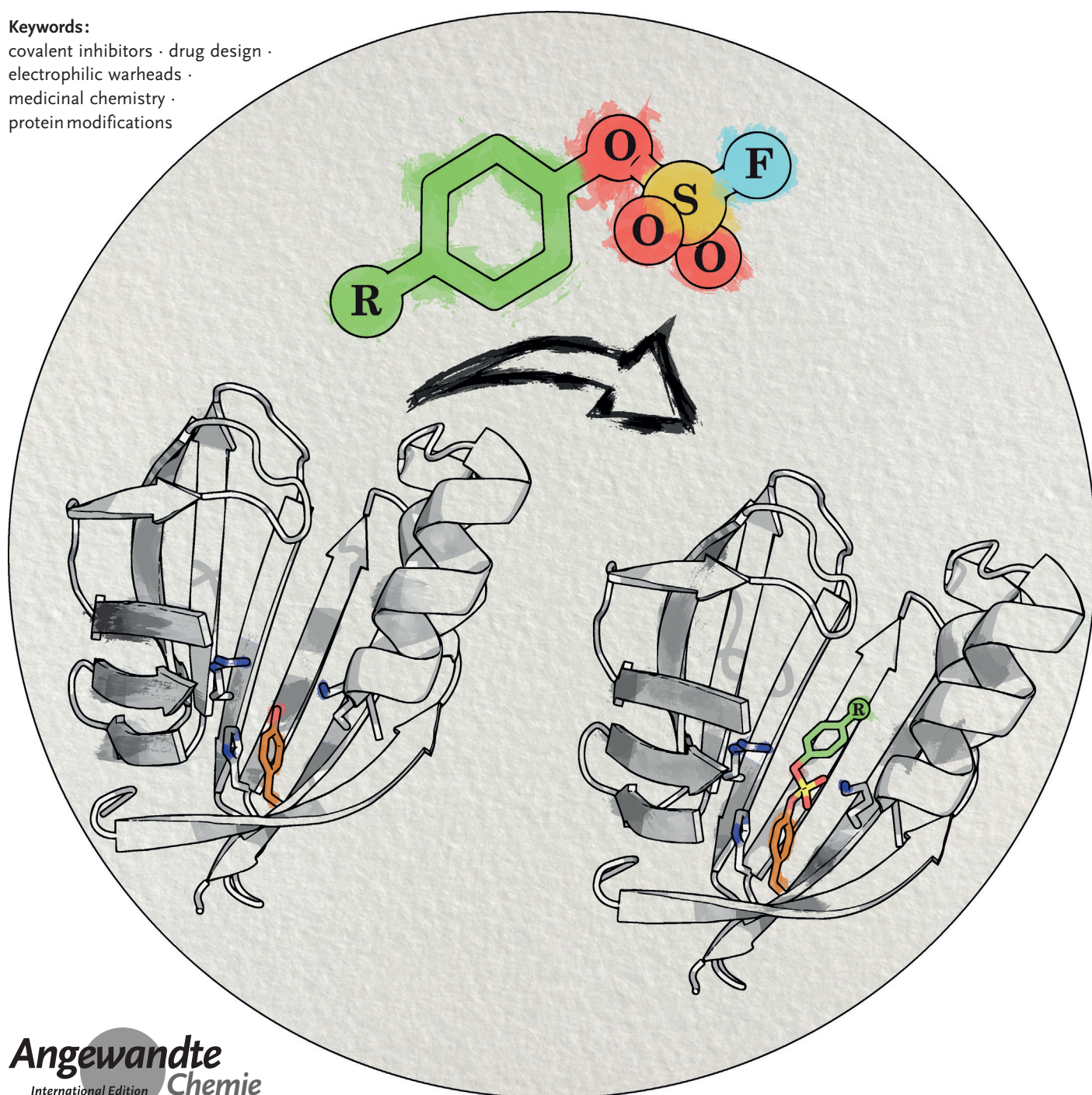


Arylfluorosulfate-Based Electrophiles for Covalent Protein Labeling: A New Addition to the Arsenal

Pablo Martín-Gago* and Christian A. Olsen*

Keywords:

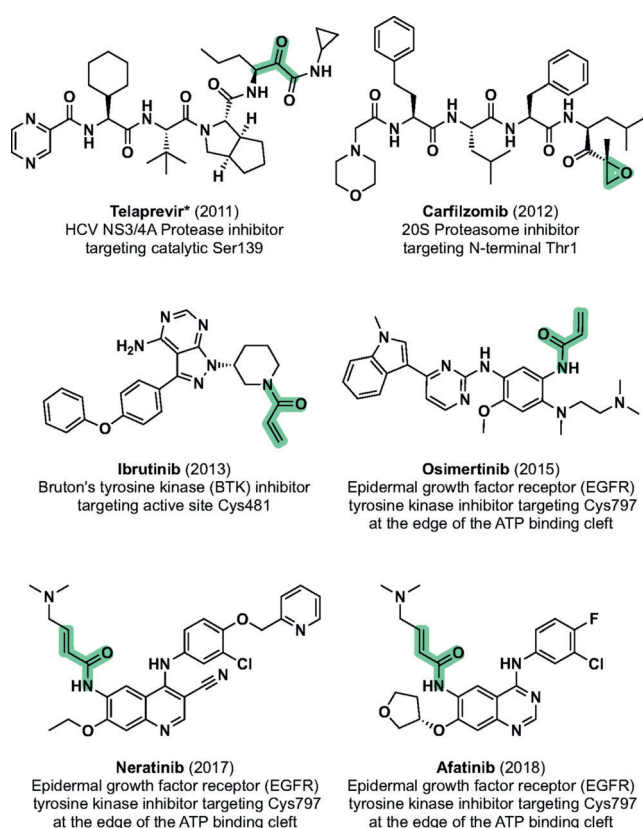
covalent inhibitors · drug design ·
electrophilic warheads ·
medicinal chemistry ·
protein modifications



Selective covalent modification of a targeted protein is a powerful tool in chemical biology and drug discovery, with applications ranging from identification and characterization of proteins and their functions to the development of targeted covalent inhibitors. Most covalent ligands contain an affinity motif and an electrophilic warhead that reacts with a nucleophilic residue of the targeted protein. Because the electrophilic warhead is prone to react and modify off-target nucleophiles, its reactivity should be balanced carefully to maximize target selectivity. Arylfluorosulfates have recently emerged as latent electrophiles for selective labeling of context-specific tyrosine and lysine residues in protein pockets. Here, we review the recent but intense introduction of arylfluorosulfates into the arsenal of available warheads for selective covalent modification of proteins. We highlight the untapped potential of this functional group for use in chemical biology and drug discovery.

1. Introduction

The use of covalent protein ligands for interrogating protein function has proven extremely helpful in chemical biology. However, translation of probes into covalent drugs was abolished from drug discovery programs for a long time,



Scheme 1. Chemical structures, FDA approval year and targeted residues of recently approved drugs that covalently bind their corresponding targets. The covalently targeted electrophiles are shaded. *Telaprevir forms a reversible hemiketal to the catalytic serine of HCV NS3/4A protease.

mainly due to concerns regarding their idiosyncratic toxicity. In recent years, however, the development of covalent inhibitors has received a renewed attention, influenced by an exponentially increasing literature of successful examples,^[1] a better understanding of mechanisms of action,^[2] refined compound selectivity,^[3] and several drug approvals (e.g., telaprevir, carfilzomib, ibrutinib, osimertinib, neratinib, and afatinib; Scheme 1).

Most covalent ligands contain an affinity motif that enables selective binding to a protein active site and, in turn, directs the reactive warhead towards a nucleophilic residue of the target protein. Ideally, this facilitates formation of a covalent, irreversible bond with just the intended functional

group among thousands of nucleophiles in the proteome. While a range of covalent ligands has been developed to target a diverse array of proteins, most of them were designed to target an accessible, suitably located nucleophilic cysteine residue in the protein pocket. The unique reactivity of cysteine allows for acrylamides to be the electrophilic functionalities of choice in these cases. These soft electrophiles require proximity to the nucleophilic, targeted cysteine to react, thus minimizing off-target labeling. It is often the case that the protein of interest lacks a cysteine residue amenable to modification, since cysteines are only found in 1.5 % of protein pockets.^[4]

Exploiting less nucleophilic residues for covalent labeling in protein binding sites has been less common, but there is a rapidly increasing number of examples of compounds capable of selective covalent targeting of tyrosine,^[5] lysine,^[6] serine,^[7] threonine,^[8] as well as glutamic acid^[9] residues in the context of activity-based proteome profiling (ABPP) probes or covalent inhibitors. The development of irreversibly binding probes and inhibitors that rely on covalent modification of Tyr, Lys or Ser residues generally requires compounds bearing considerably reactive electrophilic warheads that can form a covalent bond with the targeted residue. Additionally, the surrounding protein environment enhances the

[*] Prof. Dr. P. Martín-Gago, Prof. Dr. C. A. Olsen
Center for Biopharmaceuticals &
Department of Drug Design and Pharmacology
University of Copenhagen
Universitetsparken 2, 2100 Copenhagen (Denmark)
E-mail: pabm@lundbeck.com
cao@sund.ku.dk

The ORCID identification number(s) for the author(s) of this article can be found under:
<https://doi.org/10.1002/anie.201806037>.

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

nucleophilicity of the targeted residues by perturbing the pK_a values of their side chains. For example, the ϵ -amino groups of lysine residues display an intrinsic pK_a of 10.4, and are most likely positively charged and unavailable for covalent modification under physiological conditions.^[10] However, an adequate protein environment that effectively buries a particular lysine and/or surrounds it with the appropriate residues can contribute to depress the pK_a of its ϵ -amino group down to five units, making it readily available as a nucleophile.^[11]

Careful selection and optimization of warhead reactivity is essential to maximize target selectivity and minimize the modification of unwanted endogenous nucleophiles. Sulfonyl fluoride-based warheads have been extensively used for the chemical biology community for covalent modification of a range of proteins, and this subject has been excellently reviewed in detail by Narayanan and Jones.^[12] Sulfonyl fluorides have been shown to bind covalently to Tyr, Lys and Ser amino acids in protein binding sites, but also to Thr, Cys, and His residues (Scheme 2a).^[12,13] Sulfonyl fluoride-containing ligands have a balanced reactivity and modest aqueous stability and are incredibly useful tools in chemical biology, particularly explored as ABPP probes. However, the relatively high reactivity and promiscuity of sulfonyl fluoride-based warheads may jeopardize the efficiency of these probes in whole-cell investigations or experiments that require

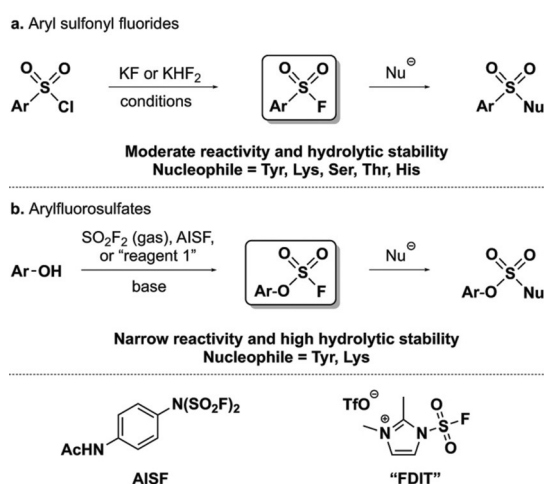
extended incubation times, which is an important limitation for the development of covalent inhibitors in drug discovery campaigns.

A much more stable and closely related sulfur(VI)-fluoride group, the arylfluorosulfates (Scheme 2b), has recently emerged as an alternative warhead for covalent modifications of context specific Tyr and Lys, but also Ser residues in protein binding sites that comprise an appropriate protein microenvironment for the sulfur(VI)-fluoride exchange (SuFEx) reaction to occur.^[14]

In this Minireview we highlight the unique properties of fluorosulfate-based warheads in the context of other options, provide examples of covalent protein modification by fluorosulfate-derived probes, and discuss their potential in probe and covalent inhibitor development for drug discovery and general chemical biology investigations.

2. Reactivity of Arylfluorosulfate-Based Warheads

Sulfonyl fluoride-derived warheads were broadly popularized in the 1960s as covalent protein modifiers after early reports describing their mechanism of action.^[15] Irreversible protein binders containing sulfonyl fluorides that take advantage of their relatively low reactivity for selective covalent inhibition of a range of proteins have been continuously developed since then.^[12] On the other hand, despite arylfluorosulfates being first reported in the 1930s,^[16] the lack of robust synthetic methods for their preparation resulted in only a few reports entertaining this functional group, and it was not adapted as warhead for covalent modification of proteins until 2015 by Kelly and co-workers.^[17] The resurgence of arylfluorosulfates was pioneered by the Sharpless group due to development of a facile method for the synthesis of these compounds (Scheme 2b).^[14a] Simply incubating phenols with sulfuryl fluoride (gas) and base selectively provides their corresponding arylfluorosulfate derivatives even in the presence of amines, carboxylates, and aliphatic alcohols. Moreover, two bench-stable “F-SO₂⁺” donor reagents have subsequently been reported for arylfluorosulfate syntheses (AISF and FDIT, Scheme 2), thus overcoming the inconvenient handling of sulfuryl fluoride.^[18] In this regard, another practical transformation of phenols to arylfluorosulfates was also developed by *ex situ* generation of SO₂F₂ by means of a two-chamber reactor.^[19]



Scheme 2. Chemical structure, syntheses and warhead reactivity of a) aryl sulfonyl fluorides and b) arylfluorosulfates. AISF = [4-(acetylamino)phenyl]-imidodisulfonyl difluoride. FDIT = 1-(fluorosulfonyl)-2,3-dimethyl-1H-imidazol-3-ium trifluoromethanesulfonate.



Pablo Martín-Gago is currently an assistant professor at Copenhagen University. He completed his Ph.D. in 2013 at the Institute for Research in Biomedicine (IRB Barcelona), under the supervision of Prof. A. Riera. After a research stay in the group of Prof. G. C. Fu at Caltech, he joined the Max Planck Institute of Molecular Physiology in Dortmund for postdoctoral studies with Prof. H. Waldmann, developing inhibitors of the KRas-PDE δ interaction.



Christian A. Olsen received his Ph.D. from the Danish University of Pharmaceutical Sciences. After independently working on the development of novel peptidomimetics, he did his postdoctoral fellowship at The Scripps Research Institute. In 2010 he returned to the Technical University of Denmark and in 2014 he accepted his current position as professor at the University of Copenhagen. He is recipient of the Lundbeck Foundation Fellowship, the EFMC prize for a young medicinal chemist in academia, and an ERC Consolidator grant.

Unlike sulfonyl fluorides, arylfluorosulfates have proved to display remarkable stability towards hydrolysis and nucleophiles in general, since they remain intact in phosphate buffer (pH 10) for up to two weeks and in neutral buffers for months. In addition, arylfluorosulfates have been shown to be virtually unreactive in the presence of amino acids or natural products containing an array of different functional groups.^[14a,20]

In the context of covalent protein modification, this superior stability of arylfluorosulfates compared to sulfonyl fluorides should lead to an even more selective protein labeling and minimized off-target modification. Granted that the ligand-binding motif provides sufficient residence time projecting the warhead towards the targeted residue, latent arylfluorosulfate electrophiles have been shown to exclusively react upon activation in particular binding sites. Thus, the surrounding protein environment requires basic residues (Arg, Lys, or His) that decrease the pK_a of the targeted nucleophilic residue and/or that facilitate fluoride ion departure (e.g. by hydrogen bonding interactions) to achieve the SuFEx reaction.^[21]

3. Arylfluorosulfate-Based Targeted Covalent Inhibitors

In the first reported attempt to use fluorosulfates for covalent protein modification, Kelly and co-workers replaced the sulfonyl fluoride-based warhead in previously reported fluorogenic transthyretin (TTR) probes^[22] with an $ArOSO_2F$ group (Scheme 3a).^[17] Unlike the sulfonyl fluoride-containing probes **1** and **2**, arylfluorosulfates **3** and **4** reacted slowly with the targeted pK_a -perturbed Lys15 ϵ -amino group of transthyretin and did not reach full conversion, due to their lower reactivity. The additional oxygen atom present in the fluorosulfate-based warhead could also contribute to the inefficient covalent binding, by decreasing the proximity of the fluorosulfate sulfur atom to the nucleophilic Lys15 side chain or by geometrically compromising the hydrogen bond-mediated stabilization of the leaving fluoride atom by neighboring residues. In addition, probes **3** and **4** proved to be essentially unreactive towards the whole proteome.

Interestingly, adducts of TTR with probe **3** or **4** were not detected after more than 8 h incubation time. Only SO_3^- -

modified TTR, the corresponding hydrolysis product (Scheme 3b), was observed by mass spectrometry. This protein-mediated hydrolysis of the $ArOSO_2$ -covalent adducts appears to be intrinsic to TTR, since hydrolysis of a TTR-**16** (Scheme 5b) conjugate with formal transfer of SO_3^- to Lys15 was also observed over several days.^[23]

The propensity and potential applicability of arylfluorosulfates for irreversible protein Lys sulfamation and/or Tyr sulfation remains to be explored in further detail. In a similar approach, attempting to overcome the limited stability of a sulfonyl fluoride-based covalent ligand of the mRNA decapping scavenger enzyme (DcpS),^[24] Jones and co-workers incorporated an arylfluorosulfate electrophile (**6**) into the previously developed DcpS covalent inhibitor **5** (Scheme 4).^[25] Remarkably, while the sulfonyl fluoride probe **5** was designed and proved to covalently label Tyr143 in the binding pocket of DcpS (Figure 1a), fluorosulfate **6** covalently labeled DcpS at the non-catalytic Ser272 (Figure 1b). The formation of the covalent adduct was shown by intact mass LC-MS and native electrospray ionization (ESI) MS analysis, which also revealed the presence of Ser272 dehydrated-DcpS, that occurs via β -elimination of the fluorosulfate-DcpS conjugate (Scheme 4). Formation of Dha272-DcpS upon covalent arylfluorosulfate binding takes place rapidly under acidic LC-MS conditions, but slowly at pH 7.4, as shown by native ESI-MS. Insights into the general dehydration tendency of Ser- and Thr-modified residues in protein binding sites using $ArOSO_2F$ -based warheads would be

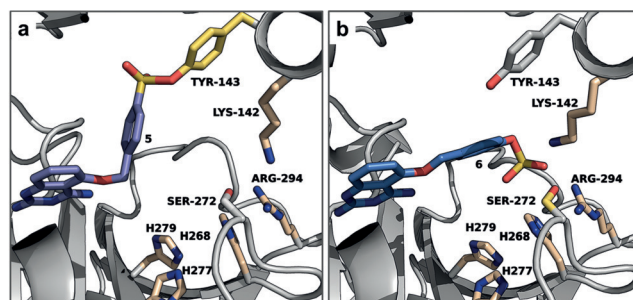
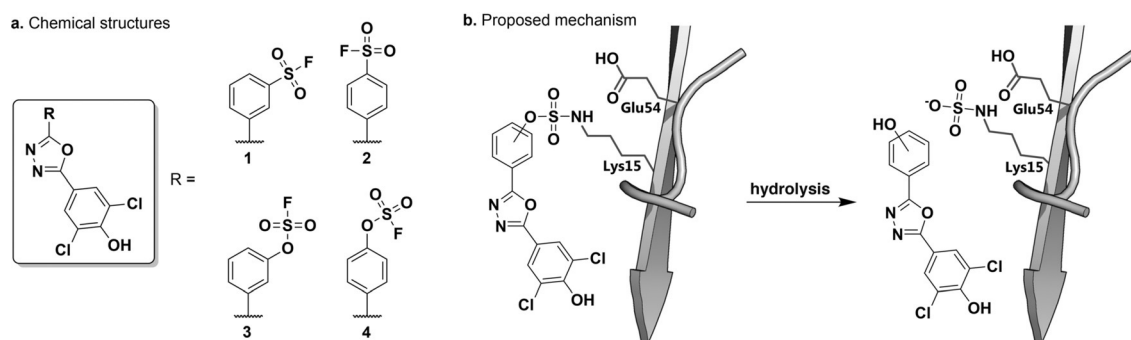
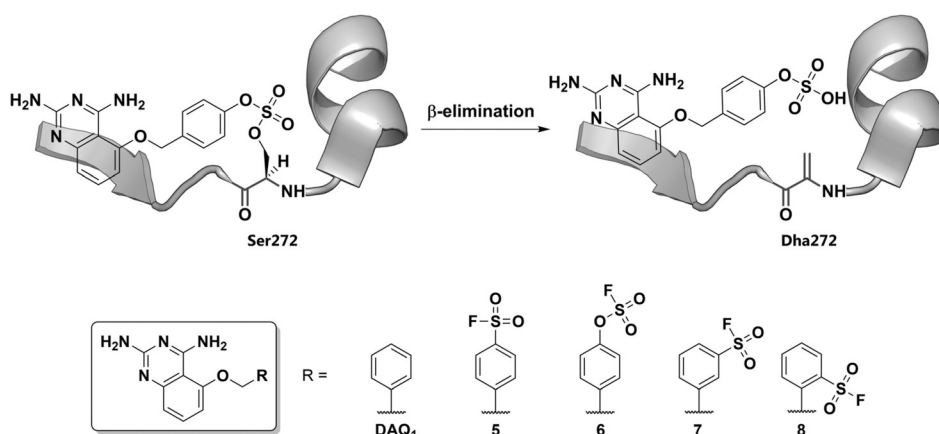


Figure 1. Binding mode of covalent DcpS targeted-inhibitors. a) Co-crystal structure of **5** and DcpS (PDB 4QDV). b) Docking pose of **6** in the DcpS binding pocket. Key residues that could catalyze the SuFEx reaction of the arylfluorosulfate with Ser272 are shown as sticks.



Scheme 3. TTR fluorogenic covalent probes. a) Chemical structures of arylfluorosulfate- and aryl sulfonyl fluoride-based TTR probes. b) Proposed mechanism for the observed formal Lys15 sulfamation. TTR = transthyretin.



Scheme 4. DcpS covalent probes. Top: proposed mechanism for the observed formal Ser272 dehydration. Bottom: chemical structures of non-covalent probe **DAQ1**, fluorosulfate- and sulfonyl fluoride-based probes.

highly valuable for the future development of arylfluorosulfate-based targeted covalent inhibitors.

While the shorter sulfonyl fluoride warhead of **5** was optimally accommodated towards the Tyr143 side chain (Figure 1a), the additional bridging oxygen atom of the arylfluorosulfate warhead in compound **6** results in improved positioning of the electrophilic sulfur(IV) atom towards Ser272 (Figure 1b).

This non-catalytic serine is neighbored by two basic amino acids (Lys142 and Arg294) and located close to three histidine residues (His268, 277, and 279). This assortment of basic side chains likely contributes to the SuFEx reaction by lowering the pK_a of the hydroxyl side chain of Ser272 and/or helping the departure of the fluoride atom. Interestingly, while the *para*-isomer **5** labels Tyr143 in the DcpS binding pocket, the *ortho*- and *meta*-regioisomers **7** and **8** (Scheme 4) were predicted to locate their electrophilic warheads differently and indeed proved to react with the proximate Tyr113 phenol side chain instead.^[24] Investigation of the corresponding *ortho*- and *meta*-arylfluorosulfate derivatives would provide further insights into their preferential positioning and reactivity. It would also provide valuable information for future attempts to design more stable and selective arylfluorosulfate-based covalent inhibitors based on known sulfonyl fluoride-derived compounds. Despite the similarity of both sulfur(VI)-fluoride groups, the length and geometry of their warheads differ due to the additional oxygen atom present in the fluorosulfate moiety, and we predict that optimization may be required on a case-by-case basis. Further characterization of the stability of these compounds revealed that **6** remained intact for up to 24 h in PBS (pH 7.4), while the analogous sulfonyl fluoride **5** was largely hydrolyzed under the same experimental conditions. For a comparable measurement of non-specific covalent protein labeling, **5** and **6** were incubated with human serum albumin (HSA), and the reactions were followed by LC-MS. While **5**-HSA and **5**($\times 2$)-HSA covalent adducts formed after 24 h incubation, covalent labeling with **6** was not detected, as a result of the superior kinetic stability and minimal promiscuity of arylfluorosulfate-based warheads.

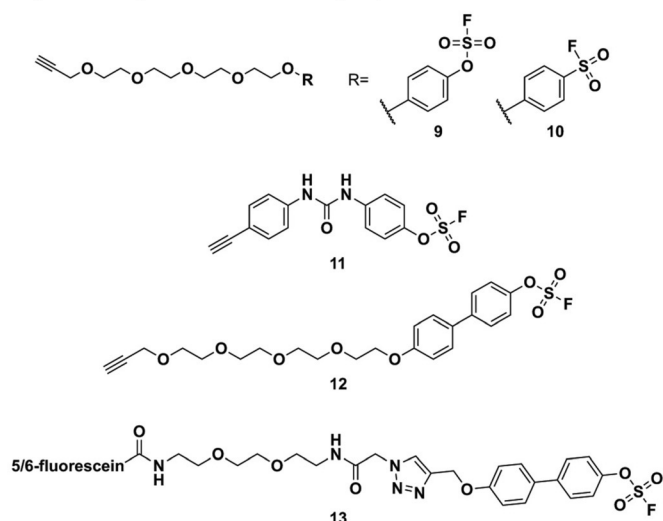
In order to sort out the plausible use of arylfluorosulfate-based warheads in drug discovery, we went a step further and provided metabolic stability and membrane permeability data of both **6** and **DAQ1** (Scheme 4), an analogous inhibitor lacking the $-\text{OSO}_2\text{F}$ functionality. Interestingly, **6** showed a reduced metabolic turnover rate in human liver microsomes. Arylfluorosulfate **6** was shown to migrate slower than **DAQ1**, but still displayed acceptable membrane permeability, although the enhanced apparent permeability from the B to A side suggests that efflux is occurring, a behavior not observed for **DAQ1**. The collection of more ADME data from potential arylfluorosulfate-based covalent inhibitors and their analogs lacking the corresponding electrophile will be important to gain insights into the potential use of fluorosulfate warheads in drug discovery research.

4. Reactivity and Selectivity of Simple Arylfluorosulfate Probes Towards the Human Proteome

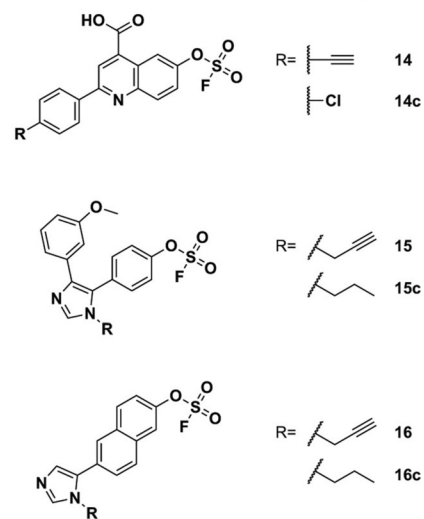
Due to their lower reactivity, arylfluorosulfate probes provide a more selective covalent labeling of proteins than the extensively used aryl sulfonyl fluorides in whole human proteome experiments.^[17] The Kelly and Sharpless groups compared the proteome reactivity of both functional groups and showed that structurally simple arylfluorosulfate-containing probes **9** and **11** (Scheme 5a) covalently modified a low number of proteins in HeLa cell culture and lysate, while the analogous aryl sulfonyl fluoride probe **10** (Scheme 5a) labeled a wide array of proteins under the same conditions.^[21]

SDS-PAGE separation of proteins from HeLa cells treated with **9** revealed that this probe selectively binds several members of the proteome in the 15 kDa range. Stable isotope labeling by amino acids in cell culture (SILAC)^[26] experiments were performed to identify these 15 kDa targets to be the fatty acid binding proteins 3 and 5 (FABP3 and FABP5) and the cellular retinoic acid binding protein 2 (CRABP2), which are members of the intracellular lipid binding protein (iLPB) family.^[27] Complete covalent adduct

a. Arylfluorosulfate probes of low structural complexity



b. Probes with affinity motifs of medium complexity



Scheme 5. Chemical structures of a) simple alkyne-functionalized fluorosulfate (**9**, **11**, **12**, and **13**) and sulfonyl fluoride (**10**) probes and b) alkyne-containing arylfluorosulfate probes (**14**, **15**, and **16**) and competitors (**14c**, **15c**, and **16c**) with intermediate structural complexity.

formation was observed upon incubation of recombinant CRABP2 (2 μM) and **9** (100 μM) after 48 h, and 80 % covalent labeling of recombinant FABP5 was achieved under the same experimental conditions. In addition, recombinant human FABP4, a structurally similar iLBP, formed 90 % covalent adduct with **9** as seen by LC-MS analysis. However, covalent binding of this probe to FABP3 could not be validated in this or other attempted experiments. In order to accelerate covalent bond formation by enhancing the potency of the affinity motif, biphenyl-containing probes **12** and **13** (Scheme 5) were synthesized and yielded quantitative covalent adduct formation with recombinant CRABP2 within 1 h. Probes **12** and **13** (100 μM) showed a remarkable and unexpected selectivity for CRABP2, since labeling of FABP3, FABP4, and FABP5 (2 μM) was not detected, even at 24 h incubation times. Selective and nearly complete covalent

labeling of overexpressed GFP-CRABP2 was achieved with the more cell-permeable probe **12** (20 μM) in living HEK293T cells overexpressing both GFP-CRABP2 and FABP5-GFP.

iLBPs contain an arylfluorosulfate-desirable conserved Arg-Tyr-Arg motif in their binding pocket (Figure 2).^[27] Tandem mass spectrometry experiments^[28] revealed that arylfluorosulfate probes bind iLBPs via a conserved Tyr residue in their fatty acid-binding sites. The Arg duo perturb the pK_a of the Tyr phenol side chain and/or stabilize the departing fluoride during the sulfur(VI)-fluoride exchange reaction. Incubation of **13** and recombinant CRABP2 in buffers ranging from pH 4.9 to 10.4 revealed that **13** labels CRABP2 in a pH-dependent manner (labeling efficiency increases with pH value), and that the phenol side chain of Tyr134 displays an apparent pK_a of 7.6, indicating pK_a perturbation of the phenol by Arg111 and/or Arg132. To gain insight into the role of the Arg pair in covalent bond formation, CRABP2 mutants Arg111Leu and Arg132Leu were incubated with **13** at different pHs. Both point mutations resulted in a markedly decreased covalent labeling even at pH 10.4, which is higher than the pK_a of standard Tyr side chains, suggesting that the neighboring arginines are directly involved in catalyzing the SuFEx reaction.

A crystal structure resulting from probe **12** in complex with CRABP2 unambiguously proved the formation of a diarylsulfate diester bond between ligand and Tyr134 and showed that the biaryl group engaged in hydrophobic interactions with several residues in the retinoic acid (RA) binding site (Figure 2).^[29] The crystal structure revealed that the Thr75-Gly78 β -turn in CRABP2 provides further space to accommodate the outer ring of **12** than other iLBPs, which explains the significant CRABP2 selectivity of the biphenyl-based probes.^[21] In addition, **12**-mediated irreversible inactivation of CRABP2 decreased the transcriptional activation of retinoic acid receptor alpha (RAR α), which regulates transcription in a ligand-dependent manner,^[30] by inhibiting retinoic acid delivery.

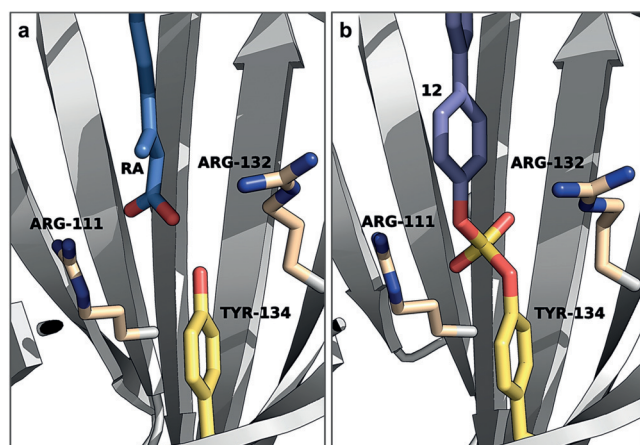


Figure 2. Binding mode of retinoic acid (RA) and **12** to CRABP2. a) Co-crystal structure of RA and CRABP2 (PDB 2FR3). b) Co-crystal structure of **12** and CRABP2 (PDB 5HZQ). Key residues are shown as sticks.

5. Inverse Drug Discovery Using Arylfluorosulfate Probes

With the knowledge that simple arylfluorosulfates display a proteome preference for iLBPs, the Sharpless, Wilson, and Kelly groups ventured into using more structurally complex arylfluorosulfate electrophile probes in an inverse drug discovery approach.^[23] In this strategy, a small number of arylfluorosulfate clickable probes with intermediate structural complexity was exposed to the human proteome and selectively labeled protein targets were identified and validated by means of mutagenesis experiments, mass spectrometry analysis, and X-ray crystallography. Three different arylfluorosulfate probes were selected, which contained different aryl motifs, functional groups, as well as H-bond donors and acceptors. These chemical features endow the probes with sufficient reversible binding energy to preferentially bind particular groups of proteins, and together with the narrow reactivity of the arylfluorosulfate electrophile, should result in a highly selective covalent labeling of different proteome members.

Living HEK293T cells and lysates were treated with alkyne containing probes **14**, **15**, and **16** alone or in the presence of an excess of competitor (**14c**, **15c**, or **16c**, respectively) previous to copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC) click reaction^[31] with either tetramethylrhodamine- or diazo biotin-azide. Labeled proteins were then investigated by fluorescence SDS-PAGE or tandem mass spectrometry after streptavidin enrichment, trypsin digestion and six-plex tandem mass tags (TMTs) labeling.^[32] Disappearance of fluorescence signal (in SDS-PAGE) or abundance (in MS-MS) of covalently modified proteins by the alkyne probes in the presence of the corresponding competitors indicate a high extent of covalent protein labeling. Therefore, enabling differentiation between complete irreversible binding of low-abundance proteins and partial modification of proteins present in higher concentrations.^[33] Covalently targeted proteome members by each probe were identified by quantitative proteomics and rated on the basis of a high competition ratio (relative abundance of a corresponding protein in an experiment where only alkyne probe vs. alkyne probe plus competitor was applied) and low *p*-values. Based on that ranking, 12 different proteins, from which structural information was already available in the Protein Data Bank, were selected for further validation studies. Only in one case, did LC-ESI-MS demonstrate that the corresponding recombinant protein did not react upon incubation with the arylfluorosulfate compound, while the recombinant proteins of the remaining 11 selected targets proved to be covalently labeled by their respective probes under the experimental conditions.

Interestingly, a diverse set of proteins was enriched by different probes, marking the predicted importance of the peculiarly distinct binding motifs to achieve selective labeling. A few protein targets were, however, shown to react also with different probes. Glutathione S-transferase π (GSTP1) was found to be enriched in the volcano plots of both **14** and **15**, and tethering experiments corroborated adduct formation in both cases by LC-ESI-MS. Mutagenesis experiments of the

two potential labeled Tyr residues^[13c,34] one at a time (Y7F and Y108F) or both at the same time (Y7F,T108F) revealed that **15** reacts preferentially with Tyr7, but Tyr108 can be labeled as well, since removal of both nucleophiles inside the GSTP1 binding site is necessary to completely abrogate covalent labeling as shown by in-gel fluorescence. The importance of the neighboring basic Arg13 residue was corroborated by showing that an R13K mutant retained fluorosulfate reactivity, while an R13Q mutation abolished covalent adduct formation.

The multidimensional protein identification technology (MuDPIT) LC-MS/MS competition ratio experiment identified another glutathione S-transferase, GSTO1, as target of **14**. We corroborated quantitative covalent adduct formation over 48 h with recombinant enzyme and a co-crystal structure proved sulfate diester bond formation to Tyr229 (Figure 3b).

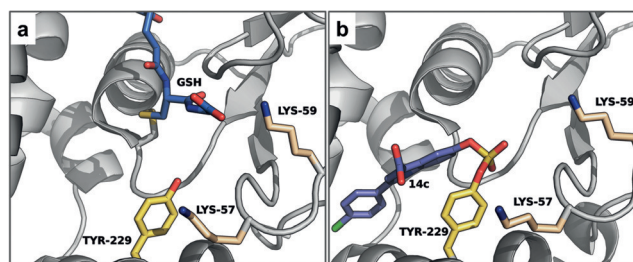


Figure 3. Binding mode of GSH and **14c** to GSTO1. a) Co-crystal structure of GSH and GSTO1 (PDB 2HVE). b) Co-crystal structure of **14c** and GSTO1 (PDB 5UI4). Key residues are shown as sticks.

Mutagenesis experiments highlighted the importance of both proximal Lys57 and Lys59 for labeling (Figure 3), since K57Q and K59Q mutations attenuated covalent adduct formation. However, covalent labeling of these mutants was rescued at pH 10.5, suggesting that the main role of the surrounding Lys residues is to lower the pK_a of the Tyr229 phenol in order to promote SuFEx reaction.

Efforts directed to validate nucleoside diphosphate kinase A (NME1) as a target of **15** led to several interesting findings. LC-ESI-MS corroborated complete formation of the NME1-**15c** covalent adduct, which the authors managed to crystallize. The resulting X-ray crystal structure revealed a sulfamate-mediated covalent bond to the Lys12 side chain, providing the first example of an X-ray crystal structure showing a fluorosulfate warhead labeling a lysine residue (Figure 4). This is particularly interesting, because a nearby Tyr52 is also available for covalent bond formation. However, only the side chain of Lys12, which is surrounded by basic Arg105 and His118 residues and predicted to be intrinsically nucleophilic,^[35] proved to partake in covalent adduct formation, since a MNE1-K12A mutant failed to react with **15**.

The remaining eight selected proteins (BLVRA, TIGAR, HSDL2, ALKBH5, TMPT, HMOX2, CRABP2, and TTR) were also validated by measuring covalent adduct formation by LC-ESI-MS and, in some cases, the labeling site was assigned by tandem mass spectrometry (MS-MS) analysis. Interestingly, **15** was found to react with thiopurine S-methyltransferase (TPMT) by binding to the solvent exposed

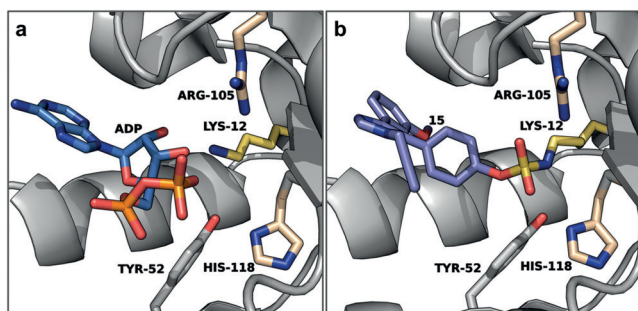


Figure 4. Binding mode of ADP and **15** to NME1. a) Co-crystal structure of ADP and NME1 (PDB 2HVE). b) Co-crystal structure of **15** and NME1 (PDB 5UEH). Key residues are shown as sticks.

Lys32 nearby the active site. However, covalent labeling of FTO, which was identified in the competition ratio experiment with probe **14**, was not observed and failed to be validated under various experimental conditions. As pointed out, this could be explained by differences between the recombinant and the endogenous protein in a cellular context, as post-translational modification of FTO or formation of a protein complex could be necessary for covalent modification to occur.^[23]

In vitro functional assays demonstrated a decreased GSTP1 and BLVRA enzymatic activity as a result of treatment with **14** and/or **15** probes. In addition, known inhibitors or endogenous ligands compete with the fluorosulfate-based probes for the protein binding sites of NME1, CRABP2, and TTR.^[36] Particularly interesting is the fact that covalent probe **14** is the first reported small molecule probe of hydroxysteroid dehydrogenase-like protein 2 (HSDL2).^[37] Several of these proteins possess a high interest for drug discovery and these encouraging results indicate that optimization of the compounds used here could provide potential covalent drug candidates for the treatment of a diverse range of diseases.

6. Summary and Outlook

The recent development of a facile method for the synthesis of arylfluorosulfates has spurred the use of this latent electrophile for covalent protein modification. The low and narrow intrinsic reactivity of fluorosulfate-based warheads assures minimal covalent labeling of off-target nucleophiles in a cellular context, making these compounds largely unreactive towards the human proteome unless rigorous conditions are met (e.g. the presence of basic residues that decrease the pK_a of the targeted nucleophilic residue and/or facilitate the departure of the fluoride ion).

Similar to the reports discussed herein, several recent examples provide additional encouraging data to start considering the use of arylfluorosulfate-targeted covalent inhibitors in drug discovery. Thus, proteins containing genetically encoded fluorosulfate-L-tyrosine (FSY) have been shown to react with strategically located proximal Tyr, Lys, and His residues on the surfaces of their interacting protein partners in live *E. coli*, and exhibited no toxicity towards *E.*

coli or mammalian cells at concentrations up to 1 mM.^[38] These results approved the use of genetically encoded FSY for covalent attachment of protein partners in vivo and suggest the value of arylfluorosulfate-based compounds for the selective covalent labeling of Tyr, Lys, or His residues on protein surfaces for targeting specific protein-protein interactions. In addition, reactive covalent docking experiments have already proved useful to predict possible covalent binding sites of arylfluorosulfate probes.^[23] This suggests that rational targeting of intrinsically reactive Tyr, Lys or even Ser and His residues could be applied in drug discovery for a given protein of interest that displays a fluorosulfate-desired protein microenvironment, especially with a high-affinity, reversible ligand in hand that can be modified. Furthermore, a platform for late-stage compound functionalization has also been developed, that allows for the in situ, selective arylfluorosulfate formation from phenol-containing compounds in a 96-well plate format, allowing the generation of arylfluorosulfate compound libraries that can be directly applied in cell-based or enzymatic assays.^[39]

To date, arylfluorosulfate-based probes have been found to modify protein pockets preferentially through tyrosine residues (in iLBPs, GSTP1, GSTO1, BLVRA, HMOX2, HSDL2, ALKBH5, and TIGAR proteins) and lysine residues (in TTR, NME1, TMPT, and HMOX2) but also serine in DcpS. We predict that the number of examples of arylfluorosulfate-derived covalent ligands will substantially increase over the coming years, which should provide further insights into the structure and substituent effects on reactivity, chemical and metabolic stability, as well as covalent target engagement selectivity of these compounds. We anticipate to witness the application of other “SuFEx-able” warheads (e.g., alkyl or aryliminosulfur oxy(di or mono)fluorides)^[40] in chemical biology research, which could lead to the development of targeted covalent inhibitors with even lower proteome-wide reactivity and different amino acid labeling preference.

These are encouraging times for the field of covalent protein modification, and yet the powerful addition of arylfluorosulfates to the catalogue of available warheads promises even more exciting developments in the fields of covalent probes and inhibitors in chemical biology and drug discovery.

Acknowledgements

We gratefully acknowledge Dr. Andreas S. Madsen for assistance with the Frontispiece graphic. This work was supported by the European Research Council (ERC-CoG-725172—SIRFUNCT; C.A.O.).

Conflict of interest

The authors declare no conflict of interest.

How to cite: *Angew. Chem. Int. Ed.* **2019**, *58*, 957–966
Angew. Chem. **2019**, *131*, 969–978

- [1] J. Singh, R. C. Petter, T. A. Baillie, A. Whitty, *Nat. Rev. Drug Discovery* **2011**, *10*, 307–317.
- [2] a) D. S. Johnson, E. Weerapana, B. F. Cravatt, *Future Med. Chem.* **2010**, *2*, 949–964; b) Z. Zhao, P. E. Bourne, *Drug Discovery Today* **2018**, *23*, 727–735.
- [3] T. A. Baillie, *Angew. Chem. Int. Ed.* **2016**, *55*, 13408–13421; *Angew. Chem.* **2016**, *128*, 13606–13619.
- [4] K. Chen, L. Kurgan, *PLoS ONE* **2009**, *4*, e4473.
- [5] A. G. Watts, I. Damager, M. L. Amaya, A. Buschiazio, P. Alzari, A. C. Frasch, S. G. Withers, *J. Am. Chem. Soc.* **2003**, *125*, 7532–7533.
- [6] a) J. Pettinger, K. Jones, C. D. Matthew, *Angew. Chem. Int. Ed.* **2017**, *56*, 15200–15209; *Angew. Chem.* **2017**, *129*, 15398–15408; b) S. Choi, D. S. T. Ong, J. W. Kelly, *J. Am. Chem. Soc.* **2010**, *132*, 16043–16051; c) A. R. Nanna, X. Li, E. Walseng, L. Pedzisa, R. S. Goydel, D. Hymel, T. R. Burke, Jr., W. R. Roush, C. Rader, *Nat. Commun.* **2017**, *8*, 1112; d) M. J. Matos, B. L. Oliveira, N. Martínez-Sáez, A. Guerreiro, P. M. S. D. Cal, J. Bertoldo, M. Maneiro, E. Perkins, J. Howard, M. J. Deery, J. M. Chalker, F. Corzana, G. Jiménez-Osés, G. J. L. Bernardes, *J. Am. Chem. Soc.* **2018**, *140*, 4004–4017.
- [7] a) F. J. Dekker, O. Rocks, N. Vartak, S. Menninger, C. Hedberg, R. Balamurugan, S. Wetzel, S. Renner, M. Gerauer, B. Schölermann, M. Rusch, J. W. Kramer, D. Rauh, G. W. Coates, L. Brunsveld, P. I. H. Bastiaens, H. Waldmann, *Nat. Chem. Biol.* **2010**, *6*, 449; b) K. Otrubova, M. Brown, M. S. McCormick, G. W. Han, S. T. O’Neal, B. F. Cravatt, R. C. Stevens, A. H. Lichtman, D. L. Boger, *J. Am. Chem. Soc.* **2013**, *135*, 6289–6299.
- [8] M. Bogyo, J. S. McMaster, M. Gaczynska, D. Tortorella, A. L. Goldberg, H. Ploegh, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 6629–6634.
- [9] P. Martín-Gago, E. K. Fansa, M. Winzker, S. Murarka, P. Janning, C. Schultz-Fademrecht, M. Baumann, A. Wittinghofer, H. Waldmann, *Cell Chem. Biol.* **2017**, *24*, 589–597.
- [10] G. Platzter, M. Okon, L. P. McIntosh, *J. Biomol. NMR* **2014**, *60*, 109–129.
- [11] a) T. K. Harris, G. J. Turner, *IUBMB Life* **2002**, *53*, 85–98; b) D. G. Isom, C. A. Castañeda, B. R. Cannon, B. E. García-Moreno, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 5260–5265; c) A. A. Yunus, C. D. Lima, *Nat. Struct. Mol. Biol.* **2006**, *13*, 491; d) I. Hiroshi, *FEBS Lett.* **2010**, *584*, 3464–3468.
- [12] A. Narayanan, L. H. Jones, *Chem. Sci.* **2015**, *6*, 2650–2659.
- [13] a) K. E. Dombrowski, Y. C. Huang, R. F. Colman, *Biochemistry* **1992**, *31*, 3785–3793; b) N. N. Gushwa, S. Kang, J. Chen, J. Taunton, *J. Am. Chem. Soc.* **2012**, *134*, 20214–20217; c) C. Gu, D. A. Shannon, T. Colby, Z. Wang, M. Shabab, S. Kumari, J. G. Villamor, C. J. McLaughlin, E. Weerapana, M. Kaiser, B. F. Cravatt, R. A. L. van der Hoorn, *Chem. Biol.* **2013**, *20*, 541–548; d) D. A. Shannon, E. Weerapana, *Curr. Opin. Chem. Biol.* **2015**, *24*, 18–26; e) O. Fadeyi, M. D. Parikh, M. Z. Chen, R. E. Kyne, Jr., A. P. Taylor, I. O’Doherty, S. E. Kaiser, S. Barbas, S. Niessen, M. Shi, S. L. Weinrich, J. C. Kath, L. H. Jones, R. P. Robinson, *ChemBioChem* **2016**, *17*, 1925–1930; f) C. Dubiella, H. Cui, M. Groll, *Angew. Chem. Int. Ed.* **2016**, *55*, 13330–13334; *Angew. Chem.* **2016**, *128*, 13524–13528; g) Q. Zhao, X. Ouyang, X. Wan, K. S. Gajiwala, J. C. Kath, L. H. Jones, A. L. Burlingame, J. Taunton, *J. Am. Chem. Soc.* **2017**, *139*, 680–685; h) C. S. Lentz, J. R. Sheldon, L. A. Crawford, R. Cooper, M. Garland, M. R. Amieva, E. Weerapana, E. P. Skaar, M. Bogyo, *Nat. Chem. Biol.* **2018**, *14*, 609–617; i) S. Guardiola, R. Prades, L. Mendieta, A. J. Brouwer, J. Streefkerk, L. Nevola, T. Tarragó, R. M. J. Liskamp, E. Giralt, *Cell Chem. Biol.* **2018**, <https://doi.org/10.1016/j.chembiol.2018.04.013>.
- [14] a) J. Dong, L. Krasnova, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2014**, *53*, 9430–9448; *Angew. Chem.* **2014**, *126*, 9584–9603; b) L. H. Jones, *ACS Med. Chem. Lett.* **2018**, *9*, 584–586.
- [15] a) B. R. Baker, W. F. Wood, *J. Med. Chem.* **1968**, *11*, 650–652; b) B. R. Baker, G. J. Lourens, *J. Med. Chem.* **1967**, *10*, 1113–1122; c) B. R. Baker, J. A. Hurlbut, *J. Med. Chem.* **1968**, *11*, 241–245; d) J. C. Powers, T. Tanaka, J. W. Harper, Y. Minematsu, L. Barker, D. Lincoln, K. V. Crumley, J. E. Fraki, N. M. Schechter, G. G. Lazarus, et al., *Biochemistry* **1985**, *24*, 2048–2058; e) D. G. Deutsch, S. Lin, W. A. G. Hill, K. L. Morse, D. Salehani, G. Arreaza, R. L. Omeir, A. Makriyannis, *Biochem. Biophys. Res. Commun.* **1997**, *231*, 217–221; f) B. R. Baker, *Ann. N. Y. Acad. Sci.* **1971**, *186*, 214–226.
- [16] W. Lange, E. Müller, *Ber. Dtsch. Chem. Ges.* **1930**, *63*, 2653–2657.
- [17] A. Baranczak, Y. Liu, S. Connelly, W.-G. H. Du, E. R. Greiner, J. C. Genereux, R. L. Wiseman, Y. S. Eisele, N. C. Bradbury, J. Dong, L. Noodleman, K. B. Sharpless, I. A. Wilson, S. E. Encalada, J. W. Kelly, *J. Am. Chem. Soc.* **2015**, *137*, 7404–7414.
- [18] a) H. Zhou, P. Mukherjee, R. Liu, E. Evrard, D. Wang, J. M. Humphrey, T. W. Butler, L. R. Hoth, J. B. Sperry, S. K. Sakata, C. J. Helal, C. W. am Ende, *Org. Lett.* **2018**, *20*, 812–815; b) T. Guo, G. Meng, X. Zhan, Q. Yang, T. Ma, L. Xu, K. B. Sharpless, J. Dong, *Angew. Chem. Int. Ed.* **2018**, *57*, 2605–2610; *Angew. Chem.* **2018**, *130*, 2635–2640.
- [19] C. Verysse, J. Demaerel, V. Bieliūnas, P. Gilles, W. M. De Borggraeve, *Org. Lett.* **2017**, *19*, 5244–5247.
- [20] H. Mukherjee, J. Debreczeni, J. Breed, S. Tentarelli, B. Aquila, J. E. Dowling, A. Whitty, N. P. Grimster, *Org. Biomol. Chem.* **2017**, *15*, 9685–9695.
- [21] W. Chen, J. Dong, L. Plate, D. E. Mortenson, G. J. Brighty, S. Li, Y. Liu, A. Galmozzi, P. S. Lee, J. J. Hulce, B. F. Cravatt, E. Saez, E. T. Powers, I. A. Wilson, K. B. Sharpless, J. W. Kelly, *J. Am. Chem. Soc.* **2016**, *138*, 7353–7364.
- [22] N. P. Grimster, S. Connelly, A. Baranczak, J. Dong, L. B. Krasnova, K. B. Sharpless, E. T. Powers, I. A. Wilson, J. W. Kelly, *J. Am. Chem. Soc.* **2013**, *135*, 5656–5668.
- [23] D. E. Mortenson, G. J. Brighty, L. Plate, G. Bare, W. Chen, S. Li, H. Wang, B. F. Cravatt, S. Forli, E. T. Powers, K. B. Sharpless, I. A. Wilson, J. W. Kelly, *J. Am. Chem. Soc.* **2018**, *140*, 200–210.
- [24] E. C. Hett, H. Xu, K. F. Geoghegan, A. Gopalsamy, R. E. Kyne, C. A. Menard, A. Narayanan, M. D. Parikh, S. Liu, L. Roberts, R. P. Robinson, M. A. Tones, L. H. Jones, *ACS Chem. Biol.* **2015**, *10*, 1094–1098.
- [25] O. O. Fadeyi, L. R. Hoth, C. Choi, X. Feng, A. Gopalsamy, E. C. Hett, R. E. Kyne, R. P. Robinson, L. H. Jones, *ACS Chem. Biol.* **2017**, *12*, 2015–2020.
- [26] S.-E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, M. Mann, *Mol. Cell. Proteomics* **2002**, *1*, 376–386.
- [27] R. L. Smathers, D. R. Petersen, *Hum. Genomics* **2011**, *5*, 170.
- [28] J. R. Yates, J. K. Eng, A. L. McCormack, D. Schieltz, *Anal. Chem.* **1995**, *67*, 1426–1436.
- [29] S. Vaezslami, E. Mathes, C. Vasileiou, B. Borhan, J. H. Geiger, *J. Mol. Biol.* **2006**, *363*, 687–701.
- [30] M. Maden, *Nat. Rev. Neurosci.* **2007**, *8*, 755.
- [31] a) C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057–3064; b) V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2002**, *41*, 2596–2599; *Angew. Chem.* **2002**, *114*, 2708–2711; c) A. E. Speers, G. C. Adam, B. F. Cravatt, *J. Am. Chem. Soc.* **2003**, *125*, 4686–4687.
- [32] L. Dayon, A. Hainard, V. Licker, N. Turck, K. Kuhn, D. F. Hochstrasser, P. R. Burkhard, J.-C. Sanchez, *Anal. Chem.* **2008**, *80*, 2921–2931.
- [33] C. Wang, E. Weerapana, M. M. Blewett, B. F. Cravatt, *Nat. Methods* **2013**, *11*, 79.
- [34] A. J. Oakley, M. L. Bello, A. Battistoni, G. Ricci, J. Rossjohn, H. O. Villar, M. W. Parker, *J. Mol. Biol.* **1997**, *274*, 84–100.

- [35] E. H. Postel, B. M. Abramczyk, M. N. Levit, S. Kyin, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 14194.
- [36] a) N. A. Malmquist, J. J. Anzinger, D. Hirzel, I. L. Buxton, *Proc. Western Pharmacol. Soc.* **2001**, *44*, 57–59; b) A. S. Budhu, N. Noy, *Mol. Cell. Biol.* **2002**, *22*, 2632–2641; c) C. E. Bulawa, S. Connelly, M. DeVit, L. Wang, C. Weigel, J. A. Fleming, J. Packman, E. T. Powers, R. L. Wiseman, T. R. Foss, I. A. Wilson, J. W. Kelly, R. Labaudinière, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 9629–9634.
- [37] C. Ruokun, X. Yake, Y. Fengdong, W. Xinting, S. Laijun, L. Xianzhi, *Tumor Biol.* **2016**, *37*, 15065–15077.
- [38] N. Wang, B. Yang, C. Fu, H. Zhu, F. Zheng, T. Kobayashi, J. Liu, S. Li, C. Ma, P. G. Wang, Q. Wang, L. Wang, *J. Am. Chem. Soc.* **2018**, *140*, 4995–4999.
- [39] Z. Liu, J. Li, S. Li, G. Li, K. B. Sharpless, P. Wu, *J. Am. Chem. Soc.* **2018**, *140*, 2919–2925.
- [40] S. Li, P. Wu, E. J. Moses, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2017**, *56*, 2903–2908; *Angew. Chem.* **2017**, *129*, 2949–2954.

Manuscript received: May 25, 2018

Revised manuscript received: July 18, 2018

Accepted manuscript online: July 19, 2018

Version of record online: October 26, 2018