A EUROPEAN JOURNAL OF CHEMICAL BIOLOGY

CHEMBIOCHEM

SYNTHETIC BIOLOGY & BIO-NANOTECHNOLOGY

Accepted Article

Title: Irreversible protein labeling via Paal-Knorr conjugation

Authors: Ramesh Dasari, James J La Clair, and Alexander Kornienko

To be cited as: *ChemBioChem* 10.1002/cbic.201700210

Link to VoR: http://dx.doi.org/10.1002/cbic.201700210

A Journal of

ChemPubSoc Europe

WILEY-VCH

www.chembiochem.org

10.1002/cbic.201700210

WILEY-VCH

Irreversible protein labeling *via* Paal-Knorr conjugation

Ramesh Dasari,[a] James J. La Clair,*[b,c] and Alexander Kornienko*[a]

Abstract: The application of new chemical reactions in a biological context has advanced bioconjugation methods for both fundamental research and commercial arenas. Recent adaptations of reactions such as Huisgen 1,3-dipolar or Diels-Alder cycloadditions have enabled the labeling of specific residues in biomolecules by the attachment of molecules carrying azides, alkynes, or strained alkenes. Although these are fundamental tools, there is a need for the discovery of reactions that can label native proteins. We now report on the adaptation of the Paal-Knorr reaction to label lysine-residues in proteins *via* pyrrole linkages.

New protein labeling methods, characterized by efficiency, short reaction times, site-selective targeting, or the use of low substrate concentrations, are in great demand.[1] Many reported labeling reagents for native proteins[2] target strongly nucleophilic cysteine residues.[3] Covalent modification of lysine (Lys) is another important strategy, although the nucleophilicity of Lys-NH_2 is diminished by its protonated nature under physiological conditions (pKa 10.5). Some recently reported Lys-targeting methods include reductive alkylation using iridium-catalyzed transfer hydrogenation,[4] coupling with diazonium terephthalates,[5] Michael addition with allylic elimination of sulphonamides,[6] sulfonylation with fluorosulfonylbenzoates,[7] and nucleophilic aromatic substitution with dichlorotriazines.[8] We now describe the advancement of a class of labels (Fig. 1) that target Lys residues by the formation of irreversible pyrrole linkages.

One of the oldest name reactions involving primary amines is the classical Paal-Knorr pyrrole synthesis that was reported independently by Paal[9] and Knorr[10] in 1884. The reaction, as depicted in Scheme 1, involves the condensation of a primary amine with 1,4-dicarbonyl compound **3** in the presence of an acid catalyst to give pyrroles **5** through the intermediary hemiaminal **4**.[11]

As the Paal-Knorr reaction proceeds under mild reaction conditions often at room temperature with the use of a mild acid, it is a method of choice for sensitive functionality. This has been well documented through a variety of synthetic applications, such as the commercial preparation of the anticholesterolemic drug atorvastatin (Lipitor),[12] or total synthesis of natural products, such as that of roseophilin reported by Trost.[13]

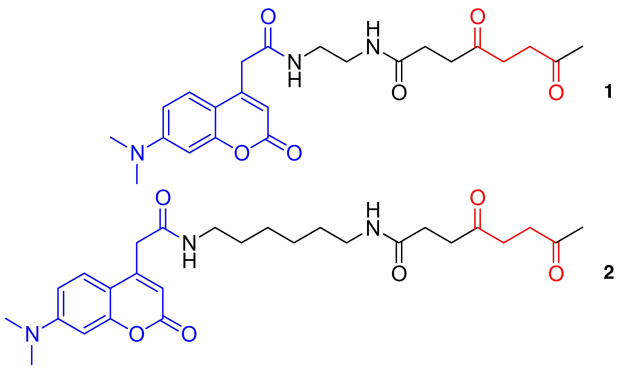
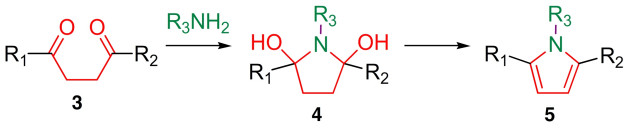


Figure 1. Structures of Paal-Knorr probes **1** and **2**. The 7-dimethylamino-4-coumarin or immunoaffinity-fluorescent (IAF) tag and 1,4-diketone warhead are colored in blue and red, respectively.



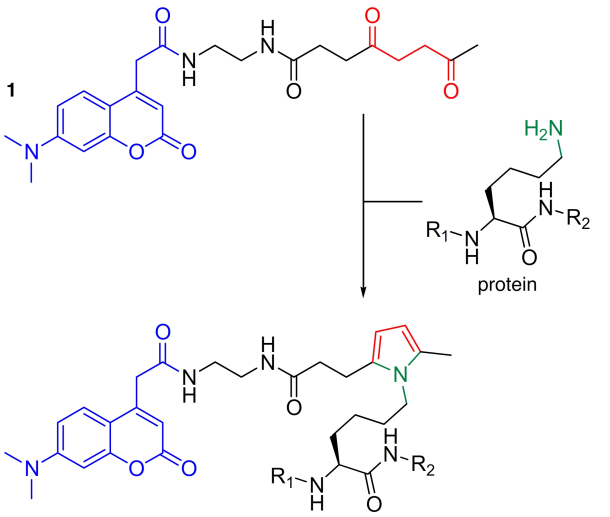
Scheme 1. The Paal-Knorr reaction. Diketone **3** condenses with an amine to generate hemiaminal **4**, which undergoes a facile double dehydration to afford pyrrole **5**. The 1,4-diketone and amine moieties are coloured in red and green, respectively.

The Paal-Knorr reaction is also biologically relevant based on its involvement in *n*-hexane-induced axonal atrophy within the central nervous system. As shown in both *in vitro* and *in vivo* systems, 2,5-hexanedione (R_1=R_2=Me, Scheme 1), a neurotoxic metabolite of *n*-hexane, undergoes a selective Paal-Knorr reaction with Lys residues (R_3NH_2, Scheme 1) of axonal cytoskeleton proteins, forming 2,5-dimethylpyrrole adducts within specific regions of neurofilaments.[14] Based on this biological precedent, it appeared to us that the Paal-Knorr reaction would be an interesting addition to the arsenal of Lys-labeling methods. We focused our efforts on demonstrating this method through probe **1**, as shown in Scheme 2.

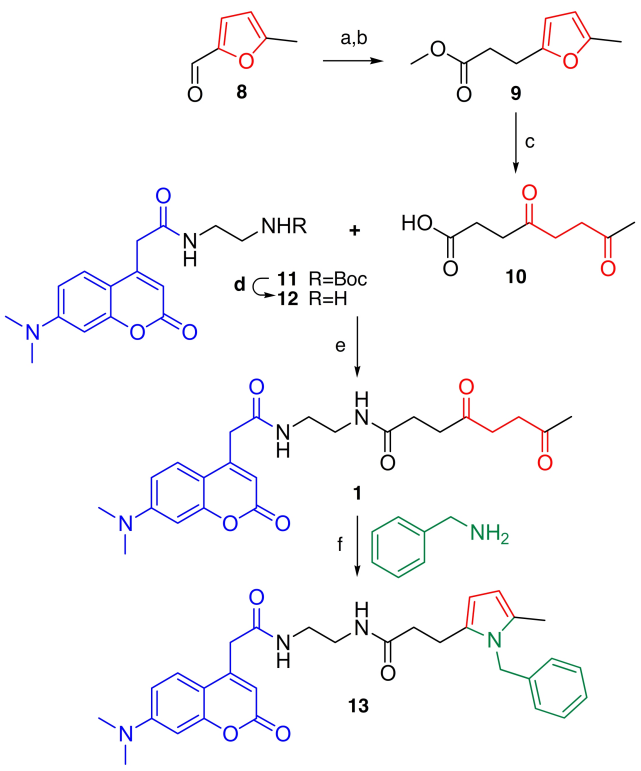
The synthesis of the Paal-Knorr probes **1** and **2** began with the Wittig reaction of 5-methylfurfural with Ph_3P=$CHCO_2Me$ to afford the corresponding α,β-unsaturated ester, which was hydrogenated to yield furan ester **9** (Scheme 3). Ester **9** was subjected to an acid-catalyzed hydrolytic opening of the furan ring with the concomitant ester hydrolysis to afford 4,7-dioxooctanoic acid (**10**) in 62% yield (Scheme 3). The resulting acid was coupled with immunoaffinity fluorescent (IAF) tag **12**[15] using EDAC and DMAP to afford the desired probe **1** in 86% yield. The absence of any detectable aldol processes under both acidic (**9** to **10**) and basic (**10** to **1**) reaction conditions attests to the stability of such aliphatic 1,4-dicarbonyl compounds and bodes well for their applications in a biological context.

[a] R. Dasari, Prof. A. Kornienko
Department of Chemistry and Biochemistry
Texas State University
San Marcos, TX 78666, United States
E-mail: a_k76@txstate.edu
[b] Dr. J. J. La Clair
Department of Chemistry and Biochemistry
University of California at San Diego
La Jolla, CA 92093, United States
[c] Dr. J. J. La Clair
Xenobe Research Institute
P. O. Box 3052
San Diego, CA 92163, United States
E-mail: i@xenobe.org

Supporting information for this article is given via a link at the end of the document.

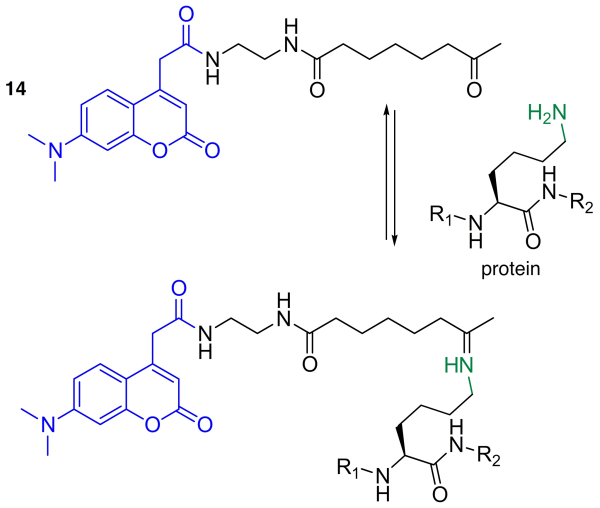


Scheme 2. A schematic representation of the Paal-Knorr labeling of a protein with probe **1**. The IAF tag, 1,4-diketone and amine are coloured blue, red and green, respectively.



Scheme 3. Synthesis of Paal-Knorr probe **1**. Reagents and conditions: a) $Ph_3P{=}CHCO_2Me$, THF, reflux, 20 h, 92%; b) H_2, Pd/C, EtOH, rt, 24 h, 78%; c) H_2SO_4 (trace), 50% aq. HOAc, 110 °C, 15 h, 62%; d) 0.5 M HCl, 1,4-dioxane, CH_2Cl_2, rt, 2 h, 98%; e) EDAC, DMAP, DMF, CH_2Cl_2, 20 h, 86%; f) AcOH, THF, rt, 5 h, 86%. The IAF tag, 1,4-diketone and amine are coloured in blue, red and green, respectively.

In contrast, we found that probe **1** was highly reactive toward primary amines as is evident by the facile condensation with benzyl amine to form pyrrole **13** in an excellent 86% yield under mildly acidic conditions at 23 °C (Scheme 3). Similar methods were then applied to prepare a chain elongated probe **2** from **10** (see Supplementary Information). This probe was prepared to further explore the linker requirements between the IAF tag and the reactive moiety. In addition, we also synthesized a mono-ketone **14** (Scheme 4, see Supplementary Information) as a control. Here, the presence of a single ketone in **14** would prevent it from undergoing the Paal-Knorr reaction, but rather lead to Schiff base formation (Scheme 4).



Scheme 4. A schematic representation of Schiff base formation with control **14**. The IAF tag and amine are coloured blue and green, respectively.

We then examined the use of probes **1** and **2** to modify two proteins (as depicted in Scheme 2), bovine serum albumin (BSA) and keyhole limpet serum (KLH), commonly conjugated via their lysine residues for antibody production.[16] As shown in Fig. 2a, probes **1** and **2** were able to label BSA and KLH. The fact that both probes operated at similar efficacy (Fig. 2a) indicated that both short linkers in **1** and longer in **2** were viable. For both proteins, the reaction was strongest with probes **1** and **2**, although a minor level of labeling was observed with the mono-ketone control **14** (Fig. 2a). After further analysis we found that this background labeling of BSA with **14** could be removed by spin-dialysis prior to gel analysis, however traces of fluorescence from **14**, likely resulting from Schiff base formation, were still observed with KLH. Comparative analyses[17] of the level of fluorescence in Fig. 2a with standards indicated that 1.4±0.1 and 2.4±0.2 tags were observed per protein for BSA and KLH, respectively, after treatment with 50 eq. of probe in PBS pH 7.2 for 24 h at 37 °C.

Experiments with other proteins revealed that the Paal-Knorr conjugation was not unique to BSA and KLH. Of interest, labeling of LC-8, a protein known to oligomerize,[19] was accompanied by increased levels of dimeric and trimeric oligomers (Fig 2c). These oligomers were not reverted by the addition of a reducing agent (BME, DTT, or TCEP), boiling the sample prior to SDS PAGE analysis, or changing of the type of gel, indicating that they may have arisen from covalent

modifications possibly involving crosslinking of the pyrrole units[20] brought into proximity by LC-8 oligomerizaton. While attempts were made to identify this by mass spectral analyses, the complexity of finding trace levels of crosslinked peptides in this sample proved too difficult. This result provides early evidence that the Paal-Knorr technique may also be used to trap oligomerization states within proteins or even act on protein complexes.

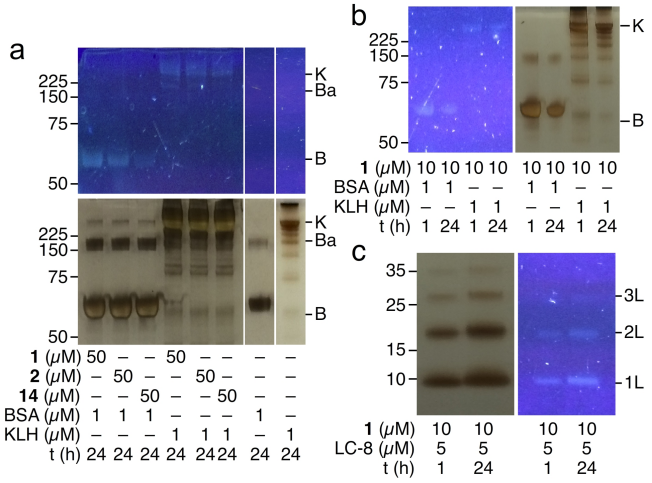


Figure 2. The Paal-Knorr labeling of BSA, KLH and LC-8. **a)** The labeling of BSA or KLH in PBS pH 7.2 with diketone probes **1** and **2** as compared to monoketone control **14**. The gel was visualized on a UV light at 280 nm (top) and then silver stained (bottom). **b)** The limit of visual detection via SDS-PAGE was approximately 1 μg of protein bearing > 1 label per protein.[17] **c)** The labeling of LC-8 with probe **1** at 1 h or 24 h. B, K, and L denote BSA, KLH and LC-8, respectively. Concentrations are provided in μM. For each lane 1 μg of protein was loaded per lane with the exception of the control lanes (right two lanes in Fig. 2a) where 0.5 μg was loaded. Aggregate bands were observed from BSA (Ba). KHL was obtained commercially and the addition bands observed during silver staining are not atypical, as it is purified from natural sources.

We then tested the limits of the method. We found that visualization by eye was limited by the treatment with a 10-fold excess of probe **1** when loading 1 μg of protein per lane on a typical SDS-PAGE gel (Fig. 2b). Using a Typhoon scanner, the labeling of BSA could be visualized after treatments with 0.5 to 2.5 eq. of probe **1** (Fig. 3a). For both, visualization improved after fixation with 10% AcOH in 30% aq. EtOH, as it removed background from unreacted **1**. The detection limit of the labeling process was further reduced by Western blot analyses using an anti-IAF monoclonal antibody (mAb).[18] As illustrated in Fig. 3b, we were able to detect labeling in bands containing greater than 1 ng of BSA treated with 5 eq. of **1**.

We also conducted studies to identify the optimal labeling conditions for BSA. We found that the reaction could be observed within 1 h at 37 °C (Fig. 2b) and labeling was not enhanced with longer treatment times, as indicated by SDS-PAGE analysis (Fig. 2b). Thus, we typically allowed the reaction to run for 6 h to 12 h to ensure completion of the pyrrole formation (Scheme 1). We also observed that reducing the pH did not influence the outcome of this reaction. While pH 4-6 is often used to accelerate Paal Knorr reactions, a reduction in the pH did not seem to adjust the ability of **1** to label BSA (Fig. 3c). Finally, we carefully checked the spectroscopic properties of gel purified **1**•BSA conjugate and found that its spectral properties matched that of **1** (Fig. 3d).

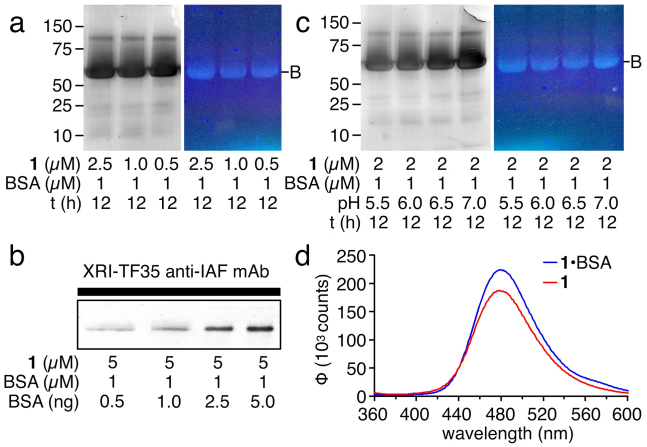


Figure 3. Paal-Knorr labeling of BSA. **a)** The labelling of BSA can be detected down to 0.5 eq. of probe **1** when scanned on a Typhoon laser scanner (Amersham). **b)** Western blot analyses can be used to further reduce the detection level of BSA labeled with **1**. **c)** Evaluation of pH effects on the labelling of BSA by probe **1**. Concentrations are provided in μM and ~1 μg of protein was loaded per lane. **d)** Fluorescence (Φ) spectra of 10 μM probe **1** and 15 BSA•**1** prepared by treatment of 1 μM BSA with 5 μM probe **1** for 12 h in PBS pH 7.2. TLC and LC-MS analysis indicated that samples of BSA•**1** were free of unreacted probe **1**.

Next, we set out to confirm that protein labeling of BSA occurs through the pyrrole forming reaction as illustrated in Scheme 2. Using 1 h and 24 h treatments at 37 °C, we conducted the labeling on 100 μg scale using 10 eq. of **1**. Once complete, we purified the resulting protein by SDS-PAGE. Samples of the bands containing **1**-labeled BSA were collected under clean conditions and dried. These samples were submitted to Trypsin-digestion and LC-MS/MS returning excellent peptide coverage with multiple long peptides (Fig. 4b).[22] We then conducted a detailed search of the identified peptides with MS-GF+[23] and MODA[24] database search algorithms that enable the identification of peptides with non-natural modifications. Within this dataset, modified peptides were identified, such as S_{310}HCIAEVEKDAIPENLPPLTADFAE-D**K**$_{336}$DV_{338} (orange, Fig. 4). Samples treated for 1 h contained peaks for both hemiaminal **4** (3567.6937 *m/z* observed, 3567.4706 *m/z* calculated) and pyrrole adduct **5** (M+407, 3532.6377 *m/z* observed, 3531.8295 *m/z* calculated). However, the sample treated for 24 h only contained the peak due to the pyrrole **5**. While detection of this peptide was possible, the low level and complexity of this detection did not allow for an analytical evaluation of the site selectivity within this reaction.

We then evaluated the stability of these conjugates. Samples of **1**-labeled BSA or **1**-labeled KLH remained fluorescent and intact at pH 4-9. Moreover, the tagged-protein did not degrade after long-term storage, as samples of **1**-labeled BSA used in Fig. 4 were stored at 4 °C for over 1 month (limit of

testing) or in lyophilized dry form at rt for over 1 month (limit of testing) as evaluated by SDS-PAGE.

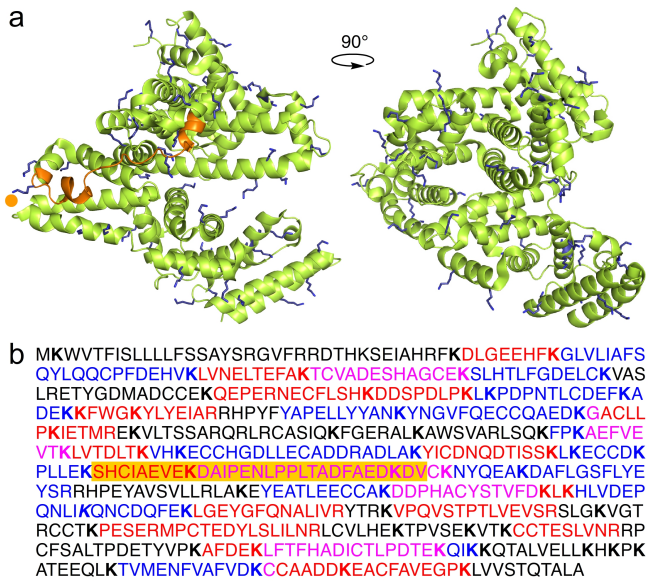


Figure 4. The labeling of BSA with probe **1**. **a)** Each molecule of BSA has 59 Lys residues, 30-35 of which have primary amines that are capable of being conjugated[21] **b)** Peptide coverage map for BSA labeled with detected peptides in red or blue and overlapping regions of these peptides in magenta. LC-MS/MS analysis confirmed the modification of Lys_{336} (orange dot) within an observed peptide (highlighted in orange). The observation of the peptide in b) does not suggest that labelling occurred on a specific Lys residue, rather, we were able to detect labelling at this position.

With stable materials in hand, we evaluated whether our conjugates were viable for cellular work. Using confocal microscopy, we observed rapid endocytosis and localization of **1**-labeled BSA within viable HCT-116 cells (Fig. 5). Subsequent staining with LysoTracker Red DND-99,[25] a red-fluorescent lysosomal organelle stain, confirmed that **1**-labeled BSA was observed in the lysosomes as previously reported for BSA conjugates in HCT-116 cells.[26] Using Western blot analyses (Supporting Fig. S1), we were able to confirm the fact that **1**-labeled BSA remained intact within the cells during this process.

In conclusion, we have demonstrated the application of the Paal-Knorr reaction for protein labeling using Lys-rich proteins as substrates (BSA, KLH or LC-8). There are several features to this reaction that make its use for protein conjugation applications and chemical biological studies attractive. First, there are no reagents required or special media. In these examples, labeling was readily observed at pH 7.2 in PBS at ambient temperature. The labeling process occurred rapidly and was completed within 24 h. Second, while we demonstrated the labeling application using an IAF tag, the ease in the synthesis of the warhead unit **10** (Scheme 3) allows tags to be appended to a wide range of agents. Third, probes containing warhead **10** are not reactive to media like many of the other functional handles such as acrylates, α-haloamides, succinimidyl esters, or sulfonyl halides, an issue that often creates significant batch-based inhomogeneity, which has recently been addressed in part with continuous-flow reactors.[27] Finally, the fact that 1,4-diketones are resistant to metabolism[28] and have demonstrated Paal-Knorr reactivity in humans[14] indicate their potential for further advance as next generation tools for chemical biological applications.

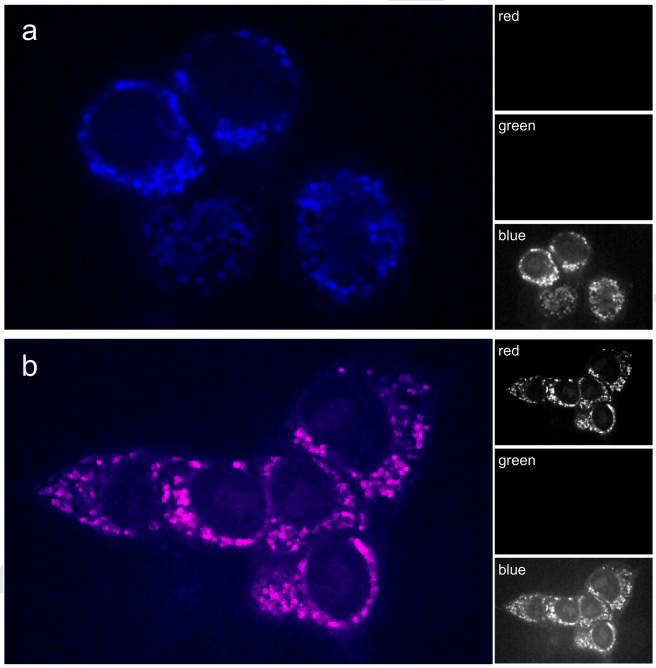


Figure 5. Cell imaging of the uptake of **1**-labeled BSA. **a)** An image of HCT-116 cells stained with **1**-labeled BSA (blue). HCT-116 cells grown to 10^5 cells/cm^2 were incubated for 3 h in media containing 25 μM **1**-labeled BSA. After this period the cells were washed with media and imaged live. **b)** Image of HCT-116 cells in a) after incubating with 2.5 μM LysoTracker Red DND-99 for 15 min prior to imaging.[23] Images were collected with blue (405 nm laser with emission filtered at 447±60 nm); green (543 nm laser and emission filtered at 593±40 nm) and red (633 nm laser with emission filtered at 692±40 nm) fluorescence. Overlay image is provided at left and selected channels at the right of each panel. We confirmed the presence of **1**-labeled BSA in the treated cells using Western blot analyses as shown in Supporting Fig. S1.

Acknowledgements

This work was generously supported by funding from the Xenobe Research Institute (to J. J. L.) and the National Cancer Institute grant CA186046 (to A. K.). We thank Majid Ghassemian (UC San Diego) for protein digestion and LC-MS/MS analyses and Nuno Bandeira (UC San Diego) for suggestions for the MS/MS data searching and analyses.

Keywords: protein labeling • Paal-Knorr reaction • Click chemistry • biocompatible reactions • bioconjugation

[1] O. Boutureira, G. J. L. Bernardes, *Chem. Rev.*, **2015**, *115*, 2174.
[2] J. I. MacDonald, H. K. Munch, T. Moore, M. B. Francis, *Nat. Chem. Biol.*, **2015**, *11*, 326.
[3] J. M. Chalker, G. J. L. Bernardes, Y. A. Lin, B. Davis, *Chem. Asian J.*, **2009**, *4*, 630.
[4] J. M. McFarland and M. B. Francis, *J. Am. Chem. Soc.*, **2005**, *127*, 13490.
[5] S. Diethelm, M. A. Schafroth, E. M. Carreira, *Org. Lett.* **2014**, *16*, 3908.

[6] F. Jinno, T. Yoneyama, A. Morohashi, T. Kondo, S. Asahi, *Biopharm. Drug Dispos.*, **2011**, *32*, 408.
[7] N. N. Gushwa, S. Kang, J. Chen, J. Taunton, *J. Am. Chem. Soc.*, **2012**, *134*, 20214.
[8] A. Shannon, R. Banerjee, E. R. Webster, D. W. Bak, C. Wang and E. Weerapana, J. *Am. Chem. Soc.,* **2014**, ***136***, 3330.
[9] C. Paal, *Chem. Ber.*, **1884**, *17*, 2756.
[10] L. Knorr, *Chem. Ber.*, **1884**, *17*, 1635.
[11] S. Abbat, D. Dhaked, M. Arfeen, P. V. Bharatam, *RSC Adv.*, **2015**, *5*, 88353.
[12] Y. V. Novozhilov, M. V. Dorogov, M. V. Blumina, A. V. Smirnov, M. Krasavin, *Chem. Cent. J.*, **2015**, *9*, 7.
[13] B. Trost and G. A. Doherty, *J. Am. Chem. Soc.*, **2000**, *122*, 3801.
[14] L. Zhang, T. Gavin, A. P. DeCaprio and R. M. LoPachin, *Toxicol. Sci.* **2010**, *117*, 180.
[15] M. D. Alexander, M. D. Burkart, M. S. Leonard, P. Portonovo, B. Liang, X. Ding, M. M. Joullié, B. M. Gulledge, J. B. Aggen, A. R. Chamberlin, J. Sandler, W. Fenical, J. Cui, S. J. Gharpure, A. Polosukhin, H. R. Zhang, P. A. Evans, A. D. Richardson, M. K. Harper, C. M. Ireland, B. G. Vong, T. P. Brady, E. A.Theodorakis and J. J. La Clair, *Chembiochem.* **2006**, *7*, 409.
[16] J. E. Coligan, J. P. Tam and J. Shao, *Curr. Protoc. Neurosci*. **2001**, *5*, Unit 5.6.
[17] Fluorescence labeling was determined in gel using densitometry measurements using parallel fluorescence with detection and SilverQuest stained gels using ImageJ, see: P. Rehbein and H. Schwalbe, *Protein Expr. Purif*. **2015**, *110*,1. b) C. T. Rueden and K. W. Eliceiri, *Biotechniques*. **2007**, *43*, 33.
[18] W. L. Yu, G. Guizzunti, T. L. Foley, M. D. Burkart MD and J. J. La Clair, *J. Nat. Prod*. **2010**, *73*, 1659.
[19] a) J. Hall, A. Hall, N. Pursifull, E. Barbar, *Biochemistry*. **2008**, *47*, 11940. b) E. Barbar, *Biochemistry*. **2008**, *47*, 503.
[20] Pyrrole polymerization can occur in oxidative or low pH environments: G. F. Smith, *Adv. Heterocycl. Chem.* **1963**, *18*, 287. b) Y. Tan and K. Ghandi, *Synt. Metals*, **2013**, *175*, 183. Its occurrence within proteins may have some utility for site-directed polymerization of proteins: c) C. Albayrak and J. R. Schwartz, *ACS Synth. Biol.* **2014**, *3*, 353.
[21] BSA has 59 Lys residues of which 30-35 are on the surface or in pockets available for labeling, as noted in: a) O. Koniev, A. Wagner, *Chem. Soc. Rev.* 2015, 44, 5495. b) K. Hirayama, S. Akashi, M. Furuya, K. I. Fukuhara, *Biochem. Biophys. Res. Comm.* **1990**, *173*, 639. c) F. Wurm, T. Steinbach, H.-A. Klok, *Chem. Comm.* **2013**, *49*, 7815.
[22] B. X. Huang, H.-Y. Kim and C. Dass, *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1237.
[23] B. A. Risk, N. J. Edwards and M. C. Giddings, *J. Proteome Res*. **2013**, *12*, 4240.
[24] a) S. Na, N. Bandeira and E. Paek, *Mol. Cell Proteomics*. **2012**, *11*, M111.010199. b) O. Horlacher, F. Lisacek, M. Müller. *J. Proteome Res.* **2016**, *15*, 721.
[25] X. Álvarez-Micó, D. D. Rocha, L. A. Guimarães, A. Ambrose, E. Chapman, L. V. Costa-Lotufo, J. J. La Clair and W. Fenical, *Chembiochem*. **2015**, *16*, 2002.
[26] K. Mellert, M. Lamla, K. Scheffzek, R. Wittig and D. Kaufmann D. *PLoS One*. **2012**, *7*, e52473.
[27] a) T. D. Mai, I. Pereiro, M. Hiraoui, J. L. Viovy, S. Descroix, M. Taverna, C. Smadja, *Analyst*. **2015**, *140*, 5891. b) M. D. Witte, T. Wu, C. P. Guimaraes, C. S. Theile, A. E. Blom, J. R. Ingram, Z. Li, L. Kundrat, S. D. Goldberg, H. L. Ploegh. *Nat. Protoc*. **2015**, *10*, 508.
[28] I. Saito, E. Shibata, J. Huang, N. Hisanaga, Y. Ono, Y. Takeuchi. *Br. J. Ind. Med*. **1991**, *48*, 568-574.

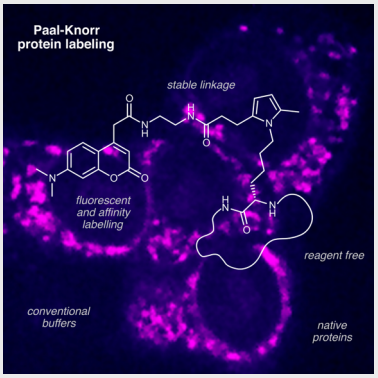
10.1002/cbic.201700210

WILEY-VCH

Entry for the Table of Contents

COMMUNICATION

The advance of a suite of bioorthogonal chemical reactions has profoundly advanced the tools available to biochemists to study proteins and other biomolecules. Here, we describe the use of the Paal-Knorr reaction to fluorescently-label proteins. Akin to the needs of a biocompatible method, the described procedures operate without reagents, catalysts, or organic solvents.



Ramesh Dasari, James J. La Clair, and Alexander Kornienko**

Page No. – Page No.

Irreversible protein labeling *via* Paal-Knorr conjugation