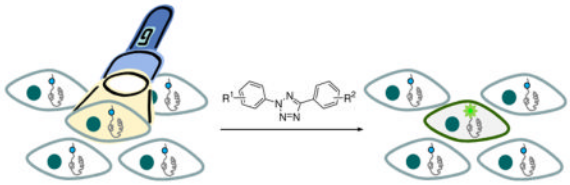
Published in final edited form as:
Acc Chem Res. 2011 September 20; 44(9): 828–839. doi:10.1021/ar200021p.

Photoinducible Bioorthogonal Chemistry: A Spatiotemporally Controllable Tool to Visualize and Perturb Proteins in Live Cells

Reyna K. V. Lim and **Qing Lin***
Department of Chemistry, State University of New York at Buffalo, Buffalo, New York 14260

Conspectus



Visualization in biology has been greatly facilitated by the use of fluorescent proteins as in-cell probes. The genes coding for these wavelength-tunable proteins can be readily fused with the DNA coding for a protein of interest, which enables direct monitoring of natural proteins in real time inside living cells. Despite their success, however, fluorescent proteins have limitations that have only begun to be addressed in the past decade through the development of bioorthogonal chemistry. In this approach, a very small bioorthogonal tag is embedded within the basic building blocks of the cell, and then a variety of external molecules can be selectively conjugated to these pre-tagged biomolecules. The result is a veritable palette of biophysical probes for the researcher to choose from.

In this Account, we review our progress in developing a photoinducible, bioorthogonal tetrazole–alkene cycloaddition reaction ("photoclick chemistry") and applying it to probe protein dynamics and function in live cells. The work described here summarizes the synthesis, structure, and reactivity studies of tetrazoles, including their optimization for applications in biology. Building on key insights from earlier reports, our initial studies of the reaction have revealed full water compatibility, high photoactivation quantum yield, tunable photoactivation wavelength, and broad substrate scope; an added benefit is the formation of fluorescent cycloadducts. Subsequent studies have shown fast reaction kinetics (up to 11.0 M^{-1} s^{-1}), with the rate depending on the HOMO energy of the nitrile imine dipole as well as the LUMO energy of the alkene dipolarophile. Moreover, through the use of photocrystallography, we have observed that the photogenerated nitrile imine adopts a bent geometry in the solid state. This observation has led to the synthesis of reactive, macrocyclic tetrazoles that contain a short "bridge" between two flanking phenyl rings.

This photoclick chemistry has been used to label proteins rapidly (within ~1 minute) both in vitro and in *E. coli*. To create an effective interface with biology, we have identified both a metabolically incorporable alkene amino acid, homoallylglycine, and a genetically encodable tetrazole amino acid, *p*-(2-tetrazole)phenylalanine. We demonstrate the utility of these two moieties, respectively, in spatiotemporally controlled imaging of newly synthesized proteins and in site-specific labeling of proteins. Additionally, we demonstrate the use of the photoclick chemistry to perturb the localization of a fluorescent protein in mammalian cells.

qinglin@buffalo.edu.

Introduction

The development of genetically encoded, wavelength-tunable fluorescent proteins (FPs) has ushered in a new era of molecular biology, allowing individual proteins to be monitored in real time in living cells and organisms.[1] While undoubtedly FPs have drastically improved our understanding of protein dynamics and function in living systems, many biological processes are not directly amenable to the genetic fusion of FPs because of the following limitations. First, because FPs are macromolecules with molecular weight typically greater than 25 kDa, fusion of FPs to a target protein may alter the target's native function.[2,3] Second, FPs are not suitable for directly reporting numerous regulatory protein posttranslational modification events such as phosphorylation, acetylation, methylation, prenylation, and glycosylation. Third, FPs are more of a stationary, observational tool rather than a perturbational tool in interrogating cellular circuitry. Thus, tools that overcome these limitations are urgently needed. Over the past decade a "bottom-up" bioorthogonal chemistry strategy has shown great promise in overcoming some of these limitations.[4] In this approach, a bioorthogonal group is built into the life's basic building blocks—amino acids, nucleotides, lipids, and sugars—as well as secondary metabolites; after their biochemical incorporation, an array of biophysical probes are then selectively conjugated to the pre-tagged biomolecules via a suitable bioorthogonal reaction.[5]

In this Account, we summarize the progress in the development of photoinducible, bioorthogonal tetrazole-alkene cycloaddition reaction, also referred to as "photoclick chemistry", and its applications in protein visualization and perturbation in live cells. We strive to provide a historical context about the incremental steps realized during the development of this reaction and the remaining challenges before its wider adoption in biology. With a heightened interest recently in using light to not only report but also *perturb* biological processes in their native environment, most notably the rise of optogenetics,[6,7] photoinducible bioorthogonal chemistry may add an invaluable tool to control defined biological events in defined cell types at defined time in intact systems.

Photoinduced Cycloaddition in Aqueous Solution

In the late 1960s, Huisgen and co-workers described the first photoinduced 1,3-dipolar cycloaddition reaction between 2,5-diphenyltetrazole (**1**) and methyl crotonate in benzene at 20 °C.[8] In their seminal study, a medium-pressure mercury lamp was used in the reaction, which led to the formation of a pair of pyrazoline regioisomers in 3:1 ratio with 78% yield (Scheme 1). Based on the stereochemistry, a concerted reaction mechanism was proposed in which upon photoirradiation, 2,5-diphenyltetrazole undergoes a facile cycloreversion reaction to release N_2 and generate *in situ* nitrile imine dipole, which then reacts with crotonate dipolarophile in a concerted manner to afford the pyrazoline cycloadducts. The presence of the short-lived nitrile imine intermediate was later established through direct spectroscopic studies *via* UV-Vis and infrared at low temperature as well as by fragmentation study of the N^{15}-labeled tetrazoles.[9] The photolysis of 2,5-diaryltetrazoles is extremely efficient under 290 nm UV irradiation with quantum yield in the range of 0.5–0.9, with electronic properties of the substituents having minimal effect.[10,11] The frontier molecular orbital calculation of the cycloaddition involving terminal alkenes indicates strong regioselectivity toward 5-substituted pyrazolines with a predominant dipole HOMO-dipolarophile LUMO interaction in the transition state.[12] A remarkable rate acceleration was observed when the cycloaddition reactions were performed in aqueous media.[13] Despite its robust mechanism, this photoinduced cycloaddition has seen very limited applications, e.g., the synthesis of benzopyrazole heterocycles[14,15] and the functionalization of polymer surfaces.[16]

Attracted by this novel mode of substrate activation, we sought to investigate whether the unique reactivity of tetrazoles could be harnessed for biological applications. To this end, 2,5-diaryltetrazoles can be readily synthesized via the Kakehi method[17] in three steps: (1) preparation of the hydrazone from aryl aldehydes and benzenesulfonylhydrazide; (2) preparation of the arene diazonium salts in situ; and (3) mixing these two components in pyridine at −20 ~ 0 °C for 3 ~ 12 hours to form the 2,5-diaryl-substituted tetrazoles (Scheme 2a). A broad range of tetrazoles have been prepared by using this procedure with overall yields of 13% to 60%.[18] In a test reaction between 2-phenyl-5-(*p*-methoxy)phenyltetrazole (**2**) and methyl methacrylate, we found that a handheld 302-nm UV lamp (UVM, 0.16 AMPS) was sufficient to drive the reaction to completion within 2 hours (Scheme 2b).[18] Notably, under this mild photoactivation condition only one regioisomers were detected by ^{1}H NMR for all the tetrazoles we tested. In addition, the reactions can be performed in a wide range of organic solvents including EtOH/H_2O (7:3) mixture.[18]

To realize the potential of this tetrazole-based cycloaddition reaction in biological systems, it is highly desirable that the reaction takes place under a brief exposure of long-wavelength UV light in order to minimize photodamage to cells. Therefore, we also screened a series of substituted diaryltetrazoles with illumination of a handheld 365-nm UV lamp (UVGL-25, 0.16 AMPS) and identified several tetrazoles containing either amino (**3** and **4**) or styryl groups (**5**) at the *para*-position of the N^2-phenyl ring, which gave moderate-to-excellent yields in the cycloaddition reactions with methyl methacrylate (Scheme 3).[19] Furthermore, two 365-nm photoactivatable diaryltetrazoles (**6** and **7** in Scheme 3) were discovered through a 'scaffold hopping' strategy in which the tetrazole moiety was allowed to "jump" from the aniline to other long-wavelength UV absorbing scaffolds such as naphthalene and coumarin.[20] Based on the X-ray structure, the long-wavelength photoactivatability of the tetrazoles was attributed to the bathochromic shift in absorbance as a result of attachment of the auxochromic group to the planar diphenyltetrazole structure.

Structure and Fate of the Photogenerated Nitrile Imine

A critical consideration in optimizing tetrazole reactivity was the electronic structure of the photogenerated, transient nitrile imines.[21] To this end, we used the photocrystallography technique and briefly exposed the tetrazole crystals to a 325-nm He-Cd laser beam at 90 K.[22] Among the six Zn-tetrazole crystals examined (Figure 1a), only Zn-tetrazole **10** crystals showed a discrete photodifference map which upon least-squares refinement gave a 13% yield of a bent nitrile imine product (Figure 1b). Concurrently, a positive electron density for escaping N_2 was also detected, providing further support for the tetrazole ring rupture in the crystal lattice. The photoreactivity of Zn-tetrazole **10** complex in crystals is consistent with the fact that tetrazole **10** has the largest void around N^3-N^4, with the nearest distance between N^3/N^4 and the surrounding atoms to be 3.32 Å and 3.21 Å, respectively (Figure 1a). In the absence of a dipolarophile, it was observed that the nitrile imine dipole underwent exclusive 1,3-addition by water to generate a hydrazonic acid intermediate which spontaneously tautomerized to yield the more stable hydrazide.[22] Based on this result, we ascribed the bent nitrile imine geometry to the 1,3-dipolar form, the major electronic structure that is likely responsible for the high reactivity of nitrile imines in aqueous medium.[23–25]

Optimization of Tetrazole Reactivity in Water

To assess the tetrazole reactivity in water, we performed a kinetic study for the photoinduced cycloaddition between a water-soluble tetrazole **14** and acrylamide in PBS buffer (Figure 2a).[26] As expected, upon 302-nm photoirradiation the transient nitrile imine intermediate was generated within 30 seconds with a first-order rate constant (k_1) of 0.14

s^{-1}; the subsequent cycloaddition with acrylamide proceeded with a second-order rate constant (k_2) of 11.0 M^{-1} s^{-1}, significantly faster than Staudinger ligation[27] and the strain-promoted azide-alkyne cycloaddition reaction.[28] However, when an unactivated alkene dipolarophile such as allyl phenyl ether was used in the cycloaddition reaction with 2-phenyl-5-(*p*-methylbenzoate)tetrazole (**15**, a precursor to compound **14**), the reaction proceeded significantly slower (k_2 = 0.00202 ± 0.00007 M^{-1} s^{-1} *vs.* 0.15 M^{-1} s^{-1} for acrylamide),[29] indicating that the rate of the cycloaddition is highly dependent on the LUMO energy of the dipolarophile.

To optimize tetrazole reactivity towards unactivated alkene dipolarophiles such as 4-penten-1-ol, we systematically tuned the HOMO energies of the nitrile imine dipoles by introducing various substituents to the phenyl rings (Figure 3a). We found that electron-donating substituents on the phenyl rings generally increase the HOMO energies of the corresponding nitrile imines, giving rise to faster cycloaddition rates.[30] For example, 2-(*p*-aniline)-5-phenyltetrazole (**16**: X = H, Y = *p*-NH_2) showed a robust cycloaddition with 4-penten-1-ol with k_2 value of 0.79 M^{-1} s^{-1}, about 200-fold faster than tetrazole **15** (X = *p*-CO_2Me, Y = H; k_2 = 0.004 M^{-1} s^{-1}). Tetrazole **16** also afforded the highest yield (~50% after only 75 sec) with the water-quenching product (~10%) and the dimerization product (~15%) to be the major side products. A plot of calculated HOMO energies, E_{HOMO}, vs. experimentally determined Log(rate) showed a good linear fit, suggesting that the rate enhancement is principally due to the HOMO-lifting effect (Figure 3b).

Recently, we also made an effort to “freeze” the photogenerated nitrile imine in the bent, reactive geometry by preparing a series of macrocyclic tetrazoles containing a short “bridge” between two flanking phenyl rings (Figure 4a) and examined their reactivities towards small-molecule alkenes and alkene-containing proteins.[31] In the test reactions with norbornene in ethyl acetate, the majority of macrocyclic tetrazoles (**17**–**20**, **22**; Figure 4a) showed higher reactivity that acyclic tetrazole **17**. Furthermore, macrocyclic tetrazole **22** showed rapid (~1 minute) functionalization of a norbornene-modified lysozyme in PBS buffer (Figure 4b). Therefore, these macrocyclic tetrazoles may potentially offer a new class of photoactivatable tetrazole reagents for bioorthogonal applications in living systems.

Photoinduced Protein Modification

The suitability of the photoinduced cycloaddition reaction for site-selective protein modification was demonstrated with a tetrazole-containing protein.[26] In this experiment, 2,5-diphenyltetrazole was first added to the C-terminus of enhanced green fluorescent protein (EGFP) through an intein-mediated chemical ligation. The resulting tetrazole-containing EGFP (EGFP-Tet) in a buffer containing 20 mM Tris, 500 mM NaCl, and 1 mM EDTA, pH 7.5, was allowed to react with 1.2 mM methyl *N*-palmityl methacrylamide under 302-nm photoirradiation for 1 minute (Figure 5a). Since the pyrazoline adduct formed is fluorescent,[32] the cycloaddition reaction can be readily monitored with in-gel fluorescence analysis. A fluorescent band was detected only for EGFP-Tet (not wild-type EGFP) and only upon UV exposure (Figure 5b), indicating that the photoinduced cycloadduct formation is very selective. Liquid chromatography-mass spectrometry analysis revealed the conversion from EGFP-Tet to EGFP-Pyr to be 52%. On the other hand, fluorescent monitoring of the cycloaddition reaction in solution revealed the appearance of an emission band for the pyrazoline adduct at around 460 nm together with a twofold increase in EGFP fluorescence intensity, which was attributed to fluorescence resonance energy transfer from pyrazoline to the EGFP fluorophore (Figure 5c). Similar level of efficiency in the photoinduced lipidation of EGFP-Tet by *N*-palmityl methacrylamide was also observed when EGFP-Tet was present in the bacterial cell lysate.[26]

Protein Visualization in Live Cells

The utility of this photoinduced cycloaddition reaction in labeling proteins selectively *in vivo* was first demonstrated in *E. coli* overexpressing an alkene-containing protein.[29] By taking advantage of the fact that the pyrazoline cycloadducts are fluorescent, we initially screened a small library of diaryltetrazoles for the selective reactions with the purified *O*-allyltyrosine-containing Z-domain proteins over the wild-type in solution. Three tetrazoles including tetrazole **15** in Figure 6a were identified that showed the selective cycloadduct formation. Subsequently, BL21(DE3) bacterial cells that overexpress either *O*-allyltyrosine-containing Z-domain proteins[33] or wild-type were incubated with 100 μM tetrazole **15** at 37 °C for 30 minutes. The cells were illuminated with a handheld 302-nm UV lamp for 4 minutes followed by overnight incubation. In the cyan fluorescent protein (CFP) channel, only the alkene-Z expressing bacteria showed strong fluorescence corresponding to the pyrazoline whereas the wild-type-Z expressing bacteria did not, indicating that the reaction is highly selective toward the genetically encoded *O*-allyltyrosine-containing protein in intact bacterial cells (Figure 6b). Moreover, in the presence of 100 μM of the more reactive tetrazole 2-(*p*-methoxy)phenyl-5-phenyltetrazole **23** (which was not included in the initial screen; structure shown in Figure 3), the *O*-allyltyrosine-containing Z-domain proteins in *E. coli* showed a rapid fluorescence development after the cells were illuminated at 302 nm for less than a minute without the need for additional incubation.[30]

A complementary strategy to genetic encoding of alkene in proteins is the metabolic approach where alkene-containing amino acids are co-translationally incorporated into proteins in a residue-specific manner.[34] For example, a number of methionine surrogates such as azidohomoalanine (AHA) and homopropargylglycine (HPG) have been successfully incorporated into proteins and subsequently employed in identifying and visualizing the newly synthesized proteins in neurons.[35,36] The main advantage of the metabolic approach is its simplicity: no genetic manipulation is required in introducing exogenous alkene groups. Nonetheless, a critical concern for visualization of exogenous alkenes in mammalian cells via photoclick chemistry is the potential interference of endogenous alkenes present in phospholipids and in lipid-derived signaling molecules such as sphingosine 1-phosphate, anandamide, linoleic acid, retinoic acid, and farnesyl pyrophosphate. Fortunately, these endogenous alkenes are predominantly internal *cis*-alkenes, which were much poorer substrates for the cycloaddition reaction as compared to the exogenous terminal alkenes.[37] Indeed, a terminal alkene, homoallylglycine (HAG), labeled proteins showed selective cycloaddition reactions with the tetrazole reagents both in the cell lysates and in live HeLa cells.[38] A tandem mass spectrometry analysis revealed that as a methionine surrogate, HAG showed an excellent co-translational activity with occupancy ranging from 42% to 85%. Importantly, substitution of Met by HAG did not appreciably alter protein function, as evidenced by both the kinetic analysis and the devoid of cytotoxic activity during cell culture.[37]

The main advantage of using an alkene reporter in tandem with photoclick chemistry for protein imaging is the potential for spatiotemporal control. To demonstrate this, the HAG-labeled HeLa cells grown on a glass cover-slip in a sealed chamber in the presence of 200 μM of tetrazole **16** (structure shown in Figure 3) were briefly (~ 5 sec) exposed to a two-photon 700 nm laser and the cellular fluorescence was recorded over a period of 1 minute with a confocal microscope equipped with a DAPI filter. We found that only the illuminated HAG-HeLa cell showed a rapid and more than twofold increase in fluorescence relative to the un-illuminated cells (Figure 7a top row; Figure 7b left plot). In contrast, only a slight increase in fluorescence was observed for the illuminated normal HeLa cell, indistinguishable from the un-illuminated cells (Figure 7a bottom row; Figure 7b, right plot). Importantly, there was no significant tetrazole- or photoirradiation-induced cellular toxicity

based on the cell viability assay. Taken together, the tandem treatment of HAG and photoclick chemistry allows facile monitoring of the newly synthesized proteins with a spatial resolution in real time without the need of additional washing step in live mammalian cells.

Perturbation of Protein Localization in Live Cells

Besides protein visualization, a distinguishing factor of bioorthogonal chemistry lies in its ability to modify and perturb proteins beyond what's known in nature. For example, covalent attachment of lipid anchors to N-Ras regulates cell signaling by controlling Ras subcellular localization; two sequential lipidations—palmitoylation at Cys-181 and farnesylation at Cys-186—are needed in order for Ras to partition into proper subcellular compartments.[39] While the affinities of the lipid anchors toward membrane structures have been determined using the lipid bilayer model,[40] it remains a challenge to visualize the dynamic, spatial effect of protein lipidation *in vivo* because of the technical difficulty in turning the lipidation enzymatic activities on and off at will.[41] With the advent of bioorthogonal chemistry, in principle it should be possible to introduce a chemically addressable bioorthogonal group into the protein of interest site-selectively, which can be subsequently manipulated *via* photoclick chemistry.

In a proof-of-concept study to validate the use of bioorthogonal reaction in perturbing protein localization,[42] enhanced green fluorescent protein (EGFP) carrying a C-terminal photoreactive tetrazole tail (**24**) was prepared. After its reactivity toward *N*-palmityl fumaric acid was confirmed *in vitro*, **24** was microinjected into HeLa cells together with Dextran-tetramethylrhodamine (MW ~70 KDa). After adding 100 μM of lipid dipolarophile, *N*-palmityl fumaric acid, to the culture medium, cells were subjected to 1-minute 302-nm UV irradiation followed by immediate image acquisitions with a confocal fluorescence microscope (Figure 8a). Punctate fluorescence was observed in the GFP channel (Figure 8b), which was absent when the lipid dipolarophile was not added (Figure 8c). Importantly, in the presence of lipid dipolarophile and before the photoirradiation, no punctate fluorescence was observed, suggesting that these intense fluorescent spots are likely the result of EGFP translocation from the cytosol to the lipid vesicles upon its chemical lipidation.

Genetic Encoding of Tetrazole Amino Acids

For efficient deployment of the photoclick chemistry to visualize and perturb biological systems, it is highly desirable that one of the reactants—tetrazole or alkene—can be genetically encoded into the target protein.[43] While several alkene amino acids have already been incorporated into proteins site-selectively in *E. coli*[33,44,45] and *Saccharomyces cerevisiae*,[46] it was only very recent that we reported the synthesis and evaluation of photoreactive tetrazole amino acids suitable for genetic encoding.[47] Among the six photoreactive tetrazole amino acids we designed, *p*-(2-tetrazole)phenylalanine (*p*-Tpa, **25**) resembles tyrosine most closely and exhibits good reactivity toward dimethyl fumarate in PBS buffer. By screening a *Methanococcus jannaschii* tyrosyl-tRNA synthetase (*Mj*TyrRS) library, we successfully identified an orthogonal aminoacyl-tRNA synthetase that specifically charges *p*-Tpa into myoglobin in response to TAG codon in *E. coli* (Figure 9a).[48] Subsequently, we demonstrated that *p*-Tpa could serve as a bioorthogonal chemical "handle" for fluorescent labeling of *p*-Tpa-encoded myoglobin via the photoclick chemistry, though a shorter 254-nm UV light is needed OR has to be employed because of the monoaryltetrazole structure of **25**. Moreover, X-ray structure of the evolved *p*-Tpa-specific aminoacyl-tRNA synthetase in complex with *p*-Tpa was solved, providing a structural basis of the altered substrate specificity (Figure 9b). Specifically, Tyr32Leu and Asp158Gly

mutations appear to be most critical in recognizing *p*-Tpa over tyrosine because these two residues in the wild-type *Mj*TyrRS are engaged in hydrogen bonding with tyrosine hydroxyl group and also occupy the tetrazole space.

Other Photoinducible Bioorthogonal Reactions and Potential Applications

In an effort to harness other photogenerated reactive species for bioorthogonal chemistry applications, we recently reported a photoinduced azirine-alkene cycloaddition reaction for efficient protein conjugation at neutral pH and room temperature in biological medium.[49] It has been long known that highly reactive nitrile ylides can be generated photochemically *via* ring opening of 2*H*-azirines and that these nitrile ylides react spontaneously with electron-deficient alkene dipolarophiles to produce Δ^1-pyrrolines.[50] After examining a series of substituted diphenylazirines, we found that a nitro-substituted azirine reacts rapidly and cleanly with dimethyl fumarate in acetonitrile/PBS buffer (1:1), giving rise to a pair of pyrroline cycloadducts (endo:exo ≈ 7:3) in less than 2 minutes. No nitrile ylide intermediate was detected by HPLC, suggesting that the cycloaddition is extremely fast and that the azirine ring-opening is likely the rate-determining step. Furthermore, photoirradiation of a mixture of an azirine-containing lysozyme (Lyso-Azi) and 100 equiv of dimethyl fumarate-linked monodisperse polyethylene glycol (mPEG-fumarate **26**, MW ≈ 5 kDa) in PBS buffer at room temperature with a 302-nm handheld UV lamp for 2 minutes led to the formation of lysozyme-mPEG conjugate linked by the pyrroline ring (Lyso-Pyr-mPEG) in 41% yield (Scheme 3).

While the focus of this Account is on the use of photoinducible bioorthogonal reactions to visualize and perturb proteins in cellular systems, it should be noted that bioorthogonal reactions can also be used to modify peptides in solution in order to fine-tune their physicochemical properties. For example, when two cognate bioorthogonal groups are appended to the peptide side chains, selective and highly efficient side chain cross-linking can be achieved.[51–53] Indeed, we have shown that the photoinduced tetrazole-alkene cycloaddition reaction can be applied to short helical peptides to reinforce peptide helicity, increase cellular uptake, and improve their biological activity.[54,55]

Conclusions and Remaining Challenges

Compared to other bioorthogonal reactions, the major advantage of photoclick chemistry is its light inducibility: the use of light confers a spatiotemporal control over the reaction initiation in living systems. With the structural and reactivity insights of the transient nitrile imines, the optimized tetrazoles now permit the direct visualization of the alkene-tagged proteins in live cells with single-cell resolution. In conjunction with the site-specific introduction of the tetrazoles into the target proteins, photoclick chemistry has a potential in functionally mimicking protein posttranslational modifications such as lipidation inside living cells. While either tetrazoles or alkenes can serve as bioorthogonal reporters with their respective reaction partners as photo-modification reagents, in general the biochemical incorporation of alkenes into biomolecules is more straightforward (due to their smaller sizes) and less likely to alter native function of the target biomolecules. With this reaction sequence, alternative photoclick reagents, e.g., more efficient two-photon-activatable tetrazoles, can be readily employed to further improve the spatiotemporal resolution without the need of re-engineering the biochemical pathways.

Taking lessons from the development of fluorescent proteins, the challenges in applying the photoinducible bioorthogonal chemistry in biology are threefold. First, there is a continuing need to optimize the properties of the tetrazole and alkene reagents, e.g., faster reaction kinetics, increased bioorthogonality, and more robust photophysical properties of the

resulting pyrazoline cycloadducts. Second, more efficient delivery systems, e.g., genetic and biochemical methods for encoding the reactive tetrazole and alkene structures into proteins, need to be developed and standardized for the wider adoption of this unique chemical tool by biologists. And finally, new reaction-driven optogenetic modalities need to be developed in order to expand the utility of this photoinducible bioorthogonal chemistry in biology beyond biomolecular visualization. With these developments, it might be possible eventually to use light to control defined molecular-scale events occurring in a specified cellular population while they remain functionally embedded within an intact tissue system with spatial and temporal resolution.

Acknowledgments

I would like to thank previous and current students, postdoctoral fellows, and collaborators who have contributed to this research. The described work was supported by the Petroleum Research Fund (PRF# 45503-G4), National Institutes of Health (R01GM 085092), and New York State Center of Excellence in Bioinformatics and Life Sciences.

References

1. Tsien RY. Constructing and exploiting the fluorescent protein paintbox (Nobel Lecture). Angew Chem Int Ed. 2009; 48:5612–5626.
2. Moritz OL, Tam BM, Papermaster DS, Nakayama T. A Functional Rhodopsin-Green Fluorescent Protein Fusion Protein Localizes Correctly in Transgenic Xenopus laevis Retinal Rods and Is Expressed in a Time-dependent Pattern. J Biol Chem. 2001; 276:28242–28251. [PubMed: 11350960]
3. Marsh DR, Holmes KD, Dekaban GA, Weaver LC. Distribution of an NMDA receptor:GFP fusion protein in sensory neurons is altered by a C-terminal construct. J Neurochem. 2001; 77:23–33. [PubMed: 11279258]
4. Sletten EM, Bertozzi CR. Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. Angew Chem Int Ed Engl. 2009; 48:6974–6998. [PubMed: 19714693]
5. Lim RK, Lin Q. Bioorthogonal chemistry: recent progress and future directions. Chem Commun. 2010; 46:1589–1600.
6. Deisseroth K. Optogenetics. Nat Methods. 2011; 8:26–29. [PubMed: 21191368]
7. Toettcher JE, Voigt CA, Weiner OD, Lim WA. The promise of optogenetics in cell biology: interrogating molecular circuits in space and time. Nat Methods. 2011; 8:35–38. [PubMed: 21191370]
8. Clovis JS, Eckell A, Huisgen R, Sustmann R. 1.3-Dipolare Cycloadditionen XXV.Der Nachweis des freien Diphenylnitrilimins als Zwischenstufe bei Cycloadditionen. Chem Ber. 1967; 100:60–70.
9. Toubro NH, Holm A. Nitrilimines. J Am Chem Soc. 1980; 102:2093–2094.
10. Weinberg P, Csongar C. Bis-2H-tetrazoles VII: Quantum yields of bis-2H-tetrazoles and studies of thermal consecutive reactions of bis-nitrile imines by time-resolved spectroscopy. J Photochem Photobiol A Chemistry. 1989; 50:11–30.
11. Lohse V, Leihkauf P, Csongar C, Tomaschewski G. Photochemistry of diarylsubstituted 2H-tetrazole. VI. Quantum yields of the photolysis of diarylsubstituted 2H-tetrazoles. J prakt Chemie. 1988; 330:406–414.
12. Houk KN, Sims J, Watts CR, Luskus LJ. The origin of reactivity, regioselectivity, and periselectivity in 1,3-dipolar clcloadditions. J Am Chem Soc. 1973; 95:7301–7315.
13. Molteni G, Orlandi M, Broggini G. Nitriimine cycloaddition in aqueous media. J Chem Soc Perkin Trans. 2000; 1:3742–3745.
14. Padwa A, Nahm S, Sato E. Intramolecular 1,3-dipolar cycloaddition reactions of alkenyl-substituted nitrile imines. J Org Chem. 1978; 43:1664–1671.
15. Meier H, Heimgartner H. Intramolecular 1,3-dipolar cycloadditions of diaryl-nitrile-imines generated from 2,5-diaryltetrazoles. Helv Chim Acta. 1985; 68:1283–1300.

16. Darkow R, Yoshikawa M, Kitao T, Tomaschewski G, Schellenberg J. Photomodification of a poly(acrylonitrile-co-butadiene-co-styrene) containing diaryltetrazolyl groups. J Polymer Sci A Polymer Chem. 1994; 32:1657–1664.
17. Ito S, Tanaka Y, Kakehi A, Kondo K. Facile Synthesis of 2,5-Disubstituted Tetrazoles by Reaction of Phenylsulfonylhydrazones with Arenediazonium Salts. Bull Chem Soc Jpn. 1976; 49:1920–1923.
18. Wang Y, Rivera Vera CI, Lin Q. Convenient Synthesis of Highly Functionalized Pyrazolines via Mild, Photoactivated 1,3-Dipolar Cycloaddition. Org Lett. 2007; 9:4155–4158. [PubMed: 17867694]
19. Wang Y, Hu WJ, Song W, Lim RKV, Lin Q. Discovery of Long-Wavelength Photoactivatable Diaryltetrazoles for Bioorthogonal 1,3-Dipolar Cycloaddition Reactions. Org Lett. 2008; 10:3725–3728. [PubMed: 18671406]
20. Yu Z, Ho LY, Wang Z, Lin Q. Discovery of New Photoactivatable Diaryltetrazoles for Photoclick Chemistry via ‘Scaffold Hopping’. Bioorg Med Chem Lett. 2011 in press.
21. Reau R, Veneziani G, Dahan F, Bertrand G. A Straightforward Synthesis of Nitrilimines; X-Ray Crystal Structure of a Nitrilimine Stabilized by Non-Heteroatom Substituents. Angew Chem Int Ed. 1992; 31:439–440.
22. Zheng SL, Wang Y, Yu Z, Lin Q, Coppens P. Direct observation of a photoinduced nonstabilized nitrile imine structure in the solid state. J Am Chem Soc. 2009; 131:18036–18037. [PubMed: 19928921]
23. Ess DH, Houk KN. Distortion/interaction energy control of 1,3-dipolar cycloaddition reactivity. J Am Chem Soc. 2007; 129:10646–10647. [PubMed: 17685614]
24. Ess DH, Houk KN. Theory of 1,3-dipolar cycloadditions: distortion/interaction and frontier molecular orbital models. J Am Chem Soc. 2008; 130:10187–10198. [PubMed: 18613669]
25. Schoenebeck F, Ess DH, Jones GO, Houk KN. Reactivity and regioselectivity in 1,3-dipolar cycloadditions of azides to strained alkynes and alkenes: a computational study. J Am Chem Soc. 2009; 131:8121–8133. [PubMed: 19459632]
26. Song W, Wang Y, Qu J, Madden MM, Lin Q. A Photoinducible 1,3-Dipolar Cycloaddition Reaction for Rapid, Selective Modification of Tetrazole-Containing Proteins. Angew Chem Int Ed. 2008; 47:2832–2835.
27. Agard NJ, Baskin JM, Prescher JA, Lo A, Bertozzi CR. A comparative study of bioorthogonal reactions with azides. ACS Chem Biol. 2006; 1:644–648. [PubMed: 17175580]
28. Baskin JM, Prescher JA, Laughlin ST, Agard NJ, Chang PV, Miller IA, Lo A, Codelli JA, Bertozzi CR. Copper-free click chemistry for dynamic in vivo imaging. Proc Natl Acad Sci USA. 2007; 104:16793–16797. [PubMed: 17942682]
29. Song W, Wang Y, Qu J, Lin Q. Selective functionalization of a genetically encoded alkene-containing protein via “photoclick chemistry” in bacterial cells. J Am Chem Soc. 2008; 130:9654–9655. [PubMed: 18593155]
30. Wang Y, Song W, Hu WJ, Lin Q. Fast Alkene Functionalization In Vivo by Photoclick Chemistry: HOMO Lifting of Nitrile Imine Dipoles. Angew Chem Int Ed. 2009; 48:5330–5333.
31. Yu Z, Lim RK, Lin Q. Synthesis of macrocyclic tetrazoles for rapid photoinduced bioorthogonal 1,3-dipolar cycloaddition reactions. Chem Eur J. 2010; 16:13325–13329.
32. Fahrni CJ, Yang L, VanDerveer DG. Tuning the photoinduced electron-transfer thermodynamics in 1,3,5-triaryl-2-pyrazoline fluorophores: X-ray structures, photophysical characterization, computational analysis, and in vivo evaluation. J Am Chem Soc. 2003; 125:3799–3812. [PubMed: 12656613]
33. Zhang Z, Wang L, Brock A, Schultz PG. The Selective Incorporation of Alkenes into Proteins in Escherichia coli. Angew Chem Int Ed. 2002; 41:2840–2842.
34. van Hest JCM, Kiick KL, Tirrell DA. Efficient Incorporation of Unsaturated Methionine Analogues into Proteins in Vivo. J Am Chem Soc. 2000; 122:1282–1288.
35. Dieterich DC, Link AJ, Graumann J, Tirrell DA, Schuman EM. Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). Proc Natl Acad Sci USA. 2006; 103:9482–9487. [PubMed: 16769897]

36. Dieterich DC, Hodas JJ, Gouzer G, Shadrin IY, Ngo JT, Triller A, Tirrell DA, Schuman EM. In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. Nat Neurosci. 2010; 13:897–905. [PubMed: 20543841]
37. Huisgen R, Knupfer H, Sustmann R, Wallbillich G, Weberndorfer V. 1,3-Dipolar cycloadditions. XXVII. Addition of diphenylnitrilimine to nonconjugate alkenes and alkynes. Steric course of reaction, orientation, and effect of substituents. Chem Ber. 1967; 100:1580–1592.
38. Song W, Wang Y, Yu Z, Vera CI, Qu J, Lin Q. A metabolic alkene reporter for spatiotemporally controlled imaging of newly synthesized proteins in Mammalian cells. ACS Chem Biol. 2010; 5:875–885. [PubMed: 20666508]
39. Hancock JF. Ras proteins: different signals from different locations. Nat Rev Mol Cell Biol. 2003; 4:373–384. [PubMed: 12728271]
40. Silvius JR, l'Heureux F. Fluorimetric evaluation of the affinities of isoprenylated peptides for lipid bilayers. Biochemistry. 1994; 33:3014–3022. [PubMed: 8130214]
41. Rocks O, Peyker A, Kahms M, Verveer PJ, Koerner C, Lumbierres M, Kuhlmann J, Waldmann H, Wittinghofer A, Bastiaens PI. An acylation cycle regulates localization and activity of palmitoylated Ras isoforms. Science. 2005; 307:1746–1752. [PubMed: 15705808]
42. Song W, Yu Z, Madden MM, Lin Q. A bioorthogonal chemistry strategy for probing protein lipidation in live cells. Mol Biosyst. 2010; 6:1576–1578. [PubMed: 20436975]
43. Liu CC, Schultz PG. Adding new chemistries to the genetic code. Annu Rev Biochem. 2010; 79:413–444. [PubMed: 20307192]
44. Wang J, Schiller SM, Schultz PG. A biosynthetic route to dehydroalanine-containing proteins. Angew Chem Int Ed. 2007; 46:6849–6851.
45. 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; 15:1187–1197. [PubMed: 19022179]
46. Ai HW, Shen W, Brustad E, Schultz PG. Genetically encoded alkenes in yeast. Angew Chem Int Ed. 2010; 49:935–937.
47. Wang Y, Lin Q. Synthesis and Evaluation of Photoreactive Tetrazole Amino Acids. Org Lett. 2009; 10:3570–3573. [PubMed: 19637915]
48. Wang J, Zhang W, Song W, Wang Y, Yu Z, Li J, Wu M, Wang L, Zang J, Lin Q. A biosynthetic route to photoclick chemistry on proteins. J Am Chem Soc. 2010; 132:14812–14818. [PubMed: 20919707]
49. Lim RK, Lin Q. Azirine ligation: fast and selective protein conjugation via photoinduced azirine-alkene cycloaddition. Chem Commun. 2010; 46:7993–7995.
50. Padwa A, Smolanoff J. Photocycloaddition of arylazirenes with electron-deficient olefins. J Am Chem Soc. 1971; 93:548–550.
51. Blackwell HE, Grubbs RH. Highly Efficient Synthesis of Covalently Cross-Linked Peptide Helices by Ring-Closing Metathesis. Angew Chem Int Ed. 1998; 37:3281–3284.
52. Schafmeister CE, Po J, Verdine GL. An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides. J Am Chem Soc. 2000; 122:5891–5892.
53. Cantel S, Isaad Ale C, Scrima M, Levy JJ, DiMarchi RD, Rovero P, Halperin JA, D'Ursi AM, Papini AM, Chorev M. Synthesis and conformational analysis of a cyclic peptide obtained via *i* to *i*+4 intramolecular side-chain to side-chain azide-alkyne 1,3-dipolar cycloaddition. J Org Chem. 2008; 73:5663–5674. [PubMed: 18489158]
54. Madden MM, Rivera Vera CI, Song W, Lin Q. Facile synthesis of stapled, structurally reinforced peptide helices via a photoinduced intramolecular 1,3-dipolar cycloaddition reaction. Chem Commun. 2009:5588–5590.
55. Madden MM, Muppidi A, Li Z, Li X, Chen J, Lin Q. Synthesis of cell-permeable stapled peptide dual inhibitors of the p53-Mdm2/Mdmx interactions via photoinduced cycloaddition. Bioorg Med Chem Lett. 2011; 21:1472–1475. [PubMed: 21277201]

Biographies

Reyna Koreen V. Lim was born in Gamu Isabela, Philippines in 1982. She obtained her B.S. in Chemistry from the University of the Philippines Diliman in 2003 where she also worked as a chemistry instructor before coming to the State University of New York at Buffalo in 2007. She is currently a Ph.D. student under the guidance of Professor Qing Lin and her research focuses on the development of bioorthogonal reactions and their applications in biology.

Qing Lin received his B.S. in Chemistry from University of Science and Technology of China in 1994, and his Ph.D. in Organic Chemistry from Yale University in 2000 under the guidance of Professor Andrew Hamilton. He then worked in Professor Peter Schultz's lab at Scripps Research Institute as a Damon Runyon Postdoctoral Fellow. After a brief stay in industry, he joined the faculty of the State University of New York at Buffalo in 2005 where he is now an Assistant Professor of Chemistry. His research interests include the development of (i) photoinducible bioorthogonal reactions, (ii) enabling chemistries for peptide-based therapeutics, and (iii) protein-based organelles for synthetic biology applications.

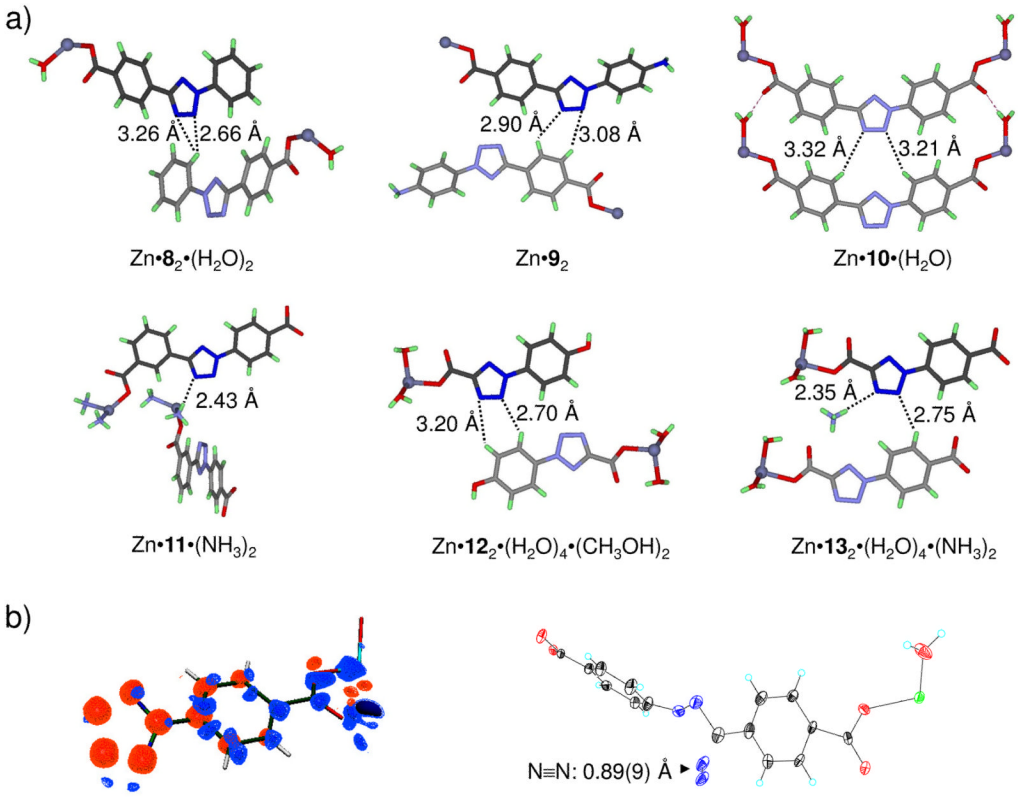


Figure 1.

Observation of a bent nitrile imine structure in the solid state: (a) Crystal structures of Zn-tetrazole complexes; Zn is shown in silver. The distances between N^3-N^4 and the nearest surrounding atoms are marked on the structures; (b) Photodifference map (left) of the Zn•**10** complex after photoirradiation based on F_o (after) - F_o (before). Blue, 2.0; light blue, 1.0; orange, −1.0; red, −2.0 e/Å^3. Only one half of the map is shown because of twofold symmetry; ORTEP representation (right) of the geometry-refined nitrile imine structure and the escaping N_2.

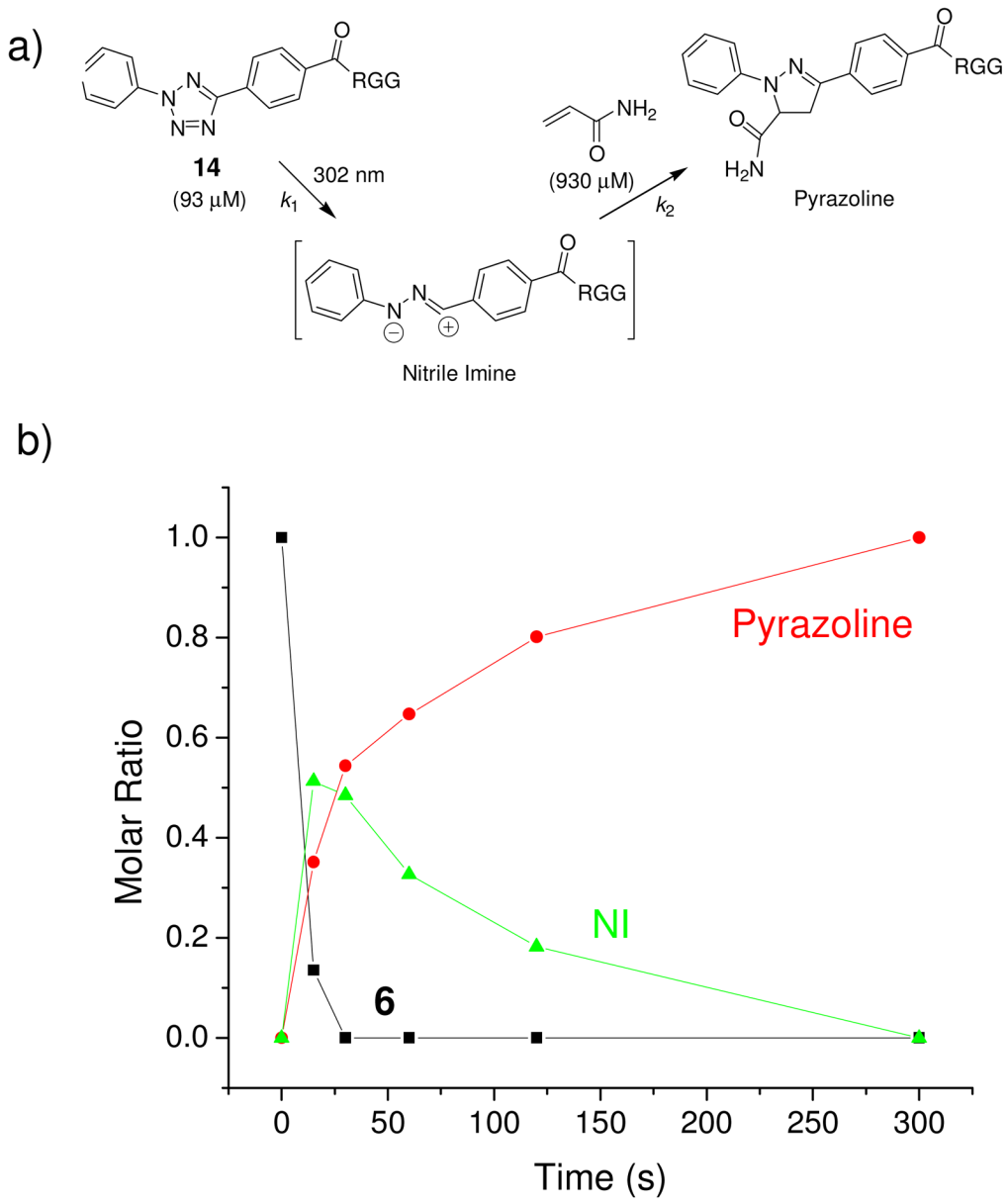


Figure 2.

Kinetic analysis of the photoinduced tetrazole-alkene cycloaddition reaction in PBS buffer at room temperature: (a) Reaction scheme showing the cycloaddition of a tetrazole-modified Arg-Gly-Gly (RGG) tripeptide **14** to acrylamide; (b) Reaction time course showing the molar ratio changes of starting material **14**, nitrile imine (NI), and product over a period of 300 seconds.

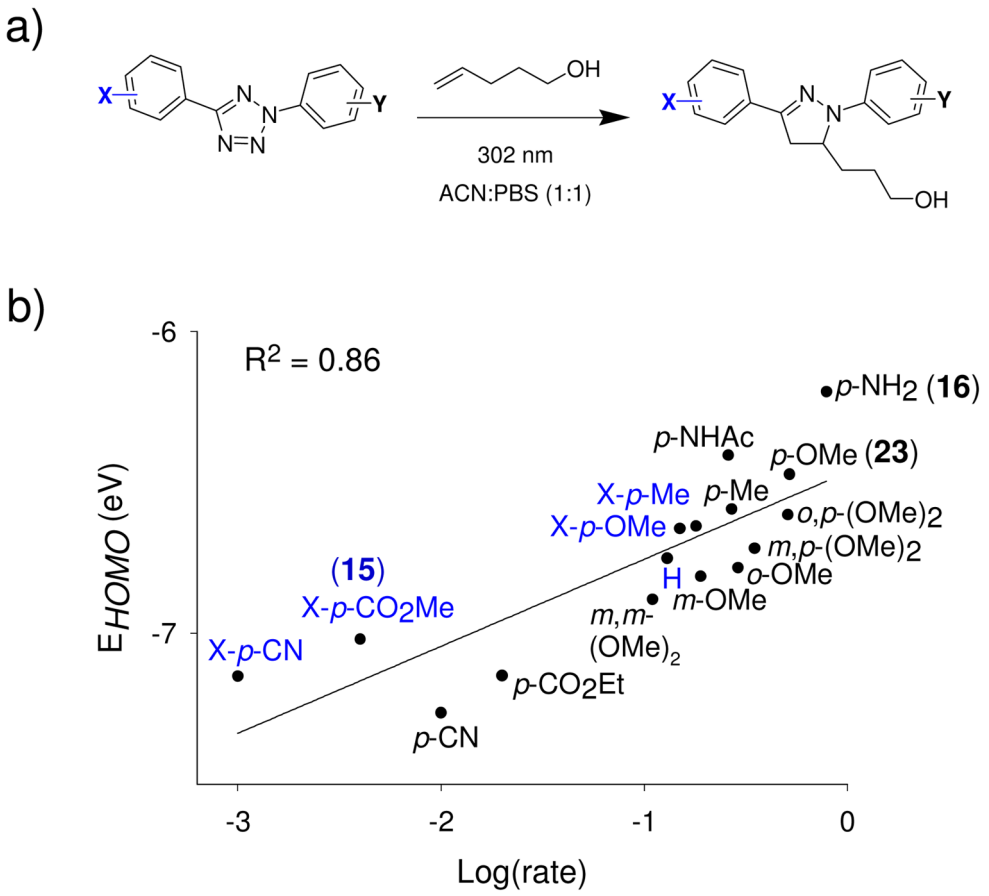


Figure 3.

The cycloaddition rates are dependent on the nitrile imine HOMO energies: (a) Scheme of the model cycloaddition reactions between substituted 2,5-diphenyltetrazoles and 4-penten-1-ol; (b) Plot of E_{HOMO} vs. Log(rate). The HOMO energies of the nitrile imines were calculated using the Hartree-Fock 3–21G model with the AM1 optimized molecular geometries. The substituents on the X-side were colored in blue while the substituents on the Y-side were colored in black.

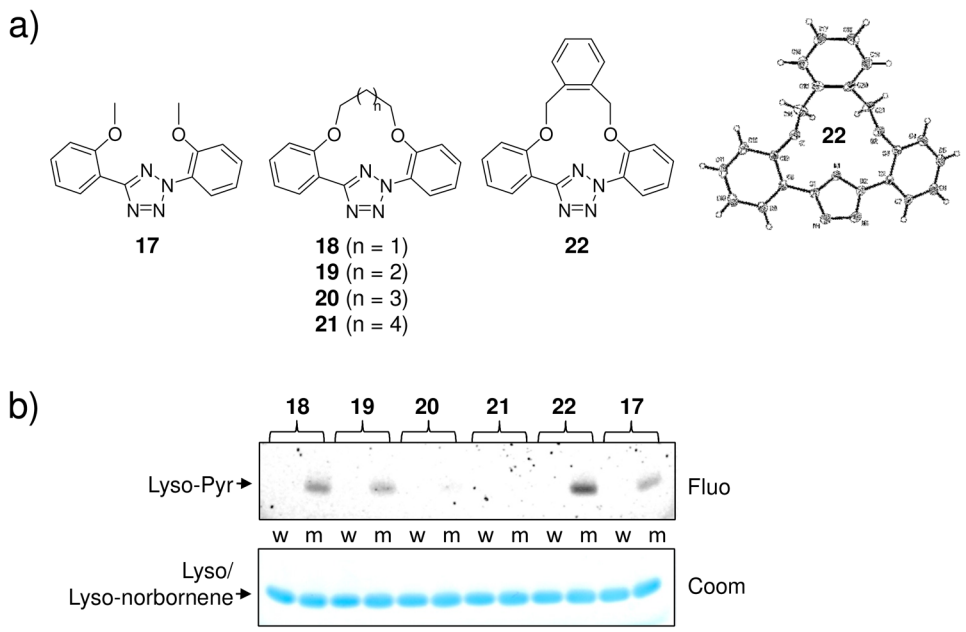


Figure 4.

Design and evaluation of macrocyclic tetrazoles: (a) Structures of macrocyclic tetrazoles. The X-ray structure of tetrazole **22** shows co-planarity of the aryl rings; (b) Photoinduced cycloaddition reaction of macrocyclic tetrazoles with norbornene-modified lysozyme (m) or wild-type lysozyme (w) in PBS buffer at 302 nm. Top panel, inverted in-gel fluorescence with λ_{ex} = 365 nm; bottom panel, Coomassie blue staining.

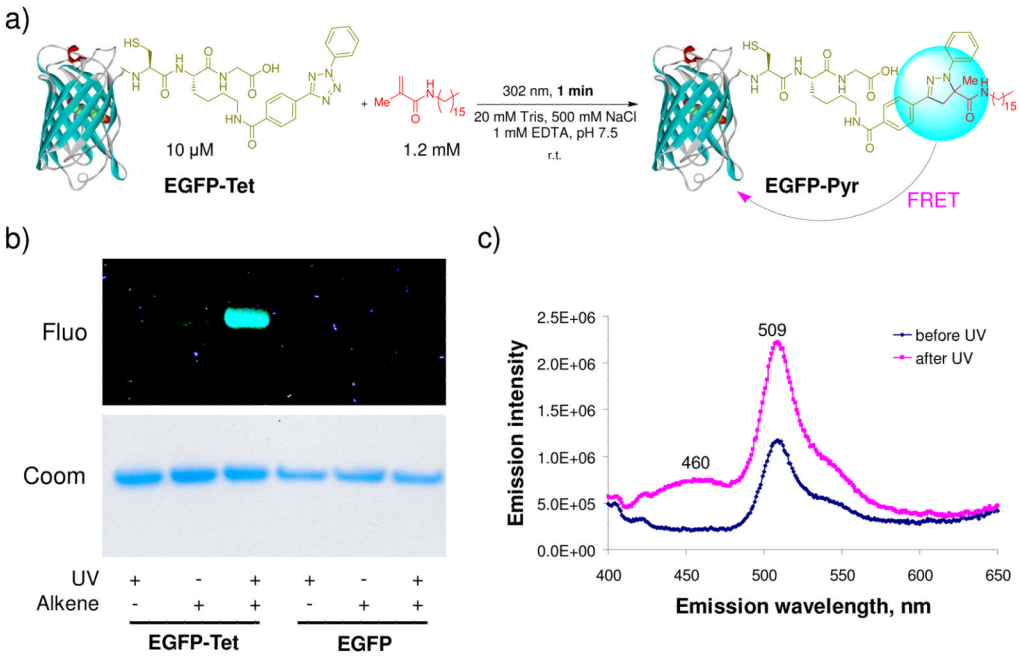


Figure 5.

Photoinduced lipidation of EGFP carrying a tetrazole motif at its C-terminus: (a) Scheme for a photoinduced lipidation by a lipid dipolarophile; (b) Fluorescent imaging (top panel, λ_{ex} = 365 nm) and Coomassie blue staining (bottom panel) of EGFP-Tet and EGFP upon photoirradiation in the presence or absence of the lipid dipolarophile. Duration of 1-min 302-nm UV irradiation was applied to the samples; (c) Fluorescence spectra of EGFP-Tet before and after UV irradiation, λ_{ex} = 370 nm.

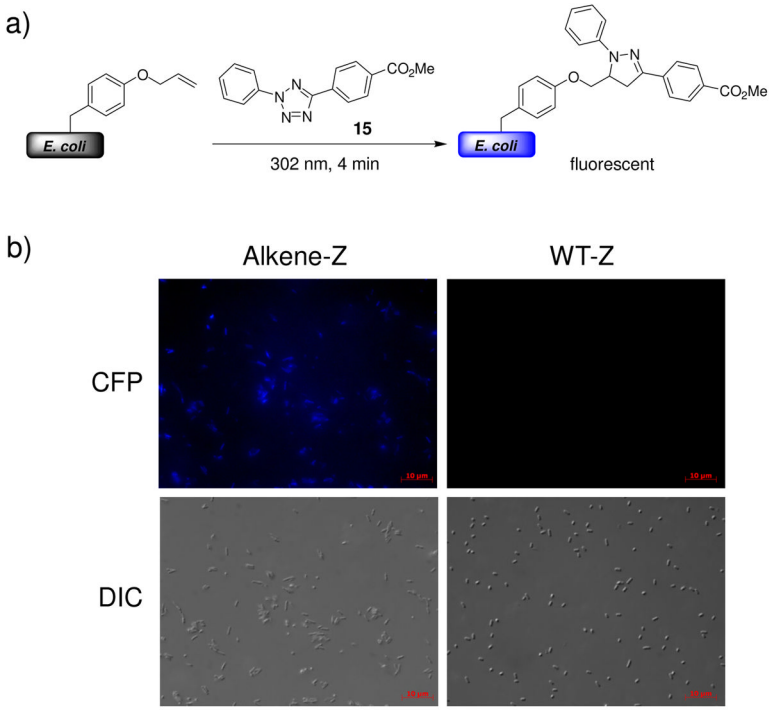


Figure 6.

Visualization of a genetically encoded *O*-allyltyrosine-containing Z-domain protein in *E. coli via* the photoinduced tetrazole-alkene cycloaddition reaction: (a) Reaction scheme; (b) CFP (cyan fluorescent protein) channel (top panels) and DIC (Differential Interference Contrast) channel (bottom panels) images of bacteria expressing either alkene-Z (left) or wt-Z (right) proteins after treatment with 100 μM of tetrazole **15**; scale bar = 10 μm.

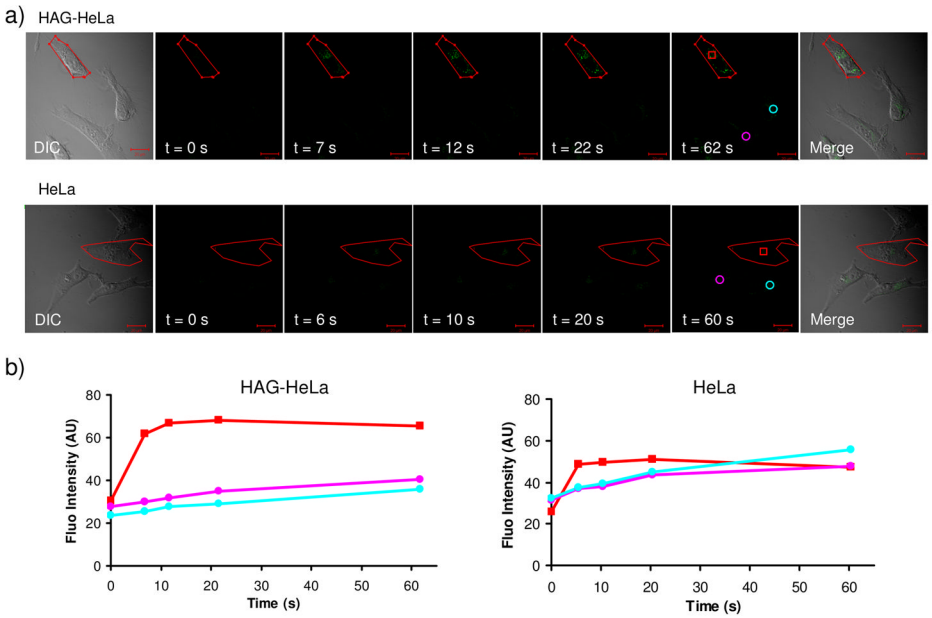


Figure 7.

Spatiotemporally controlled imaging of HAG-labeled proteins in live HeLa cells: (a) DIC (panel 1), time-lapsed fluorescence (panels 2–6), and merged DIC/fluorescence (panel 7) images of HAG-encoded HeLa cells (top row) and normal HeLa cells (bottom row) upon two-photon illumination. All cells were treated with 200 μM of tetrazole **16**. A 5-sec two-photon 700 nm laser was applied to the red encircled area in panels 1 sketched using the LSM-510 software. Scale bar = 20 μm. (b) Time courses of fluorescence development in cytosolic regions in select HeLa cells as indicated in panels 6: red square denotes the subcellular regions that were subjected to two-photon activation; cyan and purple circles denote the regions in the surrounding cells that were not subjected to two-photon activation.

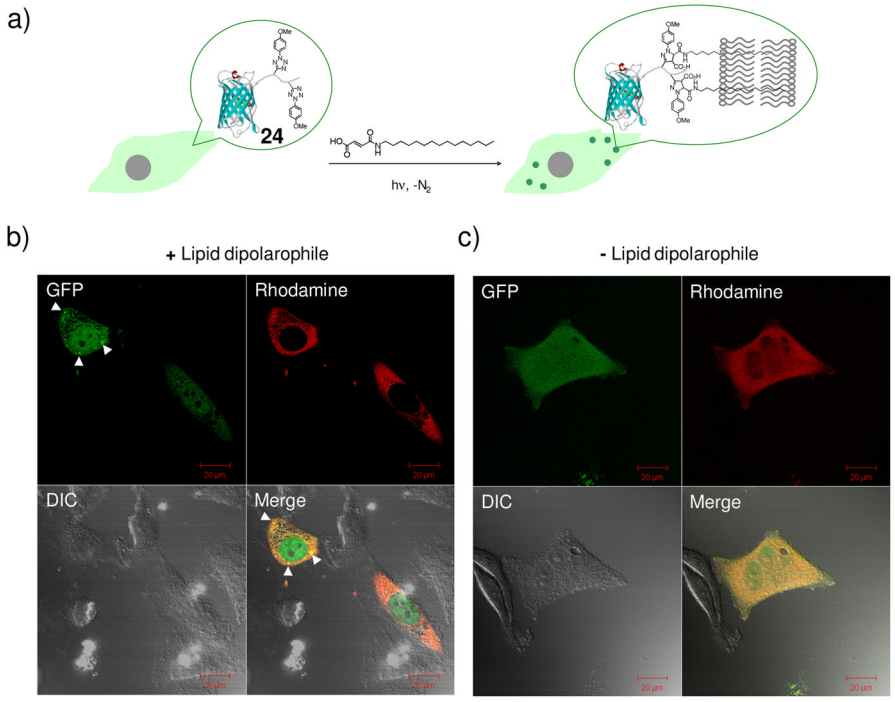


Figure 8.

Bioorthogonal chemical control of subcellular localization of **24** in live HeLa cells: (a) Scheme showing the experimental set-up; (b) Confocal micrographs of HeLa cells after the photoinduced reaction with the lipid dipolarophile showing the localization change of **24**; (c) Confocal micrographs of HeLa cells after photo illumination in the absence of lipid dipolarophile. HeLa cells were microinjected with 25 μM **24** together with 2 μM Dextran-tetramethylrhodamine (MW ~70 KDa), a red fluorescent cytosol marker. The images were acquired in separate GFP (green), DIC (Differential Interference Contrast), and rhodamine (red) channels following 1-minute 302-nm UV irradiation. The merged cell images were shown at the bottom-right panels; scale bar = 20 μm.

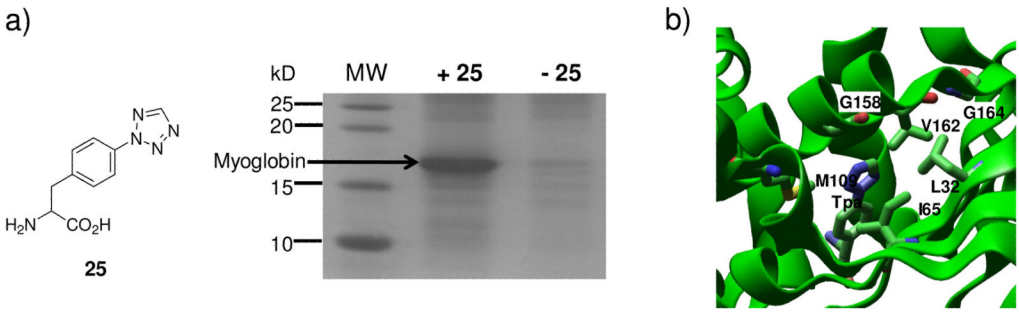
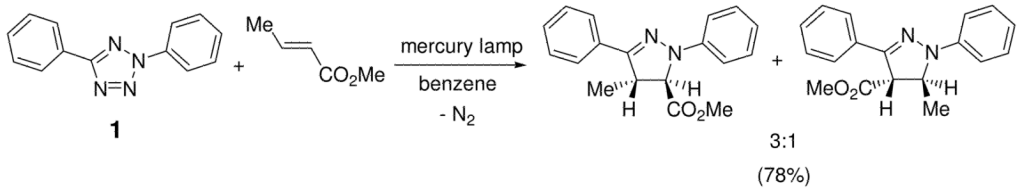


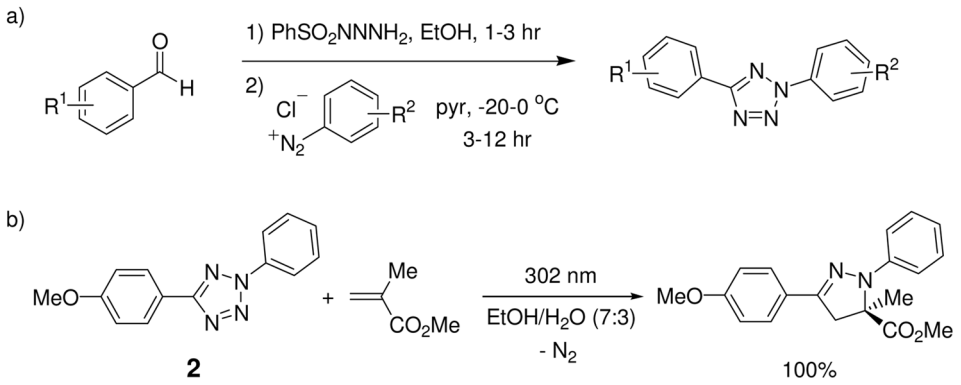
Figure 9.

Genetic incorporation of tetrazole amino acid **25** into myoglobin in *E. coli*: (a) Coomassie blue stained SDS-PAGE gel showing the expression of TAG-mutant myoglobin in the presence or absence of 1 mM *p*-Tpa (**25**); (b) A close-up view of the *p*-Tpa binding pocket with the six mutated residues of *Mj*TyrRS rendered in tube models. The complex structure has been deposited into the Protein Data Bank with access code of 3N2Y.

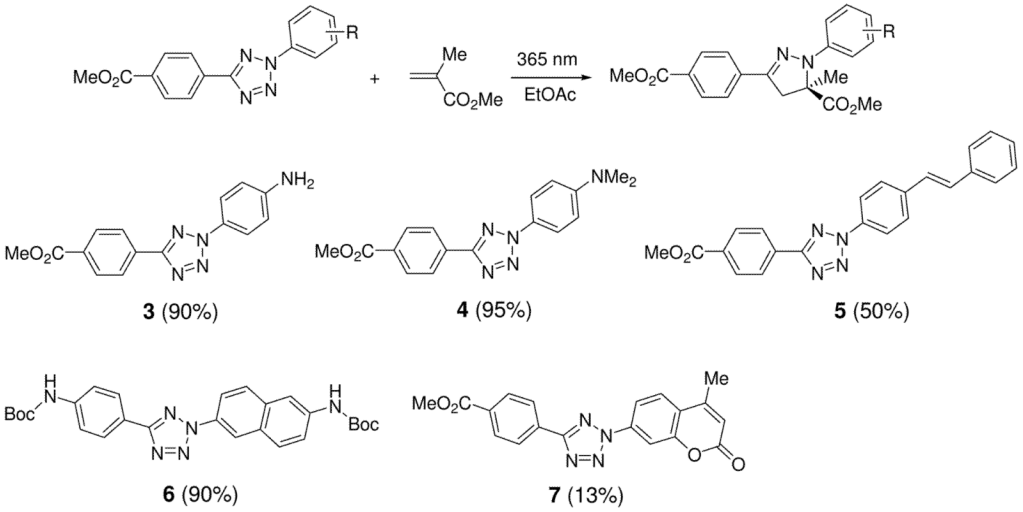


1

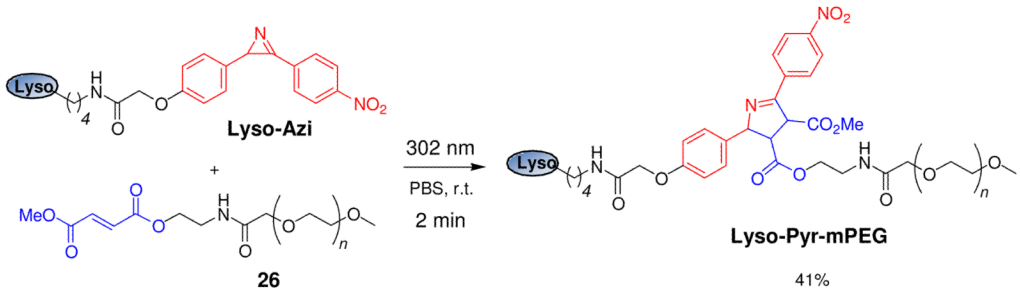
Scheme 1.



Scheme 2.



Scheme 3.



Scheme 4.
Selective PEGylation of an azirine-containing lysozyme by mPEG-fumarate **26** via a photoinduced azirine-alkene cycloaddition reaction.