4-Aminophthalimide Amino Acids as Small and Environment-Sensitive Fluorescent Probes for Transmembrane Peptides

Samantha Wörner,[a] Franziska Rönicke,[a] Anne S. Ulrich,[b] and Hans-Achim Wagenknecht*[a]

Fluorescence probing of transmembrane (TM) peptides is needed to complement state-of-the art methods—mainly oriented circular dichroism and solid-state NMR spectroscopy—and to allow imaging in living cells. Three new amino acids incorporating the solvatofluorescent 4-aminophthalimide in their side chains were synthesized in order to examine the local polarity in the α-helical TM fragment of the human epidermal growth factor receptor. It was possible to distinguish their locations, either in the hydrophobic core of the lipid bilayer or at the membrane surface, by fluorescence readout, including blue shift and increased quantum yield. An important feature is the small size of the 4-aminophthalimide chromophore. It makes one of the new amino acids approximately isosteric to tryptophan, typically used as a very small fluorescent amino acid in peptides and proteins. In contrast to the only weakly fluorescent indole system in tryptophan, the 4-aminophthalimide moiety produces a significantly more informative fluorescence readout and is selectively excited outside the biopolymer absorption range.

Fluorescent amino acids should permit insights not only into fundamental behavior of peptides and proteins, but also into their localization in cellular contexts.[1,2] Solvatochromic fluorophores are required for this task, due to their responses to changes in surrounding polarity.[1b] Fluorescent and solvatochromic α-amino acids[1b] have been used to identify interactions between peptides and protein domains[3] such as SH2 domains,[4] 14-3-3 domains,[5] and PDZ domains,[6] as well as calmodulin[7] and class II MHC proteins.[8]

In particular, 4-aminophthalimide (4AP) is a very small but powerful environment-sensitive probe, because it can be excit-

ed outside the protein and nucleic acid absorption range and exhibits polarity-dependent shifts in emission maximum and changes in fluorescence quantum yield and lifetime.[9] The 4AP system is approximately the size of the indole moiety in tryptophan, which makes it an isosteric chromophore. Therefore, it can be assumed that it does not significantly alter the peptide architecture and causes only small perturbation of the native protein conformation. So far, only the 4-(N,N-dimethylamino)-phthalimide system has been used as a reporter amino acid component in some of those recent studies.[10]

Our aim is to use the 4AP system as a constituent of solvatochromic fluorescent amino acids to examine local polarity in membrane-bound peptides and proteins, to monitor their depths of immersion and to visualize the self-assembly of transmembrane (TM) helices. This approach is urgently needed to complement state-of-the art methods, such as oriented circular dichroism (OCD)[11] and solid-state nuclear resonance,[12] that are insensitive to environment polarity and/or depth of immersion. Hence, the use of the 4AP system as a fluorophore for TM peptides should expand the repertoire of accessible structural parameters, and might even allow studying of these peptides in their native biological environments. In contrast to previously published amino acid building blocks incorporating 4-(N,N-dimethylamino)phthalimide,[13] we want to apply the non-methylated 4AP, which shows significantly higher quantum yields in polar environments. The dimethylamino group reduces the fluorescence intensity through nonradiative decay processes via twisted intramolecular charge transfer (TICT) states.[14] Herein, we present the three new and structurally different amino acid building blocks 1–3 (cf. Scheme 1), each containing a 4AP moiety in the side chain, and evaluate their use as environment-sensitive fluorescent probes in the 35-residue α-helical TM fragment of the human epidermal growth factor receptor (EGFR; residues G640–V674).[15]

All three amino acid building blocks—Fmoc-1 to Fmoc-3 (Scheme 1)—are protected for solid-phase peptide synthesis. They structurally differ in the modes of attachment of the 4AP component, to allow testing of several synthetic approaches and different linkers with regard to their fluorescence properties.

Compound 1 is an alanine-derived building block in which the 4AP moiety is directly attached to the methyl side chain, like the indole system in tryptophan. Fmoc-1 could be synthesized from Fmoc-protected (2S)-2,3-diaminopropionic acid derivative 4, which was in turn prepared from Fmoc- and Boc-protected (2S)-2,3-diaminopropionic acid [Fmoc-L-Dap(Boc)-OH] in two simple steps.[16] 4-Aminobenzene-1,2-dicarboxylic

[a] S. Wörner, F. Rönicke, Prof. H.-A. Wagenknecht
Karlsruhe Institute of Technology (KIT), Institute of Organic Chemistry
Fritz-Haber-Weg 6, 76131 Karlsruhe (Germany)
E-mail: wagenknecht@kit.edu

[b] Prof. A. S. Ulrich
Karlsruhe Institute of Technology (KIT)
IBG-2 and Institute of Organic Chemistry
Fritz-Haber-Weg 6, 76131 Karlsruhe (Germany)
E-mail: anne.ulrich@kit.edu
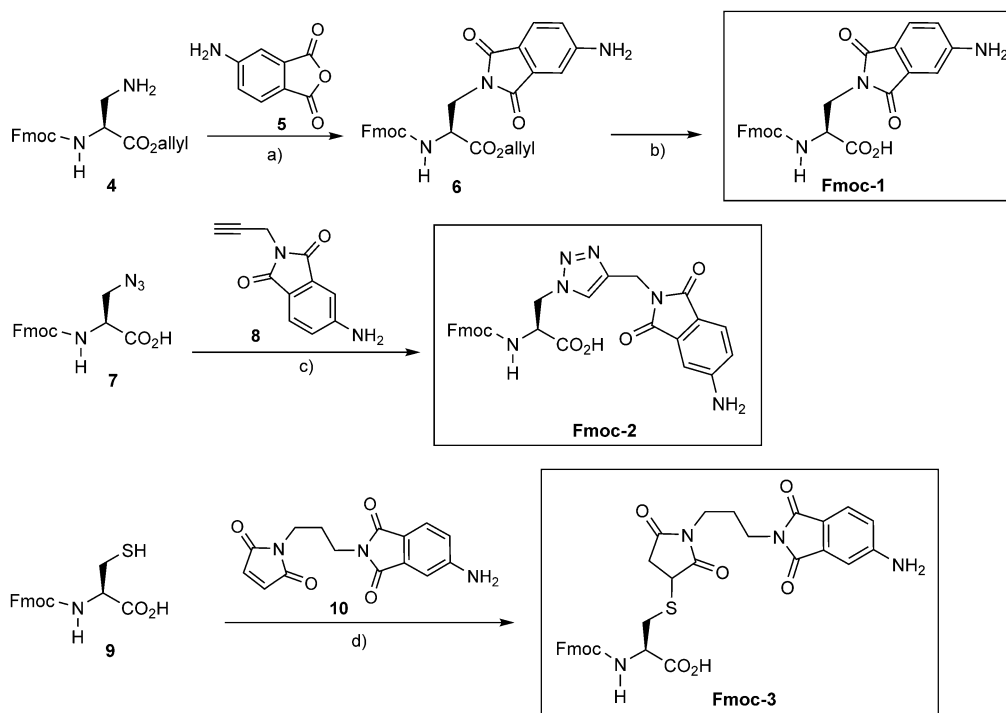
Scheme 1. Synthesis of 4AP-amino acid building blocks Fmoc-**1** to Fmoc-**3**: a) 1-hydroxybenzotriazole (HOBt), 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), ^{i}Pr$_2$NEt, RT, overnight, DMF, 71%; b) Pd(PPh$_3$)$_4$, PhSiH$_3$, RT, 1 h, CH$_2$Cl$_2$, 85%; c) sodium L-ascorbate, [Cu(CH$_3$CN)$_4$]PF$_6$, tris[(1-benzyltriazol-4-yl)methyl]amine (TBTA), RT, overnight, DMF, 92%; d) RT, 1 h, DMF, quant.

acid anhydride (**5**) was condensed with **4** through the free amino side chain in 71% yield. Finally, the allyl group of **6** was removed from the carboxy function by Pd^0 catalysis and use of PhSiH$_3$ as allyl acceptor.

The second building block Fmoc-**2** was obtained through our previous approach for attaching pyrene moieties and other chromophores[17] by means of Cu^I-catalyzed cycloaddition between Fmoc-protected β-azido-L-alanine (**7**) and the propargyl-substituted 4AP derivative **8**[18] in 92% yield.

The third building block Fmoc-**3** was accomplished through the maleimide bioconjugation ("thiol-ene") strategy.[19] Fmoc-protected L-cysteine (**9**) was coupled to the new and fluorogenic maleimide-4AP conjugate **10** in quantitative yield. Compound **10** was synthesized through a Diels–Alder/retro-Diels–Alder reaction sequence (Scheme S1 in the Supporting Information).

The solvatofluorescence properties of the amino acid building blocks Fmoc-**1** to Fmoc-**3** were determined in solvents of different polarities, ranging from the rather nonpolar ethyl acetate to water (Figure 1). These properties differ only slightly between the three variants (Figures S2–S7, Table S3). The emission maxima range from 460–463 nm in ethyl acetate to 561–568 nm in water. The fluorescence quantum yields (Φ_F) of Fmoc-**1** to Fmoc-**3** (Table S1) are in the range 0.73–0.64 in non-protic solvents, whereas in water they are very low, but still detectable (0.051 for Fmoc-**1**, 0.040 for Fmoc-**2**, and 0.043 for Fmoc-**3**). These concomitant fluorescence changes are typical for the 4AP system, though the amino acid building blocks Fmoc-**1**–**3** show quantum yields slightly higher than those of 4AP and remarkably higher than those of 4-(dimethylamino)-
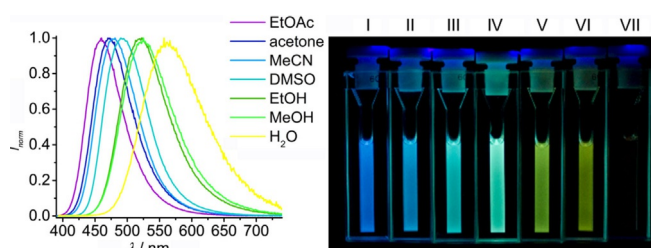

Figure 1. Fluorescence (left, excitation at λ_{abs}^{max} for each solvent) and images (right, excitation with a UV 360 nm handheld lamp) of Fmoc-**1** (0.04 mM) in different solvents (I. EtOAc. II. acetone. III. MeCN. IV. DMSO. V. EtOH. VI. MeOH. VII. H$_2$O).

phthalimide.[14] In the case of the latter chromophore, it is known that low-energy TICT states in polar environments quench the fluorescence.[9a,14]

The fluorescence properties of the maleimide-4AP conjugate Fmoc-**3** are noteworthy. Maleimides are known to quench fluorescence if they are covalently attached close to a fluorophore.[20–22] Accordingly, the reaction between **10** and maleimide **9** to form the thioalkyl succinimide Fmoc-**3** (d in Scheme 1) is accompanied by a strong increase in fluorescence intensity (Figure 2). Furthermore, the synthesis of **10** from the precursor **11** was already characterized by strong fluorescence quenching due to deprotection of the maleimide group in the retro-Diels–Alder reaction (Scheme S1). The fluorescence quantum yield (Φ_F) confirmed a 12-fold decrease from **11** to **10**, as well as a 13-fold increase to a remarkable $\Phi_F = 0.96$ after the reaction between **10** and **9** to form Fmoc-**3** (Table S2). The
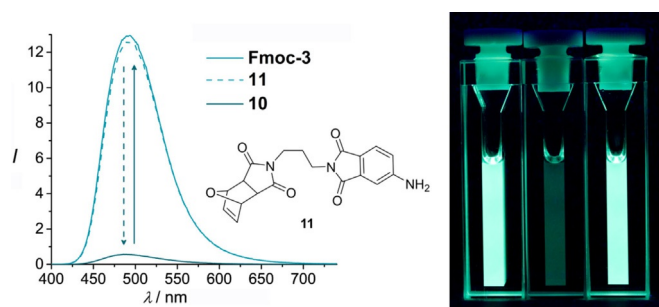
Figure 2. Fluorescence (left: $\lambda_{ex} = 383$ nm; right: images with excitation with a UV 360 nm handheld lamp) of Fmoc-**3**, **11**, and **10** (each 0.04 mM) in DMSO.

latter fluorescence increase is significant because it implies that the 4AP building block **10** is a potential fluorogenic and environment-sensitive probe for thiols and hence for cysteine residues in proteins.

As a representative TM segment of an integral membrane protein, we selected the TM segment of the EGFR.[23] This 35-residue peptide includes some residues from the extracellular juxtamembrane region (G640–P644), the hydrophobic α-helical TM domain (S645–M668), and parts of the intracellular juxta-membrane region (R669–V674). All three amino acid building blocks—Fmoc-**1** to Fmoc-**3**—were incorporated into peptides by using standard Fmoc-based solid-phase peptide synthesis protocols. The peptides were purified ($>95\%$) by reversed-phase HPLC and identified by MS (ESI-TOF). In these peptide models (sequences in Table 1), the 4AP-modified amino acid building blocks **1** and **3** were attached in the middle of the TM domain by replacement of Leu656 to obtain **EGFR1** and

Table 1. Peptide sequences of modified EGFR TM fragments incorporating the building blocks **1**–**3**.

Peptide	Sequence
EGFR$_{wt}$	H$_2$N-GPKIPSIATGMVGALLLLLVVALGIGLFMRRRHIV-COOH
EGFR1	H$_2$N-GPKIPSIATGMVGALL-**1**-LLVVALGIGLFMRRRHIV-COOH
EGFR2	H$_2$N-GPKIPSIATGMVGALL-**3**-LLVVALGIGLFMRRRHIV-COOH
EGFR3	H$_2$N-**2**-PKIPSIATGMVGALLLLLVVALGIGLFMRRRHIV-COOH
EGFR4	H$_2$N-**1**-PKIPSIATGMVGALLLLLVVALGIGLFMRRRHIV-COOH

EGFR2. In **EGFR3** and **EGFR4**, the solvatochromic building blocks **1** and **2** were attached to the N terminus (replacing G640). The solvatofluorescence properties could thus be probed and compared both in the hydrophobic core of the lipid bilayer and on the more water-accessible membrane surface.

To ensure that the artificial amino acids **1**–**3** do not affect the α-helical conformation or the membrane alignment of the TM domain, circular dichroism (CD) and oriented circular dichroism (OCD) spectra were recorded.[11] The CD spectra in MeCN/H$_2$O (1:1) each show a characteristic α-helical-like shape, with a maximum around 191 nm and two negative bands at 208 and 223 nm (Figure 3, left). This confirms that the building

blocks **1**–**3** in **EGFR1** to **EGFR4** do not perturb the global α-helical secondary structure of **EGFR$_{wt}$**.

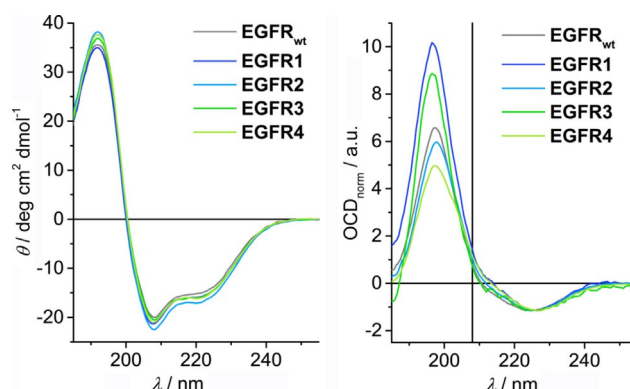OCD was used to inspect the helix orientation in macroscopically aligned membrane samples.[11] **EGFR$_{wt}$** and **EGFR1** to

Figure 3. Left: CD spectra of **EGFR$_{wt}$** to **EGFR4** (0.1 mg mL^{-1}) in MeCN/H$_2$O (1:1). Right: Oriented CD of **EGFR$_{wt}$** to **EGFR4** in POPC bilayers at a P/L ratio of 1:50. The OCD spectra (25 °C) were normalized to their negative peak at 225 nm to illustrate the similar line shapes (deviations below 200 nm are common, due to scattering of the oriented samples[24]).

EGFR4 were reconstituted in POPC bilayers with a peptide/lipid (P/L) ratio of 1:50 (mol/mol), by co-solubilization in CHCl$_3$/MeOH (1:1). The solution was spread onto a quartz glass sample holder, and the organic solvent was removed by drying and overnight vacuum, followed by full hydration in a humidity chamber. The positive OCD band at the diagnostic wavelength of 208 nm indicates that all peptides are aligned parallel to the membrane normal: that is, in the expected upright TM orientation[11] (Figure 3, right). Hence, both the CD and the OCD spectra confirm fully folded α-helical peptides in their proper TM orientation, with no significant impact from the modifications **1**–**3** in relation to **EGFR$_{wt}$**.

In order to elucidate the environment sensitivity of 4AP in the artificial amino acids, the optical properties of the labeled peptides **EGFR1** to **EGFR4** incorporated in the POPC liposomes (for preparation see the Supporting Information) were compared with those in MeOH (Figures 4 and S8), chosen due to the insolubility of the hydrophobic peptides in aqueous PBS buffer. **EGFR1** and **EGFR2**, in which building blocks **1** and **3** had been placed in the central hydrophobic position of the TM region, showed considerable changes in fluorescence readout upon reconstitution. The emission maximum was blue-shifted by 34 nm, from 517 to 483 nm, in the case of **EGFR1** and by 28 nm, from 530 to 502 nm, in that of **EGFR2**. Moreover, the fluorescence quantum yields (Φ_F) increased by a factor of 2.1 in the case of **EGFR1** and 3.3 in that of **EGFR2**, to 0.90 and 0.83, respectively, upon incorporation into liposomes (Table 2). These fluorescence changes are fully consistent with the TM orientation of the peptides in the lipid system, which places the 4AP side chains into the more hydrophobic environment inside the bilayer core. The absolute differences in fluorescence between **EGFR1** and **EGFR2** reflect the high sensitivi-
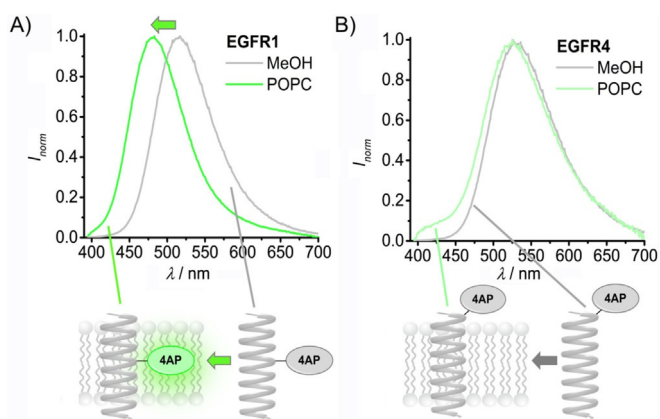
Figure 4. Observed fluorescence changes and illustration for A) **EGFR1**, and B) **EGFR4** upon incorporation into lipid vesicles, relative to in MeOH.

Table 2. Optical properties of **EGFR1** to **EGFR4** in MeOH and in POPC unilamellar vesicles.

Peptide	λ_{fluo}^{max}[nm][a]		Φ_F[b]	
	MeOH	POPC[c]	MeOH	POPC[c]
EGFR1	517	483	0.43	0.90
EGFR2	530	502	0.25	0.83
EGFR3	528	518	0.27	0.39
EGFR4	529	525	0.24	0.31

[a] $\lambda_{ex} = 380$ nm. [b] Measured by irradiation at 390 nm. [c] POPC vesicles in PBS buffer (pH 7.2) as unilamellar vesicles.

ty of the 4AP system towards the linker connecting it to the peptide. The shorter linker of **1**, in comparison with the rather long maleimide linkage of **3**, produces a more pronounced blue shift and a stronger increase in fluorescence quantum yield; these are consistent with the expected changes in polarity due to the local environments of **1** versus **3** in the peptides.

In contrast, **EGFR3** and **EGFR4** showed less pronounced fluorescence changes, as would be expected from the positioning of the 4AP building blocks at the membrane surface. With **2** at the N terminus, only a 1.4-fold increase in fluorescence quantum yield and a 10 nm blue shift from 528 to 518 nm were observed in the case of **EGFR3** upon incorporation into POPC vesicles. These changes might be due to the mobility of the rather long linker of **2**, leading to an interaction between the 4AP moiety and the lipid membrane and consequent shielding from the surrounding buffer. This is not possible with **1**, so **EGFR4** features a 1.3-fold increase in fluorescence quantum yield with only a 4 nm hypsochromic shift. These very small changes are consistent with the rigid linker moiety in the side chain of **1** maintaining the fluorophore in an aqueous environment at the membrane surface.

These results clearly show that the new amino acids **1**–**3** show distinct solvatofluorescence properties and allow characterization of TM peptides simply through fluorescence readout. In particular, the short linker in the side chain of **1** was found to be optimal for probing the TM peptides through fluorescence color and quantum yield.

To demonstrate the potential for imaging of TM peptides in living cells, HeLa cells were treated with POPC vesicles previously prepared with **EGFR1** and **EGFR4**. After 16 h, the cells were co-stained with the commercially available membrane-specific stain ConA–CF634 conjugate. The 4AP fluorescence is clearly visible in the vesicles located at the surfaces around the cells (Figure 5), but remarkably also as a diffusive fluorescence readout *in* the cellular membrane. The images also show that the cell membrane consists of both the former cell membrane
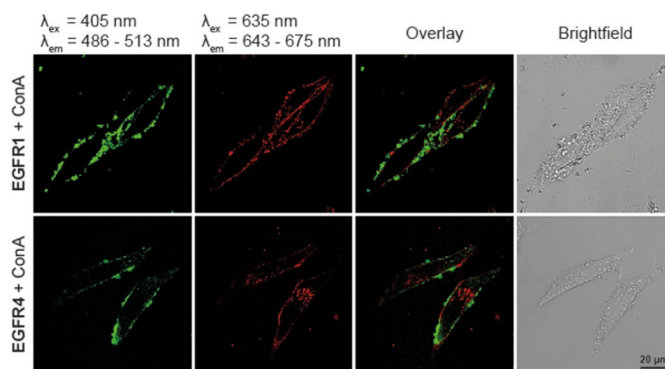

Figure 5. Live-cell confocal microscopy of HeLa cells treated with **EGFR1** and **EGFR4** ($\lambda_{ex} = 405$ nm, $\lambda_{em} = 486$–513 nm) for 16 h and co-stained with the membrane-specific stain ConA-CF633 ($\lambda_{ex} = 635$ nm, $\lambda_{em} = 643$–675 nm). Scale bar: 20 µm.

(red emission) and the former POPC vesicles (green emission by 4AP). Control experiments with empty POPC vesicles did not show the green 4AP fluorescence (Figure S9). Obviously, the 4AP-modified peptides **EGFR1** and **EGFR4** were integrated into the cell membranes of the living cells, and the fluorescence readout due to the 4AP side chain of **1** in the lipid core is suitable for imaging this TM peptide.

In conclusion, the 4AP system as an amino acid side chain modification represents a new and environment-sensitive probe for TM peptides. Three new 4AP-based amino acids were synthesized as protected building blocks Fmoc-**1** to Fmoc-**3** and are now available for incorporation into peptides. They differ in the length and type of the linker connecting the 4AP chromophore with the peptide backbone. Beside their general solvatofluorescence readout they have distinct individual advantages:

1) Compound **1** is an amino acid with a chromophore side chain that is very small in relation to other common fluorophores and nearly isosteric to that of the natural tryptophan. Unlike tryptophan, which needs to be excited in the UV-B range and shows only weak fluorescence,[2] however, **1** can be excited at the edge of the visible light range (outside the biopolymer absorbance range) and shows bright fluorescence through a large apparent Stokes shift. Also, the more recently published 4-cyanotryptophan meets these criteria only as a FRET donor.[25]

2) Compound **2** is based on preparation through Cu^{I}-cata-lyzed alkyne–azide cycloaddition. This could potentially be performed postsynthetically with azido-modified peptides.

3) Compound **3** is a strongly fluorogenic building block for thiole-ene bioconjugation, potentially applicable for selective modification of cysteine residues in proteins.

Each building block was successfully incorporated into the TM EGFR fragment G640–V674 by use of standard Fmoc-based solid-phase peptide synthesis techniques. It was possible to differentiate a position on the membrane-spanning helix deep within the hydrophobic core of the lipid bilayer (**EGFR1** and **EGFR2**) from a position at the membrane surface position (**EGFR3** and **EGFR4**) simply through fluorescence readout. The amino acids **1** and **2** each show a strong blue shift of about 30 nm and a strong two- to threefold increase in fluorescence quantum yield when introduced into the hydrophobic bilayer core. This is not the case when the amino acids **1** or **3** are incorporated at the N-terminal position of the EGFR peptide sequence, which positions the 4AP chromophore at the membrane surface. This fluorescence readout demonstrates the high potential of the new 4AP building blocks for probing TM peptides and proteins. Notable advantages are the small size of the 4AP chromophore, its excitation outside the protein and nucleic acid absorption range, and its informative and powerful fluorescence readout. Hence, it can be expected that our new synthetic amino acids will generally be applicable as environment-sensitive fluorophores to probe TM proteins, protein domains, and protein–protein interactions.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

[1] a) J. Zhang, R. E. Campbell, A. Y. Ting, R. Y. Tsien, *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 906; b) A. T. Krueger, B. Imperiali, *ChemBioChem* **2013**, *14*, 788–799; c) S. W. Hell, J. Wichmann, *Opt. Lett.* **1994**, *19*, 780–782.

[2] A. H. Harkiss, A. Sutherland, *Org. Biomol. Chem.* **2016**, *14*, 8911–8921.

[3] V. Sharma, D. S. Lawrence, *Angew. Chem. Int. Ed.* **2009**, *48*, 7290–7292; *Angew. Chem.* **2009**, *121*, 7426–7428.

[4] M. E. Vázquez, J. B. Blanco, B. Imperiali, *J. Am. Chem. Soc.* **2005**, *127*, 1300–1306.

[5] M. E. Vázquez, M. Nitz, J. Stehn, M. B. Yaffe, B. Imperiali, *J. Am. Chem. Soc.* **2003**, *125*, 10150–10151.

[6] M. Sainlos, W. S. Iskenderian, B. Imperiali, *J. Am. Chem. Soc.* **2009**, *131*, 6680–6682.

[7] G. Loving, B. Imperiali, *J. Am. Chem. Soc.* **2008**, *130*, 13630–13637.

[8] P. Venkatraman, T. T. Nguyen, M. Sainlos, O. Bilsel, S. Chitta, B. Imperiali, L. J. Stern, *Nat. Chem. Biol.* **2007**, *3*, 222–228.

[9] a) G. Saroja, B. Ramachandram, S. Saha, A. Samanta, *J. Phys. Chem. B* **1999**, *103*, 2906–2911; b) G. Saroja, T. Soujanya, B. Ramachandram, A. Samanta, *J. Fluoresc.* **1998**, *8*, 405–410; c) D. Mandal, S. Sen, D. Sukul, K. Bhattacharyya, A. K. Mandal, R. Banerjee, S. Roy, *J. Phys. Chem. B* **2002**, *106*, 10741–10747.

[10] M. Eugenio Vázquez, D. M. Rothman, B. Imperiali, *Org. Biomol. Chem.* **2004**, *2*, 1965–1966.

[11] J. Bürck, P. Wadhwani, S. Fanghänel, A. S. Ulrich, *Acc. Chem. Res.* **2016**, *49*, 184–192.

[12] E. Strandberg, A. S. Ulrich, *Concept Magn. Reson. A* **2004**, *23*, 89–120.

[13] M. E. Vazquez, D. M. Rothman, B. Imperiali, *Org. Biomol. Chem.* **2004**, *2*, 1965–1966.

[14] T. Soujanya, R. W. Fessenden, A. Samanta, *J. Phys. Chem.* **1996**, *100*, 3507–3512.

[15] E. V. Bocharov, D. M. Lesovoy, K. V. Pavlov, Y. E. Pustovalova, O. V. Bocharova, A. S. Arseniev, *Biochim. Biophys. Acta Biomembr.* **2016**, *1858*, 1254–1261.

[16] H. Waldmann, H. Kunz, *Liebigs Ann. Chem.* **1983**, 1712–1725.

[17] a) S. Hermann, D. Sack, H.-A. Wagenknecht, *Eur. J. Org. Chem.* **2018**, 2204–2207; b) S. Hermann, H.-A. Wagenknecht, *J. Pept. Sci.* **2017**, *23*, 563–566.

[18] J. Riedl, R. Pohl, N. P. Ernsting, P. Orság, M. Fojta, M. Hocek, *Chem. Sci.* **2012**, *3*, 2797–2806.

[19] J. M. J. M. Ravasco, H. Faustino, A. Trindade, P. M. P. Gois, *Chem. Eur. J.* **2019**, *25*, 43–59.

[20] G. Kokotos, C. Tzougraki, *J. Heterocycl. Chem.* **1986**, *23*, 87–92.

[21] J. Guy, K. Caron, S. Dufresne, S. W. Michnick, W. G. Skene, J. W. Keillor, *J. Am. Chem. Soc.* **2007**, *129*, 11969–11977.

[22] J. E. T. Corrie, *J. Chem. Soc. Perkin Trans. 1* **1994**, 2975–2982.

[23] E. V. Bocharov, P. E. Bragin, K. V. Pavlov, O. V. Bocharova, K. S. Mineev, A. A. Polyansky, P. E. Volynsky, R. G. Efremov, A. S. Arseniev, *Biochemistry* **2017**, *56*, 1697–1705.

[24] A. J. Miles, B. A. Wallace, *Chem. Soc. Rev.* **2016**, *45*, 4859–4872.

[25] K. Zhang, I. A. Ahmed, H. T. Kratochvil, W. F. DeGrado, F. Gai, H. Jo, *Chem. Commun.* **2019**, *55*, 5095–5098.