated to a variety of biomacromolecules including plasmids,[202] oligonucleotides,[203] and proteins[204] for the efficient delivery of aforementioned macromolecules into cells. Diversity in CPP conjugates is rapidly expanding, and has resulted in a wide variety of CPP-conjugates including heparan derived CPPs (DPV3 and DPV1047), HIV TAT, penetratin, K-FGF, Bac7, and others.[200] An HIV TAT peptide siRNA oligonucleotide conjugate has been reported to significantly increase its gene silencing efficiency.[205] The TAT sequence responsible for the cellular delivery, HIV TAT *trans*-activating factor, was later on discovered as one of the first transduction proteins.[206]

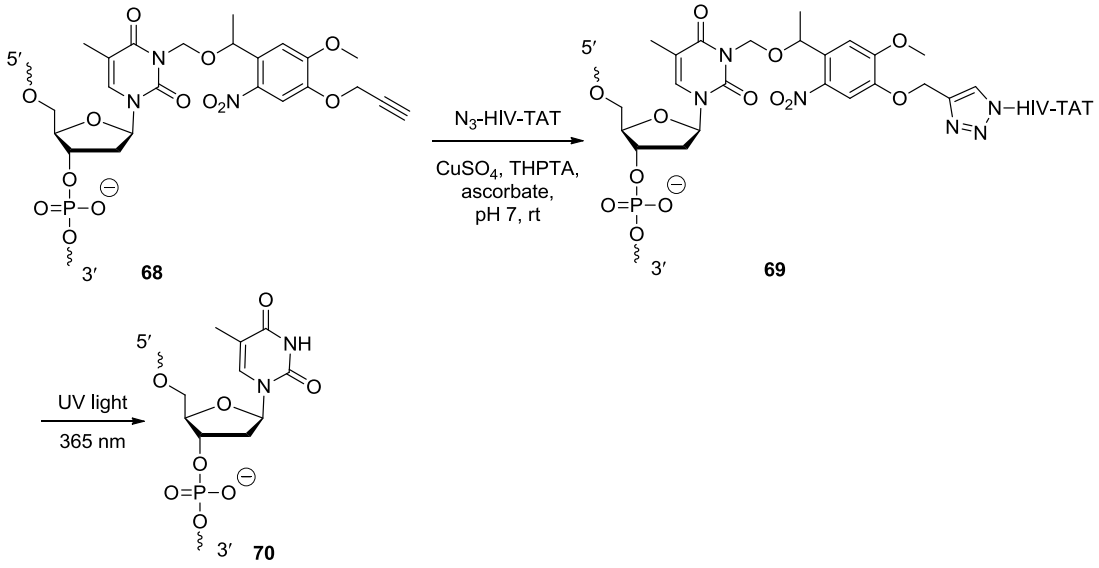
Here, in order to perform a site-specific oligonucleotide conjugation, we synthesized an azido-modified HIV TAT peptide and a PNVOM-caged thymidine phosphoramidite. The five-step synthesis of the phosphoramidite **67** started with aldehyde **62**,[207] which was prepared from vanillin (Scheme 6). Nitration of **62** with concentrated HNO_3 gave the corresponding *ortho*-nitrobenzaldehyde **63** as a yellow solid in 71% yield. Methylation of **63** with $AlMe_3$ in DCM at 0 °C provided the alcohol **64** in 70% yield, which was subsequently reacted with methylsulfide in the presence of benzoyl peroxide to furnish the thiomethylene derivative **65** in 78% yield. The sulfide **65** was activated *in situ* with sulfuryl chloride at –78 °C as the corresponding propargyl-6-nitroveratryloxymethyl chloride (PNVOM-Cl) that was then reacted (without further purification) with DMT-protected thymidine in the presence of DBU in DMF to obtain PNVOM-caged thymidine **66** in 67% yield. Finally, reaction of caged nucleoside **66** with 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite delivered the caged thymidine phosphoramidite **67** as a white solid in 90% yield.



Scheme 6. Synthesis of a propargyl-6-nitroveratryloxymethyl (PNVOM)-caged thymidine phosphoramidite **67**.

The PNVOM-caged thymidine nucleotide was then site-specifically incorporated into hairpin-protected antisense agents using standard DNA synthesis conditions by Mark O. Lively at Wake Forest University. Hairpin-protected antisense agents can inhibit gene expression without further modifications.[184, 185, 208] In this context, the hairpin loop structure protects the phosphodiester antisense agents from intracellular degradation obviating additional backbone modifications for their stabilization (e.g., phosphorothioates) that could elicit an immune response[209] or increase non-specific protein binding.[210]

The antisense agent containing three PNVOM units **68** was reacted with an azido-HIV TAT peptide through Cu(I)-catalyzed [3+2] cycloaddition reaction[211](Scheme 7) in the presence of tris(3-hydroxypropyltriazolylmethyl)amine (THPTA) as a ligand.[212, 213] The obtained caged-HIV TAT-oligonucleotide conjugate was purified using a NAP5 column.[212, 214] Gel shift assays were performed to confirm both 1) the complete conversion of the caged oligonucleotide **68** into its caged-HIV TAT-conjugate **69**, and 2) the restoration of the native antisense oligonucleotide from the conjugate **70** upon a brief UV irradiation.



Scheme 7. Conjugation of the PNVOM-caged antisense agent **68** with an azido-HIV TAT peptide to furnish a caged HIV TAT-oligonucleotide conjugate **69**, and its light-activation upon a UV irradiation. Experiments were performed by Jeane Govan.

To demonstrate an efficiency of the light activation and the delivery of the CLuc-AA-TAT, the human embryonic kidney (HEK) 293T were transfected with the reporter plasmids pGL3 and pRL-TK, encoding firefly luciferase and *Renilla* luciferase respectively in order to normalize reporter gene expression in a dual-luciferase assay. The control CNTRL and Luc-AA were either transfected into cells or simply added to the growth media. The results from these experiments (performed by Jeane Govan) showed that, in either case, the CNTRL antisense agent did not inhibit luciferase gene expression regardless of UV irradiation (Figure 17). On the other hand, a positive control Luc-AA reduces *Renilla* luciferase gene expression up to 85% relative to the CNTRL antisense agent while using a transfection reagent. Luc-AA when added to the growth media led almost no inhibition of gene expression suggest that there was no cellular uptake of Luc-AA in the absence of a transfecting reagent or a cell transduction peptide.

In contrast, when CLuc-AA and CLuc-AA-TAT were added to the cell culture media, almost no inhibition of gene expression was observed in cells that were kept in the dark. However, upon a brief UV irradiation (2 min, 365 nm), only CLuc-AA-TAT inhibited about 92% of *Renilla* luciferase expression, which is comparable to that of using a transfection reagent. These results illustrate that HIV TAT peptide covalently linked to the antisense agent facilitates cellular uptake and that light-activation restores the native form of the antisense agent to induce gene silencing.

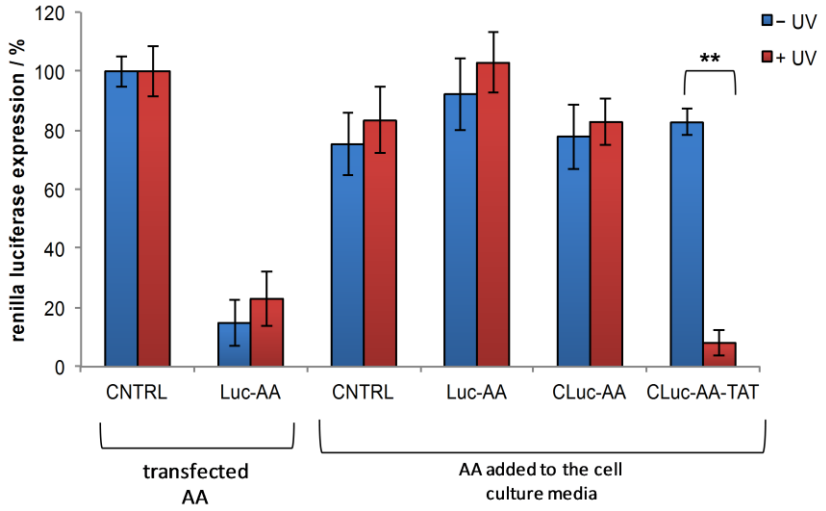


Figure 17. Light-activation of antisense agents targeting luciferase gene expression. *Renilla* luciferase expression with (left panel) or without (right panel) transfecting reagent and is normalized to firefly luciferase expression. CNTRL (a negative control antisense agent) is set to 100%. Error bars represent the standard deviations of the triplicate experimental data, (Experiments were performed by Jeane Govan).

2.3. Caged antisense oligonucleotides to control microRNA function

MicroRNAs (miRNAs), a class of about 22 nucleotides short non-coding endogenous single stranded RNAs, are negative regulators of gene expression at the post transcriptional level.[215-217] Often, miRNAs mediate the down-regulation of a gene by the sequence specific and imperfect complementary binding to the 3′ untranslated regions (UTRs) of target mRNAs thereby accelerating the degradation or the deadenylation of the mRNA.[218, 219] Generally, mRNAs with longer 3′ UTRs contain four conserved miRNA target sites, on average, and are extensively regulated by miRNAs.[220] Targets of miRNAs range from genes

that encodes various molecules including signaling molecules, enzymes, and transcription factors to RNA binding proteins to name a few.[221-223] Investigations on miRNAs have expanded substantially over the last 20 years revealing its widespread participation in biological systems.[224-228] miRBase, an online repository for all miRNAs sequences and annotation, includes over 15,000 miRNA gene loci in over 140 species and 17,000 distinct mature miRNA sequences.[229]

More than 1000 miRNAs have already been discovered in humans, which represents around 1-4% of all the genes and thereby affects up to 40-60% of the protein coding genes.[224, 230, 231] Although miRNA expression levels can differ widely for each miRNA, it is predicted that miRNAs are more abundant (up to 10,000 copies per cell) than their target mRNA genes (up to 500 copies per cell).[232] In a typical animal cell, estimated concentration of a miRNAs is at maximum up to 22 μM, which is at least ten times higher than that of target mRNA genes whose concenntrations range from 80 pM-2.2 μM.[232] The stability of a miRNA ranges from 28 h (miR-155, a less stable) to 9 days (mir-125b, the most stable one). The average stability of a miRNA is generally 10 times more than a target mRNA gene (the median half-life of which is about 10 hours). [233, 234]

After transcription from the genome, miRNAs undergo several post-transcriptional processing steps via a dedicated miRNA pathway to produce mature miRNAs through a universally linear biogenesis.[216] At first, miRNAs are transcribed by RNA polymerase II as a capped and polyadenylated large precursor called primary miRNAs. These primary miRNAs are further processed by an enzyme called Drosha with the help of a co-factor protein DGCR8 in the nucleus to form the 70 nucleotide RNA molecule called pre-miRNAs.[217, 235,

[236] The pre-miRNAs are transported into the cytoplasm by the exportin-5 complex where they are subsequently processed into an about 22 nucleotide mature miRNAs by another complex called Dicer (Figure 18).[237-239]

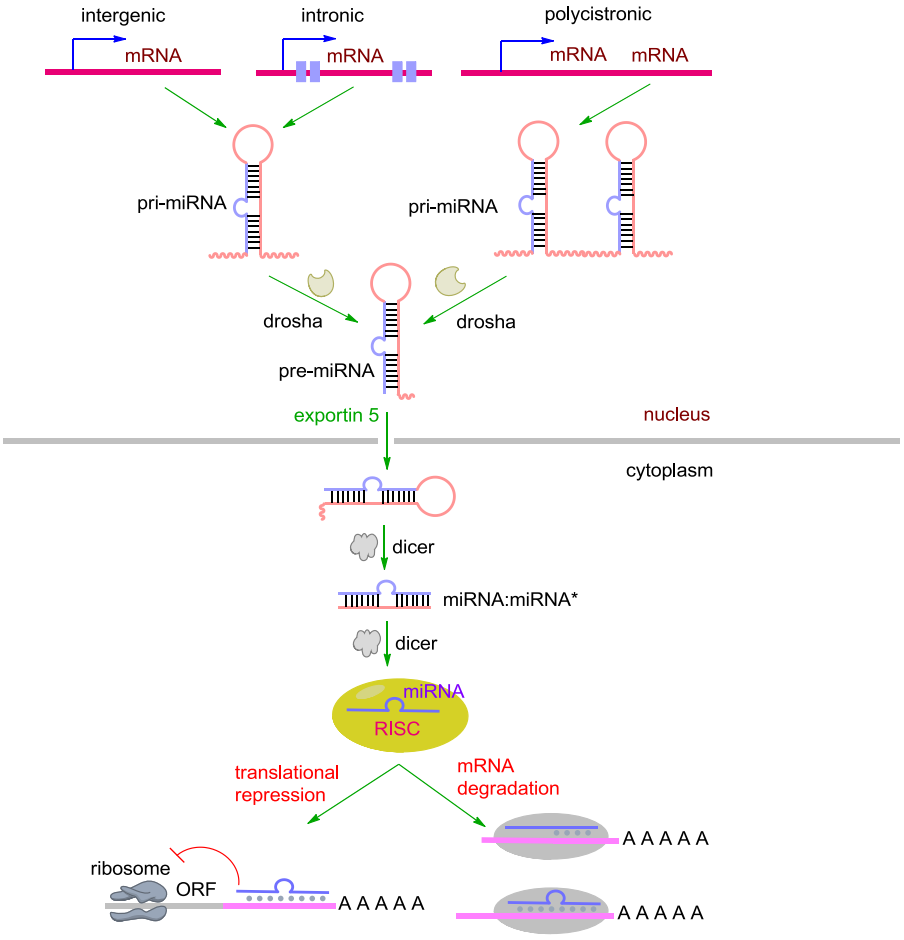


Figure 18. Biogenesis of microRNAs through a dedicated processing pathway (ORF = open reading frame). Adapted with permission from *Circ. Res.* **2011**, *108*, 219-234.

An abnormal expression level of numerous miRNAs has been linked to a wide range of human diseases including cancers,[240-242] cardiovascular diseases,[224, 243-245] defects in signaling pathways[246-249] and immune system,[250-254] metabolic disorders,[255] and viral infections.[228, 256, 257]

Tools to manipulate miRNA function are antisense oligonucleotides known as antagomirs.[13, 72] They include site-specifically modified oligonucleotides that are complementary to miRNAs and act as competitive inhibitors of the target mRNA binding sequence through stable duplex formation.[139] The chemical modifications are introduced into the antagomir structure in order to: 1) render the oligonucleotides resistant to degradation *in vivo*, 2) increase binding affinity, and 3) improve cellular delivery.[187] Such modifications include 2′ sugar modifications including 2′-*O*-methyl (2′OMe), 2′-*O*-methoxyethyl, 2′-fluoro, the addition of a bridging methylene group between the 2′-O and the 4′-C of the ribose ring (locked nucleic acid, LNA), and others.[12, 40, 44, 147]

2.3.1. Photoregulation of microRNA miR-21 and miR-122 functions using caged antagomirs

MicroRNA miR-122 is a liver specific miRNA which is necessary for hepatitis C virus replication and infectious virus production.[258 259] It has been reported that the knockdown of miR-122 with antagomirs resulted in a decrease in HCV RNA replication in human liver cells[258] and reduced HCV levels in chronically infected primates.[260] The miRNA miR-21 has been linked to several human malignancies and over expression of miR-21 has been observed in glioblastomas, breast, pancreatic, cervical, colorectal, ovarian, lung, and hepatocellular cancers.[261] MiR-21 functions as an anti-apoptotic factor in cancer cells and a

dependence of tumor growth on miR-21 presence has been demonstrated in a mouse model.[262] The involvement of miR-21 and miR-122 in human malignancies and an *in vivo* investigation on photochemical regulation make these miRNAs potential targets of new therapeutics.

Here, we hypothesized that light-activation of a caged antagomir technique can be used to control miRNA functions thereby inhibiting the gene expression. The caged antagomir has no effect on miRNA-mediated gene silencing prior to its exposure to UV light to remove the caging groups. After a brief UV irradiation, the activated antagomir forms a duplex with the target miRNA silencing its gene regulatory function.

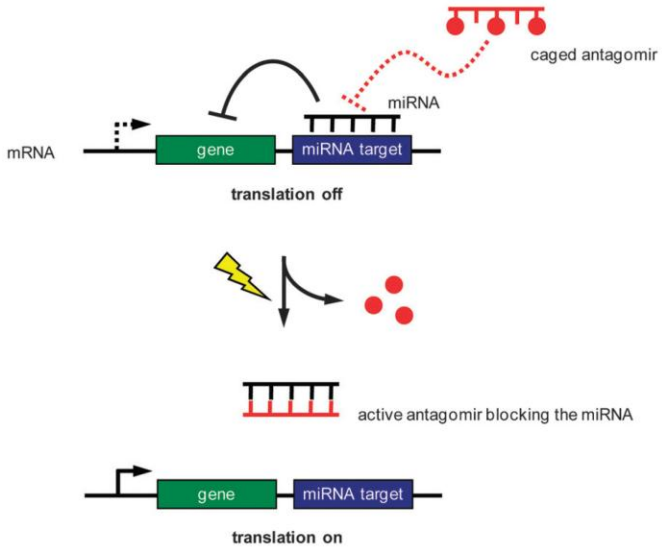
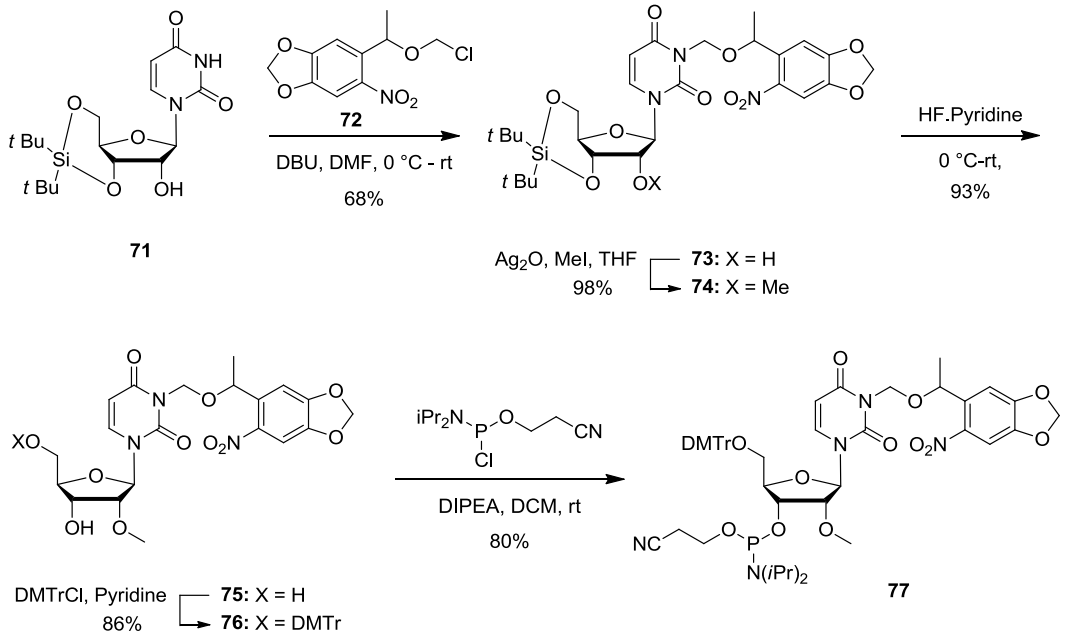


Figure 19. Photoregulation of microRNA function using a caged antagomir. Adapted with permission from *Mol. BioSyst.* **2012**, *8*, 2987-93.

Here, we prepared oligonucleotides with 2′-OMe modified nucleotides and phosphorothioate backbones with complementary sequences to the target mature miR-21 and miR-122 were prepared. 2′-OMe NPOM (6-nitropiperonyloxymethyl)-caged uridine phosphoramidite **77** was synthesized for site-specific installation of a light-removable caging group (Scheme 8). 3′,5′-*O*-(Di-*t*-butylsilanediyl)uridine[263] **71** was synthesized in one step and was subsequently reacted with NPOM chloride[264] **72** in the presence of DBU in DMF to furnish the corresponding NPOM-caged uridine **73** in 68% yield. NPOM chloride **72** was synthesized in three steps from commercial 6-nitropiperonal.[264] The NPOM caged uridine **73** was refluxed in methyl iodide in the presence of Ag_2O[265] to deliver the corresponding 2′OMe derivative **74** in 98% yield. Treatment of **74** with HF-pyridine at 0 °C in DCM yielded the NPOM caged 2′OMe uridine **75** in 93% yield. Protection of the 5′-OH group of **75** with DMTrCl in pyridine resulted in the corresponding DMT-protected caged compound **76** in 68% yield. Finally, **76** was activated with 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite to give the desired phosphoramidite **77** in 90% yield.



Scheme 8. Synthesis of an NPOM-caged 2′OMe uridine phosphoramidite **77**.

Non-caged and caged 2′OMe phosphorothioate antagomirs with 2′OMe NPOM-caged uridine groups were also synthesized (using the standard DNA synthesis protocol by James Hemphill in the Deiters group) for targeting mature miR-21 and miR-122. Four NPOM-caged 2′OMe uridine nucleotides at different positions were installed into the miR-21 caged antagomir, whereas only three 2′OMe NPOM-caged uridines were installed into the miR-122 caged antagomir. In addition, the psiCHECK-2 vector containing both a *Renilla* luciferase and an independently transcribed firefly luciferase reporter gene is chosen.[266] The complementary sequences of miR-122 or miR-21 were inserted downstream of the *Renilla* luciferase gene (Figure 19). Thus, the presence of the mature miRNA results in a decrease in the *Renilla* luciferase signal enabling the detection of endogenous miRNA levels. An active

antagomir would act as a competitive inhibitor of miRNA target gene and leading an increase in *Renilla* luciferase expression.

A Huh7 cell line that is stably transfected with the *Renilla* luciferase sensor for endogenous miR-122 was used to confirm the activity of the synthesized miR-122 antagomirs. In these in vitro experiments (performed by Colleen Connelly in the Deiters group), the Huh7-psiCHECK-miR122 cells were transfected with either the non-caged or the caged miR-122 antagomir. Following the transfection, the cells were either kept in the dark or irradiated at 365 nm and after 48 h, the relative luciferase signal was measured where *Renilla* luciferase was normalized to the firefly luciferase activity providing relative luminescence units (RLU). The non-caged miR-122 antagomir produced a 10-fold increase in RLU relative to a transfection control containing no antagomir (Figure 20). The caged miR-122 antagomir showed RLU values similar to the transfection control, indicating that the presence of three NPOM-caging groups on a 2′OMe U prevents the antagomir from inhibiting miR-122 function. Upon UV irradiation, the caging groups are removed resulting in a native miR-122 antagomir which efficiently inhibited the miR-122 and leads to a 10-fold increase in RLU.

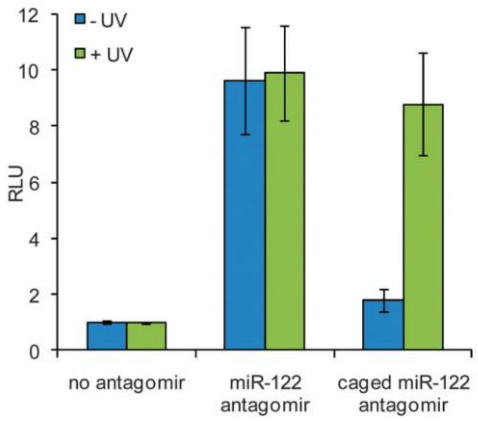


Figure 20. Light-activation of miR-122 inhibition: Inhibition of luciferase expression by light activation of a miR-122 antagomir in Huh7 cells, where cells were transfected with caged and non-caged miR-122 antagomirs and irradiated (365 nm for 2, 5, or 10 minutes) to activate antagomir function. The error bars represent standard deviations from three independent experiments. Experiments were performed by Colleen Connelly. Adapted with permission from *Mol. BioSyst.* **2012**, *8*, 2987-93.

Similarly, the miR-21 antagomirs were assayed using a Huh7 cell line transiently transfected with the *Renilla* luciferase sensor for endogenous miR-2. The Huh7 cells were co-transfected with the psiCHECK-miR21 plasmid and either the non-caged or the caged miR-21 antagomirs at 10 pmol. Following the transfection, the cells were either kept in the dark or irradiated. The relative luciferase signal was measured after 48 h, (Figure 21) showed that the non-caged miR-21 antagomir produced a 3-fold increase in RLU relative to a transfection control containing no antagomir. Prior to decaging, the caged miR-21 antagomir

showed RLU values equal to the transfection control, whereas a 3-fold increase in its values after UV irradiation, demonstrating that the miR-21 inhibitory activity is fully restored.

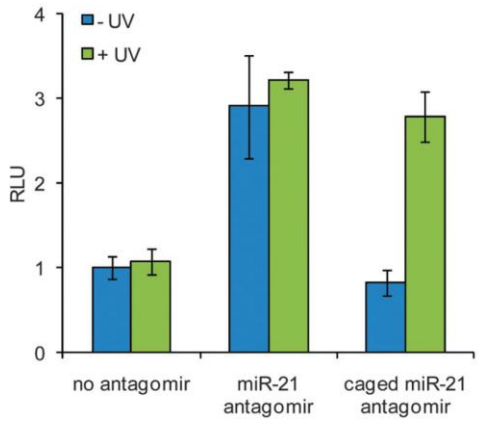


Figure 21. Light-activation of miR-21 inhibition: Inhibition of luciferase gene expression by light activation of a miR-21 antagomir in Huh7 cells, where cells were transfected with the caged and the non-caged miR-122 antagomir and irradiated (365 nm for 2, 5, or 10 minutes) to activate antagomir function. The error bars represent standard deviations from three independent experiments. Experiments were performed by Colleen Connelly. Adapted with permission from *Mol. BioSyst.* **2012**, *8*, 2987-93.

2.4. Experimental data for synthesized compounds

2-(1-(6-Nitrobenzo[d][1,3]dioxol-5-yl)ethyl)isoindoline-1,3-dione (**55**). Phthalimide (1.4 g, 9.4 mmol) and triphenylphosphine (2.90 g, 11.2 mmol) were added to a solution of the alcohol **54** (2.0 g, 9.4 mmol) in THF (25 mL) at 0 °C under an argon atmosphere. Diisopropyl azodicarboxylate (2.20 mL, 11.2 mmol) was slowly added at 0 °C and the

reaction mixture was warmed to room temperature and stirred for another 4 h. The solvent was removed under reduced pressure and the remaining residue was dissolved in CH_2Cl_2 (20 mL). The organic layer was washed with a saturated aqueous solution of $NaHCO_3$ (2 × 20 mL), brine (20 mL), dried over anhydrous Na_2SO_4, and concentrated under reduced pressure. Thus obtained residue was purified by column chromatography on silica gel using CH_2Cl_2:hexanes (1:4) to afford the dione **55** (2.4 g, 77% yield) as a white solid; mp 58-60 °C. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.79-7.76 (m, 2H), 7.69-7.66 (m, 2H), 7.34-7.31 (m, 2H), 6.08-6.04 (m,, 3H), 1.91 (d, *J* = 6.8 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 168.3, 151.8, 147.4, 143.1, 134.3, 132.5, 131.8, 123.5, 108.7, 105.4, 103.1, 46.4, 18.8. HRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{17}H_{12}N_2O_6Na$: 363.0588; found: 363.0592.

1-(6-Nitrobenzo[d][1,3]dioxol-5-yl)ethanamine (**56**). Hydrazine (0.550 mL, 17.6 mmol) was slowly added to a solution of the dione **55** (2.4 g, 7.0 mmol) in dry EtOH (30 mL) at room temperature under an argon atmosphere and the mixture was heated under reflux for an hour. The reaction mixture was then cooled to 0 °C, and diethyl ether (200 mL) was added under vigorous stirring to precipitate the byproduct 2,3-dihydrophthalazine-1,4-dione. The solid was filtered out, washed with diethyl ether (5 × 10 mL) to collect any residual product, and the solid was discarded. Then, the filtrate was washed with a saturated aqueous solution of $NaHCO_3$ (2 × 20 mL), brine (20 mL), and dried over anhydrous Na_2SO_4. The solvent was removed under reduced pressure and the obtained residue was purified by column chromatography on silica gel using CH_2Cl_2: MeOH (9:1) with 1% TEA to obtain the amine **56** (1.2 g, 81% yield) as a yellow solid; mp 98-100 °C. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.32

(s, 1H), 7.22 (s, 1H), 6.06 (dd, 2H), 4.63 (q, *J* = 6.6 Hz, 1H), 1.59 (s, 2H), 1.37 (d, *J* = 6.6 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 152.1, 146.5, 139.8, 106.5, 105.1, 102.9, 46.2, 24.8. HRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_9H_{11}N_2O_4$: 211.0713; found: 211.0717.

1-((2*R*,4*S*,5*R*)-4-((*tert*-Butyldimethylsilyl)oxy)-5-(((*tert* butyldimethylsilyl)oxy)methyl)tetrahydrofuran-2-yl)-4-((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethyl)amino)pyrimidin-2(1H)-one (**58**). DIPEA (0.72 mL, 4.4 mmol) and the amine **56** (0.58 g, 2.7 mmol) were added to a solution of the sulfonate **57** (1.0 g, 1.3 mmol) in dry DMF (5 mL) at room temperature under an argon atmosphere. The reaction mixture was heated to 90 °C overnight. The reaction mixture was cooled to room temperature, poured into a saturated aqueous solution of $NaHCO_3$ (30 mL), the layers were separated, and the aqueous layer was extracted using EtOAc (3 × 20 mL). The combined organic layers were washed with saturated aqueous solution of $NaHCO_3$ (3 × 10 mL), brine (10 mL), and dried over anhydrous Na_2SO_4. The solvent was removed under reduced pressure and the obtained residue was purified by column chromatography on silica gel using EtOAc with 1% TEA to obtain compound **58** (0.80 g, 89% yield) as a pale yellow solid; mp 125-127 °C. ^{1}H NMR (400 MHz, CD_3COCD_3): δ = 7.82 (dd, *J* = 24 Hz, 7.6 Hz, 1H), 7.58 (br, 1H), 7.43 (s, 1H), 7.14 (d, *J* = 12 Hz, 1H), 6.20-6.16 (m, overlap, 3H), 5.82 (dd, *J* = 24 Hz, 7.6 Hz, 1H), 5.73 (br, 1H), 4.47 (br, 1H), 3.90-3.83 (m, overlap, 3H), 2.31-1.99 (m, 2H), 1.53 (d, *J* = 6.8 Hz, 3H), 0.93-0.88 (m, 18 H), 0.12-0.08 (m, 12 H). ^{13}C NMR (100 MHz, CD_3COCD_3): δ = 163.7, 155.7, 155.6, 153.2, 153.1, 141.0, 139.2, 139.0, 106.9, 106.7, 105.6, 105.5, 104.2, 94.9, 94.8, 88.1, 86.3, 72.5, 72.3, 63.4, 63.3, 47.2, 47.0, 42.3, 42.2, 26.3, 26.1, 22.3, 18.9,

18.5, -4.4, -4.6, -5.2, -5.3. HRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{30}H_{49}N_4O_8Si_2$: 649.3083; found: 649.3081.

1-((2*R*,4*S*,5*R*)-4-Hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-4-((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethyl)amino)pyrimidin-2(1H)-one (**59**). TBAF (1.0 M in THF, 3.6 mL, 3.6 mmol) was slowly added to a solution of the compound **58** (0.80 g, 1.2 mmol) in dry THF (15 mL) at 0 °C. After 30 min, the ice bath was removed; the reaction mixture was warmed to room temperature and stirred for another 3 h. The solvent was removed under reduced pressure and the remaining residue was purified by column chromatography on silica gel using CH_2Cl_2: MeOH (9:1) with 1% TEA to yield compound **59** (0.48 g, 95% yield) as a white solid; mp 119-120 °C. ^{1}H NMR (300 MHz, DMSO-d_6): δ = 8.26-8.24 (m, 1H), 7.77 (d, *J* = 7.2 Hz, 1H), 7.54 (s, 1H), 7.09 (d, *J* = 3, 1H), 6.19-6.17 (m, 2H), 6.09-6.03 (m, 1H), 5.79 (d, *J* = 7.5 Hz, 1H), 5.55-5.49 (m, 1H), 5.19 (t, *J* = 3.6 Hz, 1H), 4.93 (t, *J* = 5.4 Hz, 1H), 4.16 (br, 1H), 3.74 (br, 1H), 3.51 (br, 2H), 2.12-2.01 (m, 1H), 1.94-1.82 (m, 1H), 1.45 (d, *J* = 6.8 Hz, 3H). ^{13}C NMR (100 MHz, DMSO-d_6): δ = 162.2, 154.6, 151.9, 146.4, 141.6, 140.5, 137.7, 105.7, 105.6, 104.6, 104.3, 94.3, 87.2, 84.9, 70.4, 70.3, 61.3, 45.3, 45.2, 21.9, 21.8. HRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{18}H_{21}N_4O_8$: 421.1354; found: 421.1358.

1-((2*R*,4*S*,5*R*)-5-((*bis*(4-Methoxyphenyl)(phenyl)methoxy)methyl)-4-hydroxytetrahydrofuran-2-yl)-4-((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethyl)amino)pyrimidin-2(1H)-one (**60**). DMTCl (0.19 g, 0.57 mmol) was added to a

solution of the compound **59** (0.20 g, 0.47 mmol) in dry pyridine (1.0 mL) at room temperature under an inert atmosphere. After 12 h, MeOH (0.2 mL) was added to the reaction mixture in order to quench the unreacted amount of DMTCl. The solvent was removed under reduced pressure and the remaining residue was dissolved in dichloromethane (15 ml). The organic layer was washed with 5% citric acid (2 × 5 mL), a saturated aqueous solution of $NaHCO_3$ (5 mL), brine (5 mL), and dried over anhydrous Na_2SO_4. The solvent was removed under reduced pressure and the obtained residue was purified by column chromatography on silica gel using CH_2Cl_2: MeOH (20:1) with 1% TEA to give compound **60** (0.26 g, 72% yield) as a white solid. ^{1}H NMR (300 MHz, CD_3COCD_3): δ = 8.02 (s, 1H), 7.81-7.76 (m, 1H), 7.52-7.43 (m, 3H), 7.34-28 (m, 6H), 7.24-7.11 (m, 2H), 6.89-6.86 (m, 4H), 6.24-6.15 (m, 3H), 5.74-5.65 (m, 2H), 4.50-4.47 (m, 2H), 4.01-3.99 (m, 1H), 3.78 (s, 6H), 3.35 (d, *J* = 2.7 Hz, 2H), 2.34-2.28 (m, 1H), 2.15-2.09 (m, 1H), 1.52 (d, *J* = 6.8 Hz, 3H). ^{13}C NMR (100 MHz, CD_3COCD_3): δ = 163.8, 159.6, 155.7, 153.2, 153.1, 147.6, 146.0, 143.2, 141.1, 139.1, 139.0, 136.7, 136.5, 131.0, 129.0, 128.6, 127.6, 113.9, 106.8, 106.7, 105.6, 104.2, 94.8, 87.2, 86.8, 86.4, 86.3, 79.2, 71.6, 71.5, 64.2, 55.5, 47.2, 47.0, 42.1, 42.0, 22.3. HRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{39}H_{39}N_4O_{10}$: 723.2661; found: 723.2658.

(2*R*,3*S*,5*R*)-2-((*bis*(4-Methoxyphenyl)(phenyl)methoxy)methyl)-5-(4-((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethyl)amino)-2-oxopyrimidin-1(2H)-yl)tetrahydrofuran-3-yl (2-cyanoethyl) diisopropylphosphoramidite (**61**). DIPEA (0.240 mL, 1.38 mmol) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (0.12 mL, 0.55 mmol) were added to a solution of compound **60** (0.20 g, 0.27 mmol) in dry dichloromethane (2 mL) at 0 °C under

an argon atmosphere. After 20 min, the ice bath was removed and the reaction mixture was warmed to room temperature and stirred for another 2 h. The solvent was removed under reduced pressure and the remaining residue was purified by column chromatography on silica gel using $CHCl_3$: CH_3COCH_3 (20:1) with 1% TEA to afford a white solid. The white solid was dissolved in dichloromethane (0.5 mL) and precipitated in diethyl ether (20 mL) under vigorous stirring in order to remove an excess of chlorophosphoramidite reagent that was co-eluted with the desired product. The precipitate was collected through filtration to obtain the phosphoramidite **61** (0.11 g, 45% yield) as an off-white solid. ^{1}H NMR (400 MHz, CD_3COCD_3): δ = 7.84-7.78 (m, 1H), 7.48-7.44 (m, 4H), 7.36-7.32 (m, 6H), 7.24-7.22 (m, 1H), 7.12 (d, *J* = 12 Hz, 1H), 6.90-6.87 (m, 4H), 6.27-6.16 (m, 3H), 5.74-5.71 (m, 1H), 5. 68-5.66 (m, 1H), 4.65 (br, 1H), 4.14-4.09 (m, 1H), 3.70 (br, 6H), 3.72-3.54 (m, 2H), 3.42-3.35 (m, 2H), 2.76-272 (m, 1H), 2.63-2.60 (m, 1H), 2.52-2.39 (m, 1H), 2.28-2.16 (m, 1H), 1.53 (d, *J* = 6.8 Hz, 3H), 1.27-0.83 (m, 12 H). ^{13}C NMR (100 MHz, CD_3COCD_3): δ = 163.7, 159.7, 155.4, 151.7, 145.9, 141.1, 136.5, 131.0, 129.0, 128.7, 127.7, 114.0, 106.8, 106.7, 105.6, 104.3, 94.9, 87.3, 86.4, 63.9, 59.6, 59.4, 55.5, 47.0, 46.9, 43.9, 43.8, 41.0, 24.9, 24.8, 22.2, 20.7. ^{13}P NMR (121 MHz, CD_3COCD_3, reference H_3PO_4 = 0.00 ppm): δ = 146.55, 146.50 146.39, 146.34.

5-Methoxy-2-nitro-4-(prop-2-yn-1-yloxy)benzaldehyde (**63**). The benzaldehyde **62** (3.00 g, 15.7 mmol) was added to ice-cold, concentrated HNO_3 (30 mL) and the reaction mixture was stirred for another 4 h at 0 °C. Then, the reaction mixture warmed to room temperature and was stirred overnight. The reaction mixture was poured into ice-cold water (200 mL)

under vigorous stirring and the obtained yellow precipitate was collected through filtration. The precipitate was re-dissolved in DCM (50 mL) and was washed with a saturated aqueous solution of $NaHCO_3$ (20 mL), brine (20 mL), and dried over anhydrous sodium sulfate. After filtration, the solvent was removed under reduced pressure and the remaining residue was purified by silica gel chromatography using DCM/hexanes (7:3) to deliver the corresponding nitrobenzaldehyde **63** (2.56 g, 71% yield) as a yellow solid; mp 120-123 °C. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 10.39 (s, 1H), 7.73 (s, 1H), 7.17 (s, 1H), 4.85 (s, 2H), 3.96 (s, 3H), 2.57 (s, 1H). ^{13}C NMR (75 MHz, DMSO-d6): δ = 189.2, 153.7, 150.0, 143.7, 126.3, 111.0, 109.8, 80.3, 78.5, 57.1. HRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{11}H_{10}NO_5$: 236.0553; found 236.0063.

1-(5-Methoxy-2-nitro-4-(prop-2-yn-1-yloxy)phenyl)ethanol (**64**). Trimethylaluminium (12.6 mmol, 2.0 M in hexane) was slowly added to a solution of the nitrobenzaldehyde **63** (1.9 g, 8.0 mmol) in DCM (30 mL) at 0 °C under an argon atmosphere. After stirring for 2 h at that temperature, ice cold water (5 mL) was slowly added to the reaction mixture in order to quench the unreacted amount of the reagent. Then, aqueous NaOH (40 mL, 1 M) was added and the reaction mixture was continuously stirred for another 1 h. The organic layer was separated and washed with an aqueous solution of NaOH (1 M, 2 × 20 mL), brine (20 mL), and dried over anhydrous sodium sulfate. The filtrate was concentrated under reduced pressure and the remaining residue was purified by silica gel chromatography using $CHCl_3$:CH_3COCH_3 (9:1) to furnish the alcohol **64** (2.4 g, 70% yield) as a yellow solid; mp 125-127 °C. ^{1}H NMR (300 MHz, DMSO- d_6): δ = 7.66 (s, 1H), 7.38 (s, 1H), 5.509 (d, *J* = 2.1 Hz, 2H), 5.28-5.25 (m, 1H), 4.91 (d, *J* = 1.2 Hz, 2H), 3.91 (s, 3H), 3.65-3.63 (s, 1H), 1.38 (t,

J = 4.5 Hz, 3H). ^{13}C NMR (75 MHz, acetone-d_6): δ = 154.6, 145.4, 139.3, 124.2, 110.4, 109.6, 78.4, 77.2, 65.0, 55.9, 24.8. HRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{12}H_{13}NO_5Na$: 274.0686; found 274.0731.

((1-(5-Methoxy-2-nitro-4-(prop-2-yn-1-yloxy)phenyl)ethoxy)methyl)(methyl)sulfane (**65**). Methyl sulfide (4.70 mL, 63.3 mmol) was added to a solution of the alcohol **64** (1.60 g, 6.36 mmol) in a dry acetonitrile (20 mL) under a nitrogen atmosphere at 0 °C. Benzoyl peroxide (6.10 g, 25.4 mmol) was added in 4 portions to the reaction mixture with an interval of 30 min each. After adding of the last portion of benzoyl peroxide, the mixture was stirred at the same temperature for another 3 h. An aqueous solution of NaOH (100 mL, 1 M) was added to the reaction mixture and stirring was continued overnight. The aqueous layer was extracted with ether (3 × 30 mL) and the combined ether layers were washed with an aqueous solution (1 M) of NaOH (2 × 20 mL), brine (20 mL), and dried over anhydrous sodium sulfate. The solvent was removed under reduced pressure and the remaining residue was purified by silica gel chromatography using EtOAc:hexanes (1:9) to obtain **65** (1.5 g, 78% yield) as a brown solid; mp 76-60 °C. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.74 (s, 1H), 7.21 (s, 1H), 5.54 (q, *J* = 6.3 Hz, 1H), 4.81 (d, *J* = 1.6 Hz, 2H), 4.46 (d, *J* = 6.3 Hz, 1H), 4.34 (d, *J* = 6.3 Hz, 1H), 3.97 (s, 3H), 2.57 (t, *J* = 2.4 Hz, 1H), 2.14 (s, 3H), 1.51 (d, *J* = 6.3 Hz, 3H).^{13}C NMR (75 MHz, $CDCl_3$): δ = 154.5, 145.6, 140.4, 136.0, 110.5, 109.2, 77.4, 77.2, 73.5, 70.9, 57.2, 56.6, 23.5, 14.3. HRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{14}H_{17}NO_5SNa$: 334.0720; found 334.0789.

1-((2*R*,4*S*,5*R*)-5-((*bis*(4-Methoxyphenyl)(phenyl)methoxy)methyl)-4-hydroxytetrahydrofuran-2-yl)-3-((1-(5-methoxy-2-nitro-4-(prop-2-yn-1-yloxy)phenyl)ethoxy)methyl)-5-methylpyrimidine-2,4(1H,3H)-dione (**66**). Sulfuryl chloride (22 µL, 0.26 mmol) was slowly added to a solution of the thioether **65** (75 mg, 0.24 mmol) in DCM (0.4 mL) stirred at –78 °C under an argon atmosphere. After 20 min, the solvent was removed under high vacuum at -78 °C for about 15 minutes and the reaction mixture was warmed to room temperature, followed by continued drying for another 5 h. The dry residue was dissolved in dry DMF (0.5 mL) and added to a solution containing the DMT-protected thymidine[267] (130 mg, 0.240 mmol) and DBU (107 µL, 0.720 mmol) in DMF (0.5 mL) at room temperature under an inert atmosphere. Stirring was continued for 12 h and the reaction mixture was poured into a saturated aqueous solution of $NaHCO_3$ (10 mL). The layers were separated and the aqueous layer was extracted with DCM (2 × 10 mL) and the combined organic layers were washed with a saturated aqueous solution of $NaHCO_3$ (2 × 10 mL), brine (10 mL), and dried over anhydrous sodium sulfate. The solvent was evaporated under reduced pressure and the remaining residue was purified by silica gel chromatography using DCM:acetone (9:1) with 1% TEA to provide **66** (1.5 g, 78% yield) as a brown solid; mp 93-95 °C. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.70 (s, 0.5H), 7.67 (s, 0.5H), 7.23-7.43 (m, 13 H), 6.82-6.85 (m, 4H), 6.29 (t, *J* = 6.3 Hz, 0.5H), 6.20 (t, *J* = 6.6 Hz, 0.5H), 5.38-5.49 (m, 2H), 5.26-5.29 (m, 1H), 4.74 (d, *J* = 4.0 Hz, 2 H), 4.51 (br, 1 H), 3.99-4.04 (m, 1H), 3.78 (s, 6H), 3.96 (s, 3H), 3.30-3.48 (m, 2H), 2.57 (t, *J* = 2.0 Hz, 0.5H), 2.52 (t, *J* = 2.0 Hz, 0.5H), 2.32-2.41 (m, 2H), 2.27-2.28 (m, 0.5H), 2.10-2.21 (m, 1.5H), 1.47-1.53 (m, 3H), 1.39 (s, 1.5H), 1.36 (s, 1.5 H).^{13}CNMR (75 MHz, $CDCl_3$): δ = 163.3, 158.9, 154.3, 154.0, 151.0,

150.8, 145.4, 144.4, 140.0, 139.7, 137.1, 136.8, 135.6, 135.5, 130.2, 128.3, 127.4, 113.5, 110.5, 110.3, 110.1, 109.7, 109.4, 87.1, 86.1, 85.5, 85.2, 73.6, 72.5, 72.2, 70.2, 63.6, 57.2, 56.9, 56.7, 55.5, 41.2, 24.1, 12.5. LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{44}H_{45}N_3O_{12}Na$: 830.2901; found 830.29.

(2*R*,3*S*,5*R*)-2-((*bis*(4-Methoxyphenyl)(phenyl)methoxy)methyl)-5-(3-((1-(5-methoxy-2-nitro-4-(prop-2-yn-1-yloxy)phenyl)ethoxy)methyl)-5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-3-yl (2-cyanoethyl) diisopropylphosphoramidite (**67**). DIPEA (43 µL, 0.24 mmol) and 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (28 µL, 0.12 mmol) were added to a solution of the alcohol **66** (35 mg, 0.04 mmol) in DCM (1.0 mL) at room temperature under an argon atmosphere and stirring was continued for 2 h. The solvent was evaporated under reduced pressure and the remaining residue was purified by silica gel chromatography using $CHCl_3$:acetone (20:1) containing 1% TEA to furnish the caged phosphoramidite **67** (39 mg, 90% yield) as a white solid; mp 85-87 °C. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.65-7.68 (m, 1H), 7.20-7.50 (m, 11H), 6.81-6.85 (m, 4H), 6.14-6.29 (m, 1 H), 5.37-5.47 (m, 2H), 5.23-5.31 (m, 1H), 4.72-4.77 (m, 2H), 4.54-4.65 (m, 1H), 4.07-4.19 (m, 1H), 3.93 (s, 3H), 3.71-3.89 (m, 7H), 3.25-3.66 (m, 5H), 2.37-2.71 (m, 4H), 2.08-2.26 (m, 1H), 1.50-1.54 (m, 3H), 1.30-1.36 (m, 3H), 1.14-1.18 (m, 9 H), 1.04 (d, *J* = 6.6 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 163.1. 158.7, 154.1, 154.0, 153.8, 150.8, 150.7, 145.2, 144.2, 140.1, 139.9, 139.8, 139.6, 139.6, 137.0, 136.8, 136.6, 135.3, 134.4, 130.1, 128.2, 127.2, 117.8, 117.5, 113.3, 110.4, 110.2, 110.1, 109.9, 109.5, 109.2, 86.9, 85.2, 74.0, 73.85, 73.6, 70.2, 63.3, 58.8, 58.5, 58.2,

57.9, 57.1, 57.0, 56.8, 56.5, 55.3, 43.4, 43.2, 40.1, 24.7, 23.9, 20.5, 20.4, 12.3. ^{31}P NMR (121 MHz, $CDCl_3$): δ = 149.1, 149.0, 148.6, 148.5.

1-((4a*R*,6*R*,7*R*,7a*S*)-2,2-di-*tert*-Butyl-7-hydroxytetrahydro-4H-furo[3,2-d][1,3,2]dioxasilin-6-yl)-3-((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)methyl)pyrimidine-2,4(1H,3H)-dione (**73**). DBU (697 µL, 4.50 mmol) was added to a solution of the silyl protected uridine **71** (0.88 g, 2.3 mmol) in DMF (5 mL) at room temperature under an inert atmosphere. After 30 minutes, freshly prepared NPOM chloride[264] **72** (2.76 mmol) dissolved in DMF (0.5 mL) was added to the reaction mixture and stirring was continued for 12 h. The reaction mixture was poured into a saturated solution of $NaHCO_3$ (20 mL) and the layers were separated. The aqueous layer was extracted using EtOAc (2 × 20 mL) and the combined organic layers were washed with a saturated aqueous solution of $NaHCO_3$ (3 × 20 mL), brine (15 mL), and dried over anhydrous sodium sulfate. The filtrate was concentrated and the remaining crude product was purified by silica gel chromatography using EtOAc:hexanes(1:1) with 1% TEA to obtain the NPOM-caged uridine **73** (0.95 g, 68% yield) as a yellow solid; mp 96-98 °C. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.41-7.43 (m, 1H), 7.15-7.17 (m, 1H), 7.08-7.13 (q, 1H), 6.06-6.07 (m, 2H), 5.66-5.71 (q, 1H), 5.48 (s, 0.5 H), 5.42 (s, 0.5 H), 5.21-5.33 (m, 3H), 5.09-5.12 (m, 1H), 4.27-4.45 (m, 2H), 3.93-4.15 (m, 4H), 1.48 (d, J = 4.0 Hz, 3H), 1.00-1.07 (m, 18H). ^{13}C NMR (100 MHz, DMSO- d_6): δ = 162.6, 152.6, 142.8, 140.6, 139.8, 137.8, 106.8, 105.4, 102.4, 102.5, 96.0, 95.3, 76.2, 75.2, 73.5, 69.3, 67.6, 27.8, 27.7, 24.1, 23.2, 20.8. HRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{27}H_{37}N_3O_{11}SiNa$: 630. 2090; found 630. 2088.

1-((2*R*,3*R*,4*R*,5*R*)-4-Hydroxy-5-(hydroxymethyl)-3-methoxytetrahydrofuran-2-yl)-3-((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)methyl)pyrimidine-2,4(1H,3H)-dione (**74**). NPOM-caged uridine **73** (2.1 g, 3.4 mmol) was added to methyl iodide (20 mL) stirred at room temperature under an argon atmosphere. Ag_2O (2.30 g, 10.3 mmol) was added and the reaction mixture was heated under reflux (50 °C) for 5 h. The reaction mixture was cooled to room temperature and EtOAc (20 mL) was added. Then, the reaction mixture was filtered, the residue was washed with EtOAc (10 mL), and the filtrate was collected. The filtrate was washed with a saturated aqueous solution of $NaHCO_3$ (3 × 20 mL), brine (20 mL), and dried over anhydrous sodium sulfate. The filtrate was concentrated and the obtained crude product was purified by silica gel column chromatography using EtOAc:hexanes (2:3) with 1% TEA to obtain the 2′OMe uridine **74** (2.1 g, 98% yield) as a white solid; mp 125-127 °C. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.42 (d, *J* = 1.0 Hz, 1H), 7.17 (d, *J* = 2.6 Hz, 1H), 7.11-7.15 (m, 1H), 6.06 (m, 2H), 5.66-5.70 (m, 1H), 5.54 (d, *J* = 1.6 Hz, 1H), 5.27-5.34 (m, 2H), 5.11-5.17 (m, 2H), 4.42-4.47 (m, 1H), 4.03-4.12 (m, 2H), 3.80-3.97 (m, 3H), 3.63 (s, 3H), 1.49 (d, *J* = 3.0 Hz, 3H), 1.00-1.04 (m, 18H). ^{13}C NMR (100 MHz, acetone- d_6): δ = 162.7, 162.6, 153.1, 151.3, 148.0, 142.8, 140.4, 140.0, 138.7, 106.9, 105.2, 104.3, 102.1, 92.7, 92.2, 82.8, 77.6, 75.3, 74.2, 70.4, 67.8, 59.2, 27.7, 20.9. HRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{28}H_{39}N_3O_{11}SiNa$: 644.2246; found 644.2248.

1-((2*R*,3*R*,4*R*,5*R*)-4-Hydroxy-5-(hydroxymethyl)-3-methoxytetrahydrofuran-2-yl)-3-((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)methyl)pyrimidine-2,4(1H,3H)-dione (75). HF-pyridine (330 µL, 70% HF:pyridine) diluted in dry pyridine (2 mL) was added to a solution

of the 2′OMe uridine **74** (2.0 g, 3.21 mmol) in DCM at 0 °C under a nitrogen atmosphere in a polyethylene vessel (since HF vigorously reacts with glass, avoid using glassware). After stirring the reaction mixture for 2 hours at 0 °C, DCM (40 mL) was added followed by a saturated aqueous solution of $NaHCO_3$ (30 mL) to neutralize residual amounts of HF. The organic layer was separated and washed with a saturated aqueous solution of $NaHCO_3$ (30 mL), 1 M HCl (2 × 20 mL), water (20 mL), and brine (15 mL), and was dried over anhydrous sodium sulfate. The filtrate was concentrated under reduced pressure and the remaining crude product was purified by silica gel column chromatography using DCM:MeOH (92:8) containing 1% TEA to obtain the caged 2′-OMe uridine **75** (1.4 g) as a yellow solid in 98% yield; mp 65-68 °C. ^{1}H NMR (400 MHz, acetone- d_6): δ = 8.00-8.05 (q, 1H), 7.41 (s, 1H), 7.13 (d, *J* = 6.4 Hz, 1H), 6.18-6.21 (m, 2H), 5.84 (d, *J* = 1.5 Hz, 0.5H), 5.74 (d, *J* = 1.3 Hz, 0.5H), 5.17-5.49 (m, 4H), 4.37 (t, *J* = 4.5 Hz, 1H), 4.20-4.25 (m, 1H), 3.89-3.99 (m, 3H), 3.77-3.84 (m, 2H), 3.54 (s, 1.5H), 3.50 (s, 1.5H), 1.45 (d, *J* = 3.1 Hz, 3H). ^{13}C NMR (75 MHz, acetone- d_6): δ = 162.9, 162.8, 153.1, 152.9, 151.7, 148.1, 148.0, 140.1, 140.0, 138.8, 106.9, 105.2, 104.4, 104.3, 101.5, 101.3, 88.7, 85.7, 85.4, 84.8, 84.6, 74.7, 74.3, 70.5, 70.3, 69.3, 68.9, 61.2, 60.7, 58.6, 24.2, 24.1. HRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{20}H_{23}N_3O_{11}Na$: 504.1225; found 504.1226.

1-((2*R*,3*R*,4*R*,5*R*)-5-((*bis*(4-Methoxyphenyl)(phenyl)methoxy)methyl)-4-hydroxy-3-methoxytetrahydrofuran-2-yl)-3-((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)methyl)pyrimidine-2,4(1H,3H)-dione (76). DMTrCl (1.5 g, 4.6 mmol) was added to a solution of the alcohol **75** (1.4 g, 2.9 mmol) in dry pyridine (15 mL) at room

temperature under an argon atmosphere and stirring was continued for 24 h. MeOH (3 mL) was added to the reaction mixture in order to quench the unreacted amount of DMTrCl. After 20 minutes, the solvent was removed under reduced pressure and the obtained crude product was redissolved in EtOAc (40 mL). The organic layer was washed with a 5% aqueous citric acid solution (2 × 20 mL), a saturated aqueous solution of $NaHCO_3$ (2 × 20 mL), and brine (15 mL), and dried over anhydrous sodium sulfate. The filtrate was concentrated and the remaining crude product was purified silica gel column chromatography using DCM:EtOAc (1:5) with 1% TEA to furnish the DMT protected caged 2′-OMe uridine **76** (1.9 g) as a white solid in 86% yield. ^{1}H NMR (300 MHz, acetone- d_6): δ = 7.92 (d, *J* = 3.9 Hz, 0.5H), 7.80 (d, *J* = 4.2 Hz, 0.5H), 7.43-7.48 (m, 2H), 7.26-7.38 (m, 6H), 7.12-7.20 (m, 2H), 6.90-6.94 (m, 4H), 6.12-6.22 (m, 2H), 5.82 (s, 0.5H), 5.70 (s, 0.5H), 5.40-5.47 (m, 1H), 5.05-5.32 (m, 2H), 4.36-4.43 (m, 1H), 4.02-4.11 (m, 1H), 3.77-3.81 (m, 6H), 3.60 (s, 1.5H), 3.56 (s, 1.5H), 3.46-3.52 (m, 2H), 1.44-1.47 (m, 3H). ^{13}C NMR (75 MHz, acetone- d_6): δ = 162.7, 159.8, 152.8, 151.4, 148.1, 245.8, 145.7, 139.6, 139.5, 139.09, 136.5, 136.3, 131.1, 131.0, 130.1, 129.1, 128.8, 128.3, 127.9, 106.9, 105.2, 104.4, 101.5, 101.4, 89.1, 88.9, 87.6, 84.8, 84.6, 83.8, 83.5, 74.8, 74.4, 70.6, 70.3, 69.7, 69.6, 69.4, 69.3, 62.9, 62.2, 58.8, 55.6, 24.3, 24.1. HRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{41}H_{41}N_3O_{13}Na$: 806.2532; found 806.2526.

.

(2*R*,3*R*,4*R*,5*R*)-2-((*bis*(4-Methoxyphenyl)(phenyl)methoxy)methyl)-4-methoxy-5-(3-((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)methyl)-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-3-yl (2-cyanoethyl) diisopropylphosphoramidite (77). DIPEA (0.55 mL, 3.2 mmol) and 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.284 mL, 1.27

mmol) were added to a solution of the alcohol **76** (0.50 g, 0.64 mmol) in DCM (5.0 mL) at room temperature under an argon atmosphere. After 2 hours of stirring, the reaction mixture was concentrated and the remaining crude product was purified by silica gel chromatography using EtOAc:hexanes (4:6) with 1% TEA to furnish the caged phosphoramidite **77** (0.5 g) as a white solid in 80 % yield. ^{1}H NMR (300 MHz, acetone- d_6): δ = 7.76-7.95 (m, 2H), 7.27-7.50 (m, 9H), 7.11-7.16 (m, 1H), 6.90-6.95 (m, 4H), 5.85-5.88 (m, 1H), 5.75-5.77 (m, 1H), 5.44 (d, J = 5.0 Hz, 1H), 5.04-5.31 (m, 4H), 4.57-4.66 (m, 1H), 4.46-4.54 (m, 1H), 4.20-4.27 (m, 1H), 3.91-4.01 (m, 2H), 3.80 (s, 6H), 3.61-3.74 (m, 3H), 3.48-3.58 (m, 5H), 2.607-2.65 (m, 1H), 1.45 (d, J = 3.2 Hz, 3H), 1.16-1.22 (m, 9H), 1.06-1.11 (m, 3H). ^{13}C NMR (100 MHz, acetone- d_6): δ = 162.7, 162.6, 159.9, 159.8, 153.1, 151.8, 151.5, 148.0, 145.7, 139.5, 139.4, 138.9, 138.7, 136.4, 136.2, 136.1, 131.2, 129.2, 128.8, 127.9, 114.1, 107.0, 106.9, 105.3, 105.2, 104.4, 104.2, 101.7, 101.6, 89.7, 89.5, 89.3, 87.7, 84.2, 83.4, 83.1, 82.8, 74.4, 74.2, 71.1, 71.0, 70.8, 70.2, 62.7, 62.4, 62.2, 61.9, 59.8, 59.6, 59.4, 59.1, 59.0, 58.9, 58.5, 55.6, 44.1, 43.9, 25.1, 24.9, 24.8, 24.1, 20.9, 20.8. ^{31}P NMR (161 MHz, acetone- d_6): δ = 150.8, 150.7, 150.3, 150.1.

CHAPTER 3: PHOTOSWITCHABLE AZOBENZENE NUCLEOSIDES

3. Controlled gene regulations with an azobenzene photoswitch

Azobenzene, a diphenyl substituted diazene chromophore **78**, has been explored as a photoswitchable molecular probe for biological studies. The azo compounds were discovered in 1800s as synthetic dyes, and their potential for biophysical applications is currently emerging.[268] The interest in these molecules is mainly due to their existence in a planar thermally stable *trans* (*E*) form **78** and a compact meta-stable *cis* (*Z*) conformation **79** (Figure 22a).[269] The *trans* conformation of azobenzene itself is about 10–12 kcal mol^{-1} more stable than its *cis* isomer.[268]

The *trans–cis* isomerization of azobenzene is tunable and can be achieved by exposing the *trans* isomer to UV light (365 nm) to produce the *cis* isomer and the opposite isomerization can be accomplished by using either blue light or by allowing the *cis* isomer to relax back to the thermally stable *trans* state (Figure 22b).[270] A number of theories that govern *trans-cis* isomerization have been proposed over time. However, the basis of the origin of the driving force for the molecular motion has not fully been understood and is still under investigation.[269, 270]

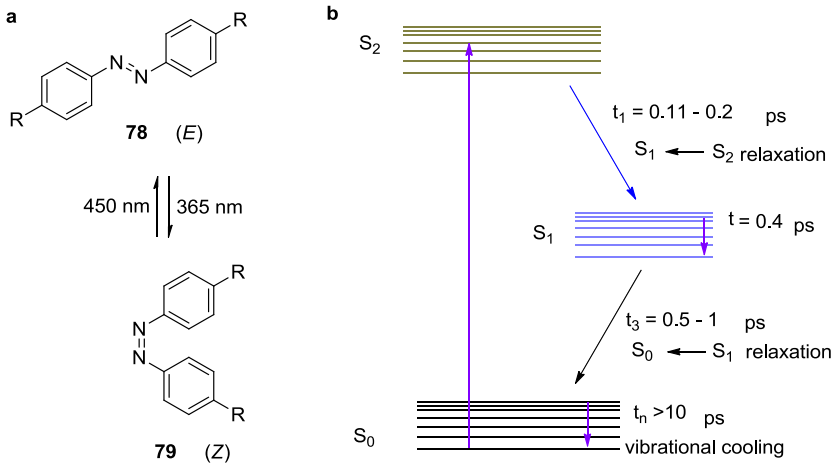


Figure 22. Photophysical properties of azobenzene: a) a *trans–cis* photoisomerization of azobenzene: b) Jablonski diagram showing excitation and relaxation with corresponding time (τ); S_0 – ground state, S_1 – first singlet excited state, S_2 – second singlet excited state; ps - picosecond.

Upon UV irradiation *trans* azobenzene is excited to singlet excited states from the ground state S_0 (Figure 22b). The *trans* → *cis* isomerization occurs during the excitation processes either via *N–N* bond rotation, inversion, concerted inversion or inversion-assisted rotation as depicted in Figure 23. The UV excitation of *trans* azobenzene to the S_2 state would cleave the *N–N* π bond that leads to the rotational isomerization followed by rapid crossover to the first excited state S_1. An inversion isomerization occurs in the S_1 state followed by $S_1 \rightarrow S_0$ decay that takes place via either concerted inversion or a rotational isomerization pathway. It has also been proposed that UV irradiation of *trans* azobenzene, in

addition to S_1 and S_2 excited states, generates S_3 and S_4 excited states that undergo relaxation without isomerization resulting in lower quantum yield.[271]

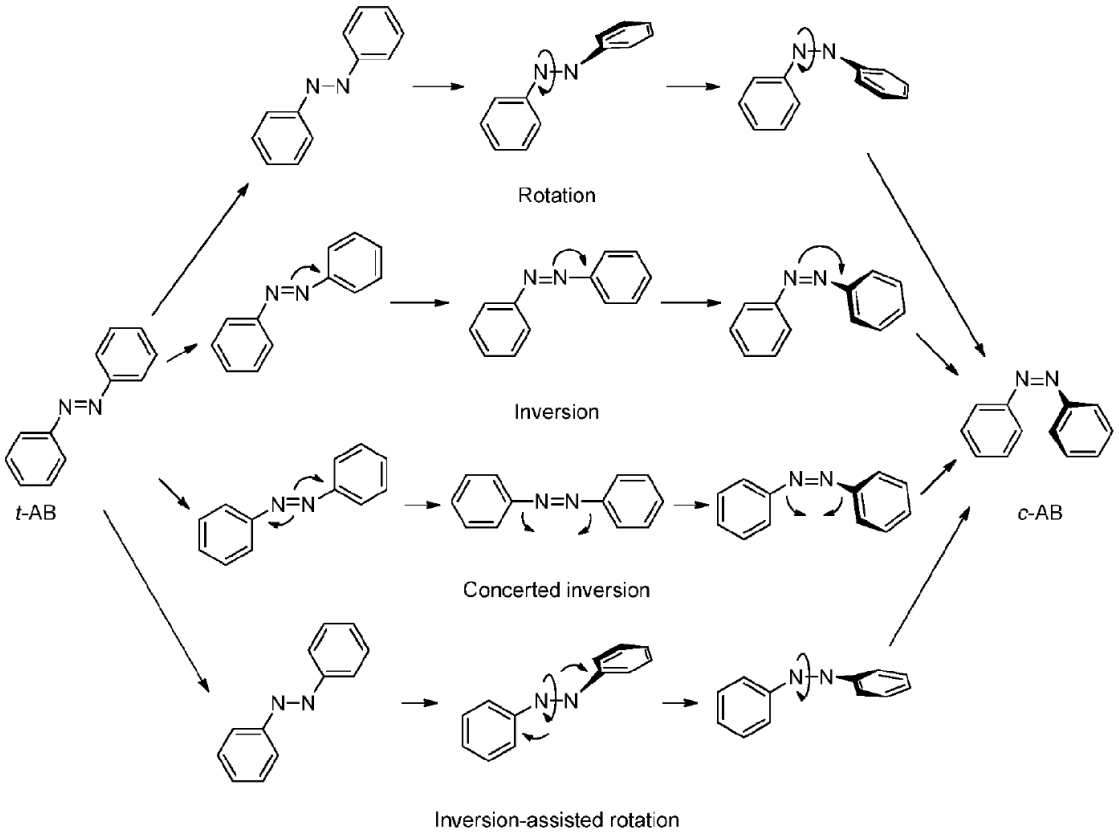


Figure 23. Proposed mechanisms of *trans* → *cis* isomerization of azobenzene. Reproduced with permission from *Chem. Soc. Rev.* **2012**, *41*, 1809-25.

Current studies on azobenzene are based on the photo-mechanical properties including light driven reversible *trans–cis* isomerization of chromophores for photo switchable devices.[272, 273] The widespread use of azobenzene to induce or regulate biochemical functions in response to light is based on the intrinsic properties of the

azobenzene chromophore.[269, 274, 275] These properties include: 1) the environment has a relatively small influence on the chromophore, e.g., as the influence of solvent polarity or viscosity on absorption and isomerization is small, 2) the chromophore possesses an outstanding photostability and undergoes minimal photobleaching, 3) the process results in high quantum yields (e.g., *trans* → *cis* up to 0.31 and *cis* → *trans* up to 0.41 in CH_3CN) with a minimal thermal effect, 4) the process can be optimized to achieve a large *cis–trans* population ratio, and 5) the isomerization can furnish ultrafast photoswitching in the order of picoseconds.

There has been a significant progress on azobenzene photoswitches for biomolecules[268, 269, 276, 277] since the first report on the isomerization of embedded azobenzene in a model membrane system.[278] These current studies on azobenzene include, but are not limited to, biological areas such as nucleic acid, peptide, proteins, and lipids.[270, 279-281] Generally, two approaches have been used to incorporate the azobenzene functionality in nucleic acids through covalent linkage.[282-284] First, azobenzenes **80**–**82** (Figure 24) can be installed into the nucleic acid backbone via linkages such as D-threoniol.[285-287] Second, properly functionalized azobenzene (for instance, with an alkyne or azide handle), may site-specifically be installed into nucleic acids through bioconjugation reactions.[282] Additionally, cationic azobenzenes have been used to manipulate nucleic acid properties via ionic interactions.[288]

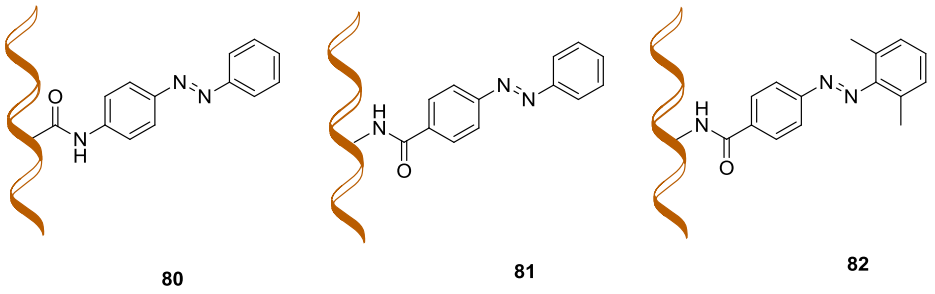


Figure 24. A cartoon representation of DNAs labeled with an azobenzene **80**–**82**.

Site-specific incorporation of the azobenzene moiety into DNA or RNA often changes the topology of the helix, thereby altering the nucleic acid binding abilities, which may be utilized for the photoregulation of gene function if the azobenzene moiety is incorporated into an antisense agent.[282, 283, 289] Investigations have shown that installation of an azobenzene moiety into a DNA oligonucleotide drastically changes its binding ability with a DNA or an RNA partner to form a duplex.[285, 290] For example, when the azo moiety in DNA adopts a *trans* form, it stabilizes its DNA binding partners through π stacking interactions with nucleo- bases.[290] On the other hand, when, the duplex is irradiated with UV light, the azo moiety changes into the *cis* form and destabilizes the duplex (Figure 25a).[290] In this context, azobenzene may be used as a photoswitch to regulate gene function. For instance, when an azo moiety is incorporated into an antisense agent, the RNase H could neither recognize the duplex nor digest it due to its unusal topology. Upon a brief UV irradiation, the *trans* azo moiety isomerizes into its *cis* form thereby destabilizing the duplex and allowing RNase H to cleave it. This notion has been validated to study the photoregulation, e.g., visible light for deactivation and UV light for activation, of RNase H

(Figure 25b).[291-294] However, activity and photoswitching of these reagents has not been demonstrated in cell culture, presumably because of the need for extended UV irradiation of approximately 10 minutes to achieve optimum isomerization.

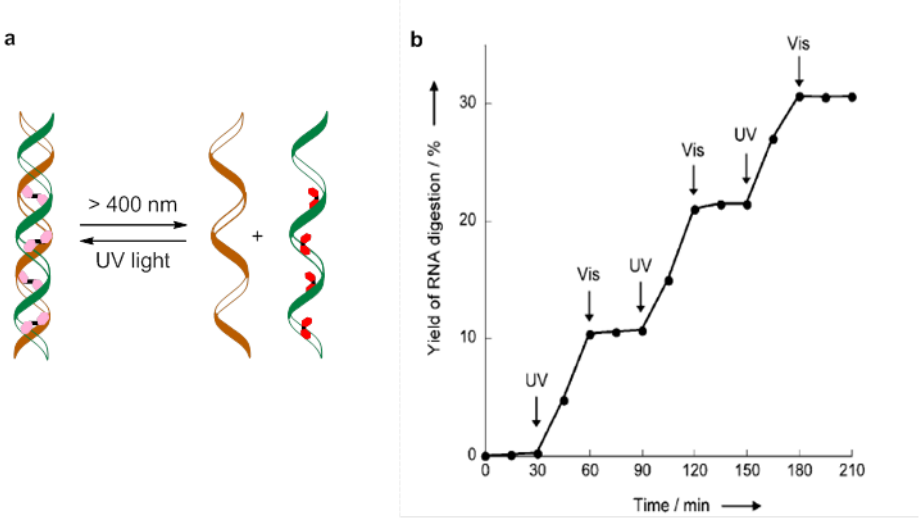


Figure 25. Photoregulation of RNase H activity using azo-antisense agent: a) an illustration of reversible duplex formation and b) RNA digestion by RNase H. Adapted with permission from *Angew. Chem. Int. Ed.* **2010**, *49*, 2167-70.

3.1. Light-activation of a diazobenzene-containing antisense agent

Here, we functionalized both ends of the diazobenzene moiety to explore its optimum effect during photoisomerization. Upon its incorporation into an oligonucleotide, the photo–mediated *cis-trans* isomerization may completely alter the base pairing landscape resulting in reversible duplex formation (Figure 26).[295] We hypothesized that this approach can address the potential *in vivo* problems of the previously discussed DNA photoswitching approach.

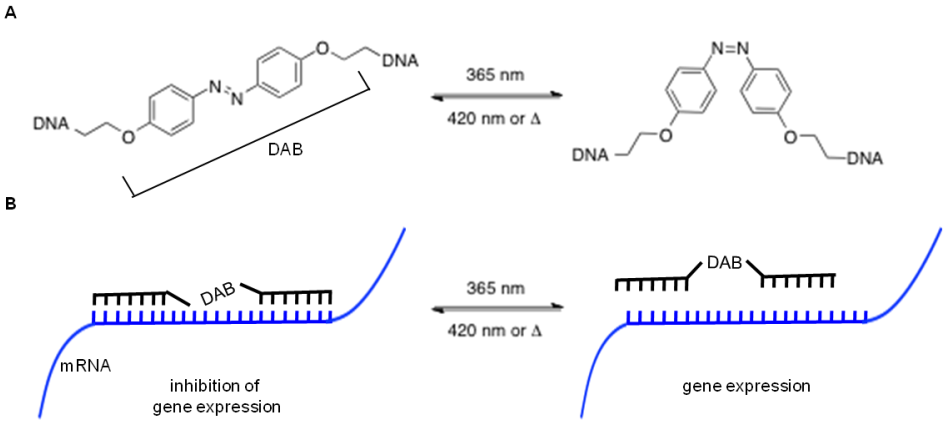
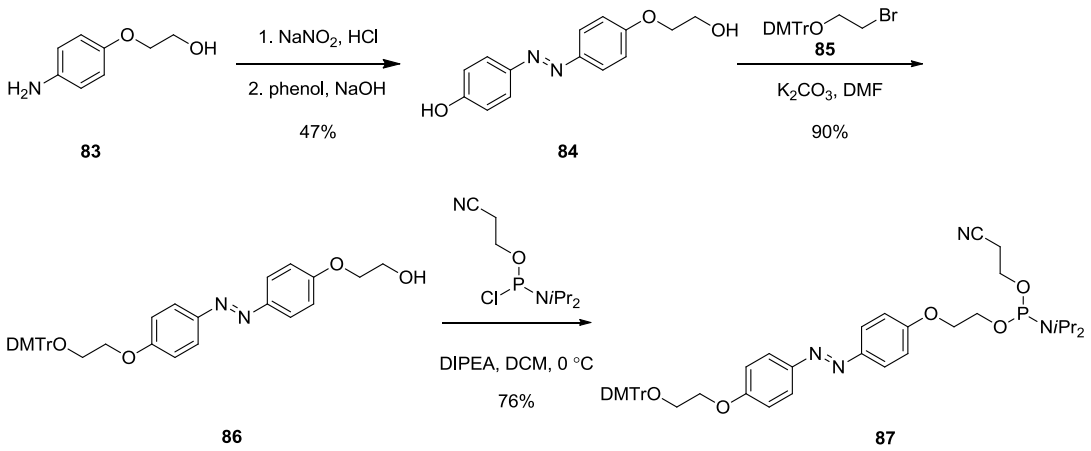


Figure 26. Light-mediated reversible gene regulation using an azo oligonucleotide: A) *trans* - *cis* diazobenzene (DAB) antisense agent, B) *trans* DAB tightly binds with mRNA (inhibition of gene expression), and *cis*-DAB loosely binds with mRNA (gene expression).

A azobenzene phosphoramidite **87** was synthesized (Scheme 9) to develop an antisense oligonucleotide with an azobenzene as a photoswitch. The phosphoramidite **87** was synthesized in three steps from alcohol **83**, which was synthesized in two steps from the commercial *para*-nitrophenol following literature procedures.[296, 297] The alcohol **83** was then treated with $NaNO_2$ in the presence of aqueous hydrochloric acid for *in situ* generation of diazonium salt which upon treating with phenol under an alkaline condition at 0 °C gave the azobenzene phenol **84** in 47% yield. All attempts in altering reaction time, acid ratio, temperature (from 0 °C to room temperature) and equivalents of phenol did not improve the yield. The phenolic hydroxyl group of the alcohol **84** was selectively reacted with the bromide **85** in the presence of K_2CO_3 in DMF to obtain the azobenzene alcohol **86** in 90% yield. The bromide **85** was prepared from commercially available bromoethanol in a single

step by following a reported procedure.[298] Finally, the alcohol **86** was activated with chlorophosphoramidite in the presence of DIPEA in DCM to obtain the desired product **87** in 76% yield.



Scheme 9. Synthesis of an azobenzene phosphoramidite **87**.

After the synthesis of the azobenzene phosphoramidite **87**, various azobenzene-containing DNA oligonucleotides (DAB) and spacer DNA strands were prepared in colabration with the Sintim Laboratory at the University of Maryland, College Park. The hybridization ability of the oligonucleotides was screened by Jeane Govan in the Deiters laboratory. The spacer DNA strands contain a linker (varying with 0-4 cytidine nucleotides) so that their remaining sequences on either side are complementary to the DAB oligonucleotide. The DAB DNA is annealed with the spacer DNA in the presence of ethidium bromide, a DNA intercalator, resulting in a fluorescence assay to detect DNA

hybridization. Here, we present the preliminary results of the hybridization assay (Figure 27), which involves the reversible *trans–cis* isomerization of the azo moiety. Initially, the azobenzene DNA binds to the spacer DNA resulting in a fluorescence signal. Upon UV irradiation, the fluorescence decreases indicating a loss of hybridization due to the change in conformation of DAB DNA caused by *trans* → *cis* isomerization of the azo moiety. After exposure to white light for 5 minutes, DAB6:DNA2, DAB5:DNA0 and DAB5:DNA4 all restored their fluorescence readings to some degree. This signifies that *cis* DAB DNA converts into the *trans* form followed by hybridization to the complementary spacer DNA. The DAB DNA sequences that show the highest photoswitchability are DNA sequences with 4-6 nucleotides flanking the DAB linker.

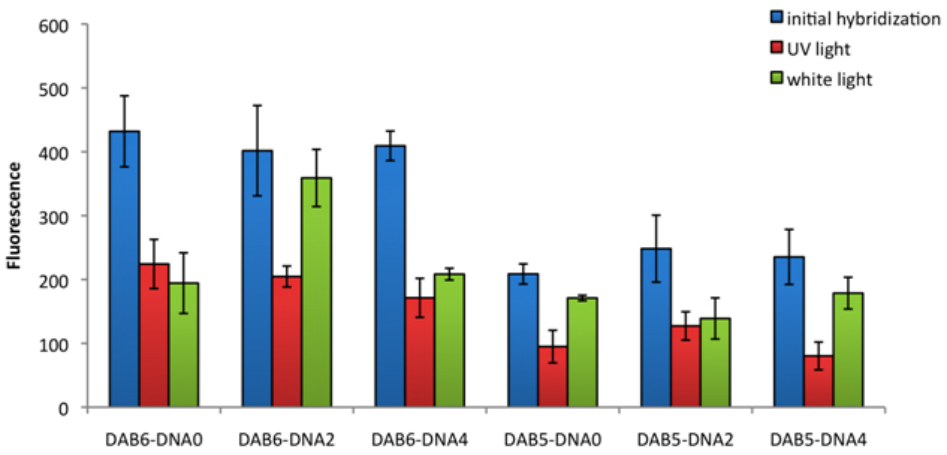


Figure 27. Fluorescent reading of DAB DNA:DNA duplex formation. Fluorescence was measured before and after irradiation with UV (2 min, 365 nm) light and white light (5 min), respectively, on a BioTek plate reader (546/590 nm). Error bars represent standard deviations from three independent experiments. Experiments were performed by Jeane Govan.

3.2. Experimental data for synthesized compounds

(*E*)-4-((4-(2-Hydroxyethoxy)phenyl)diazenyl)phenol (84). $NaNO_2$ (325 mg, 4.71 mmol) was added to the solution of the alcohol **83** (500 mg, 3.60 mmol) in aqueous HCl (4 M, 20 mL) at 0 °C. After stirring the reaction mixture for one hour, a solution of phenol (408 mg, 4.32 mmol) in aqueous NaOH (4 M, 21 mL) was slowly added to the reaction mixture at 0 °C, followed by continued stirring at room temperature for 3 h. The reaction mixture was neutralized with an aqueous solution of HCl (1 M, 5 mL) and the aqueous layer was extracted with EtOAc (3 × 20 mL). The combined organic layers were washed with water (20 mL), brine (20 mL), and dried over anhydrous Na_2SO_4. The filtrate was concentrated under reduced pressure and the remaining crude product was purified by column chromatography on silica gel using Et_2O:hexanes (9:1) to furnish the azophenol **84** as a yellow solid (410 mg, 47% yield). ^{1}H NMR (400 MHz, acetone-d_6, δ ppm): δ = 9.15 (br, 1H), 7.86–7.80 (m, 4H), 7.09 (d, J = 4.4 Hz, 2H), 6.94 (d, J = 4.4 Hz, 2H), 4.18–4.15 (m, 2H), 4.12 (s, 1H), 3.91 (s, 2H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 162.1, 160.9, 147.7, 147.1, 125.3, 125.0, 116.6, 115.7, 70.9, 61.31. HRMS-ESI (m/z) $[M+H]^+$ calcd for $C_{14}H_{15}N_2O_3$: 259.1083; found 259.1077.

(*E*)-2-(4-((4-(2-(*bis*(4-Methoxyphenyl)(phenyl)methoxy)ethoxy)phenyl)diazenyl) phenoxy)ethanol (**86**). DMT protected bromoethanol **85** (268 mg, 0.610 mmol) and K_2CO_3 (198 mg, 1.44 mmol) were added to the solution of the phenol **84** (125 mg, 0.480 mmol) in dry DMF (1 mL) stirred at room temperature under an inert atmosphere. The reaction mixture was then heated and continued stirring at 60 °C for 12 h. The reaction mixture was

cooled to room temperature, a saturated aqueous solution of $NaHCO_3$ (1 mL), and the aqueous layer was extracted using EtOAc (10 mL × 3). The combined organic layers were washed with water (20 mL), brine (20 mL), dried over anhydrous Na_2SO_4, and filtered. After filtration, the filtrate was concentrated under reduced pressure and the remaining crude product was purified by column chromatography on silica gel using a mixture of EtOAc:hexanes (3:2) to furnish the corresponding DMT protected azophenol **86** (265 mg, 90% yield) as a red foam. ^{1}H NMR (400 MHz, $CDCl_3$, δ ppm): δ = 7.90–7.88 (m, 4H), 7.50 (d, *J* = 3.6 Hz, 2H), 7.40–7.38 (m, 4H), 7.32–7.16 (m, 4H), 7.04–7.01 (m, 4H), 6.85–6.83 (m, 4H), 4.22 (t, *J* = 5.2 Hz, 2H), 4.16 (t, *J* = 4.8 Hz, 2H), 4.00 (t, *J* = 4.0 Hz, 2H), 3.97 (s, 6 H), 3.47 (t, *J* = 5.2 Hz, 2H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 161.2, 161.1, 160.7, 158.8, 147.5, 147.2, 145.0, 139.6, 136.2, 130.2, 129.3, 128.3, 128.0, 127.9, 126.9, 115.0, 114.9, 113.3, 86.3, 69.6, 67.9, 62.5, 61.6, 55.4. HRMS-ESI (m/z) $[M+Na]^+$ calcd for $C_{37}H_{36}N_2O_6Na$: 627. 2471; found 627.2466.

(*E*)-2-(4-((4-(2-(*bis*(4-Methoxyphenyl)(phenyl)methoxy)ethoxy)phenyl)diazenyl) phenoxy)ethyl (2-cyanoethyl) diisopropylphosphoramidite (**87**). DIPEA (215 µL, 1.20 mmol) and 2-cyanoethyl- *N*,*N*-diisopropylchlorophosphoramidite (110 µL, 0.490 mmol) were added to a solution of the alcohol **86** (150 mg, 0.24 mmol) in DCM (2.0 mL) stirred at room temperature under an argon atmosphere followed by continued stirring at room temperature for 2 h . The reaction mixture was concentrated under reduced pressure and the obtained crude product was directly purified by column chromatography on silica gel using a mixture of EtOAc:hexanes (1:2) with 1% TEA to obtain **87** (152 mg, 76% yield) as a red

foam. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.87–7.84 (m, 4H), 7.49–7.46 (m, 2H), 7.38–7.34 (m, 4H), 7.27–7.19 (m, 4H), 7.07–6.98 (m, 4H), 6.83–6.79 (m, 4H), 4.20–4.18 (m, 4H), 4.05-3.95 (m, 3H), 3.86–3.76 (m, 8H), 3.66–3.58 (m, 2H), 3.46–3.43 (m, 2H), 2.64–2.60 (m, 2H), 1.22–1.20 (m, 12H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 161. 1, 160.9, 158.6, 147.3, 147.2, 145.0, 136.2, 130.2, 128.3, 128.0, 126.9, 124.5, 115.0, 114.9, 113.2, 86.3, 68.4, 68.3, 67.9, 62.5, 62.2, 61.9, 58.8, 58.5, 55.4, 43.4, 43.2, 24.9, 24.8, 22.6. ^{31}P NMR (161 MHz, $CDCl_3$, δ ppm): δ = 149.75.

CHAPTER 4: PROTEIN MODIFICATION WITH UNNATURAL AMINOACIDS

4. Protein modification

Proteins of all living organisms consist of the same 20 canonical amino acids (Figure 28) encoded by the universal genetic code. The genetic code contains 64 triplet codons specifying the 20 natural amino acids and three translation-termination codons.[299] These canonical amino acids are utilized to translate genetic information from messenger RNAs to proteins by the ribosome. These processes operate with high fidelity, having an error rate of 1 in 10^4–10^5.[6] The fidelity is important to maintain the integrity of life processes. For instance, it was estimated that about 1800 inherited human diseases are caused by nonsense mutations.[300] Studies have also shown that premature termination codons produce truncated non-functional proteins, which could account for a large percentage of cases of genetic diseases, including cystic fibrosis, muscular dystrophy and cancers.[300]

Proteins are catalysts, units of signaling pathways, building blocks of cell structure, and the determinants of cellular morphology, motility, and stability.[301] Obviously, alterations in protein synthesis, structure, and functions may result in detrimental effects, such as cancers or the other biological disorders in cells and ultimately the whole organism.

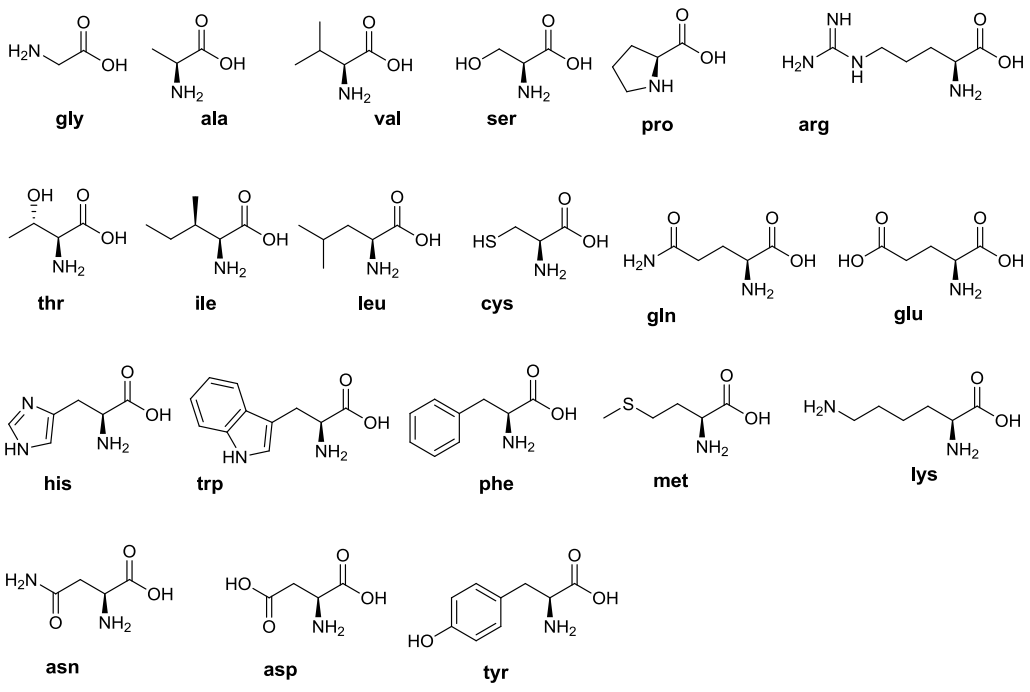


Figure 28. Canonical twenty amino acids with their abbreviated notations.

The protein data bank, one of the vast repositories in protein structures, reports over 80,000 protein structures at an atomic detail.[302] Since proteins are associated with many biological functions, studies on molecular tools for exploring protein-protein interactions may provide crucial information on gene regulations and misregulations.[303] It has been predicted that about 70,000 protein–protein interactions between 6231 proteins may occur in the human body; estimated based on protein–protein interaction data from lower eukaryotic interactomes.[304, 305] Studies also revealed that about 244 kinases, one of the largest families of the signaling molecules involved in many cellular processes, are related to disease or amplified cancer oncogenes among the 918 kinases that are present in the human genome.[306-308] Moreover, enzymes and modulator proteins (activators, inhibitors or regulators) in fact

share a significant proportion of protein–protein interactions. It was estimated that about 269 epigenetic enzymes and 4377 modulators, key players in human epigenetics, could be investigated as a subject of therapeutic potential.[309]

A variety of protein modifications techniques with or without genetic code manipulation have been used to study protein structures and functions.[310-312] They include 1) labeling with chemical reagents or dyes, 2) labeling with a peptide via native chemical ligation, 2) labeling with fusion protein, 3) labeling with fusion tag/dye, and 4) unnatural amino acids (UAAs) mutagenesis. These recently evolving methods, applicable at the molecular level, have been used to dissect various aspects of protein functions including protein–protein interaction, catalysis, or signal transductions etc.[313, 314]

4.1.Unnatural amino acid (UAA) mutagenesis

Incorporation of UAAs, amino acids that differ from the canonical 20 amino acids, can introduce a variety of molecular probes into the proteins of interest. Successful incorporation of UAAs as biophysical probes with unique physiochemical properties provide unprecedented opportunities for examining proteins *in vivo*.[315] In the past, UAA mutagenesis has been conducted via residue-specific and site-specific incorporation techniques.[316, 317]

The residue-specific technique incorporates UAA by replacing the corresponding amino acid at every position (globally) in proteins.[318, 319] Hundreds of unnatural amino acids have been investigated using the methodology.[318] A few of the recently developed residue-specific labeling techniques offer unprecedented successes to interrogate various cellular functions related to diseases.[320] They include, but not limited to, bioorthogonal non-canonical

amino acid tagging, fluorescent non-canonical amino acid tagging, and Staudinger ligation etc.[317, 321-325]

The UAAs mutagenesis via site-specific incorporation technique has four fundamental components. They include 1) an unnatural amino acid, 2) a suppressor tRNA charged with the unnatural amino acid and containing an anticodon to recognize a nonsense codon, 3) a gene of interest with a nonsense codon, and 4) a translation system containing ribosomes.[316, 326] The site-specific incorporation techniques have been extensively investigated using both cell-extract and cell-intact translational systems. These two approaches may be complementary to each other depending upon the subject of investigations since both of them have their own advantages and shortcoming.[317]

The main difference between cell-extract and cell-intact translational systems is the processes of charging a suppressor tRNA with an UAA. In the cell-extract translational system, an UAA is chemically acylated to the suppressor tRNA (*in vitro*), where as in the cell-intact translational system, an engineered orthogonal tRNA synthetase charges the UAA into the suppressor tRNA *in vivo*. The use of a cell-extract translational system is a powerful technique particularly for mapping ion channel functions,[327] exploring protease activities, and studying protein-protein interactions.[326] Nevertheless, the global application of the technique is limited because of crucial limitations, including low yield of labeled protein, need for tRNA synthesis, and need for microinjection or transfection of the tRNA or protein.[58, 326]

4.2. UAA incorporation using cell-intact translational machinery

UAA incorporation into a protein using orthogonal traslational machinery in response to a nonsense amber codon (TAG) is performed *in vivo* and considered a more advanced technique than that of the cell-extract system.[328, 329] In this method, both the tRNA and the tRNA synthetase are evolved to function as orthogonal pairs to endogenous tRNA/tRNA synthetase pairs in the cell thereby eliminating a step of chemical amino acylation of tRNA.[330] These engineered orthogonal tRNA/tRNA synthetase pairs can be coupled with cellular translational machineries without any cross talk to the endogenous tRNA/tRNA synthetase pairs.[331] Thus, to perform a site-specific incorporation of UAA into the protein, the method requires a manipulation of the corresponding genes (plasmids) that encodes all three necessary components: a tRNA, a tRNA synthetase, and a target gene with an inframe amber stop codon.[315]

In general, site-specific incorporation of UAAs using this methodology requires outstanding knowledge and skills of both chemistry as well as biology. Prior to begin a protein expression using UAA 1) an UAA is synthesized, 2) a target gene with an amber stop codon (TAG) is designed, and 3) tRNA/tRNA synthetase pairs are engineered to be orthogonal by negative and positive rounds of selection. During the translation process, the orthogonal tRNA synthetase activates an UAA and subsequently charges the tRNA without any cross reaction with endogenous tRNA/tRNA synthetase pairs (Figure 29). Finally, the charged orthogonal tRNA recognizes the nonsense stop codon and delivers the UAA into the protein thereby helping continue the translation process with high fidelity.[58, 315, 328, 332] A number of UAAs have been site-specifically incorporated into proteins in *E. coli* using

specifically evolved orthogonal tRNA/tRNA synthetase pairs that are reported elsewhere.[328, 332-335]

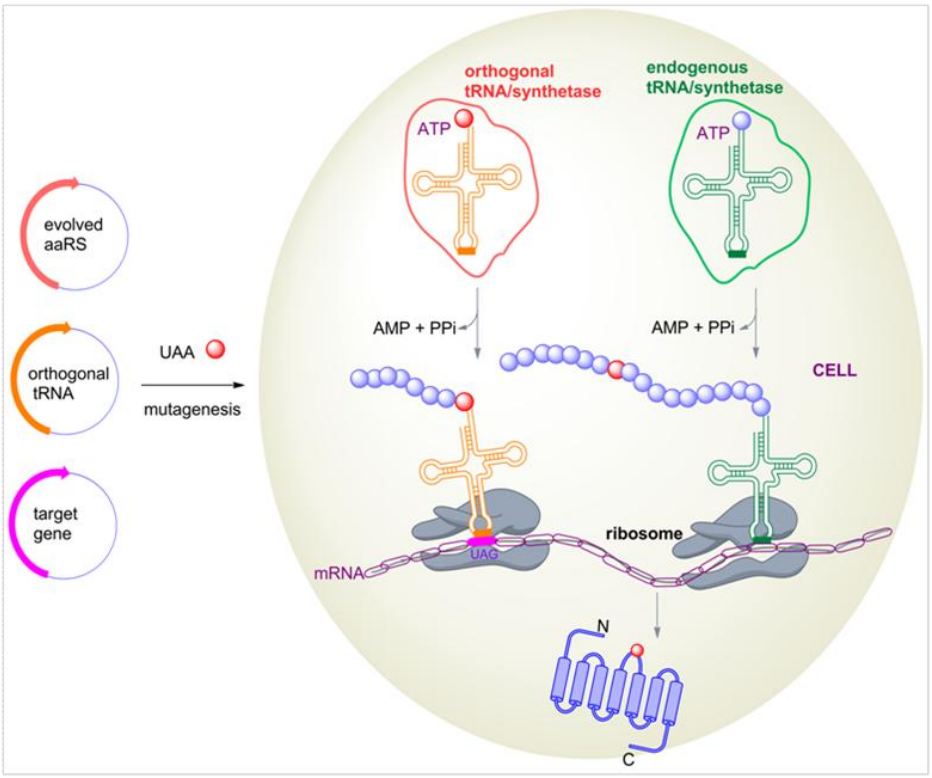


Figure 29. Site-specific incorporation of unnatural amino acids (UAAs). Adapted with permission from *Curr. Opin. Chem. Biol.* **2011**, *15*, 392-8.

Historically, the first site-specific UAA incorporation used an evolved tRNA/tRNA synthetase pair from *M. jannaschi* to encode *O*-methyl-L-tyrosine **88** (Figure 30) in response to an amber stop codon in *E. coli*.[328] Since then, biological studies via the site-specific incorporation technique have been thriving by encoding over 70 UAAs, including **89**–**98** into

proteins to study various areas, including protein-protein interactions, post translational modifications, photoregulation of gene functions and signaling pathways.[58, 322, 336-339] Most of those studies are focused on *E. coli,* and limited investigations have been performed in yeast or mammalian cells.

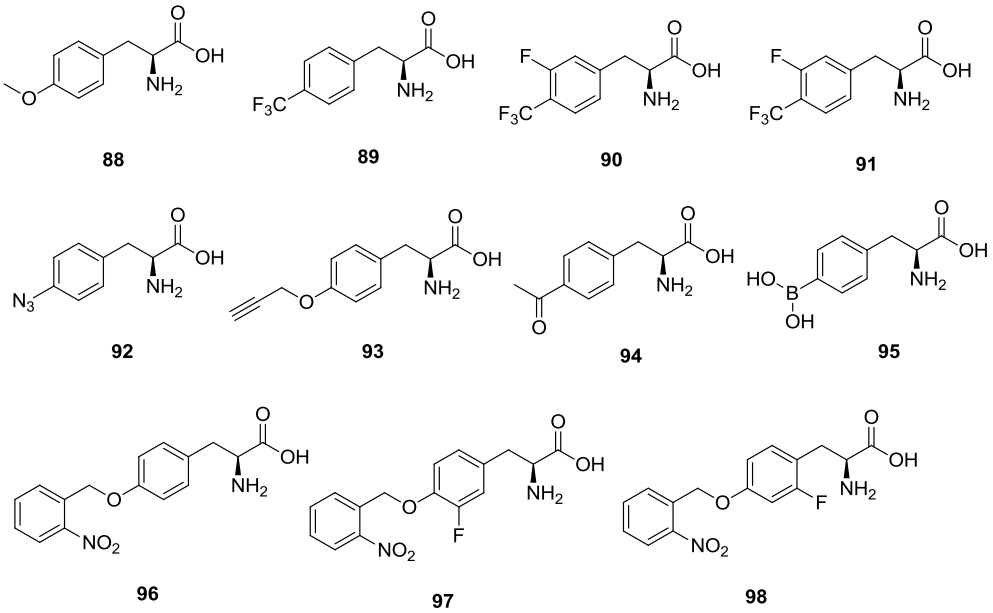


Figure 30. Genetically encoded UAAs: fluorinated UAAs **89**–**91**, clickable UAAs **92**–**93**, UAAs for cross link formation **94**–**95**, and caged UAAs **96**–**98**.

The success on the site-specific incorporation of UAAs into proteins has been well documented.[58] Yet, further investigation will be more challenging because of the inherent and technical limitations associated with the site-specific incorporation of UAAs, including 1) low yield of the labeled protein 2) extensive and lengthy procedure of tRNA/tRNA synthetase engineering 3) limited recognition flexibility of the tRNA synthetase regarding the

size of a UAA 4) nonsense-mediated mRNA decay, a eukaryotic surveillance mechanism to degrade mRNA with premature stop codon, and its consequences.[300, 322, 340]

CHAPTER 5: GENETIC CODE EXPANSION WITH LYSINE ANALOGUES

5. Genetic code expansion via orthogonal aminoacyl tRNA/tRNA synthetase pairs

Translational incorporation of unnatural amino acids (UAAs), a genetic code expansion methodology, has enabled fundamentally novel protein engineering by introducing unprecedented chemistries into proteins. Over the last twelve years, a number of UAAs have been genetically encoded in cells in response to the amber nonsense codon (TAG), the least frequently used stop codon (about 7% usage in *E. coli*),[341] by orthogonal aminoacyl tRNA/tRNA synthetase pairs.[58, 342] These newly incorporated UAAs have been utilized to study protein function including proteins involved in gene regulation, specific protein-protein interactions, and the dynamics of cell signaling, etc.[315, 343-346] Often, the site-specific incorporation of UAAs has been focused in *E. coli* because of the limitation on engineering aminoacyl tRNA/tRNA synthetase pairs that are orthogonal in other model organisms.[333, 347, 348] This might be the reason why two pairs, namely the tyrosyl tRNA/tRNA synthetase pair and pyrrolysyl tRNA/tRNA synthetase pair (*vide infra*), are investigated extensively among the twenty four known aminoacyl tRNA/tRNA synthetase pairs in *E. coli*.[349] However, the genetic encoding of UAAs has expanded into other model organisms as well, such as yeast, *Drosophila*, and mammalian cells.[331, 350-352]

5.1. Pyrrolysyl tRNA/tRNA synthetase (PylT/PylRS) pairs and its orthogonality

In the *Methanosarcina* species, an orthogonal PylT/PylRS pair encodes pyrrolysine, a lysine derivative, that is incorporated into the enzyme methyltransferase in response to an in-frame UAG codon to produce full-length (50 kDa) functional protein.[353, 354] This enzyme is

responsible for methanogenesis (the conversion of methyl amines into methane and ammonia), via the pyrrolysine residue which plays a catalytic role in activation of methylamines.[354] In fact, orthogonal PylT/PylRS pairs were found to encode pyrrolysine into all three methylamine methyltransferases namely mono-, di- and trimethyltransferases found in the archaea *M. barkeri*.[355] Later, it was discovered that other three *Methanosarcina* species, *M. mazei*, *M. acetivorans*, and *M. burtonii*, and the gram-positive bacterium *Desulfitobacterium hafniense* also use orthogonal pairs to encode pyrrolysine into methyltransferase.[353, 356-359]

Structural analysis of both the PylT and PylRS revealed a few exceptional features at the molecular level for their orthogonality to endogenous aminoacyl tRNA/tRNA synthetase pairs.[349, 359] All known synthetases are classified into two main groups in which the PylRS is categorized into the class IIC. The PylRS possesses a unique *N*-terminal tRNA binding domain, the conserved tRNA synthetase class II catalytic domain, the bulge domain, and the *C*-terminal tail (Figure 31).[349, 358, 359] In the case of class I aminoacyl synthetases such as tyrosine and lysine synthetases, *N*-terminal and *C*-terminal Rossmann folds are the catalytic domains, which are connected through an inserted domain known as connecting peptide loop (CPI). The CPI domain serves as a dimer interface in class I synthetases.

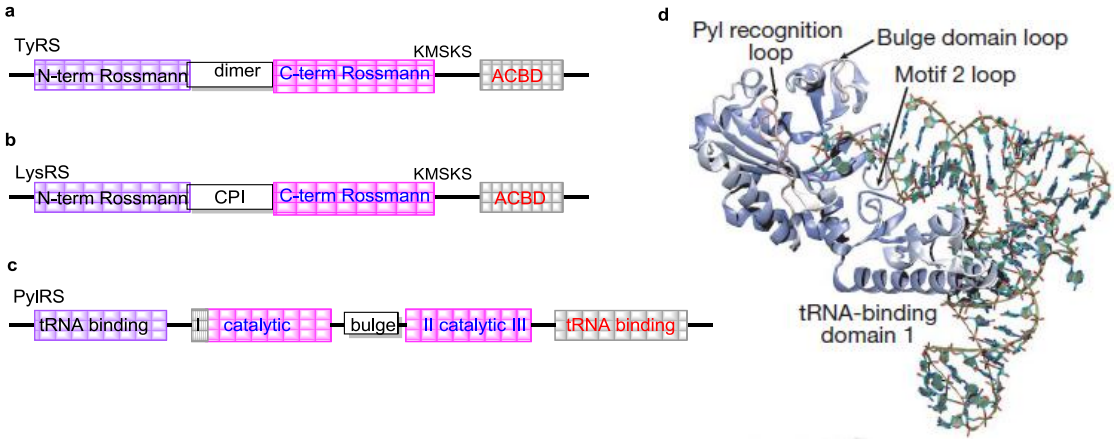


Figure 31. Synthetase structures: Primary structure and principal domains of synthetases. a) Tyrosyl synthetase (TyrRS), b) lysine synthetase (LysRS), c) pyrrolysine synthetase (PylRS); and d) interactions between the synthetase and that of tRNA in PylT-PylRS complex; CPI = connecting peptide loop, ACBD = anticodon binding domain, KMSKS = a sequence motif (lys-met-ser-lys-ser) to stabilize a loop; Figure 31d is adapted with permission from *Nature* **2009**, *457*, 1163-7.

The complementarities between the tRNA interacting domains of PlyRS and the PylT play a crucial role for their recognition specificity to each other and to the UAA. A large and a flexible hydrophobic amino acid binding pocket is a key structural feature in (wild type) PylRS for providing a broad specificity to accommodate various fairly bulky UAAs (Figure 32).[360, 361] The primary and the secondary structures of PylT possesses unique features including an extended acceptor stem, a small D-loop, and a CUA anticodon.[360] Thus, wild type orthogonal PylT/PylRS pairs have been discovered to encode several lysine analogues site-specifically into model proteins in various cell types such as *E coli*, yeast or mammalian cells (Figure 32).[362] Studies illustrated that UAA incorporation by the wild type synthetase

may also depend upon the characteristic features of the substrate (UAAs). These crucial constituents may include a carbamate functionality (at $^{\varepsilon}N$-position in lysine), a hydrophilic moiety, a specific chain length (which allow for a fit into the pocket as a spacer), and a small hydrophobic group (which can be accommodated by the flexible hydrophobic pocket).

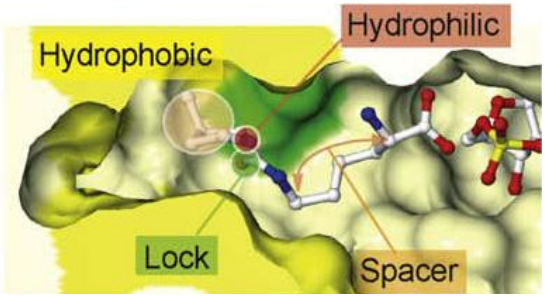


Figure 32. A pyrrolysine in the binding pocket of the PylRS, with key components of the complex indicated. Adapted with permission from *Chem. Biol.* **2008**, *15*, 1187-97.

Previously, Deiters et al. and others[362] have investigated both wild type and engineered variants of PylT/PylRS pairs to incorporate various UAAs in proteins in both pro- and eukaryotes,[360, 363] Thus, several PylT/PylRS pairs have been investigated for achieving site-specific incorporation of a wide range of uniquely functionalized amino acids. Studies also revealed a great potential for genetic code expansion through engineering the tRNA anticodon by taking advantage of the loose recognition nature of the tRNA anticodon by PylRS, to exploit a frameshift codon (e.g.; four-base or five-base) for incorporating additional UAAs into the protein.

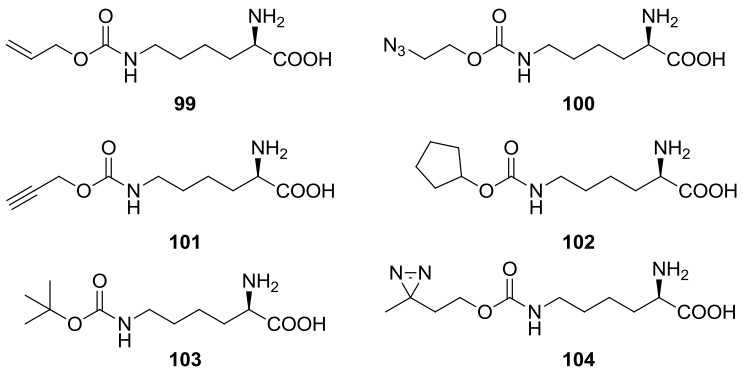


Figure 33. Lysine-based unnatural amino acids **99**–**104** that are incorporated by wild type PylT/PylRS pair.

To further continue the genetic code expansion of organisms via UAAs incorporation, here we report the synthesis of a number of lysine analogues **105**–**113** and their current status on site-specific incorporation into model proteins to achieve unique functions (*vide infra*). The UAAs may be catagorized into: 1) photocrosslinking lysine **104**, 2) caged lysine **105**–**109**, 3) fluorescent lysine **110**, 4) dithiolane lysine **111**, 5) photoswitchable lysine **112**, and 6) spin labeled lysine **113** (*vide infra*).

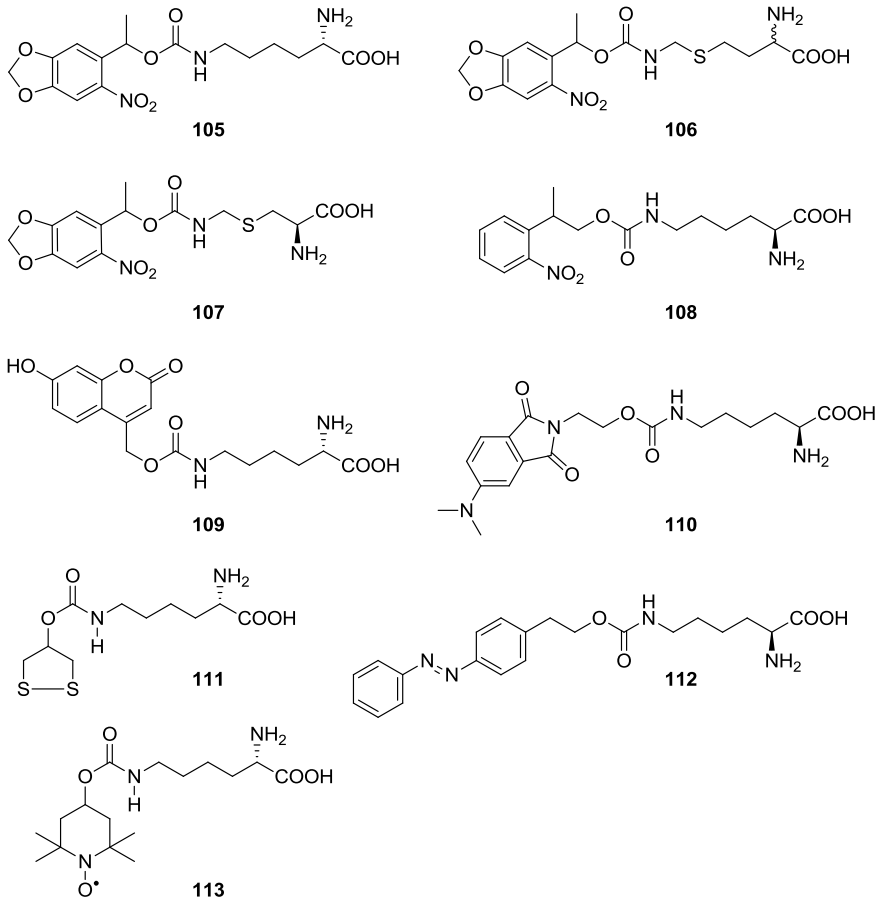


Figure 34. Lysine-based unnatural amino acids. The UAA **105–113** are incorporated by evolved PylRS variants.

5.2. Genetic code expansion with diazirine lysine and subsequent protein-protein interaction

Photocrosslinking groups have been used to study biological functions via spatial and temporal activation using light.[364-368] The key component of a photocrosslinking strategy is the photoaffinity label that, when activated by UV light of a suitable wavelength, forms a highly reactive species called a carbene.[369] The newly formed carbene may be linked to a

biomolecule within a van der Waals radius through a covalent bond.[367] Previously, the photocrosslinking technique has been successfully applied to investigate protein–lipid interactions,[369] small molecule–protein interactions,[370] oligonucleotide interactions,[371-373] oligonucleotide–protein,[374] and protein–protein interactions.[368, 375-377]

The site–specific incorporation of a photoaffinity probe using an UAA mutagenesis, *in vivo*, might be beneficial compared to that of post-translational modification with photocrosslinking reagent.[378] The methodology currently in use in the Deiters lab has the advantage that a photoaffinity probe can be incorporated site-specifically into any protein in its native environment allowing for the direct study of proteins.[379] This may enable the isolation of weak or transient interactions that are not captured by non-covalent affinity purification methods such as tandem affinity purification (TAP) tagging that are commonly used in proteomics.[380]

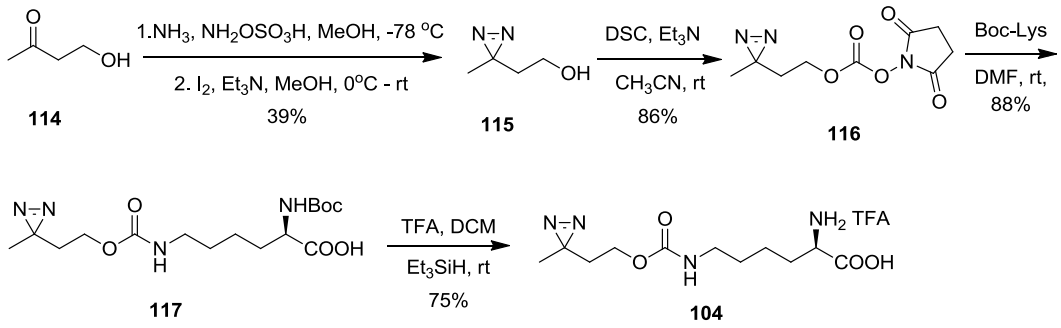
Previously, phenylalanine derivatives containing azido, benzophenone, and diazirine groups have been site-specifically incorporated into proteins in *E. coli* in response to a nonsense amber codon using engineered variants of the *M. jannaschii* tyrosyl tRNA/tRNA synthetase (*Mj*TyrT/*Mj*TyrRS) pair.[381, 382] *para*-Benzoyl-L-phenylalanine has also been incorporated site-specifically into proteins in yeast and mammalian cells using an *E. coli* TyrRS variant that was evolved in yeast to recognize the UAA.[383] The site-specific incorporation of benzophenone, another photocrosslinking molecule, has been used to investigate and define protein interactions in a wide range of biological systems.[384, 385]

The bulkiness of the benzophenone group, prolonged lifetime, low-affinity binding, and need of extended time of irradiation limit its application to some extent,[377] and when placed at a protein–protein interface may destroy the potential interaction under investigation.[378, 379] Since aromatic amino acids often constitute protein hotspots on protein surfaces, substitutions of aromatic amino acids with benzophenone may often hamper the binding event under investigation. Here, we present a brief analysis of what we reported in the synthesis of an aliphatic crosslinking amino acid called diazirine lysine and its incorporation.[378] Upon site-specific incorporation, the amino acid may serve as a complementary method to the use of *p*-benzoyl-L-phenylalanine and other aromatic or bulky photocrosslinking amino acids to investigate protein interactions.

5.2.1. Synthesis of diazirine lysine analogue

We synthesized the diazirine-lysine **104** starting from commercial 4-hydroxybutan-2-one (**114**) and converting it into the corresponding diazirine **115** in two steps (Scheme 10). Subsequent activation of the alcohol **115** with disuccinimide carbonate (DSC) in the presence of triethylamine gave the corresponding carbonate **116** in 86% yield. Reaction of **116** with Boc-lysine in DMF furnished **117** in 88% yield. The acid sensitive nature of the diazirine group posed some synthetic challenges to remove the Boc-protecting group, which happened to be more difficult than anticipated. Several common protocols (e.g., TFA:DCM (1:4, or 1:1); HCl in methanol (1.0, or 0.5 M); HCl in dioxane:DCM (1.0 M); all at 0 °C or room temperature; with or without a Me_2S as a cation scavenger) were employed and reaction progress was monitored at different time intervals. All aforementioned conditions led to

decomposition of the starting material. Through further screening of deprotection conditions,[386] a clean deprotection was achieved with a 1:20 mixture of TFA in DCE in the presence of triethylsilane as a cation scavenger. Finally, the desired amino acid **104** was obtained in 76% yield.



Scheme 10. Synthesis of the diazirine lysine **104**.

5.2.2. Incorporation of the diazirine lysine using wild type PylT/PylRS

Diazirine lysine was incorporated using wild-type PylT/PylRS into myoglobin as a model protein in *E. coli* and human embryonic kidney (HEK293) cells. In these biological studies performed by Chou in the Deiters group, a protein-protein crosslink formation was also achieved using a glutathione-S-transferase (GST) in *E coli*. The PylT/PylRS pair was transformed in *E. coli*, with pMyo4TAGpylT,[363] containing a hexahistidine-tagged myoglobin gene with an amber stop codon (TAG) at the 4 position and pBKpylS that encodes *Mb*PylRS. Protein expression was induced in culture media containing diazirine lysine **104** (1 mM) and arabinose at room temperature overnight and the protein was purified

by Ni-NTA chromatography. A moderate yield of about 2 mg diazirine myoglobin per liter of bacterial culture was obtained. The Coomassie stained SDS-PAGE of the purified protein revealed that the expressed protein is myoglobin (18 kDa), which was confirmed with a molecular weight of 18521.94 ± 0.50 Da by ESI-MS, which closely matches the expected value of 18522.26 Da.

Since the PylT/PylRS pair is also orthogonal to mammalian cells; human embryonic kidney (HEK293) cells were transfected with two constructs.[351] The first one is a reporter construct encoding an N-terminal mCherry gene, a linker with an amber stop codon (TAG), a C-terminal enhanced GFP gene (EGFP), and a HA-tag coding sequence. The second construct is a PylT/PylRS expression plasmid. As expected, only the mCherry gene was expressed regardless the presence or the absence of diazirine lysine but full-length mCherry-TAG-EGFP-HA was expressed only in the presence of diazirine lysine (Figure 35). The result was also confirmed by Western blot of mCherry-TAG-EGFP-HA with an anti-HA-tag-antibody.

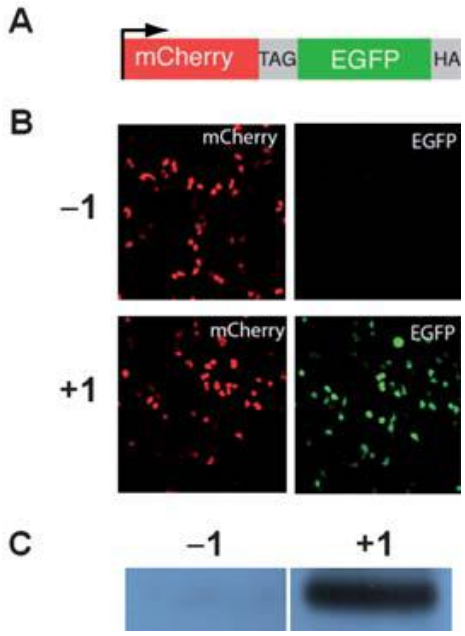


Figure 35. Diazirine lysine incorporation in mammalian cells: A) The mCherry-TAG-EGFP reporter constructs. B) Fluorescence micrographs of HEK293 cells expressing mCherry-TAG-EGFP-HA; PylT, and PylRS. A full-length protein is expressed only in the presence of diazirine lysine. C) Western blot of mCherry-TAG-EGFP-HA expression with an anti-HA-tag-antibody in the absence (left lane) or in the presence (right lane) of diazirine lysine (1). Adapted with permission from *Chem. Sci.* **2011**, *2*, 480-4.

Then, the diazirine lysine was incorporated in the glutathione-S-transferase (GST) at the 52 position, which lies at an interface between monomers in the GST dimer, to investigate protein-protein interactions. The target gene pGST52TAGpylT and the *Mb*PylRS encoding gene pBKpyls were transformed in *E coli* for expression in the presence of diazirine lysine. The expressed HA-taged-GST protein was purified by Ni-NTA

chromatography and exposed to UV light (365 nm) in order to create a photocrosslinked GST dimer. Western blot using an anti-His-tag-antibody confirmed the dimer (Figure 36). Additionally, *E. coli* cultures producing diazirine-GST were washed and resuspended in PBS buffer and either kept in the dark or irradiated with UV light (365 nm). Subsequent western blot analysis of the cellular lysate confirmed the formation of the GST dimer in the presence of UV light.

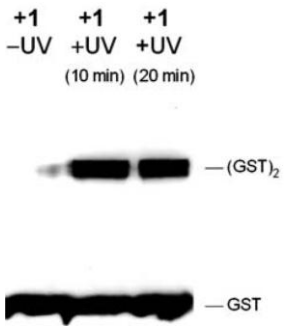


Figure 36. Diazirine lysine incorporation in GST52TAGpylT target gene: photocrosslinking of diazirine-GST to form GST dimer in the presence of diazirine lysine (1), and a UV irradiation (365 nm); left lane – no UV and right two lanes – with UV at 10 and 20 minutes respectively. Adapted with permission from *Chem. Sci.* **2011**, *2*, 480-4.

In summary, an aliphatic diazirine lysine has been synthesized, site-specifically incorporated into both bacterial and mammalian proteins, as well as utilized to create a photo-crosslink formation between the binding partners.

5.3. Genetic code expansion with caged homocysteine using PylT/PylRs pair

In general, side chains in natural amino acids are both structurally and functionally diverse. For instance, out of 20 canonical amino acids, methionine and cysteine are sulfur functionalized amino acids, another twelve amino acids contain the heteroatoms nitrogen and/or oxygen in their side chains, and the rest of the amino acids bear a hydrocarbon side chain. Both cysteine and homocysteine, a metabolized product of methionine, possess a thiol group. Therefore, they share sulfur chemistries and participate in many cellular functions through oxidation, reduction, or crosslink formation (*vide infra*).[387-389] However, unlike cysteine and methionine, homocysteine does not incorporate into proteins through translation.[390]

A cellular mechanism that governs the homocysteine and methionine interconversion and their subsequent transformation into cysteine (an essential amino acid for translation) has been an important topic for further investigation.[391] Both cysteine and homocysteine have specific biological relevance and both of the amino acids are involved in many diseases. For instance, homocysteine metabolism is linked with various diseases, including 1) diabetes, 2) neurodegenerative or neuropsychiatric diseases such as Alzheimer's disease, depression, Parkinson's disease, and so on, 3) cardiovascular diseases, 4) aging, and 5) various forms of cancers.[387, 391-393]

Proteins perform biological functions with a proper transformation of genetic information.[394] Protein dynamics and their post translational modifications pose remarkable consequences in cellular processes including unanticipated alteration in gene regulation, signaling pathways, protein folding, and enzymatic activities.[395] In order to investigate

protein function and dynamics, previously, Deiters et al. reported on site-specific incorporation of caged lysine **105** in response to amber stop codon (UAG) by using orthogonal PylT/PylRS pairs to study protein function through its post-translational light activation.[344, 351] In particular, the caged lysine is used for studying photochemical control of nuclear import of the tumor suppressor protein p53.[396]

Here, to continue the investigation of protein function through genetic code expansion, we synthesized caged homocysteines **106, 118–119** and caged cysteine **107**. Extensive literature searches (as of April 2013) did not yield any reports on homocysteine site-specific incorporation into any protein using the cellular translational machinery. However, a nitrosyl homocysteine has been reported to incorporate into proteins by using the methionine synthetase (metRS) in *E. coli*, where S-nitroso-homocysteine was activated by metRS to produce the corresponding *S*-nitroso-Hcy-tRNA complex.[390]

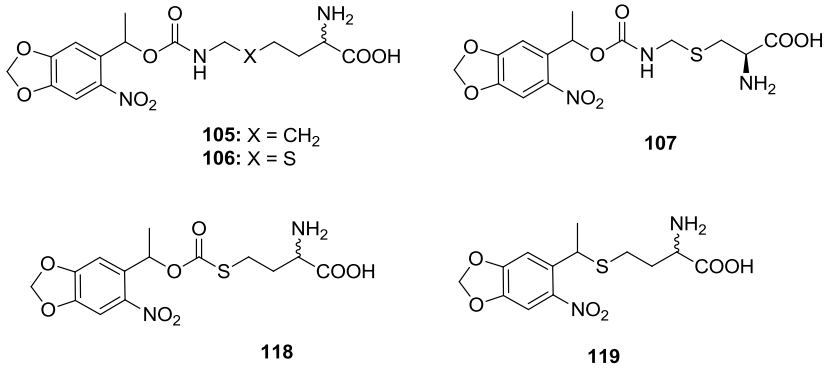
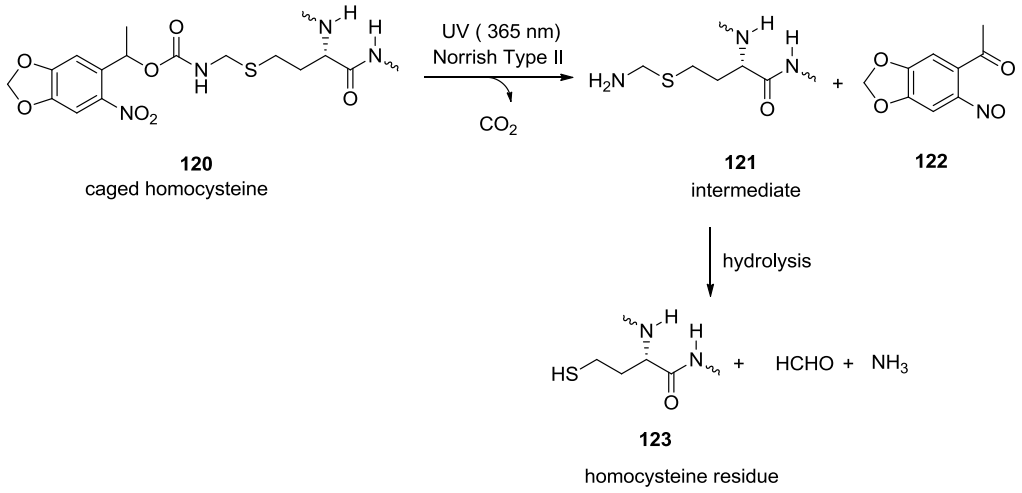


Figure 37. Caged lysine **105**, caged lysine mimics of homocysteine **106**, **118**–**119**, and caged cysteine **107**.

The caged homocysteine residue **120** in a protein undergoes photolysis (Norrish type II) upon UV irradiation (365 nm) resulting in the hemiaminal intermediate **121**. The unstable hemiaminal intermediate **121** spontaneously eliminates formaldehyde and ammonia, thereby introducing a homocysteine residue **123** into the protein (Scheme 11).

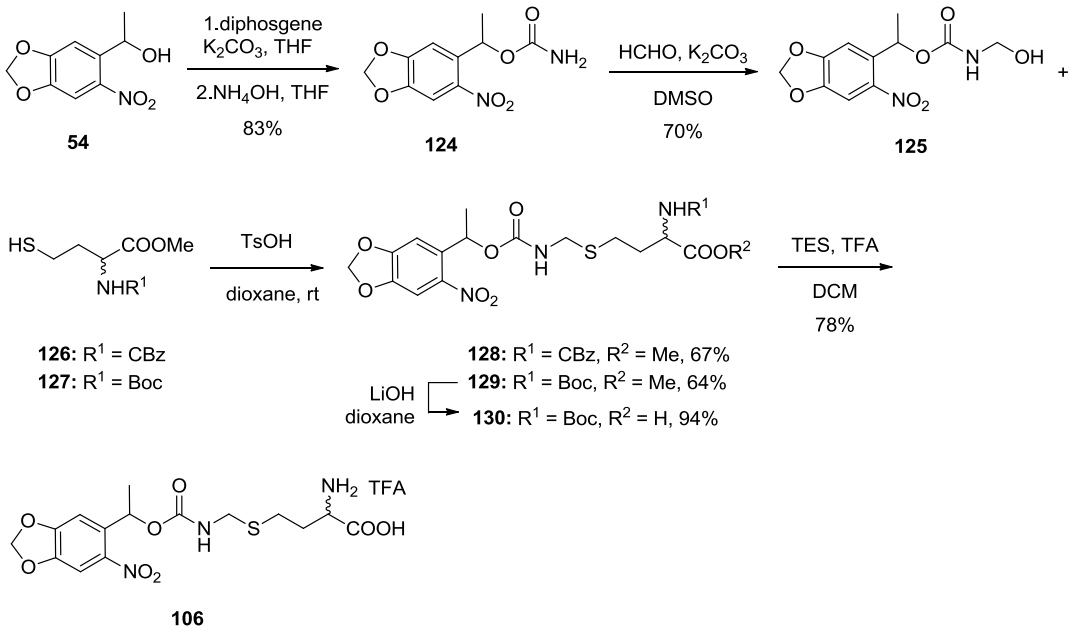


Scheme 11. UV-induced decaging of **120** incorporated in a protein.

5.3.1. Synthesis of caged homocysteine lysine analogues

Synthesis of caged homocysteine **106** was completed in five steps starting from the alcohol **54**, which was assembled following a literature procedure.[179] The alcohol **54** was treated with diphosgene and then subsequently with an excess of concentrated aqueous ammonia solution that resulted in the amide **124** in 83% yield.[397] The amide **124** was reacted with paraformaldehyde in the presence of 10 mol% K_2CO_3 in DMSO to provide the desired hydroxymethyl amide **125** in 70% yield. The alchohol **125** underwent iminium ion chemistry through treatment with TsOH (8 mol%) and the iminium intermediate, which was generated

in situ, was trapped by the thiol group of either Cbz or Boc protected homocysteine methyl esters **126** or **127**, delivering the products **128**–**129**. Both esters **126** and **127** were prepared according to reported procedures.[398, 399] The ester **129** was hydrolyzed with aqueous LiOH to the corresponding acid **130** in 94% yield. Finally, the acid **130** was deprotected with a mixture of TFA in DCM in the presence of triethylsilane (TES) to furnish the caged homocysteine **106** in 78% yield.

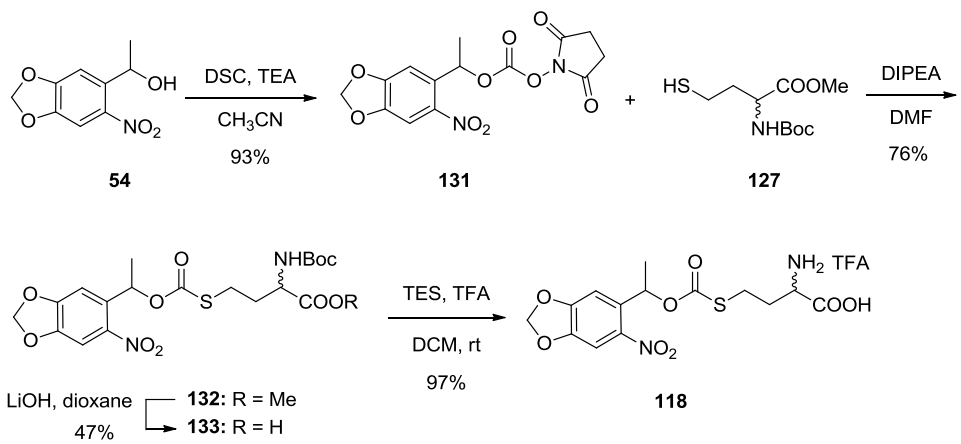


Scheme 12. Synthesis of the caged homocysteine **106**.

Decaging experiments were performed in an aqueous solution of caged homocysteine **106** in PBS buffer (0.1 mM, pH 7.4) and exposure to UV light (365 nm) at increasing time intervals (5, 10, 15 min), followed by TLC, HPLC, and mass spec analysis. All the

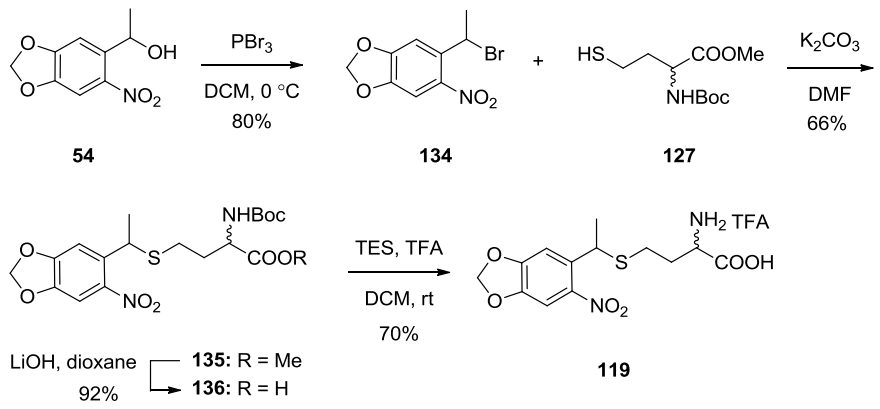
performed test experiments indicated that starting material was slowly consumed, but none of the attempted analysis (using either HPLC or mass spectrometry data) confirmed the recovery of homocysteine upon irradiation. Thus, further studies on the decaging of **106** after incorporation into protein are required to confirm the recovery of homocysteine upon photolysis.

Meanwhile, the caged homocysteine **106** was subjected to site-specific incorporation in a model protein mCherry-TAG-EGFP in mammalian cells and sfGFP in *E. coli*. Since **106** is structurally similar to the caged lysine **105**, it was incorporated by the same evolved PylRS.[344] With the success of the caged homocysteine **106** incorporation, the next step was to investigate the roles of carbamate functionality and chain length during unnatural amino acid incorporation. Therefore, we prepared the caged thiocarbonate and **118** caged homocysteine sulfide **119**. The synthesis of **118** was completed in four steps from the alcohol **54**, which was first treated with disuccinimidyl carbonate in the presence of triethyl amine in acetonitrile to obtain the succinimidyl carbonate **131** (Scheme 13). The carbonate **131** was reacted with Boc homocysteine methyl ester **127** in the presence of DIPEA in DMF to furnish the thiocarbonate **132** in76% yield, which upon hydrolysis with aqueous LiOH at 0 °C gave the corresponding acid **133** in 47% yield. The low yield was found to be mainly due to the partial hydrolysis of the thiocarbonate linkage. Finally, the thiocarbonate **133** was treated with TFA in DCM in the presence of TES as cation scavenger to provide the desired caged homocysteine thiocarbonate **118** as a yellow solid in 97% yield.



Scheme 13. Synthesis of the caged homocysteine thiocarbonate **118**.

Similarly, the synthesis of caged homocysteine sulfide **118** was accomplished in four steps from the alcohol **54** (Scheme 14). The alcohol was reacted with PBr_3 to obtain bromide **134** in 80% yield. The bromide **134** was reacted with Boc homocysteine methyl ester **127** in the presence of K_2CO_3 in DMF to deliver the sulfide **135** in 66% yield, which upon subsequent hydrolysis with LiOH gave the corresponding acid **136** in 92% yield. Finally, the thiocarbonate **136** was treated with TFA in DCM in the presence of TES as a cation scavenger to provide **119** in 70% yield.



Scheme 14. Synthesis of the caged homocysteine sulfide **119**.

Upon completion of the syntheses, the sulfide **119** and thiocarbonate homocysteine **118** were tested for incorporation using the evolved orthogonal PylRS in *E. coli*. However, neither of the caged homocysteines **118** or **119** was recognized by the synthetase. The reduction in chain length as well as the loss of carbamate functionality prevented the recognition of these amino acid derivatives by the aforementioned cellular machinery, requiring the engineering of a new PylRS through directed evolution.

5.4. Genetic code expansion with caged cysteine using PylT/PylRS pair

The caged cysteine **107** shares structural similarities with the caged homocysteine **106**, having only one less methylene unit in the side chain. Due to the successful incorporation of the caged homocysteine **106** into a protein, the incorporation of other structurally similar caged cysteines will be investigated using the same cellular mechinary.

Cysteine is the least abundant amino acid in proteins (1.9%) although it is present in almost all proteins,[400] and is one of the most investigated amino acid because of the unique chemistry of the thiol group in cellular functions.[394, 401-403] The functional roles of cysteine include 1) ubiquitination,[404, 405] 2) caspase mediated apoptosis,[406, 407] 3) maintenance of cellular redox potential,[408] 4) defense against oxidative stress,[409-413] 5) contribution in redox signaling pathways,[414, 415] and 6) protein trafficking and localization.[416-419]

The unique chemistry of the cysteine side chain residue is accountable for the aforementioned broad range of functional roles.[420] Intrinsic properties of the thiol group, such as the large atomic radius of sulfur, environmental sensitivity to pKa, variable oxidation states (Figure 38),[409, 421] and metal chelating properties offer significant advantages to accomplish diverse functions in proteins and enzymes.[400] Cysteine also offers various post transcriptional modifications including alkylation, nitrosylation, and oxidation to various forms of acids.[421]

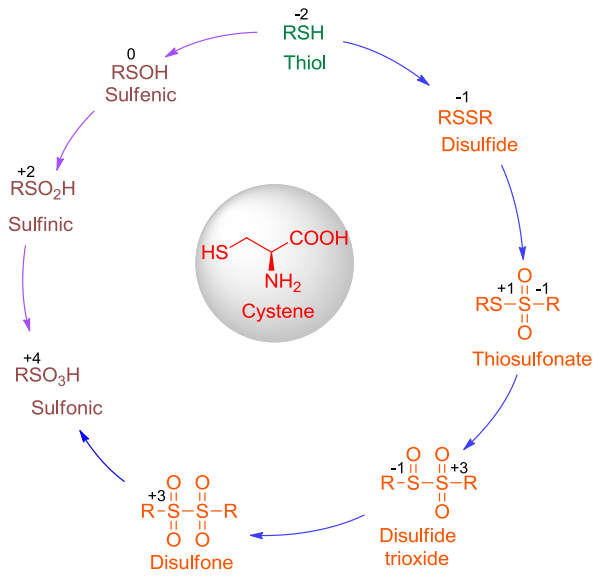


Figure 38. Cysteine chemistry illustrated in various oxidation states.

The equilibrium between a cysteine and cysteine-glutathione difulfide is important to balance cellular redox potential. Deprotonation of thiol group results in a cysteine nucleophile that is crucial for a sulfide mediated nucleophilic reaction (a key step of peptide hydrolysis). The pKa of a thiol group of cysteine is approximately 8.5 but may be lowered to 3.5 in a catalytic environment, allowing enzymatic activity over a wide pH range. The thiol group in cysteine protease forms a 'charge relay dyad' between an imidazolium ring of a histidine residue and the sulfur atom resulting in a hard cysteine nucleophile (pKa 3.5), which catalyzes the enzymatic hydrolysis of the peptide bonds.[421]

Furthermore, the thiol group can be oxidized into a more stable disulfide (S–S) bond. The disulfide bond exists in a dynamic and reversible state in the presence of glutathione or other reactive thiols.[408] Studies show that the cysteine-glutathion disulfide couple system, a

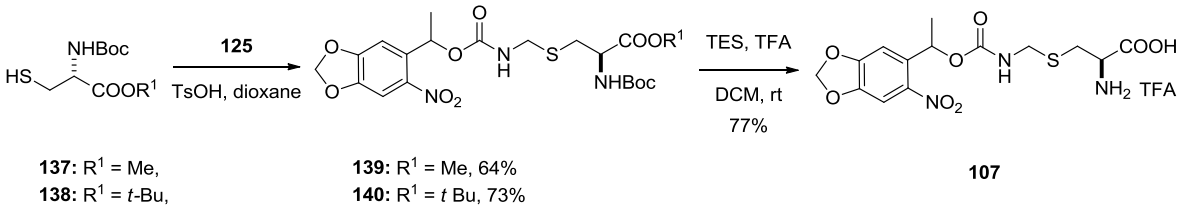
thiol based redox switch, is a key component for some gene regulation.[402] Alternatively, the formation of a disulfide bond between two cysteine residues results in intra- or intermolecular disulfide linkages that can result in a substantial change in protein folding or dimerization thereby altering protein function.[422] Since cysteine posttranslational modifications are crucial and informative on a variety of intermediates within cellular processes, much progress on protein structure and function has already been achieved through various applications of cysteine modifications.[401, 423-426]

To shed light on the protein dynamics and their roles in cells, a number of state-of-the-art methods have been developed. Recently, a quantitative method, known as isotopic tandem orthogonal proteolysis-activity-based protein profiling (TOP-ABPP), has been used to investigate reactive and nonreactive cysteine residues in a native proteome.[427] Other prevalent methods of site-specific cysteine modification include covalent modification of side chains,[428, 429] covalent labeling by enzymes,[430] affinity labeling,[431, 432] and amino acid mutagenesis via amber suppression methods.[433]

5.4.1. Synthesis of a caged cysteine analogue

In this section, we report the synthesis of a caged cysteine **107** and its site-specific incorporation into proteins. The synthesis was completed in three steps starting from previously synthesized alcohol **125**, which was used to prepare caged homocysteine **106** (Scheme 12). The alcohol **125** was reacted with Boc-protected cysteine methyl ester **137** in dioxane in the presence of a catalytic amount of TsOH, using similar conditions used for preparing caged homocysteine (*vide supra*), to get caged Boc-cysteine methyl ester **139** in

64% yield. Hydrolysis of **139** using aqueous LiOH furnished the corresponding acid (structure not shown) under unaticipated difficulties. The hydrolysis reaction not only formed product in low yield but also complicated purification due to the decomposition of the starting material. Thus, no further optimization was pursued. Alternatively, the acid-mediated coupling of Boc-protected cysteine *t*-butyl ester **138** gave the corresponding caged cysteine ester **140** in 73% yield. Finally, both Boc and *t*-butyl groups were deprotected in a single step using TFA in DCM (1:1) in the presence of TES as a cation scavenger to obtain the desired product **107** in 77% yield.



Scheme 15. Synthesis of caged cysteine **107**.

With the caged cysteine **107** in hand, the site-specific incorporation was performed by Jihe Liu in the Deiters laboratory. Because of structural similarities between caged cysteine and caged lysine, the same PylRS[344] was used. As expected, the caged cysteine **107** was recognized and successfully incorporated into protein, but with a lower yield then the caged homocysteine **107**. Further investigation of light activated protein function using both caged homocysteine and caged cysteine is in progress in the Deiters Laboratory.

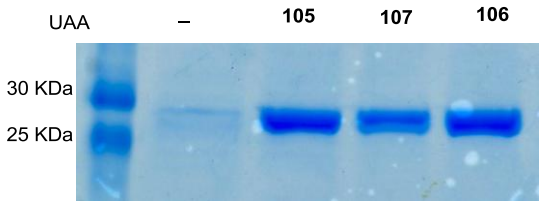
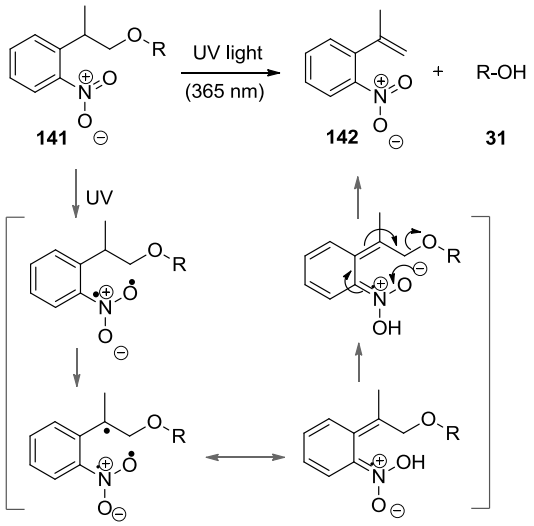


Figure 39. SDS-PAGE indicating site-specific incorporation of caged cysteine **107** and caged homocysteine **106** into sfGFP using PylRS. Caged lysine **105** and no UAA were used as positive and negative controls, respectively.

5.5. Genetic code expansion with NPP caged lysine using PylT/PylRS pair

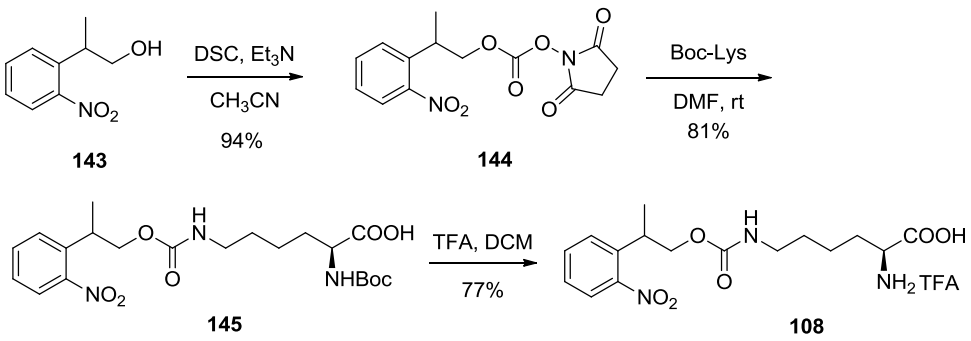
The 2-(*o*-nitrophenyl) propanol (NPP) caged lysine **108** is also a single-photon caged lysine that is similar to our previously reported NPE caged lysine **105**.[344, 396] The NPP caging group is preferred to the NPE caging group for two main reasons. First, photolysis of the NPP group is twice as fast as that of the NPE caging group.[434] Second, unlike the NPE caging group, it does not yield a toxic *o*-nitroso carbonyl byproduct upon UV irradiation. Photolysis of the NPP caging group proceeds through a β-elimination pathway resulting in an active substrate and the *o*-nitrostyrene, which is a relatively less toxic byproduct (Scheme 16).[435]



Scheme 16. Mechanism of photolysis of the NPP caged substrate.

5.5.1. Synthesis of NPP-caged lysine analogue

The NPP caged lysine **108** was synthesized in three steps starting from the alcohol **143** prepared following a literature procedure.[436] The alcohol **143** was subsequently reacted with DSC in the presence of triethyl amine in acetonitrile to give the corresponding succinimdyl carbonate **144** in 94% yield. Boc-lysine was reacted with the activated carbonate **144** in DMF at room temperature, delivering the caged lysine **145** in 81% yield. Boc-deprotection of **145** using a TFA:DCM mixture (1:1) in the presence of TES resulted in the crystalline TFA salt **108** in 92% yield.



Scheme 17. Synthesis of NPP caged lysine **108**.

The site-specific incorporation of the newly synthesized NPP-caged lysine **108** was then performed in both the Chin Laboratory at MRC-LMB and by Ji Luo in Deiters laboratory in mammalian cells. Preliminary results showed that **108** has been site-specifically incorporated into the model protein mCherry-TAG-EGFP by using an evolved PylRS variant V10 with the mutation Y271A in human embryonic kidney (HEK293) cells, which were transfected with two constructs.[351] The first one is a reporter construct with *N*-terminal mCherry gene, a linker with an amber stop codon (TAG), a *C*-terminal enhanced GFP gene (EGFP), and an HA-tag coding sequence. The second one is the PylT/PylRS expression plasmid. As expected, the mCherry gene was expressed regardless of the presence or the absence of NPP-lysine but a full-length protein mCherry-TAG-EGFP-HA was expressed only in the presence of NPP-lysine **108**, indicating specific incorporation of this unnatural amino acid (Figure 40).

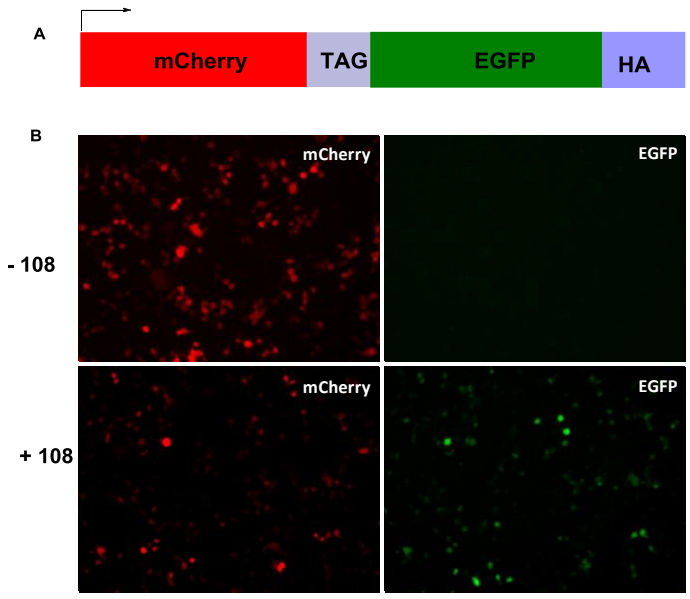


Figure 40. NPP-lysine **108** incorporation in mammalian cells. A) The mCherry-TAG-EGFP reporter construct. B) Fluorescence micrographs of HEK293 cells expressing mCherry-TAG-EGFP-HA, PyltRNA, and PylRS. A full length protein is expressed only in the presence of NPP-lysine **108**.

5.6. Genetic code expansion with coumarin-caged lysine using PylT/PylRS pair

Two-photon excitation techniques offer numerous advantages for biochemical investigation at the cellular level,[115, 116] including 1) higher tissue penetration of up to a few hundred micrometers, 2) increase in fluorescence quantum efficiency, 3) minimum photobleaching, and 4) a reduced amount of photodamage.[104, 115, 116, 437] Recent use of two-photon or multi-photon spectroscopy using fluorescent probes has dramatically improved the understanding of protein functions.[438-440]

Furuta et al. reported the first two-photon uncaging group based on a coumarin core structure.[117] They reported that the 6-bromo-7-hydroxycoumarinmethyl (**44**) caging group readily undergoes two-photon photolysis with a δ_u value 0.72 GM at 740 nm. Later Bendig et al. reported a decaging mechanism of Bhc caged substrates **44** via an ion-pair solvent-assisted heterolysis pathway (Scheme 3).[118, 119]

Coumarin based chromophores are one of the most investigated class of fluorescent labeling probes for biological studies.[100, 441-443] The protocols on coumarin functionalization and their photophysical properties are well established.[117, 444-448] There are several advantages of using coumarin-based chromophores in cellular imaging applications. In general, coumarins 1) exhibit bright and stable fluorescence in aqueous environments, 2) are fairly resistant to photobleaching, and 3) have high decaging efficiency.[449]

Lately, coumarin labeling techniques have been an appealing strategy to label biomolecules via site-specific incorporation of coumarin UAAs through unnatural amino acid mutagenesis.[450] In particular, the coumarin amino acid **146** has been genetically encoded into FtsZ, a tubulin protein in bacteria.[450] The same coumarin amino acid has also been genetically encoded into the protein colicin E1 channel domain in *E. coli* to investigate the membrane topology of the channel.[451] To further extend the scope of genetic code expansion, we report the synthesis of coumarin lysines **147**–**149** and their incorporation into proteins.

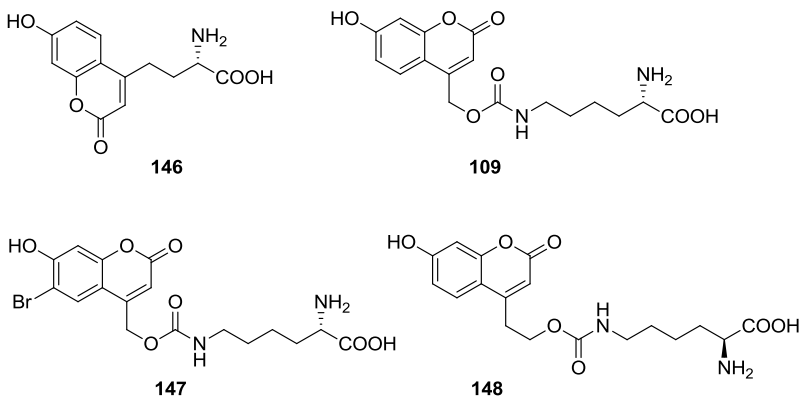


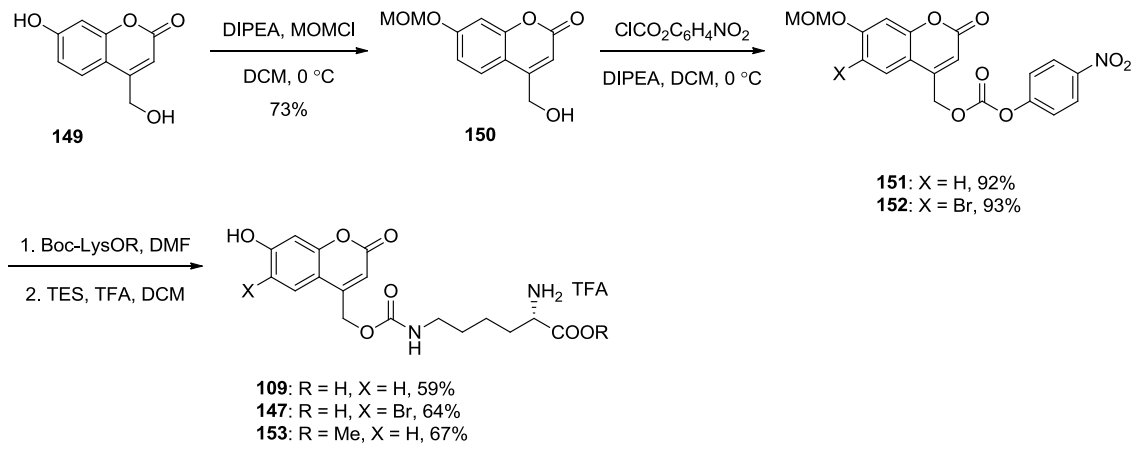
Figure 41. Newly synthesized coumarin lysine amino acids **109**, **147**–**148**, and previously reported coumarin amino acid, **146**.

5.6.1. Synthesis of coumarin lysine analogues

Hydroxycoumarin lysine **109** was synthesized in 3 steps from the reported coumarin alcohol **149**. We prepared the alcohol **149** in two steps starting from ethyl 4-chloroacetoacetate and resorcinol according to a literature procedure.[452] The alcohol **149** was then treated with MOMCl in the presence of DIPEA in DCM to achieve MOM protection of the phenolic hydroxy group, delivering **150** in 73% yield. The alcohol **150** was activated using nitrophenyl chloroformate in the presence of DIPEA in DCM to obtain the corresponding carbonate **151** in 92% yield. The carbonate **151** was then reacted with Boc-lysine in DMF followed by treatment of the crude product with TFA in DCM (1:1) with TES as a cation scavenger. Both the MOM and the Boc protecting groups were removed in a single step resulting in the hydroxycoumarin lysine amino acid **109** in 59% yield as a crystalline solid.

Hydroxycoumarin lysine **109** was successfully incorporated site-specifically into a model protein through the use of an evolved PylRS in mammalian cells by Ji Luo in the Deiters lab. However, it resulted in a low yield possibly due to the poor solubility of **109**, which was found to precipitate in cells. Thus, to increase the solubility, we also synthesized the corresponding methyl ester analogue **147**. The nitrophenyl carbonate **151** was treated with Boc-lysine methyl ester followed by deprotection of MOM and Boc groups as described for the coumarin lysine **109** above. However, the methyl ester analogue **153** was found to be toxic to cells and also did not improve the solubility.

Previous studies showed that the decaging efficiency of coumarin may be improved by having a bromo substitution on its aromatic ring (*vide supra*). Therefore, we synthesized the bromohydroxy coumarin lysine **147** from the known nitrophenyl carbonate[453] **152** (Scheme 18). The carbonate **152** was synthesized in four steps from commercially available ethyl 4-chloroacetoacetate and bromoresorcinol following a literature procedure.[454] The nitrophenyl carbonate **152** was then reacted with Boc Lysine in DMF followed by treatment of the crude product with TFA in DCM (1:1) along with TES as a cation scavenger. Both the MOM and the Boc protecting groups were removed in a single step resulting in the bromohydroxycoumarin lysine amino acid **147** in 64% yield as a white crystalline solid.



Scheme 18. Synthesis of coumarin lysines **109**, **147**, and **153**.

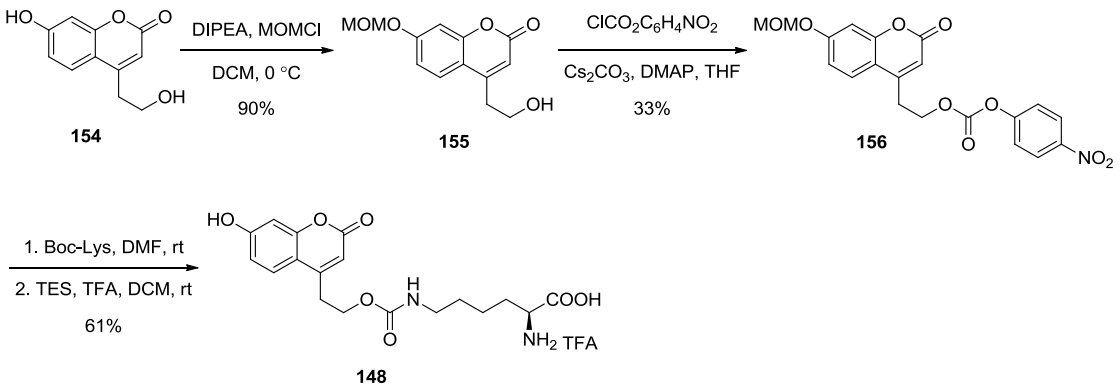
After the synthesis of **148**, Ji Luo in the Deiters lab incorporated the amino acid into the model proteins sfGFP (Y151TAG) in *E. coli* and mCherry-TAG-EGFP in mammalian (HEK293T) cells by using a PylRS with the mutations Y271A and L274M. LRMS-ESI analysis of sfGFP-**147** showed a major and a minor species with a mass of 28445.97 Da and 28524.52 Da, respectively. The minor peak indicates the incorporation of **147** in sfGFP (observed at 28524.52 Da and calculated 28523.69 Da) where as the major species represents sfGFP-**147** dependent protein with a loss of about 79 Da. However, no loss in mass was observed on LRMS-ESI analysis of the sfGFP-**109** (observed 28446.60 ± 0.50 Da and calculated 28446.28 Da). Although the exact reason for the loss of 79 Da in protein is not known at this point, the loss is consistent with the loss of a bromine atom from **147**. Further studies are in progress.

Figure 42. SDS-PAGE indicating site-specific incorporation of hydroxycoumarin lysine **109** and Bhc lysine **147** (lane 2) into sfGFP using PylT/PylRS pair.

In order to utilize coumarin fluorescence for protein labeling, coumarin fluorescent lysine **148**, a higher homologue of **109**, was synthesized. The addition of one more methylene unit at the 4-position was anticipitated to inhibit decaging but not affect its fluorescent properties. Thus, the coumarin lysine **148** was synthesized in three steps from the alcohol **154**, following a literature procedure involving two steps from resorcinol and acetonedicarboxylic acid.[455]

The alcohol **154**, as described before in Scheme 18, was treated with MOMCl in the presence of DIPEA in DCM to achieve selective MOM-protection of the phenolic hydroxy group resulting in **155** in 90% yield. The MOM-protected alcohol **155** was activated using nitrophenyl chloroformate in the presence of Cs_2CO_3 in DCM to obtain the carbonate **156** in 33% yield. Use of different bases (such as K_2CO_3 or DIPEA) and other solvents (THF or dioxane) did not improve the yield. Additionally, the poor solubility of the product in common organic solvents posed issues during purification using typical column chromatography. Attempted purification of the product by crystallization did not improve the yield either. The carbonate **156** was reacted with Boc lysine in DMF and subsequently treated with TFA in DCM (1:1) with TES as a cation scavenger, removing both the MOM

and the Boc protecting groups in a single step and delivering the amino acid **148** in 61% yield as a crystalline solid.



Scheme 19. Synthesis of coumarin lysine **148**.

The coumarin lysine **148** was then incorporated into sfGFP (Y151TAG) using the same evolved orthogonal PylRS that was used for **109** in *E. coli*. The incorporation experiments were performed by Ji Luo in the Deiters laboratory and are still in progress.

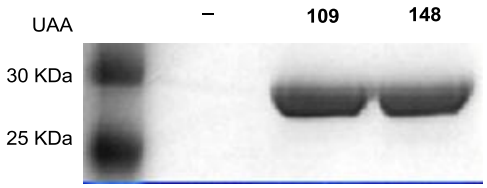


Figure 43. An SDS-PAGE gel indicating site-specific incorporation of **148** into sfGFP using PylT/PylRS. The lysine **109** and no UAA were included as positive and negative controls, respectively.

5.7. Genetic code expansion with fluorescent lysines using PylT/PylRS pair

Protein labeling with fluorescent tags has provided significant insight into protein dynamics, protein localization and protein-protein interactions.[456] Fluorescent probes have been used to investigate real time dynamics and cellular functions in *vivo*.[457] Numerous protein labeling tags or small molecules have been discovered over the decades to study cellular organization.[313, 458] These methodologies have remarkably expanded the understanding of molecular processes at the cellular level and the progress on methods of molecular imaging have illustrated an immense future prospect of labeling probes.[438, 439, 441] A number of protein and peptide tags for fluorescent labeling have been developed by assessing their advantages or disadvantages for biological relevance (*vide supra*).

While protein and peptide tags have specific advantages and disadvantages, they should generally be[457, 459] 1) small in size to minimize interference with the molecule of interest, 2) specific in order to label only the intended target, 3) quantitative and fast reacting, 4) permeable to enable labeling in cells, and 5) non-toxic. Smaller peptide tags for fluorescent labeling, such as the tetracysteine tag or His tag, have a significantly smaller size in comparison to fusion protein domain tags.[460] However, these smaller tags suffer from background staining and require substantial washing to optimize fluorescence imaging. Hence, there are many factors to consider when choosing an optimal fluorescent tag.

Instead of conducting a two-step process by 1) creation of recombinant proteins containing a labeling tag and 2) selective labeling of the protein with a suitable fluorophore, a fluorescent amino acid could be genetically encoded into the protein by using amber suppressor methods (Chapter 5.6).[328]

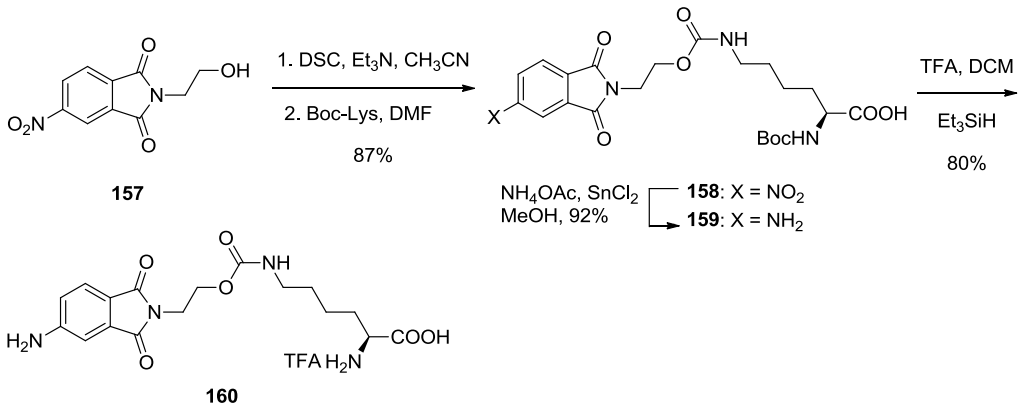
Here, we expand the fluorescent labeling of proteins via site-specific incorporation of aminophthalimides, fluorophores that are highly sensitive to their local environment.[441, 461, 462] The photophysical properties of the aminophthalimides have been widely studied because of their potential use as sensors for cellular applications.[463-465] For example, dimethyl aminophthalimide was incorporated into nucleosides for studying DNA-protein interactions by exploiting its exceptional sensitivity to the environment.[466] Both the 4-aminophthalimide **160** and 4-(dimethylamino)phthalimide **168** are highly sensitive fluorophores displaying solvatochromism, a phenomenon by which a fluorophore changes its photophysical properties such as fluorescence lifetime, emission wavelength, and quantum yield based in response to the polarity of the medium.[463]

5.7.1. Synthesis of 4-aminophthalimide and 4-(dimethylamino)phthalimide lysine analogues

Here, we report the synthesis of fluorescent lysine analogues bearing aminophthalimide probes namely 4-aminophthalimide lysine **160** and 4-(dimethylamino)phthalimide lysine **168**. Both the **160** and **168** are primarily chosen because of their 1) extraordinary sensitivity to the change in their local environment, 2) small size, which has always been advantageous for unnatural amino acid recognition by PylRS, and 3) the straight forward and high yielding synthesis.

Our synthesis of 4-aminophthalimide lysine **160** was accomplished in three steps from the alcohol **157**. The alcohol **157** was synthesized from commercially available 4-nitrophthalic acid.[467, 468] The alcohol **157** was activated with dissccinimidyl carbonate in the presence of triethylamine in acetonitrile and the resulting product, when reacted with Boc-

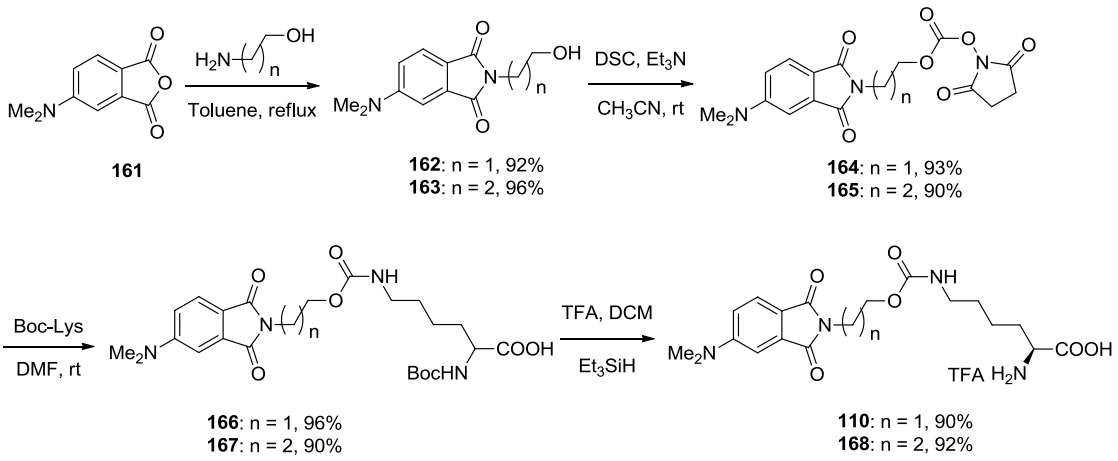
lysine, in DMF gave the corresponding 4-nitrophthalimide lysine **158** in 87% yield. The reduction of nitro group on **158** with tin(II) chloride in the presence of ammonium acetate in MeOH furnished the amino lysine **159** in 92% yield.[469] The amino lysine **159** was treated with TFA in DCM (1:1) in the presence of TES as a cation scavenger, to provide the aminophthalimide lysine **160** in 80% yield as a crystalline solid.



Scheme 20. Synthesis of 4-aminophthalimide lysine **160**.

The 4-aminophthalimide lysine **160** was then site-specifically incorporated into myoglobin using the evolved orthogonal PylRs variant EV3 in *E. coli* (Figure 44). This PylRS variant contains the mutations L274V, C313V, and M315Q. Similarly, dimethyl aminophthalimides are found to be useful in a variety of biological applications (*vide supra*). Particularly, dimethyl substitution at nitrogen substantially alters both the bulkiness (hydrophobicity) and the basicity (caused from electron donating methyl groups) thereby altering the photophysical properties.[470, 471] Thus, we investigated the synthesis of dimethyl aminophthalimide lysines **110** and **168**, and their site-specific incorporation into proteins.

Both the 4-(dimethylamino)phthalimide lysines **110** and **168** were synthesized from a known anhydride[472] **161** (Scheme 21). The only difference between **110** and **168** is the linker chain length between the fluorophore and the lysine. The alcohol **162** was prepared according to a reported procedure[473] and was subsequently reacted with DSC in the presence of triethylamine in acetonitrile to deliver the corresponding succinimidyl carbonate **164** in 93% yield. Boc-lysine was reacted with the carbonate **166** in DMF at room temperature to obtain the lysine **166** in 96% yield. Finally, the Boc-deprotection of **166** using a TFA/DCM mixture (1:1) in the presence of TES resulted in **110** in 90% yield as a white crystalline solid. The amino acid **168** was synthesized using the same procedure that was used for the synthesis of **110 (**Scheme **21**).



161

162: n = 1, 92%
163: n = 2, 96%

164: n = 1, 93%
165: n = 2, 90%

166: n = 1, 96%
167: n = 2, 90%

110: n = 1, 90%
168: n = 2, 92%

Scheme 21. Synthesis of the 4-(dimethylamino)phthalimides lysines **110** and **168**.

Both 4-(dimethylamino)phthalimides **110** and **168** were tested for their site-specific incorporation by Ji Luo in the Deiters laboratory. Lysine derivatives **110** and **168** were

successfully incorporated into myoglobin by using the same evolved orthogonal PylRS Variant EV3 in *E. coli* (Figure 44) that was used for **160** and the further studies are currently in progress.

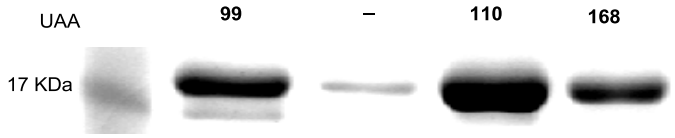


Figure 44. SDS-PAGE indicating site-specific incorporation of the fluorescent lysines **110** and **168** into myoglobin using the PylRS variant EV3. Positive (**99**) and negative (– UAA) were included as well.

5.8. Genetic code expansion with azobenzene lysine using PylT/PylRS pair

Azobenzene **78** exists in either a planar thermally stable *trans* (*E*) or a compact meta-stable *cis* (*Z*) conformation (Chapter 3, Figure 22).[269] The *trans-cis* isomerization of azobenzene can be achieved by exposing it to UV light (365 nm) and the *cis-trans* isomerization may be accomplished by using either blue light or allowing it to relax back to the thermally stable state.[270] After the first report on *cis-trans* isomerization of azobenzene in a model membrane,[278] there has been significant progress regarding the studies on azobenzene photoswitches for biomolecules.[268, 269, 276, 277]

In general, because of its exceptional photophysical properties (*vide supra*), studies on azobenzene covered a broader context including both material as well as biological areas including nucleic acids, peptides, proteins, and lipids.[270, 279-281] Azobenzene represents a versatile biological switch because of its 1) orthogonality with biological systems, 2) robustness with extraordinary photophysical properties (*vide supra*), and 3) ease of protein side chain manipulation. But they also have some limitations, such as reduction of an azo moiety in cells and low quantum yield due to its reversibility at photo steady states.[269] The azo moiety may be reduced in the cellular environment particularly because of the abundance of glutathion in the 1-10 mM range.[474] Glutathion, a tripeptide, is a defensive agent against oxidative stress such as reactive oxygen species or free radicals.[475] In particular, electron poor azobenzene derivatives (acid, amide, ester, or ketone at 4 or 4′ position) may easily be reduced in cells due to the presence of glutathion.[279]

Several approaches have been employed to integrate azobenzene photoswitches into proteins. Particularly, protein functionalization with an azobenzene may be achieved by 1) solid-phase peptide synthesis followed by native chemical ligation,[279, 476] 2) performing site-directed mutagenesis to allow thiol nucleophilic substitution with halo-azobenzene,[280, 281] 3) employing a thiol-ene click reaction,[477] 4) Staudinger-Bertozzi ligation,[478] and 5) nonsense amber suppressor method using the cellular mechinery.[479]

Accomplishments on azobenzene photoswitches further illuminate the future relevance of photoswitching strategies to monitor biological functions. Some of the recent studies on azobenzene photoswitching with light may include 1) the control of enzymatic activities of a restriction enzyme,[480] 2) the design of light controlled DNA-binding

proteins,[277] 3) the development of photochromic ion channels and AMPA receptors,[481, 482] and 4) the regulation of cell adhesion.[483]

5.8.1. Synthesis of photoswitchable azobenzene lysine analogues

We wanted to extend the scope of genetic code expansion by incorporating the photoswitchable azobenzene lysine derivatives **169–171** (Figure 45 and Scheme 22). The rationale behind each of the amino acid selections was primarily based on its structure-property relationship. It was anticipated that amino acid **169** having a carbamate at the *para* position would likely be incorporated with greater efficiency because of its small size. Yet, the activation of the 4-(phenyldiazenyl) phenol was unsuccessful, most likely due to carbonate (dimer) formation (structure and reaction scheme not shown). Attempted experiments using DSC, diphosgene, and *p*-nitrophenyl chloroformate with a variety of bases such as K_2CO_3, DIPEA, and TEA in different solvents (e.g., THF, DCM and CH_3CN) did not yield the desired carbonate intermediates (not shown), an intermediate en route to the synthesis of **169**. Alternatively, the activation of (4-(phenyldiazenyl)phenyl)methanol, a precursor to **170** using aforementioned conditions did not improve yields beyond 28% (structure not shown). From a synthetic perspective, it would be better to explore alternative candidates. Therefore, we focused on the azobenzene lysine **171**.

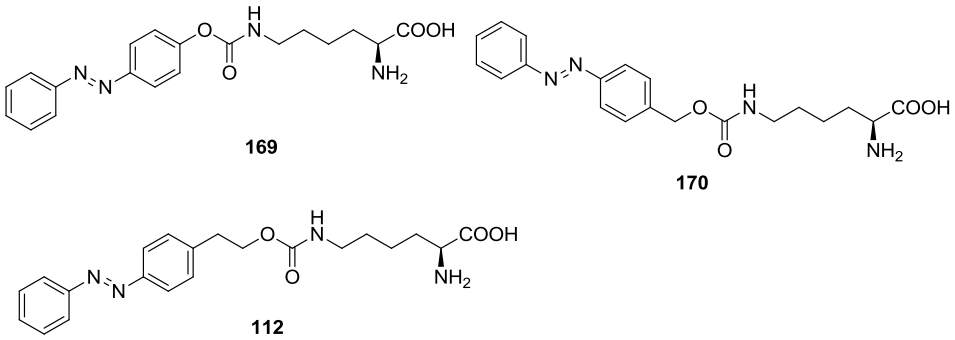
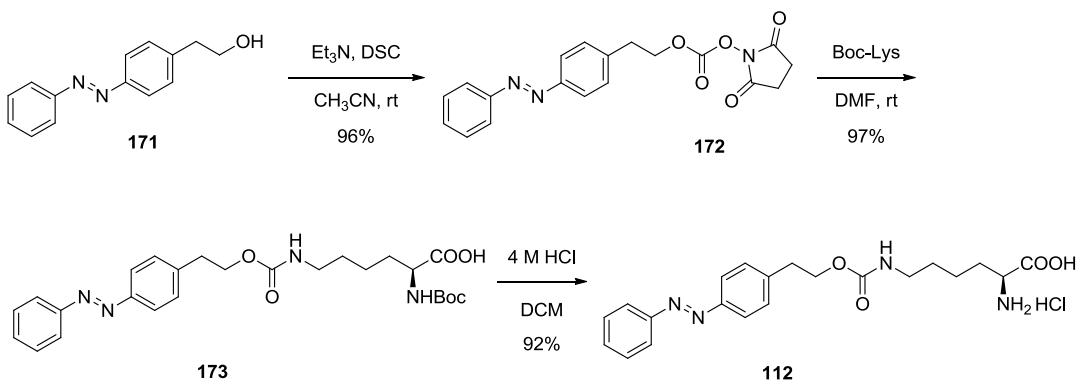


Figure 45. Azobenzene lysine analogues **112**, and **169**–**170**.

The synthesis of the azobenzene lysine **112** was accomplished in four steps starting from commercially avilable 2-(4-aminophenyl)ethanol (Scheme 22). The 2-(4-aminophenyl)ethanol was converted into the alcohol **171** using nitrosobenzene in acetic acid following a literature procedure.[484] Activation of the alcohol **171** with disuccinimide carbonate in the presence of triethylamine resulted in **172** as an orange solid in 96% yield. The activated carbonate **172** was reacted with Boc-lysine in DMF to deliver **173** in 97% yield. Our attempt to prepare a TFA salt of **112** from **173** using standard Boc deprotection conditions using TFA in DCM caused decomposition of the starting material. Finally, Boc deprotection of **173** was acomplished using 4 M HCl at room temperature, delivering the hydrochloride salt of azobenzene lysine **112** in 92% yield as a crystalline orange solid.



Scheme 22. Synthesis of the azobenzene lysine **112**.

After the synthesis of the azobenzene lysine **112**, site-specific incorporation was tested (by Ji Luo and Jihe Liu in the Deiters laboratory) into proteins using a panel of evolved PylRSs. Unfortunately, the lysine **112** was not incorporated by any of the tested synthetases. Thus, further engineering of the PylRS may be required for achieving site-specific incorporation of **112** into proteins.

5.9. Genetic code expansion with dithiolane lysine using PylT/PylRS pair

Dithiolanes are five membered heterocyclic disulfide compounds, which display potent antioxidant properties.[485, 486] Lipoic acid (**174**)**,** asparaguasic acid (**175**)**,** and tetranorlipoic acid (**176**) (see structures on Figure 46) represent a few examples of dithiolanes.[487, 488] Among them, lipoic acid (**174**) has been widely used as a model compound in a variety of studies regarding its antioxidant activities.[489, 490] Additionally, dithiolanes possess a wide range of potential applications that could be explored by studying their redox and metal chelating properties.[491]

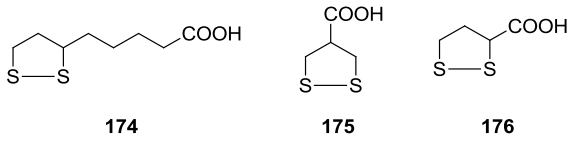


Figure 46. Lipoic acid (**174**), asparaguasic acid (**175**), and tetranorlipoic acid (**176**).

Dithiolane moities can be used to generate two gold-sulfide bonds in biomolecule-gold nanoparticle conjugates.[492-494] In general, three different approaches have been adopted to prepare biomolecule functionalized gold-nanoparticle (AuNPs) conjugates using lipoic acids (Figure 47). In the first method, lipoic acid is incubated with gold nanoparticles under suitable conditions. A stable and uniform colloidal mixture can be prepared with no aggregation of the gold particles by introducing repulsive negative charges of the carboxylate anions on nanoparticles. The carboxylic acid moieties on these nanoparticles are then coupled with molecules of biological interest (Figure 47a) to obtain functionalized nanoparticles.[495, 496] For example, a horseradish peroxidase-gold nanoparticle conjugate can be used as an electrochemical sensor.[497]

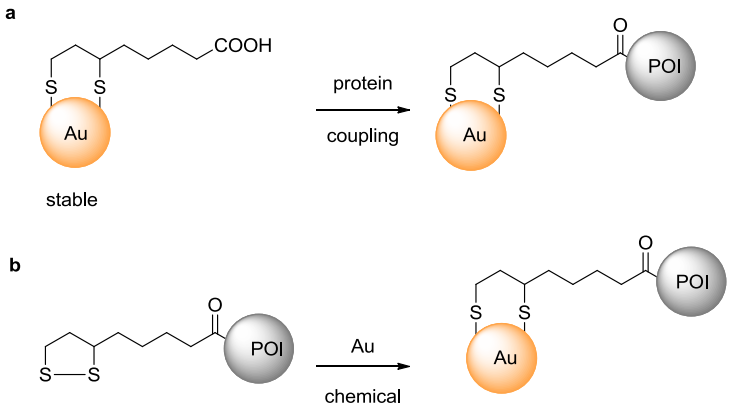


Figure 47. Strategies of immobilizing proteins on gold nanoparticles using a lipoic acid group. a) Lipoic acid-gold conjgate functionalization with protein. b) Lipoic acid functionalized protein can be immobilized onto gold nanoparticles (POI = protein of interest).

In the second approach, the lipoic acid probe is at first coupled to the molecule of interest such as a protein or nucleic acid via chemical or enzymatic methods.[498,499] The functionalized molecule is then immobilized onto gold nanoparticles under suitable conditions (Figure 47b).[500] Lipoic acid modified proteins have been conjugated with specific affinity fluorophore probes or quantum dots for the study of protein-protein interaction.[501, 502]

To our knowledge (as of April 2013), there is no report on site-specific incorporation of any dithiolane amino acid. Thus, we wanted to expand our lysine analogues by adding a new dithiolane lysine to explore the potential application of dithiolanes in proteins via unnatural amino acid mutagenesis. We proposed dithiolane lysines **111** and **177** (Figure 48) due to the relatively short linkers, which aids in the recognition of amino acids by the cellular

machinery. Upon its site-specific incorporation into proteins, the dithiolane probe may be investigated for 1) studying metal chelating properties, 2) sensing reactive oxygen species or oxidants such as hydrogen peroxide, hydroxyl radicals etc, 3) modulating thiol-redox switches, 4) inducing enzymatic activities, and 5) building protein-gold nanoparticle conjugates.

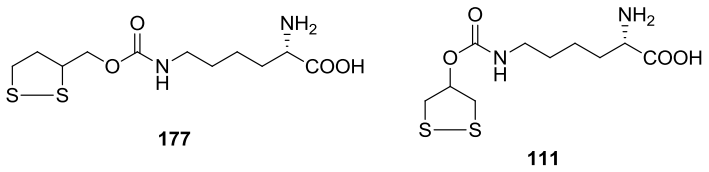
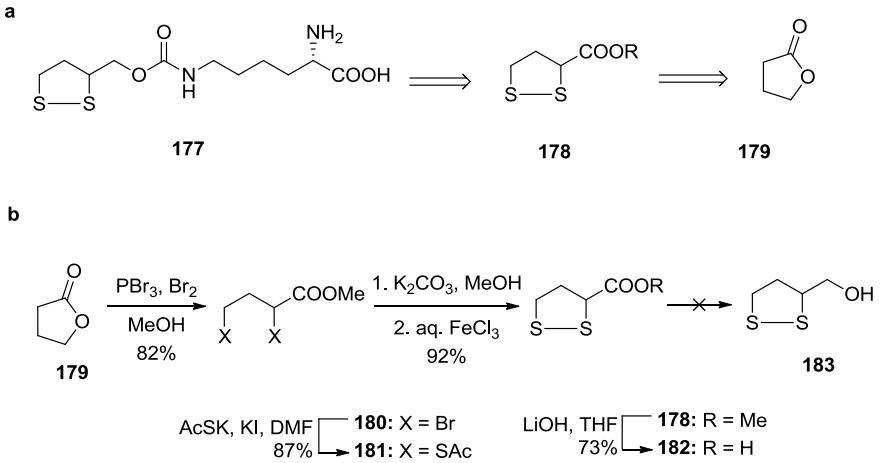


Figure 48. Proposed dithiolane lysines **111** and **177**.

5.9.1. Synthesis of dithiolane lysine analogues

We proposed that the dithiolane lysine **177** could be synthesized from the dithiolane acid **178**, which would be prepared from the commercially available lactone **179** (Scheme 23a). Thus, the synthesis commenced with the lactone **179** by reaction with Br_2 in CCl_4 in the presence of a catalytic amount of PBr_3 in MeOH under reflux, following a reported procedure,[503] to obtain the dibromide **180** in 82% yield. The dibromide **180** was then treated with potassium thioacetate in DMF to deliver the corresponding dithioacetate **181** in 87% yield.[504] The acetate groups of the dithioacetate **181** were then removed using K_2CO_3 in MeOH. The resulted crude reaction mixture was neutralized with aqueous HCl (1 M) and the reaction mixture was treated with aqueous $FeCl_3$ while stirred in a diethyl ether:water biphasic system, allowing dithiol oxidation to form the dithiolane **178** in 92% yield.[505 506]

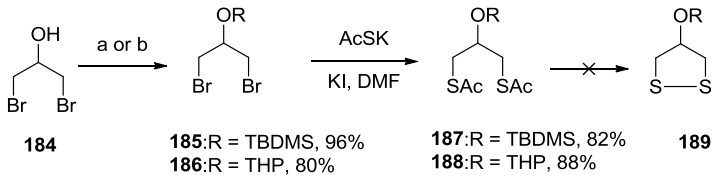
The ester **178** was hydrolyzed using aqueous LiOH to deliver the corresponding acid **182** in 73% yield. Next, we attempted the selective reduction of the carboxylic acid moiety of **182** using diborane, a chemoselective reducing agent.[507] The reduction of **182** simultaneously reduced both the acid functionality as well as the disulfide bond (structure not shown). Thus, the selective reduction of the acid moiety of **182** in the presence of such a labile disulfide linkage remained challenging.



Scheme 23. a) Retrosynthetic analysis of the proposed synthesis of dithiolane lysine **177**. b) Progress towards the synthesis of dithiolane alcohol **183**.

Alternatively, we started the synthesis of the symmetric dithiolane molecule **189** (Scheme 24) as a precursor for the desired lysine derivative **111**. Following literature procedures[508 509] the hydroxyl group of 1,3-dibromopropanol was protected either with a TBDMS or a THP group resulting in dibromide ethers **185** or **186**, respectively. The bromides **185** or **186** were separately reacted with potassium thioacetate in DMF in the

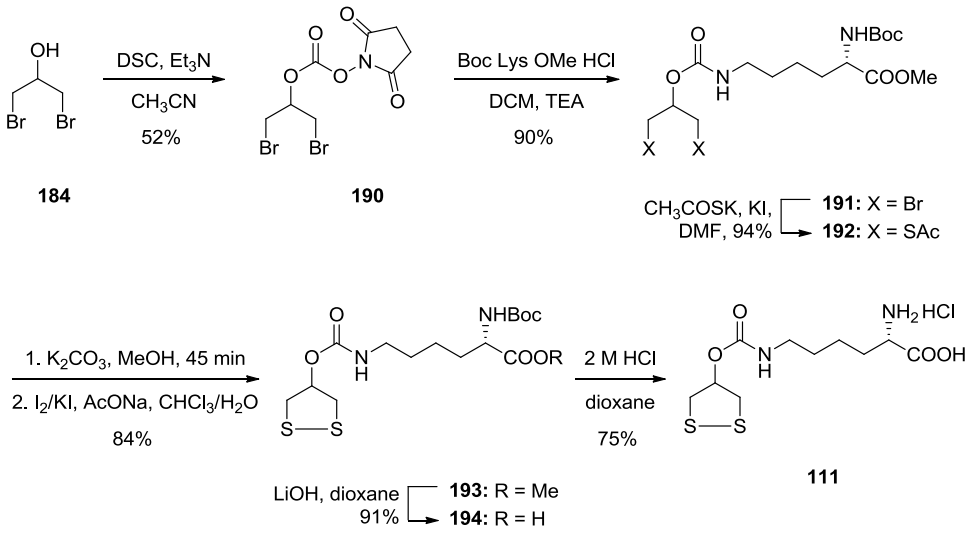
presence of KI to get the corresponding dithioacetates **187** or **188**, respectively. Both **187** and **188** were then subjected to a one-pot hydrolysis followed by oxidation as previously reported.[505 506] However, in our attempted experiments, starting materials decomposed and formed visible aggregates. We then tested an alternative synthetic route as shown in Scheme 25.



Scheme 24. Towards the synthesis of dithiolane **189**.

The outcomes of Schemes 23 and 24 indicated that electron-donating groups in these dithiolanes may not tolerate the oxidation step. With this in mind, we proposed a new approach for the synthesis of the dithiolane lysine **111** (Scheme 25), where the alcohol functionality would be protected with an electron-deficient group. Accordingly, the synthesis of the dithiolane lysine **111** was accomplished in six steps starting from 1,3-dibromopropanol. The propanol **184** was activated by reaction with disuccinimidyl carbonate in the presence of triethylamine in acetonitrile to obtain the dibromide carbonate **190** in 52% yield. The carbonate **190** was reacted with Boc-lysine methyl ester in DMF to deliver the dibromide lysine ester **191** in 90% yield. The dibromide **191** was treated with potassium thioacetate in the presence of KI in DMF to furnish the corresponding dithioacetate lysine ester **192** in 94% yield. The hydrolysis of dithioacetate groups in **192** using K_2CO_3 in MeOH

followed by oxidation using iodine in the presence of sodium acetate in a $CHCl_3$:H_2O biphasic system was performed as reported previously[510] and resulted in the desired dithiolane lysine ester **193** in 84% yield. The dithiolane ester **193** was treated with aqueous LiOH in dioxane to furnish acid **194** in 91% yield. Finally, the Boc group was deprotected by treating **194** with aqueous HCl to obtain the desired dithiolane lysine **111** as a crystalline solid in 75% yield.



Scheme 25. Synthesis of the dithiolane lysine **111**.

The newly synthesized dithiolane lysine **111** was then tested for its site-specific incorporation into proteins. In experiments performed by Jihe Liu in the Deiters lab, the PylRS variant EV13-1, containing the mutation Y349F, was used to incorporate **111** into sfGFP (Y151TAG) in *E. coli*. The SDS-PAGE shown in Figure 49 demonstrates the site-

specific incorporation of **111** into sfGFP. Further, studies including surface immobilization of sfGFP-dithiolane on the gold surface are currently being explored by Jihe Liu.

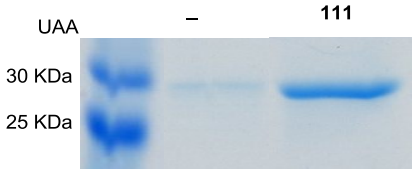


Figure 49. SDS-PAGE indicating site-specific incorporation of the dithiolane lysine **111** into sfGFP using the PylRS variant EV13-1.

5.10. Genetic code expansion with a spin labeled lysine using PylT/PylRS pair

Electron paramagnetic resonance (EPR) spectroscopy based on nitroxide spin labeling of biomolecules is one of the common techniques to study the localization and dynamics of biomolecules.[511, 512] EPR spectroscopy is exceptionally sensitive to the anisotropy of the spin probe in the presence of an external magnetic field and provides information on the local environment.[511, 513] In the presence of an external magnetic field, the spin probe provides an EPR signal with hyperfine splitting (Figure 50).[511] The line shape simulations of the EPR spectra enable investigations on global molecular tumbling as well as its conformation.[511, 514]

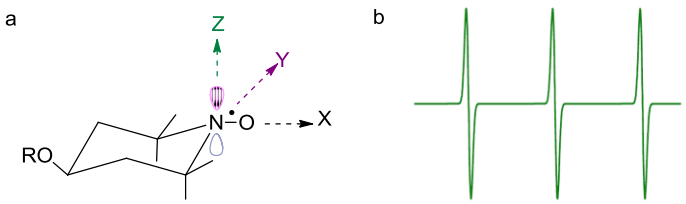


Figure 50. Illustration of anisotropy of the electron spin (in Z-axis) in a nitroxide and b) a typical EPR signal of a TEMPO free radical. R = H or substrate.

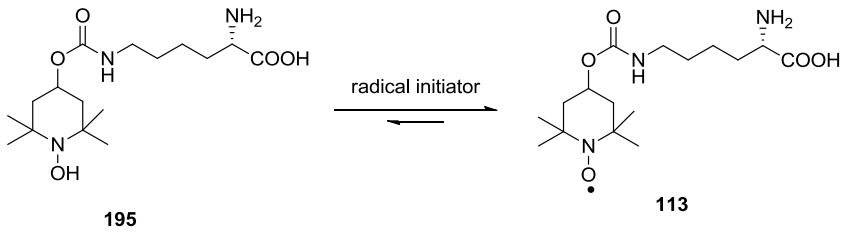
The Site Directed Spin Labeling-EPR (SDSL-EPR) method has been used to study membrane proteins for investigating their topographies, orientation onto the membrane surface, and electrostatic potentials.[512, 515] The SDSL-EPR technology has been used mainly because of 1) sufficient sensitivity to measure molecular dynamics on a nanosecond time scale,[516, 517] 2) feasible to use physiological measurement conditions,[516] 3) its ability to study biological specimens with heterogeneous states,[518, 519] SDSL-EPR technique involves introducing a cysteine residue at specific site of a protein followed by the covalent labeling with a sulfhydryl-specific spin probe.[520] SDSL-EPR technique is often complementary and sometimes superior to x-ray crystallography for exploring dynamic or transient states of macromolecules that can often be difficult to crystallize.[521]

In general, EPR benefits from larger electron spin magnetic moments compared to those of the nuclear spin in nuclear magnetic resonance (NMR) methods.[522] The EPR line shape may be simulated with fewer numbers of data sets such as free rotation in space, rotational diffusion or translation, and fewer numbers of nuclear coupling constants. NMR data sets include several variables, for example the Nuclear Overhauser Effect (NOE), *J*

couplings in multiple levels, chemical shifts, and anisotropy on rotational diffusion, causing complications in the analysis.[523]

5.10.1. Spin-labeled lysine incorporation using a PylT/PylRS pair

Recently, site-specific incorporation of a spin-labeled amino acid probe into protein via UAA mutagenesis has been reported in an *in vitro* protein expression.[524] However, incorporation in live cells may be challenging due to potential reduction of the *N*-oxide in a cellular environment. Since, TEMPO is a well-studied biophysical probe, TEMPO-labeled lysine **200** is expected to be a suitable EPR probe.[524] Here, we want to introduce spin labeled TEMPO lysine **200** and its precursor *N*-hydroxyl TEMP-lysine **199** into proteins, which could be oxidized by using either oxygen (air), an oxidizing agent such as hydrogen peroxide, or a radical initiator (Scheme 26).[525]

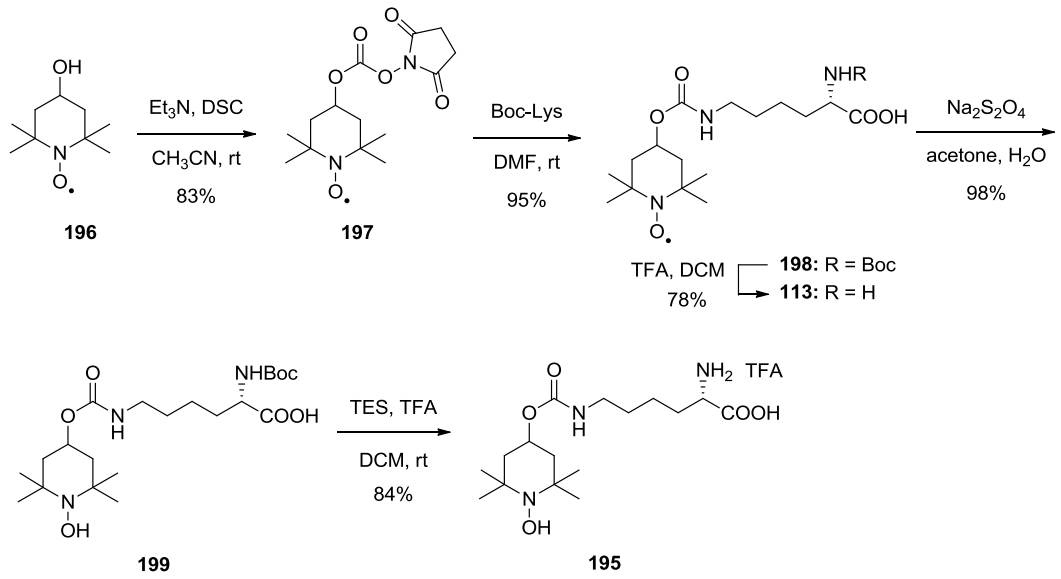


195 **113**

Scheme 26. Oxidation of TEMP(OH) lysine **195** into a spin TEMPO lysine **113**.

We synthesized the TEMP(OH) lysine **195** in four steps, starting from the commercially available TEMPO alcohol **196** (Scheme 27). Activation of the TEMPO alcohol

196 with disuccinimidyl carbonate (DSC) in the presence of triethylamine resulted in **197** in 83% yield. The TEMPO derivative **197**, upon reacting with Boc-lysine in DMF, furnished the Boc-protected TEMPO lysine **198** in 90% yield. Boc-protected **198** was then treated with a TFA:DCM (1:1) mixture at room temperature to generate the TFA salt **113** in 78% yield. In order to achieve the synthesis of **195**, nitroxide radical reduction followed by deprotection steps was performed. Reduction of the Boc-protected TEMPO lysine **198** with $Na_2S_2O_4$ in acetone:water (1:1) delivered the TEMP(OH) lysine **199** as a crystalline white solid in 95% yield.[525] Finally, Boc deprotection of **199** using a mixture of TFA:DCM (1:1) at room temperature furnished the desired amino acid **195** in 70% yield.



Scheme 27. Synthesis of TEMPO lysine **113** and TEMP(OH) lysine **195**.

The spin-labeled lysine **113** and its precursor **195** were then tested for their site-specific incorporation into proteins. In experiments performed by Ji Luo in the Deiters lab, the evolved PylRS EV3 used for fluorescent lysines **160** was investigated to incorporate **195** and **113** into sfGFP in *E. coli*. The SDS-PAGE depicted in Figure 51 demonstrates the site-specific incorporation of **113** and **195** into the protein sfGFP. Further, LRMS-ESI analysis of the sfGFP-**195** indicates a species with a mass of 28361.61Da, which is about 65 Da less than that of the calculated mass (28425.54 Da). At this point, the cause of the loss of 65 Da is not known. Further studies are in progress.

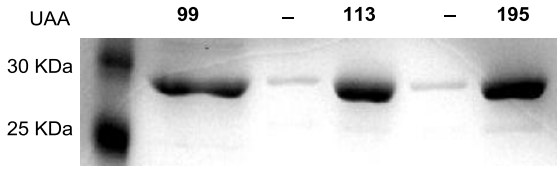


Figure 51. SDS-PAGE indicating site-specific incorporation of TEMPO lysine **113** and TEMP(OH) lysine **195** into sfGFP. Positive (**99**) and negative (– UAA) controls were included as well.

5.11. Experimental data for synthesized compounds

2-(3-Methyl-3*H*-diazirin-3-yl)-ethanol (115). 4-Hydroxy-2-butanone (**114,** 10 g, 0.11 mol) was slowly added to liquid NH_3 (60 mL) at –78 °C. Stirring was continued for 5 h at the same temperature. Next, a solution of hydroxylamine *O*-sulfonic acid (14.1 g, 0.120 mmol) in methanol (100 mL) was slowly added to the reaction mixture at –78 °C. The reaction mixture was then allowed to warm to room temperature and stirring was continued for an additional 12 h. The resulting white precipitate was filtered off and was discarded. The volume of the filtrate was reduced to ~100 mL and methanol (100 mL) and Et_3N (15 mL) were added to the mixture, followed by cooling the solution to 0 ºC in an ice bath. Iodine (~14 g) was added slowly at 0 ºC until the yellow color was persistent. The reaction mixture was allowed to warm to room temperature and stirring was continued for 2 h. The volume of the reaction mixture was reduced to 100 mL under reduced pressure, brine (200 mL) was added, and the aqueous layer was extracted with diethyl ether (3 × 20 mL). The combined organic layers were dried over anhydrous Na_2SO_4, filtered, and the volume of the filtrate was reduced to 10 mL under reduced pressure. The crude product was subjected to vacuum distillation (65 ºC, 9.0 mmHg) to obtain **115** (4.5 g, 39% yield) as a yellow oil. The analytical data matched the data reported in the literature.[526]

2,5-Dioxopyrrolidin-1-yl (2-(3-methyl-3H-diazirin-3-yl)ethyl) carbonate (116). *N*,*N*-disuccinimidyl carbonate (2.95 g, 11.5 mmol) and Et_3N (4 mL) were added to a solution of the alcohol **115** (770 mg, 7.70 mmol) in dry acetonitrile (20 mL) stirred at room temperature under an argon atmosphere. Stirring was continued for 16 h at room temperature. The solvent

was removed under reduced pressure and the crude product was purified by column chromatography on silica gel, using acetone:chloroform (20:1), to obtain **116** (2.5 g, 83%) as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ = 4.17 (t, *J* = 6.4, 2H), 3.25 (s, 4H), 1.66 (t, *J* = 6.4, 2H), 0.99 (s, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ = 168.9, 151.4, 66.5, 33.8, 25.6, 19.9. LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_9H_{11}N_3O_5Na$: 264.05, found 264.10.

(*R*)-2-((*tert*-Butoxycarbonyl)amino)-6-(((2-(3-methyl-3H-diazirin-3-yl)ethoxy)carbonyl)amino)hexanoic acid (117). Boc-lysine (2.70 g, 11.2 mmol) was added to a solution of the carbonate **116** (1.8 g, 7.4 mmol) in dry DMF (25 mL) at room temperature, under an argon atmosphere. Stirring was continued for 30 h. The reaction mixture was poured into water (100 mL) and stirring was continued for 10 minutes. The aqueous layer was extracted with diethyl ether (3 × 20 mL) and the combined organic layers were washed with brine (2 × 15 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated under reduced pressure to obtain **117** (2.4 g, 88%) as a colorless, viscous oil. ^{1}H NMR (300 MHz, $CDCl_3$) δ = 10.30 (s, br, 1H), 6.31 (s, 1H), 5.21 (s, 1H), 4.89 (s, br, 1H), 4.24 (br, 1H), 3.99-3.92 (m, 2H), 3.11 (m, br 2H), 1.77-1.37 (m, 17H), 0.98 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ = 176.6, 156.7, 156.0, 80.2, 60.1, 53.3, 40.7, 34.3, 32.2, 29.4, 28.5, 22.5, 19.9. LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{16}H_{28}N_4O_6Na$: 395.19, found 395.19.

(*R*)-2-Amino-6-(((2-(3-methyl-3H-diazirin-3-yl)ethoxy)carbonyl)amino)hexanoic acid TFA salt, (104). Boc-aminoacid **117** (600 mg, 1.60 mmol) and triethylsilane (513 µL, 3.20 mmol) were added to a solution of 5% TFA in dichloroethane (25 mL), stirred at room

temperature under an argon atmosphere. Stirring was continued for 12 h. The solvent was removed under reduced pressure and the remaining residue was dissolved in MeOH (1.5 mL). The solution was added dropwise to diethyl ether (100 mL) under vigorous stirring to precipitate the crude product. The crdue product was redissolved in MeOH (1.5 mL) and the precipitation process was performed twice to obtain **104** (472 mg, 76%) as a white crystalline solid. ^{1}H NMR (300 MHz, D_2O) δ = 4.02-3.98 (t, *J* = 5.7, 2H), 3.74-3.70 (t, *J* = 6.0, 1H), 3.14-3.10 (t, *J* = 6.0, 2H), 1.75-1.70 (m, 2H), 1.53-1.50 (t, *J* = 5.7, 2H), 1.42-1.32 (m, 2H), 1.29-1.24 (m, 2H), 1.03 (s, 3H). ^{13}C NMR (75 MHz, D_2O) δ = 60.6, 54.6, 40.0, 33.4, 30.1, 28.7, 26.5, 25.1, 23.3, 21.7, 18.8. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{11}H_{21}N_4O_4$: 273.30, found 273.15.

1-(6-Nitrobenzo[d][1,3]dioxol-5-yl)ethyl carbamate (**124**). Diphosgene (1.20 mL, 10.4 mmol) and K_2CO_3 (2.50 g, 18.8 mmol) were added to a solution of the alcohol **54** (2.0 g, 9.4 mmol) in dry THF (30 mL) stirred at 0 °C under an inert atmosphere. Stirring was continued for 12 h and the reaction mixture was allowed to warm to room temperature. The reaction mixture was concentrated under reduced pressure, DCM (50 mL) was added, and the organic layer was and washed with brine (2 × 20 mL). The organic layer was dried over anhydrous Na_2SO_4 and filtered. The filtrate was concentrated under reduced pressure and the residue was dissolved in THF (20 mL). The solution was cooled to 0 °C in an ice bath followed by addition of a concentrated aqueous solution of ammonium hydroxide (20 mL, 14.8 M) under vigorous stirring to precipitate the crude product. Stirring was continued for 2 h at the same temperature to induce complete precipitation. The volatiles were then removed under reduced

pressure to minimize the loss of product because of its partial solubility in THF and the precipitate was collected by filtration. The precipitate was washed with water (5 × 30 mL) and dried under high vacuum to furnish the carbamate **124** (1.7 g, yield 89%) as a yellow solid; mp 189-190 °C. ^{1}H NMR (400 MHz, DMSO- d_6): δ = 7.57 (s, 1 H), 7.08 (s, 1H), 6.75 (br, 1H), 6.51 (br, 1H), 6.23-5.91 (m, 1H), 1.46 (d, J = 3.4 Hz, 3H). ^{13}C NMR (100 MHz, DMSO- d_6): δ = 155.6, 152.2, 146.9, 141.1, 135.8, 105.4, 104.6, 103.5, 66.7, 21.7. Due to the instability of **124**, no mass spec data was obtained.

1-(6-Nitrobenzo[d][1,3]dioxol-5-yl)ethyl (hydroxymethyl)carbamate (**125**).

Paraformaldehyde (130 mg, 4.32 mmol) and K_2CO_3 (67 mg, 0.48 mmol) were added to a solution of the carbamate **124** (1.00 g, 3.93 mmol) in dry DMSO (10 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 2 h. EtOAc (40 mL) and brine (40 mL) were added to the reaction mixture. The organic layer was separated and the aqueous layer was extracted with EtOAc (2 × 20 mL). The organic layers were combined, washed with brine (2 × 20 mL), dried over anhydrous Na_2SO_4, and filtered The filtrate was concentrated and the remaining crude product was purified by column chromatography on silica gel using acetone:DCM (1:7) to furnish **125** (780 mg, yield 70%) as a white solid. ^{1}HNMR (300 MHz, DMSO-d_6): δ = 7.98 (t, J = 6.3 Hz, 1H), 7.57 (s, 1H), 7.09 (s, 1H), 6.22 (s, 2H), 6.02-5.54 (m, 1H), 5.57 (t, J = 6.6 Hz, 1H), 4.35 (t, J = 6.6 Hz, 1H), 1.50 (s, J = 3.1 Hz, 3H). ^{13}C NMR (75 MHz, DMSO-d_6): δ = 155.0, 152.2, 147.0, 141.2, 135.4, 105.5, 104.6, 103.5, 67.1, 64.3, 21.7. LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{11}H_{12}N_2O_7Na$: 307.05; found 307.00.

Methyl 2-(((benzyloxy)carbonyl)amino)-4-(((((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)carbonyl)amino)methyl)thio)butanoate (**128**). A solution of NaOMe (200 µL, 5.4 M in MeOH) was added to a suspension of Cbz-protected thiolactone carbamate in MeOH (1 mL), stirred at room temperature under an inert atmosphere. Stirring was continued for 30 minutes, the reaction mixture was neutralized with aqueous citric acid (5%), and the aqueous layer was extracted with EtOAc (2 × 10 mL). The combined organic layers were washed with a saturated aqueous solution of $NaHCO_3$ (2 × 10 mL), and brine (10 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated under reduced pressure to obtain the Cbz-protected homocysteine methyl ester **126** (212 mg, yield 94 %) as a colorless oil. The crude product **126** was used in the next step without further purification. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.37-7.25 (m, 5H), 5.28 (br, 1H), 5.11 (s, 2H), 4.55-4.54 (m, 1H), 3,76 (s, 3H), 2.60-2.53 (m, 2H), 2.13-1.95 (m, 2H), 1.5 (s, 1H). The Cbz-protected homocysteine methyl ester **126** (44 mg, 0.15 mmol) and a catalytic amount (10 mol%) of TsOH were added to a solution of the caged alcohol **125** (37 mg, 0.13 mmol) in dry dioxane (2 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 10 h. EtOAc (10 mL) was added to the reaction mixture, the organic layer was washed with a saturated aqueous solution of $NaHCO_3$ (2 × 5 mL) and brine (5 mL), dried over Na_2SO_4, and filtered. The filtrate was concentrated under reduced pressure and the crude product was purified by column chromatography on silica gel using Et_2O:DCM (1:4) to furnish **128** (48 mg, yield 67%) as an oil. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.46-7.44 (m, 1H), 7.36-7.26 (m, 4H), 6.99 (s, 1H), 6.38-6.25 (m, 1H), 6.10-6.08 (m, 2H), 5.52 (s, 1H), 5.37 (br, 1H), 5.12-5.10 (m, 2H), 4.40-4.21 (m, 3H), 3.72 (s, 3H), 2.62-2.55 (m, 2H), 2.12-1.99 (m, 2H), 1.54-

1.51 (m, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 172.8, 152.8, 147.6, 141.9, 136.3, 128.6, 106.2, 105.7, 103.5, 70.0, 67.5, 53.3, 53.07, 43.2, 33.0, 27.2, 22.5. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{24}H_{28}N_3O_{10}S$: 550.14; found 550.00.

Methyl 2-((*tert*-butoxycarbonyl)amino)-4-(((((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)carbonyl)amino)methyl)thio)butanoate (**129**). Boc-homocysteine methyl ester **127** (270 mg, 1.03 mmol) and TsOH (16 mg, 0.08 mmol) were added to a solution of the alcohol **125** (1.00 g, 3.93 mmol) in dry dioxane (5 mL), stirred at room temperature under an inert atmosphere. Stirring was continued for another 14 h, EtOAc (20 mL) was added to the reaction mixture, and the organic layer was washed with a saturated aqueous solution of $NaHCO_3$ (2 × 20 mL), and brine (2 × 20 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated under reduced pressure and the crude product was purified by column chromatography on silica gel using Et_2O:DCM (1:4) to yield **129** (316 mg, yield 71%) as an oil. ^{1}HNMR (300 MHz, $CDCl_3$): δ = 7.48 (s, 1 H), 6.99 (s, 1H), 6.31-6.25 (q, 1H), 6.11 (s, 2H), 5.46 (br, 1H), 5.19 (br, 1H), 4.38-4.16 (m, 4H), 3.75 (s, 3H), 2.73-2.54 (m, 2H), 2.07-1.88 (m, 2H), 1.58 (d, *J* = 3.3 Hz, 3H), 1.44 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 172.9, 155.2, 152.6, 147.3, 141.6, 136.3, 105.5, 105.3, 105.2, 80.4, 69.6, 65.9, 52.7, 43.3, 34.6, 32.7, 28.5, 27.1, 22.3. LRMS-ESI (*m/z*) $[M+Na]^+$ m/z calcd for $C_{21}H_{29}N_3O_{10}SNa$: 538.14; found 538.10.

2-((*tert*-Butoxycarbonyl)amino)-4-(((((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)carbonyl)amino)methyl)thio)butanoic acid (**130**). An aqueous solution of LiOH (2 mL, 2 M) was added to a solution of methyl ester **129** (300 mg, 0.580 mol) in dry dioxane (2 mL) stirred at 0 °C under an inert atmosphere. Stirring was continued for 30 minutes. Water (20 mL) was added to the reaction mixture, the aqueous layer was washed with Et_2O (10 mL), and the organic layer was discarded. Next, the aqueous layer was acidified to pH 3.0 with an aqueous solution of citric acid (5%). The product remaining in the aqueous layer was then extracted with EtOAc (2 × 20 mL). The combined organic layers were washed with brine (2 × 20 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated to dryness under reduced pressure to obtain **130** (275 mg, yield 94%) as a yellow foam. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.48 (s, 1 H), 7.00 (s, 1H), 6.27-6.26 (m, 1H), 6.10 (s, 2H), 5.57 (s, 1H), 5.46-5.36 (m, 1H), 4.36-4.21 (m, 3H), 2.78-2.65 (br, 2H), 2.11-1.96 (br, 2H), 1.57 (d, *J* = 3.2 Hz, 3H), 1.4 (s, 9H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 176.0, 152.6, 147.3, 141.6, 136.2, 105.8, 105.4, 103.3, 70.7, 69.7, 65.9, 52.6, 43.4, 32.5, 28.5, 27.3, 22.3. LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{20}H_{27}N_3O_{10}SNa$: 524.13; found 524.10.

2-Amino-4-(((((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)carbonyl)amino)methyl)thio)butanoic acid (**106**). Triethylsilane (644 µL, 2.34 mmol) and caged Boc-homocysteine **130** (1.0 g, 1.9 mmol) were added to TFA:DCM (1:1, 18 mL) and stirred at room temperature for 45 min under an inert atmosphere. The reaction mixture was concentrated under reduced pressure and the residue was dissolved in MeOH (10 mL). To remove the residual amount of TFA completely from the reaction mixture, the

mixture was concentrated again under reduced pressure and the residue was redissolved in MeOH. The process was subsequently repeated three times. Finally, the solid was dissolved in MeOH (2 mL) and added drop wise to Et_2O (200 mL) under vigorous stirring to precipitate the crude product. The process of dissolving the product followed by precipitation was repeated twice to obtain the caged homocysteine TFA salt **106** (800 mg, yield 78%) as a crystalline white solid. ^{1}HNMR (400 MHz, CD_3OD): δ = 7.48 (d, *J* = 2.8 Hz, 1H), 7.11-7.09 (m, 1H), 6.20-6.13 (m, 3H), 4.25-4.14 (m, 2H), 3.86-3.77 (m, 2H), 2.87 (br, 1H), 2.69-1.67 (br, 2H), 2.22 (br, 1H), 2.07 (br, 1H), 1.63-1.59 (m, 3H). ^{13}C NMR (100 MHz, CD_3OD): δ = 172.9, 157.7, 154.1, 148.9, 143.0, 137.1, 106.6, 105.9, 105.0, 70.3, 53.9, 43.4, 32.1, 27.2, 22.4. LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{15}H_{20}N_3O_8S$: 402.09; found 402.09.

Methyl 2-((*tert*-butoxycarbonyl)amino)-4-(((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)carbonyl)thio)butanoate (**132**). Triethylamine (1.90 mL, 14.1 mmol) and disuccinimidyl carbonate (1.80 g, 7.10 mmol) were added to the solution of the alcohol **54** (1.00 g, 4.73 mmol) stirred at room temperature under an inert atmosphere. Stirring was continued for 12 h. Next, the solvent was removed under reduced pressure and the crude product was directely purified, by column chromatography on silica gel using $CHCl_3$:CH_3COCH_3 (12:1), to furnish succinimidyl carbonate **131** (1.6 g, yield 93%) as a white solid. ^{1}HNMR (300 MHz, $CDCl_3$): δ = 7.60 (s, 1H), 7.20 (s, 1H), 6.50 (t, *J* = 3.3 Hz , 1H), 6.24 (s, 2H), 2.89 (s, 4H), 1.83 (d, *J* = 3.3 Hz, 3H). Boc-homocysteine methyl ester **127** (421 mg, 1.69 mmol) and DIPEA (800 µL, 4.59 mmol) were added to a solution of the succinimidyl carbonate **131** (500 mg, 1.53 mmol) in dry DMF (5 mL) stirred at room

temperature under an inert atmosphere. Stirring was continued for 14 h. The reaction mixture was then concentrated under reduced pressure and EtOAc (20 mL) was added to dissolve the remaining crude product. The organic layer was washed with a saturated aqueous solution of $NaHCO_3$ (2 × 20 mL), brine (2 × 20 mL), dried over anhydrous Na_2SO_4 and filtered. The filtrate was concentrated under reduced pressure and the remaining crude product was purified by column chromatography on silica gel using EtOAc:hexanes (4:6) to furnish **132** (568 mg, yield 76%) as a yellow solid. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.48 (s, 1 H), 7.00 (s, 1H), 6.27-6.26 (m, 1H), 6.12 (s, 2H), 5.08 (d, *J* = 3.8 Hz, 1H), 4.34 (br, 1H), 3.75-3.72 (m, 3H), 2.19-2.73 (m, 2H), 2.15-2.12 (m, 1H), 1.96-1.87 (m, 1H) 1.62 (d, *J* = 3.8 Hz, 3H), 1.42 (s, 9H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 172.6, 170.0, 155.5, 152.8, 147.6, 141.5, 135.1, 105.9, 105.5, 103.3, 80.4, 71.6, 52.7, 33.3, 28.4, 27.2, 22.3. LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{20}H_{26}N_2O_{10}SNa$: 509.12; found 509.12.

2-((*tert*-Butoxycarbonyl)amino)-4-(((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)carbonyl)thio)butanoic acid (**133**). An aqueous solution of LiOH (4 mL, 2 M) was added to a solution of the methyl ester **132** (500 mg, 1.02 mol) in dioxane (4 mL) stirred at 0 °C, and stirring was continued for 1 h. Water (30 mL) was added and the aqueous layer was washed with Et_2O (2 × 20 mL) to remove any unreacted starting material. The aqueous layer was acidified to pH 3.0 using an aqueous solution of citric acid (5%) and extracted with EtOAc (2 × 20 mL). The combined organic layers were washed with brine (2 × 20 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated to dryness under reduced pressure to obtain **133** (275 mg, yield 94%) as an oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.28-

2-26 (m, 2H), 6.10 (s, 1H), 6.27-6.26 (m, 1H), 5.01 (br, 1H), 4.75-4.73 (m, 1H), 4.33 (br, 1H), 2.49-2.36 (m, 2H), 2.04-1.80 (m, 2H), 1.53 (d, *J* = 3.4 Hz, 3H), 1.43 (s, 9H).

2-Amino-4-(((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)carbonyl)thio)butanoic acid (**118**). Triethylsilane (155 µL, 0.970 mmol) and caged Boc-homocysteine **133** (230 mg, 0.480 mmol) were added to a solution of TFA:DCM (1:1, 3 mL) and stirred at room temperature for 45 minutes. The reaction mixture was concentrated under reduced pressure. The obtained residue was redissolved in MeOH (2 mL) followed by concentration under reduced pressure. The process of MeOH addition followed by concentration was repeated twice to remove any residual amount of TFA. Finally, MeOH (500 µL) was added to the crude product and the solution was added dropwise into Et_2O (50 mL) under vigorous stirring to precipitate the crude product. The crude product was redissolved in MeOH (1.5 mL) and the process of precipitation was repeated once more to obtain the caged homocysteine TFA salt **118** (220 mg, yield 93%) as a white crystalline solid. ^{1}H NMR (400 MHz, CD_3OD): δ = 7.51 (s, 1H), 7.08 (s, 1H), 6.43-6.41 (m, 2H), 6.16 (s, 2H), 4.00-3.95 (m,1H), 3.07-2.90 (m, 2H), 2.25-2.12 (m, 2H), 1.66 (d, *J* = 3.2 Hz, 3H). ^{13}C NMR (100 MHz, CD_3OD): δ = 172.1, 154.8, 149.9, 143.9, 135.8, 107.4, 106.6, 105.7, 73.6, 33.1, 28.9, 22.7. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{14}H_{17}N_2O_8S$: 372.06; found 373.07.

5-(1-Bromoethyl)-6-nitrobenzo[d][1,3]dioxole (**134**).A solution of PBr_3 (1.4 mL, 1.0 M in THF) was added to a solution of the alcohol **54** (200 mg, 0.94 mmol) in dry DCM (2 mL) stirred at 0 °C under an inert atmosphere. The reaction mixture was warmed to room

temperature and the stirring was continued for 12 h. DCM (10 mL) was added to the reaction mixture and the organic layer was washed with a saturated aqueous solution of $NaHCO_3$ (3 × 10 mL), brine (10 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated under reduced pressure and under high vacuum to obtain the bromide **134** (205 mg, yield 80%) as a brown solid, which was directly used for the next step without further purification. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.36-7.24 (m, 2H), 6.14 (s, 2H), 5.91-5.88 (m, 1H), 2.04 (s, 2H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 108.8, 105.1, 103.4, 42.9, 27.7. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_9H_9BrNO_4$: 273.97; found 274.27.

Methyl 2-((*tert*-butoxycarbonyl)amino)-4-((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethyl)thio)butanoate (**135**). Boc-homocysteine methyl ester **127** (231 mg, 0.870 mmol) and K_2CO_3 (302 mg, 2.19 mmol) were added to a solution of nitropiperonyl bromide **134** (200 mg, 0.730 mmol) in dry DMF (1 mL) stirred at room temperature under an inert atmosphere. The reaction mixture was contineously stirred for 12 h. Then, EtOAc (20 mL) was added to the reaction mixture and the organic layer washed with a saturated aqueous solution of $NaHCO_3$ (2 × 20 mL), brine (10 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated and the remaining residue was purified by column chromatography, on silica gel using EtOAc:hexanes, (1:4) to yield **135** (323 mg, yield 66%) as an oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.28-7.23 (m, 2H), 6.11-6.09 (m, 2H), 4.99 (br, 1H), 4.73-4.69 (m, 1H), 4.30 (br, 1H), 3.71 (s, 3H), 2.38-2.33 (m, 2H), 2.04-1.98 (m, 1H), 1.81-1.72 (m,1H), 1.55-1.43 (m, 3H), 1.42 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 172.8,

155.4, 152.1, 146.9, 143.4, 136.3, 108.2, 104.8, 103.1, 52.6, 39.0, 38.5, 32.8, 28.4, 27.9, 27.6, 23.2.

2-((*tert*-Butoxycarbonyl)amino)-4-((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethyl)thio)butanoic acid (**136**). An aqueous solution of LiOH (2 mL, 2 M) was added to a solution of the methyl ester **135** (202 mg, 0.45 mmol) in dioxane (2 mL) stirred at 0 °C. Stirring was continued for 1 h, the reaction mixture was diluted with water (30 mL) and the aqueous layer was washed with Et_2O (2 × 20 mL). The organic layer was discarded and the aqueous layer was acidified to pH 3.0 with aqueous solution of citric acid (5%). The product remaining in the aqueous layer was extracted using EtOAc (2 × 10 mL). The combined organic layers were washed with brine (2 × 20 mL), dried over anhydrous Na_2SO_4 and filtered. The filtrate was concentrated under reduced pressure to obtain **136** (181 mg, yield 92%) as a foam. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.27-7.24 (m, 2H), 6.10-6.07 (m, 2H), 5.01 (br, 1H), 4.77-4.71 (m, 1H), 4.32 (br, 1H), 2.48-2.37 (m, 2H), 2.10-1.78 (m, 2H), 1.54-1.45 (m, 3H), 1.43 (s, 9H). LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{18}H_{24}N_2O_8$ SNa: 451.11; found 451.11.

2-Amino-4-((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethyl)thio)butanoic acid (**119**). Triethylsilane (128 μL, 0.790 mmol) and caged Boc-homocysteine **136** (171 mg, 0.390 mmol) were added to a solution of TFA:DCM (1:1, 2 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 45 minutes. The reaction mixture was concentrated under reduced pressure, redissolved in MeOH (2 mL), and reconcentrated. The

process of co-evaporation was repeated twice to remove residual amount of TFA. Finally, the crude product was dissolved in MeOH (500 µL) and precipitated in Et_2O (50 mL) under vigorous stirring. The precipitate was separated by filtration, redissolved in MeOH, and reprecipitated in Et_2O. The process of dissolution followed by precipitation was repeated once more to obtain the caged homocysteine TFA salt **119** (172 mg, yield 70%) as a crystalline white solid; mp 82-85 °C. ^{1}H NMR (300 MHz, CD_3OD): δ = 7.34-7.30 (m, 2H), 6.12 (s, 2H), 4.77-4.72 (m, 1H), 3.80-3.75 (m, 1H), 2.57-2.48 (m, 2H), 2.09-1.94 (m, 2H), 1.54 (d, *J* = 4.0 Hz, 3H). ^{13}C NMR (75 MHz, CD_3OD): δ = 173.2, 153.9, 148.9, 145.1, 137.0, 109.1, 106.1, 105.9, 105.1, 54.5, 40.23, 32.3, 28.6, 28.5, 23.5. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{13}H_{17}N_2O_6S$: 329.08; found 329.08.

(2*R*)-Methyl 2-((*tert*-butoxycarbonyl)amino)-3-(((((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)carbonyl)amino)methyl)thio)propanoate (**139**). Boc-cysteine methyl ester **137** (794 mg, 3.37 mmol) and TsOH (54 mg, 0.28 mmol) were added to a solution of the alcohol **125** (800 mg, 2.81 mmol) in dry dioxane (10 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 14 h, the reaction mixture was diluted with EtOAc (30 mL), the organic layer was washed with a saturated aqueous solution of $NaHCO_3$ (2 × 30 mL) and brine (2 × 20 mL), and then dried over anhydrous Na_2SO_4 and filtered. The filtrate was concentrated under reduced pressure and the remaining crude product was purified by column chromatography on silica gel using Et_2O:DCM (1:9) to furnish **139** (900 mg, yield 64%) as an oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.47-7.46 (m, 1H), 7.02-6.96 (m, 1H), 6.29-6.22 (m, 1H), 6.10 (s, 2H), 5.76 (br, 1H), 5.37 (br, 1H), 4.53-4.20 (m, 3H), 3.76 (s, 3H), 3.17-

2.86 (m, 2H), 1.57 (d, *J* = 3.1 Hz, 3H), 1.45 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 171.7, 155.7, 155.3, 152.6, 147.3, 141.6, 136.0, 105.9, 105.4, 103.2, 80.7, 69.6, 53.7, 52.9, 48.2, 44.4, 41.4, 36.6, 34.3, 33.2, 28.5, 22.3.

(**2*R*)-*tert*-Butyl 2-((*tert*-butoxycarbonyl)amino)-3-(((((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)carbonyl)amino)methyl)thio)propanoate** (**140**). Boc-cysteine *t*-butyl ester **138** (421 mg, 1.52 mmol) and TsOH (24 mg, 0.12 mmol) were added to a solution of the alcohol **125** (360 mg, 1.26 mmol) in dry dioxane (5 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 12 h. The reaction mixture was diluted with EtOAc (20 mL); the organic layer was washed with a saturated aqueous solution of $NaHCO_3$ (2 × 20 mL), brine (20 mL), dried over anhydrous Na_2SO_4 and filtered. The filtrate was concentrated under reduced pressure and the obtained crude product was purified by column chromatography on silica gel using Et_2O:DCM (1:9) to deliver **140** (505 mg, yield 73%) as an oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.47 (s, 1H), 7.02 (s, 1H), 6.28 (br, 1H), 6.09 (s, 2H), 5.79 (br, 1H), 5.35 (br, 1H), 4.36-4.17 (m, 3H), 3.04-3.00 (m, 1H), 2.92-2.89 (m, 1H), 1.58 (br, 3H), 1.56 (s, 18H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 170.1, 154.9, 152.6, 147.2, 141.5, 136.4, 105.8, 105.4, 103.2, 83.1, 80.5, 69.5, 54.4, 44.6, 34.7, 28.5, 28.1, 22.4. LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{23}H_{33}N_3O_{10}SNa$: 566.1788; found 566.1784.

(2*R*)-2-Amino-3-(((((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)carbonyl)amino)methyl)thio)propanoic acid (**107**). Triethylsilane (311 µL, 1.90 mmol) and the caged Boc-cysteine **140** (525 mg, 0.960 mmol) were added to a solution of

TFA:DCM (1:1, 10 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 45 minutes. The, the reaction mixture was concentrated under reduced pressure and the obtained residue was redissolved in MeOH (5 mL); and the process dissolution and concentration was repeated twice to remove residual amounts of TFA. Finally, the crude product was dissolved in MeOH (1 mL) and and the solution was added dropwise to Et_2O (100 mL) under vigorous stirring to precipitate the crude product. The precipitate was redissolved in MeOH (1 mL) and the process of precipitation was repeated once more to obtain the caged homocysteine TFA salt **107** (370 mg, yield 77%) as a crystalline hygroscopic white solid. ^{1}H NMR (300 MHz, CD_3OD): δ = 7.48 (s, 1H), 7.10 (s, 1H), 6.23-6.16 (m, 1H), 6.14 (s, 2H), 4.30-4.20 (m, 2H), 3.89-3.85 (m, 1H), 3.24-3.15 (m, 1H), 2.99-2.91 (m, 1H), 1.57 (d, *J* = 3.1 Hz, 3H). ^{13}C NMR (75 MHz, CD_3OD): δ = 172.4, 157.6, 154.2, 149.0, 143.1, 137.2, 106.7, 106.0, 105.1, 70.5, 54.9, 44.1, 32.7, 28.3, 22.5. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{14}H_{18,}N_3O_8S$: 379.14; found 379.15.

2,5-Dioxopyrrolidin-1-yl (2-(2-nitrophenyl)propyl) carbonate (**144**). Triethyl amine (2.30 mL, 16.5 mmol) and disuccinimidyl carbonate (2.10 g, 8.28 mmol) were added to a solution of the alcohol **143** (1.00 g, 5.52 mmol) in dry acetonitrile (20 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for another 12 h. Then, the reaction mixture was concentrated under reduced pressure and the obtained crude product was directly purified by silica gel column chromatography using acetone:$CHCl_3$ (1:9) mixture to furnish the corresponding carbonate **144** (1.6 g, yield 94%) as a white solid; mp 150-153 °C. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.80-7.72 (m, 1H), 7.59-7.56 (m, 1H), 7.48-

7.46 (m, 1H), 7.40-7.38 (m, 1H), 4.50-4.47 (m, 2H), 3.77-3.75 (m, 1H), 2.78 (s, 4H), 1.40 (d, J = 3.6 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 168.7, 151.6, 136.0, 133.3, 128.7, 128.2, 124.7, 74.6, 33.5, 25.6, 17.72. Due to the instability of **144**, no mass spec data was obtained.

(2*S*)-2-((*tert*-Butoxycarbonyl)amino)-6-(((2-(2-nitrophenyl)propoxy)carbonyl)amino)hexanoic acid (**145**). Boc-lysine (1.40 g, 5.96 mmol) was added to a solution of the carbonate **144** (1.6 g, 4.9 mmol) in dry DMF (10 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 30 h. Then, the reaction mixture was poured into water (100 mL) and the aqueous layer was extracted using EtOAc (3 × 20 mL). The combined organic layers were washed with water (3 × 20 mL), brine (20 mL), dried over anhydrous Na_2SO_4, and filtered. The solvent was removed under reduced pressure and the remaining product was dried under a high vacuum to obtain the caged Boc-lysine **145** (1.8 g, yield 81%) as a foam. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.71-7.69 (m, 1H), 7.54-7.53 (m, 1H), 7.46-7.44 (m, 1H), 7.34 (br, 1H), 6.36 (s, 0.5H), 6.23 (s, 0.5H), 5.27 (s, 1H), 4.83 (br, 1H), 4.23-4.21 (m, 2H), 4.08-4.07 (m, 1H), 3.08-2.96 (br, 2H), 1.80-1.64 (m, 2H), 1.41-1.30 (m, 13H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 176.7, 156.6, 156.0, 150.9, 137.5, 132.8, 128.3, 128.2, 127.7, 127.5, 124.2, 80.2, 69.6, 68.9, 53.3, 40.7, 33.4, 32.1, 29.5, 28.3, 25.7, 22.5, 18.0, 17.6. LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{16}H_{23,}N_3O_6Na$: 476.20; found 476.10.

(2*S*)-2-Amino-6-(((2-(2-nitrophenyl)propoxy)carbonyl)amino)hexanoic acid (**108**). Triethylsilane (46 µL, 0.28 mmol) and the caged Boc-lysine **145** (64 mg, 0.14 mmol) were added to a solution of TFA:DCM (1:1, 2 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 45 minutes. Then the reaction mixture was concentrated under reduced pressure and the obtained residue was redissolved in MeOH (200 µL). The process of dissolution of the solid followed by removal of solvents was repeated twice to remove residual amount of TFA. Finally, the crude product was dissolved in MeOH (200 µL) and the solution was added dropwise to Et_2O (10 mL) under vigorous stirring to precipitate the crude product. The precipitate was collected through filtration and redissolved in MeOH (200 µL). Then the processes of precipitation was performed once more to obtain the caged lysine TFA salt **108** (370 mg, yield 77%) as a white solid. ^{1}H NMR (400 MHz, CD_3OD): δ = 7.75-7.73 (m, 1H), 7.62-7.60 (m, 2H), 7.43-7.40 (m, 1H), 4.25-4.15 (m, 2H), 3.57-3.54 (m, 2H), 3.05 (br, 2H), 1.88-1.76 (m, 2H), 1.48-1.32 (m, 7H). ^{13}C NMR (100 MHz, CD_3OD): δ = 174.2, 158.8, 152.1, 138.5, 133.8, 129.6, 128.7, 125.0, 69.5, 55.9, 41.38, 35.0, 31.9, 30.5, 23.5, 18.3. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{16}H_{24,}N_3O_6$: 354.15; found 354.16.

4-(Hydroxymethyl)-7-(methoxymethoxy)-2H-chromen-2-one (**150**). Diisopropylethyl amine (1.80 mL, 10.4 mmol) and MOMCl (175 µL, 2.29 mmol) were added to a solution of the 7-hydroxy coumarin **149** (400 mg, 2.08 mmol) in dry DCM (5 mL) stirred at 0 °C under an inert atmosphere. The reaction mixture was allowed to warm to room temperature and stirring was continued for 12 h. Next, the reaction mixture was diluted with DCM (20 mL),

the organic layer was washed with a saturated aqueous solution of $NaHCO_3$ (3 × 20 mL) and brine (20 mL), and then dried over anhydrous Na_2SO_4. The solvent was removed under reduced pressure to obtain MOM-protected coumarin **150** (491 mg, yield 79%) as a white solid; mp 147-150 °C. ^{1}H NMR (300 MHz, CD_3OD): δ = 7.54 (d, *J* = 4.6 Hz, 1H), 6.97 (s, 1H), 6.38 (s, 1H), 5.23 (s, 2H), 4.78 (s, 2H), 3.42 (s, 3H). ^{13}C NMR (75 MHz, CD_3OD): δ = 164.3, 162.4, 158.7, 156.9, 126.7, 115.4, 113.8, 109.8, 105.1, 96.2, 61.4, 57.2. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{12}H_{13}O_5$: 237.0768; found 237.0760.

(7-(Methoxymethoxy)-2-oxo-2H-chromen-4-yl)methyl (4-nitrophenyl) carbonate (**151**). Diisopropylethyl amine (936 µL, 5.35 mmol) and *p*-nitrophenyl chloroformate (430 mg, 2.14 mmol) were added to a solution of the coumarin alcohol **150** (353 mg, 1.07 mmol) in dry DCM (5 mL) stirred at 0 °C under an inert atmosphere. The reaction mixture was warmed to room temperature and stirring was continued for 12 h. The reaction mixture was concentrated under reduced pressure and the residue was dissolved in EtOAc (20 mL). The organic layer was washed with a saturated aqueous solution of $NaHCO_3$ solution (3 × 20 mL) and brine (20 mL), and dried over anhydrous Na_2SO_4. The filtrate was concentrated and the crude product was precipitated in Et_2O:hexanes (1:1) mixture to obtain the coumarin carbonate **151** (395 mg, yield 92%) as a white solid; mp 138-140 °C. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 8.32-8.29 (m, 2H), 7.46-7.41 (m, 3H), 7.07-6.99 (m, 2H), 6.46 (s, 1H), 5.45 (s, 1H), 5.25 (s, 2H), 3.4 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 160.8, 155.3, 147.5, 125.7, 124.6, 124.5, 121.9, 113.9, 111.4, 104.5, 94.6, 65.7, 56.6. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{19}H_{16}NO_9$: 402.08; found 402.08.

(*S*)-2-Amino-6-((((7-hydroxy-2-oxo-2H-chromen-4-yl)methoxy)carbonyl)amino)hexanoic acid (**109**). Boc-lys (311 mg, 1.26 mmol) was added to a solution of the coumarin carbonate **151** (390 mg, 0.970 mmol) in dry DMF (5 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 12 h, the reaction mixture was poured into water (50 mL) and the product in the aqueous layer was extracted using EtOAc (2 × 20 mL). The combined organic layers were washed with water (3 × 20 mL) and brine (20 mL), and dried over anhydrous Na_2SO_4. The filtrate was concentrated under reduced pressure and the solid product was dried under a high vacuum. The crude product was used to the next step without further purification. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.40 (d, *J* = 4.3 Hz, 1H), 7.00-6.96 (m, 2H), 6.33 (s, 1H), 5.36-5.31 (m, 1H), 5.29-5.22 (m, 4H), 4.28 (br, 1H), 3.47 (s, 3H), 3.25-3.23 (m, 2H), 1.82-1.70 (m, 2H), 1.57-1.24 (s, 13H). The crude compound was treated with triethylsilane (350 µL, 2.17 mmol) and a solution of TFA:DCM (1:1, 10 mL) at room temperature under an inert atmosphere. Stirring was continued for 40 minutes. The reaction mixture was concentrated under reduced pressure. The obtained residue was dissolved in MeOH (2 mL) and concentrated under reduced pressure. The process of adding MeOH followed by the removal of solvents was repeated three times to remove residual amount of TFA. Finally, the crude product was dissolved in MeOH (1 mL), precipitated in Et_2O (50 mL) under vigorous stirring. The precipitate was collected, and the processes of dissolution followed by precipitation was repeated twice to furnish coumarin lysine **109** (275 mg, yield 59%) as a crystalline white solid; mp 215-218 °C. ^{1}H NMR (300 MHz, DMSO- d_6): δ = 7.56-7.48 (m, 2H), 6.82-6.76 (m, 2H), 6.10 (s, 1H), 5.44 (s, 1H), 3.57 (t, *J* = 3.0 Hz, 1H), 3.01 (br, 2H), 1.73 (br, 2H), 1.41

(br, 4H). ^{13}C NMR (75 MHz, DMSO- d_6): δ = 172.0, 162.6, 160.9, 160.7, 156.0, 155.6, 152.7, 150.3, 126.3, 113.9, 109.9, 107.6, 103.1, 61.5, 53.5, 30.8, 29.6, 22.6. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{17}H_{21}N_2O_7$: 365.13; found 365.13.

(*S*)-2-Amino-6-((((6-bromo-7-hydroxy-2-oxo-2H-chromen-4-yl)methoxy)carbonyl)amino)hexanoic acid (**147**). Boc-lysine (400 mg, 1.62 mmol) was added to a solution of the bromocoumarin carbonate **152** (600 mg, 1.25 mmol) in dry DMF (7 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 30 h. Next, the reaction mixture was poured into water (100 mL) and the product in the aqueous layer was extracted using EtOAc (3 × 20 mL). The combined organic layers were washed with water (3 × 20 mL) and brine (20 mL), dried over anhydrous Na_2SO_4 and filtered. The solvent was removed to obtain crude product as a solid (637 mg). The crude product was subjected to the next step synthesis without further purification. ^{1}H NMR (300 MHz, DMSO-d_6): δ = 7.97 (s, 1H), 7.58 (br, 1H), 7.15 (br, 1H), 6.28 (s, 1H), 5.43 (s, 2H), 5.27 (s, 2H), 3.80 (br, 1H), 3.42 (s, 3H), 3.20 (br, 2H), 1.65-1.42 (m, 2H), 1.36-1.34 (s, 13H). Triethylsilane (377 µL, 2.34 mmol) and the Bhc Boc-lysine were added to a solution of TFA:DCM (1:1, 14 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 45 minutes. The reaction mixture was concentrated under reduced pressure and the obtained residue was dissolved in MeOH (5 mL) and then re-concentrated. The process was repeated three times to remove any residual TFA. The crude product was dissolved in MeOH (1 mL) and precipitated in Et_2O (150 mL) under vigorous stirring. The precipitate was collected, redissolved in MeOH, and precipitated. The processes was repeated twice to

obtain the Bhc-lysine TFA salt **147** (450 mg, yield 64%) as a crystalline white solid. ^{1}H NMR (300 MHz, DMSO-d_6): δ = 7.82 (s, 1H), 7.03 (s, 1H), 6.14 (s, 1H), 5.22 (s, 2H), 3.57 (t, *J* = 6.0 Hz, 13H), 3.00-2.98 (m, 2H), 1.74-1.69 (br, 2H), 1.42 (br, 4H). ^{13}C NMR (75 MHz, DMSOd$_6$): δ = 171.1, 159.8, 157.9, 155.2, 151.3, 128.2, 110.1, 108.15, 106.4, 103.3, 60.8, 52.6, 30.0, 28.9, 21.9. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{17}H_{20}BrN_2O_7$: 443.04; found 443.11.

(*S*)-Methyl 2-amino-6-((((7-hydroxy-2-oxo-2H-chromen-4-yl)methoxy)carbonyl)amino)hexanoate (**153**). Boc-lysine methy ester hydrochloride (71 mg, 0.23 mmol) and TEA (138 µL, 0.79 mmol) were added to a solution of the coumarin carbonate **151** (80 mg, 0.199 mmol) in dry DMF (1.0 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 12 h, the reaction mixture was diluted with EtOAc (10 mL), the organic layer was washed with water (2 × 10 mL) and brine (10 mL), and dried over anhydrous Na_2SO_4 and filtered. The solvent was removed under reduced pressure to obtain a crude product that was subjected to the next step of synthesis without further purification. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.40 (d, *J* = 4.3 Hz, 1H), 7.00-6.95 (m, 2H), 6.32 (s, 1H), 5.36-5.28 (m, 1H), 5.24-5.21 (s, 4H), 4.28 (br, 1H), 3.72 (s, 3H), 3.46 (s, 3H), 3.22 (br, 2H), 1.82-1.53 (m, 4H), 1.48-1.34 (m, 11H). The crude product was treated with triethylsilane (30 µL, 0.18 mmol) and a soluion of TFA:DCM (1:1, 1 mL) at room temperature. Stirring was continued for 45 minutes and the reaction mixture was concentrated under reduced pressure. The residue was dissolved in MeOH (1 mL) and concentrated under reduced pressure to remove both the solvent and TFA. The process was

repeated three times to remove any residual TFA. The crude product was dissolved in MeOH (200 µL) and precipitated into Et_2O (5.0 mL) under vigorous stirring to obtain the coumarin lysine methyl ester **153** (37 mg, 67% yield) as a crystalline white solid. ^{1}H NMR (400 MHz, CD_3OD): δ = 7.52 (d, *J* = 4.4 Hz, 1H), 7.81 (d, *J* = 4.2 Hz, 1H), 6.18 (m, 1H), 5.29 (s, 2h), 4.08-4.05 (m, 1H), 3.82 (s, 3H), 3.18-3.15 (m, 2H), 1.98-1.81 (m, 2H), 1.59-1.41 (m, 4H). ^{13}C NMR (100 MHz, CD_3OD): δ = 170.9, 163.5, 163.1, 157.1, 156.7, 153.8, 126.3, 114.5, 110.8, 108.3, 103.7, 62.6, 53.8, 41.2, 31.1, 30.2, 23.1.

4-(2-Hydroxyethyl)-7-(methoxymethoxy)-2H-chromen-2-one (**155**). Diisopropylethyl amine (251 µL, 1.44 mmol) and MOMCl (46 µL, 0.58 mmol) were added to a solution of the 7-hydroxy coumarin alcohol **154** (100 mg, 0.48 mmol) in dry DCM (2 mL) stirred at 0 °C under an inert atmosphere. The reaction mixture was allowed to warm to room temperature and the stirring was continued for 12 h. The reaction mixture was diluted with DCM (10 mL), the organic layer was washed with a saturated aqueous solution of $NaHCO_3$ solution (3 × 10 mL), brine (5 mL), dried over anhydrous Na_2SO_4, and filtered. The solvent was removed under reduced pressure to furnish the MOM-protected coumarin alcohol **155** (121 mg, yield 90%) as a low melting solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.55 (d, *J* = 4.3 Hz, 1 H), 6.96 (d, *J* = 4.3 Hz, 1 H), 6.21 (s, 1H), 5.21 (s, 2H), 3.97 (t, *J* = 6.3 Hz, 1 H), 3.46 (s, 3H), 2.99 (t, *J* = 6.3 Hz, 1 H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 161.5, 160.3, 155.4, 153.5, 125.7, 113.9, 113.5, 112.6, 104.2, 94.5, 60.9, 56.5, 34.9. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{13}H_{15}O_5$: 251.09; found 251.09.

2-(7-(Methoxymethoxy)-2-oxo-2H-chromen-4-yl)ethyl (4-nitrophenyl) carbonate (**156**). Cs_2CO_3 (253 mg, 0.780 mmol), DMAP (63 mg, 0.52 mmol), and *p*-nitrophenyl chloroformate (156 mg, 0.780 mmol) were added to a solution of the coumarin alcohol **155** (353 mg, 1.07 mmol) in dry THF (5 mL) stirred at 0 °C under an inert atmosphere. The reaction mixture was allowed to warm to room temperature and the stirring was continued for 12 h. The reaction mixture was filtered to discard any solid precipitate and the filtrate was concentrated under reduced pressure. EtOAc (15 mL) was added to the residue and the organic layer was washed with a saturated aqueous solution of $NaHCO_3$ (3 × 10 mL) and brine (20 mL). The organic layer was dried over anhydrous Na_2SO_4, filtered, and the filtrate was concentrated and the obtained crude product was purified by column chromatography on silica gel using an EtOAc:hexanes (1:1) mixture to obtain the coumarin carbonate **156** (70 mg, 33% yield) as a crystalline white solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 8.28 (d, *J* = 4.6 Hz, 2H), 7.56 (d, *J* = 4.5 Hz, 1H), 7.35 (d, *J* = 4.6 Hz, 2H), 7.05-6.99 (m, 2H), 6.25 (s, 1H), 5.23 (s, 2H), 4.59 (t, *J* = 6.6 Hz, 2H), 3.48 (s, 3H), 3.21 (t, *J* = 6.6 Hz, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 160.9, 160.6, 155.5, 152.6, 150.9, 125.6, 125.2, 122.0, 113.8, 113.2, 104.5, 94.5, 66.57, 56.6, 31.0. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{20}H_{18}NO_9$: 416.0903; found 416.0903.

(*S*)-2-Amino-6-((((7-hydroxy-2-oxo-2H-chromen-4-yl)methoxy)carbonyl)amino)hexanoic acid (**148**). Boc-lysine (57 mg, 0.23 mmol) was added to a solution of the coumarin carbonate **156** (74 mg, 0.17 mmol) in dry DMF (1.5 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 30 h, the

reaction mixture was poured into water (20 mL), and the aqueous layer was extracted using EtOAc (2 × 15 mL). The combined organic layers were washed with water (3 × 15 mL), brine (10 mL), then dried over anhydrous Na_2SO_4 and filtered. The filtrate was concentrated to dryness to obtain the crude product. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.56 (d, *J* = 4.2 Hz, 1H), 7.01-6.92 (m, 2H), 6.22 (s, 1H), 5.50 (br, 1H), 5.22 (s, 2H), 4.37-4.32 (m, 3H), 3.47 (s, 3H), 3.18 (br, 2H), 3.08-3.00 (m, 2H), 1.82-1.51 (m, 2H), 1.49-1.23 (s, 13H). Without further purification, the crude compound was subjected to the next step by treatment with triethylsilane (55 µL, 0.34 mmol) and a solution of TFA:DCM (1:1, 2 mL) at room temperature under an inert atmosphere. Stirring was continued for 40 minutes. The reaction mixture was concentrated under reduced pressure to remove both the solvent and TFA. The residue was redissolved in DCM (2 mL) followed by concentration under reduced pressure and the process of adding DCM followed by the removal of solvents was repeated three times to remove any residual amounts of TFA. The crude product was dissolved in MeOH (300 µL) and precipitated in Et_2O (15 mL) under vigorous stirring. The precipitate was collected, re-dissolved in MeOH (300 µL), and precipitated again into Et_2O. The process was repeated twice more to furnish the coumarin lysine TFA salt **148** (53 mg, 61% yield) as a crystalline white solid. ^{1}H NMR (300 MHz, CD_3OD): δ = 7.67 (d, *J* = 4.3 Hz, 1H), 6.84-6.81 (m, 1H), 6.71 (s, 1H), 6.13 (s, 1H), 4.37-4.33 (m, 2H), 3.81-3.77 (m, 1H), 3.11-3.08 (m, 2H), 1.89-1.87 (m, 2H), 1.49 (br, 4H). ^{13}C NMR (75 MHz, CD_3OD): δ = 171.7, 162.8, 161.8, 155.7, 154.6, 126.1, 113.3, 111.8, 110.5, 102.5, 62.3, 53.5, 40.0, 31.4, 30.3, 29.2, 22.1. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{18}H_{22}N_2O_7$: 379.15; found 379.15.

(*S*)-2-((*tert*-Butoxycarbonyl)amino)-6-(((2-(5-nitro-1,3-dioxoisoindolin-2-yl)ethoxy)carbonyl)amino)hexanoic acid (**158**). Triethylamine (2.20 mL, 16.0 mmol) and disuccinimidyl carbonate (983 mg, 3.89 mmol) were added to a solution of the alcohol **157** (757 mg, 3.20 mmol) in dry acetonitrile (10 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 12 h. The reaction mixture was concentrated under reduced pressure and the residue was dried under high vacuum for 3 h to obtain a crude product that was used in the next step without purification. A small portion of the crude product was purified by column chromatography on activated silica gel using acetone:DCM (1:12) with 1% TEA to deliver corresponding carbonate as a solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 8.67 (s, 1H), 8.61 (d, *J* = 3.9 Hz, 1H), 8.07 (d, *J* = 4.3 Hz, 1H), 4.55 (t, *J* = 4.8 Hz, 2H), 4.10 (t, *J* = 4.8 Hz, 2H), 2.78 (s, 4H). The crude mixture was dissolved in DMF (6 mL) at room temperature under an inert atmosphere. Boc-lysine (1.1 g, 4.8 mmol) was added to the solution under vigorous stirring and stirring was continued for 12 h. The reaction mixture was poured into water (100 mL) and the aqueous layer was extracted using EtOAc (3 × 20 mL). The combined organic layers were washed with water (5 × 20 mL), brine (20 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was dried under reduced pressure and high vacuum to obtain the phthalimide Boc-lysine **158** (1.4 g, 87% yield) as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.62-8.56 (m, 2H), 8.04 (d, *J* = 3.7 Hz, 1H), 7.24 (br, 1H), 6.19 (br, 1H), 5.29 (br, 1H), 4.91 (br, 1H), 4.29-4.27 (m, 2H), 3.99-3.95 (m, 3H), 3.07 (br, 2H), 1.77-1.60 (m, 2H), 1.43-1.21 (m, 13 H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 176.4, 166.3, 166.0, 156.4, 151.9, 136.6, 133.6, 129.6, 124.9, 119.0, 80.3, 62.0, 60.5, 53.3, 41.4, 40.7, 38.6,

32.0, 29.4, 28.5, 22.4. LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{22}H_{28}N_4O_{10}Na$: 531.17; found 531.17.

(*S*)-6-(((2-(5-Amino-1,3-dioxoisoindolin-2-yl)ethoxy)carbonyl)amino)-2-((tert-butoxycarbonyl)amino)hexanoic acid (**159**). Ammonium acetate (174 mg, 2.26 mmol) and $SnCl_2$ (257 mg, 1.35 mmol) were added to a solution of the nitrophthaliimide Boc-lysine **158** (115 mg, 0.22 mmol) in dry MeOH (1.5 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 20 h. EtOAc (20 mL) and water (10 mL) were added to the reaction mixture and stirring was continued for an additional 20 min. The organic layer was separated, washed with water (3 × 10 mL) and brine (10 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated under reduced pressure and the remaining residue was dried under high vacuum to furnish the amino phthaliimide Boc-lysine **159** (102 mg, 92% yield) as an orange solid. ^{1}H NMR (300 MHz, CD_3OD): δ = 7.60 (d, *J* = 4.0 Hz, 1H), 7.24 (s, 1H), 7.10 (d, *J* = 4.2 Hz, 1H), 4.22-4.07 (m, 2H), 4.02 (br, 1H), 3.83-3.81 (m, 2H), 2.98-2.96 (br, 2H), 1.73-1.55 (m, 2H), 1.48-1.26 (m, 13H). ^{13}C NMR (100 MHz, $DMSOd_6$): δ = 174.9, 168.8, 168.3, 157.3, 156.2, 155.5, 135.1, 134.3, 125.5, 125.0, 117.2, 106.1, 106.0, 78.6, 61.3, 54.0, 37.7, 31.0, 29.5, 28.8, 23.5. LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{22}H_{30}N_4O_8Na$: 501.20; found 501.19.

(*S*)-2-Amino-6-(((2-(5-amino-1,3-dioxoisoindolin-2-yl)ethoxy)carbonyl)amino)hexanoic acid (**160**). Triethylsilane (67 µL, 0.41 mmol) and **159** (100 mg, 0.20 mmol) were added to a solution of TFA:DCM (1:1, 1 mL) stirred at room temperature under an inert atmosphere.

Stirring was continued for 40 minutes and the reaction mixture was concentrated under reduced pressure. The obtained residue was dissolved in MeOH (500 µL), and then concentrated under reduced pressure to remove both the solvent and TFA. The process was repeated twice to remove any residual TFA. Finally, the crude product was dissolved in MeOH (500 µL) and precipitated into Et_2O (5 mL) under vigorous stirring. The solid was collected, re-dissolved in MeOH and precipitated into Et_2O. The process of dissolving followed by precipitation was repeated once more to furnish the amino phthalimide lysine TFA salt **160** (102 mg, 80% yield) as a crystalline yellow solid. ^{1}H NMR (400 MHz, CD_3OD): δ = 8.10 (br, 2H), 7.45 (d, *J* = 4.0 Hz, 1H), 7.23 (br, 1H), 7.07 (s, 1H), 6.89 (s, 1H), 6.76 (d, *J* = 4.0 Hz, 1H), 6.47 (br, 2H), 4.10-4.07 (m, 2H), 3.68-3.66 (m, 3H), 2.87 (br, 2H), 1.76-1.60 (br, 2H), 1.38-1.26 (m, 4H). ^{13}C NMR (100 MHz, DMSO- d_6): δ = 168.8, 168.4, 156.5, 155.6, 135.1, 125.5, 117.2, 107.6, 65.6, 53.3, 37.5, 30.6, 29.5, 22.3. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{17}H_{22}N_4O_6$: 379.16; found 379.16.

5-(Dimethylamino)-2-(2-hydroxyethyl)isoindoline-1,3-dione (**162**). 2-Aminoethanol (38 µL, 0.62 mmol) was added to a solution of the dimethyl aminophthalic anhydride **161** (100 mg, 0.52 mmol) in dry toluene (3 mL) stirred at room temperature under an inert atmosphere. The reaction vessel, equipped with a Dean-Stark apparatus, was heated to reflux for 7 h. Next, the reaction mixture was filtered while still warm, and the filtrate was cooled to room temperature. The filtrate was concentrated under reduced pressure, and the residue was dissolved in DCM (10 mL). The organic layer was washed with a saturated aqueous solution of $NaHCO_3$ (2 × 5 mL), brine (5 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate

was concentrated to dryness under reduced pressure to yield the alcohol **162** (109 mg, 89% yield) as a crystalline yellow solid; mp 71-74 °C. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.55 (d, *J* = 4.2 Hz, 1H), 6.97 (s, 1H), 6.70 (d, *J* = 5.6 Hz, 1H), 3.79-3.77 (m, 4H), 3.05 (s, 6H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 169.8, 169.4, 154.4, 134.6, 125.0, 117.3, 114.7, 105.8, 61.4, 40.8.

5-Dimethylamino)-2-(3-hydroxypropyl)isoindoline-1,3-dione (**163**). The alcohol **163** was synthesized using the same protocol that was used for the synthesis of the alcohol **162**. Yellow solid (92% yield); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.61 (d, *J* = 4.8 Hz, 1H), 7.04 (s, 1H), 6.76 (d, *J* = 5.4 Hz, 1H), 3.76 (t, *J* = 6.4 Hz, 2H), 3.56 (t, *J* = 6.4 Hz, 2H), 3.10 (s, 6H), 1.82-1.79 (m, 2H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 170.3, 170.0, 154.8, 135.1, 125.4, 117.7, 115.1, 106.3, 59.2, 41.03, 34.1, 31.9. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{13}H_{17}N_2O_3$: 249.12; found 249.12.

2-(5-(Dimethylamino)-1,3-dioxoisoindolin-2-yl)ethyl (2,5-dioxopyrrolidin-1-yl) carbonate (**164**). Triethylamine (236 µL, 1.70 mmol) and disuccinimidyl carbonate (174 mg, 0.680 mmol) were added to a solution of the alcohol **162** (80 mg, 0.34 mmol) in dry acetonitrile (2 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 12 h. The reaction mixture was concentrated and the remaining residue was dissolved in $CHCl_3$ (2 mL) for purification by column chromatography on activated silica gel, using acetone:DCM (1:12), to obtain the corresponding carbonate **164** (122 mg, 95% yield) as a white solid; mp 55-57 °C. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.61 (d, *J* = 4.2 Hz, 1H), 7.02 (s, 1H), 6.76 (d, *J* = 4.2 Hz, 1H), 4.45 (t, *J* = 5.4 Hz, 2H), 3.95 (t, *J* = 5.4 Hz, 2H),

3.08 (s, 6H), 2.79 (s, 4H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 169.1, 168.9, 154.8, 151.5, 134.9, 125.4, 117.6, 115.2, 106.3, 68.7, 46.1, 40.9, 36.3, 26.0. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{17}H_{18}N_2O_7$: 376.11; found 376.11.

3-(5-(Dimethylamino)-1,3-dioxoisoindolin-2-yl)propyl (2,5-dioxopyrrolidin-1-yl) carbonate (**165**). The carbonate **165** was synthesized using the same protocol that was used for the synthesis of the carbonate **164**. Yellow solid (96% yield); ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.61 (d, *J* = 4.3 Hz, 1H), 7.05 (s, 1H), 6.76 (d, *J* = 5.4 Hz, 1H), 4.34 (t, *J* = 6.6 Hz, 2H), 3.73 (t, *J* = 6.6 Hz, 2H), 3.09 (s, 6H), 2.82 (s, 4H), 2.15-2.04 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 168.8, 154.5, 134.7, 125.1, 117.5, 114.8, 106.0, 69.0, 45.9, 40.6, 34.2, 27.9, 25.6. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{18}H_{20}N_2O_3$: 390.13; found 390.13.

(*S*)-2-((*tert*-Butoxycarbonyl)amino)-6-(((2-(5-(dimethylamino)-1,3-dioxoisoindolin-2-yl)ethoxy)carbonyl)amino)hexanoic acid (**166**). Boc-lysine (118 mg, 0.480 mmol) was added to a solution of the succinimidyl carbonate **164** (122 mg, 0.320 mmol) in dry DMF (1 mL) stirred at room temperature under an inert atmosphere. Stirring was cotinued for 16 h. The reaction mixture was poured into water (15 mL) and the aqueous layer was extracted using EtOAc (3 × 10 mL). The combined organic layers were washed with water (3 × 10 mL), brine (5 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated to dryness under reduced pressure and then under high vacuum to obtain the phthalimide lysine **166** (150 mg, 90% yield) as a white solid; mp 62-64 °C. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.60 (d, *J* = 4.3 Hz, 1H), 7.03 (s, 1H), 6.75 (d, *J* = 4.3 Hz, 1H), 6.09 (br, 1H), 5.37 (br,

1H), 4.93 (s, 1H), 4.24 (br, 2H), 4.51 (s, 1H), 3.88-3.74 (m, 2H), 3.08 (s, 6H), 1.77-1.69 (m, 2H), 1.40-1.22 (m, 13H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 176.03, 169.4, 169.0, 156.5, 156.0, 154.6, 134.7, 125.1, 117.5, 114.8, 106.0, 80.2, 63.2, 62.5, 53.4, 40.7, 37.4, 31.8, 29.4, 28.5, 22.5. LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{24}H_{34}N_4O_8Na$: 529.22; found 529.22.

(*S*)-2-((*tert*-Butoxycarbonyl)amino)-6-(((3-(5-(dimethylamino)-1,3-dioxoisoindolin-2-yl)propoxy)carbonyl)amino)hexanoic acid (**167**). The Boc-lysine **167** was synthesized using the same protocol that was used for the synthesis of the aminophthalimide Boc-lysine **165**. White solid (91% yield); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.59 (d, *J* = 4.2 Hz, 1H), 7.02 (s, 1H), 6.74 (d, *J* = 4.2 Hz, 1H), 6.11 (br, 1H), 5.40 (br, 1H), 4.86 (s, 1H), 4.28 (s, 1H), 4.11-4.05 (m, 2H), 3.70-3.67 (m, 2H), 3.12-3.07 (m, 8H), 1.94-1.93 (m, 2H), 1.80-1.71 (m, 2H), 1.47-1.23 (m, 13H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 175.9, 169.4, 169.0, 156.8, 156.0, 154.5, 134.8, 125.0, 117.6, 114.7, 105.9, 80.1, 62.4, 53.3, 40.6, 34.9, 32.0, 29.5, 28.4, 22.3. LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{25}H_{36}N_4O_8Na$: 543.24; found 543.24.

(*S*)-2-Amino-6-(((2-(5-(dimethylamino)-1,3-dioxoisoindolin-2-yl)ethoxy)carbonyl)amino)hexanoic acid (**110**). Triethylsilane (820 µL, 5.13 mmol) and the Boc-lysine **166** (1.30 g, 2.56 mmol) were stirred for 40 min in TFA:DCM (1:1, 15 mL) at room temperature under an inert atmosphere. The reaction mixture was concentrated under reduced pressure and the obtained residue was dissolved in MeOH (10 mL) followed by evaporation of all volatiles. The process was repeated twice to remove any residual TFA. Finally, the crude product was dissolved in MeOH (3 mL) and the solution was added

dropwise to Et_2O (200 mL) under vigorous stirring to precipitate the product. The process of dissolving followed by precipitation was repeated once more to obtain the phthalimide lysine TFA salt **110** (1.2 g, 92% yield) as a crystalline yellow solid. ^{1}H NMR (300 MHz, CD_3OD): δ = 7.61 (d, *J* = 4.2 Hz, 1H), 7.09 (s, 1H), 6.94 (d, *J* = 4.2 Hz, 1H), 4.23 (t, *J* = 5.1 Hz, 1H), 3.85-3.81 (m, 4H), 3.14 (s, 6H), 3.04-3.02 (br, 2H), 1.93-1.78 (br, 2H), 1.51-1.28 (m, 4H). ^{13}C NMR (75 MHz, CD_3OD): δ = 170.6, 169.3, 169.0, 157.5, 154.9, 134.6, 124.5, 117.0, 114.7, 105.3, 61.7, 52.6, 39.9, 39.3, 37.1, 29.9, 29.1, 21.8. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{19}H_{27}N_4O_6$: 407.20; found 407.20.

(*S*)-2-Amino-6-(((2-(5-amino-1,3-dioxoisoindolin-2-yl)ethoxy)carbonyl)amino)hexanoic acid (**168**). The phthalimide lysine TFA salt **168** was synthesized using the same protocol that was used for the synthesis of the aminophthalimide lysine TFA salt **167**. Yellow solid (92% yield); ^{1}H NMR (300 MHz, CD_3OD): δ = 7.59 (d, *J* = 4.3 Hz, 1H), 7.06 (s, 1H), 6.90 (d, *J* = 5.5 Hz, 1H), 4.06-3.95 (m, 3H), 3.68 (t, *J* = 6.6 Hz, 2H), 3.29-3.04 (m, 8H), 1.98-1.89 (m, 4H), 1.51-1.42 (m, 4H). ^{13}C NMR (75 MHz, CD_3OD): δ = 170.6, 169.4, 169.2, 157.8, 154.9, 134.7, 124.4, 117.0, 114.7, 105.3, 62.2, 52.6, 39.3, 34.5, 30.0, 28.2, 28.1, 21.9. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{20}H_{29}N_4O_6$: 421.20; found 421.20.

(*E*)-2,5-Dioxopyrrolidin-1-yl 4-(phenyldiazenyl)phenethyl carbonate (**172**). Triethylamine (5.20 mL, 37.6 mmol) and disuccinimidyl carbonate (2.80 g, 11.3 mmol) were added to a solution of the alcohol **171** (1.70 g, 7.51 mmol) in dry acetonitrile (40 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 10 h. The reaction

mixture was concentrated and the crude product was purified by column chromatography on silica gel using EtOAc:DCM (1:13) to furnish the corresponding carbonate **172** (2.6 g, 96% yield) as an orange solid. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.92-7.89 (m, 4H), 7.54-7.47 (m, 3H), 7.39-7.38 (m, 2H), 4.57-4.53 (m, 2H), 3.16-3.12 (m, 2H), 2.81 (s, 4H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 168.8, 152.8, 151.9, 139.6, 131.2, 129.9, 129.3, 13.4, 123.0, 71.3, 34.8, 25.6. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{19}H_{18}N_3O_5$: 368.12; found 368.09.

(*S*,*E*)-2-((*tert*-Butoxycarbonyl)amino)-6-(((4-(phenyldiazenyl)phenethoxy)carbonyl)amino)hexanoic acid (**173**). Boc-lysine (1.90 g, 7.84 mmol) was added to a solution of the succinimidyl carbonate **172** (2.40 g, 6.53 mmol) in dry DMF (15 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 30 h. The reaction mixture was poured into water (100 mL) and the product remaining in the aqueous layer was extracted using EtOAc (2 × 20 mL). The combined organic layers were washed with water (3 × 20 mL), brine (20 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated under reduced pressure and under high vacuum to obtain azobenzene Boc-lysine **173** (3.0 g, 97% yield) as an orange solid. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.82-7.76 (m, 4H), 7.44-7.37 (m, 3H), 7.27-7.26 (m, 2H), 6.30 (br, 1H), 5.31 (s, 1H), 4.90 (br, 1H), 4.24-4.21 (m, 3H), 3.08-3.06 (m 2H), 2.92-2.90 (m, 2H), 1.85-1.69 (m, 2H), 1.41-1.34 (m, 13H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 176.5, 156.9, 152.8, 151.6, 141.7, 131.1, 129.8, 129.3, 123.2, 123.0, 80.3, 65.9, 65.1, 53.4, 40.7, 35.6, 32.1, 29.6, 29.3, 28.5, 22.5, 22.2. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{26}H_{35}N_4O_6$: 499.25; found 499.22.

(2*S*)-2-Amino-6-(((2-(2-nitrophenyl)propoxy)carbonyl)amino)hexanoic acid (**112**). A solution of dry HCl (3 mL, 4 M in dioxane) was added to the solution of the azobenzene Boc-lysine **173** (1.50 g, 3.01 mmol) in dry DCM (12 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 2 h. The reaction mixture was concentrated under reduced pressure and the obtained residue was dissolved in MeOH (10 mL) followed by evaporation of the volatiles. The crude product was dissolved in MeOH (5 mL) and and the solution was added dropwise to Et_2O (200 mL) under vigorous stirring to precipitate the crude product. The precipitate was collected and dissolved in MeOH (5 mL) and reprecipitated into Et_2O (200 mL) to provide the azobenzene lysine hydrochloride **112** (1.2 g, 92% yield) as an orange solid. ^{1}H NMR (300 MHz, DMSO-d_6): δ = 8.38 (br, 3H), 7.89-7.83 (m, 4H), 7.62-7.58 (m, 3H), 7.49-7.46 (m, 2H), 7.16 (s, 1H), 4.22-4.20 (m, 2H), 3.84 (br, 1H), 2.98-2.94 (m, 4H), 1.76 (br, 2H), 1.39-1.36 (m, 4H). ^{13}C NMR (75 MHz, DMSO-d_6): δ = 171.7, 156.8, 152.6, 151.2, 143.1, 132.1, 130.6, 130.1, 123.2, 64.5, 52.5, 35.4, 30.3, 29.5, 22.2. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{21}H_{27}N_4O_4$: 399.19; found 399.17.

Methyl 2,4-dibromobutanoate (**180**). Bromine (2.20 mL, 43.2 mmol) and PBr_3 (70 μL, 0.72 mmol) were added to a solution of the butyrolactone **179** (3.00 mL, 39.3 mmol) in dry MeOH (16 mL) stirred at room temperature under an inert atmosphere. The reaction mixture was heated to reflux for 1 h, cooled to room temperature, and MeOH (16 mL) was added. A saturated aqueous solution of $NaHSO_3$ (5 mL) was slowly added to the reaction mixture under vigorous stirring to decolorise the bromine color, and then brine (20 mL) was added. The aqueous layer was extracted using EtOAc (3 × 20 mL). The combined organic layers

were washed with brine (20 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated and the remaining product was purified by column chromatography on silica gel using hexanes:Et_2O (9:1) to deliver the dibromide **180** (8.3 g, 82% yield) as a colorless oil. The analytical data matched the data reported in the literature.[527]

Methyl 2,4-*bis*(acetylthio)butanoate (**181**). Potassium thioacetate (5.50 g, 48.4 mmol) and KI (1.60 g, 9.65 mmol) were added to the solution of the dibromide **180** (3.00 mL, 39.3 mmol) in dry DMF (10 mL) stirred at room temperature under an inert atmosphere. Stirring was contiuned for 24 h. The reaction mixture was poured into water (100 mL) and the product in the aqueous layer was extracted using EtOAc (3 × 20 mL). The combined organic layers were washed with a saturated aqueous solution of $NaHCO_3$ (2 × 30 mL), brine (20 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated under reduced pressure and the obtained residue was purified by column chromatography on silica gel using EtOAc:hexanes (2:8) to furnish the dithioacetate **181** (4.2 g, 87% yield) as a dark brown oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 4.24-4.22 (m, 1H), 3.78 (s, 3H), 2.93-2.81 (m, 2H), 2.38 (s, 3H), 2.27 (s, 3H), 2.23-1.92 (m, 2H). The analytical data matched the data reported in the literature.[503]

Methyl 1,2-dithiolane-3-carboxylate (**178**). Potassium carbonate (1.30 g, 9.99 mmol) was added to a solution of the dithioacetate **181** (500 mg, 1.99 mmol) in dry MeOH (4 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 40 minutes. The reaction mixture was neutralized with aqueous HCl (1.0 M) and an aqueous solution (10

mL) of $FeCl_3$•$6H_2O$ (1.00 g, 3.98 mmol), and Et_2O (20 mL) were added to the reaction mixture under stirring, and the resulting mixture was contineously stirred for another 3 h while oxygen was bubbled through it using ballon-needle technique. The organic layer was separated and the aqueous layer was extracted with Et_2O (2 × 10 mL). The combined organic layers were washed with a saturated aqueous solution of $NaHCO_3$ (2 × 30 mL), brine (20 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated under reduced pressure and dried under high vacuum to obtain the dithilane **178** (300 mg, 92% yield) as a dark brown oil. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.18-4.15 (m, 1H), 3.71 (s, 3H), 3.31-3.16 (m, 1H), 3.17-2.12 (m, 1H), 2.68-2.64 (m, 1H), 2.35-2.30 (m, 1H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 171.9, 52.9, 52.7, 41.0, 36.5. The analytical data matched the data reported in the literature.[503]

1,2-Dithiolane-3-carboxylic acid (**182**). An aqueous solution of LiOH (2.0 mL, 2 M) was added to a solution of the dithiolane methyl ester **178** (300 mg, 1.82 mmol) in dry THF (5 mL) stirred at 0 °C. Stirring was continued for 2 h at the same temperature. The reaction mixture was acidified using aqueous HCl (1 M) to pH 4.0 at 0 °C, the aqueous layer was extracted with EtOAc (3 × 10 mL) and the combined organic layers were washed with water (20 mL), brine (20 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated under reduced pressure and dried under a high vacuum to furnish the dithiolane acid **182** (201 mg, yield 73%). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.21-4.18 (m, 1H), 3.38-3.32 (m, 1H), 3.20-3.14 (m, 1H), 2.66-2.61 (m, 1H), 2.41-2.36 (m, 1H). ^{13}C NMR (100 MHz,

$CDCl_3$): δ = 177.6, 52.5, 41.3, 36.1. The analytical data matched the data reported in the literature.[503]

***S,S*'-(2-((*tert*-Butyldimethylsilyl)oxy)propane-1,3-diyl) diethanethioate** (**187**)**.** Potassium thioacetate (400 mg, 3.50 mmol) and KI (52 g, 0.31 mmol) were added to the solution of the dibromide **185** (525 mg, 1.59 mmol) in dry DMF (2 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 24 h, the reaction mixture was poured into water (20 mL), and the product in the aqueous layer was extracted using EtOAc (2 × 20 mL). The combined organic layers were washed with a saturated aqueous solution of $NaHCO_3$ (3 × 20 mL), brine (20 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated under reduced pressure and dried under a high vacuum to get the dithioacetate **187** (4.2 g, 87% yield) as a brown oil. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 3.89-3.83 (m, 1H), 3.06-3.01 (m, 4H), 2.34 (s, 6H), 0.91-0.86 (m, 9H), 0.11 (s, 6H). No further analytical data were collected since this approach was discontinued.

***S,S*'-(2-((Tetrahydro-2H-pyran-2-yl)oxy)propane-1,3-diyl) diethanethioate** (**188**). Potassium thioacetate (261 mg, 2.29 mmol) and KI (76 g, 0.45 mmol) were added to the solution of the dibromide **186** (275 mg, 0.91 mmol) in dry DMF (1 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 24 h, the reaction mixture was poured into water (20 mL), and the product in the aqueous layer was extracted with EtOAc (2 × 10 mL). The combined organic layers were washed with a saturated aqueous solution of $NaHCO_3$ (10 × 20 mL), brine (10 mL), dried over anhydrous Na_2SO_4 and then

filtered. The filtrate was concentrated under reduced pressure and then dried under a high vacuum to get dithioacetate **188** (236 mg, 88% yield) as a brown oil. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.78 (s, 1H), 3.94-3.87 (m, 2H), 3.53 (br, 1H), 3.20-3.07 (m, 4H), 2.36-2.31 (m, 6H), 1.78-1.53 (m, 6H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 195.4, 195.2, 98.02, 74.2, 62.7, 33.4, 31.8, 30.7, 25.5, 19.5, 14.4. No further analytical data were collected since this approach was discontinued.

1,3-Dibromopropan-2-yl-(2,5-dioxopyrrolidin-1-yl) carbonate (**190**). Disuccinimidyl carbonate (1.0 g, 4.2 mmol) and DIPEA (1.3 mL, 9.6 mmol) were added to the solution of 1, 3- dibromopropanol (705 mg, 3.20 mmol) in dry acetonitrile (10 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 2 h, the reaction mixture was concentrated under reduced pressure, and the residue was purified by column chromatography on silica gel using EtOAc:hexanes (4:6) to get **190** (572 mg, 52% yield) as a low melting solid. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.09-5.06 (m, 1 H), 3.68 (d, *J* = 6.2 Hz, 4H), 2.82 (s, 4H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 168.4, 78.3, 29.9, 25.7. Due to the instability of **190**, no mass spec data was obtained.

(*S*)-Methyl 2-((*tert*-butoxycarbonyl)amino)-6-((((1,3-dibromopropan-2-yl)oxy)carbonyl)amino)hexanoate (**191**). Boc-lysine methyl ester hydrochloride (465 mg, 1.56 mmol) and TEA (657 µL, 4.60 mmol) were added to the solution of the succinimidyl carbonate **190** (560 mg, 1.56 mmol) in dry DCM (30 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 12 h, the reaction mixture was diluted with

DCM (10 mL) and the organic layer was washed with a saturated aqueous solution of $NaHCO_3$ (2 × 10 mL), brine (10 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated under reduced pressure and the remaining product was purified by column chromatography on silica gel using EtOAc:hexanes (1:1) to furnish the dibromide **191** (710 mg, 90% yield) as a colorless thick oil. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.06-4.99 (m, 2H), 4.92 (br, 1H), 4.26 (br, 1H), 3.71 (s, 3H), 3.61-3.55 (m, 4H), 3.18-3.13 (m, 2H), 1.79-1.66 (m, 1H), 1.66-1.32 (m, 14H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 155.5, 71.5, 53.3, 52.6, 41.0, 32.7, 32.2, 29.3, 28.6, 22.6. Due to the instability of **191**, no mass spec data was obtained.

(*S*)-Methyl 6-((((1,3-*bis*(acetylthio)propan-2-yl)oxy)carbonyl)amino)-2-((tert-butoxycarbonyl)amino)hexanoate (**192**). Potassium acetate (650 mg, 5.69 mmol) and KI (207 mg, 1.25 mmol) were added to a solution of the lysine methyl ester **191** (1.30 g, 2.58 mmol) in dry DMF (10 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 24 h, the reaction mixture was diluted with DCM (20 mL) and the organic layer was washed with a saturated aqueous solution of $NaHCO_3$ (2 × 10 mL), brine (10 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated and the remaining product was purified by silica gel chromatography using EtOAc:hexanes (1:1) to obtain the diacetate **192** (1.2 g, 94% yield) as a dark brown thick oil. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.06, (br, 1H), 4.19-4.88 (m, 1H), 4.70 (br, 1H), 4.25 (br, 1H), 3.71 (s, 3H), 3.23-3.02 (m, 6H), 2.32, (s, 6H), 1.78-1.59 (m, 2H), 1.51-1.32 (m, 13H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 155.6, 71.6, 53.4, 52.5, 40.9, 32.6, 32.1, 30.7, 29.5, 28.6, 22.6. LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{20}H_{34}N_2O_6S_2$ Na: 517.16; found 517.16.

(*S*)-Methyl 6-((((1,2-dithiolan-4-yl)oxy)carbonyl)amino)-2-((tert-butoxycarbonyl)amino)hexanoate (**193**). K_2CO_3 (1.3 g, 9.7 mmol) was added to a solution of the diacetate **192** (1.20 g, 2.40 mmol) in dry MeOH (15 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 45 minutes, the solvent was evaporated and, the remaining residue was resuspended in EtOAc (50 mL). Water (20 mL) was added to the mixture under stirring and the organic layer was separated. The organic layer was washed with water (2 × 20 mL) and brine (20 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated and the crude product was dissolved in $CHCl_3$ (120 mL). Water (50 mL) was added to the solution to establish a biphasic system, and then the reaction mixture was cooled to 0 °C. Sodium acetate (604 mg, 7.2 mmol) and a solution (0.1 M) of iodine in $CHCl_3$ were added slowly to the reaction mixture under stirring until a yellow persistent color appeared. Next, the organic layer was separated, washed with aqueous $Na_2S_2O_3$ (20 mL, 10%), a saturated aqueous solution of $NaHCO_3$ (2 × 20 mL), brine (10 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated under reduced pressure and the product was purified by column chromatography on silica gel using EtOAc:hexanes (1:1) to obtain the dithiolane **193** (979 mg, 84% yield) as a thick oil. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.67-5.66 (m, 1H), 5.25 (br, 1H), 4.92 (br, 1H), 4.28 (br, 1H), 3.72 (s, 3H), 3.34-3.22 (m, 1H), 3.21-3.07 (m, 5H), 1.78-1.65 (m, 1H), 1.63-1.40 (m, 14H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 173.3, 155.6, 80.1, 78.4, 53.2, 52.5, 45.3, 40.7, 32.5, 29.3, 28.5, 22.8, 22.5. LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{16}H_{28}N_2O_6S_2Na$: 431.12; found 431.12.

(*S*)-6-((((1,2-Dithiolan-4-yl)oxy)carbonyl)amino)-2-((tert-butoxycarbonyl)amino)hexanoic acid (**194**). An aqueous solution of LiOH (4.0 mL, 2 M) was added to a solution of the dithiolane **193** (815 mg, 1.99 mmol) in dry dioxane (5 mL) stirred at room temperature. Stirring was continued for 1 h and the reaction mixture was acidified to pH 4.0 using dilute aqueous HCl (8 mL, 1 M) at 0 °C. The aqueous layer was extracted with EtOAc (3 × 30 mL), the combined organic layers were washed with water (2 × 20 mL), brine (20 mL), dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated under reduced pressure and the obtained product was dried under a high vacuum to furnish the dithiolane acid **194** (715 mg, 91% yield). ^{1}H NMR (300 MHz, $CDCl_3$): δ = 5.66 (s, 1H), 5.22 (br, 1H), 5.02 (s, 1H), 4.27 (br, 1H), 3.29-3.15 (m, 6H), 1.82-1.67 (m, 2H), 1.58-1.26 (m, 13H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 177.0, 156.2, 78.9, 53.70, 45.72, 41.13, 32.48, 29.82, 28.94, 22.96. LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{15}H_{26}N_2O_6S_2Na$: 417.1123; found 417.1140.

(*S*)-6-((((1,2-Dithiolan-4-yl)oxy)carbonyl)amino)-2-aminohexanoic acid hydrochloride (**111**). A solution of dry HCl (4.0 mL, 4 M in dioxane) was added slowly to a solution of the dithiolane **194** (715 mg, 1.84 mmol) in dry dioxane (4.0 mL) stirred at room temperature. Stirring was continued for 4 h. The reaction mixture was concentrated under reduced pressure and the obtained residue was redissolved in MeOH (1.0 mL). The solution was added dropwise to Et_2O (250 mL) under vigorous stirring to precipitate the product. The precipitate was redissolved in MeOH (1 mL), and the precipitation process was repeated three times to obtain the dithiolane lysine **111** (450 mg, 75% yield) as a white solid. ^{1}H NMR

(300 MHz, CD_3OD): δ = 5.16 (br, 1H), 4.03-3.90 (m, 1H), 3.20-3.06 (m, 6H), 1.92-1.89 (br, 2H), 1.53-1.41 (m, 4H). ^{13}C NMR (75 MHz, CD_3OD): δ =158.1, 118.6, 79.82, 77.7, 68.6, 53.8, 45.8, 43.2, 41.1, 37.4, 31.3, 30.3, 28.9, 23.4. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{10}H_{19}N_2O_4S_2$: 295.07; found 295.07.

2,5-Dioxopyrrolidin-1-yl (1-oxyl-2,2,6,6-tetramethylpiperidin-4-yl) carbonate (**197**). Triethylamine (2.40 mL, 17.4 mmol) and disuccinimidyl carbonate (2.20 g, 8.70 mmol) were added to the solution of **196** (1.00 g, 5.80 mmol) in dry acetonitrile (20 mL) stirred at room temperature under an inert atmosphere. The reaction mixture was stirred for 10 h, the solvent was removed under reduced pressure, and the obtained residue was directly purified by column chromatography on silica gel using a acetone:DCM (1:20) to obtain the corresponding carbonate **197** (1.5 g, 83% yield) as an orange solid; mp 163-165 °C. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{14}H_{22}N_2O_6$: 314.14; found 314.14. NMR data could not be collected due to the radical nature of the amino acid.

(*S*)-2-Amino-6-((((1-hydroxy-2,2,6,6-tetramethylpiperidin-4-yl)oxy)carbonyl)amino)hexanoic acid (**198**). Boc-lysine (1.80 g, 7.32 mmol) was added to the solution of the succinimidyl carbonate **197** (1.50 g, 4.88 mmol) in dry DMF (15 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 18 h and the reaction mixture was poured into water (100 mL) under stirring. EtOAc (30 mL) was added to the aqueous mixture to dissolve the precipitate. The combined organic layer were washed with water (3 × 20 mL), brine (20 mL), dried over anhydrous Na_2SO_4, and filtered.

The filtrate was concentrated under reduced pressure and dried under high vacuum to yield TEMPO Boc-lysine **198** (2.0 g, 95% yield) as an orange, low melting solid. The crude product was subjected to the next step without further purification. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{21}H_{39}N_3O_7$ $[M+H]^+$: 445.27; found 445.27. Then, the Boc-lysine **198** (780 mg, 1.76 mmol) was added to a solution (14 mL) of TFA:DCM (1:1) stirred at room temperature under an inert atmosphere. Stirring was continued for 45 minutes. The reaction mixture was concentrated under reduced pressure and the obtained residue was redissolved in MeOH (5 mL), and concentrated. The process was repeated two more times to remove any residual TFA. MeOH (1.5 mL) was added to dissolve the crude product and the solution was added dropwise to Et_2O (100 mL) under vigorous stirring to precipitate the crude product. The precipitate was redissolved in MeOH (1.5 mL) and the process of precipitation was repeated once more to obtain the TEMPO lysine TFA salt **113** (610 mg, 78% yield) as a white solid. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{16}H_{31}N_3O_5$: 345.22; found 345.22. NMR data could not be collected due to the radical nature of the amino acid.

(*S*)-2-((*tert*-Butoxycarbonyl)amino)-6-((((1-hydroxy-2,2,6,6-tetramethylpiperidin-4-yl)oxy)carbonyl)amino)hexanoic acid (**199**). Sodium dithionite (1.40 g, 8.21 mmol) was added to a solution of the Boc-lysine **198** (1.40 g, 3.28 mmol) in acetone:H_2O (1:1, 30 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for another 30 minutes and the reaction mixture was concentrated under reduced pressure. Water (20 mL) and EtOAc (30 mL) were added and the organic layer was separated. The organic layer was washed with water (2 × 20 mL), brine (20 mL), dried over anhydrous Na_2SO_4, and filtered.

The filtrate was concentrated under reduced pressure and dried under a high vacuum to obtain TEMP (OH) Boc-lysine **199** (1.4 g, 98% yield) as a white solid. ^{1}H NMR (300 MHz, CD_3OD): δ = 5.36 (s, 1H), 4.99 (br, 1H), 4.12-4.10 (m, 1H), 3.14 (br, 2H), 2.16 (br, 2H), 1.82-1.62 (m, 3H), 1.41-1.20 (m, 24H). ^{13}C NMR (100 MHz, CD_3OD): δ = 178.67, 155.9, 69.8, 66.0, 64.2, 54.6, 54.0, 42.6, 41.0, 33.2, 32.1, 29.9, 29.7, 29.5, 28.7, 28.1, 22.6, 21.2. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{21}H_{40}N_3O_7$: 446.28; found 446.20.

(*S*)-2-Amino-6-((((1-hydroxy-2,2,6,6-tetramethylpiperidin-4-yl)oxy)carbonyl)amino)hexanoic acid (**195**). Triethylsilane (932 µL, 5.80 mmol) and TEMP(OH) Boc-lysine **199** (64 mg, 0.14 mmol) were added to a solution of TFA:DCM (1:1, 20 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 45 minutes and the reaction mixture was concentrated under reduced pressure. The obtained residue was dissolved in MeOH (10 mL) and the evaporation process was repeated twice to remove any residual TFA. MeOH (2 mL) was added and the solution was added dropwise to Et_2O (150 mL) under vigorous stirring to precipitate the product. The precipitation process was repeated once more to obtain the TEMP(OH) lysine TFA salt **195** (1.1 g, 84% yield) as a white solid. ^{1}H NMR (300 MHz, D_2O): δ = 5.05 (br, 1H), 3.79 (s, 1H), 3.10 (br, 2H), 2.31-2.27 (m, 2H), 1.85 (br, 4H), 1.54-1.42 (m, 16H). ^{13}C NMR (100 MHz, CD_3OD): δ = 173.1, 163.2, 157.6, 117.8, 114.9, 68.4, 66.9, 64.3, 53.5, 48.9, 41.2, 39.9, 38.7, 29.7, 28.5, 28.4, 27.6, 27.3, 21.5, 21.1, 19.5. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{16}H_{31}N_3O_5$: 346.23; found 346.30.

6. Tyrosine in protein function

Tyrosine, a non-essential amino acid for animals, is synthesized from the essential amino acid phenylalanine by the enzyme phenylalanine hydroxylase.[528, 529] The tyrosine residue offers exceptional chemistries in proteins by virtue of 1) an amphipatic side chain that can be compatible to both hydrophilic and hydrophobic environments,[530] 2) a polar hydroxyl group (with a pKa of approximately 10) that often interacts via hydrogen bonding and acid-base reactions,[531] and 3) a tyrosyl moiety that has been linked to various redox reactions via stabilizing a free radical.[532, 533] Thus, the various biochemical reactions of a tyrosine residue allow for ionic, non-covalent (via a π-π interaction or Van der Waals forces) and covalent interactions. Specifically, the tyrosine residue introduces acid-base chemistry, cation-π interactions, cross-links formation, and metal chelation in protein.[531, 532, 534] Additionally, the rigid structural character of a tyrosine side chain leads to both high affinity and specific interactions.[530] Thus, the prevalence of tyrosine residue chemistry in a wide range of biological processes emphasizes its role in protein function.[535, 536]

Tyrosine residues in proteins are involved in a variety of cellular processes, including proliferation, intercellular communication, cell motility, apoptosis, and post-translational modifications (PTMs).[535, 537] Recent studies revealed the importance of PTMs at tyrosine residues in proteins via phosphorylation,[535, 538, 539] sulphation,[540, 541] and nitration[542, 543]. For instance, phosphorylation at tyrosine makes up 62% of all the PTMs that occurs in the human body.[544] This reaction is mainly maintained by a cooperative action of protein tyrosine kinases (PTKs), which represent 90 genes, and protein tyrosine phosphatases (PTPs), which

are encoded by 107 genes.[535] Phosphorylation of a tyrosine residue by an angiogenic receptor tyrosine kinase, also called vascular endothelial growth factor receptor 2 (VEGFR2), enhances kinase stability.[545, 546] Studies showed that hypoxia-induced tyrosine phosphorylation of Bcl-2 proteins, anti-apoptotic proteins named as B-cell lymphoma 2 family proteins (Bcl-2), can induce apoptosis.[547-549]

Tyrosine participates in various enzymatic reactions to maintain protein homeostasis.[550] Therefore, misregulation of tyrosine PTMs results in various diseases such as cancer,[551, 552] neural diseases and cytopathogenesis.[535, 536, 553-555] A few examples of proteins that utilize tyrosine residues are myoglobin,[556] dehaloperoxidase, [557-559] and cyclooxygenase.[560-562] Previously, biophysical probes such as **88–92** that are structurally similar to tyrosine but spectroscopically different have been incorporated into proteins through metabolic labeling,[563-565] and amber suppressor methodology (*vide supra*).

Site-specific incorporation of *O*-methyl tyrosine **88** into proteins using an evolved tyrosyl tRNA/tRNA synthetase pair in *E. coli* was reported almost a decade ago.[328] Since the first successful incorporation of *O*-methyl tyrosine, over 50 different unnatural amino acids with diverse functionalities,[352] **88–98** to name a few (Figure 30),[315, 566-568] have been incorporated using engineered orthogonal tyrosyltRNA/tRNA synthetase pairs in *E. coli*, yeast and mammalian cells. The amber suppressor technique has been applied to study protein-protein interactions, post-translational modifications, and signaling pathways.[569-571] These studies have mainly been focused in *E. coli* along with a limited number of investigations accomplished in yeast or mammalian cells;[572] because of a low incorporation

efficiency of the orthogonal tyrosyl tRNA/tRNA synthetase pairs in higher organisms.[58, 335, 569]

Previously, caged tyrosine analogues have been incorporated into proteins using amber suppressor methods. For examples, *o*-NB tyrosine **96** has been genetically incorporated by Deiters et al. into the active site of β-galactosidase in *E. coli* using the *M. jannaschii* tyrosyl $tRNA_{CUA}$/tyrosyl-tRNA synthetase pair.[334] The β-galactosidase activity of the mutant was then restored upon a brief UV irradiation. This method may be useful for the photoregulation of a number of biological processes such as transcription, signal transduction, and cellular trafficking. Similarly, site-specific incorporation of a caged fluorotyrosine **97**–**98** in *E. coli* was reported.[573] Isotope labeled (^{15}N) *o*-nitrobenzyl tyrosine was genetically encoded in the active site of the thioesterase domain of the human fatty acid synthetase (FAS-TE; 33kD) for studying protein function and dynamics by NMR.[574] Additionally, Deiters et. al also reported the spatial and temporal control of gene function in mammalian cells through the site–specific incorporation of *o*-nitrobenzyl tyrosine into Cre recombinase.[67]

6.1. Synthesis of a caged thiotyrosine and its site-specific incorporation into proteins

The SH-group in thiotyrosine is a unique biochemical handle that can be used for achieving selective protein labeling via bioconjugation reactions. Additionally, dimerization or crosslinking between two different protein domains through disulfide linkages may be investigated. Previously, thiotyrosine residues have been incorporated into peptides through solid-phase peptide synthesis, and the resulting peptides, such as angiotensin II analogues,

have been investigated for their agonistic activities.[575] In addition, thiotyrosine residues have been incorporated at two specific positions in polyphenylalanine peptides to establish a disulfide crosslink for studying protein folding via light (270 nm) activation.[576] However, to the best of our knowledge, there has been no investigation on site-specific incorporation of thiotyrosine **209** or caged thiotyrosine **205** into proteins to date. We envisioned that site-specific incorporation of caged thiotyrosine **205** into proteins may provide a new avenue to explore sulfur chemistry in cells. The SH-group, restored after the decaging of *o*-NB thiotyrosine **205** using UV light, could be explored to investigate the aforementioned potential.

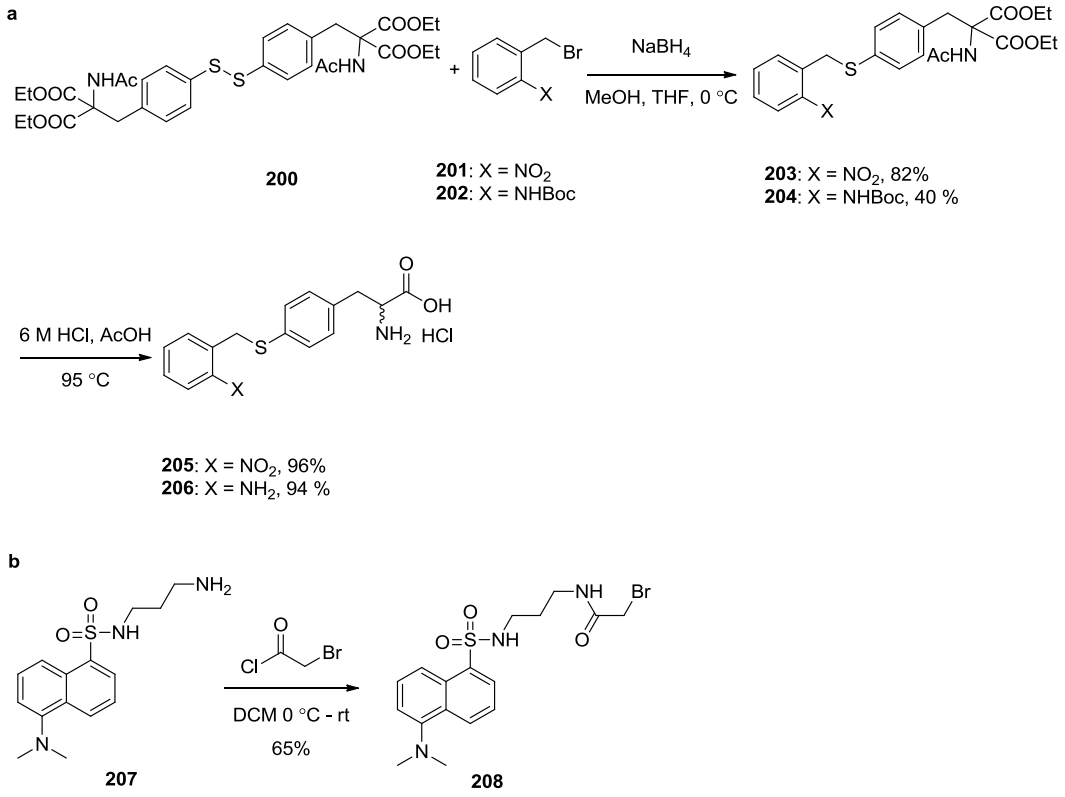
We synthesized both the *o*-nitrobenzyl thiotyrosine **205** and a corresponding *o*-amino benzyl tyrosine **206** (Scheme 28a). The amino benzyl analogue **206** serves as a control for subsequent mass spectrometry analysis. Both the *o*-nitrobenzyl thiotyrosine **205** and the *o*-aminobenzyl tyrosine **206** were prepared from the disulfide ester **200** which was synthesized in two steps starting from toluene disulfide as reported in the literature.[577] Reduction of a disulfide bond of **200** in the presence of $NaBH_4$ followed by alkylation with *o*-nitrobenzylbromide **201** or Boc protected aminobenzyl bromide **202** in a THF:EtOH mixture (1:1) at room temperature furnished corresponding sulfide **203** (82% yield) or **204** (40% yield). Boc protected aminobenzyl bromide **202** was synthesized from commercially available aminobenzyl alcohol in two steps following a reported procedure.[578] Finally, the acetamidomalonate **203** or **204** was heated to reflux in a mixture of aqueous 6 M HCl: CH_3COOH (1:1) to obtain the corresponding amino acid **205** (96% yield) or **206** (94% yield) as crystalline hydrochlorides. A dansyl amine fluorophore **207,** synthesized following a

literature procedure,[579] was then reacted with bromoacetyl chloride in DCM to deliver the desired dansyl fluorophore **208** in 65% yield (Scheme 28b).

Upon the completion of caged thiotyrosine **205** synthesis, decaging experiments were performed either in MeOH or in PBS buffer at pH 7.0 through UV irradiation (365 nm). A solution of caged thiotyrosine (0.1 mM in PBS buffer) was subjected to UV irradiation (365 nm) for 5, 10, or 15 minutes using a transilluminator, and the products were analyzed by HPLC. Results showed that nearly half of the caged thiotyrosine **205** was consumed upon 5 minutes of UV irradiation. Irradiation of the sample for up to 10 minutes showed a gradual decrease in the starting material, but did not result in complete consumption of the starting material. Although the HPLC chromatogram indicated a decrease in starting material upon irradiation, the anticipated decaged thiotyrosine was not observed.

Thus, thiotyrosine **209** was prepared (Scheme 29) as a reference compound for the aforementioned decaging study. The disulfide ester **200** was treated with $NaBH_4$ in MeOH at room temperature to reduce the disulfide bond, providing the corresponding thiol intermediate, which was subsequently heated to reflux in a solution of acetic acid in aqueous HCl (6 M) to deliver thiotyrosine **209**. The thiotyrosine **209** solution (0.1 mM in PBS buffer) was analyzed by HPLC under the same conditions mentioned before. Under these conditions (0.1 mM thiotyrosine in PBS buffer at pH 7.0), detection of the thiotyrosine was very poor and lacked a well-resolved peak. Increasing the concentration of thiotyrosine to 0.5 mM resulted in a better signal in the chromatogram. With these observations, we anticipate that the decaging profile of the caged thiotyrosine **205** at 0.5 mM would result in the detection of thiotyrosine. Therefore, thiotyrosine **205** was irradiated with UV light under the same

conditions as before in 0.5 mM solution. However, no clear decaging event was observed through HPLC analysis.



Scheme 28. a) Synthesis of the caged thiotyrosine **205**, the *o*-amino benzyl tyrosine **206**, and b) the dansylfluorophore **208**.

The incorporation of **205** into a superfolder green fluorescent protein (sfGFP)[580] in response to an amber stop codon (TAG) using the *Mj*Tyr tRNA and E10 *Mj*TyrRS variant were tested by Dr. Liu in the Cropp laboratory at Virginia Commeonwealth University. The E10 *Mj*TyrRS variant has mutations at seven active-site residues (Y32G, L65G, H70M,

F108G, D158S, I159M, and Ll62G and it can incorporate *o*-nitrobenzyl tyrosine and its fluoro derivatives but not natural tyrosine into proteins in *E. coli*.[334, 573] The ability to selectively incorporate the caged thiotyrosine **205** may allow to further probe the function of active-site tyrosine residues. The expression resulted in a moderate **205**-dependent protein production (Figure 52, lane 3). Further characterization of the purified protein via ESI-MS suggested that the expressed protein contains a major species with a mass 30 Da less than the expected value. Although the detailed reason for the loss of 30 Da in mass was not understood, it appears consistent with the reduction of a nitro group to an amino group. These modifications could occur at either the pre-translational level (free amino acids) or the post-translational level (proteins containing UAAs).[581, 582]

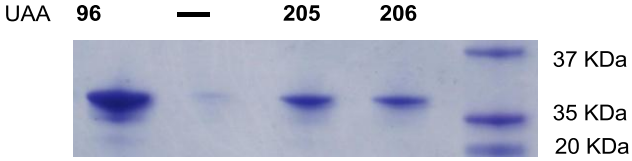


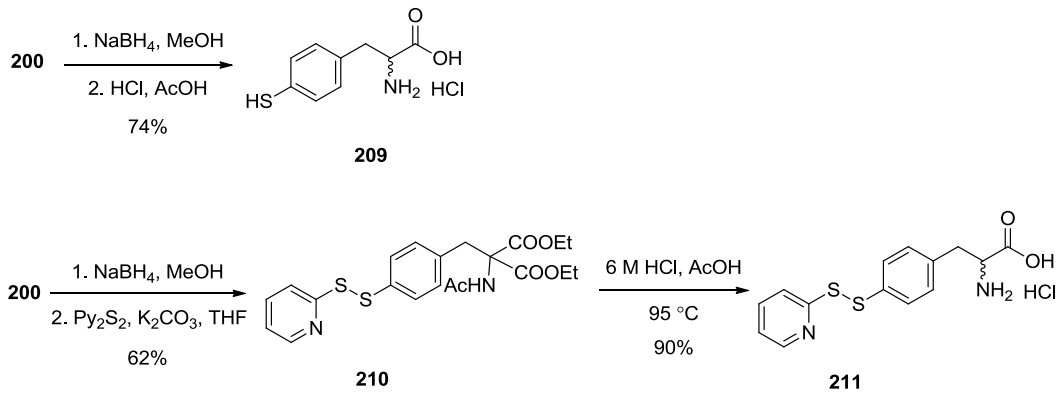
Figure 52. SDS-PAGE gel indicating site-specific incorporation of **205** into sfGFP using *Mj*TyrRS variant E10. The experiment was performed by Jia Liu, Cropp lab, VCU.

To investigate the 30 Da mass loss, regarding the possibility of the reduction of an *o*-nitro group, **206** was incorporated into sfGFP (Figure 52, lane 4) using the same protocol used for **206**. Indeed, **206** was recognized as a substrate by the *Mj*TyrRS variant. Together with their expression and ESI-MS results may indicate that the mis-incorporation of pre-

translationally modified UAA might be the cause of the observed **205**-30 Da protein. To address this issue, the expression was performed in the presence of dicoumarol,[583] an inhibitor for nitroreductases; but the reduction was not inhibited. Alternatively, orthogonal pyrrolysyl tRNA/ tRNA synthetase pairs have been investigated in the Cropp laboratory to identify two *Methanosarcina mazei* PylRS variants that could selectively incorporate **205** into sfGFP. Further studies are in progress.

Thus, a pyridinium disulfide tyrosine **211**, which could deliver a thiotyrosine upon disulfide bond reduction, was synthesized from compound **200** in two steps (Scheme 29). After incorporation into proteins, the pyridinium disulfide of **211** can be reduced simply by treating cells with a mild reducing agent such as dithiothreitol (DTT) or tris(2-carboxyethyl)phosphine (TCEP) , thereby restoring thiol functionality and releasing 2-mercaptopyridone.[584] One of the main advantages of using pyridinium pyridyl disulfide, as compared to that of aliphatic or other aromatic disulfide analogues, is the formation of 2-mercaptopyridone as a predominant tautomer upon reduction.[585] Studies showed that the possibility of involvement of 2-mercaptopyridone for crosslink formation with cysteine or disulfide exchange in an aqueous environment is negligible.[586] It was anticipated that the pyridinium disulfide tyrosine **211** may be incorporated into proteins by tRNA/tRNA synthetase pairs because of the structural similarity, To prepare **211**, the disulfide ester **200** was treated with sodium borohydride in methanol to provide the corresponding thiol, which was subsequently reacted with commercial 2,2′-dipyridyl disulfide (Py_2S_2) in THF to furnish the pyridinium disulfide ester **210** in 62% yield.[587] Finally, the disulfide tyrosine ester **210** was stirred in a solution of aqueous HCl (6 M) in glacial acetic acid and the mixture was

heated to generate pyridinium disulfide tyrosine **210** in 90% yield. The site-specific incorporation of **211** was tested using several PylRS variants by Jihe Liu in the Deiters laboratory. Unfortunately, no incorporation of the amino acid **211** has been achieved to date.

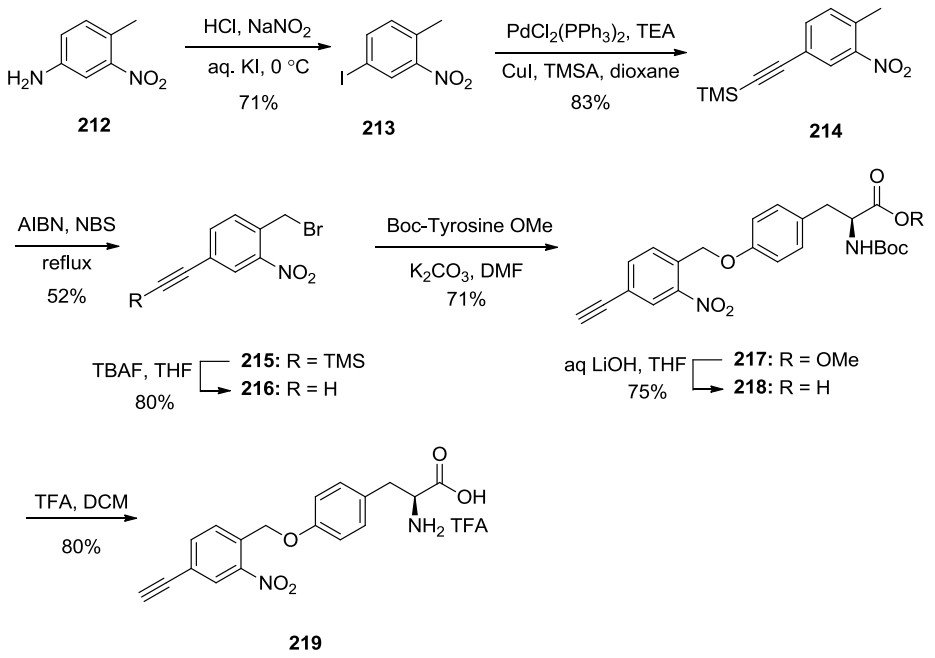


Scheme 29. Synthesis of thiotyrosine (**209**) and the pyridinium disulfide tyrosine **211**.

6.2. Synthesis of the para-ethyne-caged tyrosine and its site-specific incorporation attempts

A terminal alkyne moiety in proteins is a potential site for a copper catalyzed alkyne-azide cycloaddition reaction.[588, 589] Here, we installed an acetylene moiety, the smallest alkyne functionality to impose minimal steric bulk, at the 4-position of the *o*-NB caging group to furnish **219** (Scheme 30). The synthesis of 4-ethynyl-*o*-nitrobenzyl tyrosine (**219**) was accomplished in six steps starting from the commercially available 4-amino-2-nitrotoluene (**212**). The amine group of **212** was converted to an iodide by reaction with $NaNO_2$ in the presence of aqueous HCl at 0 $^{\circ}$C for the *in situ* generation of a diazonium salt, which upon reaction with aqueous KI was converted to the iodide **213** in 71% yield.[590] The iodide **213** was reacted with trimethylsilyl acetylene (TMSA) using Pd catalysis to provide

compound **214** in 82% yield.[591] Benzylic bromination of **214** in the presence of NBS and AIBN in carbon tetrachloride at 90 °C gave the bromide **215** in 52% yield. Attempted optimizations by employing different ratios of NBS or liquid bromine (Br_2) as the brominating reagent in either CCl_4 or acetonitrile did not improve the yield. TMS group deprotection using TBAF in THF at -78 °C gave the desired bromide **216** in 80% yield. The bromide **216** was then reacted with Boc-Tyr OMe in the presence of K_2CO_3 in DMF at room temperature to deliver the ester **217** in 71% yield. Methyl ester hydrolysis of **217** using aqueous LiOH in THF resulted in the corresponding acid **218**, which upon treating with TFA:DCM (1:1) at room temperature gave a white, crystalline TFA salt of 4-ethynyl-*o*-nitrobenzyl tyrosine **219** in 80% yield.



Scheme 30. Synthesis of the *p*-ethyne-caged tyrosine **219**.

The site-specific incorporation of **219** was tested using several PylRS variants by Jihe Liu in the Deiters group. Unfortunately, the amino acid **219** could not be incorporated.

6.3. Synthesis of caged azatyrosine and its site-specific incorporation

Azatyrosines, isolated from *Streptomyces chibanesis*,[592] **220** and **222** (Figure 53) have been studied in a variety of chemical and biological processes.[593-595] Studies also showed that the azatyrosine moiety undergoes pH dependent tautomerization.[596] In addition, azatyrosine exists predominantly in a lactam form in a polar solvent such as water, whereas the lactim tautomer predominates in non-polar or hydrophobic medium. Therefore, the solvent polarity plays a crucial role in altering its equilibrium state (Figure 53).[597] It was envisioned that the structural uniqueness of the hydroxy pyridine moiety (and its pyridone tautomer) is linked with the distinct biochemical features of azatyrosines.[598] In the past, physico-chemical properties of the 2-hydroxy pyridine (pKa = 11.6) functionality have been studied in detail including, proton transfer,[599] tautamerization,[596] hydrogen bonding,[599] metal coordination,[600] antibiotic activity,[594] and anticancer activities.[601-603]

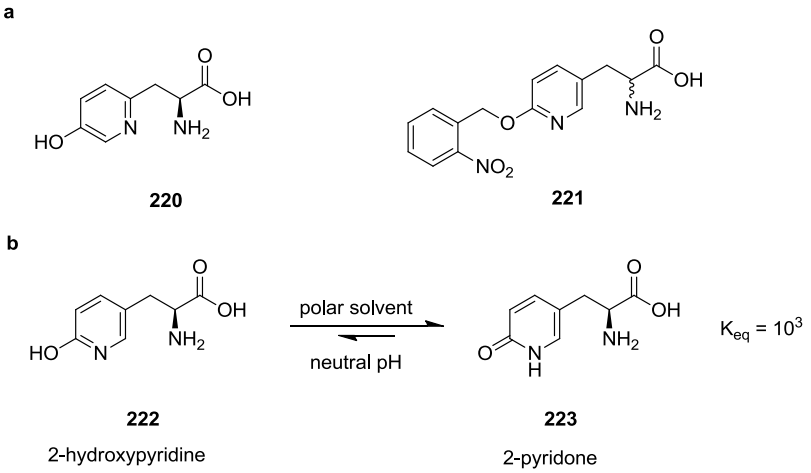
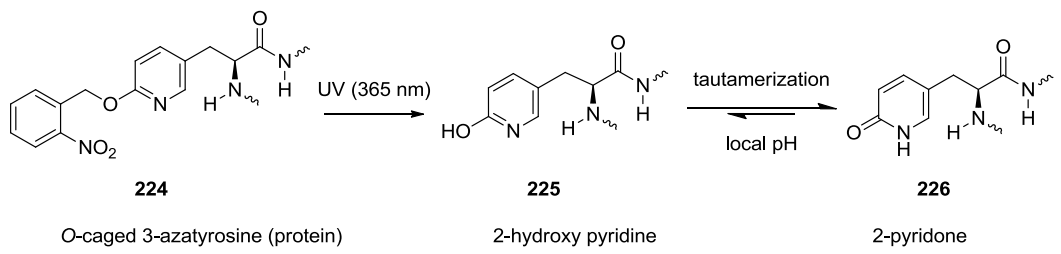


Figure 53. Azatyrosine analogues a) 2-azatyrosine **220**, 3-azatyrosine **222**, and caged azatyrosine **221**; and b) azatyrosine in equilibrium between 2-hydroxy pyridine **222** and 2-pyridone **223** tautomers; K_{eq} = equilibrium constant.

Investigations the properties of the pyridone moiety have revealed potential applications towards protein functions and dynamics. The characteristic features of 2-pyridone chemistry may be mediated via 1) acid-base reaction,[599, 604] 2) conjugated dual hydrogen bonding,[605, 606] 3) metal coordination,[607] and 4) non-covalent interactions.[608]

We envisioned that the caged 3-azatyrosine **221** may be used to investigate protein function performing site-specific incorporation. After site-specific incorporation, upon a brief UV irradiation, **221** undergoes photolysis thereby generating **222**, which spontaneously tautamerizes into the thermodynamically more stable 2-pyridone tautomer **223**. The equilibrium state is maintained based on the microenvironment such as polarity of the medium or the local pH (Scheme 31). In the past, only a few methods for the synthesis of

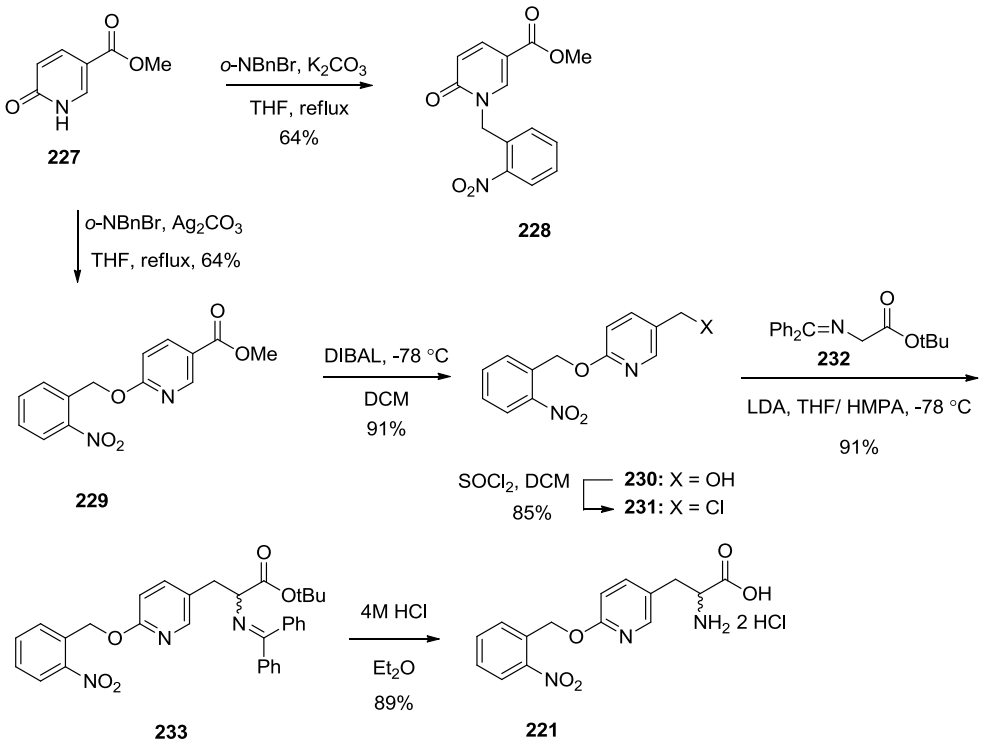
azatyrosine analogues have been reported.[609, 610] Here, we report the preparation of the caged azatyrosine to achieve photochemical control on protein function.



Scheme 31. A schematic of light-activation of the caged azatyrosine resulting in a 2-hydroxy pyridine **225** in equilibrium with a 2-pyridone **226**.

The synthesis of caged azatyrosine **221** is completed in seven steps from the known pyridone methyl ester **227** (Scheme 32). A selective *O*-alkylation of **227** with *o*-nitrobenzyl bromide **201** was achieved in 64% yield using Ag_2CO_3 in THF when heated under reflux.[611] The use of inorganic bases such as K_2CO_3, Na_2CO_3, or Cs_2CO_3 in various solvents such as THF and acetonitrile at room temperature or under reflux resulted in only the *N*-alkyl product **228**. It has been shown that pyridone undergoes *N*-alkylation in the presence of alkali metal carbonates such as K_2CO_3, Na_2CO_3, or Cs_2CO_3, whereas *O*-alkylation occurs when using Ag_2CO_3 as the base.[611, 612] We confirmed this regio-selectivity in our synthesis and the observed *N*- vs *O*-alkylation in pyridone **227** were consistent with literature reports. We observed that benzylic protons of **228** (*o*-$NBCH_2N$ at 5.86 ppm) are more de-shielded as compared to those of **229** (*o*-$NBCH_2O$ at 5.82 ppm). The deshielding of the benzylic protons of **228** as compared to that of **229** can be attributed to the presence of carbonyl group in **228**.

Besides, the 2-pyridone protons on **228** are found to be slightly de-shielded as well, compare to than that of the 2-hydroxy pyridine protons on **229** (see Experimental Section). Furthermore, the *N*-alkyl ester **228** did not undergo decaging on using UV 365 nm in 0.1 mM in PBS buffer. Thus, only the *O*-alkyl product was pursued. The *O*-alkyl ester **229** was reduced to the corresponding alcohol **230** with DIBAL in DCM at -78 $^{\circ}$C. The alcohol **230** was subsequently reacted with thionyl chloride to provide the chloride **231** in 85% yield. The chloride **231** was reacted with the known *t*-butyl imine glycinate **232** in the presence of LDA in THF:HMPA at -78 $^{\circ}$C to deliver the caged ester **233** in 91% yield. The reaction was found to be low yielding without the use of HMPA. Finally both *t*-butyl protecting groups were removed when the ester **233** was treated with aqueous HCl (4 M), furnishing the caged azatyrosine **221** in 89% yield.



Scheme **32**. Synthesis of the caged azatyrosine **221**.

The decaging experiment was performed with **221** (0.1 mM in PBS buffer, at pH 7.0 with UV irradiation at 365 nm for either 5 or 10 min) and samples were analyzed by HPLC. The caged azatyrosine **221** was completely consumed in 10 minutes of UV irradiation, but it could not be assessed if azatyrosine was formed. The site-specific incorporation of caged azatyrosine **221** into sfGFP using the *M. barkeri* PylRS mutant EV16-3 was performed by Jihe Liu; and **221**-dependent protein expression was observed (Figure 54). The PylRS mutant EV16-3 has the mutations Y349F, N311A, and C313A. The SDS-PAGE analysis indicated that the PylRS mutant EV16-3 incorporates the caged azatyrosine **221**, resulting in a sfGFP-

azatyrosine expression. The protein expression has been further confirmed by LRMS-ESI analysis (observed at 28398.56 and calculated 28398.85 Da).

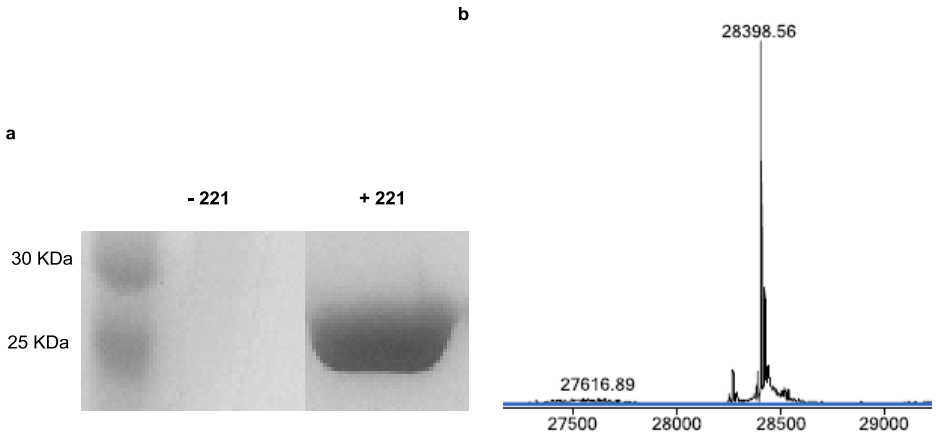


Figure 54. Caged azatyrosine **221** incorporation into a sfGFP by the PylRS mutant EV16-3. a) SDS-PAGE indicates the sfGFP expression only in the presence of **221**, b) ESI-MS spectrum indicating that **221** is incorporated into the protein. These experiments were performed by Jihe Liu in the Deiters lab.

6.4. Experimental data for synthesized compounds

Diethyl 2-acetamido-2-(4-((2-nitrobenzyl)thio)benzyl)malonate (**203**). Sodium borohydride (640 mg, 1.74 mmol) was added to a solution of the acetamidomalonate **200** (200 mg, 0.29 mmol) in THF:MeOH (1:1, 5 mL) stirred at 0 °C under an inert atmosphere. After 2 h, a solution of *o*-nitrobenzyl bromide (**201**, 138 mg, 0.630 mmol) in THF (1 mL) was added to the reaction mixture at 0 °C. Stirring was continued overnight and the reaction mixture was allowed to warm to room temperature. A saturated aqueous solution of NH_4Cl (3 mL) was added and the solution was concentrated under reduced pressure to remove organic solvents. EtOAc (20 mL) was added to the aqueous mixture and the organic layer was washed with a saturated aqueous solution of $NaHCO_3$ (2 × 20 mL), brine (20 mL), and dried over anhydrous Na_2SO_4. The filtrate was concentrated under vacuum and the obtained residue was purified by column chromatography on silica gel using EtOAc:hexanes (1:1) to obtain **203** (230 mg, yield 82%) as a foam. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.90 (d, *J* = 6.6 Hz, 1H), 7.22 (m, 7.45-7.31, 2H), 7.19 (d, *J* = 3.9 Hz, 2H), 7.10 (d, *J* = 3.9 Hz, 2H), 6.83 (s, *J* = 3.9 Hz, 2H), 6.52 (s, 3H), 4.39 (s, 1H), 4.24-4.17 (m, 4 H), 3.59 (s, 2H), 2.02 (s, 3H), 1.36-1.25 (m, 6H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 169.92, 169. 26, 167.38, 166.48, 148.52, 134.88, 133.46, 133.38, 133.11, 132.07, 132.00, 130.54, 128.27, 125.25, 67.11, 62.76, 62.63, 37.38, 37.19, 23.06, 22.76, 14.07. HRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{23}H_{26}N_2O_7SNa$: 497.1358; found: 497.1349.

Diethyl 2-acetamido-2-(4-((2-((*tert*-butoxycarbonyl)amino)benzyl)thio)benzyl)malonate (**204**). Sodium borohydride (50 mg, 1.3 mmol) was added to a solution of acetamidomalonate

200 (180 mg, 0.260 mmol) in THF:MeOH (1:1, 2 mL) stirred at 0 °C under an inert atmosphere. After 90 minutes, a solution of Boc-protected aminobenzyl bromide **202** (76 mg, 0.26 mmol) in THF (1 mL) was added at 0 °C. The reaction mixture was allowed to warm to room temperature and the stirring was continued overnight. The reaction mixture was quenched with a saturated aqueous solution of NH_4Cl (2 mL). The solution was concentrated under reduced pressure to remove THF. EtOAc (10 mL) was added to the remaining aqueous solution and the organic layer was separated and washed with a saturated aqueous solution of $NaHCO_3$ (2 × 10 mL), brine (10 mL), and dried over anhydrous Na_2SO_4. After filtration, the filtrate was concentrated followed by purification using column chromatography on silica gel with EtOAc:hexanes (3:7) to obtain the desired product **204** (109 mg, yield 40%) as a solid. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.71 (br, 1H), 7.26-7.18 (m, 3H), 6.96-6.87 (m, 5H), 6.56 (br, 1H), 4.27-4.21 (m, 4H), 4.01 (s, 2H), 3.59 (s, 2H), 2.00 (s, 3H), 1.51 (s, 9H), 1.27 (t, *J* = 7.2 Hz, 6H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 169.25, 167.48, 153.38, 136.77, 134.90, 133.44, 131.90, 130.60, 130.57, 128.49, 124.02, 123.08, 67.22, 62.08, 37.49, 37.32, 28.46, 23.10, 14.13. LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{28}H_{36}N_2O_7SNa$: 567.2243; found: 567.2201.

2-Amino-3-(4-((2-nitrobenzyl)thio)phenyl)propanoic acid hydrochloride (**205**). Diethyl acetamidomalonate (**205**, 200 mg, 0.420 mmol) was added to a solution of acetic acid in 6 M HCl (1:1, 2 mL) at room temperature. The suspension was then heated to 95 °C under stirring for 30 h. The solvent was removed under reduced pressure and the remaining product was dried under high vacuum for 20 h to furnish the amino acid **205** (150 mg, yield 96%) as a

brown solid. ^{1}H NMR (300 MHz, DMSO- d_6): δ = 8.44 (s, 2H), 8.00 (d, *J* =3.9 Hz, 1H), 7.62-7.59 (m, 2H), 7.54-7.48 (m, 2H), 7.28 (d, *J* = 4.0 Hz, 2H), 7.20 (d, *J* = 4.0 Hz, 2H), 4.50 (s, 2H), 4.12 (br, 1H), 3.08 (d, *J* = 3.1 Hz, 2 H). ^{13}C NMR (100 MHz, DMSO- d_6): δ = 170.20, 148.30, 133.60, 133.57, 133.49, 132.57, 132.11, 130.31, 129.53, 128.90, 125.17, 53.03, 35.06, 34.06. HRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{16}H_{17}N_2O_4S$: 333.0909; found: 333.0901.

2-Amino-3-(4-((2-aminobenzyl)thio)phenyl)propanoic acid hydrochloride (**206**). Diethyl acetamidomalonate (**204**, 90 mg, 0.16 mmol) was added to a solution of acetic acid in 6 M HCl (1:1, 6 mL) at room temperature and the reaction mixture was heated to 90 °C under stirring for 24 h. The solvent was removed under reduced pressure and the obtained product was dried under a high vacuum resulting in the amino acid **206** (58 mg, 94%) as a white solid. ^{1}H NMR (400 MHz, D_2O): δ = 7.36-7.08 (m, 8H), 4.22-4.21 (m, 3H), 3.24-3.10 (m, 2H). ^{13}C NMR (100 MHz, D_2O): δ = 171.73, 134.04, 132.60, 132.38, 131.82, 131.15, 130.22, 129.26, 129.25, 128.59, 124.17, 54.30, 35.40, 34.62. HRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{16}H_{18}N_2O_2SNa$: 325.0981; found: 325.0986.

2-Bromo-*N*-(3-(5-(dimethylamino)naphthalene-1-sulfonamido)propyl)acetamide (**208**). A solution of the dansyl amine **207** (40 mg, 0.13 mmol) in DCM (1.0 mL) was added dropwise to a solution of bromoacetyl chloride (16.2 µL, 0.190 mmol) in DCM (3.0 mL) stirred at 0 °C under an inert atmosphere. The reaction mixture was warmed to room temperature and stirring was continued overnight. The reaction mixture was diluted with

DCM (10 mL), washed with a saturated aqueous solution of $NaHCO_3$ (2 × 10 mL), brine (10 mL), and dried over anhydrous Na_2SO_4. After filtration, the filtrate was concentrated under reduced pressure and the remaining residue was purified by column chromatography on silica gel using acetone:DCM (1:9) delivering the fluorophore **208** (37 mg, 65%) as a solid. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.52 (d, *J* = 4.6 Hz 1H), 8.28-8.20 (m, 2H), 7.58-7.47 (m, 2H), 7.16 (d, *J* = 3.9 Hz, 1H), 6.82 (br, 1H), 6.73 (br, 1H), 5.56-5.49 (m, 1H), 3.96 (s, 1H), 3.77 (s, 1H), 3.35-3.27 (m, 2H), 2.92-2.86 (m, 6H), 1.63-1.55 (m, 2H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 167.08, 166.80, 135.51, 130.13, 129.71, 129.63, 128.47, 142.15, 121.50, 116.14, 45.99, 42.78, 40.14, 36.83, 36.55, 29.63, 29.56, 29.22. HRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{17}H_{22}BrN_3O_3SNa$: 450.0457; found: 450.0460.

2-Amino-3-(4-mercaptophenyl)propanoic acid (209). Sodium borohydride (20 mg, 0.51 mmol) was added to a solution of the acetamidomalonate **200** (35 mg, 0.05 mmol) in EtOH (2 mL) stirred at 0 °C under an inert atmosphere. Stirring was continued for 1 h at the same temperature. The reaction mixture was acidified to pH 4 using an aqueous solution of HCl (1 M) and the product in the aqueous layer was extracted using EtOAc (2 × 5 mL). The combined organic layers were washed with a saturated aqueous solution of $NaHCO_3$ (2 × 5 mL) and brine (5 mL), and then dried over anhydrous Na_2SO_4. After filtration, the filtrate was concentrated under reduced pressure and the product was dried under high vacuum to obtain the thiol. The thiol was dissolved in a solution of CH_3COOH/6M HCl (1:1, 300 µL) at room temperature and the solution was heated to reflux for 30 h. The reaction mixture was then cooled down to room temperature and dried under high vacuum for 20 h to provide

thiotyrosine **209** (13.7 mg, yield 58%) as brown solid. The analytical data matched the data reported in the literature.[575]

Diethyl 2-acetamido-2-(4-(pyridin-2-yldisulfanyl)benzyl)malonate (210). Sodium borohydride (55 mg, 1.4 mmol) was added to a solution of the acetamidomalonate **200** (500 mg, 0.730 mmol) in MeOH (5 mL) stirred at 0 °C under an inert atmosphere and continued stirring for 1 h at the same temperature. The reaction mixture was neutralized using aqueous solution of HCl (1 M) and the resulting aqueous layer was extracted using EtOAc (2 × 20 mL). The combined organic layers were washed with a saturated aqueous solution of $NaHCO_3$ (2 × 20 mL), brine (20 mL), and dried over anhydrous Na_2SO_4. After filtration, the filtrate was concentrated under reduced pressure followed by drying under high vacuum for 2 h to obtain the thiol, which was used in next step without further purification. A solution of the thiol **000** in THF (1 mL) and K_2CO_3 (101 mg, 0.73 mmol) were added to a solution of 2,2′-dipyridyl disulfide (209 mg, 0.94 mmol) in THF (10 mL) at room temperature and stirring was continued for 12 h. The solvent was removed under reduced pressure and the remaining residue was dissolved in EtOAc (20 mL), washed with a saturated aqueous solution of $NaHCO_3$ (2 × 20 mL), brine (20 mL), and dried over anhydrous Na_2SO_4. After filtration, the filtrate was concentrated and the obtained residue was purified by column chromatography on silica gel using EtOAc:hexanes (1:1) delivering **210** (230 mg, yield 82%) as a foam. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 8.45 (d, *J* = 2.4 Hz, 1H), 7.62-7.60 (m, 1), 7.39-7.37 (m, 2H), 7.10-7.07 (m, 1H), 6.93-6.91 (m, 2H), 6.55 (s, 1H), 4.27-4.19 (m, 4H), 3.59 (s, 2H), 1.99 (s, 3H), 1.27-1.23 (m, 6H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 169.2, 167.4, 159.6,

149.6, 137.4, 137.1, 134.7, 130.7, 127.3, 121.0, 119.7, 67.1, 62.8, 37.4, 23.1, 14.1. LRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{21}H_{24}N_2O_5S_2Na$: 471.1024; found 471.1022.

2-Amino-3-(4-(pyridin-2-yldisulfanyl)phenyl)propanoic acid (**211**). Diethyl acetamidomalonate (**210**, 200 mg, 0.44 mmol) was added to a mixture of acetic acid/6 M HCl (1:1, 2 mL) at room temperature and the reaction mixture was heated to 90 °C for 30 h. The solvent was removed under reduced pressure and the obtained product was dried under high vacuum furnishing the amino acid **211** (150 mg, yield 90%) as a white solid. ^{1}H NMR (400 MHz, D_2O): δ = 8.45 (br, 1H), 8.80 (br, 1H), 8.12(m, 1H), 7.60 (m, 1H), 7.42-7.38 (m, 2H), 7.14-7.11 (m, 2H), 4.11 (br, 1H), 3.18-2.95 (m, 2H). ^{13}C NMR (75 MHz, D_2O): δ = 199.0, 118.9, 171.2, 143.2, 143.1, 142.9, 135.6, 135.1, 130.8, 130.3, 124.9, 124.5, 53.9, 35.2. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{14}H_{15}N_2O_2S_2$: 307.04; found 306.96.

Trimethyl((4-methyl-3-nitrophenyl)ethynyl)silane (**214**). TEA (396 µL, 2.85 mmol), $PdCl_2(PPh_3)_2$ (67 mg, 0.09 mmol), CuI (36 mg, 0.19 mmol), and TMSA (405 µL, 2.85 mmol), were added to a solution of the iodide **213** (500 mg, 1.90 mmol) in dioxane (10 mL) stirred at room temperature under an inert atmosphere, and stirring was continued for 5 h. The reaction mixture was concentrated under reduced pressure and the obtained residue was resuspended in DCM (20 mL). The mixture was filtered through diatomacious earth (Fisher Scientific). The filtrate was washed with a saturated aqueous solution of $NaHCO_3$ (2 × 10 mL), brine (10 mL), and dried over anhydrous Na_2SO_4. After filtration, the filtrate was concentrated under reduced pressure and the crude product was purified by column

chromatography on silica gel using DCM:hexanes (1:20), delivering the silane **214** (740 mg, yield 83%) as a white solid; mp 51-52 °C. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 8.05 (s, 1H), 7.55 (d, *J* = 4.6 Hz, 1H), 7.20 (d, *J* = 3.9 Hz, 1H), 2.59 (s, 3H), 0.27 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 136.1, 133.9, 132.9, 128.1, 122.6, 102.4, 96.8, 20.7, 0.02. Due to the instability of **214**, no mass spec data was obtained.

((4-Bromomethyl-3-nitrophenyl)ethynyl)trimethylsilane (**215**). NBS (113 mg, 0.640 mmol) and AIBN (10 mg, 0.06 mmol) were added to a solution of the silane **214** (100 mg, 0.42 mmol) in anhydrous CCl_4 (2 mL) stirred at room temperature under an inert atmosphere. The reaction mixture was heated to reflux for 12 h, cooled to room temperature, and diluted with DCM (10 mL). The organic layer was washed with a saturated aqueous solution of $NaHCO_3$ (2 × 10 mL), brine (10 mL), and dried over anhydrous Na_2SO_4. After filtration, the filtrate was concentrated under reduced pressure and the remaining crude product was purified by column chromatography on silica gel using Et_2O:hexanes (1:20) to deliver the bromide **215** (70 mg, yield 52%) as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.81 (s, 1H), 7.64 (d, *J* = 3.2 Hz, 1H), 7.49 (d, *J* = 5.2 Hz, 1H), 4.80 (s, 2H), 0.27 (s, 9H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 136.6, 136.1, 132.9, 132.7, 128.9, 128.1, 125.4, 101.8, 99.1, 28.6, -0.05. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{12}H_{15}BrNO_2Si$: 312.99; found: 313.14.

1-Bromomethyl-4-ethynyl-2-nitrobenzene (**216**). A solution of TBAF (3.50 mL, 3.53 mmol, 1 M in THF) was added to a solution of the silane **215** (735 mg, 2.35 mmol) in dry THF (7 mL) stirred at -78 °C under an inert atmosphere and stirring was continued for 1 h.

The reaction mixture was warmed to room temperature, and concentrated under reduced pressure. The obtained residue was dissolved in Et_2O (20 mL), washed with a saturated aqueous solution of $NaHCO_3$ (2 × 10 mL), brine (10 mL), and dried over anhydrous Na_2SO_4. After filtration, the filtrate was concentrated followed by purification using column chromatography on silica gel with Et_2O:hexanes (1:20), furnishing the bromide **216** (456 mg, yield 80%) as a tan solid; mp 40-42 °C. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 8.13 (s, 1H), 7.70-7.67 (m, 1H), 7.55-7.52 (m, 1H), 4.8 (s, 2H), 3.26 (s, 1H). ^{13}C NMR (75 MHz, $CDCl_3$): δ = 136.9, 136.8, 132.8, 129.1, 129.0, 81.1, 28.5. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_9H_7BrNO_2$: 239.96; found: 239.11.

(*S*)-Methyl 2-(*tert*-butoxycarbonyl)amino-3-(4-((4-ethynyl-2-nitrobenzyl)oxy)phenyl)propanoate (**217**). Boc-tyrosine methyl ester (44 mg, 0.15 mmol) and K_2CO_3 (51 mg, 0.37 mmol) were added to a solution of the bromide **216** (30 mg, 0.12 mmol) in dry DMF (2 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 12 h at the same temperature and the reaction mixture was poured into EtOAc (10 mL). The organic layer was washed with water (3 × 10 mL), brine (10 mL), and dried over anhydrous Na_2SO_4. After filtration, the filtrate was concentrated under reduced pressure and the crude product was purified by column chromatography on silica gel using EtOAc:hexanes (3:7) to obtain the caged tyrosine **217** (41 mg, yield 76%) as a white solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 8.27-8.23 (m, 1H), 7.88-7.85 (m, 1H), 7.76-7.72 (m, 1H), 7.06 (d, *J* = 4.2 Hz, 2H), 6.90 (d, *J* = 4.2 Hz, 1H), 5.45 (s, 2H), 4.97 (d, *J* = 4.2 Hz, 1H), 4.55 (d, *J* = 3.9 Hz, 1H), 3.71 (s, 3H), 3.21-3.01 (m, 2H), 1.46 (s, 9H). ^{13}C NMR (75 MHz,

$CDCl_3$): δ = 172.5, 157.2, 146.8, 137.3, 137.2, 134.6, 130.78, 130.4, 129.3, 128.8, 128.6, 123.0, 115.1, 81.0, 80.2, 66.8, 54.6, 52.4, 37.7, 28.5.

2-(*tert*-Butoxycarbonyl)amino-3-(4-((4-ethynyl-2-nitrobenzyl)oxy)phenyl)propanoic acid (**218**). An aqueous solution of LiOH (1 mL, 2 M) was added to a solution of the caged tyrosine methyl ester **217** (25 mg, 0.05 mmol) in dioxane (1 mL) stirred at 0 °C under an inert atmosphere. Stirring was continued for 30 minutes at the same temperature, the reaction mixture was acidified with a citric acid solution (5%) to pH 4, and the aqueous layer was extracted using EtOAc (2 × 5 mL). The combined organic layers were washed with water (2 × 10 mL), brine (10 mL), and dried over anhydrous Na_2SO_4. After filtration, the filtrate was concentrated under reduced pressure and the remaining product was dried under high vacuum to furnish the caged **218** (18 mg, yield 75%) as a thick oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 8.14 (s, 1H), 7.61 (d, *J* = 4.0 Hz, 1H), 7.35 (d, *J* = 4.3 Hz, 1H), 6.93 (d, *J* = 4.3 Hz, 1H), 6.67 (d, *J* = 4.0 Hz, 1H), 5.45 (s, 2H), 4.96 (br, 2H), 4.58 (br, 1H), 3.16 (s, 1H), 2.96 (d, *J* = 3.3 Hz, 2H), 1.36 (s, 9H). LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{23}H_{25}N_2O_7$: 441.1662; found: 441.1620. A ^{13}C NMR was not measured due to an insufficient amount of material.

(*S*)-2-Amino-3-(4-((4-ethynyl-2-nitrobenzyl)oxy)phenyl)propanoic acid (**219**). Et_3SiH (150 µL, 0.930 mmol) and the caged Boc-tyrosine **218** (205 mg, 0.460 mmol) were added to a solution of TFA:DCM (1:1, 1.5 mL) stirred at room temperature under an inert atmosphere. Stirring was continued for 1 h and the reaction mixture was concentrated under reduced pressure. The obtained residue was dissolved in MeOH (1 mL) and concentrated, and the

process was repeated two more times to remove residual amounts of TFA. The crude product was dissolved in MeOH (500 µL) and precipitated into Et_2O (10 mL) under vigorous stirring. The precipitate was washed with Et_2O (3 mL) and dried under reduced pressure to obtain the caged tyrosine TFA salt **219** (180 mg, yield 85%) as a white solid. ^{1}H NMR (400 MHz, CD_3OD): δ = 8.16 (s, 1H), 7.72 (d, *J* = 4.0 Hz, 1H), 7.31 (d, *J* = 4.2 Hz, 1H), 7.04 (d, *J* = 4.0 Hz, 1H), 7.73 (d, *J* = 4.2 Hz, 1H), 5.55 (q, 2H), 4.35 (t, *J* = 3.4 Hz, 1H), 3.80 (s, 1H), 3.12 (d, *J* = 3.4 Hz, 2H). ^{13}C NMR (100 MHz, CD_3OD): δ = 168.8, 157.2, 136.6, 130.3, 130.0, 127.9, 124.4, 115.7, 81.0, 64.1, 54.2, 35.6. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{18}H_{17}N_2O_5$: 341.1137; found 341.1135

Methyl 1-(2-nitrobenzyl)-6-oxo-1,6-dihydropyridine-3-carboxylate (**228**). *o*-Nitrobenzyl bromide (169 mg, 0.780 mmol) and K_2CO_3 (90 mg, 0.65 mmol) were added to a solution of the pyridone methyl ester **227** (100 mg, 0.65 mmol) in dry THF (3.0 mL) stirred at room temperature under an inert atmosphere. The reaction mixture was heated to reflux for 10 h and then cooled to room temperature. EtOAc (5.0 mL) and water (10 mL) were added to the mixture under stirring. The organic layer was separated, washed with a saturated aqueous solution of $NaHCO_3$ (2 × 5 mL), brine (5 mL), and dried over anhydrous Na_2SO_4. After filtration, the filtrate was concentrated under reduced pressure and the crude product was purified by column chromatography on silica gel using EtOAc:hexanes (3:2) to furnish the corresponding *N*-substituted oxo-dihydropyridine **228** (130 mg, 70% yield) as a white solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 8.76 (s, 1H), 8.21-8.10 (m, 2H), 7.68-7.60 (m, 1H), 6.90-6.87 (m, 1H), 5.86 (s, 2H), 3.9 (s, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 164.6, 162.5,

143.5, 139.3, 134.3, 131.0, 129.2, 129.0, 125.7, 120.4, 110.7, 52.4, 51.0. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{14}H_{13}N_2O_5$: 289.08; found 289.08.

Methyl 6-((2-nitrobenzyl)oxy)nicotinate (**229**). *o*-Nitrobenzyl bromide (987 mg, 4.57 mmol) and silver carbonate (1.00 g, 3.65 mmol) were added to a solution of the pyridone methyl ester **227** (700 mg, 7.51 mmol) in dry THF (40 mL) stirred at room temperature under an inert atmosphere. The reaction mixture was heated to reflux for 12 h, cooled to room temperature, re-suspended in DCM (20 mL), and filtered to discard silver carbonate. The filtrate was washed with a saturated aqueous solution of $NaHCO_3$ (2 × 20 mL), brine (20 mL), and dried over anhydrous Na_2SO_4. After filtration, the solvent was removed under reduced pressure followed by purification using column chromatography on silica gel with EtOAc:hexanes (1:9) containing 1% TEA to furnish the corresponding methyl ester **229** (650 mg, yield 50%) as a white solid; mp 91-93 °C. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.73 (s, 1H), 8.17-8.14 (m, 1H), 8.08-8.06 (m, 1H), 7.67-7.65 (m, 1H), 7.61-7.57 (m, 1H), 7.45-7.41 (m, 1H), 6.86-6.83 (m, 1H), 5.82 (s, 2H), 3.86 (s, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ =165.8, 150.0, 140.1, 133.8, 129.0, 128.7, 128.5, 125.0, 120.5, 110.8, 65.0, 52.2. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{14}H_{13}N_2O_5$: 289.0824; found 289.0826.

(6-((2-Nitrobenzyl)oxy)pyridin-3-yl)methanol (**230**). DIBAL (12.20 mL, 12.28 mmol, 1 M in THF) was added to a solution of the methyl ester **229** (1.60 g, 5.58 mmol) in dry DCM (20 mL) stirred at -78 °C under an inert atmosphere Stirring was continued for 7 h at the same temperature and the reaction mixture was diluted with DCM (30 mL), slowly quenched with

MeOH (5.0 mL) at -78 °C, and was warmed to room temperature. A saturated aqueous solution of $NaHCO_3$ (5 mL) was added to the reaction mixture. The suspension was filtered through the diatomaceous earth and the filtrate was washed with a saturated aqueous solution of $NaHCO_3$ (2 × 20 mL), brine (20 mL), and dried over anhydrous Na_2SO_4. After filtration, the solvent was removed under reduced pressure and the remaining crude product was purified by column chromatography on silica gel with DCM:EtOAc (4:1) containing 1% TEA to deliver the alcohol **230** (1.2 g, yield 81%) as a white solid; mp 70-71 °C. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.06-8.03 (d, *J* = 4.2 Hz, 1H), 7.97-7.96 (s, 1H), 7.67-7.65 (m, 1H), 7.61-7.54 (m, 2H), 7.42-7.37 (m, 1H), 6.79 (d, *J* = 4.2 Hz, 1H), 5.71 (s, 2H), 4.53 (s, 2H), 2.54 (s, 1H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 163.0, 146.1, 139.2, 134.4, 134.1, 130.2, 129.3, 128.6, 125.3, 111.4, 64.8, 62.7. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{13}H_{13}N_2O_4$: 261.08; found 261.09.

5-Chloromethyl-2-((2-nitrobenzyl)oxy)pyridine (**231**). $SOCl_2$ (1.00 mL, 13.7 mmol) was added drop wise to a solution of the alcohol **230** (1.20 g, 4.61 mmol) in dry DCM (20 mL) stirred at 0 °C under an inert atmosphere. The reaction mixture was warmed to room temperature and stirring was continued for 12 h. The reaction mixture was diluted with DCM (10 mL), slowly quenched with water (20 mL), and neutralized with a saturated aqueous solution of $NaHCO_3$. The organic layer was separated, washed with a saturated aqueous solution of $NaHCO_3$ (2 × 20 mL), brine (20 mL), and dried over anhydrous Na_2SO_4. After filtration, the solvent was removed under reduced pressure and the remaining crude product was purified by column chromatography on silica gel using hexanes:EtOAc (9:1) to obtain

the chloride **231** (1.1 g, yield 85%) as a white solid; mp 73-75 °C. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 9.09-8.07 (m, 2H), 7.67-7.59 (m, 3H), 7.46-7.41 (m, 1H), 6.85-6.83 (m, 1H), 5.77 (s, 2H), 4.51 (s, 2H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 163.0, 146.9, 139.8, 134.0, 133.8, 129.1, 128.4, 127.1, 125.0, 111.5, 64.6, 43.3. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{13}H_{12}ClN_2O_3$: 279.0536; found 279.0535.

***tert*-Butyl 2-(diphenylmethylene)amino-3-(6-((2-nitrobenzyl)oxy)pyridin-3-yl)propanoate** (**233**). *n*-Butyl lithium (2.5 mL, 2.58 mmol, 1 M in hexanes) was added to a solution of diisopropyl amine (394 µL, 2.79 mmol) in THF:HMPA (2:1, 6 mL) stirred at -78 °C under an inert atmosphere and stirring was continued for 30 minutes at the same temperature. Diphenylmethylene glycine *t*-butyl ester (**232**, 823 mg, 2.79 mmol) was added and stirring was continued at -78 °C for 1 h. The chloride **231** (600 mg, 2.15 mmol) was added and the reaction mixture was allowed to warm to room temperature and stirring was continued overnight. The reaction mixture was diluted with EtOAc (20 mL) and quenched with a saturated aqueous solution of NH_4Cl (5.0 mL) at 0 °C. The organic layer was separated, washed with water (2 × 20 mL), brine (20 mL), and dried over anhydrous Na_2SO_4. After filtration, the solvent was removed under reduced pressure and the remaining crude product was purified by column chromatography on silica gel using Et_2O:hexanes (3:7) to obtain the corresponding glycinate **233** (1.0 g, yield 91%) as a white solid; mp 120-122 °C. ^{1}H NMR (300 MHz, $CDCl_3$): δ = 8.07 (d, *J* = 4.6 Hz, 1H), 7.81 (s, 1H), 7.68-7.53 (m, 4H), 7.50-7.26 (m, 8H), 6.71-6.68 (m, 3H), 5.75 (s, 2H), 4.08-4.04 (m, 1H), 3.12-3.08 (m, 2H), 1.43 (s, 9H). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 170.5, 161.7, 147.5, 140.7, 136.3, 134.4,

133.6, 130.4, 129.0, 128.8, 128.6, 128.4, 128.2, 128.1, 127.6, 127.3, 124.9, 110.3, 81.5, 67.5, 64.2, 35.9, 28.1. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{32}H_{32}N_3O_5$: 538.2342; found 358.2352.

2-Amino-3-(6-((2-nitrobenzyl)oxy)pyridin-3-yl)propanoic acid (**221**). An aqueous solution of HCl (15 mL, 4.0 M) was added drop-wise to a solution of the *t*-butyl glycinate **233** (1.00 g, 1.86 mmol) in Et_2O (20 mL) at room temperature. The reaction mixture was stirred for 5 h and the aqueous layer was separated and washed with Et_2O (2 × 20 mL). To remove any benzophenone byproduct, leaving the crude product in the aqueous layer, more Et_2O (20 mL) was added and the biphasic mixture was stirred for 24 h at room temperature. The aqueous layer was separated, concentrated under reduced pressure to dryness, and dried further under a high vacuum for 8 h to obtain the caged azatyrosine **221** (650 mg, 89%) as a white crystalline solid. ^{1}H NMR (400 MHz, D_2O): δ = 8.24-8.18 (m, 2H), 8.13-8.11 (m, 1H), 7.70-7.67 (m, 2H), 7.54 (br, 1H), 7.43-7.41 (m, 1H), 5.74 (s, 2H), 4.31-4.29 (m, 1H), 3.33-3.29 (m, 2H). ^{13}C NMR (100 MHz, D_2O): δ = 170.7, 159.7, 146.7, 140.0, 134.9, 129.9, 129.1, 126.3, 125.6, 111.9, 69.5, 53.2, 31.7. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{15}H_{16}N_3O_5$: 318.1090; found 318.1010.

CHAPTER 7: CAGED PEPTIDES FOR THE PHOTOREGULATION OF GPCRS

7. Peptides in cell signaling

Peptides, polymers of amino acids, typically with fewer than 50 amino acids, have been investigated in various purposes including 1) additives in foods and cosmetics,[613] 2) antimicrobials,[614, 615] 3) drugs for various diseases including cancers,[616] and 3) ligands for studying cell signaling.[617, 618] Therefore, investigations on peptide chemistries have intensified, in both academia and the pharmaceutical industries, through further investigation of their benefits and drawbacks[619-622] Peptide chemistry represents a beneficial avenue for future drug discovery because of the high specificity, high potency, low off-targets side effects, and high chemical as well as biological diversity of the molecules.[620] Currently, more than 100 peptide-based drugs are on the market and over 400 are in advanced preclinical stages.[623] It was estimated that the current market for protein/peptide drugs is over 40 billion USD per year, which covers about 10% of the entire pharmaceutical market.[623] Between 2009-2011, peptide-based drugs made up 8% of the total drugs approved by the US Food and Drug administration (FDA).[623-626] However, a few crucial limitations of peptide drugs such as low systemic stability, high renal clearance, poor membrane permeability, and high manufacturing cost need to be addressed to explore peptide therapy.[623]

Peptides, often synthesized either by traditional synthesis[627, 628] or solid phase synthesis,[629] have been used to interrogate cell signaling,[630] probe viral antigens,[631] enzymatic activities,[632] and neural disorders.[633, 634] For an example, several peptides have been investigated to explore G-protein-coupled receptors (GPCRs) protein functioning in cell

signaling pathways. About 1200 genes encode GPCRs (4.5% of total genes) in humans and they are the targets of 36% of all current marketed drugs.[635-638]

Thus, a clear understanding of the GPCR functioning in cell signaling is important to address the above mentioned receptor related issues. Here, we focus on two different peptides namely met-enkephalin (a pentapeptide), and an *N*-terminal endogenous hexapeptide - SLIGKV (ser-leu-iso-gly-lys-val) found in transmembrane proteins. Met-enkephalin stimulates the opioid receptors producing a signaling cascade (vide infra), to inhibit pain messages in pain pathways[639] whereas SLIGK is involved in protease-activated receptor 2 (PAR2) activity.

We prepared two caged peptides: 1) the caged hexapeptide SLIGK*V **235** (with the caging group installed at the lysine residue, K*) and 2) caged met-enkephalin **237** (with the caging group installed at the amine moiety of tyrosine). We envisioned that the caged hexapeptide **235** can be explored as a light-activated agonist to investigate PAR2 function. Similarly, the caged met-enkephalin **237** may be used to study opioid receptors, a class of GPCR receptors.

*7.1.Caged hexapeptide SLIGK*V in the investigation of PAR2 activation*

PAR2 is regulated by self-activation of tethered sequence SLIGKV in response to endogenous or exogenous proteases.[640] Site-specific proteolysis of the *N*-terminal chain of PAR2 furnish the short sequence SLIGKV that spontaneously binds to PAR2 intramolecularly and activates downstream signaling cascades and maintains both the cellular homeostasis (Figure 55).[635] Extensive studies on the enzyme-activated receptor PAR2 show

that exogenous synthetic peptides may be applied to regulate its functions. However, the success rate is low as only a limited numbers of short synthetic peptides (6-8 amino acids and devoid of the non-specific and proteolytic effects) could produce a significant potency as compared to the native ones in human and rodent PAR2.[641-643] Thus, further investigation of PAR2 using light-activated peptide could be beneficial to study its function in spatio-temporal fashion.

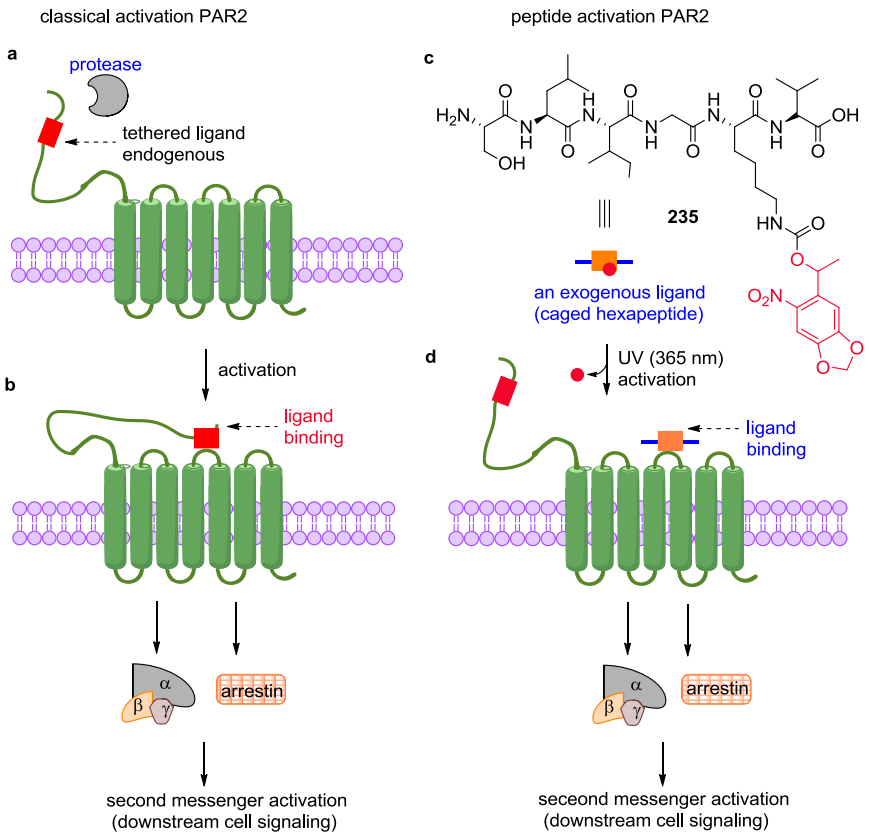
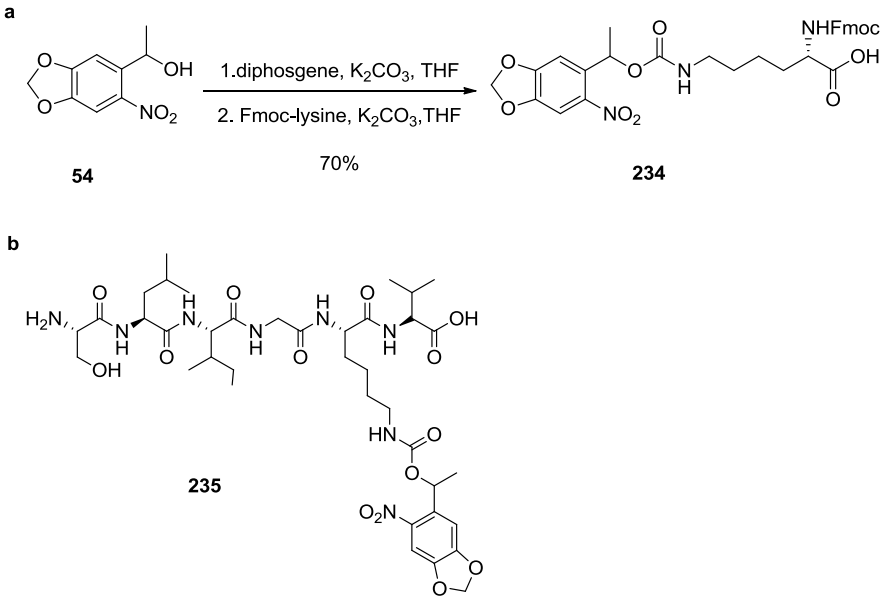


Figure 55. Schematics of mechanism of PAR2 activation: a) protease site-specifically cleaves the tethered sequence activating the native ligand, b) PAR2 activation thereby contributing a signaling cascade, c) an exogenous peptide SLIGK*V (K* = caged lysine), and d) light-activation of PAR2.

Previously, light-cleavable caged peptides have been applied to achieve spatiotemporal control of protein functions and dynamics.[633, 644-646] These caged peptides have been prepared in a post-synthesis reaction via residue specific modification of the peptide.[647-649] Alternatively, a caged amino acid residue can be incorporated during peptide synthesis.[644, 650, 651] To the best of our knowledge, none of the aforementioned approaches has

been employed in PAR2 studies using caged peptides. Here, we report the synthesis of a caged Fmoc-lysine **234** and its subsequent incorporation to the hexapeptide SLIGK*V **235** (K* = caged lysine,Scheme 33) through a solid phase peptide synthesis. Thus, the obtained caged peptide **235** could then used to study the light-activation of protease-activated receptor 2 (PAR2).



Scheme 33. Peptide caging: a) synthesis of a caged Fmoc-lysine **234,** and b) caged hexapeptide (SLIGK*V) **235**; K* = caged lysine residue.

Caged Fmoc-lysine **234** was prepared from the alcohol **54** in a single step in 70% yield. The caged lysine **234** was then incorporated into the peptide following the Fmoc based solid phase peptide synthesis protocol[629] to obtain caged peptide (SLIGK*V) **235**. PAR2 activation using the caged peptide **235** was investigated by Kalyn Brown in the Deiters

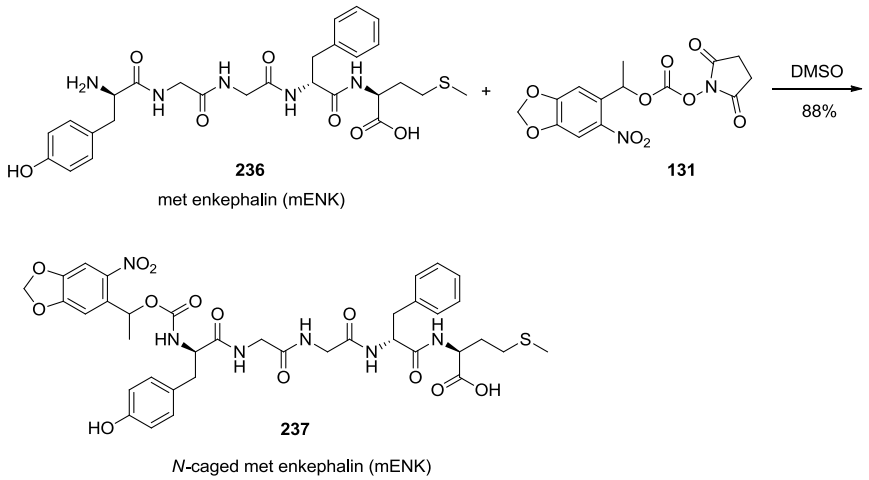
group. The non-caged analogue of SLIGKV was used as a control in KNRK cells. Cells transfected with PAR2-EGFP 24 h prior to the experiments were treated with either the activating peptide SLIGKV or the caged peptide **235** (100 μM), and cells were observed under a Zeiss Observer Z1 microscope. Twenty minutes after SLIGKV addition, translocation of PAR2-EGFP into perinuclear vessels was observed. This result was comparable to the literature report.[652] In cells treated with caged peptide **235**, perinuclear translocation of PAR2-EGFP was observed within 35 minutes. Thus, the presence of a caging group at the lysine residue in peptide **235** did not seem to inhibit the activation of PAR2. Next, the cells were irradiated with UV light (365 nm). No significant change in activity after irradiation (10 seconds, 365 nm) was observed. These results suggest that caging the peptide on the lysine residue is insufficient to inhibit the activation and subsequent translocation of PAR2-EGFP. To achieve photochemical control of PAR2 activation using a caging strategy, a thorough structural-activity relationship (SAR) is necessary in future studies.

7.2. Light-activated caged met-enkephalin

Met-enkephalin, a pentapeptide opioid motif (tyr-gly-gly-phe-met), stimulates the opioid receptors producing a signaling cascade to inhibit pain messages in pain pathways.[639] Opioid receptors, expressed throughout the peripherial (nociceptive pathways) as well as central nervous (CNS) system (brain), are involved in hematopoiesis,[653] pain control,[654] and emotional responses.[655, 656] Previous studies have shown that met-enkephalin could be applied to study either delta (δ) or mu (μ) receptors among the three delta, (δ), mu (μ), and

kappa (κ) opioid receptors.[657] Moreover, the agonist met-enkephalin is found to be associated with the regulation of reactive oxygen species, tumor formation etc.[658, 659]

Here, we hypothesized that the caged met-enkephalin **237** may remain inactive due to protection at the *N*-terminal amino functionality with a caging group. The caging group can be removed upon a brief UV irradiation, thereby releasing a native agonist that may restore activity. Therefore, the specially designed caged met-enkephalin may be a useful molecule to study opioid receptor function with high temporal resolution. Thus, we synthesized the caged met-enkephalin **237** in three steps starting from the alcohol **54**. To begin the synthesis, the alcohol **54** was converted to the corresponding chloroformate by using diphosgene followed by reacting with met-enkephalin **236** in the presence of K_2CO_3 in THF (Scheme not shown). But the reaction did not reach completion after 24 h, resulting in the mixture of crude product along with a large fraction of starting materials. We then chose DSC as an alternative reagent for the activation of the alcohol **54** to prepare the succinimidyl intermediate **131** (Scheme 13). The intermediate **131** was reacted with met-enkephalin **236** in DMSO at room temperature for 24 h to obtain the desired caged met-enkephalin **237** as a crystalline solid (Scheme 34). The product was confirmed by mass spectrometry and the purity was assessed by HPLC. The biological studies using caged met-enkephalin are under current investigations in the Sombers laboratory at NCSU.



Scheme 34. Synthesis of *N*-caged met-enkephalin **237**.

7.3. Experimental data for the synthesized compounds

(2*S*)-2-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-6-(((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)carbonyl)amino)hexanoic acid (234). Diphosgene (420 mL, 3.55 mmol) and K_2CO_3 (977 mg, 7.08 mmol) were added to a solution of the alcohol **54** (500 mg, 2.36 mmol) in THF (10 mL) stirred at 0 °C under an inert atmosphere. The reaction mixture was warmed to room temperature and stirring was continued for 12 h. The solvent was removed under reduced pressure and the remaining residue was dissolved in DCM (20 mL). The organic layer was washed with brine (2 × 20 mL), dried over anhydrous Na_2SO_4, and filtrated. The filtrate was concentrated and the remaining product chloroformate was dried under high vacuum for 1 h. A solution of the chloroformate in dry THF (2 mL) was transferred to a solution of Fmoc-lysine (956 mg, 2.59 mmol) and K_2CO_3 (358 mg, 2.59 mmol) in THF:H_2O (4:1, 20 mL) stirred at 0 °C. The reaction mixture warmed to room temperature and stirring was continued for 12 h. The reaction mixture was concentrated under reduced pressure to remove THF, water (30 mL) was added, and the mixture was neutralized with aqueous HCl (1 M). The aqueous layer was extracted with EtOAc (3 × 20 mL) and the combined organic layers were washed with brine (20 mL), dried over anhydrous Na_2SO_4, and filtered. After filtration, the solvent was removed under reduced pressure, and the remaining residue was purified by column chromatography on silica gel using EtOAc to furnish the caged Fmoc-lysine **234** (1 g, yield 70%) as a yellow solid. ^{1}H NMR (300 MHz, $CDCl_3$, δ ppm): δ = 7.73 (d, *J* = 3.6 Hz, 2H), 7.57 (br, 2H), 7.44–7.25 (m, 5H), 6.96 (s, 1H), 6.23 (d, *J* = 3.1 Hz, 1H), 6.02 (s, 2H), 5.6 (br, 1H), 4.9 (br, 1H), 4.40–4.36 (m, 3H), 4.18 (s, 1H), 3.10 (br, 2H), 1.86–1.23 (m, 10H). ^{13}C NMR (400 MHz, $CDCl_3$): δ = 155.8, 152.5, 147.2, 144.0, 143.9, 141.5,

136.6, 127.9, 127.3, 125.3, 125.0, 120.1, 105.7, 105.4, 103.2, 69.2, 67.3, 53.8, 47.3, 40.7, 32.0, 29.4, 22.3.

(2*R*,5*S*,11*S*,14*S*,17*S*)-17-Amino-11-(*sec*-butyl)-18-hydroxy-14-isobutyl-2-isopropyl-5-(4-(((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)carbonyl)amino)butyl)-4,7,10,13,16-pentaoxo-3,6,9,12,15-pentaazaoctadecan-1-oic acid (235). The synthesis of the caged peptide **235** was performed following the solid phase peptide synthesis protocol starting from the commercial Fmoc-Val-Wang resin (from Sigma Aldrich) and was completed as described in the following steps.

Swelling and Fmoc-deprotection: Fmoc-val-wang resin (50 mg, 0.04 mmol) was taken in a caped syringe (NORM-JECT®,3 mL) in dry DCM (1mL) and allowed to swell for one hour. (A syringe was used as the reaction vessel till the final cleavage stage). The cap was removed and placed into a stopcock valve of the vacuum manifold (Vac-man® laboratory manifold) and washing of the Fmoc-val-resin via gravity filtration was performed using DCM (0.5 mL) and dry DMF (0.5 mL) respectively. The syringe was removed from the vacuum, caped with a syringe tip cap, and a rubber septum securely. Then, a solution (0.5mL) of piperidine:DMF (1:5) was added to the resin at room temperature under an inert atmosphere and reaction vessel was placed on a shaker for 20 minutes. As mentioned before, the solvent was removed and the resin (residue) was washed with dry DMF (0.5 mL) through gravity filtration. The resin was subjected for second round of deprotection by treating the resin with piperidine:DMF solution (0. 5 mL) for 20 minutes and washing with DMF (0.5 mL) to ensure the complete Fmoc deprotection and the product was dried under high vacuum.

Coupling: DCM (0.5 mL) was added to the resin to get swelling in the syringe as described before (all the deprotection and the coupling reactions were performed in the same syringe consecutively). Then, Fmoc amino acid (0.12 mmol), HBTU (0.12 mmol), and DIPEA (0.16 mmol) were added to the mixture and the reaction vessel (syringe) was placed on the shaker for 90 minutes. A few crystal of resin beds were taken out to perform a Kaiser test. Kaiser solutions were made as the reported procedure[629] and the test was performed to ensure the complete consumption of free amine while coupling. In the Kaiser test, a primary amine results in a blue color solution (a positive test) when heated with the Kaiser reagent over a boiling water bath for 2 minutes, where as the amide results in no color change (negative test). After the completion of the coupling reaction, the peptide on resin was washed with DMF (2 × 0.5 mL), MeOH (2 × 0.5 mL), DCM (2 × 0.5 mL), DMF (2 × 0.5 mL), and (2 × 0.5 mL) respectively, and dried under reduced pressure for 2 h. (For this particular coupling step, Fmoc-caged lysine **234** was used). The same procedures mentioned above were performed to attach all of the remaining amino acids, by performing both deprotection and coupling steps, to prepare the hexapeptide (SLIGK*V).

Cleavage and purification: Solid supported peptide was transferred to a vial and DCM (1 mL) was added to swell resin for 1 h. Then, a cocktail (300 µL) of TFA:TES:H_2O (95: 2.5:2.5) was added to the mixture, and stirred for 30 minutes at room temperature. After the cleavage, the resin was filtered out and discarded. The filtrate was concentrated under reduced pressure, redissolved in MeOH (0.5 mL), and concentrated remove the residual amount of water and TFA. Finally the crude product was dissolved in MeOH (200 µL), precipitated in Et_2O under stirring, and centrifuged (5000 rpm, 5 min) to get the desired

peptide **235** (17.7 mg, 52% over all yield). The product was confirmed by mass spectrometry and its purity was assessed by HPLC. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{38}H_{61}N_8O_{14}$: 853.43; found 853.42.

(6*R*,15*R*,18*S*)-15-Benzyl-6-(4-hydroxybenzyl)-18-(2-(methylthio)ethyl)-2-(6-nitrobenzo[d][1,3]dioxol-5-yl)-4,7,10,13,16-pentaoxo-3-oxa-5,8,11,14,17-pentaazanonadecan-19-oic acid (237). The succinimidyl carbonate **131** (1.35 mg, 0.003 mmol) was added to the solution of met-enkephalin **236** (2.00 mg, 0.003 mmol) in dry DMSO (300 µL) at room temperature under an inert atmosphere and continued stirring for 12 h. The solvent was removed under high vacuum for 10 h. Water (0.5 mL) was added to dissolve the residue followed by chloroform (0.6 mL). The mixture was stirred overnight at room temperature, which resulted in a precipitation of a product in the chloroform layer. The liquid (both water and chloroform) was drained, the solid was collected, and washed with chloroform (0.5 mL) to remove any soluble impurities. The solid was dried under high vacuum to obtain caged met-enkephalin **237** (2. 5 mg, 88% yield). The product was confirmed by mass spectrometry and its purity was assessed by HPLC. LRMS-ESI (*m/z*) $[M+H]^+$ calcd for $C_{37}H_{43}N_6O_{13}S$: 811.26; found 811.40.

CONCLUSION

In conclusion, this dissertation presents syntheses of new light-activated phosphoramidites and oligonucleotides. These caged oligonucleotides have been used to study gene regulation in transcription and translation processes. The caged 2′-deoxycytidine is used for preparing caged TFO that was used for light-mediated transcription regulation. Caged 2′- *O*-methyl uridine was used to study controlled spatiotemporal regulation of micro RNA function using light. In addition, a wide variety of new unnatural caged amino acids have been synthesized in this research work. Our newly synthesized analogues include a variety of lysine derivatives with a distinct functional moiety at the $^{\varepsilon}N$ position and different tyrosine derivatives. A diazirine moiety has been introduced to lysine as a photocrosslinking group to study protein-protein interactions. Besides, one-and two-photon caging groups, fluorescent probes, a dithiolane unit, or a spin labeled probe have also been installed to lysine. Furthermore, selective photo-decaging of these amino acid residues in proteins upon UV irradiation to release caging groups and restore biological activities have been achieved that help investigate protein structure, dynamics, localization, and biomolecular interactions in a real-time fashion.

REFERENCES

1. Dahm, R., Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. *Human Genetics* **2008**, *122*, 565-581.

2. Watson, J.; Crick, F., Molecular Structure Of Nucleic Acids - A Structure for deoxyribose nucleic acid. *Nature* **1953**, *171*, 737-738.

3. Watson, J.; Maaloe, O., Nucleic Acid Transfer from parental to progeny bacteriophage. *Biochimica Et Biophysica Acta* **1953**, *10*, 432-442.

4. Hirao, I., Unnatural base pair systems for DNA/RNA-based biotechnology. *Curr Opin Chem Biol* **2006**, *10*, 622-7.

5. Crick, F., Central Dogma of Molecular Biology. *Nature* **1970**, *227*, 561-&.

6. Roy, H.; Ibba, M., Molecular biology: sticky end in protein synthesis. *Nature* 2006, 443, 41-2.

7. Iwai, S., Chemical synthesis of oligonucleotides containing damaged bases for biological studies. *Nucleosides Nucleotides Nucleic Acids* **2006**, *25*, 561-82.

8. Eckert, K. A.; Kunkel, T. A., DNA polymerase fidelity and the polymerase chain reaction. *PCR Methods Appl* **1991**, *1*, 17-24.

9. Temin, H. M., Reverse transcription in the eukaryotic genome: retroviruses, pararetroviruses, retrotransposons, and retrotranscripts. *Mol Biol Evol* **1985**, *2*, 455-68.

10. Emerman, M.; Panganiban, A. T.; Temin, H. M., Insertion of genes into retrovirus genomes and of retrovirus DNA into cell genomes. *Prog Med Virol* **1985**, *32*, 174-88.

11. Ling, J.; Roy, H.; Ibba, M., Mechanism of tRNA-dependent editing in translational quality control. *Proc Natl Acad Sci U S A* **2007**, *104*, 72-7.

12. Deleavey, G. F.; Damha, M. J., Designing chemically modified oligonucleotides for targeted gene silencing. *Chem Biol* **2012**, *19*, 937-54.

13. Kole, R.; Krainer, A. R.; Altman, S., RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat Rev Drug Discov* **2012**, *11*, 125-40.

14. Sinsheimer, R. L., Recombinant DNA. *Annu Rev Biochem* **1977**, *46*, 415-38.

15. Jones, M.; Fayerman, J., Industrial applications of recombinant -DNA technology. *J. Chem. Edu.* **1987**, *64*, 337-339.

16. Jackson, D. A.; Symons, R. H.; Berg, P., Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of Escherichia coli. *Proc Natl Acad Sci U S A* **1972**, *69*, 2904-9.

17. Arnheima, N.; Erlich, H., Polymerase chain-reaction strategy. *Annu. Rev. Biochem.* 1992, 61, 131-156.

18. Arnheima, N., The polymerase chain-reaction - principle, applications and advantages. *Abstracts of Papers of the American Chemical Society* **1992**, *203*, 64-BTEC.

19. VanGuilder, H.; Vrana, K.; Freeman, W., Twenty-five years of quantitative PCR for gene expression analysis. *Biotechniques* **2008**, *44*, 619-626.

20. Mardis, E. R., Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet* **2008**, *9*, 387-402.

21. Mardis, E. R., The impact of next-generation sequencing technology on genetics. *Trends Genet* **2008**, *24*, 133-41.

22. Tian, J.; Ma, K.; Saaem, I., Advancing high-throughput gene synthesis technology. *Mol Biosyst* **2009**, *5*, 714-22.

23. Kim, V. N.; Han, J.; Siomi, M. C., Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* **2009**, *10*, 126-39.

24. Siomi, H.; Siomi, M. C., On the road to reading the RNA-interference code. *Nature* **2009**, *457*, 396-404.

25. Moazed, D., Small RNAs in transcriptional gene silencing and genome defence. *Nature* **2009**, *457*, 413-20.

26. Dorsam, R. T.; Gutkind, J. S., G-protein-coupled receptors and cancer. *Nat Rev Cancer* **2007**, *7*, 79-94.

27. Ferrao, R.; Li, J.; Bergamin, E.; Wu, H., Structural insights into the assembly of large oligomeric signalosomes in the Toll-like receptor-interleukin-1 receptor superfamily. *Sci Signal* **2012**, *5*, re3.

28. Zoncu, R.; Efeyan, A.; Sabatini, D. M., mTOR: from growth signal integration to cancer, diabetes and ageing. *Nat Rev Mol Cell Biol* **2011**, *12*, 21-35.

29. Fong, Y. W.; Cattoglio, C.; Yamaguchi, T.; Tjian, R., Transcriptional regulation by coactivators in embryonic stem cells. *Trends Cell Biol* **2012**, *22*, 292-8.

30. Revyakin, A.; Zhang, Z.; Coleman, R. A.; Li, Y.; Inouye, C.; Lucas, J. K.; Park, S. R.; Chu, S.; Tjian, R., Transcription initiation by human RNA polymerase II visualized at single-molecule resolution. *Genes Dev* **2012**, *26*, 1691-702.

31. D'Alessio, J. A.; Wright, K. J.; Tjian, R., Shifting players and paradigms in cell-specific transcription. *Mol Cell* **2009**, *36*, 924-31.

32. O'Connor, C. J.; Laraia, L.; Spring, D. R., Chemical genetics. *Chem Soc Rev* **2011**, *40*, 4332-45.

33. Crews, C. M.; Splittgerber, U., Chemical genetics: exploring and controlling cellular processes with chemical probes. *Trends Biochem Sci* **1999**, *24*, 317-20.

34. Riedl, J.; Menova, P.; Pohl, R.; Orsag, P.; Fojta, M.; Hocek, M., GFP-like Fluorophores as DNA Labels for Studying DNA-Protein Interactions. *J.Org. Chem.* **2012**, *77*, 8287-8293.

35. Riedl, J.; Pohl, R.; Rulíšek, L.; Hocek, M., Synthesis and photophysical properties of biaryl-substituted nucleos(t)ides. Polymerase synthesis of DNA probes bearing solvatochromic and pH-sensitive dual fluorescent and 19F NMR labels. *J. Org. Chem.* **2012**, *77*, 1026-44.

36. Kalachova, L.; Pohl, R.; Hocek, M., Synthesis of nucleoside mono- and triphosphates bearing oligopyridine ligands, their incorporation into DNA and complexation with transition metals. *Org Biomol Chem* **2012**, *10*, 49-55.

37. Chiu, Y. L.; Rana, T. M., siRNA function in RNAi: a chemical modification analysis. *RNA* **2003**, *9*, 1034-48.

38. Shukla, S.; Sumaria, C. S.; Pradeepkumar, P. I., Exploring chemical modifications for siRNA therapeutics: a structural and functional outlook. *ChemMedChem* **2010**, *5*, 328-49.

39. Kool, E. T., Preorganization of DNA: Design Principles for Improving Nucleic Acid Recognition by Synthetic Oligonucleotides. *Chem Rev* **1997**, *97*, 1473-1488.

40. Zhou, C.; Chattopadhyaya, J., Intramolecular free-radical cyclization reactions on pentose sugars for the synthesis of carba-LNA and carba-ENA and the application of their modified oligonucleotides as potential RNA targeted therapeutics. *Chem Rev* **2012**, *112*, 3808-32.

41. Hirao, I., Unnatural base pair systems for DNA/RNA-based biotechnology. *Curr Opin Chem Biol* **2006**, *10*, 622-7.

42. Hirao, I.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Kawai, R.; Sato, A.; Harada, Y.; Yokoyama, S., An unnatural hydrophobic base pair system: site-specific incorporation of nucleotide analogs into DNA and RNA. *Nat Methods* **2006**, *3*, 729-35.

43. Govan, J. M.; McIver, A. L.; Deiters, A., Stabilization and photochemical regulation of antisense agents through PEGylation. *Bioconjug Chem* **2011**, *22*, 2136-42.

44. Jain, M. L.; Bruice, P. Y.; Szabó, I. E.; Bruice, T. C., Incorporation of positively charged linkages into DNA and RNA backbones: a novel strategy for antigene and antisense agents. *Chem Rev* **2012**, *112*, 1284-309.

45. Bergstrom, D., Unnatural nucleosides with unusual base pairing properties. *Curr Protoc Nucleic Acid Chem* **2009**, Chapter 1, Unit 1.4.

46. Singh, Y.; Murat, P.; Defrancq, E., Recent developments in oligonucleotide conjugation. *Chem Soc Rev* **2010**, *39*, 2054-70.

47. Chan, J. H.; Lim, S.; Wong, W. S., Antisense oligonucleotides: from design to therapeutic application. *Clin Exp Pharmacol Physiol* **2006**, *33*, 533-40.

48. Kurreck, J., Antisense technologies. Improvement through novel chemical modifications. *Eur J Biochem* **2003**, *270*, 1628-44.

49. Crooke, S. T., Progress in antisense technology. *Annu Rev Med* **2004**, *55*, 61-95.

50. Lorenz, P.; Baker, B. F.; Bennett, C. F.; Spector, D. L., Phosphorothioate antisense oligonucleotides induce the formation of nuclear bodies. *Mol Biol Cell* **1998**, *9*, 1007-23.

51. Dias, N.; Stein, C. A., Antisense oligonucleotides: basic concepts and mechanisms. *Mol Cancer Ther* **2002**, *1*, 347-55.

52. Eisen, J. S.; Smith, J. C., Controlling morpholino experiments: don't stop making antisense. *Development* **2008**, *135*, 1735-43.

53. Bedell, V. M.; Westcot, S. E.; Ekker, S. C., Lessons from morpholino-based screening in zebrafish. *Brief Funct Genomics* **2011**, *10*, 181-8.

54. Gleave, M. E.; Monia, B. P., Antisense therapy for cancer. *Nat Rev Cancer* **2005**, *5*, 468-79.

55. Mikhailopulo, I. A.; Miroshnikov, A. I., New trends in nucleoside biotechnology. *Acta Naturae* **2010**, *2*, 36-59.

56. Carell, T.; Brandmayr, C.; Hienzsch, A.; Müller, M.; Pearson, D.; Reiter, V.; Thoma, I.; Thumbs, P.; Wagner, M., Structure and function of noncanonical nucleobases. *Angew Chem Int Ed Engl* **2012**, *51*, 7110-31.

57. Sinkeldam, R.; Greco, N.; Tor, Y., Fluorescent Analogs of Biomolecular Building Blocks: Design, Properties, and Applications. *Chemical Reviews* **2010**, *110*, 2579-2619.

58. Liu, C.; Schultz, P., Adding new chemistries to the genetic code. *Annu Rev Biochem* **2010**, *79*, 413-44.

59. Drahl, C.; Cravatt, B. F.; Sorensen, E. J., Protein-reactive natural products. *Angew Chem Int Ed Engl* **2005**, *44*, 5788-809.

60. Karikó, K.; Weissman, D., Naturally occurring nucleoside modifications suppress the immunostimulatory activity of RNA: implication for therapeutic RNA development. *Curr Opin Drug Discov Devel* **2007**, *10*, 523-32.

61. Young, D.; Lusic, H.; Lively, M.; Yoder, J.; Deiters, A., Gene silencing in mammalian cells with light-activated antisense agents. *Chembiochem* **2008**, *9*, 2937-40.

62. Deiters, A., Principles and applications of the photochemical control of cellular processes. *Chembiochem* **2010**, *11*, 47-53.

63. Dmochowski, I.; Tang, X., Taking control of gene expression with light-activated oligonucleotides. *Biotechniques* **2007**, *43*, 161, 163, 165 passim.

64. Singh, Y.; Murat, P.; Defrancq, E., Recent developments in oligonucleotide conjugation. *Chem Soc Rev* **2010**, *39*, 2054-70.

65. Young, D.; Lively, M.; Deiters, A., Activation and Deactivation of DNAzyme and Antisense Function with Light for the Photochemical Regulation of Gene Expression in Mammalian Cells. *J. Am. Chem. Soc.* **2010**, 6183-6193.

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

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

68. Young, D.; Govan, J.; Lively, M.; Deiters, A., Photochemical regulation of restriction endonuclease activity. *Chembiochem* **2009**, *10*, 1612-6.

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

70. Keefe, A. D.; Pai, S.; Ellington, A., Aptamers as therapeutics. *Nat Rev Drug Discov* **2010**, *9*, 537-50.

71. Govan, J. M.; Lively, M. O.; Deiters, A., Photochemical control of DNA decoy function enables precise regulation of nuclear factor κB activity. *J Am Chem Soc* **2011**, *133*, 13176-82.

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

73. Prokup, A.; Hemphill, J.; Deiters, A., DNA computation: a photochemically controlled AND gate. *J Am Chem Soc* **2012**, *134*, 3810-5.

74. Chen, X.; Ellington, A. D., Shaping up nucleic acid computation. *Curr Opin Biotechnol* **2010**, *21*, 392-400.

75. Lee, H.; Larson, D.; Lawrence, D., Illuminating the chemistry of life: design, synthesis, and applications of "caged" and related photoresponsive compounds. *ACS Chem Biol* **2009**, *4*, 409-27.

76. Kaplan, J.; Forbush, B. r.; Hoffman, J., Rapid photolytic release of adenosine 5'-triphosphate from a protected analogue: utilization by the Na:K pump of human red blood cell ghosts. *Biochemistry* **1978**, *17*, 1929-35.

77. Shah, S.; Jain, P. K.; Kala, A.; Karunakaran, D.; Friedman, S. H., Light-activated RNA interference using double-stranded siRNA precursors modified using a remarkable regiospecificity of diazo-based photolabile groups. *Nucleic Acids Res* **2009**, *37*, 4508-17.

78. Monroe, W. T.; McQuain, M. M.; Chang, M. S.; Alexander, J. S.; Haselton, F. R., Targeting expression with light using caged DNA. *J Biol Chem* **1999**, *274*, 20895-900.

79. Furuta, T.; Takeuchi, H.; Isozaki, M.; Takahashi, Y.; Kanehara, M.; Sugimoto, M.; Watanabe, T.; Noguchi, K.; Dore, T. M.; Kurahashi, T.; Iwamura, M.; Tsien, R. Y., Bhc-cNMPs as either water-soluble or membrane-permeant photoreleasable cyclic nucleotides for both one- and two-photon excitation. *Chembiochem* **2004**, *5*, 1119-28.

80. Ando, H.; Furuta, T.; Tsien, R. Y.; Okamoto, H., Photo-mediated gene activation using caged RNA/DNA in zebrafish embryos. *Nat Genet* **2001**, *28*, 317-25.

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

82. Chaulk, S.; MacMillan, A., Synthesis of oligo-RNAs with photocaged adenosine 2'-hydroxyls. *Nat Protoc* **2007**, *2*, 1052-8.

83. Lusic, H.; Young, D.; Lively, M.; Deiters, A., Photochemical DNA activation. *Org Lett* **2007**, *9*, 1903-6.

84. Deiters, A.; Garner, R. A.; Lusic, H.; Govan, J. M.; Dush, M.; Nascone-Yoder, N. M.; Yoder, J. A., Photocaged morpholino oligomers for the light-regulation of gene function in zebrafish and Xenopus embryos. *J Am Chem Soc* **2010**, *132*, 15644-50.

85. Mayer, G.; Heckel, A., Biologically active molecules with a "light switch". *Angew Chem Int Ed Engl* **2006**, *45*, 4900-21.

86. Deiters, A., Light activation as a method of regulating and studying gene expression. *Curr Opin Chem Biol* **2009**, *13*, 678-686.

87. Riggsbee, C.; Deiters, A., Recent advances in the photochemical control of protein function. *Trends Biotechnol* **2010**, *28*, 468-475.

88. Young, D. D.; Deiters, A., Photochemical control of biological processes. *Org Biomol Chem* **2007**, *5*, 999-1005.

89. Homsher, E.; Millar, N., Caged compounds and striated muscle contraction. *Annu Rev Physiol* **1990**, *52*, 875-96.

90. McCray, J.; Trentham, D., Properties and uses of photoreactive caged compounds. *Annu Rev Biophys Biophys Chem* **1989**, *18*, 239-70.

91. Ellis-Davies, G., Caged compounds: photorelease technology for control of cellular chemistry and physiology. *Nat Methods* **2007**, *4*, 619-28.

92. Anslyn, E. V.; Dougherty, D. A., *Modern Physical Organic Chemistry*. University Science Books: United States of America, **2004**; p 1099.

93. Davis, M.; Kragor, C.; Reddie, K.; Wilson, H.; Zhu, Y.; Dore, T., Substituent effects on the sensitivity of a quinoline photoremovable protecting group to one- and two-photon excitation. *J Org Chem* **2009**, *74*, 1721-9.

94. Lusic, H.; Deiters, A., A new photocaging group for aromatic *N*-heterocycles. *Synthesis* **2006**, 2147-2150.

95. Solomek, T.; Mercier, S.; Bally, T.; Bochet, C. G., Photolysis of ortho-nitrobenzylic derivatives: the importance of the leaving group. *Photochem Photobiol Sci* **2012**, *11*, 548-55.

96. Aujard, I.; Benbrahim, C.; Gouget, M.; Ruel, O.; Baudin, J.; Neveu, P.; Jullien, L., o-Nitrobenzyl photolabile protecting groups with red-shifted absorption: Syntheses and uncaging cross-sections for one- and two-photon excitation. *Chemistry-A European Journal* **2006**, *12*, 6865-6879.

97. Klán, P.; Solomek, T.; Bochet, C. G.; Blanc, A.; Givens, R.; Rubina, M.; Popik, V.; Kostikov, A.; Wirz, J., Photoremovable protecting groups in chemistry and biology: reaction mechanisms and efficacy. *Chem Rev* **2013**, *113*, 119-91.

98. Il'ichev, Y. V.; Schwörer, M. A.; Wirz, J., Photochemical reaction mechanisms of 2-nitrobenzyl compounds: methyl ethers and caged ATP. *J Am Chem Soc* **2004**, *126*, 4581-95.

99. Givens, R. S.; Rubina, M.; Wirz, J., Applications of p-hydroxyphenacyl (pHP) and coumarin-4-ylmethyl photoremovable protecting groups. *Photochem Photobiol Sci* **2012**, *11*, 472-88.

100. Casey, J. P.; Blidner, R. A.; Monroe, W. T., Caged siRNAs for spatiotemporal control of gene silencing. *Mol Pharm* **2009**, *6*, 669-85.

101. Lusic, H.; Uprety, R.; Deiters, A., Improved synthesis of the two-photon caging group 3-nitro-2-ethyldibenzofuran and its application to a caged thymidine phosphoramidite. *Org Lett* **2010**, *12*, 916-9.

102. Brieke, C.; Rohrbach, F.; Gottschalk, A.; Mayer, G.; Heckel, A., Light-controlled tools. *Angew Chem Int Ed Engl* **2012**, *51*, 8446-76.

103. Hagen, V.; Bendig, J.; Frings, S.; Eckardt, T.; Helm, S.; Reuter, D.; Kaupp, U. B., Highly Efficient and Ultrafast Phototriggers for cAMP and cGMP by Using Long-Wavelength UV/Vis-Activation This work was supported by the Deutsche Forschungsgemeinschaft, the European Union, and the Fonds der Chemischen Industrie. We thank B. Dekowski and J. Loßmann for technical assistance and S. Hecht for proof reading. *Angew Chem Int Ed Engl* **2001**, *40*, 1045-1048.

104. Fisher, W. G.; Partridge, W. P.; Dees, C.; Wachter, E. A., Simultaneous two-photon activation of type-I photodynamic therapy agents. *Photochem Photobiol* **1997**, *66*, 141-55.

105. Bochet, C., Photolabile protecting groups and linkers. *J. Chem Soc-Per Trans 1* **2002**, 125-142.

106. Pelliccioli, A. P.; Wirz, J., Photoremovable protecting groups: reaction mechanisms and applications. *Photochem Photobiol Sci* **2002**, *1*, 441-58.

107. Tatsu, Y.; Nishigaki, T.; Darszon, A.; Yumoto, N., A caged sperm-activating peptide that has a photocleavable protecting group on the backbone amide. *Febs Letters* **2002**, *525*, 20-24.

108. Adams, S.; Kao, J.; Grynkiewicz, G.; Minta, A.; Tsien, R., Biologically useful chelators that release Ca^{2+} uopn illumination. *J. Am. Chem. Soc.* **1988**, *110*, 3212-3220.

109. Holmes, C. P., Model Studies for New o-Nitrobenzyl Photolabile Linkers: Substituent Effects on the Rates of Photochemical Cleavage. *J Org Chem* **1997**, *62*, 2370-2380.

110. Jockusch, S.; Li, Z.; Ju, J.; Turro, N., Two-photon excitation induced fluorescence of a trifluorophore-labeled DNA. *Photochem Photobiol* **2005**, *81*, 238-41.

111. Kaiser, W.; Garrett, C., 2- Photon excitation in CAF2 - EU2+. *Physical Review Letters* 1961, 229-&.

112. Zhu, Y.; Pavlos, C.; Toscano, J.; Dore, T., 8-Bromo-7-hydroxyquinoline as a photoremovable protecting group for physiological use: mechanism and scope. *J Am Chem Soc* **2006**, *128*, 4267-76.

113. So, P.; Dong, C.; Masters, B.; Berland, K., Two-photon excitation fluorescence microscopy. *Annu. Rev. Biomed. Eng.* **2000**, *2*, 399-429.

114. Zinter, J.; Levene, M., Maximizing fluorescence collection efficiency in multiphoton microscopy. *Optics Express* **2011**, *19*, 15348-15362.

115. Rubart, M., Two-photon microscopy of cells and tissue. *Cir Res* **2004**, *95*, 1154-1166.

116. Diaspro, A.; Bianchini, P.; Vicidomini, G.; Faretta, M.; Ramoino, P.; Usai, C., Multi-photon excitation microscopy. *Biomed. Eng. Online* **2006**, 5.

117. Furuta, T.; Wang, S.; Dantzker, J.; Dore, T.; Bybee, W.; Callaway, E.; Denk, W.; Tsien, R., Brominated 7-hydroxycoumarin-4-ylmethyls: Photolabile protecting groups with biologically useful cross-sections for two photon photolysis. *Proc. Natl. Acad. Sci U.S.A.* **1999**, *96*, 1193-1200.

118. Schmidt, R.; Geissler, D.; Hagen, V.; Bendig, J., Mechanism of photocleavage of (coumarin-4-yl)methyl esters. *J Phys Chem A* **2007**, *111*, 5768-74.

119. Schade, B.; Hagen, V.; Schmidt, R.; Herbrich, R.; Krause, E.; Eckardt, T.; Bendig, J., Deactivation behavior and excited-state properties of (coumarin-4-yl)methyl derivatives. 1. Photocleavage of (7-methoxycoumarin-4-yl)methyl-caged acids with fluorescence enhancement. *J.Org. Chem.* **1999**, *64*, 9109-9117.

120. Zhu, Y.; Pavlos, C. M.; Toscano, J. P.; Dore, T. M., 8-Bromo-7-hydroxyquinoline as a photoremovable protecting group for physiological use: mechanism and scope. *J Am Chem Soc* **2006**, *128*, 4267-76.

121. Patpi, S.; Pulipati, L.; Yogeeswari, P.; Sriram, D.; Jain, N.; Sridhar, B.; Murthy, R.; Devi, T.; Kalivendi, S.; Kantevari, S., Design, Synthesis, and Structure-Activity Correlations of Novel Dibenzo[b,d]furan, Dibenzo[b,d]thiophene, and N-Methylcarbazole Clubbed 1,2,3-Triazoles as Potent Inhibitors of Mycobacterium tuberculosis. *J Med Chem* **2012**, *55*, 3911-3922.

122. Kantevari, S.; Buskila, Y.; Ellis-Davies, G., Synthesis and characterization of cell-permeant 6-nitrodibenzofuranyl-caged IP3. *Photochem. & Photobiol Sci* **2012**, *11*, 508-513.

123. Papageorgiou, G.; Ogden, D.; Barth, A.; Corrie, J., Photorelease of carboxylic acids from 1-acyl-7-nitroindolines in aqueous solution: Rapid and efficient photorelease of L-glutamate. *J.Am. Chem. Soc.* **1999**, *121*, 6503-6504.

124. Warther, D.; Gug, S.; Specht, A.; Bolze, F.; Nicoud, J. F.; Mourot, A.; Goeldner, M., Two-photon uncaging: New prospects in neuroscience and cellular biology. *Bioorg Med Chem* **2010**, *18*, 7753-8.

125. Caruthers, M. H.; Beaucage, S. L.; Efcavitch, J. W.; Fisher, E. F.; Matteucci, M. D.; Stabinsky, Y., New chemical methods for synthesizing polynucleotides. *Nucleic Acids Symp Ser* **1980**, 215-23.

126. Caruthers, M. H.; Beaucage, S. L.; Becker, C.; Efcavitch, J. W.; Fisher, E. F.; Galluppi, G.; Goldman, R.; deHaseth, P.; Matteucci, M.; McBride, L., Deoxyoligonucleotide synthesis via the phosphoramidite method. *Gene Amplif Anal* **1983**, *3*, 1-26.

127. Tian, W.; Wang, Y., Mechanisms of Staudinger reactions within density functional theory. *J. Org. Chem.* **2004**, *69*, 4299-4308.

128. Lusic; Hrvoje. Control of biological processes with unnatural nucleosides, amino acids, and small molecule probes. Ph. D. Dissertation, North Carolina State University, Raleigh, **2009**.

129. Van Der Kelen, K.; Beyaert, R.; Inzé, D.; De Veylder, L., Translational control of eukaryotic gene expression. *Crit Rev Biochem Mol Biol* **2009**, *44*, 143-68.

130. Berckmans, B.; De Veylder, L., Transcriptional control of the cell cycle. *Curr Opin Plant Biol* **2009**, *12*, 599-605.

131. Pickart, M. A.; Klee, E. W.; Nielsen, A. L.; Sivasubbu, S.; Mendenhall, E. M.; Bill, B. R.; Chen, E.; Eckfeldt, C. E.; Knowlton, M.; Robu, M. E.; Larson, J. D.; Deng, Y.; Schimmenti, L. A.; Ellis, L. B.; Verfaillie, C. M.; Hammerschmidt, M.; Farber, S. A.; Ekker, S. C., Genome-wide reverse genetics framework to identify novel functions of the vertebrate secretome. *PLoS One* **2006**, *1*, e104.

132. Lawson, N. D.; Wolfe, S. A., Forward and reverse genetic approaches for the analysis of vertebrate development in the zebrafish. *Dev Cell* **2011**, *21*, 48-64.

133. Alonso, J. M.; Ecker, J. R., Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in Arabidopsis. *Nat Rev Genet* **2006**, *7*, 524-36.

134. Hardy, S.; Legagneux, V.; Audic, Y.; Paillard, L., Reverse genetics in eukaryotes. *Biol Cell* **2010**, *102*, 561-80.

135. Adams, D. J.; van der Weyden, L., Contemporary approaches for modifying the mouse genome. *Physiol Genomics* **2008**, *34*, 225-38.

136. Jain, P. K.; Shah, S.; Friedman, S. H., Patterning of gene expression using new photolabile groups applied to light activated RNAi. *J Am Chem Soc* **2011**, *133*, 440-6.

137. Govan, J. M.; Uprety, R.; Hemphill, J.; Lively, M. O.; Deiters, A., Regulation of transcription through light-activation and light-deactivation of triplex-forming oligonucleotides in mammalian cells. *ACS Chem Biol* **2012**, *7*, 1247-56.

138. Young, D.; Edwards, W.; Lusic, H.; Lively, M.; Deiters, A., Light-triggered polymerase chain reaction. *Chem. Comm.* **2008**, 462-464.

139. Freier, S. M.; Altmann, K. H., The ups and downs of nucleic acid duplex stability: structure-stability studies on chemically-modified DNA:RNA duplexes. *Nucl Acids Res* **1997**, *25*, 4429-43.

140. Mayer, G.; Müller, J.; Mack, T.; Freitag, D. F.; Höver, T.; Pötzsch, B.; Heckel, A., Differential regulation of protein subdomain activity with caged bivalent ligands. *Chembiochem* **2009**, *10*, 654-7.

141. Chaulk, S.; MacMillan, A., Caged RNA: photo-control of a ribozyme reaction. *Nucl Acids Res* 1998, 26, 3173-8.

142. Frank-Kamenetskii, M. D.; Mirkin, S. M., Triplex DNA structures. *Annu Rev Biochem* **1995**, *64*, 65-95.

143. Goñi, J. R.; Vaquerizas, J. M.; Dopazo, J.; Orozco, M., Exploring the reasons for the large density of triplex-forming oligonucleotide target sequences in the human regulatory regions. *BMC Genomics* **2006**, *7*, 63.

144. Mirkin, S. M.; Frank-Kamenetskii, M. D., H-DNA and related structures. *Annu Rev Biophys Biomol Struct* **1994**, *23*, 541-76.

145. Wang, G.; Vasquez, K. M., Naturally occurring H-DNA-forming sequences are mutagenic in mammalian cells. *Proc Natl Acad Sci U S A* **2004**, *101*, 13448-53.

146. Por, E.; Byun, H. J.; Lee, E. J.; Lim, J. H.; Jung, S. Y.; Park, I.; Kim, Y. M.; Jeoung, D. I.; Lee, H., The cancer/testis antigen CAGE with oncogenic potential stimulates cell proliferation by up-regulating cyclins D1 and E in an AP-1- and E2F-dependent manner. *J Biol Chem* **2010**, *285*, 14475-85.

147. Watts, J. K.; Deleavey, G. F.; Damha, M. J., Chemically modified siRNA: tools and applications. *Drug Discov Today* **2008**, *13*, 842-55.

148. Young, D. D.; Edwards, W. F.; Lusic, H.; Lively, M. O.; Deiters, A., Light-triggered polymerase chain reaction. *Chem Commun (Camb)* **2008**, 462-4.

149. Nguyen, Q. N.; Chavli, R. V.; Marques, J. T.; Conrad, P. G.; Wang, D.; He, W.; Belisle, B. E.; Zhang, A.; Pastor, L. M.; Witney, F. R.; Morris, M.; Heitz, F.; Divita, G.; Williams, B. R.; McMaster, G. K., Light controllable siRNAs regulate gene suppression and phenotypes in cells. *Biochim Biophys Acta* **2006**, *1758*, 394-403.

150. Lusic, H.; Lively, M. O.; Deiters, A., Light-activated deoxyguanosine: photochemical regulation of peroxidase activity. *Mol Biosyst* **2008**, *4*, 508-11.

151. Heckel, A.; Mayer, G., Light regulation of aptamer activity: an anti-thrombin aptamer with caged thymidine nucleobases. *J Am Chem Soc* **2005**, *127*, 822-3.

152. Knauert, M. P.; Glazer, P. M., Triplex forming oligonucleotides: sequence-specific tools for gene targeting. *Hum Mol Genet* **2001**, *10*, 2243-51.

153. Vasquez, K. M.; Narayanan, L.; Glazer, P. M., Specific mutations induced by triplex-forming oligonucleotides in mice. *Science* **2000**, *290*, 530-3.

154. Vasquez, K. M.; Glazer, P. M., Triplex-forming oligonucleotides: principles and applications. *Q Rev Biophys* **2002**, *35*, 89-107.

155. Simon, P.; Cannata, F.; Concordet, J. P.; Giovannangeli, C., Targeting DNA with triplex-forming oligonucleotides to modify gene sequence. *Biochimie* **2008**, *90*, 1109-16.

156. Wu, Q.; Gaddis, S. S.; MacLeod, M. C.; Walborg, E. F.; Thames, H. D.; DiGiovanni, J.; Vasquez, K. M., High-affinity triplex-forming oligonucleotide target sequences in mammalian genomes. *Mol Carcinog* **2007**, *46*, 15-23.

157. Jain, A.; Magistri, M.; Napoli, S.; Cabone, G. M.; Catapano, C. V., Mechanisms of triplex DNA-mediated inhibition of transcription initiation in cells. *Biochimie* **2010**, *92*, 317-320.

158. Schleifman, E.; Chin, J.; Glazer, P., Triplex-mediated gene modification. *Methods Mol Biol* **2008**, *435*, 175-90.

159. Jain, A.; Wang, G.; Vasquez, K. M., DNA triple helices: biological consequences and therapeutic potential. *Biochimie* **2008**, *90*, 1117-30.

160. Guntaka, R. V.; Varma, B. R.; Weber, K. T., Triplex-forming oligonucleotides as modulators of gene expression. *Int J Biochem Cell Biol* **2003**, *35*, 22-31.

161. Chin, J. Y.; Glazer, P. M., Repair of DNA lesions associated with triplex-forming oligonucleotides. *Mol Carcinog* **2009**, *48*, 389-99.

162. Diviacco, S.; Rapozzi, V.; Xodo, L.; Helene, C.; Quadrifoglio, F.; Giovannangeli, C., Site-directed inhibition of DNA replication by triple helix formation. *FASEB J* **2001**, 15, 2660-8.

163. Oussedik, K.; François, J. C.; Halby, L.; Senamaud-Beaufort, C.; Toutirais, G.; Dallavalle, S.; Pommier, Y.; Pisano, C.; Arimondo, P. B., Sequence-specific targeting of IGF-I and IGF-IR genes by camptothecins. *FASEB J* **2010**, *24*, 2235-44.

164. Liu, Y.; Nairn, R. S.; Vasquez, K. M., Processing of triplex-directed psoralen DNA interstrand crosslinks by recombination mechanisms. *Nucl Acids Res* **2008**, *36*, 4680-8.

165. Patterson, A.; Caprio, F.; Vallée-Bélisle, A.; Moscone, D.; Plaxco, K. W.; Palleschi, G.; Ricci, F., Using Triplex-Forming Oligonucleotide Probes for the Reagentless, Electrochemical Detection of Double-Stranded DNA. *Anal Chem* **2010**.

166. Heckel, A.; Buff, M. C.; Raddatz, M. S.; Müller, J.; Pötzsch, B.; Mayer, G., An anticoagulant with light-triggered antidote activity. *Angew Chem Int Ed Engl* **2006**, *45*, 6748-50.

167. Bochet, C., Wavelength-selective cleavage of photolabile protecting groups. *Tetrahedron Lett.* **2000**, *41*, 6341-6346.

168. Park, S.; Bando, T.; Shinohara, K.; Nishijima, S.; Sugiyama, H., Photocontrollable sequence-specific DNA alkylation by a pyrrole-imidazole polyamide seco-CBI conjugate. *Bioconjug Chem* **2011**, *22*, 120-4.

169. Lusic, H.; Young, D. D.; Lively, M. O.; Deiters, A., Photochemical DNA activation. *Org. Lett.* **2007**, *9*, 1903-1906.

170. Buchakjian, M. R.; Kornbluth, S., The engine driving the ship: metabolic steering of cell proliferation and death. *Nat Rev Mol Cell Biol* **2010**, *11*, 715-27.

171. Caldon, C. E.; Sutherland, R. L.; Musgrove, E., Cell cycle proteins in epithelial cell differentiation: implications for breast cancer. *Cell Cycle* **2010**, *9*, 1918-28.

172. Diehl, J. A., Cycling to cancer with cyclin D1. *Cancer Biol Ther* **2002**, *1*, 226-31.

173. Fu, M.; Wang, C.; Li, Z.; Sakamaki, T.; Pestell, R., Minireview: Cyclin D1: Normal and abnormal functions. *Endocrinology* **2004**, *145*, 5439-5447.

174. D'Amico, M.; Wu, K.; Fu, M.; Rao, M.; Albanese, C.; Russell, R.; Lian, H.; Bregman, D.; White, M.; Pestell, R., The inhibitor of cyclin-dependent kinase 4a/alternative reading frame (INK4a/ARF) locus encoded proteins p16(INK4a) and p19(ARF) repress cyclin D1 transcription through distinct cis elements. *Cancer Research* **2004**, *64*, 4122-4130.

175. Kim, J. K.; Diehl, J. A., Nuclear cyclin D1: an oncogenic driver in human cancer. *J Cell Physiol* **2009**, *220*, 292-6.

176. Tashiro, E.; Tsuchiya, A.; Imoto, M., Functions of cyclin D1 as an oncogene and regulation of cyclin D1 expression. *Cancer Sci* **2007**, *98*, 629-35.

177. Kim, H. G.; Miller, D. M., A novel triplex-forming oligonucleotide targeted to human cyclin D1 (bcl-1, proto-oncogene) promoter inhibits transcription in HeLa cells. *Biochemistry* **1998**, *37*, 2666-2672.

178. Shuai, X.; Han, G.; Wang, G., Effects of cyclin D1 antisense oligodeoxyneucleotides on the growth and expression of G1 phase regulators in gastric carcinoma cells. *J Huazhong Univ Sci Technolog Med Sci* **2003**, *23*, 396-8, 406.

179. Anderson, E.; Brown, T.; Picken, D., Novel photocleavable universal support for oligonucleotide synthesis. *Nucleos Nucleot Nucl* **2003**, *22*, 1403-6.

180. Höbartner, C.; Silverman, S. K., Modulation of RNA tertiary folding by incorporation of caged nucleotides. *Angew Chem Int Ed Engl* **2005**, *44*, 7305-9.

181. Schäfer, F.; Joshi, K. B.; Fichte, M. A.; Mack, T.; Wachtveitl, J.; Heckel, A., Wavelength-selective uncaging of dA and dC residues. *Org Lett* **2011**, *13*, 1450-3.

182. Young, D. D.; Edwards, W. F.; Lusic, H.; Lively, M. O.; Deiters, A., Light-triggered polymerase chain reaction. *Chem. Comm.* **2008**, 462-464.

183. Young, D. D.; Lusic, H.; Lively, M. O.; Yoder, J. A.; Deiters, A., Gene Silencing in Mammalian Cells with Light-Activated Antisense Agents. *Chembiochem* **2008**, *9*, 2937-2940.

184. Young, D.; Lively, M.; Deiters, A., Activation and deactivation of DNAzyme and antisense function with light for the photochemical regulation of gene expression in mammalian cells. *J Am Chem Soc* **2010**, *132*, 6183-93.

185. Abdelgany, A.; Wood, M.; Beeson, D., Hairpin DNAzymes: a new toot for efficient cellular gene silencing. *J Gen Med* **2007**, *9*, 727-738.

186. Rinn, J. L.; Chang, H. Y., Genome regulation by long noncoding RNAs. *Annu Rev Biochem* **2012**, *81*, 145-66.

187. Bennett, C. F.; Swayze, E. E., RNA targeting therapeutics: molecular mechanisms of antisense oligonucleotides as a therapeutic platform. *Annu Rev Pharmacol Toxicol* **2010**, *50*, 259-93.

188. Castanotto, D.; Rossi, J. J., The promises and pitfalls of RNA-interference-based therapeutics. *Nature* **2009**, *457*, 426-33.

189. Sun, G.; Rossi, J. J., Problems associated with reporter assays in RNAi studies. *RNA Biol* **2009**, *6*, 406-11.

190. Laufer, S. D.; Restle, T., Peptide-mediated cellular delivery of oligonucleotide-based therapeutics in vitro: quantitative evaluation of overall efficacy employing easy to handle reporter systems. *Curr Pharm Des* **2008**, *14*, 3637-55.

191. Koppelhus, U.; Nielsen, P. E., Cellular delivery of peptide nucleic acid (PNA). *Adv Drug Deliv Rev* **2003**, *55*, 267-80.

192. Akinc, A.; Thomas, M.; Klibanov, A. M.; Langer, R., Exploring polyethylenimine-mediated DNA transfection and the proton sponge hypothesis. *J Gene Med* **2005**, *7*, 657-63.

193. Stewart, K. M.; Horton, K. L.; Kelley, S. O., Cell-penetrating peptides as delivery vehicles for biology and medicine. *Org Biomol Chem* **2008**, *6*, 2242-55.

194. van den Berg, A.; Dowdy, S. F., Protein transduction domain delivery of therapeutic macromolecules. *Curr Opin Biotechnol* **2011**, *22*, 888-93.

195. Juliano, R. L.; Ming, X.; Nakagawa, O., The chemistry and biology of oligonucleotide conjugates. *Acc Chem Res* **2012**, *45*, 1067-76.

196. Guterstam, P.; Madani, F.; Hirose, H.; Takeuchi, T.; Futaki, S.; El Andaloussi, S.; Gräslund, A.; Langel, U., Elucidating cell-penetrating peptide mechanisms of action for membrane interaction, cellular uptake, and translocation utilizing the hydrophobic counter-anion pyrenebutyrate. *Biochim Biophys Acta* **2009**, *1788*, 2509-17.

197. Wolfrum, C.; Shi, S.; Jayaprakash, K. N.; Jayaraman, M.; Wang, G.; Pandey, R. K.; Rajeev, K. G.; Nakayama, T.; Charrise, K.; Ndungo, E. M.; Zimmermann, T.; Koteliansky, V.; Manoharan, M.; Stoffel, M., Mechanisms and optimization of in vivo delivery of lipophilic siRNAs. *Nature Biotech.* **2007**, **25**, 1149-1157.

198. Juliano, R.; Alam, M. R.; Dixit, V.; Kang, H., Mechanisms and strategies for effective delivery of antisense and siRNA oligonucleotides. *Nucleic Acids Res* 2008, 36, 4158-71.

199. Heitz, F.; Morris, M. C.; Divita, G., Twenty years of cell-penetrating peptides: from molecular mechanisms to therapeutics. *Br J Pharmacol* **2009**, *157*, 195-206.

200. Milletti, F., Cell-penetrating peptides: classes, origin, and current landscape. *Drug Discov Today* **2012**.

201. Schmidt, N.; Mishra, A.; Lai, G. H.; Wong, G. C., Arginine-rich cell-penetrating peptides. *FEBS Lett* **2010**, *584*, 1806-13.

202. Mann, A.; Thakur, G.; Shukla, V.; Singh, A. K.; Khanduri, R.; Naik, R.; Jiang, Y.; Kalra, N.; Dwarakanath, B. S.; Langel, U.; Ganguli, M., Differences in DNA condensation and release by lysine and arginine homopeptides govern their DNA delivery efficiencies. *Mol Pharm* **2011**, *8*, 1729-41.

203. Turner, J. J.; Jones, S.; Fabani, M. M.; Ivanova, G.; Arzumanov, A. A.; Gait, M. J., RNA targeting with peptide conjugates of oligonucleotides, siRNA and PNA. *Blood Cells Mol Dis* **2007**, *38*, 1-7.

204. El-Andaloussi, S.; Johansson, H. J.; Holm, T.; Langel, U., A novel cell-penetrating peptide, M918, for efficient delivery of proteins and peptide nucleic acids. *Mol Ther* **2007**, *15*, 1820-6.

205. Chiu, Y. L.; Ali, A.; Chu, C. Y.; Cao, H.; Rana, T. M., Visualizing a correlation between siRNA localization, cellular uptake, and RNAi in living cells. *Chem. Biol.* **2004**, *11*, 1165-1175.

206. Vivès, E.; Brodin, P.; Lebleu, B., A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. *J Biol Chem* **1997**, *272*, 16010-7.

207. Kamal, A.; Prabhakar, S.; Janaki Ramaiah, M.; Venkat Reddy, P.; Ratna Reddy, C.; Mallareddy, A.; Shankaraiah, N.; Lakshmi Narayan Reddy, T.; Pushpavalli, S. N.; Pal-Bhadra, M., Synthesis and anticancer activity of chalcone-pyrrolobenzodiazepine conjugates linked via 1,2,3-triazole ring side-armed with alkane spacers. *Eur J Med Chem* **2011**, *46*, 3820-31.

208. Govan, J. M.; Lively, M. O.; Deiters, A., Photochemical control of DNA decoy function enables precise regulation of nuclear factor κB activity. *J Am Chem Soc* **2011**, *133*, 13176-82.

209. Agrawal, S., Importance of nucleotide sequence and chemical modifications of antisense oligonucleotides. *Biochim Biophys Acta* **1999**, *1489*, 53-68.

210. Brown, D. A.; Kang, S. H.; Gryaznov, S. M.; DeDionisio, L.; Heidenreich, O.; Sullivan, S.; Xu, X.; Nerenberg, M. I., Effect of phosphorothioate modification of oligodeoxynucleotides on specific protein binding. *J Biol Chem* **1994**, *269*, 26801-5.

211. Lallana, E.; Riguera, R.; Fernandez-Megia, E., Reliable and efficient procedures for the conjugation of biomolecules through Huisgen azide-alkyne cycloadditions. *Angew Chem Int Ed Engl* **2011**, *50*, 8794-804.

212. Hong, V.; Presolski, S. I.; Ma, C.; Finn, M. G., Analysis and optimization of copper-catalyzed azide-alkyne cycloaddition for bioconjugation. *Angew Chem Int Ed Engl* **2009**, *48*, 9879-83.

213. Presolski, S. I.; Hong, V.; Cho, S. H.; Finn, M. G., Tailored ligand acceleration of the Cu-catalyzed azide-alkyne cycloaddition reaction: practical and mechanistic implications. *J Am Chem Soc* **2010**, *132*, 14570-6.

214. Shelbourne, M.; Brown, T.; El-Sagheer, A. H., Fast and efficient DNA crosslinking and multiple orthogonal labelling by copper-free click chemistry. *Chem Commun (Camb)* **2012**, *48*, 11184-6.

215. Kota, S. K.; Balasubramanian, S., Cancer therapy via modulation of micro RNA levels: a promising future. *Drug Discov Today* **2010**, *15*, 733-40.

216. Winter, J.; Jung, S.; Keller, S.; Gregory, R. I.; Diederichs, S., Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat Cell Biol* **2009**, *11*, 228-34.

217. Ji, X., The mechanism of RNase III action: how dicer dices. *Curr Top Microbiol Immunol* **2008**, *320*, 99-116.

218. Chatterjee, S.; Pal, J. K., Role of 5'- and 3'-untranslated regions of mRNAs in human diseases. *Biol Cell* **2009**, *101*, 251-62.

219. Didiano, D.; Hobert, O., Molecular architecture of a miRNA-regulated 3' UTR. *RNA* **2008**, *14*, 1297-317.

220. Sharp, P. A., The centrality of RNA. *Cell* **2009**, *136*, 577-80.

221. Janga, S. C.; Vallabhaneni, S., MicroRNAs as post-transcriptional machines and their interplay with cellular networks. *Adv Exp Med Biol* **2011**, *722*, 59-74.

222. Dumelin, C. E.; Chen, Y.; Leconte, A. M.; Chen, Y. G.; Liu, D. R., Discovery and biological characterization of geranylated RNA in bacteria. *Nat Chem Biol* **2012**, *8*, 913-9.

223. Wang, J.; Lu, M.; Qiu, C.; Cui, Q., TransmiR: a transcription factor-microRNA regulation database. *Nucl Acids Res* **2010**, *38*, D119-22.

224. van Rooij, E.; Olson, E. N., MicroRNA therapeutics for cardiovascular disease: opportunities and obstacles. *Nat Rev Drug Discov* **2012**, *11*, 860-72.

225. Andorfer, C. A.; Necela, B. M.; Thompson, E. A.; Perez, E. A., MicroRNA signatures: clinical biomarkers for the diagnosis and treatment of breast cancer. *Trends Mol Med* **2011**, *17*, 313-9.

226. Boeri, M.; Verri, C.; Conte, D.; Roz, L.; Modena, P.; Facchinetti, F.; Calabrò, E.; Croce, C. M.; Pastorino, U.; Sozzi, G., MicroRNA signatures in tissues and plasma predict development and prognosis of computed tomography detected lung cancer. *Proc Natl Acad Sci U S A* **2011**, *108*, 3713-8.

227. Seila, A. C.; Sharp, P. A., Small RNAs tell big stories in Whistler. *Nat Cell Biol* **2008**, *10*, 630-3.

228. Henke, J. I.; Goergen, D.; Zheng, J.; Song, Y.; Schüttler, C. G.; Fehr, C.; Jünemann, C.; Niepmann, M., microRNA-122 stimulates translation of hepatitis C virus RNA. *EMBO J* **2008**, *27*, 3300-10.

229. Kozomara, A.; Griffiths-Jones, S., miRBase: integrating microRNA annotation and deep-sequencing data. *Nucl Acids Res* 2011, 39, D152-7.

230. Deiters, A., Small molecule modifiers of the microRNA and RNA interference pathway. *AAPS J* **2010**, *12*, 51-60.

231. Gambari, R.; Fabbri, E.; Borgatti, M.; Lampronti, I.; Finotti, A.; Brognara, E.; Bianchi, N.; Manicardi, A.; Marchelli, R.; Corradini, R., Targeting microRNAs involved in human diseases: a novel approach for modification of gene expression and drug development. *Biochem Pharmacol* **2011**, *82*, 1416-29.

232. Ragan, C.; Zuker, M.; Ragan, M. A., Quantitative prediction of miRNA-mRNA interaction based on equilibrium concentrations. *PLoS Comput Biol* **2011**, *7*, e1001090.

233. Bail, S.; Swerdel, M.; Liu, H.; Jiao, X.; Goff, L. A.; Hart, R. P.; Kiledjian, M., Differential regulation of microRNA stability. *RNA* **2010**, *16*, 1032-9.

234. Gantier, M. P.; McCoy, C. E.; Rusinova, I.; Saulep, D.; Wang, D.; Xu, D.; Irving, A. T.; Behlke, M. A.; Hertzog, P. J.; Mackay, F.; Williams, B. R., Analysis of microRNA turnover in mammalian cells following Dicer1 ablation. *Nucl Acids Res* **2011**, *39*, 5692-703.

235. Mueller, G. A.; Miller, M. T.; Derose, E. F.; Ghosh, M.; London, R. E.; Hall, T. M., Solution structure of the Drosha double-stranded RNA-binding domain. *Silence* **2010**, *1*, 2.

236. Sohn, S. Y.; Bae, W. J.; Kim, J. J.; Yeom, K. H.; Kim, V. N.; Cho, Y., Crystal structure of human DGCR8 core. *Nat Struct Mol Biol* **2007**, *14*, 847-53.

237. Köhler, A.; Hurt, E., Exporting RNA from the nucleus to the cytoplasm. *Nat Rev Mol Cell Biol* **2007**, *8*, 761-73.

238. Macrae, I. J.; Zhou, K.; Li, F.; Repic, A.; Brooks, A. N.; Cande, W. Z.; Adams, P. D.; Doudna, J. A., Structural basis for double-stranded RNA processing by Dicer. *Science* **2006**, *311*, 195-8.

239. Macrae, I. J.; Li, F.; Zhou, K.; Cande, W. Z.; Doudna, J. A., Structure of Dicer and mechanistic implications for RNAi. *Cold Spring Harb Symp Quant Biol* **2006**, *71*, 73-80.

240. Chan, S. P.; Slack, F. J., microRNA-mediated silencing inside P-bodies. *RNA Biol* **2006**, 3, 97-100.

241. Garzon, R.; Marcucci, G.; Croce, C. M., Targeting microRNAs in cancer: rationale, strategies and challenges. *Nat Rev Drug Discov* **2010**, *9*, 775-89.

242. Shi, X. B.; Tepper, C. G.; deVere White, R. W., Cancerous miRNAs and their regulation. *Cell Cycle* **2008**, *7*, 1529-38.

243. Dangwal, S.; Bang, C.; Thum, T., Novel techniques and targets in cardiovascular microRNA research. *Cardiovasc Res* **2012**, *93*, 545-54.

244. Bang, C.; Fiedler, J.; Thum, T., Cardiovascular importance of the microRNA-23/27/24 family. *Microcirculation* **2012**, *19*, 208-14.

245. Dangwal, S.; Hartmann, D.; Thum, T., MicroRNAs deciding cardiac stem cell fate. *J Mol Cell Cardiol* **2012**, *53*, 747-8.

246. Pottier, N.; Maurin, T.; Chevalier, B.; Puisségur, M. P.; Lebrigand, K.; Robbe-Sermesant, K.; Bertero, T.; Lino Cardenas, C. L.; Courcot, E.; Rios, G.; Fourre, S.; Lo-Guidice, J. M.; Marcet, B.; Cardinaud, B.; Barbry, P.; Mari, B., Identification of keratinocyte growth factor as a target of microRNA-155 in lung fibroblasts: implication in epithelial-mesenchymal interactions. *PLoS One* **2009**, *4*, e6718.

247. Mendell, J. T.; Olson, E. N., MicroRNAs in stress signaling and human disease. *Cell* **2012**, *148*, 1172-87.

248. Hata, A.; Davis, B. N., Control of microRNA biogenesis by TGFbeta signaling pathway-A novel role of Smads in the nucleus. *Cytokine Growth Factor Rev* **2009**, *20*, 517-21.

249. Dinger, M. E.; Mercer, T. R.; Mattick, J. S., RNAs as extracellular signaling molecules. *J Mol Endocrinol* **2008**, *40*, 151-9.

250. Schmid, C. A.; Craig, V. J.; Müller, A.; Flori, M., The Role of microRNAs in the Pathogenesis and Treatment of Hematopoietic Malignancies. *Curr Pharm Des* **2013**, *19*, 1201-10.

251. Yin, Q.; Wang, X.; McBride, J.; Fewell, C.; Flemington, E., B-cell receptor activation induces BIC/miR-155 expression through a conserved AP-1 element. *J Biol Chem* **2008**, *283*, 2654-62.

252. Leng, R. X.; Pan, H. F.; Qin, W. Z.; Chen, G. M.; Ye, D. Q., Role of microRNA-155 in autoimmunity. *Cytokine Growth Factor Rev* **2011**, *22*, 141-7.

253. He, M.; Xu, Z.; Ding, T.; Kuang, D. M.; Zheng, L., MicroRNA-155 regulates inflammatory cytokine production in tumor-associated macrophages via targeting C/EBPbeta. *Cell Mol Immunol* **2009**, *6*, 343-52.

254. O'Connell, R. M.; Taganov, K. D.; Boldin, M. P.; Cheng, G.; Baltimore, D., MicroRNA-155 is induced during the macrophage inflammatory response. *Proc Natl Acad Sci U S A* **2007**, *104*, 1604-9.

255. Rottiers, V.; Näär, A. M., MicroRNAs in metabolism and metabolic disorders. *Nat Rev Mol Cell Biol* **2012**, *13*, 239-50.

256. Jangra, R. K.; Yi, M.; Lemon, S. M., Regulation of hepatitis C virus translation and infectious virus production by the microRNA miR-122. *J Virol* **2010**, *84*, 6615-25.

257. Lee, C. H.; Kim, J. H.; Kim, H. W.; Myung, H.; Lee, S. W., Hepatitis C virus replication-specific inhibition of microRNA activity with self-cleavable allosteric ribozyme. *Nucleic Acid Ther* **2012**, *22*, 17-29.

258. Jopling, C.; Yi, M.; Lancaster, A.; Lemon, S.; Sarnow, P., Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science* **2005**, *309*, 1577-81.

259. Esau, C.; Davis, S.; Murray, S.; Yu, X.; Pandey, S.; Pear, M.; Watts, L.; Booten, S.; Graham, M.; McKay, R.; Subramaniam, A.; Propp, S.; Lollo, B.; Freier, S.; Bennett, C.; Bhanot, S.; Monia, B., miR-122 regulation of lipid metabolism revealed by in vivo antisense targeting. *Cell Metab* **2006**, *3*, 87-98.

260. Lanford, R.; Hildebrandt-Eriksen, E.; Petri, A.; Persson, R.; Lindow, M.; Munk, M.; Kauppinen, S.; Ørum, H., Therapeutic silencing of microRNA-122 in primates with chronic hepatitis C virus infection. *Science* **2010**, *327*, 198-201.

261. Tong, A.; Nemunaitis, J., Modulation of miRNA activity in human cancer: a new paradigm for cancer gene therapy? *Cancer Gene Ther* **2008**, *15*, 341-55.

262. Medina, P. P.; Nolde, M.; Slack, F. J., OncomiR addiction in an in vivo model of microRNA-21-induced pre-B-cell lymphoma. *Nature* **2010**, *467*, 86-90.

263. Furusawa, K.; Ueno, K.; Katsura, T., Synthesis and restricted conformation of 3',5'-O-(di-tert-Butylsilanediyl)ribonucleosides. *Chem. Lett.* **1990**, 97-100.

264. Lusic, H.; Deiters, A., A new photocaging group for aromatic N-heterocycles. *Synthesis* **2006**, *8*, 2147-2150.

265. Akiyama, T.; Nishimoto, H.; Ozaki, S., he selective protectionof uridine with a para-methoxybenzyl chloride- A Synthesis of 2'-O-Methyluridine. *Bull.Chem. Soc.Jpn* **1990**, *63*, 3356-3357.

266. Gumireddy, K.; Young, D. D.; Xiong, X.; Hogenesch, J. B.; Huang, Q.; Deiters, A., Small-molecule inhibitors of microrna miR-21 function. *Angew Chem Int Ed Engl* **2008**, *47*, 7482-4.

267. Denny, W. A.; Leupin, W.; Kearns, D. R., Simplified Liquid-Phase Preparation of Four Decadeoxyribonucleotides and their Preliminary Spectroscopic Characterization. *Helv Chem Acta* **1982**, *65*, 2372-2393.

268. Beharry, A. A.; Woolley, G. A., Azobenzene photoswitches for biomolecules. *Chem Soc Rev* **2011**, *40*, 4422-37.

269. Bandara, H. M.; Burdette, S. C., Photoisomerization in different classes of azobenzene. *Chem Soc Rev* **2012**, *41*, 1809-25.

270. Yager, K.; Barrett, C., Novel photo-switching using azobenzene functional materials. *J. Photochem. Photobiol. A* **2006**, *182*, 250-261.

271. Schultz, T.; Quenneville, J.; Levine, B.; Toniolo, A.; Martinez, T.; Lochbrunner, S.; Schmitt, M.; Shaffer, J.; Zgierski, M.; Stolow, A., Mechanism and dynamics of azobenzene photoisomerization. *J.Am. Chem. Soc.* **2003**, *125*, 8098-8099.

272. Mahimwalla, Z.; Yager, K.; Mamiya, J.; Shishido, A.; Priimagi, A.; Barrett, C., Azobenzene photomechanics: prospects and potential applications. *Polymer Bulletin* **2012**, *69*, 967-1006.

273. Willner, I.; Rubin, S., Control of the structure and functions of biomaterials by light. *Angew Chem Int Ed Engl* **1996**, *35*, 367-385.

274. Renner, C.; Moroder, L., Azobenzene as conformational switch in model peptides. *Chembiochem* **2006**, *7*, 868-78.

275. Yager, K.; Tanchak, O.; Godbout, C.; Fritzsche, H.; Barrett, C., Photomechanical effects in azo-polymers studied by neutron reflectometry. *Macromolecules* **2006**, *39*, 9311-9319.

276. Renner, C.; Moroder, L., Azobenzene as conformational switch in model peptides. *Chembiochem* **2006**, *7*, 868-78.

277. Fan, H. Y.; Morgan, S. A.; Brechun, K. E.; Chen, Y. Y.; Jaikaran, A. S.; Woolley, G. A., Improving a designed photocontrolled DNA-binding protein. *Biochemistry* **2011**, 50, 1226-37.

278. Balasubramanian, D.; Subramani, S.; Kumar, C., Modification of a model membrane structure by embedded photochrome. *Nature* **1975**, *254*, 252-4.

279. Zeyat, G.; Rück-Braun, K., Building photoswitchable 3,4'-AMPB peptides: Probing chemical ligation methods with reducible azobenzene thioesters. *Beilstein J Org Chem* **2012**, *8*, 890-6.

280. Kusebauch, U.; Cadamuro, S. A.; Musiol, H. J.; Lenz, M. O.; Wachtveitl, J.; Moroder, L.; Renner, C., Photocontrolled folding and unfolding of a collagen triple helix. *Angew Chem Int Ed Engl* **2006**, *45*, 7015-8.

281. Beharry, A. A.; Wong, L.; Tropepe, V.; Woolley, G. A., Fluorescence imaging of azobenzene photoswitching in vivo. *Angew Chem Int Ed Engl* **2011**, *50*, 1325-7.

282. Yang, Y. Y.; Grammel, M.; Raghavan, A. S.; Charron, G.; Hang, H. C., Comparative analysis of cleavable azobenzene-based affinity tags for bioorthogonal chemical proteomics. *Chem Biol* **2010**, *17*, 1212-22.

283. Kuzuya, A.; Tanaka, K.; Komiyama, M., Photoswitching of site-selective RNA scission by sequential incorporation of azobenzene and acridine residues in a DNA oligomer. *J Nucleic Acids* **2011**, *2011*, 162452.

284. Asanuma, H.; Liang, X.; Nishioka, H.; Matsunaga, D.; Liu, M.; Komiyama, M., Synthesis of azobenzene-tethered DNA for reversible photo-regulation of DNA functions: hybridization and transcription. *Nat Protoc* **2007**, *2*, 203-12.

285. Ito, H.; Liang, X.; Nishioka, H.; Asanuma, H., Construction of photoresponsive RNA for photoswitching RNA hybridization. *Org Biomol Chem* **2010**, *8*, 5519-24.

286. Phillips, J. A.; Liu, H.; O'Donoghue, M. B.; Xiong, X.; Wang, R.; You, M.; Sefah, K.; Tan, W., Using azobenzene incorporated DNA aptamers to probe molecular binding interactions. *Bioconjug Chem* **2011**, *22*, 282-8.

287. Liang, X.; Wakuda, R.; Fujioka, K.; Asanuma, H., Photoregulation of DNA transcription by using photoresponsive T7 promoters and clarification of its mechanism. *FEBS J* **2010**, *277*, 1551-61.

288. Zinchenko, A. A.; Tanahashi, M.; Murata, S., Photochemical modulation of DNA conformation by organic dications. *Chembiochem* **2012**, *13*, 105-11.

289. Liang, X.; Yoshida, T.; Asanuma, H.; Komiyama, M., Photo-regulation of DNA/RNA duplex formation by azobenzene-tethered DNA towards antisense strategy. *Nucleic Acids Symp Ser* **2000**, 277-8.

290. Itoh, H.; Nishioka, H.; Liang, X.; Asanu, H., Development of photoresponsive RNA towards photoswitching of RNA functions. *Nucleic Acids Symp Ser (Oxf)* **2007**, 171-2.

291. Liu, M.; Asanuma, H.; Komiyama, M., Real time monitoring of the interaction of T7 RNA polymerase with azobenzene-tethered T7 promoter by biosensor. *Nucleic Acids Symp Ser (Oxf)* **2004**, 221-2.

292. Zhao, J.; Zhou, J.; Liu, M.; Asanuma, H.; Komiyama, M., Photoregulation of in vitro transcription/translation of GFP by tethering an azobenzene to T7 promoter. *Nucleic Acids Symp Ser (Oxf)* **2004**, 199-200.

293. Zhou, M.; Liang, X.; Mochizuki, T.; Asanuma, H., A light-driven DNA nanomachine for the efficient photoswitching of RNA digestion. *Angew Chem Int Ed Engl* **2010**, *49*, 2167-70.

294. Matsunaga, D.; Asanuma, H.; Komiyama, M., Photoregulation of RNA digestion by RNase H with azobenzene-tethered DNA. *J Am Chem Soc* **2004**, *126*, 11452-3.

295. Wu, L.; Koumoto, K.; Sugimoto, N., Reversible stability switching of a hairpin DNA via a photo-responsive linker unit. *Chemical Communications* **2009**, 1915-1917.

296. Steller, K.; Letsinge.RL, Photoinduced substitution.. Effects of distant substituents on photoinduced aromatic substitution reactions.. *J. Org. Chem.* **1970**, *35*, 308

297. Wang, L.; Yang, F.; Yang, X.; Guan, X.; Hu, C.; Liu, T.; He, Q.; Yang, B.; Hu, Y., Synthesis and biological evaluation of new 4 beta-anilino-4 '-O-demethyl-4-desoxypodophyllotoxin derivatives as potential antitumor agents. *Eur J Med. Chem.* **2011**, *46*, 285-296.

298. Borstel; Reid, v. Prodrugs activated by targeted catalytic proteins. 2001.

299. Schwarzer, D., Hacking the genetic code of mammalian cells. *Chembiochem* **2009**, *10*, 1602-4.

300. Nudelman, I.; Glikin, D.; Smolkin, B.; Hainrichson, M.; Belakhov, V.; Baasov, T., Repairing faulty genes by aminoglycosides: development of new derivatives of geneticin (G418) with enhanced suppression of diseases-causing nonsense mutations. *Bioorg Med Chem* **2010**, *18*, 3735-46.

301. Liao, J., Protein and cellular engineering with unnatural amino acids. *Biotechnol Prog* **2007**, *23*, 28-31.

302. Dill, K. A.; MacCallum, J. L., The protein-folding problem, 50 years on. *Science* **2012**, *338*, 1042-6.

303. Moreira, I. S.; Fernandes, P. A.; Ramos, M. J., Hot spots--a review of the protein-protein interface determinant amino-acid residues. *Proteins* **2007**, *68*, 803-12.

304. Rhodes, D. R.; Tomlins, S. A.; Varambally, S.; Mahavisno, V.; Barrette, T.; Kalyana-Sundaram, S.; Ghosh, D.; Pandey, A.; Chinnaiyan, A. M., Probabilistic model of the human protein-protein interaction network. *Nat Biotechnol* **2005**, *23*, 951-9.

305. Chautard, E.; Thierry-Mieg, N.; Ricard-Blum, S., Interaction networks: from protein functions to drug discovery. A review. *Pathol Biol (Paris)* **2009**, *57*, 324-33.

306. Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S., The protein kinase complement of the human genome. *Science* **2002**, *298*, 1912-34.

307. Anamika, K.; Garnier, N.; Srinivasan, N., Functional diversity of human protein kinase splice variants marks significant expansion of human kinome. *BMC Genomics* **2009**, *10*, 622.

308. Deshmukh, K.; Anamika, K.; Srinivasan, N., Evolution of domain combinations in protein kinases and its implications for functional diversity. *Prog Biophys Mol Biol* 2010, 102, 1-15.

309. Huang, Z.; Jiang, H.; Liu, X.; Chen, Y.; Wong, J.; Wang, Q.; Huang, W.; Shi, T.; Zhang, J., HEMD: an integrated tool of human epigenetic enzymes and chemical modulators for therapeutics. *PLoS One* **2012**, *7*, e39917.

310. Wu, Y. W.; Goody, R. S., Probing protein function by chemical modification. *J Pept Sci* **2010**, *16*, 514-23.

311. Kalesh, K. A.; Shi, H.; Ge, J.; Yao, S. Q., The use of click chemistry in the emerging field of catalomics. *Org Biomol Chem* **2010**, *8*, 1749-62.

312. Sunbul, M.; Yin, J., Site specific protein labeling by enzymatic posttranslational modification. *Org Biomol Chem* **2009**, *7*, 3361-71.

313. Lin, M. Z.; Wang, L., Selective labeling of proteins with chemical probes in living cells. *Physiology (Bethesda)* **2008**, *23*, 131-41.

314. Sletten, E. M.; Bertozzi, C. R., Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angew Chem Int Ed Engl* **2009**, *48*, 6974-98.

315. Jones, D.; Cellitti, S.; Hao, X.; Zhang, Q.; Jahnz, M.; Summerer, D.; Schultz, P.; Uno, T.; Geierstanger, B., Site-specific labeling of proteins with NMR-active unnatural amino acids. *J Biomol NMR* **2010**, *46*, 89-100.

316. Beene, D.; Dougherty, D.; Lester, H., Unnatural amino acid mutagenesis in mapping ion channel function. *Curr Opin Neurobiol* **2003**, *13*, 264-70.

317. Johnson, J. A.; Lu, Y. Y.; Van Deventer, J. A.; Tirrell, D. A., Residue-specific incorporation of non-canonical amino acids into proteins: recent developments and applications. *Curr Opin Chem Biol* **2010**, *14*, 774-80.

318. Bacher, J. M.; Bull, J. J.; Ellington, A. D., Evolution of phage with chemically ambiguous proteomes. *BMC Evol Biol* **2003**, *3*, 24.

319. Singh-Blom, A.; Hughes, R. A.; Ellington, A. D., Residue-specific incorporation of unnatural amino acids into proteins in vitro and in vivo. *Methods Mol Biol* **2013**, *978*, 93-114.

320. Tcherkezian, J.; Brittis, P.; Thomas, F.; Roux, P.; Flanagan, J., Transmembrane Receptor DCC Associates with Protein Synthesis Machinery and Regulates Translation. *Cell* **2010**, *141*, 632-644.

321. Ngo, J. T.; Tirrell, D. A., Noncanonical amino acids in the interrogation of cellular protein synthesis. *Acc Chem Res* **2011**, *44*, 677-85.

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

323. Yonemoto, I. T.; Tippmann, E. M., The juggernauts of biology. *Bioessays* **2010**, *32*, 314-21.

324. Ngo, J. T.; Babin, B. M.; Champion, J. A.; Schuman, E. M.; Tirrell, D. A., State-selective metabolic labeling of cellular proteins. *ACS Chem Biol* **2012**, *7*, 1326-30.

325. Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R., Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc Natl Acad Sci U S A* **2002**, *99*, 19-24.

326. England, P. M., Unnatural amino acid mutagenesis: a precise tool for probing protein structure and function. *Biochemistry* **2004**, *43*, 11623-9.

327. Tsuji, F. I.; Inouye, S.; Goto, T.; Sakaki, Y., Site-specific mutagenesis of the calcium-binding photoprotein aequorin. *Proc Natl Acad Sci U S A* **1986**, *83*, 8107-11.

328. Wang, L.; Brock, A.; Herberich, B.; Schultz, P., Expanding the genetic code of Escherichia coli. *Science* **2001**, *292*, 498-500.

329. Kawakami, T.; Murakami, H., Genetically encoded libraries of nonstandard peptides. *J Nucleic Acids* **2012**, *2012*, 713510.

330. Wang, L.; Schultz, P., A general approach for the generation of orthogonal tRNAs. *Chem Biol* **2001**, *8*, 883-90.

331. Young, T.; Schultz, P., Beyond the canonical 20 amino acids: expanding the genetic lexicon. *J Biol Chem* **2010**, *285*, 11039-44.

332. Voloshchuk, N.; Montclare, J., Incorporation of unnatural amino acids for synthetic biology. *Mol Biosyst* **2010**, *6*, 65-80.

333. Wang, L.; Schultz, P., Expanding the genetic code. *Chem. Comm.* **2002**, *1*-11.

334. Deiters, A.; Groff, D.; Ryu, Y.; Xie, J.; Schultz, P., A genetically encoded photocaged tyrosine. *Angew Chem Int Ed Engl* **2006**, *45*, 2728-31.

335. Iraha, F.; Oki, K.; Kobayashi, T.; Ohno, S.; Yokogawa, T.; Nishikawa, K.; Yokoyama, S.; Sakamoto, K., Functional replacement of the endogenous tyrosyl-tRNA synthetase-tRNATyr pair by the archaeal tyrosine pair in Escherichia coli for genetic code expansion. *Nucl Acids Res* **2010**, *38*, 3682-91.

336. Davis, L.; Chin, J. W., Designer proteins: applications of genetic code expansion in cell biology. *Nat Rev Mol Cell Biol* **2012**, *13*, 168-82.

337. Lang, K.; Davis, L.; Torres-Kolbus, J.; Chou, C.; Deiters, A.; Chin, J. W., Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction. *Nat Chem* **2012**, *4*, 298-304.

338. Lang, K.; Davis, L.; Wallace, S.; Mahesh, M.; Cox, D. J.; Blackman, M. L.; Fox, J. M.; Chin, J. W., Genetic Encoding of bicyclononynes and trans-cyclooctenes for site-specific protein labeling in vitro and in live mammalian cells via rapid fluorogenic Diels-Alder reactions. *J Am Chem Soc* **2012**, *134*, 10317-20.

339. Daggett, K. A.; Sakmar, T. P., Site-specific in vitro and in vivo incorporation of molecular probes to study G-protein-coupled receptors. *Curr Opin Chem Biol* **2011**, *15*, 392-8.

340. Chang, Y. F.; Imam, J. S.; Wilkinson, M. F., The nonsense-mediated decay RNA surveillance pathway. *Annu Rev Biochem* **2007**, *76*, 51-74.

341. Liu, W.; Brock, A.; Chen, S.; Schultz, P. G., Genetic incorporation of unnatural amino acids into proteins in mammalian cells. *Nat Methods* **2007**, *4*, 239-44.

342. Chin, J. W., Reprogramming the genetic code. *EMBO J* **2011**, *30*, 2312-24.

343. de Graaf, A. J.; Kooijman, M.; Hennink, W. E.; Mastrobattista, E., Nonnatural amino acids for site-specific protein conjugation. *Bioconjug Chem* **2009**, *20*, 1281-95.

344. Hancock, S. M.; Uprety, R.; Deiters, A.; Chin, J. W., Expanding the genetic code of yeast for incorporation of diverse unnatural amino acids via a pyrrolysyl-tRNA synthetase/tRNA pair. *J Am Chem Soc* **2010**, *132*, 14819-24.

345. Schmidt, M. J.; Summerer, D., A need for speed: genetic encoding of rapid cycloaddition chemistries for protein labelling in living cells. *Chembiochem* **2012**, *13*, 1553-7.

346. Summerer, D.; Chen, S.; Wu, N.; Deiters, A.; Chin, J.; Schultz, P., A genetically encoded fluorescent amino acid. *Proc Natl Acad Sci U S A* **2006**, *103*, 9785-9.

347. Wang, L.; Xie, J.; Schultz, P., Expanding the genetic code. *Annu Rev Biophys Biomol Struct* **2006**, *35*, 225-49.

348. Young, T.; Ahmad, I.; Yin, J.; Schultz, P., An enhanced system for unnatural amino acid mutagenesis in E. coli. *J Mol Biol* **2010**, *395*, 361-74.

349. Perona, J. J.; Hadd, A., Structural diversity and protein engineering of the aminoacyl-tRNA synthetases. *Biochemistry* **2012**, *51*, 8705-29.

350. Bianco, A.; Townsley, F. M.; Greiss, S.; Lang, K.; Chin, J. W., Expanding the genetic code of Drosophila melanogaster. *Nat Chem Biol* **2012**, *8*, 748-50.

351. Gautier, A.; Nguyen, D. P.; Lusic, H.; An, W.; Deiters, A.; Chin, J. W., Genetically encoded photocontrol of protein localization in mammalian cells. *J Am Chem Soc* **2010**, *132*, 4086-8.

352. Thibodeaux, G. N.; Liang, X.; Moncivais, K.; Umeda, A.; Singer, O.; Alfonta, L.; Zhang, Z. J., Transforming a pair of orthogonal tRNA-aminoacyl-tRNA synthetase from Archaea to function in mammalian cells. *PLoS One* **2010**, *5*, e11263.

353. Krzycki, J. A., Function of genetically encoded pyrrolysine in corrinoid-dependent methylamine methyltransferases. *Curr Opin Chem Biol* **2004**, *8*, 484-91.

354. Gaston, M. A.; Jiang, R.; Krzycki, J. A., Functional context, biosynthesis, and genetic encoding of pyrrolysine. *Curr Opin Microbiol* **2011**, *14*, 342-9.

355. Srinivasan, G.; James, C. M.; Krzycki, J. A., Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science* **2002**, *296*, 1459-62.

356. Hao, B.; Zhao, G.; Kang, P. T.; Soares, J. A.; Ferguson, T. K.; Gallucci, J.; Krzycki, J. A.; Chan, M. K., Reactivity and chemical synthesis of L-pyrrolysine- the 22(nd) genetically encoded amino acid. *Chem Biol* **2004**, *11*, 1317-24.

357. Gaston, M. A.; Zhang, L.; Green-Church, K. B.; Krzycki, J. A., The complete biosynthesis of the genetically encoded amino acid pyrrolysine from lysine. *Nature* **2011**, *471*, 647-50.

358. Mahapatra, A.; Srinivasan, G.; Richter, K. B.; Meyer, A.; Lienard, T.; Zhang, J. K.; Zhao, G.; Kang, P. T.; Chan, M.; Gottschalk, G.; Metcalf, W. W.; Krzycki, J. A., Class I and class II lysyl-tRNA synthetase mutants and the genetic encoding of pyrrolysine in Methanosarcina spp. *Mol Microbiol* **2007**, *64*, 1306-18.

359. Nozawa, K.; O'Donoghue, P.; Gundllapalli, S.; Araiso, Y.; Ishitani, R.; Umehara, T.; Söll, D.; Nureki, O., Pyrrolysyl-tRNA synthetase-tRNA(Pyl) structure reveals the molecular basis of orthogonality. *Nature* **2009**, *457*, 1163-7.

360. Yanagisawa, T.; Ishii, R.; Fukunaga, R.; Kobayashi, T.; Sakamoto, K.; Yokoyama, S., Multistep engineering of pyrrolysyl-tRNA synthetase to genetically encode N(epsilon)-(o-azidobenzyloxycarbonyl) lysine for site-specific protein modification. *Chem Biol* **2008**, *15*, 1187-97.

361. Kobayashi, T.; Yanagisawa, T.; Sakamoto, K.; Yokoyama, S., Recognition of non-alpha-amino substrates by pyrrolysyl-tRNA synthetase. *J Mol Biol* **2009**, *385*, 1352-60.

362. Fekner, T.; Li, X.; Chan, M., Pyrrolysine Analogs for Translational Incorporation into Proteins. *Eur.J. Org Chem* **2010**, 4171-4179.

363. Neumann, H.; Peak-Chew, S.; Chin, J., Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. *Nat Chem Biol* **2008**, *4*, 232-4.

364. Brunner, J., New photolabeling and crosslinking methods. *Annu Rev Biochem* **1993**, *62*, 483-514.

365. Kapfer, I.; Jacques, P.; Toubal, H.; Goeldner, M. P., Comparative photoaffinity labeling study between azidophenyl, difluoroazidophenyl, and tetrafluoroazidophenyl derivatives for the GABA-gated chloride channels. *Bioconjug Chem* **1995**, *6*, 109-14.

366. Foucaud, B.; Gambs, S.; Schmid, N.; Behr, J. P.; Goeldner, M., A photoaffinity probe for the polyamine site regulating the NMDA receptor. *Eur J Pharmacol* **1995**, *289*, 471-7.

367. Hatanaka, Y.; Sadakane, Y., Photoaffinity labeling in drug discovery and developments: chemical gateway for entering proteomic frontier. *Curr Top Med Chem* **2002**, *2*, 271-88.

368. Das, J., Aliphatic diazirines as photoaffinity probes for proteins: recent developments. *Chem Rev* **2011**, *111*, 4405-17.

369. Knoll, F.; Kolter, T.; Sandhoff, K., Sphingolipid photoaffinity labels. *Methods Enzymol* **2000**, *311*, 568-600.

370. Bond, M. R.; Zhang, H.; Vu, P. D.; Kohler, J. J., Photocrosslinking of glycoconjugates using metabolically incorporated diazirine-containing sugars. *Nat Protoc* **2009**, *4*, 1044-63.

371. Lentzen, O.; Constant, J. F.; Defrancq, E.; Prévost, M.; Schumm, S.; Moucheron, C.; Dumy, P.; Kirsch-De Mesmaeker, A., Photocrosslinking in ruthenium-labelled duplex oligonucleotides. *Chembiochem* **2003**, *4*, 195-202.

372. Lentzen, O.; Constant, J. F.; Defrancq, E.; Moucheron, C.; Dumy, P.; Kirsch-De Mesmaeker, A., Photoadduct leading to crosslinking in Ru(II)-derivatized oligonucleotides. *Nucleos Nucleot Nucl* **2003**, *22*, 1487-9.

373. Lentzen, O.; Defrancq, E.; Constant, J. F.; Schumm, S.; García-Fresnadillo, D.; Moucheron, C.; Dumy, P.; Kirsch-De Mesmaeker, A., Determination of DNA guanine sites forming photo-adducts with Ru(II)-labeled oligonucleotides; DNA polymerase inhibition by the resulting photo-crosslinking. *J Biol Inorg Chem* **2004**, *9*, 100-8.

374. Geyer, H.; Geyer, R.; Pingoud, V., A novel strategy for the identification of protein-DNA contacts by photocrosslinking and mass spectrometry. *Nucl Acids Res* **2004**, *32*, e132.

375. Suchanek, M.; Radzikowska, A.; Thiele, C., Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells. *Nat Methods* **2005**, *2*, 261-7.

376. Preston, G. W.; Wilson, A. J., Photo-induced covalent cross-linking for the analysis of biomolecular interactions. *Chem Soc Rev* **2013**, *42*, 3289-301.

377. Dormán, G.; Prestwich, G. D., Benzophenone photophores in biochemistry. *Biochemistry* **1994**, *33*, 5661-73.

378. Chou, C.; Uprety, R.; Davis, L.; Chin, J.; Deiters, A., Genetically encoding an aliphatic diazirine for protein photocrosslinking. *Chemical Science* **2011**, *2*, 480-483.

379. Dubinsky, L.; Krom, B. P.; Meijler, M. M., Diazirine based photoaffinity labeling. *Bioorg Med Chem* **2012**, *20*, 554-70.

380. Rigaut, G.; Shevchenko, A.; Rutz, B.; Wilm, M.; Mann, M.; Séraphin, B., A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **1999**, *17*, 1030-2.

381. Chin, J.; Martin, A.; King, D.; Wang, L.; Schultz, P., Addition of a photocrosslinking amino acid to the genetic code of Escherichiacoli. *Proc Natl Acad Sci U S A* **2002**, *99*, 11020-4.

382. Tippmann, E.; Liu, W.; Summerer, D.; Mack, A.; Schultz, P., A genetically encoded diazirine photocrosslinker in Escherichia coli. *Chembiochem* **2007**, *8*, 2210-4.

383. Chin, J.; Cropp, T.; Anderson, J.; Mukherji, M.; Zhang, Z.; Schultz, P., An expanded eukaryotic genetic code. *Science* **2003**, *301*, 964-7.

384. Braig, D.; Bär, C.; Thumfart, J. O.; Koch, H. G., Two cooperating helices constitute the lipid-binding domain of the bacterial SRP receptor. *J Mol Biol* **2009**, *390*, 401-13.

385. Okuda, S.; Tokuda, H., Model of mouth-to-mouth transfer of bacterial lipoproteins through inner membrane LolC, periplasmic LolA, and outer membrane LolB. *Proc Natl Acad Sci U S A* **2009**, *106*, 5877-82.

386. Metha, A.; Jaouhari, R.; Benson , T.; Douglas, K., Improved Efficiency and Selectivity in Peptide-Synthesis - Uuse of Triethylsilane as a Carbocation Scavenger in Deprotection of tert-Butyl Esters and tert-Butoxycarbonyl -Protected Sites . *Tetrahedron Lett.* **1992**, *33*, 5441-5444.

387. Vanek, V.; Budesínský, M.; Kabeleová, P.; Sanda, M.; Kozísek, M.; Hanclová, I.; Mládková, J.; Brynda, J.; Rosenberg, I.; Koutmos, M.; Garrow, T. A.; Jirácek, J., Structure-activity study of new inhibitors of human betaine-homocysteine S-methyltransferase. *J Med Chem* **2009**, *52*, 3652-65.

388. Saavedra, O. M.; Isakovic, L.; Llewellyn, D. B.; Zhan, L.; Bernstein, N.; Claridge, S.; Raeppel, F.; Vaisburg, A.; Elowe, N.; Petschner, A. J.; Rahil, J.; Beaulieu, N.; MacLeod, A. R.; Delorme, D.; Besterman, J. M.; Wahhab, A., SAR around (l)-S-adenosyl-l-homocysteine, an inhibitor of human DNA methyltransferase (DNMT) enzymes. *Bioorg Med Chem Lett* **2009**, *19*, 2747-51.

389. Isakovic, L.; Saavedra, O. M.; Llewellyn, D. B.; Claridge, S.; Zhan, L.; Bernstein, N.; Vaisburg, A.; Elowe, N.; Petschner, A. J.; Rahil, J.; Beaulieu, N.; Gauthier, F.; MacLeod, A. R.; Delorme, D.; Besterman, J. M.; Wahhab, A., Constrained (l-)-S-adenosyl-l-homocysteine

(SAH) analogues as DNA methyltransferase inhibitors. *Bioorg Med Chem Lett* **2009**, *19*, 2742-6.

390. Jakubowski, H., Translational incorporation of S-nitrosohomocysteine into protein. *J Biol Chem* **2000**, *275*, 21813-6.

391. Stanger, O.; Fowler, B.; Piertzik, K.; Huemer, M.; Haschke-Becher, E.; Semmler, A.; Lorenzl, S.; Linnebank, M., Homocysteine, folate and vitamin B12 in neuropsychiatric diseases: review and treatment recommendations. *Expert Rev Neurother* **2009**, *9*, 1393-412.

392. Ford, E. S.; Smith, S. J.; Stroup, D. F.; Steinberg, K. K.; Mueller, P. W.; Thacker, S. B., Homocyst(e)ine and cardiovascular disease: a systematic review of the evidence with special emphasis on case-control studies and nested case-control studies. *Int J Epidemiol* **2002**, *31*, 59-70.

393. Jakubowski, H.; Zhang, L.; Bardeguez, A.; Aviv, A., Homocysteine thiolactone and protein homocysteinylation in human endothelial cells: implications for atherosclerosis. *Circ Res* **2000**, *87*, 45-51.

394. Baslé, E.; Joubert, N.; Pucheault, M., Protein chemical modification on endogenous amino acids. *Chem Biol* **2010**, *17*, 213-27.

395. Deribe, Y. L.; Pawson, T.; Dikic, I., Post-translational modifications in signal integration. *Nat Struct Mol Biol* **2010**, *17*, 666-72.

396. Gautier, A.; Nguyen, D.; Lusic, H.; An, W.; Deiters, A.; Chin, J., Genetically encoded photocontrol of protein localization in mammalian cells. *J Am Chem Soc* **2010**, *132*, 4086-8.

397. Ding, C.; Silverman, R., Synthesis of *N*-carbobenzyloxy-*N*,*N*-acetals. *Synthetic Communications* **1993**, *23*, 1467-1471.

398. Weitz, I. S.; Pellegrini, M.; Mierke, D. F.; Chorev, M., Synthesis of a Trisubstituted 1,4-Diazepin-3-one-Based Dipeptidomimetic as a Novel Molecular Scaffold. *J Org Chem* **1997**, *62*, 2527-2534.

399. McInnis, C. E.; Blackwell, H. E., Thiolactone modulators of quorum sensing revealed through library design and screening. *Bioorg Med Chem* **2011**, *19*, 4820-8.

400. Pace, N. J.; Weerapana, E., Diverse functional roles of reactive cysteines. *ACS Chem Biol* **2013**, *8*, 283-96.

401. Kim, Y.; Ho, S. O.; Gassman, N. R.; Korlann, Y.; Landorf, E. V.; Collart, F. R.; Weiss, S., Efficient site-specific labeling of proteins via cysteines. *Bioconjug Chem* **2008**, *19*, 786-91.

402. Antelmann, H.; Helmann, J. D., Thiol-based redox switches and gene regulation. *Antioxid Redox Signal* **2011**, *14*, 1049-63.

403. Reddie, K. G.; Carroll, K. S., Expanding the functional diversity of proteins through cysteine oxidation. *Curr Opin Chem Biol* **2008**, *12*, 746-54.

404. Spasser, L.; Brik, A., Chemistry and biology of the ubiquitin signal. *Angew Chem Int Ed Engl* **2012**, *51*, 6840-62.

405. Tasaki, T.; Kwon, Y. T., The mammalian N-end rule pathway: new insights into its components and physiological roles. *Trends Biochem Sci* **2007**, *32*, 520-8.

406. Lamkanfi, M.; Festjens, N.; Declercq, W.; Vanden Berghe, T.; Vandenabeele, P., Caspases in cell survival, proliferation and differentiation. *Cell Death Differ* **2007**, *14*, 44-55.

407. Riedl, S. J.; Shi, Y., Molecular mechanisms of caspase regulation during apoptosis. *Nat Rev Mol Cell Biol* **2004**, *5*, 897-907.

408. Kemp, M.; Go, Y. M.; Jones, D. P., Nonequilibrium thermodynamics of thiol/disulfide redox systems: a perspective on redox systems biology. *Free Radic Biol Med* **2008**, *44*, 921-37.

409. Netto, L. E.; de Oliveira, M. A.; Monteiro, G.; Demasi, A. P.; Cussiol, J. R.; Discola, K. F.; Demasi, M.; Silva, G. M.; Alves, S. V.; Faria, V. G.; Horta, B. B., Reactive cysteine in proteins: protein folding, antioxidant defense, redox signaling and more. *Comp Biochem Physiol C Toxicol Pharmacol* **2007**, *146*, 180-93.

410. Shao, L. E.; Tanaka, T.; Gribi, R.; Yu, J., Thioredoxin-related regulation of NO/NOS activities. *Ann N Y Acad Sci* **2002**, *962*, 140-50.

411. Jones, D. P., Radical-free biology of oxidative stress. *Am J Physiol Cell Physiol* **2008**, *295*, C849-68.

412. Balaban, R. S.; Nemoto, S.; Finkel, T., Mitochondria, oxidants, and aging. *Cell* **2005**, *120*, 483-95.

413. Finkel, T., Radical medicine: treating ageing to cure disease. *Nat Rev Mol Cell Biol* **2005**, *6*, 971-6.

414. Delaunay, A.; Pflieger, D.; Barrault, M. B.; Vinh, J.; Toledano, M. B., A thiol peroxidase is an H2O2 receptor and redox-transducer in gene activation. *Cell* **2002**, *111*, 471-81.

415. Collins, Y.; Chouchani, E. T.; James, A. M.; Menger, K. E.; Cochemé, H. M.; Murphy, M. P., Mitochondrial redox signalling at a glance. *J Cell Sci* **2012**, *125*, 801-6.

416. Bastin, G.; Singh, K.; Dissanayake, K.; Mighiu, A. S.; Nurmohamed, A.; Heximer, S. P., Amino-terminal cysteine residues differentially influence RGS4 protein plasma membrane targeting, intracellular trafficking, and function. *J Biol Chem* **2012**, *287*, 28966-74.

417. Greaves, J.; Chamberlain, L. H., Palmitoylation-dependent protein sorting. *J Cell Biol* **2007**, *176*, 249-54.

418. Smotrys, J. E.; Linder, M. E., Palmitoylation of intracellular signaling proteins: regulation and function. *Annu Rev Biochem* **2004**, *73*, 559-87.

419. Higdon, A.; Diers, A. R.; Oh, J. Y.; Landar, A.; Darley-Usmar, V. M., Cell signalling by reactive lipid species: new concepts and molecular mechanisms. *Biochem J* **2012**, *442*, 453-64.

420. Lindahl, M.; Mata-Cabana, A.; Kieselbach, T., The disulfide proteome and other reactive cysteine proteomes: analysis and functional significance. *Antioxid Redox Signal* **2011**, *14*, 2581-642.

421. Giles, N.; Giles, G.; Jacob, C., Multiple roles of cysteine in biocatalysis. *Biochem Bioph Res Co.* **2003**, *300*, 1-4.

422. Fass, D., Disulfide bonding in protein biophysics. *Annu Rev Biophys* **2012**, *41*, 63-79.

423. Giron, P.; Dayon, L.; Sanchez, J. C., Cysteine tagging for MS-based proteomics. *Mass Spectrom Rev* **2011**, *30*, 366-95.

424. Giron, P.; Dayon, L.; David, F.; Sanchez, J. C.; Rose, K., Enrichment of N-terminal cysteinyl-peptides by covalent capture. *J Proteomics* **2009**, *71*, 647-61.

425. Giron, P.; Dayon, L.; Mihala, N.; Sanchez, J. C.; Rose, K., Cysteine-reactive covalent capture tags for enrichment of cysteine-containing peptides. *Rapid Commun Mass Spectrom* **2009**, *23*, 3377-86.

426. O'Hare, H. M.; Johnsson, K.; Gautier, A., Chemical probes shed light on protein function. *Curr Opin Struct Biol* **2007**, *17*, 488-94.

427. Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F., Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **2010**, *468*, 790-5.

428. Holmes, K. L.; Lantz, L. M., Protein labeling with fluorescent probes. *Methods Cell Biol* **2001**, *63*, 185-204.

429. Scheck, R. A.; Francis, M. B., Regioselective labeling of antibodies through N-terminal transamination. *ACS Chem Biol* **2007**, *2*, 247-51.

430. Duckworth, B. P.; Zhang, Z.; Hosokawa, A.; Distefano, M. D., Selective labeling of proteins by using protein farnesyltransferase. *Chembiochem* **2007**, *8*, 98-105.

431. Nagahara, N.; Sawada, N.; Nakagawa, T., Affinity labeling of a catalytic site, cysteine(247), in rat mercaptopyruvate sulfurtransferase by chloropyruvate as an analog of a substrate. *Biochimie* **2004**, *86*, 723-9.

432. Adams, S. R.; Campbell, R. E.; Gross, L. A.; Martin, B. R.; Walkup, G. K.; Yao, Y.; Llopis, J.; Tsien, R. Y., New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications. *J Am Chem Soc* **2002**, *124*, 6063-76.

433. Wu, N.; Deiters, A.; Cropp, T.; King, D.; Schultz, P., A genetically encoded photocaged amino acid. *J Am Chem Soc* **2004**, *126*, 14306-7.

434. Wöll, D.; Laimgruber, S.; Galetskaya, M.; Smirnova, J.; Pfleiderer, W.; Heinz, B.; Gilch, P.; Steiner, U., On the mechanism of intramolecular sensitization of photocleavage of the 2-(2-nitrophenyl)propoxycarbonyl (NPPOC) protecting group. *J Am Chem Soc* **2007**, *129*, 12148-58.

435. Wöll, D.; Smirnova, J.; Galetskaya, M.; Prykota, T.; Bühler, J.; Stengele, K.; Pfleiderer, W.; Steiner, U., Intramolecular sensitization of photocleavage of the photolabile 2-(2-nitrophenyl)propoxycarbonyl (NPPOC) protecting group: photoproducts and photokinetics of the release of nucleosides. *Chemistry* **2008**, *14*, 6490-7.

436. Bhushan, K. R.; DeLisia, C.; Laursenb, R. A., Synthesis of photolabile 2-(2-nitrophenyl)propyloxycarbonyl protected amino acids. *Tetrahedron Lett.* 2003, 44, 8585-8588.

437. Mertz, J., Nonlinear microscopy: new techniques and applications. *Curr Opin Neurobiol* **2004**, *14*, 610-616.

438. Li, G. W.; Xie, X. S., Central dogma at the single-molecule level in living cells. *Nature* **2011**, *475*, 308-15.

439. Giepmans, B. N.; Adams, S. R.; Ellisman, M. H.; Tsien, R. Y., The fluorescent toolbox for assessing protein location and function. *Science* **2006**, *312*, 217-24.

440. Yu, B.; Kim, K.; So, P.; Blankschtein, D.; Langer, R., Topographic heterogeneity in transdermal transport revealed by high-speed two-photon microscopy: Determination of representative skin sample sizes. *J Invest Dermatol* **2002**, *118*, 1085-1088.

441. Gonçalves, M. S., Fluorescent labeling of biomolecules with organic probes. *Chem Rev* **2009**, *109*, 190-212.

442. Katritzky, A. R.; Narindoshvili, T., Fluorescent amino acids: advances in protein-extrinsic fluorophores. *Org Biomol Chem* **2009**, *7*, 627-34.

443. Hrdlovic, P.; Donovalova, J.; Stankovicova, H.; Gaplovsky, A., Influence of polarity of solvents on the spectral properties of bichromophoric coumarins. *Molecules* **2010**, *15*, 8915-32.

444. Suzuki, A.; Watanabe, T.; Kawamoto, M.; Nishiyama, K.; Yamashita, H.; Ishii, M.; Iwamura, M.; Furuta, T., Coumarin-4-ylmethoxycarbonyls as phototriggers for alcohols and phenols. *Org. Lett.* **2003**, *5*, 4867-4870.

445. Furuta, T.; Momotake, A.; Sugimoto, M.; Hatayama, M.; Torigai, H.; Iwamura, M., Acyloxycoumarinylmethyl-caged cAMP, the photolabile and membrane-permeable derivative of cAMP that effectively stimulates pigment-dispersion response of melanophores. *Biochem Biophys Res Commun* **1996**, 228, 193-8.

446. Braun, M.; Dittrich, T., Synthesis of the fluorescent amino acid rac-(7-hydroxycoumarin-4-yl)ethylglycine. *Beilstein J Org Chem* **2010**, 6.

447. de Melo, J.; Fernandes, P., Spectroscopy and photophysics of 4-and 7-hydroxycoumarins and their thione analogs. *J Mol Struct* **2001**, *565*, 69-78.

448. Fonseca, A. S.; Gonçalves, M. S.; Costa, S. P., A photoactivable amino acid based on a novel functional coumarin-6-yl-alanine. *Amino Acids* **2012**, *43*, 2329-38.

449. Li, W. H.; Zheng, G., Photoactivatable fluorophores and techniques for biological imaging applications. *Photochem Photobiol Sci* **2012**, *11*, 460-71.

450. Charbon, G.; Brustad, E.; Scott, K. A.; Wang, J.; Løbner-Olesen, A.; Schultz, P. G.; Jacobs-Wagner, C.; Chapman, E., Subcellular protein localization by using a genetically encoded fluorescent amino acid. *Chembiochem* **2011**, *12*, 1818-21.

451. Ho, D.; Lugo, M. R.; Lomize, A. L.; Pogozheva, I. D.; Singh, S. P.; Schwan, A. L.; Merrill, A. R., Membrane topology of the colicin E1 channel using genetically encoded fluorescence. *Biochemistry* **2011**, *50*, 4830-42.

452. Campos-Toimil, M.; Orallo, F.; Santana, L.; Uriarte, E., Synthesis and vasorelaxant activity of new coumarin and furocoumarin derivatives. *Bioorg.Med. Chem. Lett.* 2002, 12, 783-786.

453. Montgomery, H.; Perdicakis, B.; Fishlock, D.; Lajoie, G.; Jervis, E.; Guillemette, J., Photo-control of nitric oxide synthase activity using a caged isoform specific inhibitor. *Bioorg.Med. Chem.* **2002**, *10*, 1919-1927.

454. Bourbon, P.; Peng, Q.; Ferraudi, G.; Stauffacher, C.; Wiest, O.; Helquist, P., Synthesis, Photophysical, Photochemical, and Computational Studies of Coumarin-Labeled Nicotinamide Derivatives. *J. Org. Chem.* **2012**, *77*, 2756-2762.

455. Zhu, Q.; Uttamchandani, M.; Li, D.; Lesaicherre, M.; Yao, S., Enzymatic profiling system in a small-molecule microarray. *Org. Lett.* **2003**, *5*, 1257-1260.

456. Soh, N., Selective chemical labeling of proteins with small fluorescent molecules based on metal-chelation methodology. *Sensors* **2008**, *8*, 1004-1024.

457. Wombacher, R.; Cornish, V. W., Chemical tags: applications in live cell fluorescence imaging. *J Biophotonics* **2011**, *4*, 391-402.

458. Vendrell, M.; Zhai, D.; Er, J. C.; Chang, Y. T., Combinatorial strategies in fluorescent probe development. *Chem Rev* **2012**, *112*, 4391-420.

459. Jing, C.; Cornish, V. W., Chemical tags for labeling proteins inside living cells. *Acc Chem Res* **2011**, *44*, 784-92.

460. Jung, D.; Min, K.; Jung, J.; Jang, W.; Kwon, Y., Chemical biology-based approaches on fluorescent labeling of proteins in live cells. *Mol.BioSys.* **2013**, *9*, 862-872.

461. Goguen, B. N.; Loving, G. S.; Imperiali, B., Development of a fluorogenic sensor for activated Cdc42. *Bioorg Med Chem Lett* **2011**, *21*, 5058-61.

462. Yang, M.; Song, Y.; Zhang, M.; Lin, S.; Hao, Z.; Liang, Y.; Zhang, D.; Chen, P. R., Converting a solvatochromic fluorophore into a protein-based pH indicator for extreme acidity. *Angew Chem Int Ed Engl* **2012**, *51*, 7674-9.

463. Loving, G. S.; Sainlos, M.; Imperiali, B., Monitoring protein interactions and dynamics with solvatochromic fluorophores. *Trends Biotechnol* **2010**, *28*, 73-83.

464. Vázquez, M. E.; Blanco, J. B.; Imperiali, B., Photophysics and biological applications of the environment-sensitive fluorophore 6-N,N-dimethylamino-2,3-naphthalimide. *J Am Chem Soc* **2005**, *127*, 1300-6.

465. Eugenio Vazquez, M.; Rothman, D. M.; Imperiali, B., A new environment-sensitive fluorescent amino acid for Fmoc-based solid phase peptide synthesis. *Org Biomol Chem* **2004**, *2*, 1965-6.

466. Riedl, J.; Pohl, R.; Ernsting, N.; Orsag, P.; Fojta, M.; Hocek, M., Labelling of nucleosides and oligonucleotides by solvatochromic 4-aminophthalimide fluorophore for studying DNA-protein interactions. *Chemical Science* **2012**, *3*, 2797-2806.

467. Sikoraiova, J.; Marchalin, S.; Daich, A.; Decroix, B., Acid-mediated intramolecular cationic cyclization using an oxygen atom as internal nucleophile: synthesis of substituted oxazolo-, oxazino- and oxazepinoisoindolinones. *Tetrahedron Lett.* 2002, 43, 4747-4751.

468. Soine, T., The beta-hydroxy and beta-halo ethyl-3-nitro and 4-nitro phthalimides. *J. Am. Pharm. Assoc.-Sci. Ed*. **1946**, *35*, 42-43.

469. Suzuki, T.; Ando, T.; Tsuchiya, K.; Fukazawa, N.; Saito, A.; Mariko, Y.; Yamashita, T.; Nakanishi, O., Synthesis and histone deacetylase inhibitory activity of new benzamide derivatives. *J. Med. Chem.* **1999**, *42*, 3001-3003.

470. Katritzky, A.; Narindoshvili, T., Fluorescent amino acids: advances in protein-extrinsic fluorophores. *Org Biomol Chem* **2009**, *7*, 627-34.

471. Saroja, G.; Ramachandram, B.; Saha, S.; Samanta, A., The fluorescence response of a structurally modified 4-aminophthalimide derivative covalently attached to a fatty acid in homogeneous and micellar environments. *J. Phy .Chem. B* **1999**, *103*, 2906-2911.

472. Vazquez, M.; Rothman, D.; Imperiali, B., A new environment-sensitive fluorescent amino acid for Fmoc-based solid phase peptide synthesis. *Org. Biomol. Chem.* **2004**, *2*, 1965-1966.

473. Pascale, R.; Carocci, A.; Catalano, A.; Lentini, G.; Spagnoletta, A.; Cavalluzzi, M.; De Santis, F.; De Palma, A.; Scalera, V.; Franchini, C., New N-(phenoxydecyl) phthalimide derivatives displaying potent inhibition activity towards alpha-glucosidase. *Bioorg. Med. Chem.* 2010, 18, 5903-5914.

474. Samanta, S.; Woolley, G. A., Bis-azobenzene crosslinkers for photocontrol of peptide structure. *Chembiochem* **2011**, *12*, 1712-23.

475. Pompella, A.; Visvikis, A.; Paolicchi, A.; De Tata, V.; Casini, A. F., The changing faces of glutathione, a cellular protagonist. *Biochem Pharmacol* **2003**, *66*, 1499-503.

476. Deeg, A. A.; Schrader, T. E.; Kempter, S.; Pfizer, J.; Moroder, L.; Zinth, W., Light-triggered aggregation and disassembly of amyloid-like structures. *Chemphyschem* **2011**, *12*, 559-62.

477. Hoppmann, C.; Kühne, R.; Beyermann, M., Intramolecular bridges formed by photoswitchable click amino acids. *Beilstein J Org Chem* **2012**, *8*, 884-9.

478. Szymański, W.; Wu, B.; Poloni, C.; Janssen, D. B.; Feringa, B. L., Azobenzene photoswitches for staudinger-bertozzi ligation. *Angew Chem Int Ed Engl* **2013**, *52*, 2068-72.

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

480. Schierling, B.; Noël, A. J.; Wende, W.; Hien, l. T.; Volkov, E.; Kubareva, E.; Oretskaya, T.; Kokkinidis, M.; Römpp, A.; Spengler, B.; Pingoud, A., Controlling the enzymatic activity of a restriction enzyme by light. *Proc Natl Acad Sci U S A* **2010**, *107*, 1361-6.

481. Stawski, P.; Sumser, M.; Trauner, D., A photochromic agonist of AMPA receptors. *Angew Chem Int Ed Engl* **2012**, *51*, 5748-51.

482. Mourot, A.; Fehrentz, T.; Le Feuvre, Y.; Smith, C. M.; Herold, C.; Dalkara, D.; Nagy, F.; Trauner, D.; Kramer, R. H., Rapid optical control of nociception with an ion-channel photoswitch. *Nat Methods* **2012**, *9*, 396-402.

483. Goulet-Hanssens, A.; Lai Wing Sun, K.; Kennedy, T. E.; Barrett, C. J., Photoreversible surfaces to regulate cell adhesion. *Biomacromolecules* **2012**, *13*, 2958-63.

484. Yu, B.; Shirai, Y.; Tour, J., Syntheses of new functionalized azobenzenes for potential molecular electronic devices. *Tetrahedron* **2006**, *62*, 10303-10310.

485. Guillonneau, C.; Charton, Y.; Ginot, Y. M.; Fouquier-d'Hérouël, M. V.; Bertrand, M.; Lockhart, B.; Lestage, P.; Goldstein, S., Synthesis and pharmacological evaluation of new 1,2-dithiolane based antioxidants. *Eur J Med Chem* **2003**, *38*, 1-11.

486. Szeląg, M.; Mikulski, D.; Molski, M., Quantum-chemical investigation of the structure and the antioxidant properties of α-lipoic acid and its metabolites. *J Mol Model* **2012**, *18*, 2907-16.

487. Jansen, E., The isolation and identification of 2,2'-Dithiolisobutyric acid from asparagu. *J. Biol. Chem.* **1948**, *176*, 657-664.

488. Reed, L.; Gunsalus, I.; Schnakenberg, G.; Soper, Q.; Boaz, H.; Kern, S.; Parke, T., Isolation, characterization and structure of alpha-lipoic acid. *J. Am. Chem. Soc.* **1953**, *75*, 1267-1270.

489. Rock, C. O., Opening a new path to lipoic acid. *J Bacteriol* **2009**, *191*, 6782-4.

490. Schupke, H.; Hempel, R.; Peter, G.; Hermann, R.; Wessel, K.; Engel, J.; Kronbach, T., New metabolic pathways of alpha-lipoic acid. *Drug Metab Dispos* **2001**, *29*, 855-62.

491. Sigel, H., The hydrophobic and metal-ion coordinating properties of alpha lipoic acid-an example of intramolecular equilibria in metal-ion complexes. *Angew. Chem.Int.Ed. Eng.* **1982**, *21*, 389-400.

492. Susumu, K.; Mei, B. C.; Mattoussi, H., Multifunctional ligands based on dihydrolipoic acid and polyethylene glycol to promote biocompatibility of quantum dots. *Nat Protoc* **2009**, *4*, 424-36.

493. Abad, J. M.; Mertens, S. F.; Pita, M.; Fernández, V. M.; Schiffrin, D. J., Functionalization of thioctic acid-capped gold nanoparticles for specific immobilization of histidine-tagged proteins. *J Am Chem Soc* **2005**, *127*, 5689-94.

494. Frasconi, M.; Mazzei, F.; Ferri, T., Protein immobilization at gold-thiol surfaces and potential for biosensing. *Anal Bioanal Chem* **2010**, *398*, 1545-64.

495. Sharma, J.; Chhabra, R.; Andersen, C. S.; Gothelf, K. V.; Yan, H.; Liu, Y., Toward reliable gold nanoparticle patterning on self-assembled DNA nanoscaffold. *J Am Chem Soc* **2008**, *130*, 7820-1.

496. Tan, Y. H.; Schallom, J. R.; Ganesh, N. V.; Fujikawa, K.; Demchenko, A. V.; Stine, K. J., Characterization of protein immobilization on nanoporous gold using atomic force microscopy and scanning electron microscopy. *Nanoscale* **2011**, *3*, 3395-407.

497. Ahirwal, G. K.; Mitra, C. K., Direct electrochemistry of horseradish peroxidase-gold nanoparticles conjugate. *Sensors (Basel)* **2009**, *9*, 881-94.

498. Dougan, J. A.; Karlsson, C.; Smith, W. E.; Graham, D., Enhanced oligonucleotide-nanoparticle conjugate stability using thioctic acid modified oligonucleotides. *Nucl Acids Res* **2007**, *35*, 3668-75.

499. Fernández-Suárez, M.; Baruah, H.; Martínez-Hernández, L.; Xie, K. T.; Baskin, J. M.; Bertozzi, C. R.; Ting, A. Y., Redirecting lipoic acid ligase for cell surface protein labeling with small-molecule probes. *Nat Biotechnol* **2007**, *25*, 1483-7.

500. Ha, T. H.; Jung, S. O.; Lee, J. M.; Lee, K. Y.; Lee, Y.; Park, J. S.; Chung, B. H., Oriented immobilization of antibodies with GST-fused multiple Fc-specific B-domains on a gold surface. *Anal Chem* **2007**, *79*, 546-56.

501. Liu, D. S.; Loh, K. H.; Lam, S. S.; White, K. A.; Ting, A. Y., Imaging trans-cellular neurexin-neuroligin interactions by enzymatic probe ligation. *PLoS One* **2013**, *8*, e52823.

502. Ali, M. M.; Kang, D. K.; Tsang, K.; Fu, M.; Karp, J. M.; Zhao, W., Cell-surface sensors: lighting the cellular environment. *Wiley Interdiscip Rev Nanomed Nanobiotechnol* **2012**, *4*, 547-61.

503. Claeson, G., 1,2-Dithiolane-3-carboxylic acid. *Acta Chemica Scandinavica* **1955**, *9*, 178-180.

504. Johary, N.; Owen, L., Dithiols.17. S-benzyl derivatives of 2-3-Dimercaptopropanol and 1-3-dimercaptopropan-2-ol. *J. Chem. Soc.* 1955, 1302-1307.

505. Ghosh, S.; Martin, J.; Fried, J., A total synthesis of the methyl ester of the 9,11-dithia analog of 13,14-dehydro-PGH2.. *J. Org. Chem.* **1987**, *52*, 862-876.

506. Pattenden, G.; Shuker, A., Synthetic approaches towards the novel 1,3-Dioxo-1,2-dithiolane moiety in the antitumor antibiotic substance leinamycin. *J. Chem. Soc.-Perkin Trans 1* **1992**, 1215-1221.

507. Krishnamurthy, S.; Thompson, K., Selective reduction of functionalized carboxylic acids with Borane-methyl sulfide-convenient undergraduate organic experiment. *J.Chem. Ed.* **1977**, *54*, 778-779.

508. Axenrod, T.; Watnick, C.; Yazdekhasti, H.; Dave, P., Synthesis of 1,3,3-Trinitroazetidine via the oxidative nitrolysis of *N*-*p*-Tosyl-3-azetidinone oxime. *J. Org.Chem.* **1995**, *60*, 1959-1964.

509. Adamsen, T.; Grierson, J.; Krohn, K., A new synthesis of the labeling precursor for [F-18]-fluoromisonidazole. *J. Label. Compd. Radiopharm.* **2005**, *48*, 923-927.

510. Morera, E.; Lucente, G.; Ortar, G.; Nalli, M.; Mazza, F.; Gavuzzo, E.; Spisani, S., Exploring the interest of 1,2-dithiolane ring system in peptide chemistry. Synthesis of a chemotactic tripeptide and x-ray crystal structure of a 4-amino-1,2-dithiolane-4-carboxylic acid derivative. *Bioorg Med Chem* **2002**, *10*, 147-57.

511. Nesmelov, Y. E.; Thomas, D. D., Protein structural dynamics revealed by site-directed spin labeling and multifrequency EPR. *Biophys Rev* **2010**, *2*, 91-99.

512. Altenbach, C.; Froncisz, W.; Hyde, J. S.; Hubbell, W. L., Conformation of spin-labeled melittin at membrane surfaces investigated by pulse saturation recovery and continuous wave power saturation electron paramagnetic resonance. *Biophys J* **1989**, *56*, 1183-91.

513. Fleissner, M. R.; Bridges, M. D.; Brooks, E. K.; Cascio, D.; Kálai, T.; Hideg, K.; Hubbell, W. L., Structure and dynamics of a conformationally constrained nitroxide side chain and applications in EPR spectroscopy. *Proc Natl Acad Sci U S A* **2011**, *108*, 16241-6.

514. Marsh, D., Electron spin resonance in membrane research: protein-lipid interactions from challenging beginnings to state of the art. *Eur Biophys J* **2010**, *39*, 513-25.

515. Altenbach, C.; Flitsch, S. L.; Khorana, H. G.; Hubbell, W. L., Structural studies on transmembrane proteins. 2. Spin labeling of bacteriorhodopsin mutants at unique cysteines. *Biochemistry* **1989**, *28*, 7806-12.

516. Klare, J. P.; Steinhoff, H. J., Spin labeling EPR. *Photosynth Res* **2009**, *102*, 377-90.

517. Bordignon, E., Site-directed spin labeling of membrane proteins. *Top Curr Chem* **2012**, *321*, 121-57.

518. Strancar, J.; Koklic, T.; Arsov, Z.; Filipic, B.; Stopar, D.; Hemminga, M. A., Spin label EPR-based characterization of biosystem complexity. *J Chem Inf Model* **2005**, *45*, 394-406.

519. Kavalenka, A. A.; Spruijt, R. B.; Wolfs, C. J.; Strancar, J.; Croce, R.; Hemminga, M. A.; van Amerongen, H., Site-directed spin-labeling study of the light-harvesting complex CP29. *Biophys J* **2009**, *96*, 3620-8.

520. Klug, C.; Feix, J.; Correia, J.; Detrich, H., Methods and applications of site-directed spin Labeling EPR Spectroscopy. *Biophysical Tools For Biologists: Vol 1 in Vitro Techniques* **2008**, *84*, 617-658.

521. Möbius, K.; Savitsky, A.; Wegener, C.; Plato, M.; Fuchs, M.; Schnegg, A.; Dubinskii, A. A.; Grishin, Y. A.; Grigor'ev, I. A.; Kühn, M.; Duché, D.; Zimmermann, H.; Steinhoff, H. J., Combining high-field EPR with site-directed spin labeling reveals unique information on proteins in action. *Magn Reson Chem* **2005**, *43* Spec no., S4-S19.

522. Möbius, K.; Savitsky, A.; Schnegg, A.; Plato, M.; Fuchst, M., High-field EPR spectroscopy applied to biological systems: characterization of molecular switches for electron and ion transfer. *Phys Chem Chem Phys* **2005**, *7*, 19-42.

523. Strancar, J.; Kavalenka, A.; Urbancic, I.; Ljubetic, A.; Hemminga, M. A., SDSL-ESR-based protein structure characterization. *Eur Biophys J* **2010**, *39*, 499-511.

524. Cornish, V.; Benson, D.; Altenbach, C.; Hideg, K.; Hubbell, W.; Schultz, P., Site-specific incorporation of biophysical probes into proteins. *Proc Natl Acad Sci U S A* **1994**, *91*, 2910-4.

525. Wu, A.; Mader, E.; Datta, A.; Hrovat, D.; Borden, W.; Mayer, J., Nitroxyl Radical Plus Hydroxylamine Pseudo Self-Exchange Reactions: Tunneling in Hydrogen Atom Transfer. *J. Am. Chem. Soc.* **2009**, *131*, 11985-11997.

526. Church, R.; Weiss, M., Diazirines:. Synthesis and properties of small functionalized diazirine molecules-some observations on reaction of a diazirine with Iodine-Iodine ion system. *J. Org. Chem.* **1970**, *35*, 2465.

527. Thai, K.; Wang, L.; Dudding, T.; Bilodeau, F.; Gravel, M., NHC-Catalyzed Intramolecular Redox Amidation for the Synthesis of Functionalized Lactams. *Org. Lett.* **2010**, *12*, 5708-5711.

528. Braus, G. H., Aromatic amino acid biosynthesis in the yeast Saccharomyces cerevisiae: a model system for the regulation of a eukaryotic biosynthetic pathway. *Microbiol Rev* **1991**, *55*, 349-70.

529. Sprecher, M.; Srinivasan, P. R.; Sprinson, D. B.; Davis, B. D., The biosynthesis of tyrosine from labeled glucose in Escherichia coli. *Biochemistry* **1965**, *4*, 2855-60.

530. Koide, S.; Sidhu, S. S., The importance of being tyrosine: lessons in molecular recognition from minimalist synthetic binding proteins. *ACS Chem Biol* **2009**, *4*, 325-34.

531. Whiteson, K. L.; Chen, Y.; Chopra, N.; Raymond, A. C.; Rice, P. A., Identification of a potential general acid/base in the reversible phosphoryl transfer reactions catalyzed by tyrosine recombinases: Flp H305. *Chem Biol* **2007**, *14*, 121-9.

532. Abu Tarboush, N.; Jensen, L. M.; Feng, M.; Tachikawa, H.; Wilmot, C. M.; Davidson, V. L., Functional importance of tyrosine 294 and the catalytic selectivity for the bis-Fe(IV) state of MauG revealed by replacement of this axial heme ligand with histidine. *Biochemistry* **2010**, *49*, 9783-91.

533. Rouzer, C. A.; Marnett, L. J., Cyclooxygenases: structural and functional insights. *J Lipid Res* **2009**, *50* Suppl, S29-34.

534. Ali, F. E.; Barnham, K. J.; Barrow, C. J.; Separovic, F., Metal catalyzed oxidation of tyrosine residues by different oxidation systems of copper/hydrogen peroxide. *J Inorg Biochem* **2004**, *98*, 173-84.

535. Chiarugi, P.; Buricchi, F., Protein tyrosine phosphorylation and reversible oxidation: two cross-talking posttranslation modifications. *Antioxid Redox Signal* **2007**, *9*, 1-24.

536. Shenoy, S. R.; Jayaram, B., Proteins: sequence to structure and function--current status. *Curr Protein Pept Sci* **2010**, *11*, 498-514.

537. Arora, A.; Scholar, E. M., Role of tyrosine kinase inhibitors in cancer therapy. *J Pharmacol Exp Ther* **2005**, *315*, 971-9.

538. Cozzone, A. J., Protein phosphorylation in prokaryotes. *Annu Rev Microbiol* **1988**, *42*, 97-125.

539. Cozzone, A. J., Post-translational modification of proteins by reversible phosphorylation in prokaryotes. *Biochimie* **1998**, *80*, 43-8.

540. Huttner, W. B., Sulphation of tyrosine residues-a widespread modification of proteins. *Nature* **1982**, *299*, 273-6.

541. Liu, M. C.; Lipmann, F., Decrease of tyrosine-O-sulfate-containing proteins found in rat fibroblasts infected with Rous sarcoma virus or Fujinami sarcoma virus. *Proc Natl Acad Sci U S A* **1984**, *81*, 3695-8.

542. Begara-Morales, J. C.; Chaki, M.; Sánchez-Calvo, B.; Mata-Pérez, C.; Leterrier, M.; Palma, J. M.; Barroso, J. B.; Corpas, F. J., Protein tyrosine nitration in pea roots during development and senescence. *J Exp Bot* **2013**, *64*, 1121-34.

543. Corpas, F. J.; Palma, J. M.; Del Río, L. A.; Barroso, J. B., Protein tyrosine nitration in higher plants grown under natural and stress conditions. *Front Plant Sci* **2013**, *4*, 29.

544. Mathivanan, S.; Ahmed, M.; Ahn, N. G.; Alexandre, H.; Amanchy, R.; Andrews, P. C.; Bader, J. S.; Balgley, B. M.; Bantscheff, M.; Bennett, K. L.; Björling, E.; Blagoev, B.; Bose, R.; Brahmachari, S. K.; Burlingame, A. S.; Bustelo, X. R.; Cagney, G.; Cantin, G. T.; Cardasis, H. L.; Celis, J. E.; Chaerkady, R.; Chu, F.; Cole, P. A.; Costello, C. E.; Cotter, R. J.; Crockett, D.; DeLany, J. P.; De Marzo, A. M.; DeSouza, L. V.; Deutsch, E. W.; Dransfield, E.; Drewes, G.; Droit, A.; Dunn, M. J.; Elenitoba-Johnson, K.; Ewing, R. M.; Van Eyk, J.; Faca, V.; Falkner, J.; Fang, X.; Fenselau, C.; Figeys, D.; Gagné, P.; Gelfi, C.; Gevaert, K.; Gimble, J. M.; Gnad, F.; Goel, R.; Gromov, P.; Hanash, S. M.; Hancock, W. S.; Harsha, H. C.; Hart, G.; Hays, F.; He, F.; Hebbar, P.; Helsens, K.; Hermeking, H.; Hide, W.; Hjernø, K.; Hochstrasser, D. F.; Hofmann, O.; Horn, D. M.; Hruban, R. H.; Ibarrola, N.; James, P.; Jensen, O. N.; Jensen, P. H.; Jung, P.; Kandasamy, K.; Kheterpal, I.; Kikuno, R. F.; Korf, U.; Körner, R.; Kuster, B.; Kwon, M. S.; Lee, H. J.; Lee, Y. J.; Lefevre, M.; Lehvaslaiho, M.; Lescuyer, P.; Levander, F.; Lim, M. S.; Löbke, C.; Loo, J. A.; Mann, M.; Martens, L.; Martinez-Heredia, J.; McComb, M.; McRedmond, J.; Mehrle, A.; Menon, R.;

Miller, C. A.; Mischak, H.; Mohan, S. S.; Mohmood, R.; Molina, H.; Moran, M. F.; Morgan, J. D.; Moritz, R.; Morzel, M.; Muddiman, D. C.; Nalli, A.; Navarro, J. D.; Neubert, T. A.; Ohara, O.; Oliva, R.; Omenn, G. S.; Oyama, M.; Paik, Y. K.; Pennington, K.; Pepperkok, R.; Periaswamy, B.; Petricoin, E. F.; Poirier, G. G.; Prasad, T. S.; Purvine, S. O.; Rahiman, B. A.; Ramachandran, P.; Ramachandra, Y. L.; Rice, R. H.; Rick, J.; Ronnholm, R. H.; Salonen, J.; Sanchez, J. C.; Sayd, T.; Seshi, B.; Shankari, K.; Sheng, S. J.; Shetty, V.; Shivakumar, K.; Simpson, R. J.; Sirdeshmukh, R.; Siu, K. W.; Smith, J. C.; Smith, R. D.; States, D. J.; Sugano, S.; Sullivan, M.; Superti-Furga, G.; Takatalo, M.; Thongboonkerd, V.; Trinidad, J. C.; Uhlen, M.; Vandekerckhove, J.; Vasilescu, J.; Veenstra, T. D.; Vidal-Taboada, J. M.; Vihinen, M.; Wait, R.; Wang, X.; Wiemann, S.; Wu, B.; Xu, T.; Yates, J. R.; Zhong, J.; Zhou, M.; Zhu, Y.; Zurbig, P.; Pandey, A., Human Proteinpedia enables sharing of human protein data. *Nat Biotechnol* **2008**, *26*, 164-7.

545. Meyer, R. D.; Srinivasan, S.; Singh, A. J.; Mahoney, J. E.; Gharahassanlou, K. R.; Rahimi, N., PEST motif serine and tyrosine phosphorylation controls vascular endothelial growth factor receptor 2 stability and downregulation. *Mol Cell Biol* **2011**, *31*, 2010-25.

546. Mishra, O. P.; Ashraf, Q. M.; Delivoria-Papadopoulos, M., Hypoxia-induced activation of epidermal growth factor receptor (EGFR) kinase in the cerebral cortex of newborn piglets: the role of nitric oxide. *Neurochem Res* **2010**, *35*, 1471-7.

547. Mudduluru, M.; Zubrow, A. B.; Ashraf, Q. M.; Delivoria-Papadopoulos, M.; Mishra, O. P., Tyrosine phosphorylation of apoptotic proteins during hyperoxia in mitochondria of the cerebral cortex of newborn piglets. *Neurochem Res* **2010**, *35*, 1003-9.

548. Delivoria-Papadopoulos, M.; Mishra, O. P., Mechanism of post-translational modification by tyrosine phosphorylation of apoptotic proteins during hypoxia in the cerebral cortex of newborn piglets. *Neurochem Res* **2010**, *35*, 76-84.

549. Gross, A.; McDonnell, J. M.; Korsmeyer, S. J., BCL-2 family members and the mitochondria in apoptosis. *Genes Dev* **1999**, *13*, 1899-911.

550. Vendruscolo, M.; Dobson, C. M., Quantitative approaches to defining normal and aberrant protein homeostasis. *Faraday Discuss* **2009**, *143*, 277-91; discussion 359-72.

551. Bogdanova, O. V.; Kot, L. I.; Lavrova, K. V.; Bogdanov, V. B.; Sloan, E. K.; Beregova, T. V.; Ostapchenko, L. I., Modulation of protein tyrosine phosphorylation in gastric mucosa during re-epithelization processes. *World J Biol Chem* **2010**, *1*, 338-47.

552. Okamura, T.; Singh, S.; Buolamwini, J.; Haystead, T.; Friedman, H.; Bigner, D.; Ali-Osman, F., Tyrosine phosphorylation of the human glutathione S-transferase P1 by epidermal growth factor receptor. *J Biol Chem* **2009**, *284*, 16979-89.

553. Yokawa, K.; Kagenishi, T.; Goto, K.; Kawano, T., Free tyrosine and tyrosine-rich peptide-dependent superoxide generation catalyzed by a copper-binding, threonine-rich neurotoxic peptide derived from prion protein. *Int J Biol Sci* **2009**, *5*, 53-63.

554. Yuan, J.; Zhang, J.; Wong, B. W.; Si, X.; Wong, J.; Yang, D.; Luo, H., Inhibition of glycogen synthase kinase 3beta suppresses coxsackievirus-induced cytopathic effect and apoptosis via stabilization of beta-catenin. *Cell Death Differ* **2005**, *12*, 1097-106.

555. Chun, T. W.; Chadwick, K.; Margolick, J.; Siliciano, R. F., Differential susceptibility of naive and memory CD4+ T cells to the cytopathic effects of infection with human immunodeficiency virus type 1 strain LAI. *J Virol* **1997**, *71*, 4436-44.

556. Witting, P. K.; Mauk, A. G.; Lay, P. A., Role of tyrosine-103 in myoglobin peroxidase activity: kinetic and steady-state studies on the reaction of wild-type and variant recombinant human myoglobins with H_2O_2. *Biochemistry* **2002**, *41*, 11495-503.

557. Franzen, S.; Thompson, M. K.; Ghiladi, R. A., The dehaloperoxidase paradox. *Biochim Biophys Acta* **2012**, *1824*, 578-88.

558. Thompson, M. K.; Franzen, S.; Ghiladi, R. A.; Reeder, B. J.; Svistunenko, D. A., Compound ES of Dehaloperoxidase Decays via Two Alternative Pathways Depending on the Conformation of the Distal Histidine. *J Am Chem Soc* **2010**, *132*, 17501-10.

559. D'Antonio, J.; D'Antonio, E. L.; Thompson, M. K.; Bowden, E. F.; Franzen, S.; Smirnova, T.; Ghiladi, R. A., Spectroscopic and mechanistic investigations of dehaloperoxidase B from Amphitrite ornata. *Biochemistry* **2010**, *49*, 6600-16.

560. Parfenova, H.; Balabanova, L.; Leffler, C. W., Posttranslational regulation of cyclooxygenase by tyrosine phosphorylation in cerebral endothelial cells. *Am J Physiol* **1998**, *274*, C72-81.

561. Simmons, D. L.; Botting, R. M.; Hla, T., Cyclooxygenase isozymes: the biology of prostaglandin synthesis and inhibition. *Pharmacol Rev* **2004**, *56*, 387-437.

562. Picot, D.; Loll, P. J.; Garavito, R. M., The X-ray crystal structure of the membrane protein prostaglandin H2 synthase-1. *Nature* **1994**, *367*, 243-9.

563. Beynon, R. J.; Pratt, J. M., Metabolic labeling of proteins for proteomics. *Mol Cell Proteomics* **2005**, *4*, 857-72.

564. Becker, G. W., Stable isotopic labeling of proteins for quantitative proteomic applications. *Brief Funct Genomic Proteomic* **2008**, *7*, 371-82.

565. Becker, C. F.; Lausecker, K.; Balog, M.; Kálai, T.; Hideg, K.; Steinhoff, H. J.; Engelhard, M., Incorporation of spin-labelled amino acids into proteins. *Magn Reson Chem* **2005**, *43* Spec no., S34-9.

566. Bundy, B.; Swartz, J., Site-specific incorporation of p-propargyloxyphenylalanine in a cell-free environment for direct protein-protein click conjugation. *Bioconjug Chem* **2010**, *21*, 255-63.

567. Deiters, A.; Geierstanger, B.; Schultz, P., Site-specific in vivo labeling of proteins for NMR studies. *Chembiochem* **2005**, *6*, 55-8.

568. Deiters, A.; Schultz, P., In vivo incorporation of an alkyne into proteins in Escherichia coli. *Bioorg Med Chem Lett* **2005**, *15*, 1521-4.

569. Xie, J.; Schultz, P., A chemical toolkit for proteins--an expanded genetic code. *Nat Rev Mol Cell Biol* **2006**, *7*, 775-82.

570. Fleissner, M. R.; Brustad, E. M.; Kálai, T.; Altenbach, C.; Cascio, D.; Peters, F. B.; Hideg, K.; Peuker, S.; Schultz, P. G.; Hubbell, W. L., Site-directed spin labeling of a genetically encoded unnatural amino acid. *Proc Natl Acad Sci U S A* **2009**, *106*, 21637-42.

571. Neumann, H.; Wang, K.; Davis, L.; Garcia-Alai, M.; Chin, J., Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* **2010**, *464*, 441-4.

572. Takimoto, J. K.; Adams, K. L.; Xiang, Z.; Wang, L., Improving orthogonal tRNA-synthetase recognition for efficient unnatural amino acid incorporation and application in mammalian cells. *Mol Biosyst* **2009**, *5*, 931-4.

573. Wilkins, B.; Marionni, S.; Young, D.; Liu, J.; Wang, Y.; Di Salvo, M.; Deiters, A.; Cropp, T., Site-specific incorporation of fluorotyrosines into proteins in Escherichia coli by photochemical disguise. *Biochemistry* **2010**, *49*, 1557-9.

574. Cellitti, S.; Jones, D.; Lagpacan, L.; Hao, X.; Zhang, Q.; Hu, H.; Brittain, S.; Brinker, A.; Caldwell, J.; Bursulaya, B.; Spraggon, G.; Brock, A.; Ryu, Y.; Uno, T.; Schultz, P.; Geierstanger, B., In vivo incorporation of unnatural amino acids to probe structure, dynamics, and ligand binding in a large protein by nuclear magnetic resonance spectroscopy. *J Am Chem Soc* **2008**, *130*, 9268-81.

575. Escher, E.; Bernier, M.; Parent, P., Angiotensin -II analogs Synthesis and incorporation of the sulfur-containing aromatic amino-acids - L-(4'-SH)Phe, L-(4'-SO_2NH_2)Phe, L-(4'-SO_3-)Phe and L-(4'-S-CH_3)Phe. *Helv Chim Acta* **1983**, *66*, 1355-1365.

576. Lu, H.; Volk, M.; Kholodenko, Y.; Gooding, E.; Hochstrasser, R.; DeGrado, W., Aminothiotyrosine disulfide, an optical trigger for initiation of protein folding. *J. Am. Chem. Soc.***1997**, *119*, 7173-7180.

577. Miao, Z.; Liu, J.; Norman, T.; Driver, R. New non-natural amino acids useful for preparation of modified or non-modified, non-natural amino acid polypeptides, useful in e.g. treatment of a disorder, condition or disease. PCT Int. Appl. WO2006069246 Jun 29, **2006**.

578. Mukai, C.; Kobayashi, M.; Kubota, S.; Takahashi, Y.; Kitagaki, S., Construction of azacycles based on endo-mode cyclization of allenes. *J.Org. Chem.* **2004**, *69*, 2128-2136.

579. Sikkema, F.; Comellas-Aragones, M.; Fokkink, R.; Verduin, B.; Cornelissen, J.; Nolte, R., Monodisperse polymer-virus hybrid nanoparticles. *Org. Biomol. Chem.* **2007**, *5*, 54-57.

580. Andrews, B.; Schoenfish, A.; Roy, M.; Waldo, G.; Jennings, P., The rough energy landscape of superfolder GFP is linked to the chromophore. *J. Mol. Biol.* **2007**, *373*, 476-490.

581. Peterson, F.; Mason, R.; Hovsepian, J.; Holtzman, J., Oxygen-sensitive and oxygen-insensitive nitro-reduction by E. coli and rat hepatic microsome. *J. Biol.Chem.* **1979**, *254*, 4009-4014.

582. Meyer, D.; Witholt, B.; Schmid, A., Suitability of recombinant Escherichia coli and Pseudomonas putida strains for selective biotransformation of m-nitrotoluene by xylene monooxygenase. *Appl.Environ. Microbiol.* **2005**, *71*, 6624-6632.

583. Johansson, E.; Parkinson, G.; Denny, W.; Neidle, S., Studies on the nitroreductase prodrug-activating system. Crystal structures of complexes with the inhibitor dicoumarol and dinitrobenzamide prodrugs and of the enzyme active form. *J. Med. Chem.* **2003**, *46*, 4009-4020.

584. West, K.; Bake, K.; Otto, S., Dynamic disulfide combinatorial libraries of cages in water. *Org. Lett.* **2005**, *7*, 2615-2618.

585. Beak, P.; Covington, J.; White, J., Quantitative model of solvent effects on hydroxypyridine-pyridone and mercaptopyridine-thiopyridone equilibria- correlation with reaction-field and hydrogen-bonding effects. *J.Org. Chem.***1980**, *45*, 1347-1353.

586. Moran, D.; Sukcharoenphon, K.; Puchta, R.; Schaefer, H.; Schleyer, P.; Hoff, C., 2-pyridinethiol/2-pyridinethione tautomeric equilibrium. A comparative experimental and computational study. *J. Org. Chem.* **2002**, *67*, 9061-9069.

587. Dugave, C.; Kessler, P., Synthesis of activated disulfide adducts containing a 4-Diazocyclohexa-2,5-dienone precursor for photoaffinity labelling. *Tetrahedron Lett.* **1994**, *35*, 9557-9560.

588. Kolb, H.; Sharpless, K., The growing impact of click chemistry on drug discovery. *Drug Discov Today* **2003**, *8*, 1128-37.

589. Moses, J.; Moorhouse, A., The growing applications of click chemistry. *Chem Soc Rev* **2007**, *36*, 1249-62.

590. Harris, R. r.; Stabler, R.; Repke, D.; Kress, J.; Walker, K.; Martin, R.; Brothers, J.; Ilnicka, M.; Lee, S.; Mirzadegan, T., Highly potent, non-basic 5-HT6 ligands. Site mutagenesis evidence for a second binding mode at 5-HT6 for antagonism. *Bioorg Med Chem Lett* **2010**, *20*, 3436-40.

591. Elangovan, A.; Yang, S.; Lin, J.; Kao, K.; Ho, T., Synthesis and electrogenerated chemiluminescence of donor-substituted phenylquinolinylethynes and phenylisoquinolinylethynes: effect of positional isomerism. *Org Biomol Chem* **2004**, 2, 1597-602.

592. Inouye, S.; Shomura, T.; Tsuruoka, T.; Ogawa, Y.; Watanabe, H.; Yoshida, J.; Niida, T., L- Beta-(5-Hydroxy-2-Pyridyl)-Alanine and L-Beta-(3-Hydroxyureido)-Alanine fromStreptomyces. *Chem.Pharm. Bull.* **1975**, *23*, 2669-2677.

593. Monden, Y.; Hamano-Takaku, F.; Shindo-Okada, N.; Nishimura, S.; Maruta, H., Azatryosine - Mechanism of action for conversion of transformed phenotype to normal. *Anticancer Molecules: Structure, Function, and Design* **1999**, *886*, 109-121.

594. Campa, M.; Glickman, J.; Yamamoto, K.; Chang, K., The antibiotic azatyrosine supresses progesterone or [VAL(12)]P21 HA-RAS insulin-like growth factor -I-induced germinal vesicle breakdown and tyrosine phosphorylation of Xenopus mitogen-activated protein-kinase in oocytes. *Proc.Natl. Acad. Sci. U.S.A.* **1992**, *89*, 7654-7658.

595. Hamano-Takaku, F.; Iwama, T.; Saito-Yano, S.; Takaku, K.; Monden, Y.; Kitabatake, M.; Soll, D.; Nishimura, S., A mutant Escherichia coli tyrosyl-tRNA synthetase utilizes the unnatural amino acid azatyrosine more efficiently than tyrosine. *J. Biol. Chem.* **2000**, *275*, 40324-40328.

596. Barone, V.; Adamo, C., Proton-transfer in excited electronic states-environmental effects on tautomerization of 2-Pyridone *J. Photochem. Photobiol.Chem.* **1994**, *80*, 211-219.

597. Burke, T. J.; Yao, Z.; Ye, B.; Miyoshi, K.; Otaka, A.; Wu, L.; Zhang, Z., Phospho-Azatyrosine, a less effective protein-tyrosine phosphatase substrate than phosphotyrosine. *Bioorg Med Chem Lett* **2001**, *11*, 1265-8.

598. Abou-Zied, O.; Al-Shihi, O.; Achilefu, S.; Bornhop, D.; Raghavachari, R., Steady-state and time-resolved fluorescence investigation of 2-pyridone and 3-pyridone in solution and their specific binding to human serum albumin. *Mol. Prob. Biomed. Appl. Ii* **2008**, 6867.

599. Field, M.; Hillier, I., Nondissociative proton-transfer in 2-Pyridone-2-Hydroxypyridine - an abinitio molecular-orbital study. *J. Chem. Soc.-Perkin Trans. 2* **1987**, 617-622.

600. Chou, P.; Wei, C.; Hung, F., Conjugated dual hydrogen bonds mediating 2-pyridone/2-hydroxypyridine tautomerism. *J. Phy. Chem. B* **1997**, *101*, 9119-9126.

601. Watkins, E.; Phillips, R., Inhibition of tyrosine phenol-lyase from Citrobacter freundii by 2-azatyrosine and 3-azatyrosine. *Biochemistry* **2001**, *40*, 14862-8.

602. Izawa, M.; Takayama, S.; Shindo-Okada, N.; Doi, S.; Kimura, M.; Katsuki, M.; Nishimura, S., Inhibition of chemical carcinogenesis in vivo by azatyrosine. *Cancer Res* 1992, 52, 1628-30.

603. Benoit, R.; Eiseman, J.; Jacobs, S.; Kyprianou, N., Reversion of human prostate tumorigenic growth by azatyrosine. *Urology* **1995**, *46*, 370-377.

604. Casasnovas, R.; Frau, J.; Ortega-Castro, J.; Salva, A.; Donoso, J.; Munoz, F., Absolute and relative pK(a) calculations of mono and diprotic pyridines by quantum methods. *J. Mol. Str.-Theochem* **2009**, *912*, 5-12.

605. Loppinet-Serani, A.; Charbonnier, F.; Rolando, C.; Huc, I., Role of lactam vs. lactim tautomers in 2(1H)-pyridone catalysis of aromatic nucleophilic substitution. *J. Chem. Soc.-Perkin Trans. 2* **1998**, 937-942.

606. Wu, D.; Liu, L.; Liu, G.; Jia, D., Ab initio/DFT and AIM studies on dual hydrogen-bonded complexes of 2-hydroxypyridine/2-pyridone tautomerism. *J. Phy. Chem. A.* **2007**, *111*, 5244-5252.

607. Rawson, J.; Winpenny, R., The coordination chemistry of chemistry of 2-Pyridone and its derivatives. *Coord. Chem. Rev.* **1995**, *139*, 313-374.

608. Yang, M.; Szyc, Ł.; Dreyer, J.; Nibbering, E. T.; Elsaesser, T., The hydrogen-bonded 2-pyridone dimer model system. 2. Femtosecond mid-infrared pump-probe study. *J Phys Chem A* **2010**, *114*, 12195-201.

609. Germain, H.; Harris, C.; Renaud, F.; Warin, N., Efficient Large-Scale Synthesis of Boc-L-azatyrosine. *Syn. Comm.* **2009**, 523-530.

610. Myers, A.; Gleason, J., A Practical Synthesis of L-Azatyrosine. *J Org Chem* **1996**, *61*, 813-815.

611. Mayer, A.; Leumann, C. J., Pyrrolidino DNA with Bases Corresponding to the 2-Oxo Deletion Mutants of Thymine and Cytosine: Synthesis and Triplex-Forming Properties. *Eur. J. Org. Chem* **2007**, 4038-4049.

612. Sugahara, M.; Moritani, Y.; Kuroda, T.; Kondo, K.; Shimadzu, H.; Ukita, T., An efficient synthesis of the anti-asthmatic agent T-440: A selective N-alkylation of 2-pyridone. *Chem. Pharm. Bull.* **2000**, *48*, 589-591.

613. Mills, S.; Stanton, C.; Hill, C.; Ross, R. P., New developments and applications of bacteriocins and peptides in foods. *Annu Rev Food Sci Technol* **2011**, *2*, 299-329.

614. Fjell, C.; Hiss, J.; Hancock, R.; Schneider, G., Designing antimicrobial peptides: form follows function. *Nat. Rev. Drug Discov.* **2012**, *11*, 37-51.

615. Oh, J.; Lee, K., Synthesis of novel unnatural amino acid as a building block and its incorporation into an antimicrobial peptide. *Bioorg. Med. Chem.* **1999**, *7*, 2985-2990.

616. Brubaker, P.; Drucker, D., Minireview: Glucagon-like peptides regulate cell proliferation and apoptosis in the pancreas, gut, and central nervous system. *Endocrinology* **2004**, *145*, 2653-2659.

617. Choi, K.; Swierczewska, M.; Lee, S.; Chen, X., Protease-Activated Drug Development. *Theranostics* **2012**, *2*, 156-178.

618. Lovshin, J.; Huang, Q.; Seaberg, R.; Brubaker, P.; Drucker, D., Extrahypothalamic expression of the glucagon-like peptide-2 receptor is coupled to reduction of glutamate-induced cell death in cultured hippocampal cells. *Endocrinology* **2004**, *145*, 3495-3506.

619. Pattabiraman, V. R.; Bode, J. W., Rethinking amide bond synthesis. *Nature* **2011**, *480*, 471-9.

620. Marx, V., Watching peptide drugs grow up. *Chem. Eng. News* **2005**, 83, 17

621. Kaga, C.; Okochi, M.; Tomita, Y.; Kato, R.; Honda, H., Computationally assisted screening and design of cell-interactive peptides by a cell-based assay using peptide arrays and a fuzzy neural network algorithm. *Biotechniques* **2008**, *44*, 393-402.

622. Salvador, L.; Elofsson, M.; Kihlberg, J., Preparation of building blocks for glycopeptide synthesis by glycosylation of Fmoc amino-acids having unprotected carboxyl groups. *Tetrahedron* **1995**, *51*, 5643-5656.

623. Craik, D.; Fairlie, D.; Liras, S.; Price, D., The Future of Peptide-based Drugs. *Chem. Biol.Drug Des.* **2013**, *81*, 136-147.

624. Albericio, F.; Kruger, H., Therapeutic peptides FOREWORD. *Future Medicinal Chemistry* **2012**, *4*, 1527-1531.

625. Mullard, A., 2012 FDA drug approvals. *Nat. Rev. Drug Discov.* **2013**, *12*, 87-90.

626. Mullard, A., 2011 FDA drug approvals The US FDA approved 30 new therapeutics last year, including 11 first-in-class agents. *Nat. Rev. Drug Discov.* **2012**, *11*, 91-94.

627. Chen, H.; Feng, Y.; Xu, Z.; Ye, T., The total syntheses and reassignment of stereochemistry of dragonamide. *Tetrahedron* **2005**, *61*, 11132-11140.

628. Wang, S.; Gisin, B.; Winter, D.; Makofske, R.; Kulesha, I.; Tzugraki, C.; Meienhofer, J., Facile synthesis of amino-acid and peptide esters under mild conditions via cesium salts. *J. Org. Chem.***1977**, *42*, 1286-1290.

629. Amblard, M.; Fehrentz, J.; Martinez, J.; Subra, G., Methods and Protocols of modern solid phase peptide synthesis. *Molecular Biotechnology* **2006**, *33*, 239-254.

630. Stalheim, L.; Ding, Y.; Gullapalli, A.; Paing, M.; Wolfe, B.; Morris, D.; Trejo, J., Multiple independent functions of arrestins in the regulation of protease-activated receptor-2 signaling and trafficking. *Mol. Pharmacol.* **2005**, *67*, 78-87.

631. Geysen, H.; Meloen, R.; Barteling, S., Use of peptide-synthesis to probe viral-antigens forepitopes to a resolution of a single amino-acid. *Proc.Natl. Acad.Sci.U.S.A.Biol. Sci.* **1984**, *81*, 3998-4002.

632. Schatz, P., Use of peptidelibraries to map the substrate-specificity of a peptide-modifying enzyme- - A 13 residue consensus peptide specifies biotinylation in E. coli.. *Bio-Technology* **1993**, *11*, 1138-1143.

633. Tatsu, Y.; Shigeri, Y.; Sogabe, S.; Yumoto, N.; Yoshikawa, S., Solid-phase synthesis of caged peptides using tyrosine modified with a photocleavable protecting group: Application to the synthesis of caged neuropeptide. *Biochem.Biophys. Res. Comm.* **1996**, *227*, 688-693.

634. Kan, T.; Tominari, Y.; Rikimaru, K.; Morohashi, Y.; Natsugari, H.; Tomita, T.; Iwatsubo, T.; Fukuyama, T., Parallel synthesis of DAPT derivatives and their gamma-secretase-inhibitory activity. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1983-1985.

635. Shpacovitch, V.; Feld, M.; Hollenberg, M.; Luger, T.; Steinhoff, M., Role of protease-activated receptors in inflammatory responses, innate and adaptive immunity. *J. Leukocyte Biology* **2008**, *83*, 1309-1322.

636. Rask-Andersen, M.; Almen, M.; Schioth, H., Trends in the exploitation of novel drug targets. *Nat. Rev. Drug Discov.* **2011**, *10*, 579-590.

637. Pham, T.; Kriwacki, R.; Parrill, A., Peptide design and structural characterization of a GPCR loop mimetic. *Biopolymers* **2007**, *86*, 298-310.

638. Marchese, A.; Paing, M.; Temple, B.; Trejo, J., G protein-coupled receptor sorting to endosomes and lysosomes. *Ann. Rev. Pharmacol. Toxicol.* **2008**, *48*, 601-629.

639. Pradhan, A.; Befort, K.; Nozaki, C.; Gaveriaux-Ruff, C.; Kieffer, B., The delta opioid receptor: an evolving target for the treatment of brain disorders. *Trends in Pharmacological Sciences* **2011**, *32*, 581-590.

640. Ramachandran, R.; Noorbakhsh, F.; DeFea, K.; Hollenberg, M., Targeting proteinase-activated receptors: therapeutic potential and challenges. *Nat. Rev. Drug Discov.* **2012**, *11*, 69-86.

641. Maryanoff, B.; Santulli, R.; McComsey, D.; Hoekstra, W.; Hoey, K.; Smith, C.; Addo, M.; Darrow, A.; Andrade-Gordon, P., Protease-activated receptor-2 (PAR-2): Structure-function study of receptor activation by diverse peptides related to tethered-ligand epitopes. *Arch. Biochem.Biophy.* **2001**, *386*, 195-204.

642. Devlin, M.; Pfeiffer, B.; Flanagan, B.; Beyer, R.; Cocks, T.; Fairlie, D., Hepta and octapeptide agonists of protease-activated receptor 2. *J Pept Sci* **2007**, *13*, 856-861.

643. Boitano, S.; Flynn, A.; Schulz, S.; Hoffman, J.; Price, T.; Vagner, J., Potent Agonists of the Protease Activated Receptor 2 (PAR(2)). *J. Med. Chem.* **2011**, *54*, 1308-1313.

644. Shigeri, Y.; Tatsu, Y.; Yumoto, N., Synthesis and application of caged peptides and proteins. *Pharmacol Ther* **2001**, *91*, 85-92.

645. Curley, K.; Lawrence, D., Light-activated proteins. *Curr Opin. Chem. Biol.* **1999**, *3*, 84-88.

646. Curley, K.; Lawrence, D., Caged regulators of signaling pathways. *Pharmacol Ther* **1999**, *82*, 347-354.

647. Marriott, G., Caged protein conjugates and light-directed generation of protein-activity-preparation, photoactivation, and spectroscopic characterization of caged G-actin conjugates. *Biochemistry* **1994**, *33*, 9092-9097.

648. Zou, K.; Miller, W.; Givens, R.; Bayley, H., Caged thiophosphotyrosine peptides. *Angew. Chem-Int. Ed.* **2001**, *40*, 3049-3051.

649. Bayley, H.; Chang, C. Y.; Miller, W. T.; Niblack, B.; Pan, P., Caged peptides and proteins by targeted chemical modification. *Methods Enzymol* **1998**, *291*, 117-35.

650. Walker, J.; Gilbert, S.; Drummond, R.; Yamada, M.; Sreekumar, R.; Carraway, R.; Ikebe, M.; Fay, F., Signaling pathways underlying eosinophil cell motility revealed by using caged peptides. *Proc. Natl. Acad. Sci U. S. A.* **1998**, *95*, 1568-1573.

651. Marriott, G.; Walker, J., Caged peptides and proteins: new probes to study polypeptide function in complex biological systems. *Trends in Plant Science* **1999**, *4*, 330-334.

652. Bohm, S.; Khitin, L.; Grady, E.; Aponte, G.; Payan, D.; Bunnett, N., Mechanisms of desensitization and resensitization of proteinase-activated receptor-2. *J. Biol. Chem.* **1996**, *271*, 22003-22016.

653. Boranic, M.; KrizanacBengez, L.; Gabrilovac, J.; Marotti, T.; Breljak, D., Enkephalins in hematopoiesis. *Biomed. Pharmacother.* **1997**, *51*, 29-37.

654. Gaveriaux-Ruff, C.; Nozaki, C.; Nadal, X.; Hever, X. C.; Weibel, R.; Matifas, A.; Reiss, D.; Filliol, D.; Nassar, M. A.; Wood, J. N.; Maldonado, R.; Kieffer, B. L., Genetic ablation of delta opioid receptors in nociceptive sensory neurons increases chronic pain and abolishes opioid analgesia. *Pain* **2011**, *152*, 1238-48.

655. Daniel, A.; Ortega, L.; Papini, M., Role of the opioid system in incentive downshift situations. *Neurobiology of Learning and Memory* **2009**, *92*, 439-450.

656. Sibinga, N.; Goldstein, A., Opioid-peptides and opioid receptors in cells of the Immune-system . *Ann. Rev. Immunol.* **1988**, *6*, 219-249.

657. Jordan, B.; Cvejic, S.; Devi, L., Opioids and their complicated receptor complexes. *Neuropsychopharmacology* **2000**, *23*, S5-S18.

658. Ohmori, H.; Fujii, K.; Sasahira, T.; Luo, Y.; Isobe, M.; Tatsumoto, N.; Kuniyasu, H., Methionine-enkephalin secreted by human colorectal cancer cells suppresses T lymphocytes. *Cancer Sci.* **2009**, *100*, 497-502.

659. Marotti, T.; Balog, T.; Mazuran, R.; Rocic, B., The role of cytokines in MET-enkephalin-modulated nitric oxide release. *Neuropeptides* **1998**, *32*, 57-62.