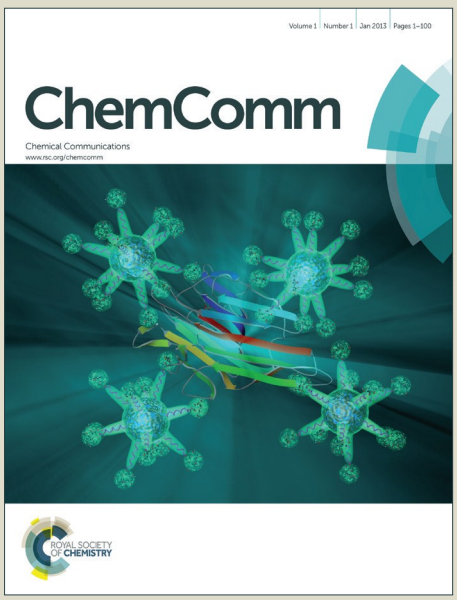
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Journal Name

COMMUNICATION

Rational Design of Reversible and Irreversible Cysteine Sulfenic Acid-Targeted Linear C-Nucleophiles

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

DOI: 10.1039/x0xx00000x

www.rsc.org/

Vinayak Gupta[a] and Kate S. Carroll*[, a]

Concerns about off-target effects has motivated the development of reversible covalent inhibition strategies for targeting cysteine. However, such strategies have not been reported for the unique cysteine oxoform, sulfenic acid. Herein, we have designed and identified linear C-nucleophiles that react selectively with cysteine sulfenic acid. The resulting thioether adducts exhibit reversibility ranging from minutes to days under reducing conditions, showing the feasibility of tuning C-nucleophile reactivity across a wide range of time scales.

Oxidation of a protein cysteine thiol (Cys-SH) to sulfenic acid (Cys-SOH) by endogenous reactive oxygen species (ROS), including H_2O_2, is a reversible post-translational modification that plays a central role in regulating the activity, location, and function of a wide variety of proteins.[1] For example, Paulsen *et al.* reported that a critical active site cysteine (Cys797) in the epidermal growth factor receptor (EGFR) is *S*-sulfenylated by H_2O_2 induced by growth factor stimulation, which enhances its intrinsic tyrosine kinase activity.[2] Recently, we have shown that SIRT6 exerts redox control of HIF1A transcriptional activity *via* sulfenyl-mediated complex formation.[3, 4] Aberrant Cys-SOH formation during oxidative stress or chronic inflammation has also been linked to human diseases, including cancer.[5, 6]

The FDA approval of kinase inhibitors such as Afatinib and Ibrutinib targeting cysteine residues led to the re-emergence of covalent inhibition as a drug design strategy.[7] Concerns about off-target modification has motivated the exploration of reversible covalent chemistry. Several such strategies have been reported, including bromomaleimide-based reagents that are labile under reducing conditions (**Fig. S1A**) and electron-deficient olefins, like acrylamides that rapidly and reversibly react with cysteine (**Fig. S1B**).[8, 9] However, the aforementioned reversible covalent inhibitors are not capable of reacting with Cys-SOH under physiological conditions and cysteine oxidation is expected to impact their pharmacology. For this reason, we have proposed a complementary strategy in which the sulfenyl form of therapeutically important proteins could be exploited to develop a new class of irreversible and reversible covalent inhibitors targeting this unique cysteine oxoform.[1, 2, 10-12]

Due to the electrophilic as well as nucleophilic character of the sulphur atom in Cys-SOH, detection methods exploiting both chemical features have been reported.[1, 12] However, under physiological conditions, it is the electrophilic reactivity of Cys-SOH that dominates and the vast majority of probes for selective chemical detection are C-nucleophiles based on cyclic 1,3-carbonyl scaffolds, such as dimedone **(R = R' = Me, n = 1)** (**Fig. S1C**). Such chemical probes have been widely employed for qualitative and quantitative study of the sulfenome. [1, 3, 13] Although dimedone and closely related probes are selective under aqueous conditions, they are collectively hampered by slow reaction kinetics when compared to competing biological fates of Cys-SOH (*e.g.,* thiol-based reduction to form a disulfide or over-oxidation to sulfinic acid).[1]

To overcome the issue of slow reactivity, we recently developed a mass spectrometry-based assay and dipeptide sulfenic acid model **(1)** that enabled us to screen a library of *cyclic* C-nucleophiles and identify novel compounds with >200-fold enhanced reactivity, compared to dimedone.[12] In general, cyclic C-nucleophiles form stable covalent adducts with Cys-SOH, a feature that renders such compounds useful for more comprehensive analyses of the sulfenome and as warheads for the development of *irreversible* covalent inhibitors that target Cys-SOH.[2, 11] On the other hand, the irreversible nature of this reaction may increase the chances of forming permanent adducts with undesired proteins and precludes applications in which target release is favoured. *Reversible* covalent probes of Cys-SOH are yet unknown and represents a significant hurdle in the development of nucleophilic inhibitors.

Herein, we have employed our LC/MS assay to screen six different classes **(3-8)** of *linear C-nucleophiles* **(Fig. 1)**. By

[a] *Department of Chemistry, The Scripps Research Institute, Jupiter, Florida, 33458, USA. Email: kcarroll@scripps.edu*

† Footnotes relating to the title and/or authors should appear here.

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

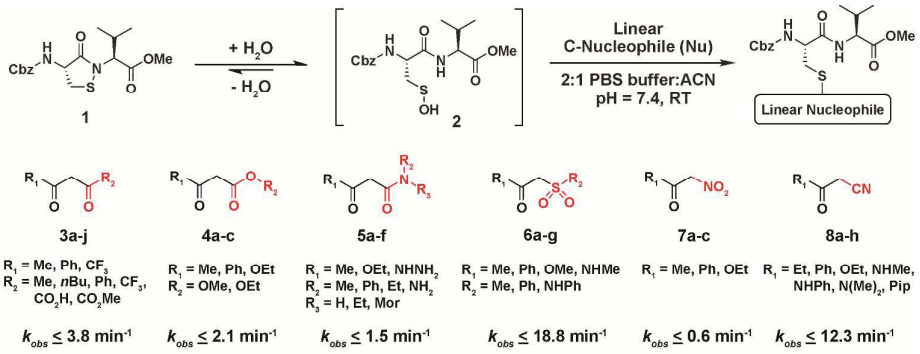


Fig. 1 Structure and reaction of various classes of linear C-nucleophiles with sulfenic acid **2**.

contrast to cyclic 1,3-dicarbonyls, the keto-enol equilibrium is predominantly shifted towards keto form in linear 1,3-dicarbonyls **(3)**.[14] We reasoned that this shift towards keto tautomer should manifest as a difference in reactivity towards Cys-SOH and in the stability of the corresponding thioether adduct. Moreover, a systematic study of linear C-nucleophiles would also increase the diversity of chemical tools for use as probes and inhibitors of sulfenic acid-modified proteins.

Initial support for our hypothesis was obtained from the reactivity of 2,4-pentanedione (**3a**) towards sulfenic acid **2**. Notably, **3a** exhibited a 4:1 ratio of keto:enol tautomers and reacted with k_{obs} = 3.8 min^{-1}, almost 200-fold higher than its 5-membered cyclic analog (*i.e.,* cyclopentane-1,3-dione, k_{obs} = 0.02 min^{-1})[12]. Next, we explored the effects of R_1 and R_2 substitution of **3** with different alkyl, aryl and electron-withdrawing groups (EWGs), such as $-CO_2H$, $-CO_2Me$, $-CF_3$. In general, while mono alkylated/arylated substrates **3b-c** gave k_{obs} similar to **3a**, diarylated analog **3d** or analogs with EWGs in place of methyl, like **3e-j**, reduced k_{obs} **(Fig. S2)**. This decrease in activity can be attributed to the shift in keto-enol equilibrium towards the enol tautomer and enhanced carbanion stability at C-3. Moreover, substitution by highly EWGs, such as $-CF_3$, may lead to hydration of the carbonyl and complete loss of reactivity due an increase in C-3 carbanion p*K*a).[15] Subsequently, for nucleophiles of structural class **4** and **5** we postulated that replacing one or both carbonyls with a carboxylate ester, amide or hydrazide functionality would increase the pK_a of the nucleophilic C-centre. Consistent with this hypothesis, we observed a marked decrease in reactive carbanion formation at physiological pH and a corresponding reduction in reactivity (**Fig. S2**).

An apparent trend in C-nucleophile reactivity towards Cys-SOH is that carbanion destabilization increases its reactivity, at least within a specific p*K*a range (*i.e.,* 4-10).[12] Along these lines, we theorized that replacement of one carbonyl group with a sulfone with concomitant partial loss of resonance would also increase reactivity even further. For this reason, we evaluated linear C-nucleophiles in which one carbonyl was replaced with a sulfone functional group **(6)**. Gratifyingly, a k_{obs} of 18.8 min^{-1} or 25-fold increase relative to dimedone (k_{obs} = 0.8 min^{-1}) was observed for the reaction between **6d** and sulfenic acid **2** (**Fig. S2**). Next, we evaluated the reactivity of molecules in which one or both carbonyls were replaced by an EWG such as nitro ($-NO_2$, **7**) and cyano (-CN, **8**). Given that the pK_as of **7a-c** are predicted to be similar to 1,3-cyclohexanedione (~5), it was not too surprising that the nitro-functionalized nucleophiles exhibited similar k_{obs} with sulfenic acid **2**. Fortunately, with predicted pK_a of ~10 and corresponding $k_{obs} \leq 20$ min^{-1}, cyano-substituted nucleophiles **8a-h** gave more promising results (**Fig. S2**). The above trends highlight the fine balance between keto-enol tautomerism and carbanion stability (*i.e.,* more/less resonance forms) for linear C-nucleophile reactivity towards Cys-SOH.

In contrast to our study of cyclic C-nucleophiles where nucleophiles exhibiting > 200-fold enhancement in reactivity relative to dimedone were identified, the best C-nucleophile identified in initial screening of acyclic compounds exhibited a modest 25-fold rate enhancement compared to dimedone. Therefore, we decided to take a rational design approach in which structural elements of the most reactive linear C-nucleophiles were combined to produce *"designer nucleophiles"* with superior k_{obs} with Cys-SOH. To this end, we combined the phenylsulfonyl group of **6b** (k_{obs} = 8.7 min^{-1}) and nitro group of **7b** (k_{obs} = 0.5 min^{-1}) to create phenylsulfonylnitromethane **(9)**. The reaction of **9** with sulfenic acid **2** gave a k_{obs} of 7.8 min^{-1}, which is similar to that of **6b**. Although this initial compound did not afford better reaction kinetics, as discussed below, a nucleophile modelled on **9** did prove highly useful in the detection of protein Cys-SOH in cell lysates. In subsequent experiments, we combined the phenylsulfonyl moiety from **6b** with nitrile moiety from malononitrile **8h** (k_{obs} = 20.0 min^{-1}) to produce methyl- and phenylsulfonylacetonitrile **(10a-b)**. Both **10a** and **10b** formed thioether adducts with sulfenic acid **2** with k_{obs} equal to 28.6 min^{-1} and 78.0 min^{-1}, respectively (**Fig. 2A**). The findings were very encouraging and indicate the potential to design linear C-nucleophiles with favourable reactivity profiles.

While scanning the literature for linear C-nucleophiles, we came across molecules such as the natural product, curcumin **(11)** and anti-rheumatoid arthritis drug, tofacitinib **(12)** (**Fig. 2B**). Curcumin is a diarylheptanoid with numerous biomedical properties.[16] Of note, **11** possess an electrophilic as well as a

nucleophilic reaction site. Although covalent bond formation between Cys-SH and the α,β-unsaturated ketone group of curcumin is logical and recently reported,[17] we reasoned that the 1,3-dicarbonyl centre should also be able to react with Cys-SOH. Indeed, **11** successfully reacted with **2** exhibiting a k_{obs} of 1.9 min^{-1}. Moreover, as cyanoacetamido- based nucleophiles **8f-g** reacted with sulfenic acid **2**, we wondered whether **12** (containing a similar nucleophilic warhead) might also form a covalent bond with Cys-SOH. Consistent with our hypothesis, **12** effectively formed the expected thioether adduct with **2** (k_{obs} = 0.4 min^{-1}) (**Fig. 2B**).

In subsequent experiments, we examined the chemical selectivity of selected nucleophiles (**3a, 4a, 6b, 8h, 9, 10b, 11, 12**) in control experiments by profiling their reactivity with biofunctional groups such as lysine (amine), serine (alcohol), cysteine (thiol) and sulfinic acid (**Scheme S3**). Importantly, all acylic C-nucleophiles, with the exception of malononitrile **8h**, gave no indication of cross-reactivity with other relevant functional groups. Next, we examined the reactivity (and selectivity) of linear C-nucleophiles (**3a, 4a, 6b, 8h, 9, 10b, 11, 12**) towards a protein Cys-SOH. For these studies, we used our well-established model, a Cys64Ser Cys82Ser variant of the thiol peroxidase, Gpx3.[12] Incubation of Gpx3-SOH (prepared *in situ* by treating 10 μM Gpx3-SH with 1.5 eq of H_2O_2) with the aforementioned linear C-nucleophiles, including curcumin **11** and tofacitinib **12**, led to covalent modification of Gpx3-SOH. Control experiments clearly indicated that linear C-nucleophiles did not react with reduced Gpx3 (**Scheme S4**).

Given the differences among linear C-nucleophiles with respect to reactivity, pK_a, and carbanion stability it is possible that the resulting thioether adducts also display a range of stability under reducing conditions. As cell contains millimolar concentrations of glutathione (GSH) we tested the ability of this low-molecular weight thiol as well as dithiothreitol (DTT) and the phosphine-based reagent, TCEP to reduce the nascent covalent bond formed between **2** and representative members

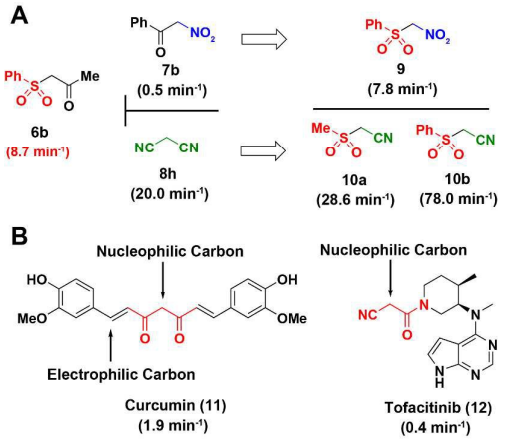
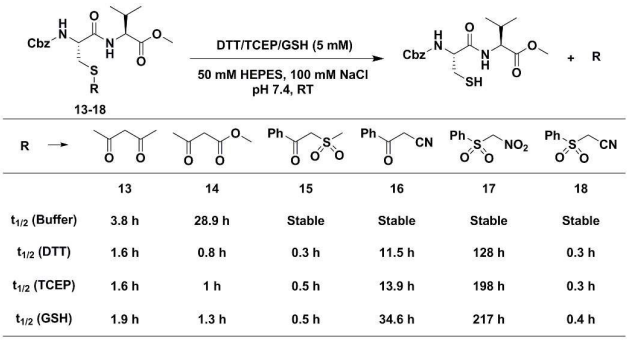


Fig. 2 (A) Strategy for rational design of linear C-nucleophiles exhibiting superior reactivity with Cys-SOH. (B) Structure and putative reaction sties of curcumin **(11)** and tofacitinib **(12)** towards sulfenic acid **2**.



R →	13	14	15	16	17	18
$t_{1/2}$ (Buffer)	3.8 h	28.9 h	Stable	Stable	Stable	Stable
$t_{1/2}$ (DTT)	1.6 h	0.8 h	0.3 h	11.5 h	128 h	0.3 h
$t_{1/2}$ (TCEP)	1.6 h	1 h	0.5 h	13.9 h	198 h	0.3 h
$t_{1/2}$ (GSH)	1.9 h	1.3 h	0.5 h	34.6 h	217 h	0.4 h

Fig. 3 Stability of thioethers **13 - 18** under reducing conditions.

of our linear C-nucleophile library. Accordingly, thioethers **13** - **18** were prepared (**Scheme S5**) and examined for their stability under reducing conditions (**Fig. 3** and **Scheme S6**). Thioether adducts **13** and **14** were unstable, ostensibly owing to carbonyl hydration, elimination of corresponding acetic acid unit, and net loss of a methylketo (–COMe) group (**Scheme S7**).[15] More interesting, however, was the reversibility of many thioether bonds under reducing conditions to give the corresponding dipeptide-SH and linear C-nucleophile. To better characterize such reactions, pseudo 1st order rate constants (k_{obs}) were determined for the reduction of all thioether adducts (**Table S1**). Interestingly, thioether adducts generated from our 'designer nucleophiles' exhibited distinct stability profiles. For example, thioether **15** displayed excellent stability under all reducing conditions, with k_{obs} ranging from 0.003 h^{-1} to 0.005 h^{-1} ($t_{1/2}$ = 6 – 10 days). By contrast, thioether **18** was the most unstable adduct with k_{obs} ranging from 1.9 h^{-1} to 2.5 h^{-1} ($t_{1/2}$ = 17 – 22 min), representing a 500-fold difference in stability **(Fig. 3)**. All remaining thioether adducts gave stability profiles in between these two extremes (**Table S1**). Subsequently, we studied the stability of thioether adducts in the Gpx3-SOH model. Analogous to the above analysis with dipeptide sulfenic acid **2**, the Gpx3-S-Nu adduct resulting from the reaction of Gpx3-SOH and **9** was the most stable and that of **10b** was the least stable (**Scheme S8**).

Having characterized the proficiency and reversibility of linear C-nucleophiles with Cys-SOH in dipeptide- and protein-SOH model systems, we next investigated labelling of protein sulfenic acids in whole cell lysates. To facilitate detection, **9** and **10b** were each modified to incorporate an alkyne handle (for 'clicking' with biotin azide and visualization by SDS-PAGE) without adversely affecting their reactivity profiles (**Fig. 4A**). For comparison, we chose cyclic C-nucleophile, DYn-2 **(19)**, an alkyne-functionalized probe known for its ability to form stable thioether adducts with Cys-SOH.[2-4] Initial studies confirmed the reactivity and reversibility of NO_2-alkyne **(20)** and CN-alkyne **(21)** in the less complex Gpx3-SOH model. While all three nucleophiles formed a thioether adduct with Gpx3-SOH, the product of DYn-2 and **20** reactions were stable after GSH treatment, whereas the covalent product of **21** was completely reversible (**Scheme S9**). Subsequently, we probed HeLa cell lysate with DYn-2, NO_2-alkyne **20** and CN-alkyne **21** for protein S-sulfenylation (**Fig. 4B**). Compared to DYn-2, both NO_2-alkyne

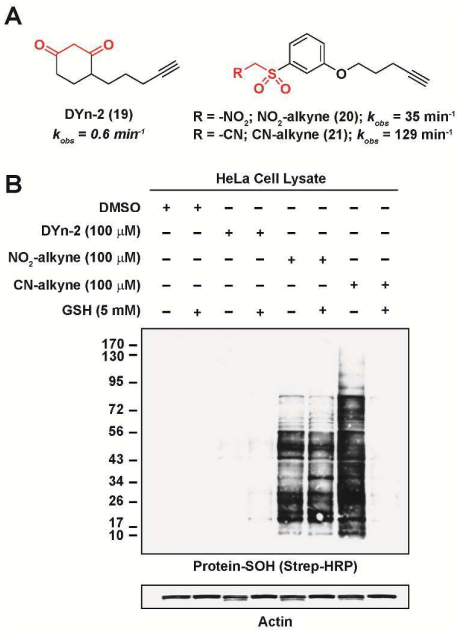


Fig. 4 Stability of protein-S-Nu thioether bonds under reducing conditions. **(a)** Alkyne-functionlized irreversible and reversible C-nucleophiles; **(b)** Gel-based reactivity profiling of irreversible and reversible C-nucleophiles in HeLa cell lysate.

and CN-alkyne showed robust protein-SOH labelling at 100 μM probe concentration, an observation that is fully consistent with reaction rate constants obtained from sulfenic acid **2**. Even more importantly, while CN-alkyne was more efficient than NO_2-alkyne at capturing the potein-SOH, labelling by **21** was totally reversible. On the other hand, protein-SOH labelling by **20** was robust and stable under reducing conditions.

We have shown that linear C-nucleophiles can be used for efficient and selective modification of Cys-SOH in peptides, proteins and lysates (k_{obs} = 0.05 - 129 min^{-1}). Moreover, we have studied the reversibility of the resulting thioether adducts and shown that linear C-nucleophiles, such as **21**, form reversible covalent bonds. To the best of our knowledge, this work is the first to report the reversibility of such thioether adducts, a key feature that can be exploited in the development of small-molecule probes and inhibitors that target Cys-SOH within proteins. In another important aspect of this study, we have established parameters that must berigorously tested in the development of chemical probes for Cys-SOH. For example, we reveal the unstable nature of thioether adducts formed by β-ketoester-based probes [18], with direct implications for the use of these and related probes for Cys-SOH detection and proteomic analyses. Based on the reactivity and reversibility trends observed in our screen, we successfully designed two novel alkyne-functionalized linear C-nucleophile probes that react rapidly and selectively with Cys-SOH. Notably, we show that NO_2-alkyne **20** forms a stable adduct, whereas the bond formed by reaction with CN-alkyne **21** is reversible in reducing conditions.

Hitherto unknown, in the present study we also show the reactivity of the natural product, curcumin with Cys-SOH. Our observations, along with the knowledge that curcumin can also function as a Michael-acceptor with Cys-SH, highlight its dual reactivity and might account for the many intriguing biological activities reported for this natural product. Furthermore, we show that Tofacitinib, the active ingredient in Pfizer's oral anti-inflammatory drug 'Xeljanz' prescribed for rheumatoid arthritis,[19] reacts with Cys-SOH. Tofacitinib was initially reported as a selective inhibitor of JAK3 (Janus Kinase 3) in comparison to other kinases of this family, JAK1 and JAK2.[20] Sequence alignment of JAK family kinases demonstrates a high degree of homology in the active-site/inhibitor-binding pocket, suggesting that Tofacitinib should inhibit JAK1-3 with similar potency. Interestingly, a distinguishing feature of JAK3 is the presence of Cys_{909} in the 'hinge' region, whereas JAK1/2 have a serine residue at that position.[21] Since Cys_{909} of JAK3 is located at the structural position corresponding to EGFR Cys797 [22] there is a high probability that this cysteine is redox sensitive and may become oxidized to sulfenic acid in cells. Such collective observations raise the possibility that Tofacitinib selectivity towards JAK3 relative to JAK1/2 may be due, at least in part, to covalent reversible interaction with Cys_{909}-SOH. More studies are required to evaluate this possibility and will be reported in due course.

Notes and references

1. V. Gupta and K. S. Carroll, *Biochim. Biophys. Acta*, 2014, **1840**, 847-875.
2. C. E. Paulsen, et al., *Nat Chem Biol*, 2012, **8**, 57-64.
3. J. Yang, et al., *Nat Commun*, 2014, **5**, 4776.
4. J. Yang, et al., *Nat Protoc*, 2015, **10**, 1022-1037.
5. Y. H. Seo and K. S. Carroll, *Proc Natl Acad Sci U S A*, 2009, **106**, 16163-16168.
6. E. Zito, et al., *Mol. Cell*, 2012, **48**, 39-51.
7. J. Singh, et al., *Nat Rev Drug Discov*, 2011, **10**, 307-317.
8. M. E. B. Smith, et al., *J. Am. Chem. Soc.*, 2010, **132**, 1960-1965.
9. J. M. Bradshaw, et al., *Nat Chem Biol*, 2015, **11**, 525-531.
10. T. H. Truong and K. S. Carroll, *Biochemistry*, 2012, **51**, 9954-9965.
11. T. H. Truong and K. S. Carroll, *Crit. Rev. Biochem. Mol. Biol.*, 2013, **48**, 332-356.
12. V. Gupta and K. S. Carroll, *Chem. Sci.*, 2016, **7**, 400-415.
13. P. Martinez-Acedo, et al., *J. Mass Spectrom.*, 2014, **49**, 257-265.
14. S. G. Mills and P. Beak, *J. Org. Chem.*, 1985, **50**, 1216-1224.
15. K. V. Ruddraraju, et al., *J. Org. Chem.*, 2015, **80**, 12015-12026.
16. V. P. Menon and A. R. Sudheer, *Adv Exp Med Biol*, 2007, **595**, 105-125.
17. D. Abegg, et al., *Angew. Chem. Int. Ed.*, 2015, **54**, 10852-10857.
18. J. Qian, et al., *Chem. Commun.*, 2012, **48**, 4091-4093.
19. *Nat Rev Drug Discov*, 2012, **11**, 895-895.
20. P. S. Changelian, et al., *Science*, 2003, **302**, 875-878.
21. J. D. Clark, et al., *J. Med. Chem.*, 2014, **57**, 5023-5038.
22. J. Singh, et al., *Curr. Opin. Chem. Biol.*, 2010, **14**, 475-480.