A Journal of the Gesellschaft Deutscher Chemiker

Angewandte International Edition Chemie

GDCh

www.angewandte.org

Accepted Article

Title: Rational tuning of fluorobenzene probes for cysteine-selective protein modification

Authors: Ahmed M. Embaby, Sanne Schoffelen, Christian Kofoed, Meldal Morten, and Frederik Diness

To be cited as: *Angew. Chem. Int. Ed.* 10.1002/anie.201712589
Angew. Chem. 10.1002/ange.201712589

Link to VoR: http://dx.doi.org/10.1002/anie.201712589
http://dx.doi.org/10.1002/ange.201712589

WILEY-VCH

10.1002/anie.201712589

WILEY-VCH

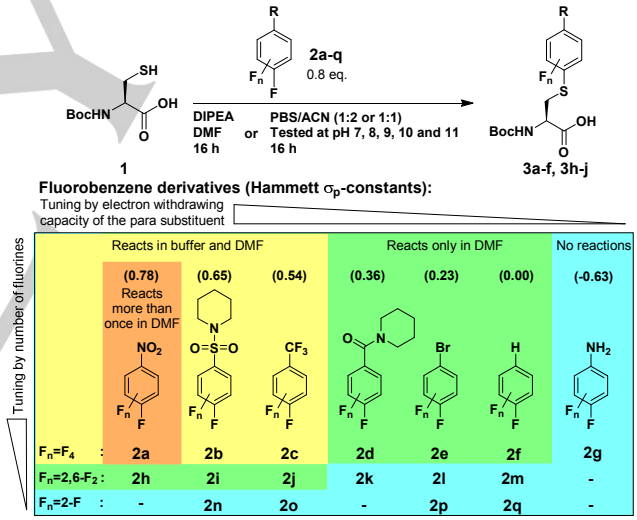
Rational tuning of fluorobenzene probes for cysteine-selective protein modification

Ahmed M. Embaby,[a] Sanne Schoffelen,[a] Christian Kofoed,[a] Morten Meldal[a],* and Frederik Diness[a],*

Abstract: *Fluorobenzene probes for protein profiling through selective cysteine labeling have been developed by rational reactivity tuning. Tuning was achieved by selecting an electron-withdrawing para-substituent combined with variation of the number of fluorine substituents. Optimized probes chemo-selectively arylated cysteine residues in proteins under aqueous conditions. Probes linked to azide, biotin or a fluorophore were applicable to labeling of eGFP and albumin. Selective inhibition of cysteine proteases was also demonstrated with the probes. Additionally, probes were tuned for site selective labeling among cysteine residues and for activity based protein profiling in cell lysates.*

Application of chemical tools for covalent modification of proteins is a powerful approach for identification, quantification and regulation of these biomolecules.[1] This principle is utilized in techniques such as ELISA, microarray screening, protein pull-down and activity based protein profiling (ABPP).[2] It is also an established therapeutic concept, exemplified by the marketed pharmaceuticals aspirin, penicillin, omeprazole and neratinib.[3] Cysteine is among the reactive target amino acids, and in spite of its low occurrence in proteins (1.9%), it is often located in functionally important sites.[4] In addition to stabilizing proteins through disulfide bonds, cysteines have diverse roles in metabolic processes, e.g. in catalysis, allosteric regulation, and metal binding.[5] Numerous electrophiles designed to probe cysteine residues have been described,[6] including many applied to cysteine proteases e.g. chloro-, and acylo-methylketones,[7] epoxides,[8] sulfonate esters,[9] haloacetamides[10] and Michael acceptors[11]. In recent years, also *S*-arylation has been explored as a method for modifying cysteine.[12] Notwithstanding these efforts, cross reactivity with other nucleophilic amino acid side chains, insufficient reactivity and lack of selectivity among cysteine residues are still unsolved issues.[9,13] It is noteworthy that the S-H bond has low dissociation energy,[14] and that the acidity, and thereby the reactivity of cysteines in proteins greatly depends on the local environment. At the surface of proteins, cysteine thiols have a pKa of ~8.5; whereas the pKa may be as low as 2.5 for a catalytic thiol in an active site.[15] Hence, a general methodology for rational tuning of probes towards selective modification of a given cysteine residue of interests is needed. Based on experience with nucleophilic aromatic substitution (S_NAr) reactions,[16] we hypothesized that a two dimensional reactivity tuning would provide such a system. The tuning relied, in one dimension, on the electron withdrawing capacity of the para-substituent on fluorobenzene and, in the second dimension, on varying the number of fluorine substituents. The reactivity of fluorobenzene derivatives towards *N*-tert-butyloxycarbonyl (Boc) cysteine (**1**) was investigated as a model system for exploring the concept (Scheme 1).

Scheme 1. Fluorobenzenes' reactivity towards cysteine.



Both promoting conditions using diisopropylethyl amine (DIPEA) in dimethylformamide (DMF) and less promoting conditions using a buffered mixture of water and acetonitrile (ACN) were tested. The reactivity of the pentafluorobenzene derivatives (**2a-g**) could be directly linked to the electron withdrawing capacity of the last substituent. Substituents with a Hammett σ_p-constant above 0.5 lead to reaction in both buffer and DMF.[17] A σ_p-constant between 0.5 and 0.2 (**2d-e**) provided reaction in DMF but not in buffer. Pentafluorobenzene **2f** (σ_p-constant of H = 0.0) reacted only slowly in DMF, while pentafluoroaniline **2g** (σ_p-constant of $-NH_2$ = -0.63) was completely unreactive. Reducing the number of fluorine substituents from five to three significantly reduced the reactivity of all the derivatives. The sulfonylamide- and trifluoromethyl-substituted trifluorobenzenes (**2i** and **2j**) were not reactive with **1** in buffer solution in contrast to their corresponding pentafluorobenzene derivatives (**2b** and **2c**). The amidoyl-substituted trifluorobenzene **2k** did not react with **1** in DMF while the amidoyl-substituted pentafluorobenzene **2d** was

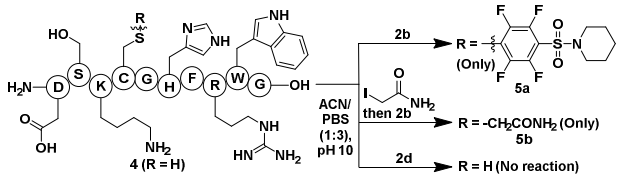
[a] Mr. Ahmed M. Embaby, Dr. Sanne Schoffelen, Mr. Christian Kofoed, Prof. Dr. Morten Meldal* and Prof. Dr. Frederik Diness*
Center for Evolutionary Chemical Biology, Department of Chemistry
University of Copenhagen
Universitetsparken 5, DK2100 Copenhagen, Denmark
E-mail: fdi@chem.ku.dk, meldal@nano.ku.dk.

Supporting information for this article is given via a link at the end of the document.((Please delete this text if not appropriate))

reactive. Finally, the reactivity was reduced even further by having only two fluorine substituents. Thus, the sulfonylamide or trifluoromethyl-substituted difluorobenzene (**2n** and **2o**) were not reactive at all even in DMF. All the compounds that initially displayed reactivity in buffer at pH 11 were further tested for reactivity at pH 7, 8, 9 and 10, and even though the reaction rates were reduced, conversion still occurred at these pH values. All reactions, except for the one with **2a** in DMF, afforded one major adduct as confirmed by product isolation.

Compound **2b** was further studied for selectivity towards cysteine over other amino acids in aqueous buffer. Peptide **4** containing the most nucleophilic protein functionalities was mixed with an excess of **2b** which gave only the monosubstituted adduct **5a** (Scheme 2). Reaction of **4** with iodoacetamide led to blocking of cysteine (**5b**) which prevented further reaction with probe **2b** added afterwards. In comparison, the less activated probe **2d** did not react under the same conditions.

Scheme 2. Chemo-selective arylation of cysteine in peptide.



Functionalized derivatives of **2b** were synthesized by reacting pentafluorophenylsulfonyl chloride with amino-functionalized PEG linkers containing a tag, such as an azide group for click chemistry (CuAAC or SPAAC), the sulforhodamine B fluorophore or biotin affinity probe (**6a**, **7**, or **8,** Figure 1). The generated tag-containing probes were reacted with enhanced green fluorescent protein (eGFP) comprising two cysteines (Cys49 and Cys71), not involved in disulfide bonds;[18] or with bovine serum albumin (BSA), which has 17 conserved disulfide bonds and one free thiol (Cys34).[19] All three types of reporter-linked probes were capable of labeling eGFP and BSA in a concentration-dependent manner and proved compatible with the functional click, the fluorescence or the affinity tag used for detection.

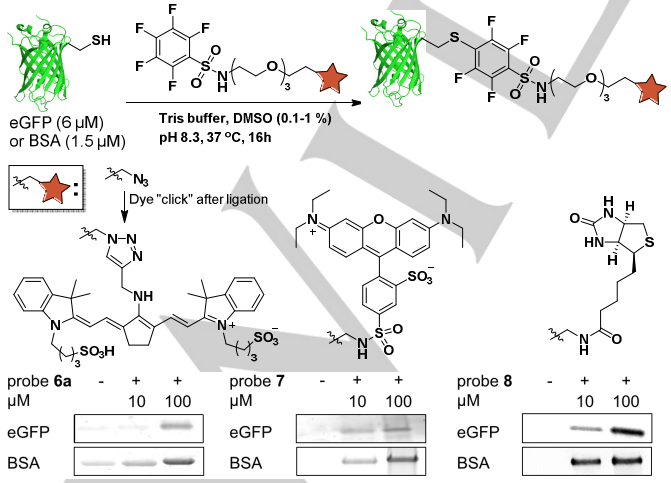


Figure 1. Labeling of eGFP and BSA with pentafluorophenylsulfonamide probes.

The methodology could also be adopted for selective inhibition of cysteine proteases over other classes of proteases, demonstrated herein by an on-bead inhibition screening assay towards the cysteine protease papain or the serine protease subtilisin (Figure 2). FRET substrates for the two proteases were synthesized on PEGA1900 polymer beads. The enzymes were pre-incubated with the potential inhibitor (100 equiv.) for 1 h at 37 °C. These mixtures were then added to the substrate beads. After 30 minutes, beads started to light up in wells with papain alone or papain pre-incubated with **2d**, **2g** or **2i** (Figure 2 and Figure S20) while the beads in the wells with papain pre-incubated with more reactive compounds **2a** or **2b** remained dark, even after 24 h. In a parallel experiment with subtilisin under identical conditions, none of the compounds were able to inhibit the protease activity.

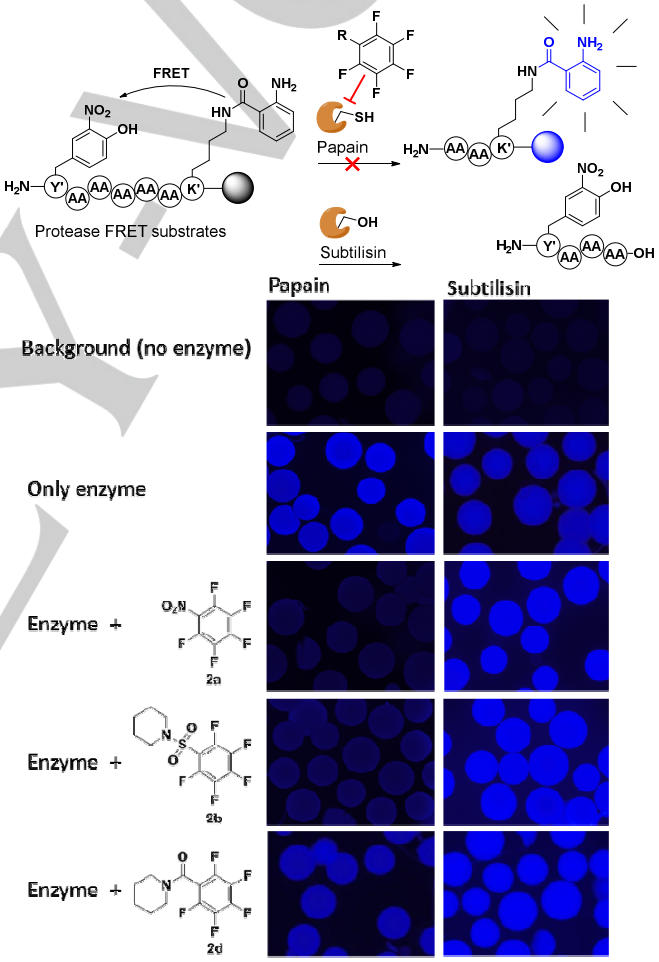


Figure 2. On-bead papain and subtilisin inhibition assays of probe **2a**, **2b** and **2d**. Schematic reaction and pictures of assay beads.

Three fluorobenzene derivatives with different reactivities (**2b**, **2i** and **2k**) were selected for investigation of their ability to react selectively with one among several cysteine residues in the same protein. Tobacco Etch Virus (TEV) protease (His-tagged variant) was chosen as a model protease as it contains a free cysteine exposed on the surface (Cys129) and another in the catalytic site (Cys170).[20] Native TEV protease was incubated with fluoroaryl compounds **2b**, **2i** or **2k** (~50 equiv.) (Figure 3).

10.1002/anie.201712589

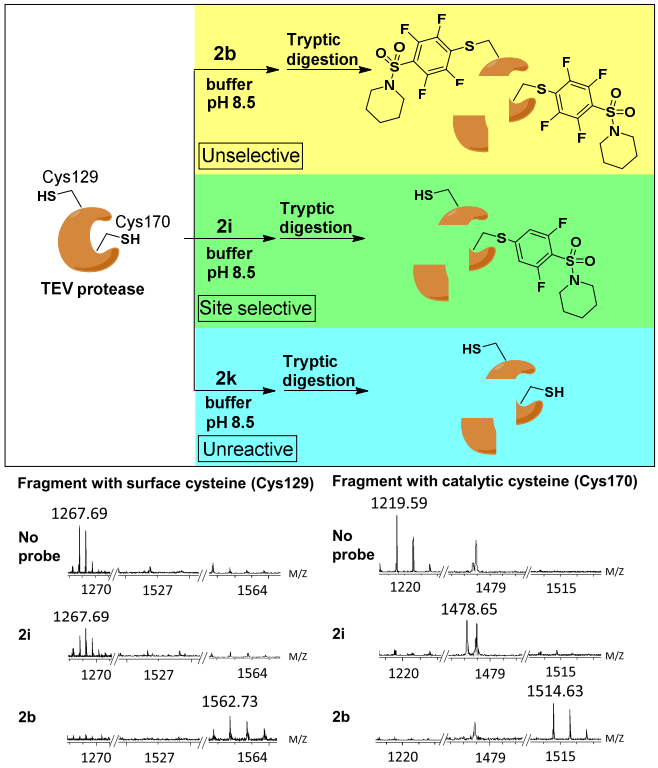


Figure 3. Screening of probes for site specific cysteine arylation in TEV protease. Schematic reaction and MS spectra showing peaks corresponding to cysteine-containing fragments of non-modified TEV protease and TEV protease treated with probes **2b** and **2i**.

After incubation for 10 h with the most reactive probe **2b**, a mixture of adducts were detected, with the highest mass corresponding to arylation of all free cysteine residues (Figure S24). Gratifyingly, when TEV protease was incubated with the less reactive probe **2i**, only the monosubstituted adduct was identified (Figure S25). Lastly, TEV protease was not modified by the least reactive probe **2k** under the same conditions. After incubation for 24 h, the selective arylation of the cysteine in the catalytic site by probe **2i** was confirmed by high-resolution mass spectrometry of all samples after tryptic digestion (Figure 3). In the native TEV protease, the peptide fragment containing the unmodified cysteine in the active site (Cys170) was found at 1219.59 Da and that of the fragment with a surface exposed cysteine (Cys129) at 1267.68 Da. Only the signal at 1219.59 Da disappeared upon treatment of TEV protease with **2i**, and a new mass peak at 1478.65 corresponding to the modified fragment appeared. In this sample the fragment with the surface exposed cysteine (Cys129) remained unchanged and a signal of the probe-functionalized fragment with an expected mass peak at 1526.72 Da was not detected. In contrast, both the 1219.59 Da and the 1267.68 Da signals completely disappeared upon treatment of TEV protease with compound **2b**. In this reaction two new mass peaks corresponding to modified peptide fragments appeared at 1514.62 Da and at 1562.73 Da. As a control experiment, both fragments were found to be modified after TEV treatment with iodoacetamide (Figure S27-29). The difference between the selective probe **2i** and the non-selective probe **2b** is the number of fluorine substituents.

The concept was also applied in activity-based protein profiling (ABPP). Native or heat-denatured TEV protease was treated with an azide-functionalized probe (**6a**-**d**, 10 µM, Figure 4) at 37 °C for 16 h and then visualized by a cyanine "click" dye. Following gel electrophoresis, the scanning demonstrated the ability of probe **6b** and **6c** to label TEV protease in an activity-dependent manner, as these only labeled the native protease. The reduced reactivity of the denaturated protease may be explained by loss of the local activating eviroment in the active site. In contrast to **6b** and **6c**, probe **6a** was able to label both active and denatured TEV protease, while probe **6d** did not label any of the samples. This clearly demonstrates the concept of rational reactivity tuning. Thus, the excessively reactive flourobenzene probe **6a** could be "tuned down" by either replacing the sulfonamide with the less electron-withdrawing amide **6c** or by removing two fluoro-substituents (**6b**). However, combining both effects reduces the reactivity of the probe (**6d)** to a level that prohibits labeling of even the active protease.

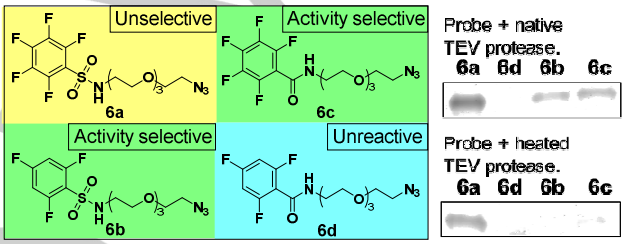


Figure 4. Activity-based protein profiling of TEV. Structures of **6a**-**d** and their labeling pattern of native or heat denatured TEV protease with the probes.

To demonstrate the application of the fluorobenzenes to ABPP in a cellular environment, bacterial cell lysates were treated with probe **6a, 6b** and **6d** with or without pretreatment with iodoacetamide. This revealed that the most reactive probe **6a**, labeled many protein bands with great intensity, whereas probe **6b** appeared much more selective. Probe **6d**, the least reactive, was not able to label any protein in the cell lysate. The pretreatment with iodoacetamide completely inhibited any labeling with the probes, supporting that the probes target the free cysteine residues. Moreover, probe **6b** was clearly able to label TEV protease within a spiked cell lysate (Figure 5).

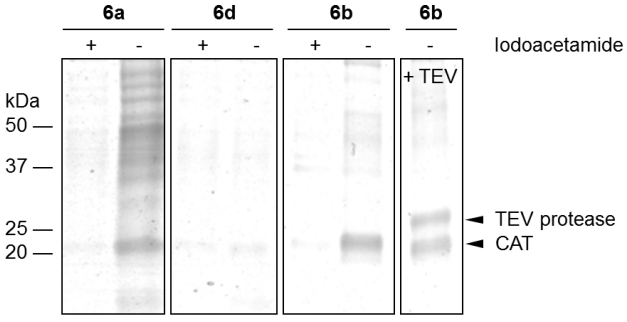


Figure 5. The selectivity of probes **6a**, **6b** and **6d** towards proteins in cell lysate. Labeling of TEV protease (spiked) and identification of chloramphenicol acetyl transferase (CAT) by probe **6b**. Reaction conditions: Pre-incubation of cell lysate/Tris buffer 1:1 (10 µL), with or without iodoacetamide (0.5 mM), 37 °C, 2 h, then **6a**, **6b** or **6d** in Tris buffer (10 µL, 20 µM), 37 °C, 16 h.

In the gels, another major protein around 25 kDa was also labeled by probe **6a** or **6b**. This band was isolated and analyzed by mass spectrometry after tryptic digestion. Subsequently, the protein was identified to be chloramphenicol acetyltransferase (CAT), which is an enzyme containing a cysteine (Cys31) located in the chloramphenicol binding site and susceptible to covalent modification.[21] Retrospectively, detecting this protein should be expected since the applied B834[DE3]pLysS bacteria contain the pLysS plasmid with the gene encoding for CAT.

In summary, it has been demonstrated that probes for site selective cysteine labeling, enzyme inhibition and protein profiling can be developed by rational and accurate tuning of the electron deficiency of fluorobenzene derivatives. The reactivity of the flourobenzene derivatives towards cysteine proved to correlate with the Hammett σ_p-constant of the para-substituent and with the number of additional fluorine substituents. Based on this knowledge, probes for selective cysteine labeling over all other amino acids were developed. The probes were demonstrated to be applicable for labeling of proteins with free cysteine residues such as BSA and eGFP. Chemoselective inhibition of a cysteine protease over a serine protease was also demonstrated. Developing less reactive probes allowed discrimination among cysteine residues with different physicochemical properties in the TEV protease, giving arylation only at the cysteine residue in the catalytic site in an activity-dependent manner. These probes were used for activity based protein profiling and for identifying proteins containing reactive cysteine residues in cellular lysates.

Acknowledgements

We are grateful to The University of Copenhagen for support through the UCPH – lighthouse program (Center for Evolutionary Chemical Biology).

Keywords: Fluorobenzene • Arylation • Protein Modification • Enzyme • Inhibitor

[1] a) B. F. Cravatt, A. T. Wright, J. W. Kozarich, *Annu. Rev. Biochem.* **2008**, *77*, 383; b) Y. Takaoka, A. Ojida, I. Hamachi, *Angew. Chem.* **2013**, *125*,; *Angew. Chem. Int. Ed.* **2013**, *52*, 4088.

[2] a) Y. Liu, M. P. Patricelli, B. F. Cravatt, *Proc. Natl. Acad. Sci. U.S.A.* **1999**, 96, 14694; b) G. C. Adam, E. J. Sorensen, B. F. Cravatt, *Nat. Biotechnol.* **2002**, 20, 805; c) E. F. Jansen, A. L. Curl, A. K. J. Balls, *Biol. Chem.* **1951**, 190, 557; (d) N. A. Thornberry, E. P. Peterson, J. J. Zhao, A. D. Howard, P. R. Griffin, K. T. Chapman, *Biochemistry* **1994**, 33, 3934; e) J. C. Powers, J. L. Asgian, Ö. D. Ekici, K. E. James, *Chem. Rev.* **2002**, 102, 4639; f) L. I. Willems, W. A. Van der Linden, N. Li, K. Y. Li, N. Liu, S. Hoogendoorn, G. A. Van Der Marel, B. I. Florea, H. S. Overkleeft, *Acc. of Chem. Res.* **2011**, 44, 718; g) D. K. Nomura, M. M. Dix, B. F. Cravatt, *Nat. Rev. Cancer* **2010**, 10, 630.

[3] J. Singh, R. C. Petter, T. A. Baillie, A. Whitty, *Nat. Rev. Drug. Discov.* **2011**, *10*, 307.

[4] S. M. Marino, V. N. Gladyshev, *J. Mol. Biol.* **2010**, *404*, 902.

[5] N. J. Pace, E. Weerapana, *ACS. Chem. Biol.* **2013**, *8*, 283.

[6] a) E. Basle, N. Joubert, M. Pucheault, *Chem. Biol.* **2010**, *17*, 213; b) O. Boutureira, G. J. Bernardes, *Chem. Rev.* **2015**, *115*, 2174.

[7] a) L. Faleiro, R. Kobayashi, H. Fearnhead, Y. Lazebnik, *EMBO. J.* **1997**, *16*, 2271; b) A. B. Berger, M. D. Witte, J. B. Denault, A. M. Sadaghiani, K. M. Sexton, G.S. Salvesen, M. Bogyo, *Mol. Cell.* **2006**, *23*, 509.

[8] D. Greenbaum, K. F. Medzihradszky, A. Burlingame, M. Bogyo, *Chem. Biol.* **2000**, *7*, 569.

[9] G. C. Adam, B. F. Cravatt, E. J. Sorensen, *Chem. Biol.* **2001**, *8*, 81.

[10] a) K. T. Barglow, B. F. Cravatt, *Chem. Biol.* **2004**, *11*, 1523; b) K. T. Barglow, B. F. Cravatt, *Angew. Chem.* **2006**, *118*, 7568; *Angew. Chem. Int. Ed.* **2006**, *45*, 7408; c) K. T. Barglow, K. S. Saikatendu, M. H. Bracey, R. Huey, G. M. Morris, A. J. Olson, R. C. Stevens, B. F. Cravatt, *Biochemistry* **2008**, *47*, 13514; d) E. Weerapana, C. Wang, G. M. Simon, F. Richter, S. Khare, M. B. Dillon, D. A. Bachovchin, K. Mowen, D. Baker, B. F. Cravatt, *Nature* **2010**, *468*, 790; e) N. J. Pace, E. Weerapana, *ACS. Chem. Biol.* **2014**, *9*, 258.

[11] a) H. L. Wong, D. C. Liebler, *Chem. Res. Toxico.* **2008**, *21*, 796; b) J. A. Doorn, D. R. Petersen, *Chem. Res. Toxicol.* **2002**, *15*, 1445; c) R. M. Lopachin, T. Gavin, D. R. Petersen, D. S. Barber, *Chem. Res. Toxicol*. **2009**, *22*, 1499; d) S. G. Kathman, Z. Xu, A. V. Statsyuk, *J. Med. Chem*. **2014**, *57*, 4969.

[12] a) C. Zhang, A. M. Spokoyny, Y. Zou, M. D. Simon, B. L. Pentelute, *Angew. Chem.* **2013**, 125, 14251; *Angew. Chem. Int. Ed*. **2013**, *52*, 14001; b) A. M. Spokoyny, Y. Zou, J. J. Ling, H. Yu, Y. Lin, B. L. Pentelute, *J. Am. Chem. Soc.* **2013**, *135*, 5946; c) A. D. de Araujo, H. N. Hoang, W. M. Kok, F. Diness, P. Gupta, T. A. Hill, R. W. Driver, A. D. Price, S. Liras, D. P. Fairlie, Angew. Chem. **2014**; *126*, 7085; *Angew. Chem. Int. Ed*. **2014**, *53*, 6965; d) D. A. Shannon, R. Banerjee, E. R. Webster, D. W. Bak, C. Wang, E. Weerapana, *J. Am. Chem. Soc.* **2014**, *136*, 3330; e) S. P. Brown, A. B. Smith III, *J. Am. Chem. Soc.* **2015**, *137*, 4034; f) E. V. Vinogradova, C. Zhang, A. M. Spokoyny, B. L. Pentelute, S. L. Buchwald, *Nature* **2015**, *526*, 687. g) C. Zhang, M. Welborn, T. Zhu, N. J. Yang, M. S. Santos, T. Van Voorhis, B. L. Pentelute, *Nat. Chem.* **2016**, *8*, 120; (h) S. Kalhor-Monfared, M. R. Jafari, J. T. Patterson, P. I. Kitov, J. J. Dwyer, J. M. Nussb, R. Derda, *Chem. Sci.* **2016**, 7, 3785; (i) D. Gimenez, A. Dose, N. L. Robson, G. Sandford, S. L. Cobb, C. R. Coxon, *Org. Biomol. Chem*. **2017**, 15, 4081.

[13] a) M. J. Evans, A. Saghatelian, E. J. Sorensen, B. F. Cravatt, *Nat. Biotechnol*. **2005**, *23*, 1303; b) M. J. Evans, G. M. Morris, J. Wu, A. J. Olson, E. J. Sorensen, B. F. Cravatt, *Mol. Biosyst*. **2007**, *3*, 495; c) E. Weerapana, G. M. Simon, B. F. Cravatt, *Nat. Chem. Biol*. **2008**, *4*, 405; d) W. T. Lowther, D. A. McMillen, A. M. Orville, B. W. Matthews, *Proc. Natl. Acad. Sci.* **1998**, *95*, 12153.

[14] G. Bulaj, T. Kortemme, D. P. Goldenberg, Biochemistry **1998**, 37, 8965.

[15] Z. Y. Zhang, J. E. Dixon. *Biochemistry* **1993**, *32*, 9340.

[16] a) F. Diness, D. P. Fairlie, *Angew. Chem.* **2012**, *124*, 8136; *Angew. Chem. Int. Ed.* **2012**, *51*, 8012; b) F. Diness, M. Begtrup, *Org. Lett.* **2014**, *16*, 3130; c) C. B. Jacobsen, M. Meldal, F. Diness *Chem. Eur. J.* **2017**, *23*, 846.

[17] C. Hansch, A. Leo, R. W. Taft, *Chem. Rev.* **1991**, *91*, 165.

[18] F. Yang, L. G. Moss, G. N. Jr. Philips, *Nat. Biotechnol.* **1996**, *14*, 1246.

[19] K. A. Majorek, P. J. Porebski, A. Dayal, M. D. Zimmerman, K. Jablonska, A. J. Stewart, M. Chruszcz, W. Minor, *Mol. Immunol.* **2012**, *52*, 174.

[20] a) J. Phan, A. Zdanov, A. G. Evdokimov, J. E. Tropea, H. K. Peters III, R. B. Kapust, M. Li, A. Wlodawer, D. S. Waugh, *J. Biol. Chem.* **2002**, *277*, 50564; b) R. B. Kapust, D. S. Waugh, *Protein Expr. Purif.* **2000**, *19*, 312.

[21] A. Lewendon, W. V. Shaw, *Biochem. J.* **1990**, 272, 499.

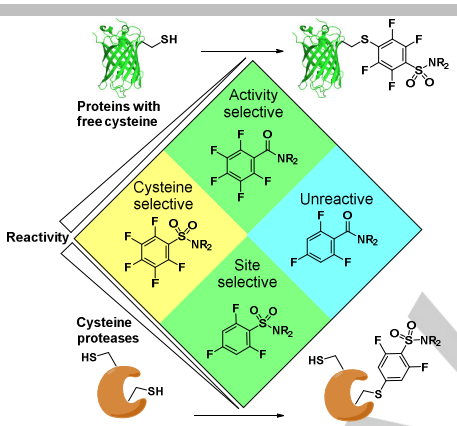
10.1002/anie.201712589

WILEY-VCH

Entry for the Table of Contents

COMMUNICATION

The reactivity of chemical probes for selective cysteine S-arylation in proteins has been accurately tuned by rational selection of substituents in a two-dimensional fashion. The optimized probes are demonstrated applicable to protein labelling of GFP and albumin, selective inhibition of cysteine proteases, selective functionalization among cysteine residues and for activity based protein profiling in cell lysates.



Ahmed M. Embaby, Sanne Schoffelen, Christian Kofoed, Morten Meldal and Frederik Diness**

Page No. – Page No.

Rational tuning of fluorobenzene probes for cysteine-selective protein modification