ChemComm
Chemical Communications

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: A. Gupta, B. Garreffi and M. Guo, *Chem. Commun.*, 2020, DOI: 10.1039/D0CC03643A.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/chemcomm

View Article Online
DOI: 10.1039/D0CC03643A

COMMUNICATION

Facile Synthesis of a Novel Genetically Encodable Fluorescent α-Amino Acid Emitting Greenish Blue Light

Aakash Gupta, [a] Brian P. Garreffi, [a] and Maolin Guo *[ab]

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

We report the facile synthesis and characterization of a novel fluorescent α-amino acid 4-phenanthracen-9-yl-L-phenylalanine (Phen-AA) (5) that emits greenish blue light in the visible region. This genetically encodable L-α-amino acid has excellent photostability with a 75% quantum yield. It readily gets into human cells, being clearly imaged upon 405 nm laser excitation. The synthetic procedure is resistant to racemization and only involves three simple steps which use mild conditions and generate the Phen-AA in reasonably good yield. It may find broad applications in research, biotechnology, and the pharmaceutical industry.

Synthetic biology which harnesses innovative approaches for engineering new biological molecules/systems or re-designing existing ones for useful purposes has been revolutionizing scientific research as well as biotechnology and the pharmaceutical industry[1]. As one of its most promising areas, site-specific incorporation of functional unnatural amino acids into peptides or proteins has a wide range of applications in biochemical research and the biopharmaceutical industry[2]. Proteins or peptides can be modified with new physical, chemical or biological properties via the inclusion of unnatural amino acids. These unnatural amino acids resemble the natural building blocks of proteins but contain distinct functional groups[2]. Expanding in amino acid building blocks enable peptides or proteins to have new or modified functions, providing great opportunities for innovative research or novel medicines[2]. Of particular interest is the incorporation of fluorescent probes as a mechanism to track protein function, transport, and folding[3].

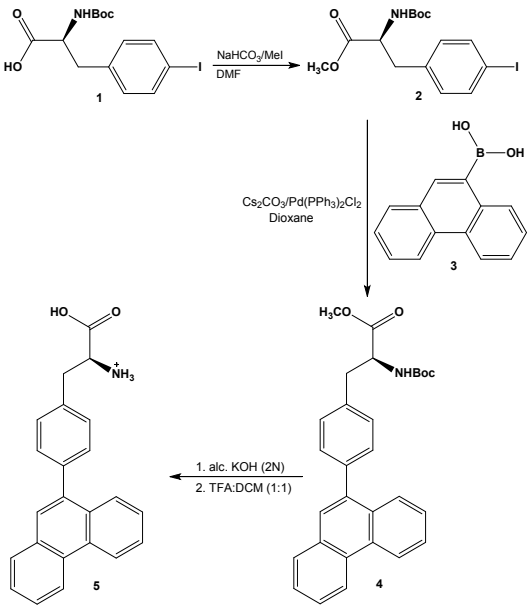
The recent advances in microscopy and the utility of highly sensitive fluorescence-based approaches have significantly accelerated the development in research fields such as molecular and cellular biology, biophysics, biotechnology and medicine[4]. Fluorescent fusion proteins (such as Green Fluorescent Proteins, GFPs), and the more recent enzymatic "tags" (such as the SNAP-tag or CLIP-tag) which are genetically encoded onto the *N*- or *C*-terminus of the protein of interest have enabled the visualization of proteins in living systems[4]. However, these fusion-based technologies suffer from restriction to a potentially essential terminus of the protein of interest, and the fluorophores are quite large (≥20 kDa) which can casue them to interfere with the expression, localization, and stability or function of the protein to which they are attached[4]. The use of intrinsic fluorescent α-amino acids, such as tryptophan (Trp) which absorbs/emits in the UV wavelength region and has low fluorescence quantum yield and poor photostability, relies on a relative abundance within the protein and multiple residues in different environments which can complicate the spectroscopy and is hardly useful for single-molecule measurements and imaging applications[3a,5,6k]. A promising alternative approach is the utilization of unnatural fluorescent α-amino acids[3a,5,6]. These α-amino acid analogues are very small in comparison with the GFPs or tags, and can therefore be placed precisely at targeted positions within a peptide or protein, and can therefore act as relatively nonperturbing replacements for the native residues, instead of being restricted to the termini, thereby maintaining the overall native structure and function of a target peptide or protein[3a]. Moreover, the photophysical and chemical properties of unnatural fluorescent α-amino acids can be purposely tuned and they can be incorporated into peptides or proteins via traditional solid phase peptide synthesis (SPPS), expressed protein ligation (EPL) or the more recent genetically encoded unnatural amino acid incorporation techniques[2,3a,5]. Once the unnatural fluorescent α-amino acids are synthesized, new applications in peptide or protein incorporation could be readily adopted[2,3a,5].

Unnatural fluorescent α-amino acids (FAA) have gained particular interests in the field of chemical biology due to their versatile applications such as being used as fluorescent probes

[a] *Department of Chemistry and Biochemistry, UMass Cranberry Health Research Center, University of Massachusetts Dartmouth, 285 Old Westport Road, North Dartmouth, MA 02747 (USA), E-mail: mguo@umassd.edu.*
[b] *Department of Chemistry, University of Massachusetts Amherst, 710 North Pleasant Street, Amherst, MA 01003 (USA), E-mail: mguo@umass.edu.*
† Footnotes relating to the title and/or authors should appear here.
Electronic Supplementary Information (ESI) available: Details on chemicals and instrumentation are provided in the Supporting Information.
See DOI: 10.1039/x0xx00000x

to study protein dynamics, folding/unfolding, local conformation changes, biomolecular interactions including peptide-peptide, peptide-membrane, protein-protein, and protein-nucleic acids interactions[6], as well as in probing peptidoglycan synthesis and bacterial growth *in situ*[7] and live-cell imaging[8]. However, in applications where an unnatural fluorescent α-amino acid reporter is required or desirable, the choice is rather limited[6,8,9], owing to their high cost and lack of availability[10]. Since unnatural α-amino acids typically have at least one stereocenter and possess two or more reactive functional groups (the amine and carboxylate), they are challenging synthetic targets and thus their synthetic routes usually require multiple steps[5,8,9,11]. Nevertheless, they have proved to be promising in broad applications, increasing their attraction as synthetic targets[3a,5-10], using both chemical[5-9] and enzymatic approaches[10,11]. Still, the majority of the reported unnatural fluorescent α-amino acids do not absorb or emit light in the visible wavelength region which is a much desired property of the FAAs for live-cell or *in-vivo* imaging applications[8,9,10].

Over the past decade, our laboratory has been developing novel fluorescent molecules using a chemical biology approach to enable the study of real-time biological phenomena in live cells[3b,12]. Recently, we have been exploring real-time protein imaging in live cells via genetically-encoded incorporation of an unnatural fluorescent amino acid into a protein of interest[13]. Synthesis of novel FAAs has thus become a need to expand our fluorescent amino acid toolbox. We have previously demonstrated that the polycyclic aromatic anthracene ring can be fused on a pro-chelator, exhibiting blue fluorescence in aqueous solution[14]. In this work, via a facile three-step synthesis, we have coupled the anthracene ring to L-phenylalanine, creating a novel fluorescent α-amino acid that emits greenish blue fluorescence.



Scheme 1. Synthesis route of 4-Phenanthracen-9-yl-L-phenylalanine (**5**)

Herein, we report a facile synthesis and initial spectroscopic characterization of a novel fluorescent α-amino acid 4-phenanthracen-9-yl-L-phenylalanine (**5**) (Phen-AA) that shines light in the visible spectrum. As shown in the synthesis scheme (**Scheme 1**), we commenced with the synthesis of N-(*tert*-Butoxycarbonyl)-4-iodo-L-phenylalanine methyl ester intermediate (Boc-Phe(4-I)-OMe) (**2**) by protecting the caboxylate group through methylation of the commercially available N-(*tert*-butoxycarbonyl)-4-iodo-L-phenylalanine (**1**). The next step of our synthetic strategy involves Suzuki-Miyaura cross-coupling reaction as the key step, which couples α-amino acids bearing vinyl or aryl halide side-chains with polyaromatic boronic acids[6h,15], and here to give the N-(Boc)-4-(9-phenanthracenyl)-L-phenylalanine methyl ester (**4**) intermediate. De-protection of the methyl and Boc protecting groups of **4** produced the final L-α-amino acid 4-Phenanthracen-9-yl-L-phenylalanine (**5**) (SI).

A few optimizations/modifications were made in the synthesis. Our initial approach of the Suzuki-Miyaura cross-coupling reaction of **2** with 9–Phenanthracenyl boronic acid (**3**) using the traditional THF–toluene (1:1 v/v) solvent combining Na_2CO_3 as the base[6h,i] to synthesize **4** was low yielding. Consumption of the starting material as monitored by TLC was observed after changing the solvent to dioxane (less toxic and high boiling point) as the aprotic solvent and the base to Cs_2CO_3 which is more soluble in organic solvents than other common bases[15b]. Another modification was to use the strong base KOH (KOH dissolved in methanol) solution in step 3 instead of the commonly used LiOH[16] which gave incomplete deprotection. Isolation of the 4-Phenanthracen-9-yl-L-phenylalanine (**5**) in ~37 % over three steps was achieved using preparation HPLC (**Figure S1**). Characterization was accomplished using ^{1}H and ^{13}C NMR and HR MS. ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.94 (d, 1H), 8.87 (d, 1H), 8.02 (d, 1H), 7.91 (d, 1H), 7.75 – 7.59 (m, 5H), 7.46 (s, 4H), 3.51 (m, 1H), 3.28 (dd, 1H), 2.98 (dd, 1H). ^{13}C NMR (101 MHz, DMSO-d_6): δ 170.08, 138.52, 138.46, 137.66, 131.55, 130.86, 130.64, 130.15, 129.97, 129.79, 129.12, 127.56, 127.53, 127.39, 127.25, 126.86, 123.82, 123.24, 56.14, 37.42. HRMS (ESI): calcd for $[C_{23}H_{19}NO_2 + H]^+$: 342.1416, found: *m/z*: 342.1278, [minor dimer at m/z 683.3774]. $[\alpha]_\lambda^T$ = - 26.36 (SI). Our synthetic approach has two major features: First, as the starting material is an L-α-phenylalanine derivative, and the synthetic procedure does not involve any reactions on the α-C, the final α -amino acid product remains in the L-form (SI) which can be incorporated into biosynthesis pathways and thus is genetically encodable. The conservation of chiral purity is likely related to the electron donating property of the *tert*-butoxycarbonyl (Boc) protecting group which can block the α-amino function and thus resists to racemization [18]. Second, the reactions in this synthetic approach proceeded under mild conditions as well as in reasonably good yield. Not only could this synthesis be potentially performed on large scale, but it also gives the synthetic route the advantage of being able to be conducted in biology laboratories which usually lack advanced synthetic facilities.

Next, we investigated the photo proterties of the novel amino acid 4-Phenanthracen-9-yl-L-phenylalanine (**5**). As seen in **Figure 1**, the UV-Vis spectrum of the Phen-AA displayed absorption peaks at ~258 nm (ϵ=6.21×10^4 M^{-1}cm^{-1}) and 301 nm (ϵ=1.56×10^4 M^{-1}cm^{-1}), with two weaker bands at 334 nm (ϵ=7.85×10^2 M^{-1}cm^{-1}) and 351 nm (ϵ=5.51×10^2 M^{-1}cm^{-1}). The spectrum shows the characteristic peak of phenylalanine at ~260 nm, and ~260 nm, ~300, ~334 and ~351 nm for the conjugated phenanthracenyl moiety[14,18].

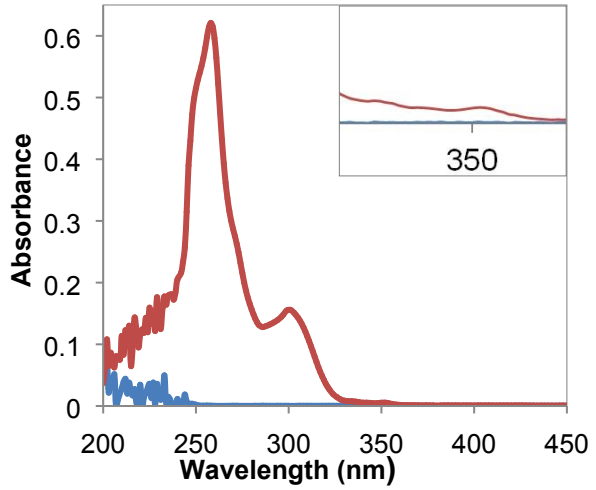


Figure 1. UV-Vis spectrum of 4-Phenanthracen-9-yl-L-phneylalanine in DMSO:water (1:1, 10 µM). Inset shows the zoom-in absorption bands at 334 and 351 nm.

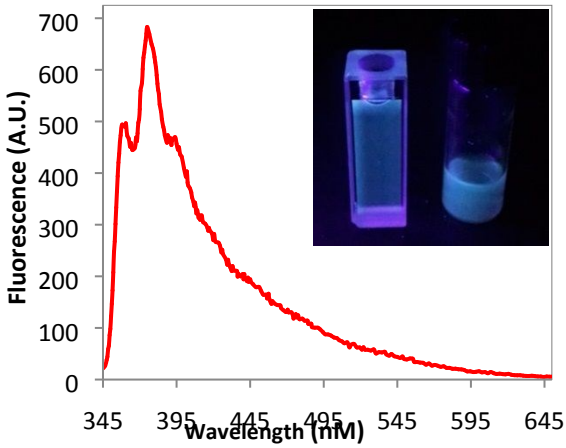


Figure 2. Fluorescence spectrum of 4-Phenanthracen-9-yl-L-phneylalanine (20 µM) in acetonitrile: water (1:1). Inset is an image of 4-Phenanthracen-9-yl phenylalanine (**5**) in methanol solution in a quartz cuvette and a glass sample vial under UV illumination at 365 nm, showing the blue light emission.

The fluorescence spectrum of 4-Phenanthracen-9-yl-L-phenylalanine (**5**)(**Figure 2**) shows interesting emission peaks at 359 nm, 375 nm, 400 nm, and shoulders ~445 nm typical for the conjugated phenanthracenyl moiety[14,18] when excited at 334 nm. It's interesting that the fluorescence eimission band at ~445 nm falls into the visible wavelength, tailing beyond ~640 nm (Figure 2). Upon illumination with a UV lamp at wavelength 365 nm, the 4-Phenanthracen-9-yl-L-phenylalanine (**5**) samples show greenish light emission as seen by naked eyes and blue light emission on captured images (shown in **Figure 2 inset**).

Photo-stability of the Phen-AA (**5**) was investigated by subjecting the amino acid to repeated excitations and recorded an emission spectrum every 180 seconds over 5 h. Therefore, a spectrum was recorded 120 seconds after the previous spectrum until 100 spectra were recorded. As seen in in **Fig. S5** (SI), little decrease in fluorescence intensity (~6.7% at 421 nm) was observed, demonstrating that Phen-AA has excellent photo-stability and is resistant to photo-bleaching. Quantum yield measurement using rhodamine B as a standard gives an excellent quantum yield of 0.75 in DMSO (SI).

Next, we investigated the effect of solvent on the fluorescence spectrum of Phen-AA. The fluorescence spectra of Phen-AA in different solvents (**Fig. S6** (SI)) two emission peaks (421 nm and 438 nm) with a tail extended to ~650 nm in all the solvents except DMSO and acetone, in which a third band at ~541 nm is more distinguishable. The fluorescence of the main peaks decreases in intensity in the order of DMSO, DMF, MeOH, EtOH, acetone, isopropanol and DMSO-SPB buffer (1:1, pH 7.0, 0.05 M). As solvatochromism is an extremely complex coupling of many different interactions and dynamical processes, such as solubility, polarization effects, solvent-induced structural modifications, solute-solvent hydrogen-bonding interactions, and solute aggregation [17], no further analysis on the nature of the solvatochromic effects was performed.

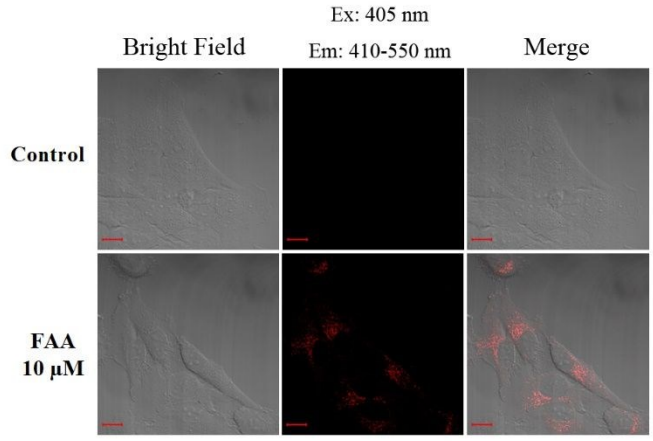


Figure 3. Confocal microscopy images (with DIC) of human Hela cells incubated with/without 10 µM Phen-AA. Scale bar, 20 µm.

Finally, we investigated the biocompatibility and potential in bioimaging of the new fluorescent amino acid Phen-AA in live cells. As seen in **Fig. 3**, Phen-AA readily gets into human Hela cells and can be clearly imaged (the red fluorescent signals in Phen-AA loaded cells) under a confocal microscope upon excitation at 405 nm. No change in morphology or death of the cells was observed during the imaging process, indicating good biocompatibility and low toxicity of the Phen-AA.

The fluorescence spectrum and the captured fluorescent emission images unambiguously demonstrated that the novel amino acid 4-Phenanthracen-9-yl-L-phenylalanine (**5**) is a fluorescent amino acid that emits in the UV as well as the visible region (Figs. 3, S5 and S6). The fluorescence of Phen-AA is quite

strong with excellent quantum yield (0.75). Considering the fact that the anthracene moiety has a long life time of ~35 ns in water[19a], the fluorescent property of the Phen-AA(**5**) is promising. Moreover, it has good biocompatibility and readily gets into human cells which can be clearly imaged upon 405 nm laser excitation. Furthermore, anthracene-based dyes are capable of two-photon absorption[19b], it may thus be possible to use visible light at longer wavelength or even NIR light to excite Phen-AA via two-photon technology in bioimaging studies to avoid UV damage.

In summary, we have created a novel genetically encodable fluorescent α-amino acid 4-phenanthracen-9-yl-L-phenylalanine (Phen-AA) (**5**) that emits greenish blue light in the visible region and can be clearly imaged in live cells upon 405 nm laser excitation. The synthetic procedure only involves three simple steps and all the reactions in our synthetic approach proceeded under mild conditions as well as in reasonably good yield. The general approach demonstrated here could also be applied to the synthesis of other new L-phenylalanine derivatives with novel structure and functions, which may find broad applications in scientific research, biotechnology and the pharmaceutical industry. Further characterization of this novel fluorescent amino acid and exploration of its biological applications are in progress in our laboratory.

We thank Drs B. Yan and Y. Wei for technical assistance and the NSF (CHE-1229339), NIH (1R15GM126576-01), and the Orphan Disease Center/Loulou foundation for funding.

Conflicts of interest

The development described in this manuscript is the subject of a patent application (M.G., A.G. and B.P.G. are the inventors).

Notes and references

1 a) M. El Karoui, M. Hoyos-Flight and L. Fletcher, *Front. Bioeng. Biotechnol.* 2019, **7**,175; b) F. Y. Bueso, and M. Tangney, *Trends in Biotechnology*, 2017, 35(5), 373.

2 a) C. J. Noren, S. J. Anthony-Cahill, M. C. Griffith, and P. G. Schultz, *Science*, 1989, **24**, 182; b) M. Kang, Y. Lu, S. Chen, and F. Tian, *Curr. Opin. Chem. Biol.* 2018, **46**, 123; c) N. D. Pham, R. B. Parker, and J. J. Kohler, *Curr. Opin. Chem. Biol.* 2013, **17**, 90; d) B. Yang, S. Tang, C. Ma, S. T. Li, G. C. Shao, B. Dang, W. F. DeGrado, M. Q. Dong, P. G. Wang, S. Ding, and L. Wang, *Nat. Commun*. 2017, **8**, 2240, e) L. Jakob, A. Gust, D. Grohmann, *Biochem. Biophys. Rep*. 2019, **17**, 1; f) L. Wang, *N Biotechnol,* 2017, **38(Pt A)**:16.

3 a) A. T. Krueger and B. Imperiali, *ChemBioChem,* 2013, **14**, 788; b) Y. Jiao, Y. Zhang, Y. Wei, Z. Liu, W. An, and M. Guo, *ChemBioChem,* 2012, **13**, 2335.

4 a) R. Y. Tsien, *Annu. Rev. Biochem*. 1998, **67**, 509; b) S. J. Remingtopn, *Protein Sci*. 2011, **20(9)**, 1509; c) R. Pantoja, E. A. Rodriguez, M. I. Dibas, D. A. Dougherty, and H. A. Lester, *Biophysical J*. 2009, **96**, 226.

5 a) A. H. Harkiss and A. Sutherland, *Org. Biomol. Chem.,* 2016, **14**, 8911; b) Z. Cheng, E. Kuru, A. Sachdeva, M. Vebdrell, *Nat. Rev. Chem.*, 2020, **4**, 275.

6 a) J. Wang, J. Xie, and P. G. Schultz, *J. Am. Chem. Soc.* 2006, **128**, 8738; b) D. Summerer, S. Chen, N. Wu, A. Deiters, J. W. Chin, and P. G. Schultz, *Proc. Natl. Acad. Sci.* USA 2006, **103**, 9785; c) Q. Wang, S. M. Cahill, M. Blumenstein, and D. S. Lawrence, *J. Am. Chem. Soc.* 2006, **128**, 1808; d) A. M. Reynolds, B. R. Sculimbrene, and B. Imperiali, *Bioconjugate Chem.* 2008, **19**, 588; e) H. S. Lee, J. Guo, E. A. Lemke, R. D. Dimla, and P. G. Schultz, *J. Am. Chem. Soc*. 2009, **131**, 12921; f) J. M. Goldberg, L. C. Speight, M. W. Fegley and E. J. Petersson, *J. Am. Chem. Soc.* 2012, **134**, 6088; g) S. Chen, L. Wang, N. E. Fahmi, S. J. Benkovic and S. M. Hecht, *J. Am. Chem. Soc.* 2012, **134**, 18883; h) S. Chen, N. E. Fahmi, L. Wang, C. Bhattacharya, S. J. Benkovic,, and S. M. Hecht, *J. Am. Chem. Soc*. 2013, **135**, 12924; i) J. S. Lampkowski, D. M. Uthappa and D. D. Young, *Bioorg. Med. Chem. Lett.* 2015, **25**, 5277; j) W. Ko, S. Kim, S. Lee, K. Jo and H. S. Lee, *RSC Adv*. 2016, **6**, 78661; k) K. Zhang, I. A. Ahmed, H. T. Kratochvil, W. F. DeGrado, F. Gai and H. Jo, *Chem. Commun.,* 2019, **55**, 5095; l) S.-M. Huang, F. Yang, B.-Y. Cai, Q.-T. He, Q. Liu, C.-X. Qu, M.-J. Han, W. Kong, Y.-L. Jia, F. Li, X. Yu, J.-P. Sun, J. Wang, *Anal. Chem*. 2019, **91**, 14936.

7 a) E. Kuru, H. V. Hughes, P. J. Brown, E. Hall, S. Tekkam, F. Cava, M. A. de Pedro, Y. V. Brun and M. S. VanNieuwenhze, *Angew. Chem., Int. Ed.*, 2012, **51(50)**, 12519; (b) E. Kuru, S. Tekkam, E. Hall, Y. V. Brun and M. S. Van Nieuwenhze, *Nat. Protoc*., 2015, **10**, 33.

8 L. Mendive-Tapia, R. Subiros-Funosas, C. Zhao, F. Albericio, N. D. Read, R. Lavilla, and M. Vendrell, *Nat. Protoc*., 2017, **12**, 1588.

9 a) M. R. Hilairea, I. A. Ahmedb, C. W. Lin, H. Joc, W. F. DeGradoc, and F. Gai, *PNAS.* 2017, **114**, 6005, b) Z. Xiang and L. Wang, *J. Org. Chem*. 2011, **76(15)**, 6367; X. Xu, X. Hu, J. Wang, *Beilstein J. Org. Chem.* 2013**, 9**, 254.

10 E. J. Watkins, P. J. Almhjell, and F. H. Arnold, *ChemBioChem.* 2020, **21**, 80.

11 C. E. Boville, D. K. Romney, P. J. Almhjell, M. Sieben, and F. H. Arnold, *J. Org. Chem*. 2018, **83**, 7447;

12 a) Y. Wei, Y. Zhang, Z. Liu, and M. Guo, *Chem. Commun.,* 2010, **46**, 4472; b) Y. Wei, Z. Aydin, Y. Zhang, Z. Liu, M. Guo, *Chembiochem*, 2012, **13(11)**, 1569; c) S. Maiti, Z. Aydin, Y. Zhang, and M. Guo, *Dalton Trans*. 2015, **44**, 8942; d) M. Ozdemir, Y. Zhang, and M. Guo, *Inorg. Chem. Commun.* 2018, **90**, 73; e) B. P. Garreffi, M. Guo, C. Tokranova, T. Castracane, and I. A, Levitsky, *Sensors and Actuators B: Chemical*, 2018, **276**, 466; f) Z. Aydin, B. Yan, Y. Wei, and M. Guo, *Chem. Commun.*, 2020, 56, 6043.

13 Z. Liu in Genetically encoded incorporation of unnatural fluorescent amino acids into proteins for molecular imaging in living cells, PhD dissertation, University of Massachusetts Lowell/Dartmouth. 2016, pp. 1.

14 Y. Wei, and M. Guo, *Chem. Commun*., 2009, 1413.

15 a) S. Kotha, and K. Lahiri, *Biopolymers*, 2003, **69**, 517; b) C. Zhang, J. Huang, M. L. Trudell, and S. P. Nolan, *J. Org. Chem.* 1999, **64(11)**, 3804.

16 J. M. Khurana, S. Chauhan, and G. Bansal, *Monatsh. Chem.* 2004, **135**, 83.

17 A. Marini, A. Munoz-Losa, A. Biancardi, and B. Mennucci, *J. Phys. Chem. B,* 2010, **114**, 17128.

18 a) A. I. Llobet, M. Alvarez, and F. Albericio, *Chem. Rev.* 2009, **109**, 2455; b) M. Bodanszky, *Principles of Peptide Synthesis*, Springer-Verlag Berlin Heidelberg, 1993, pp. 171

19 a) Z. Wang, D. M. Friedrich, M. R. Beversluis, S. L. Hemmer, A. G. Joly, M. H. Huesemann, M. J. Truex, R. G. Riley, and C. J. Thompson, *Environ. Sci. Technol.* 2001, 35, 2710; b) T. Xue, D. Zhao, T. Hao, X. Li, T. Wang and J. Nie, *New J. Chem.*, 2019,**43**, 6737.

