

Cite this: *Chem. Sci.*, 2021, 12, 12098

All publication charges for this article have been paid for by the Royal Society of Chemistry

A direct-to-biology high-throughput chemistry approach to reactive fragment screening†

Ross P. Thomas,^{ab} Rachel E. Heap,^c Francesca Zappacosta,^c Emma K. Grant,^a Peter Pogány,^{id a} Stephen Besley,^a David J. Fallon,^a Michael M. Hann,^{id a} David House,^a Nicholas C. O. Tomkinson^{id b} and Jacob T. Bush^{id *a}

Methods for rapid identification of chemical tools are essential for the validation of emerging targets and to provide medicinal chemistry starting points for the development of new medicines. Here, we report a screening platform that combines 'direct-to-biology' high-throughput chemistry (D2B-HTC) with photoreactive fragments. The platform enabled the rapid synthesis of >1000 PhotoAffinity Bits (HTC-PhABits) in 384-well plates in 24 h and their subsequent screening as crude reaction products with a protein target without purification. Screening the HTC-PhABit library with carbonic anhydrase I (CAI) afforded 7 hits (0.7% hit rate), which were found to covalently crosslink in the Zn²⁺ binding pocket. A powerful advantage of the D2B-HTC screening platform is the ability to rapidly perform iterative design–make–test cycles, accelerating the development and optimisation of chemical tools and medicinal chemistry starting points with little investment of resource.

Received 29th June 2021
Accepted 5th August 2021

DOI: 10.1039/d1sc03551g

rsc.li/chemical-science

Introduction

Human genetics and functional genomic studies are providing insights into potential therapeutic targets on an unprecedented scale.^{1–5} Validation of these targets in early-stage discovery relies on the rapid identification of tool molecules that can inform on the mechanism of action best suited to therapeutic intervention.⁶ Reactive fragment-based technologies have recently emerged as a powerful approach for the identification of tool molecules.^{7–12} Reactive fragments leverage the ability of relatively small compounds (<300 Da) to efficiently explore chemical space, coupled to a reactive functionality to enable the capture and robust detection of transient fragment-protein interactions.^{13–18} Crucially, covalent bond formation opens up a suite of follow-up studies, such as site identification and cellular target engagement, providing valuable tools for the study of emerging targets.^{19–21}

Initial reports of reactive fragments have focussed on the use of cysteine-reactive electrophilic fragment libraries (e.g. α,β -unsaturated carboxylic esters and amides).^{7–12,17,54} These platforms have enabled the development of numerous chemical probes, including for previously unliganded proteins.^{8,9} A limitation of this approach is the requirement for a cysteine residue

in the proximity of the binding pocket. More recently, platforms have been developed employing photoreactive groups.^{22–26} Irradiation by UV light converts the photo-labile functionality into a highly reactive intermediate (typically a carbene/nitrene) that covalently inserts into a proximal amino acid residue.^{27–29,56} This expands the number of protein pockets amenable to reactive fragment screening, greatly enhancing the applicability of such strategies. This approach was employed by Parker, Cravatt and co-workers to enable the mapping of fragment-protein interactions in cells using fully functionalised fragments (FFFs).^{30,31} Building upon this technology, we recently reported a photoreactive fragment screening platform (PhABits) which enabled the rapid identification of tools for purified proteins of interest to support early-stage biomedical research (Fig. 1a).³²

A key challenge for reactive fragment platforms, and indeed any chemical tool/medicinal chemistry campaign, is the optimisation of initial hits to generate more potent and selective tool compounds.^{14,55} In these early medicinal chemistry design–make–test cycles, compound synthesis presents the major bottleneck. We anticipated an opportunity to accelerate hit optimisation through the combination of high-throughput chemistry (HTC) with direct-to-biology (D2B) screening. The approach would involve the synthesis of whole libraries of reactive fragments in a 384-well plate-based format followed by direct screening of the crude products with a protein target.^{33–39} Crucially, the intact protein LC-MS screening method provides a quality control for the chemical identity of any hit fragments, as determined by the observed protein mass shift. Thus reactive fragment screening would overcome the challenges

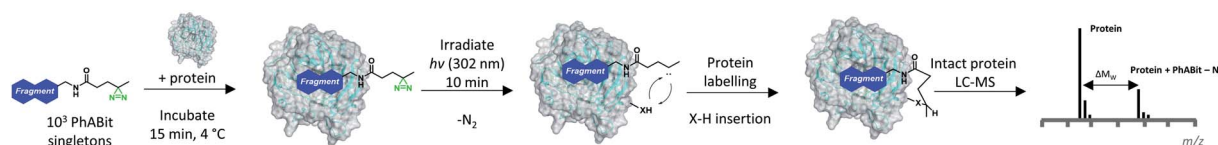
^aGlaxoSmithKline, Gunnels Wood Road, Stevenage, Hertfordshire, SG1 2NY, UK. E-mail: jacob.x.bush@gsk.com

^bDepartment of Pure and Applied Chemistry, University of Strathclyde, 295 Cathedral Street, Glasgow, G1 1XL, UK

^cGlaxoSmithKline, South Collegeville Road, Collegeville, PA 19426, USA

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d1sc03551g

a Previous work: The PhotoAffinity Bit (PhABit) screening platform



b This work: Direct-to-biology high-throughput chemistry (D2B-HTC) PhABit screening

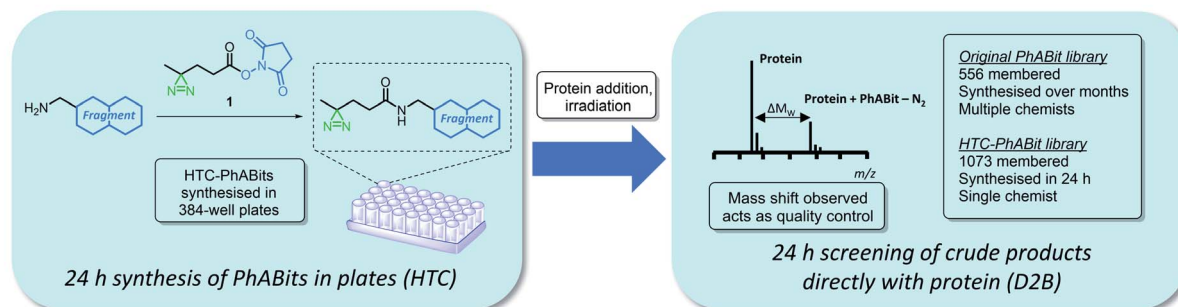


Fig. 1 (a) Schematic representation of previous work describing the PhotoAffinity Bit (PhABit) screening technology. (b) D2B-HTC PhABit screening platform. PhABits can be synthesised in 384-well plates in 24 h and screened as crude reaction products with a protein target.

encountered in historical D2B-HTC approaches, where the prevalence of false positives due to low levels of impurities in the reaction mixtures has prohibited widespread adoption.^{40–42}

Herein, we report a D2B-HTC photoreactive fragment screening platform (HTC-PhABits). A library of >1000 HTC-PhABits were synthesised in 384-well plates and screened directly without purification by irradiation in the presence of a purified protein of interest (Fig. 1b). An LC-MS analytical method allowed robust identification of light-induced covalent crosslinking events, leading to expedient identification of fragment hits. Crucially, HTC was subsequently used to synthesise and screen a second-generation library of 100 hit analogues to improve hit rate and crosslinking yield. HTC-PhABits thus offer a significant advance on our previously reported 556-membered PhABit library, where synthesis and purification were conducted over a number of months and hit optimisation was hindered by the resource requirements of hit analogue synthesis. We expect this platform to substantially accelerate PhABit synthesis and screening and enable rapid design-make-test cycles with minimal investment of resource and in short timeframes, thereby avoiding the compound synthesis bottleneck. Coupled with informative hit follow-up studies, such as binding site identification, we anticipate the platform will be widely applicable for identification of tool molecules for emerging protein targets of interest.

Results and discussion

Development of a high-throughput PhABit synthesis

A D2B-HTC approach to reactive fragment screening could provide an efficient means to grow and optimise reactive fragment libraries. Such a platform requires robust, high yielding, DMSO-compatible reactions, obviating the need for

purification. Furthermore, the reactions should be biocompatible with the protein screening step and form innocuous by-products (Fig. 2a).

The succinimide-activated (OSu) amide coupling provided a suitable option, as a high yielding reaction that affords *N*-hydroxysuccinimide as the only by-product.⁴³ It was anticipated that the diazirine OSu-ester **1** could be added to a plate containing fragment amines to afford the desired HTC-PhABits in sufficient purity to allow direct screening against targets of interest. Diazirine **1** was selected as the photoreactive functionality based on the high reactivity of the associated carbene, which limits non-specific labelling.^{24,28,44}

Reaction conditions were optimised to enable coupling of the diazirine OSu-ester **1** with a library of amines in 384-well plate format at room temperature. DMSO was selected as the reaction solvent to ensure direct compatibility with biochemical assays. The selection of base for the reaction was governed by both miscibility with DMSO and boiling point, as evaporation rates can be a concern in plate-based syntheses (Fig. 2a).⁴⁵ Common bases, triethylamine and *N,N*-diisopropylethylamine, were found to be unsuitable for the reaction as they were immiscible with DMSO. *N*-Methylmorpholine (NMM) exhibited good miscibility, however, was found to evaporate rapidly from DMSO in a 384-well plate format (b.p. = 116 °C, half-life = 150 min, recorded at ambient temperature) (Fig. S1†). Therefore, *N*-ethylmorpholine (NEM) was employed, which retained good miscibility with DMSO and exhibited a slower evaporation rate (b.p. = 138 °C, half-life = 270 min, recorded at ambient temperature) (Fig. S1†). Trial reactions were performed on a selection of 4 representative amines (1 eq. amine (S1–4), 1.1 eq. **1**, 3 eq. NEM, 100 mM in DMSO, total volume = 80 μL). Alkyl amines *p*-methoxybenzylamine (S1) and *p*-nitrobenzylamine (S2) afforded complete conversion to the desired product after



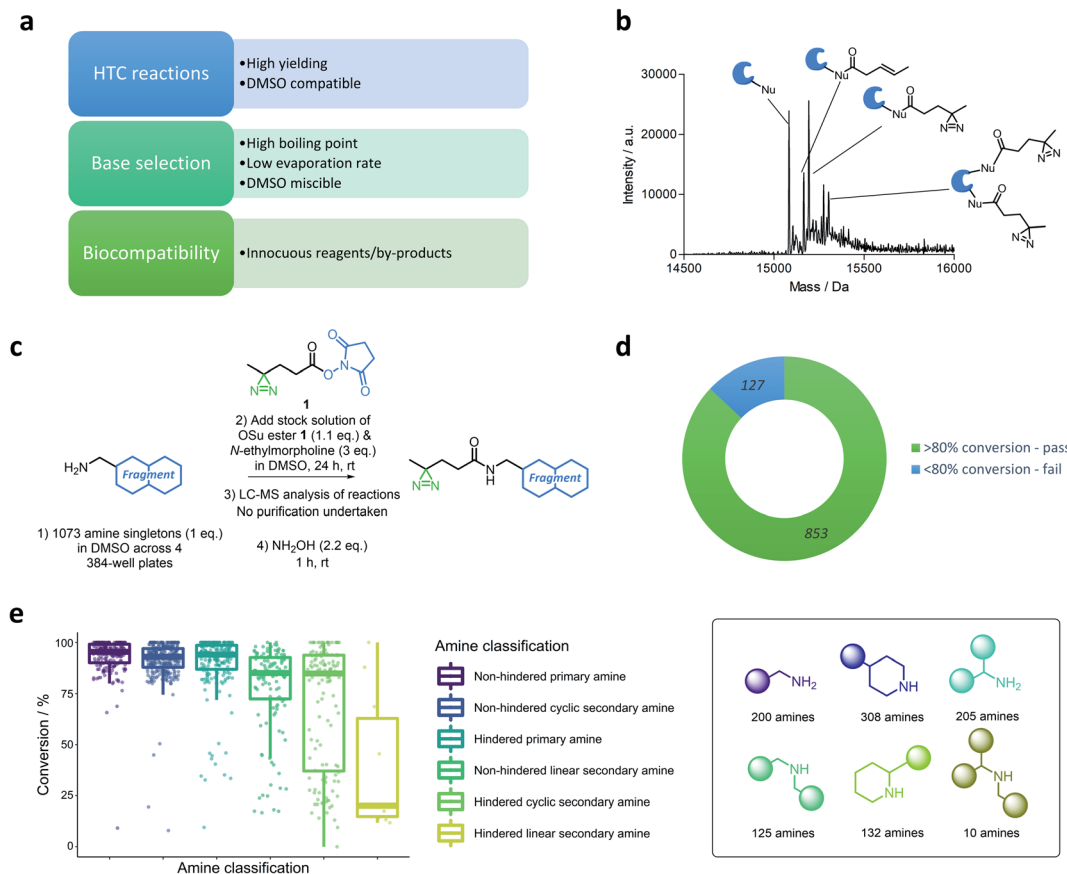


Fig. 2 (a) Criteria required for the development of a high-throughput chemistry direct-to-biology workflow. (b) Mass spectrum displaying the non-specific adducts formed between an exemplar protein (BRD4-BD1) and an irradiated (302 nm, 10 min) HTC-PhABit containing unreacted succinimide activated ester **1**. (c) Protocol for the in-plate synthesis of a library of high-throughput PhABits (HTC-PhABits). 1073 PhABits were synthesised without purification across 4 × 384-well plates. (d) Success rate of the HTC-PhABit synthetic protocol, omitting the 54 impure amine starting materials and the 39 reactions for which conversion was unable to be calculated. (e) Boxplot for the conversion of the HTC-PhABit reactions grouped by the classification of the amine. Amines were classified according to the following categories: hindered vs. unhindered, secondary vs. primary and linear vs. cyclic.

24 hours at room temperature, while aniline (**S3**) and *N*-methylaniline (**S4**) displayed poor conversion (Table S1†). Aryl amines were therefore excluded from this library of amine fragments.

Initial studies on the biocompatibility of these crude reaction products with protein screening indicated that unreacted succinimide ester **1** could react non-specifically with nucleophiles on the protein surface forming multiple adducts, confounding the mass spectrometry analysis (Fig. 2b). Therefore, a quenching step was introduced to remove unreacted OSu-ester **1** by reaction with hydroxylamine ($\text{NH}_2\text{OH}_{(\text{aq})}$, 2.2 eq., 1 h, rt).

Library synthesis

With the D2B-HTC reaction conditions optimised (Fig. 2c), a library of 1073 alkyl-amine fragments (mean M_w 211 Da) were selected from the GSK compound collection using clustering on molecular fingerprints (ECFP4) to maximise diversity (for library properties see Fig. S2†). This library included 288 alkylamines present in our previously reported PhABit library.³²

The amine fragments were plated across 4 × 384-well plates (10 mM DMSO) and subjected to LC-MS analysis to assess initial purity. This resulted in 54 amines being removed from subsequent analysis due to insufficient purity. To each well was added a stock solution of compound **1** (1.1 eq.) and *N*-ethylmorpholine (3 eq.) in DMSO (Fig. 2c, final HTC-PhABit concentration 5 or 8 mM). The plates were sealed and incubated at room temperature for 24 hours. The wells were then analysed directly by LC-MS to assess conversion. The success of the reaction was determined by purity (>80% purity was classified as a pass (626 reactions)). HTC-PhABits with purity <80% were further investigated to determine the cause of the low LC-MS purity according to the flow diagram provided in the ESI (Fig. S3†). This more in depth analysis rescued 227 reactions, to afford a total of 853 successful reactions. The conversion was unable to be determined for 39 reactions due to either failed LC-MS or the product not being visible by LC-MS. Therefore, conversion was calculated for a total of 980 reactions. The HTC reaction success rate was therefore 87% when considering only amines for which a conversion figure could be determined and omitting those



with impure amine starting materials (Fig. 2d, for full breakdown see Fig. S4†).

To investigate the influence of amine structure on the success rate of the synthetic procedure, the amines were grouped into 6 categories according to: primary/secondary, cyclic/linear and hindered/unhindered (amines with ≥ 2 substituents α to the nitrogen were classified as hindered) (Fig. 2e). Primary amines, both hindered and unhindered, afforded high levels of conversion with few failures (median conversion = $94 \pm 14\%$ ($n = 205$) and $95 \pm 9\%$ ($n = 200$) respectively). Unhindered cyclic/linear secondary amines also afforded good conversions (median conversion = $93 \pm 9\%$ ($n = 308$) and $85 \pm 21\%$ ($n = 125$) respectively), while cyclic/linear secondary amines featuring an α -substitution gave the largest variation in conversion and a significant number of failures (median conversion = $85 \pm 30\%$ ($n = 132$) and $20 \pm 34\%$ ($n = 10$) respectively). The low conversions of hindered secondary amines can be rationalised by sterics, while the lower conversions observed with linear *versus* cyclic secondary amines can be attributed to steric and conformational factors. It is possible that these amines could be rescued through the use of alternative coupling conditions or elevated temperature.

Direct-to-biology HTC-PhABit screen against human carbonic anhydrase I

Human carbonic anhydrase I (CAI) was selected for validation of the HTC-PhABit library as a proof of concept on account of the availability of reported crystallographic information and known inhibitors for validation studies. CAI is a member of a family of Zn^{2+} binding metalloenzymes that catalyse the physiological conversion of CO_2 to HCO_3^- , a key respiratory process and method of pH regulation.^{46,47} A number of inhibitors of carbonic anhydrase isozymes have been reported, which commonly feature a primary sulfonamide group.^{46,47} Carbonic anhydrase inhibitors have been used to treat glaucoma, oedema, cancer and can act as diuretics.⁴⁶

CAI (1 μM , PBS) was incubated with the HTC-PhABit library (100 μM) in 384-well plates before irradiation with UV light (10 min, 302 nm). The HTC-PhABit library was used assuming quantitative conversion during library synthesis. Following irradiation, the plates were directly analysed by intact protein LC-MS to identify any wells where photo-induced crosslinking had occurred. The majority of wells gave spectra that contained a single peak corresponding to unmodified protein, which provided validation of the biocompatibility of the HTC protocol. Hit fragments were characterised by those that displayed $>1.5\%$ crosslinking yield. Such hits contained a second peak at a mass shift corresponding to the mass of [PhABit- N_2], which enabled identification of these as binding hits and also provided confirmation of the chemical identity of the hit. It is important to note that the absolute value of the crosslinking yield is not indicative of binding affinity, but a composite of the binding affinity and the crosslinking efficiency of the carbene intermediate generated upon irradiation. Thus, crosslinking yields were not used to rank PhABit affinity. Of the 1073 HTC-PhABits screened, 7 displayed crosslinking yields $>1.5\%$,

corresponding to a 0.7% hit rate (Fig. 3a, S5 and Table S2†). Many of these hits contained aryl primary sulfonamide functionalities (2, 3, 5, 7 and 8, Fig. 3a) which are known to coordinate to the Zn^{2+} ion in the active site and therefore set up reversible binding prior to covalent crosslinking, while some novel chemotypes were also observed (4 and 6, Fig. 3a).⁴⁶

Concentration response and competition assay

Hit follow-up was performed to further validate the detected HTC-PhABit hits and to subsequently compare the behaviour of hits derived from D2B-HTC *versus* traditional synthesis and purification. Competition studies with ethoxzolamide (K_i 25 nM) were carried out to elucidate whether the 7 HTC-PhABit hits were binding in the same binding site.⁴⁶ The 7 unpurified HTC-PhABits (100 μM) were incubated and irradiated in the presence of either ethoxzolamide (100 μM) or DMSO as a control. All 7 HTC-PhABits (2–8), including novel structures 4 and 6, were displaced by ethoxzolamide, suggesting these HTC-PhABits were binding in the Zn^{2+} pocket (Fig. 3b and Table S3†).

To determine the binding affinity of the 7 HTC-PhABit hits, the unpurified hits were incubated at varying concentrations (0–100 μM) with CAI protein (1 μM , PBS) and irradiated (302 nm, 10 min). Crosslinking yields, as determined by intact protein LC-MS analysis, were plotted *vs.* concentration (Fig. S6†). Only 5 of the 7 HTC-PhABits screened in concentration-response displayed a measurable dissociation constant ($\text{p}K_D = 4.1\text{--}6.3$, Fig. 3c). These 5 HTC-PhABits (2, 3, 5, 7 and 8) all contained a primary sulfonamide. The strong binding affinities for 2, 7, and 8 are consistent with K_i values of known fragment-like sulfonamides. For example, ethoxzolamide, acetazolamide and methazolamide ($M_w < 300$ Da) have reported K_i values in the range 25–250 nM.⁴⁶ HTC-PhABits 4 and 6 showed a linear relationship between concentration and crosslinking yield, suggesting a K_D value >100 μM (Fig. S6†).

To investigate the consequence of screening with HTC-generated crude products, the concentration-response study was repeated following re-synthesis and purification of the 5 sulfonamide hits (2, 3, 5, 7 and 8) (Fig. 3c–f). There was good correlation between the $\text{p}K_D$ values obtained for HTC-PhABits and the purified PhABits (Fig. 3f, $R^2 = 0.78$). It was noted that the HTC-PhABit study underestimated the binding affinity in 3/5 cases, likely due to lower purity. However, the overall correlation provides support for the utility of the D2B-HTC protocol in rapidly characterising hits, prior to any subsequent re-synthesis of compounds of interest.

The crosslinking yields of PhABits 2, 3, 5, 7 and 8 with carbonic anhydrase II (CAII) were investigated to explore isoform selectivity (Table S4†). All PhABits were found to crosslink to this related isoform, and subsequent concentration-response studies differentiated between ‘pan-CAI/II’ fragments, and those that offered selectivity for a single isoform (Fig. 3c and S8†).

Site of crosslinking identification

A key advantage of covalent crosslinking is the ability to determine the site of binding by leveraging LC-MS/MS analysis. CAII



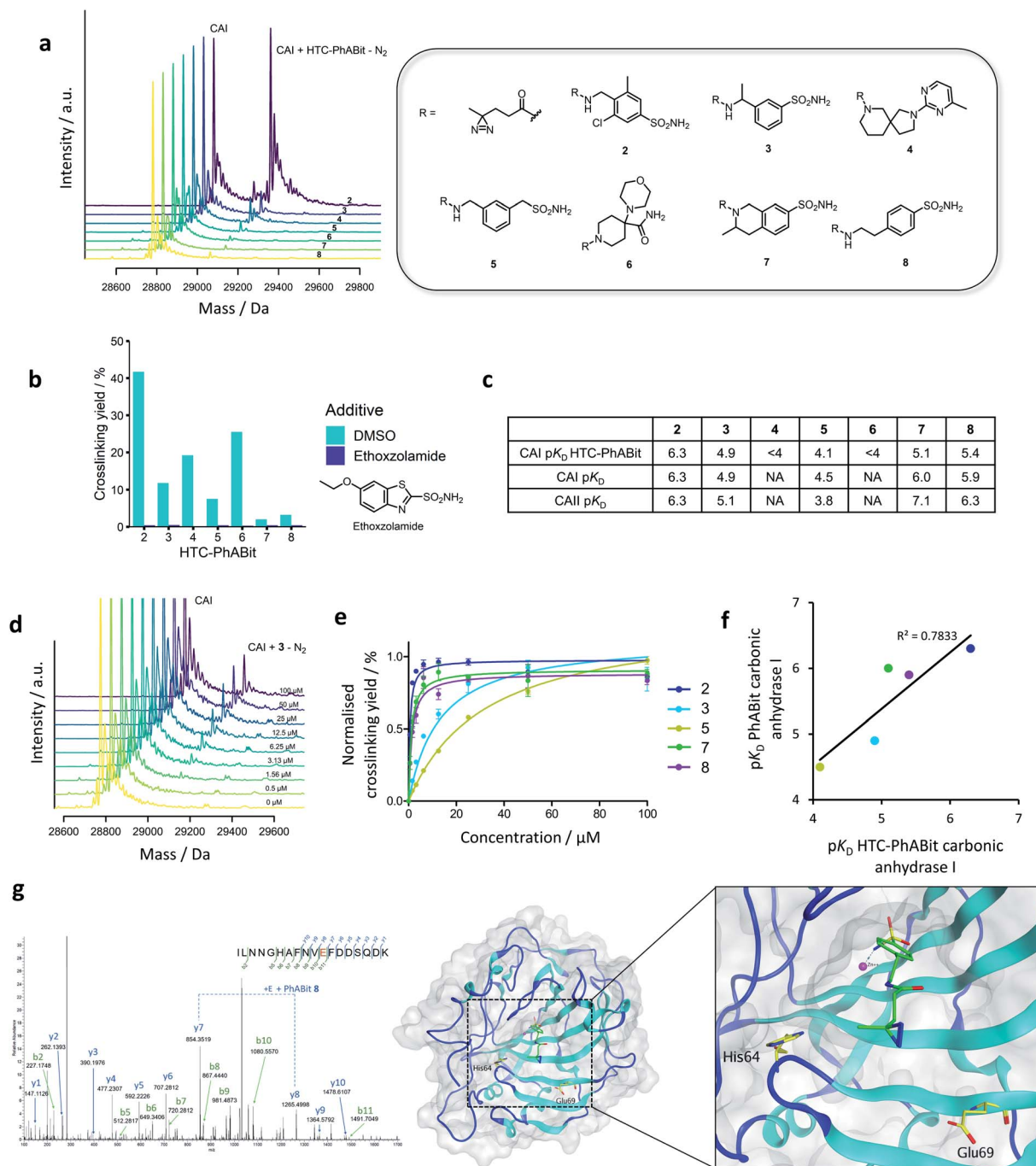


Fig. 3 (a) Summary of the single shot screen of the HTC-PhABit library against carbonic anhydrase I (CAI), showing the structures and corresponding mass spectra for 7 hits. (b) Crosslinking competition study with known CA inhibitor, ethoxzolamide. (c) pK_D values of the five sulfonamide containing PhABits as both crude products and purified with carbonic anhydrase I & II (concentration-response data recorded in triplicate). (d) Exemplar concentration-response mass spectra for PhABit 3. (e) Normalised concentration-response curves for the five purified sulfonamide containing PhABits with carbonic anhydrase I. (f) Comparison of the pK_D values obtained through the concentration-response studies of the five sulfonamide containing hits with CAI as both crude products and purified compounds. (g) Left: LC-MS/MS spectra of the peptide 59 ILNNGHAFNVEDDSQDK $_{76}$ crosslinked to **8** indicating Glu69 as one of the sites of crosslinking; Right: the X-ray crystal structure of carbonic anhydrase II (PDB: 3CAJ) virtually docked with PhABit **8**, highlighted are the two residues that were modified by all PhABits (His64 and Glu69).

labelled with a PhABit (**3**, **5**, **7** and **8**) was subjected to digestion with trypsin, followed by LC-MS/MS analysis. CAII was selected for this study due to higher levels of crosslinking observed for

the hit PhABits. For all four compounds, a data-dependent analysis identified peptide 59–76 ILNNGHAFNVEDDSQDK as carrying the modification (282.1038 Da for **3**, **5**, **8** and



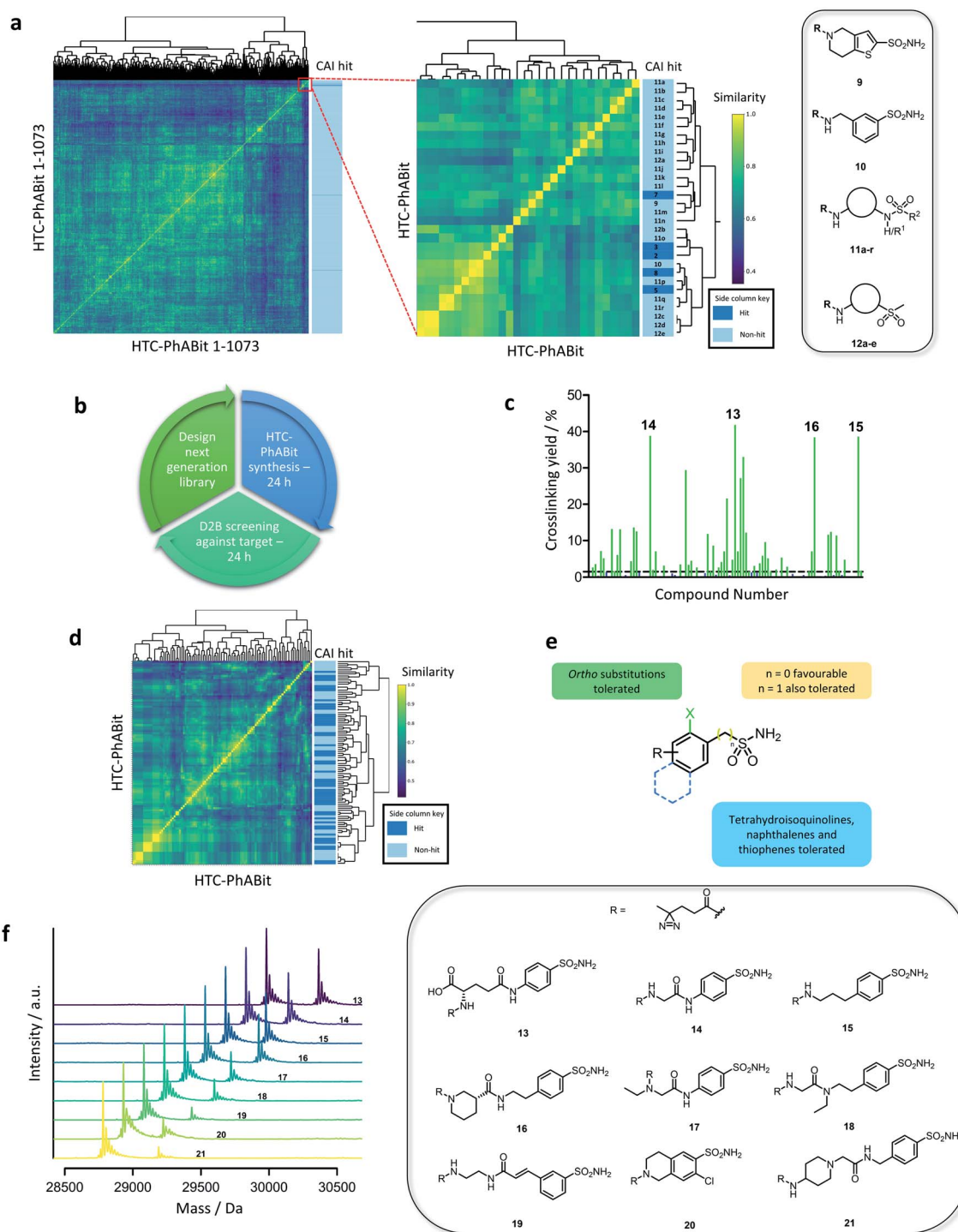


Fig. 4 (a) Visualisation of the library through similarity matrix and heatmap generation with a side column indicating crosslinking to CAI (dark blue: hit) along with structures of non-crosslinking primary sulfonamides (**9–10**) and generalised structures of non-crosslinking secondary/tertiary sulfonamides (**11a–r**) and sulfones (**12a–e**). Dendrograms link structurally similar compounds and are a result of the hierarchical clustering method used. (b) Proposed iterative design–make–test hit ID protocol utilising the D2B HTC-PhABit platform to identify binders. (c) Crosslinking yields of the 100 HTC-PhABits with CAI (green: >1.5% crosslinking, blue: <1.5% crosslinking). (d) Heatmap displaying the similarity between the members of the 100-membered second-generation HTC-PhABit library with a side column indicating crosslinking to CAI (dark blue: hit). (e) Generalised structure of a hit primary sulfonamide summarising the tolerated features for binding in the Zn^{2+} binding pocket. (f) Structures and corresponding mass spectra for the 9 highest crosslinking hits from the single shot screen.

308.1195 Da for 7). No modified peptides were observed in the non-irradiated controls. To determine the specific residue of modification, the peptide digests were subjected to parallel

reaction monitoring (PRM) LC-MS/MS, targeting the $[\text{M} + 3\text{H}]^{3+}$ ion of the 59–76 modified peptide. Manual interpretation of the PRM MS/MS spectra identified His64 and Glu69 as major sites

of modification for the four hit compounds (for example spectra see Fig. 3g).

The residues His64 and Glu69 are located at the entrance of the Zn²⁺ binding pocket. A virtual docking of PhABit 8, based on the reported binding mode of ethoxzolamide (PDB: 3CAJ), highlighted the two residues are suitably positioned for cross-linking to the docked PhABit (Fig. 3g).⁴⁸ Interestingly, the tetrapeptide ₆₅AFNV₆₈ between the two modified residues was largely unmodified by the PhABits (only PhABit 7 labelled Ala65 in very low stoichiometry) (Fig. S9†). The lack of labelling observed for these four non-polar amino acids highlights a potential selectivity of the activated diazirine towards particular residues. This is consistent with previous reports that carbenes more frequently label polar residues.⁴⁹ Semi-quantitative analysis of relative crosslinking yields indicated that all PhABits displayed highest labelling at Glu69, with the exception of PhABit 5 which displayed a higher labelling of His64.

Iterative library synthesis using direct-to-biology high-throughput chemistry

A key advantage of D2B-HTC is the ability to synthesise iterative libraries of hit-analogues in short timeframes to identify more potent binders, as well as improve target selectivity through off-target screening. It is anticipated that early stage discovery could be supplemented by rapid D2B-HTC design-make-test cycles to generate large volumes of data for protein targets, enabling rapid exploration of SAR and evolution towards optimised tools.

Interpretation of the initial CAI screen and elucidation of any structure activity relationships (SAR) emerging was required for the design and selection of a second-generation HTC-PhABit library. Therefore, a similarity matrix of the HTC-PhABit library was constructed and visualised as a heatmap. Using software package RCDK in R-Studio, the library of 1073 HTC-PhABits were converted to molecular fingerprints (MACCS) and the similarity (Tanimoto) was calculated between all library members.^{50–53} Hierarchical clustering based on structural features afforded a 1073 × 1073 similarity matrix, which was visualised as a heatmap (Fig. 4a). A region of close similarity (inset Fig. 4a, S10 and Table S5†) was identified that contained 7 primary sulfonamides, including hits (2, 3, 5, 7 and 8) and non-hits (9 and 10, Fig. 4a), as well as several secondary/tertiary sulfonamides (11a–r) and aryl sulfones (12a–e) each of which did not crosslink to the protein in the initial screen. Upon closer inspection, compound 9 showed minor crosslinking (1.0%), which did not meet the threshold to be classified as a hit (>1.5%). Compound 10, a close analogue of hit 3, showed no crosslinking, suggesting the benzylic methyl group was either crucial to potency or generated a conformation favourable to crosslinking.

Examination of the structures of the other HTC-PhABits within this region of the heatmap provided further insight into the SAR around the hits. Numerous secondary/tertiary sulfonamides (11a–r) and aryl sulfones (12a–e) were present that did not display crosslinking to CAI (Fig. 4a), highlighting

the primary sulfonamide as a privileged chemotype for carbonic anhydrase, a feature consistent with previously reported inhibitors.⁴⁶

Taking this information together, a 100-member library of hit-analogues was designed based on the structures of initial hits 2, 3, 5, 7 and 8 comprising 83 primary sulfonamides as well as 17 secondary/tertiary sulfonamides and sulfones as negative controls (Fig. 4b) (for library properties see Fig. S11†). The library was synthesised and screened against CAI following the established D2B-HTC protocol, which yielded a significant improvement in hit rate (52 hits >1.5% crosslinking) (Fig. 4c). The crosslinking yields were also improved with 17 fragments giving greater than 10% labelling. This is particularly impactful to the PhABit platform, in providing a means to overcome the challenges associated with the often low crosslinking yields of photoreactive fragments. D2B-HTC cycles can thus enable optimisation toward PhABits that give high crosslinking yields, allowing them to serve as valuable tools for the study of proteins of interest.

Inspection of the hit and non-hit fragment structures provided further insights into the SAR within the Zn²⁺ binding pocket (Fig. 4d–f). All 52 hits contained a primary sulfonamide, while none of the secondary/tertiary sulfonamides and sulfones displayed crosslinking. The majority of the hits (77%) featured a primary sulfonamide directly bonded to a phenyl ring. Other ring systems were also tolerated such as tetrahydroisoquinoline sulfonamide (7 hits), thiophene sulfonamide (1 hit), benzyl sulfonamide (1 hit) and naphthalene sulfonamide (1 hit). Several hits (19%) featured an *ortho* substituent, suggesting additional space in the pocket adjacent to the Zn²⁺ ion. Conversely, inspection of the non-hits identified 1*H*-indole-5-sulfonamide, 1*H*-indole-5-benzylsulfonamide and 1*H*-isoindoline-5-sulfonamide as potentially non-binding primary sulfonamides. However, a limitation of analysing negative hits from photoaffinity labelling experiments is the potential for false negatives due to poor crosslinking.

Overall, the application of iterative design-make-test cycles to the D2B-HTC-PhABit platform enabled the identification of a large number of new elaborated hits. Furthermore, a higher proportion of these hits displayed >10% crosslinking making them suitable chemical tools for downstream labelling experiments. It has also been demonstrated that detailed SAR patterns can be elucidated by analysing the results from the iterative screening process.

Conclusion

The D2B-HTC platform developed here provides a powerful approach to enable the rapid iterative synthesis of large libraries (10³) of photoreactive fragments and subsequent screening with single purified protein targets of interest directly as crude reaction products, without any purification. Such a screening platform is anticipated to enable the facile identification and optimisation of hit fragments into more potent and selective tool molecules, overcoming the significant synthetic bottleneck in previous reactive fragment screening platforms. The HTC-PhABit approach exploits the intrinsic quality control



of the mass spectrometric read-out, confirming the identity of hit fragments and thereby overcoming the challenges associated with false positives, which have hindered D2B-HTC platforms to date.

The HTC-PhABit approach enabled the rapid generation of >1000 PhABits that were directly screened against carbonic anhydrase I. The screening process identified multiple micromolar hits, which were validated by re-synthesis and purification. Follow-up studies enabled determination of potency and site of binding. In a second screening cycle, a library of 100 hit analogues was rapidly synthesised and screened to yield 52 additional hits with improved crosslinking yields. This follow-up screen provided multiple insights into the SAR with very little investment of resource relative to traditional medicinal chemistry design-make-test cycles. With the rate and scale of data generation offered by D2B-HTC, we envision that this platform will provide an opportunity to rapidly identify more potent and selective tool molecules for emerging protein targets. In order to fully realise the potential of this approach, a key next step will be to expand the scope of D2B-HTC-compatible reactions, providing multiple avenues for the development of high-quality tools.

Author contributions

RPT, REH, SB and FZ performed experiments and analysed data. RPT, EKG, JTB and NCOT drafted the paper. PP carried out computational studies. JTB, NCOT, MMH and DH contributed to experiment design. DJF developed code for data visualisation.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors thank the GlaxoSmithKline/University of Strathclyde Collaborative PhD programme for funding and chemical resources and the EPSRC for funding via Prosperity Partnership EP/S035990/1.

References

- 1 L. H. Jones, *Expert Opin. Drug Discovery*, 2016, **11**, 623–625.
- 2 L. H. Jones and M. E. Bunnage, *Nat. Rev. Drug Discovery*, 2017, **16**, 285–296.
- 3 J. A. Doudna and E. Charpentier, *Science*, 2014, **346**, 1077.
- 4 R. Kramer and D. Cohen, *Nat. Rev. Drug Discovery*, 2004, **3**, 965–972.
- 5 B. Haley and F. Roudnicki, *Cancer Cell*, 2020, **38**, 31–43.
- 6 J. P. Hughes, S. Rees, S. B. Kalindjian and K. L. Philpott, *Br. J. Pharmacol.*, 2011, **162**, 1239–1249.
- 7 H. Johansson, Y.-C. Isabella Tsai, K. Fantom, C.-W. Chung, S. Kümper, L. Martino, D. A. Thomas, H. C. Eberl, M. Muelbauer, D. House and K. Rittinger, *J. Am. Chem. Soc.*, 2019, **141**, 2703–2712.
- 8 E. Resnick, A. Bradley, J. Gan, A. Douangamath, T. Krojer, R. Sethi, P. P. Geurink, A. Aimon, G. Amitai, D. Bellini, J. Bennett, M. Fairhead, O. Fedorov, R. Gabizon, J. Gan, J. Guo, A. Plotnikov, N. Reznik, G. F. Ruda, L. Díaz-Sáez, V. M. Straub, T. Szommer, S. Velupillai, D. Zaidman, Y. Zhang, A. R. Coker, C. G. Dowson, H. M. Barr, C. Wang, K. V. M. Huber, P. E. Brennan, H. Ovaa, F. von Delft and N. London, *J. Am. Chem. Soc.*, 2019, **141**, 8951–8968.
- 9 C. Dubiella, B. J. Pinch, K. Koikawa, D. Zaidman, E. Poon, T. D. Manz, B. Nabet, S. He, E. Resnick, A. Rogel, E. M. Langer, C. J. Daniel, H.-S. Seo, Y. Chen, G. Adelmant, S. Sharifzadeh, S. B. Ficarro, Y. Jamin, B. M. da Costa, M. W. Zimmerman, X. Lian, S. Kibe, S. Kozono, Z. M. Doctor, C. M. Browne, A. Yang, L. Stoler-Barak, R. B. Shah, N. E. Vangos, E. A. Geffken, R. Oren, E. Koide, S. Sidi, Z. Shulman, C. Wang, J. A. Marto, S. Dhe-Paganon, T. Look, X. Z. Zhou, K. P. Lu, R. C. Sears, L. Chesler, N. S. Gray and N. London, *Nat. Chem. Biol.*, 2021, DOI: 10.1038/s41589-021-00786-7.
- 10 S. G. Kathman, Z. Xu and A. V. Statsyuk, *J. Med. Chem.*, 2014, **57**, 4969–4974.
- 11 S. G. Kathman and A. V. Statsyuk, *MedChemComm*, 2016, **7**, 576–585.
- 12 M. D. Olp, D. J. Sprague, C. J. Goetz, S. G. Kathman, S. L. Wynia-Smith, S. Shishodia, S. B. Summers, Z. Xu, A. V. Statsyuk and B. C. Smith, *ACS Chem. Biol.*, 2020, **15**, 1036–1049.
- 13 D. E. Scott, A. G. Coyne, S. A. Hudson and C. Abell, *Biochemistry*, 2012, **51**, 4990–5003.
- 14 C. W. Murray and D. C. Rees, *Nat. Chem.*, 2009, **1**, 187–192.
- 15 M. Congreve, R. Carr, C. Murray and H. Jhoti, *Drug Discovery Today*, 2003, **8**, 876–877.
- 16 A. R. Leach, M. M. Hann, J. N. Burrows and E. J. Griffen, *Mol. Biosyst.*, 2006, **2**, 429–446.
- 17 C. Jöst, C. Nitsche, T. Scholz, L. Roux and C. D. Klein, *J. Med. Chem.*, 2014, **57**, 7590–7599.
- 18 A. Keeley, L. Petri, P. Ábrányi-Balogh and G. M. Keserü, *Drug Discovery Today*, 2020, **25**, 983–996.
- 19 K. M. Backus, B. E. Correia, K. M. Lum, S. Forli, B. D. Horning, G. E. González-Páez, S. Chatterjee, B. R. Lanning, J. R. Teijaro, A. J. Olson, D. W. Wolan and B. F. Cravatt, *Nature*, 2016, **534**, 570–574.
- 20 P. R. A. Zanon, L. Lewald and S. M. Hacker, *Angew. Chem., Int. Ed.*, 2020, **59**, 2829–2836.
- 21 E. V. Vinogradova, X. Zhang, D. Remillard, D. C. Lazar, R. M. Suci, Y. Wang, G. Bianco, Y. Yamashita, V. M. Crowley, M. A. Schafroth, M. Yokoyama, D. B. Konrad, K. M. Lum, G. M. Simon, E. K. Kemper, M. R. Lazear, S. Yin, M. M. Blewett, M. M. Dix, N. Nguyen, M. N. Shokhirev, E. N. Chin, L. L. Lairson, B. Melillo, S. L. Schreiber, S. Forli, J. R. Teijaro and B. F. Cravatt, *Cell*, 2020, **182**, 1009–1026.
- 22 E. K. Grant, D. J. Fallon, H. C. Eberl, K. G. M. Fantom, F. Zappacosta, C. Messenger, N. C. O. Tomkinson and J. T. Bush, *Angew. Chem., Int. Ed.*, 2019, **58**, 17322–17327.



- 23 J. T. Bush, L. J. Walport, J. F. McGouran, I. K. H. Leung, G. Berridge, S. S. van Berkel, A. Basak, B. M. Kessler and C. J. Schofield, *Chem. Sci.*, 2013, **4**, 4115–4120.
- 24 K. Sakurai, S. Ozawa, R. Yamada, T. Yasui and S. Mizuno, *ChemBioChem*, 2014, **15**, 1399–1403.
- 25 J. Park, M. Koh, J. Y. Koo, S. Lee and S. B. Park, *ACS Chem. Biol.*, 2016, **11**, 44–52.
- 26 B. Zhao and K. Burgess, *ACS Med. Chem. Lett.*, 2018, **9**, 155–158.
- 27 S.-S. Ge, B. Chen, Y.-Y. Wu, Q.-S. Long, Y.-L. Zhao, P.-Y. Wang and S. Yang, *RSC Adv.*, 2018, **8**, 29428–29454.
- 28 J. R. Hill and A. A. B. Robertson, *J. Med. Chem.*, 2018, **61**, 6945–6963.
- 29 L. Dubinsky, B. P. Krom and M. M. Meijler, *Bioorg. Med. Chem.*, 2012, **20**, 554–570.
- 30 C. G. Parker, A. Galmozzi, Y. Wang, B. E. Correia, K. Sasaki, C. M. Joslyn, A. S. Kim, C. L. Cavallaro, R. M. Lawrence, S. R. Johnson, I. Narvaiza, E. Saez and B. F. Cravatt, *Cell*, 2017, **168**, 527–541.
- 31 Y. Wang, M. M. Dix, G. Bianco, J. R. Remsberg, H.-Y. Lee, M. Kalocsay, S. P. Gygi, S. Forli, G. Vite, R. M. Lawrence, C. G. Parker and B. F. Cravatt, *Nat. Chem.*, 2019, **11**, 1113–1123.
- 32 E. K. Grant, D. J. Fallon, M. M. Hann, K. G. M. Fantom, C. Quinn, F. Zappacosta, R. S. Annan, C.-w. Chung, P. Bamborough, D. P. Dixon, P. Stacey, D. House, V. K. Patel, N. C. O. Tomkinson and J. Bush, *Angew. Chem., Int. Ed.*, 2020, **59**, 21096–21105.
- 33 M. Benz, M. R. Molla, A. Böser, A. Rosenfeld and P. A. Levkin, *Nat. Commun.*, 2019, **10**, 2879.
- 34 J. B. Murray, S. D. Roughley, N. Matassova and P. A. Brough, *J. Med. Chem.*, 2014, **57**, 2845–2850.
- 35 L. Liu, *J. Med. Chem.*, 2014, **57**, 2843–2844.
- 36 A. Leggott, J. E. Clarke, S. Chow, S. L. Warriner, A. J. O'Neill and A. Nelson, *Chem. Commun.*, 2020, **56**, 8047–8050.
- 37 G. Karageorgis, S. Warriner and A. Nelson, *Nat. Chem.*, 2014, **6**, 872–876.
- 38 Z. Liu, J. Li, S. Li, G. Li, K. B. Sharpless and P. Wu, *J. Am. Chem. Soc.*, 2018, **140**, 2919–2925.
- 39 S. Kitamura, Q. Zheng, J. L. Woehl, A. Solania, E. Chen, N. Dillon, M. V. Hull, M. Kotaniguchi, J. R. Cappiello, S. Kitamura, V. Nizet, K. B. Sharpless and D. W. Wolan, *J. Am. Chem. Soc.*, 2020, **142**, 10899–10904.
- 40 J. L. Dahlin and M. A. Walters, *Future Med. Chem.*, 2014, **6**, 1265–1290.
- 41 J. C. Hermann, Y. Chen, C. Wartchow, J. Menke, L. Gao, S. K. Gleason, N.-E. Haynes, N. Scott, A. Petersen, S. Gabriel, B. Vu, K. M. George, A. Narayanan, S. H. Li, H. Qian, N. Beatini, L. Niu and Q.-F. Gan, *ACS Med. Chem. Lett.*, 2013, **4**, 197–200.
- 42 Y. Shin, J. W. Jeong, R. P. Wurz, P. Achanta, T. Arvedson, M. D. Bartberger, I. D. G. Campuzano, R. Fucini, S. K. Hansen, J. Ingersoll, J. S. Iwig, J. R. Lipford, V. Ma, D. J. Kopecky, J. McCarter, T. San Miguel, C. Mohr, S. Sabet, A. Y. Saiki, A. Sawayama, S. Sethofer, C. M. Tegley, L. P. Volak, K. Yang, B. A. Lanman, D. A. Erlanson and V. J. Cee, *ACS Med. Chem. Lett.*, 2019, **10**, 1302–1308.
- 43 G. W. Anderson, J. E. Zimmerman and F. M. Callahan, *J. Am. Chem. Soc.*, 1963, **85**, 3039.
- 44 J. Wang, J. Kubicki, H. Peng and M. S. Platz, *J. Am. Chem. Soc.*, 2008, **130**, 6604–6609.
- 45 E. Berthier, J. Warrick, H. Yu and D. J. Beebe, *Lab Chip*, 2008, **8**, 852–859.
- 46 C. T. Supuran, *Nat. Rev. Drug Discovery*, 2008, **7**, 168–181.
- 47 C. L. Lomelino, C. T. Supuran and R. McKenna, *Int. J. Mol. Sci.*, 2016, **17**, 1–13.
- 48 A. Di Fiore, C. Pedone, J. Antel, H. Waldeck, A. Witte, M. Wurl, A. Scozzafava, C. T. Supuran and G. De Simone, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 2669–2674.
- 49 D. S. Ziemianowicz, R. Bomgarden, C. Etienne and D. C. Schriemer, *J. Am. Soc. Mass Spectrom.*, 2017, **28**, 2011–2021.
- 50 A. Cereto-Massagué, M. J. Ojeda, C. Valls, M. Mulero, S. Garcia-Vallvé and G. Pujadas, *Methods*, 2015, **71**, 58–63.
- 51 J. L. Durant, B. A. Leland, D. R. Henry and J. G. Nourse, *J. Chem. Inf. Comput. Sci.*, 2002, **42**, 1273–1280.
- 52 A. Voicu, N. Duteanu, M. Voicu, D. Vlad and V. Dumitrascu, *J. Cheminf.*, 2020, **12**, 1–8.
- 53 M. A. Fligner, J. S. Verducci and P. E. Blower, *Technometrics*, 2002, **44**, 110–119.
- 54 J. T. Bush, R. K. Leśniak, T. L. Yeh, R. Belle, H. Kramer, A. Tumbler, R. Chowdhury, E. Flashman, J. Mecinović and C. J. Schofield, *Chem. Commun.*, 2019, **55**, 1020–1023.
- 55 J. T. Bush, P. Pogany, S. D. Pickett and D. V. S. Green, *J. Med. Chem.*, 2020, **63**(20), 11964–11971.
- 56 D. J. Fallon, S. Lehmann, C. Chung, A. H. Phillipou, C. Eberl, K. G. M. Fantom, F. Zappacosta, V. K. Patel, M. Bantscheff, C. J. Schofield, N. C. O. Tomkinson and J. T. Bush, *Chem.–Eur. J.*, 2021, DOI: 10.1002/chem.202102036.

