UNIVERSITY OF COPENHAGEN

https://doi.org/10.17952/37EPS.2024.P2210

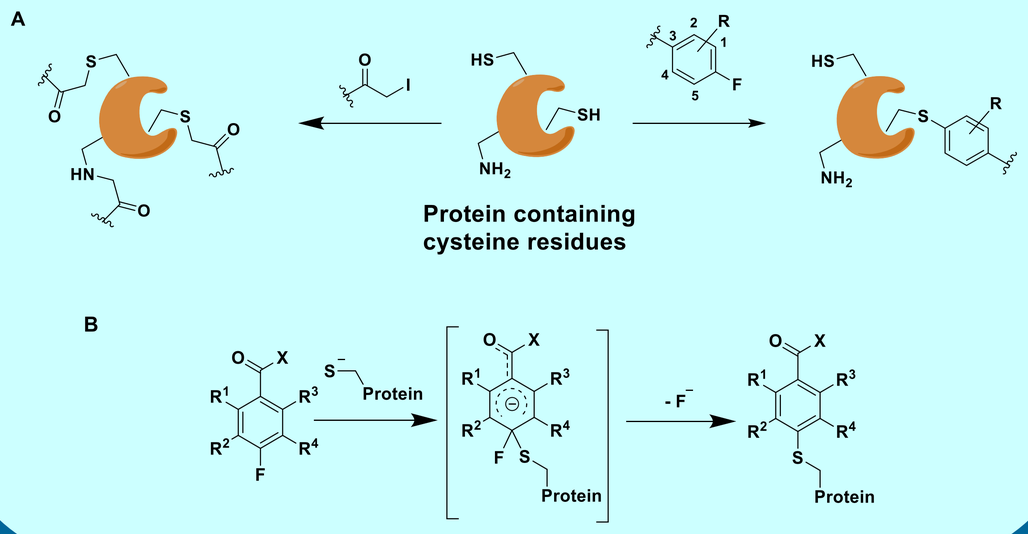
Chemoselective arylation of amino acids, peptides and proteins

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

Center for Evolutionary Chemical Biology, Department of Chemistry, University of Copenhagen, Copenhagen, Denmark. *fdi@chem.ku.dk

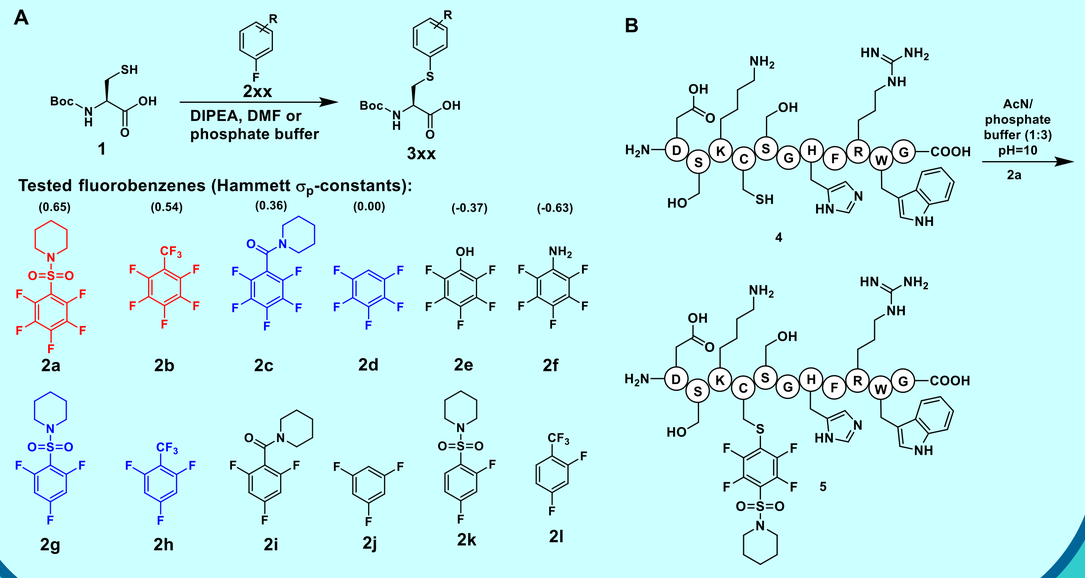
1. Introduction

Methods for identification and regulation of cysteine containing proteins are essential as these proteins have key roles in diseases, including parasite infections, osteoporosis and cancer. Within this area many chemical tools for covalent protein modification have been developed. However, the traditional applied electrophiles (e.g. iodoacetamide), have in general little room for tuning of the reactive carbon. Benzene, on the other hand, is interesting as a core for developing selective reactive probes, as it provides five positions for tuning of reactivity, and attachment of other functionalities of interests, in addition to the electrophilic carbon. Based on our experience with nucleophilic aromatic substitution (S_NAr) reactions, we hypothesized that molecular tools for selective cysteine labeling in water could be developed through fine tuning the reactivity of fluorobenzenes by variation of number of fluorine atoms in combination with other substituents.[1,2,3]



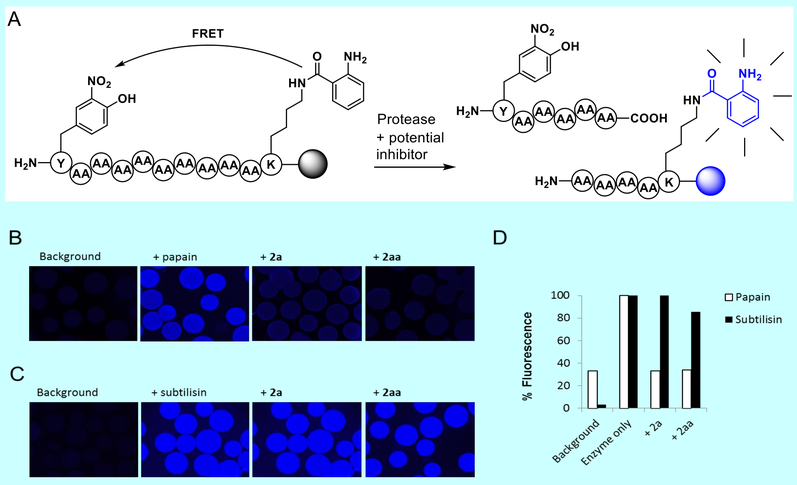
2. Reactivity screening of fluorobenzenes

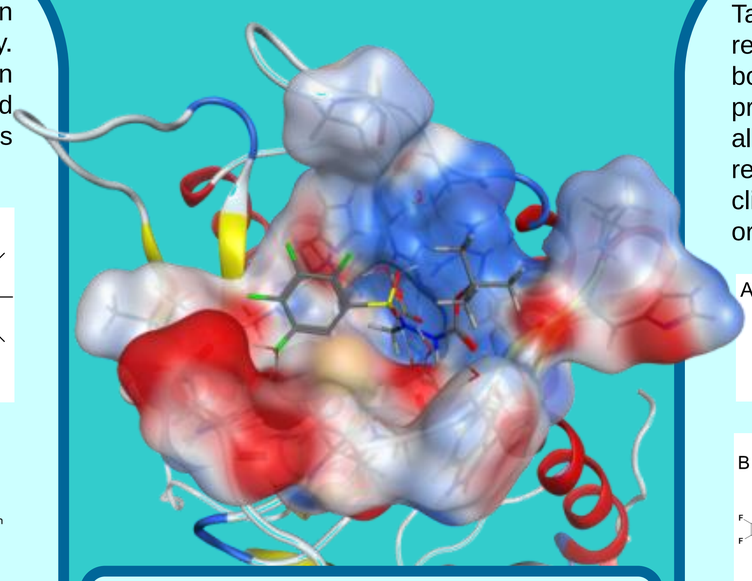
As a model system, the reactivity of fluorobenzene derivatives towards *tert*-butyloxycarbonyl protected cysteine was tested. Under promoting reaction conditions using DMF and DIPEA the compounds presented in red and blue reacted. The remaining derivatives (in black) did not react. Under less promoting reaction condition, in aqueous buffer only the examples presented in red reacted. The reactivity of the fluorobenzene derivatives could be linked to the electron withdrawing capacity (Hammett σ_p-constant) of the substituents as well as the number of fluoro substituents. To test for selectivity towards cysteine over other amino acids, peptide **4** was reacted with an excess (3 eq.) of **2a** under aqueous condition (see Scheme 3). Only a the monosubstituted adduct **5** was observed and identified by LC-MS and MSMS.



3. Selective cysteine protease inhibition

Selective inhibition of papain (cysteine protease) over subtilisin (serine protease) was explored using an on-bead inhibition assay. The enzymes were pre-incubated with the fluorobenzenes and then added to polymer bound FRET substrates. Compound **2a** and pentafluoronitrobenzene (**2aa**) completely inhibited papain, whereas subtilisin was still fully active under identical conditions.

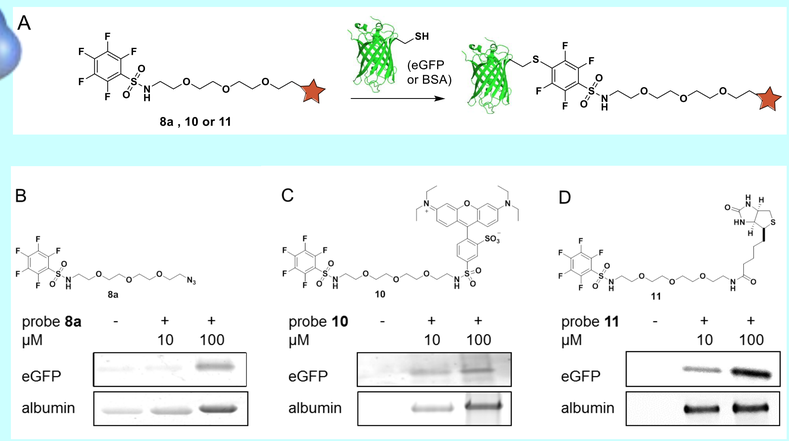




Model of cysteine protease with covalent bound inhibition

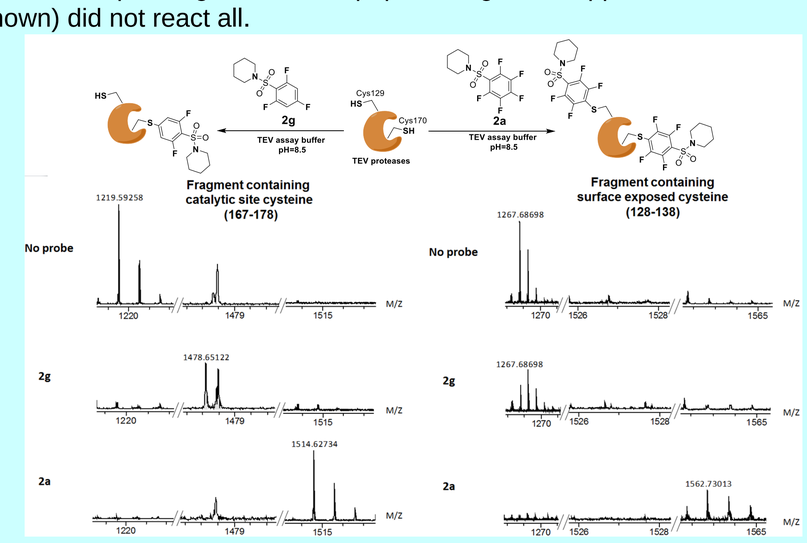
4. Protein labelling

Tag-containing probes based on **2a** were synthesized and then reacted with enhanced green fluorescent protein (eGFP) or with bovine serum albumin (BSA). All three types of reporter-linked probes were capable of labeling the free thiols of eGFP and albumin in a concentration-dependent manner. Gratifyingly, the reactive fluorobenzene probes did not interfere with the functional click, the fluorescence or the affinity tag under the protein labeling or the detection conditions.



5. Selective reaction among cysteine residues

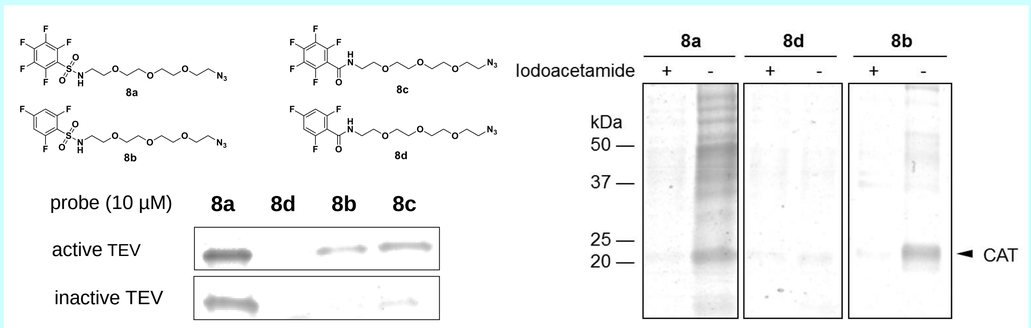
Three fluorobenzene derivatives with different reactivities (**2a**, **2g** and **2i**) were investigated for their ability to react selectively with only one of the unpaired cysteine residues in Tobacco Etch Virus (TEV) protease. After tryptic digest of the native TEV protease (negative control), the peptide fragment with cysteine in the active site (Cys170) was found at 1219.59 and that of the fragment with a surface exposed cysteine (Cys129) at 1267.68 Da. The signal at 1219.59 Da disappeared upon pre-treatment with **2g**, and a new mass peak at 1478.65 corresponding to the modified fragment appeared. The fragment with Cys129 remained unchanged. In contrast, both of the native signals completely disappeared upon reaction with compound **2a** and the mass peaks corresponding to modified peptide fragments appeared. On the other hand **2i** (not shown) did not react all.



6. Activity Based Protein Profiling

To investigate the applicability of fluorobenzenes for activity-based protein profiling (ABPP), native or heat-denatured TEV protease was treated with compound **8a**, **8b**, **8c** or **8d**. Probe **8b** and **8c** proved able to label TEV protease in an activity-dependent manner. In contrast, probe **8a** was able to label both active and denatured TEV protease, while probe **8d** did not label any of the samples.

The probes were also tested towards bacterial cell lysates, with or without pretreatment with iodoacetamide. The most reactive probe (**8a**), labeled many bands with great intensity, whereas probe **8b** appeared much more selective and probe **8d** did not label any protein in the cell lysate. The pre-treatment with iodoacetamide inhibited the labeling with the active probes completely, supporting that these probes selectively target cysteine residues. A major labeled protein band around 25 kDa was analyzed by mass spectrometry after tryptic digestion. Gratifyingly, the protein proved to be chloramphenicol acetyltransferase (CAT), an enzyme containing a cysteine in the catalytic site.



1) Embaby, A. M.; Schoffelen, S.: Kofoed, C.; Meldal, M.; Diness, F. *Angewandte Chemie International Edition,* **2018**, *57*, 8022-8026
2) Diness, F.; Fairlie, D. P. *Angewandte Chemie International Edition*, **2012**, *51*, 8012-8016
3) de Araujo, A. D.; Hoang, H. N.; Kok, W. M.; Diness, F.; Gupta, P.; Hill, T. A.; Driver, R. W.; Price, D. A.; Liras, S.; Fairlie, D. P. *Angewandte Chemie International Edition*, **2014**, *53*, 6965-6969