

1. Title

Defining the Translatome: Bioorthogonal Noncanonical Amino Acid Tagging

2. Summary of Recent Advances

The complexity of biological systems is a source of both endless fascination and frustration. Improving the methods to isolate and visualize species of interest from biological mixtures is often the most well-recognized contribution at the chemistry-biology interface. By the clever combination of two well known phenomena – a set of simple exothermic reactions known as click chemistry, and a set of biologically inert tags known as bioorthogonal labels – a powerful new analytical technique known as bioorthogonal noncanonical amino acid tagging (BONCAT) has been developed. This technique allows for direct analysis of the proteome at the translational level, thus capturing the increasingly important details of post-translational modification. BONCAT is robust, extensible, customizable, and highly optimized; it promises to become one of the most widely used analytical techniques in proteomics.

3. Introduction

One of the greatest challenges in modern experimental biology is the difficulty in separating, visualizing, and characterizing selected components from biological systems. This problem is often at the forefront of chemical biology, as chemistry is the source of many of the most robust techniques for separation and analysis. Unfortunately, many of these techniques are not scalable enough to deal with the complexity of biological mixtures, which necessarily contain orders of magnitude more species than even the most complicated synthetic reaction. Furthermore, techniques that are sufficiently scalable often involve reagents or conditions that are too harsh to be useful in biological systems. Therefore, the development of scalable purification and labeling techniques that function under mild conditions is one of the greatest methodological challenges in chemical biology.

This search experienced an unexpected boon in 2001 when Nobel laureate and synthetic chemist K. Barry Sharpless introduced the concept of “click chemistry” [1]. Defined more by a philosophy than by a specific set of reactions, click chemistry aims to model laboratory synthetic chemistry after biological systems, which assemble incredibly complex molecules from a relatively small selection of building blocks under physiological conditions. In order for a reaction to be considered part of click chemistry, it must have very high yield, occur readily at room or slightly elevated temperature, need no solvent or only water, produce only inert byproducts, and require only minimal separation or purification. Although it was originally developed with organic

synthesis and drug development applications in mind, click chemistry quickly found favor among chemical biologists for its obvious potential to translate easily into biological systems. Research into biological click chemistry has been rapid and diverse, and has led to a number of reactions that are suitable for living systems [2•].

By far the most famous click reaction is the 1,3-dipolar cycloaddition between an azide and a terminal alkyne to yield a [1,2,3]-triazole, as seen in **Figure 1**. Originally described by the German chemist Rolf Huisgen (and named the Azide-Alkyne Huisgen Cycloaddition in his honor), the reaction goes to completion on the time scale of hours to produce >90% yields at slightly elevated temperatures in water [3]. Subsequent work by several groups led to the reaction's modern incarnation, which uses a copper(I) catalyst to regioselectively generate the 1,4 isomer at room temperature, either in aqueous solution or on a solid support [4,5].

The Huisgen reaction is of special interest to chemical biology because of its use of the azide functional group, which has the distinction of being almost completely abiotic (found only in some species of algae) [6]. However, despite its complete absence from endogenous sources, the azide group is well tolerated and nontoxic to most prokaryotes and eukaryotes, including *E. coli*, yeast, and mammals [6]. For these reasons, the azide group is typically referred to as being bioorthogonal, indicating that it does not perturb biological function in any way. This unique combination of chemical reactivity and biological inertness made the azide group an ideal candidate for introducing the power of click chemistry to the problem of separation and characterization of biological mixtures.

One of the first methods investigated for the introduction of azide groups into biological systems was the use of unnatural amino acids [7]. The most common of these has been azidohomoalanine (Aha), shown in **Figure 2**. Simple inspection reveals that Aha is closely related to methionine, possessing similar molecular geometry, atomic mass, and chemical properties of the R-group. Given these similarities, it is easy to hypothesize that translational incorporation of Aha is a feasible mechanism for introduction of azide functionality into intact proteins, and indeed that was proven to be the case in *E. coli* grown in the absence of methionine and the presence of Aha [8]. Subsequent development of this method led to it being coupled with the copper(I)-catalyzed Huisgen reaction [9], which was then followed by the discovery of a more efficient copper(I) source that allowed the reaction to proceed with high yield at the 4°C temperature that is ideal for most bioanalytical work [10]. Given these advances and others, the spectrum of bioorthogonal azide reactions has become broad [11•].

The combination of Aha translation into intact proteins followed by a Huisgen reaction on the resulting azide-tagged protein is a powerful analytical technique that would come to be called bioorthogonal noncanonical amino acid tagging (BONCAT). Given its ability to facilitate quantitative proteomic analysis at the translational rather than transcription level – that is, its ability to elucidate the “translatome” – the BONCAT technique has seen rapid and varied development in the past two years.

4. Application to Mammalian Cells

The first logical step in the development of BONCAT would be to prove its effectiveness in higher organisms, which was accomplished by Dietrich *et al.* in 2006 [12••]. Using human embryonic kidney 293 cells grown in media lacking methionine but supplemented with Aha, the authors demonstrated that incorporation of Aha into human proteins is unbiased, nontoxic, and does not noticeably perturb cellular function. Using extensive control experiments, they also demonstrated that translation with Aha is specific only for proteins synthesized after the introduction of Aha, and that its incorporation does not alter global protein synthesis kinetics. This highlights one of the greatest strengths of BONCAT, which is that it can provide temporal analysis of translational changes. By growing cells with Aha in the absence of methionine, then introducing a perturbation of interest – environmental stress, cell cycle progression, or addition of an exogenous chemical – a map can be generated showing proteins translated in response to the perturbation. This holds exciting possibilities for elucidating the mechanisms of a variety of processes, including cell cycle control and pharmaceutical action.

Dietrich *et al.* also used the copper(I)-catalyzed Huisgen reaction to link newly synthesized proteins to a biotin handle, permitting separation of the proteins by affinity chromatography for analysis by standard liquid chromatograph / tandem mass spectrometry (LC/MS/MS) techniques. The entire BONCAT process is represented graphically in **Figure 3**, which illustrates another significant strength of the technique: the modularity of the ligated alkyne. Because of the exothermicity and efficiency of the copper(I)-catalyzed Huisgen reaction, the structure of the alkyne is almost irrelevant to the yield of the reaction. This permits the extensive functionalization of the ligand seen here: a biotin handle for capture with streptavidin, a FLAG tag for visualization, and trypsin cleavage sites for release from the chromatography column. This modularity allows one to customize the alkyne to maximize ease of processing, increasing the likelihood that the method can be scaled up for high-throughput style analyses. As a result of these developments, the basic technique has been extensively optimized [13,14,15•].

5. Fluorescent Applications

A similarly logical extension of the basic BONCAT technique is to fuse it with one of the imaging or visualization methods that have become ubiquitous in modern biology. Beatty *et al.* accomplished such a fusion with fluorescence in 2006 [16•]. Using mouse embryo fibroblast cells, they first showed that Aha is not the only amino acid analogue that can be used in BONCAT by looking at incorporation of the structurally similar homopropargylglycine (Hpg). As seen in **Figure 2**, Hpg is closely related to both methionine and Aha, and it too is translated by native protein machinery with no adverse effects on cellular function or protein synthesis kinetics. The resulting Hpg-tagged proteins can be functionalized just like the Aha-tagged ones by adding an azide-bearing (rather than an alkyne-bearing) ligand. The authors were subsequently able to ligate Hpg-tagged proteins to commercially available fluorophores, thus providing the first mechanism for selectively visualizing newly synthesized proteins at a whole cell level. This method holds a significant advantage over traditional recombinant fluorescent protein strategies in that it does not require a priori knowledge of the proteins of interest. In addition, incorporation of a fluorophore into the ligated species does not necessitate removing any of the functionalities discussed earlier, thus increasing the size of the “toolkit” available for translational study. The method proved robust and popular enough to be commercialized almost immediately, and BONCAT fluorescent kits are currently available [17].

An exciting modification to this original technique was put forth by Beatty *et al.* in 2008 [18••]. By combining both the Aha and Hpg tagging methods with Rat-1 fibroblasts, they demonstrated that newly synthesized proteins could be tagged either simultaneously or sequentially by both Aha and Hpg. In addition, they also proved that the resulting species could be selectively tagged by using different colored fluorophores attached to the alkyne- and azide-bearing ligands. The resulting two-color labeling method creates many exciting possibilities, most notably the promise of even finer temporal resolution in translational profiling. By sequentially tagging with Aha then Hpg while perturbing a cell population with a stimulus of interest, then ligating the resulting proteins with two different fluorophores, information could be gained about the time course of the cell's response. Alternatively, intact cells could be visualized using the same technique to gain information about protein localization. While this technique is still in its infancy, the potentially profound applications are clear.

6. Copper-Free Variant

A potentially revolutionary development in this field is another improvement upon the Huisgen cycloaddition underlying the analytical portion of the technique. In 2007, Baskin *et al.* showed

that they could generate a nearly quantitative Huisgen reaction without a copper(I) catalyst. [19••]. By using the difluorinated cyclooctyne (DIFO) shown in **Figure 4**, they were able to perform copper-free Huisgen reactions that completed in time scales of minutes, taking advantage of the angle strain in the cyclic alkyne to “spring load” the reaction with high exothermicity. Importantly, they also demonstrated that the DIFO species was biologically inert. In addition, they coupled this discovery with the incorporation of either a fluorophore or biotin into the ligating species, proving that there is no loss of modularity or functionality in the transition from linear alkyne species to the cyclooctyne. They further demonstrated that the copper-free reaction showed an equal amount of azide-specific labeling as the catalyzed version, as well as virtually no non-specific labeling. These results were surprising because they stood in direct contrast to all previous attempts to eliminate the copper(I). Methodologically, they were exciting because the cytotoxicity of the copper(I) was the only barrier to using BONCAT for real-time analysis of dynamic translational processes (as opposed to snapshots at specific time points or averages over multiple time points). Using their copper-free Huisgen, the authors were able to perform such a dynamic BONCAT study, allowing for time lapsed photography of labeled cell surface glycan trafficking. This represents the most advanced use of BONCAT, and could potentially yield valuable data for decades to come.

7. Conclusions

The techniques and applications of BONCAT lie at the heart of chemical biology, arising from the application of chemical principles to help simplify and analyze biological systems. In this regard, it is one of the most promising analytical techniques to emerge in the past decade. While the technique is currently far from ubiquitous, its varied applications and extensive opportunity for customization make it popular among those at the forefront of proteomics. The ability to provide both qualitative (fluorescence visualization) and quantitative (isolation and LC/MS/MS) descriptions of protein changes, in environmentally-defined snap shots or in real time, all from a single technique is revolutionary. As researchers continue to optimize the reactions involved and add new functional modules, the popularity and power of the technique will inevitably climb.

8. Figures

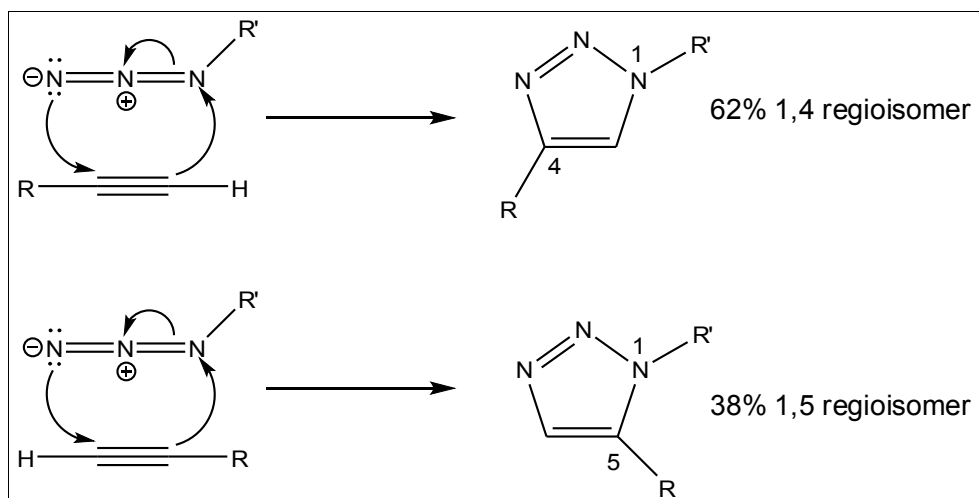


Figure 1 – Mechanism of the Azide-Alkyne Huisgen Cycloaddition

This figure shows the concerted electron-pushing mechanism for the standard Huisgen 1,3-dipolar cycloaddition between an azide and a terminal alkyne. This figure is the author's own creation.

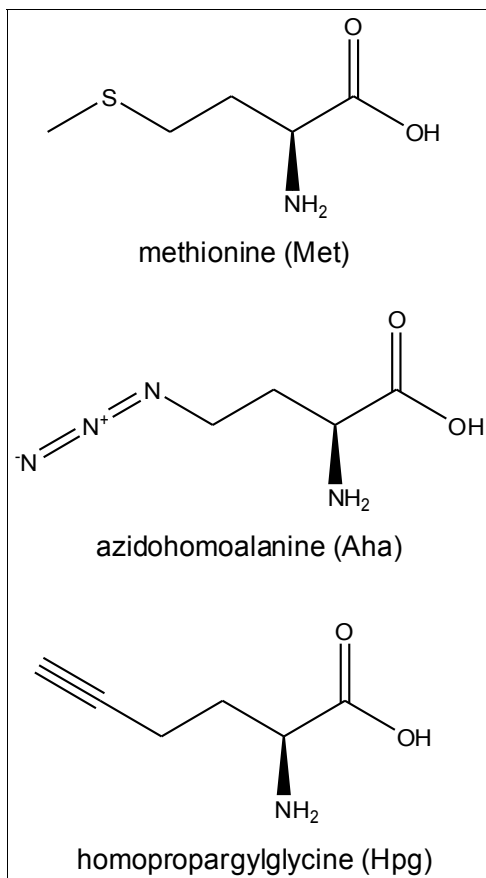


Figure 2 – Methionine, Azidohomoalanine, and Homopropargylglycine

This figure shows the structural similarity between the standard amino acid methionine and its bioorthogonally-tagged analogues, azidohomoalanine and homopropargylglycine. This figure is the author's own creation.

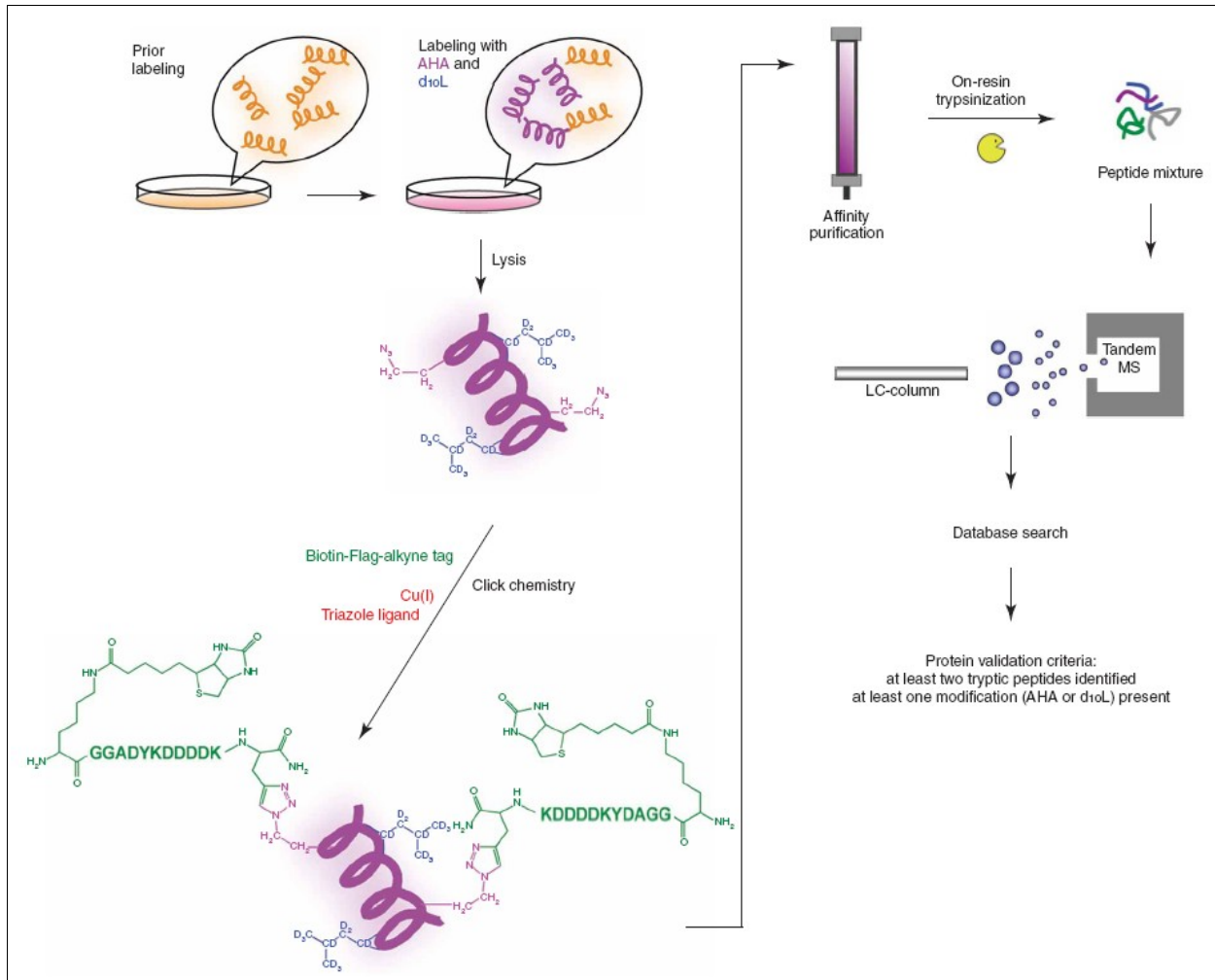


Figure 3 – Schematic of the BONCAT Protocol

This figure shows the complete process of BONCAT analysis, from labeling cells with Aha to collecting proteins with affinity chromatography and identifying them by LC/MS/MS. The deuterated leucine (d₁₀L) is added as a tracer to confirm the success of BONCAT. This figure is taken from Dietrich *et al.* [15••].

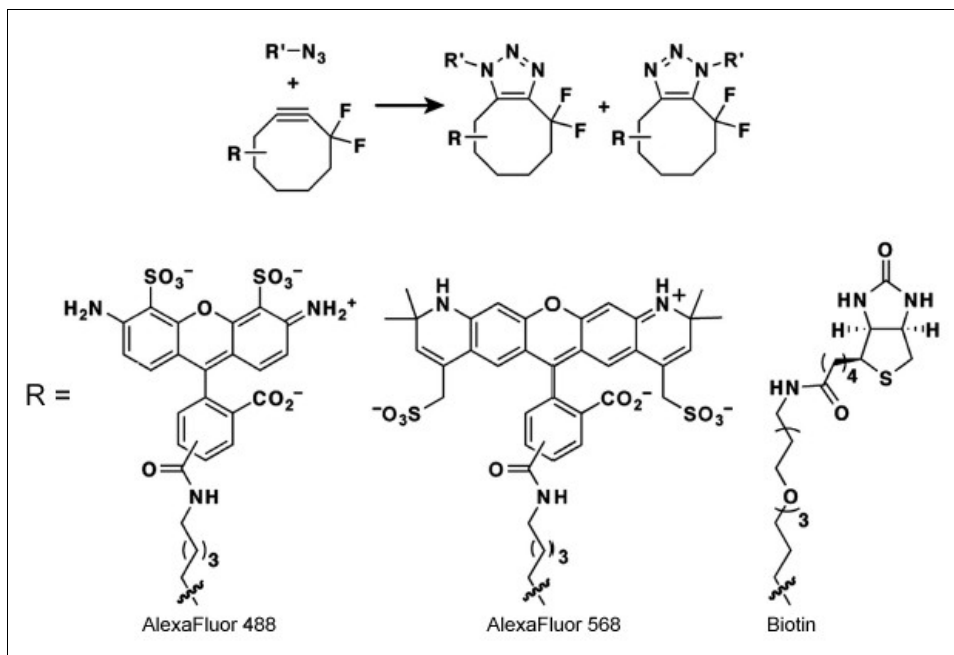


Figure 4 – Copper-free Huisgen Cycloaddition

This figure shows the structure of the difluorinated cyclooctyne used in the copper-free Huisgen cycloaddition, as well as the structures of the fluorophores and affinity tags that were incorporated. This figure is adapted from Baskin *et al.* [19••].

9. References

- of special interest
- of outstanding interest

1. Kolb HC, Finn MG, Sharpless KB: **Click chemistry: diverse chemical function from a few good reactions**. *Angew Chem Int Ed* 2001, **40**:2004-2021.

•2. Baskin JM, Bertozzi CR: **Bioorthogonal click chemistry: covalent labeling in living systems**. *QSAR Comb Sci* 2007, **26**:1211-1219.

This review frames click chemistry within the philosophy of chemical biology, and provides a thorough overview of major click reactions, reporter/visualization techniques, and considerations for *in vitro* versus *in vivo* applications.

3. Huisgen R, Padwa A (Ed.): *1,3-Dipolar Cycloaddition Chemistry*, Wiley; 1984.

4. Tornøe CW, Christensen C, Meldal M: **Peptidotriazoles on solid phase: [1,2,3]-triazoles by regioselective copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides**. *J Org Chem* 2002, **67**:3057-3064.

5. Rostovtev VV, Green LG, Fokin VV, Sharpless KB: **A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes**. *Angew Chem Int Ed* 2002, **41**: 2596-2599.

6. Griffin RJ: **The medicinal chemistry of the azido group**. *Prog Med Chem* 1994, **31**: 121-232.

7. Link AJ, Mock ML, Tirrell DA: **Non-canonical amino acids in protein engineering**. *Cur Op Biotech* 2003, **14**: 603-609.

8. Kiick KL, Saxon T, Tirrell DA, Bertozzi CR: **Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation**. *PNAS* 2002, **99**: 19-24.

9. Link AJ, Tirrell DA: **Cell surface labeling of *Escherichia coli* via copper(I)-catalyzed [3+2] cycloaddition**. *J Am Chem Soc* 2003, **125**: 11164-11165.

10. Link AJ, Vink MKS, Tirrell DA: **Presentation and detection of azide functionality in bacterial cell surface proteins**. *J Am Chem Soc* 2004, **126**: 10598-10602.

•11. Agard NJ, Baskin JM, Prescher JA, Lo A, Bertozzi CR: **A comparative study of bioorthogonal reactions with azides**. *ACS Chem Bio* 2006, **1**: 644-648.

This review covers the major classes of azide-bearing biomolecules and their associated reactions, and presents an organized list of the strengths, weaknesses, and methodological details/references for each.

••12. Dietrich 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)**. *PNAS* 2006, **103**: 9482-9487.

Using human embryonic kidney 293 cells, the authors showed for the first time that incorporation of Aha into mammalian proteins is unbiased and nontoxic. They subsequently used the labeling to identify 195 proteins synthesized in a 2-hour window.

13. Link AJ, Vink MKS, Tirrell DA: **Preparation of the functionalizable methionine surrogate azidohomoalanine via copper-catalyzed diazo transfer.** *Nat Proto* 2007, **2**: 1879-1883.

14. Link AJ, Vink MKS, Tirrell DA: **Synthesis of the functionalizable methionine surrogate azidohomoalanine using Boc-homoserine as precursor.** *Nat Proto* 2007, **2**: 1884-1887.

•15. Dietrich DC, Lee JJ, Link AJ, Graumann J, Tirrell DA, Schuman EM: **Labeling, detection and identification of newly synthesized proteomes with bioorthogonal non-canonical amino-acid tagging.** *Nat Proto* 2007, **2**: 532-540.

The authors provide thorough instructions on the application of BONCAT, including example control experiments, highlights of potential pitfalls, and steps for troubleshooting various reactoins.

•16. Beatty KE, Liu JC, Xie F, Dietrich DC, Schuman EM, Wang Q, Tirrell DA: **Fluorescence visualization of newly synthesized proteins in mammalian cells.** *Angew Chem Int Ed* 2006, **45**: 7364-7367.

Using mouse embryo fibroblasts, the authors demonstrated that incorporation of Hpg into mammalian proteins is also unbiaed and nontoxic, and they showed that both Aha and Hpg could be selectively labeled with fluorescent tags to visualize newly synthesized proteins.

17. Bordun J (Ed.): *BioProbes* 56. Invitrogen, 2008.

••18. Beatty KE, Tirrell DA: **Two-color labeling of temporally defined protein populations in mammalian cells.** *Bioorg Med Chem Lett* 2008, in press, doi:10.1016/j.bmcl.2008.08.046

Using Rat-1 fibroblasts, the authors showed that incorporation of Aha and Hpg into proteins could be accomplished both simultaneously and sequentially, and that the resulting tagged proteins could be selectively labeled with two different fluorophores.

••19. 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.** *PNAS* 2007, **104**: 16793-16797.

Using a difluorinated cyclooctyne instead of a terminal alkyne, the authors showed that a Huisgen cycloaddition could be accomplished under physiological conditions in minutes in the absence of a copper(I) catalyst. This provides new possibilities for BONCAT in living systems by avoiding the cytotoxic copper(I) species.