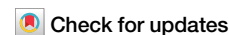


<https://doi.org/10.1038/s42004-024-01252-w>

Mapping protein binding sites by photoreactive fragment pharmacophores



Péter Ábrányi-Balogh ^{1,2,3,19}, Dávid Bajusz ^{1,2,19}, Zoltán Orgován^{1,2,19}, Aaron B. Keeley ¹, László Petri ^{1,2}, Nikolett Péczka^{1,2,3}, Tibor Viktor Szalai^{1,2,4}, Gyula Pálffy ⁵, Márton Gadanez ^{2,5,6}, Emma K. Grant ⁷, Tímea Imre^{2,8}, Tamás Takács^{2,9,10}, Ivan Randelović ^{11,12}, Marcell Baranyi^{12,13}, András Marton^{14,15}, Gitta Schlosser¹⁶, Qirat F. Ashraf¹⁷, Elvin D. de Araujo ¹⁸, Tamás Karancsi^{14,15}, László Buday^{2,9}, József Tóvári ¹¹, András Perczel ^{2,5}, Jacob T. Bush ⁷ & György M. Keserű ^{1,2,3} ✉

Fragment screening is a popular strategy of generating viable chemical starting points especially for challenging targets. Although fragments provide a better coverage of chemical space and they have typically higher chance of binding, their weak affinity necessitates highly sensitive biophysical assays. Here, we introduce a screening concept that combines evolutionary optimized fragment pharmacophores with the use of a photoaffinity handle that enables high hit rates by LC-MS-based detection. The sensitivity of our screening protocol was further improved by a target-conjugated photocatalyst. We have designed, synthesized, and screened 100 diazirine-tagged fragments against three benchmark and three therapeutically relevant protein targets of different tractability. Our therapeutic targets included a conventional enzyme, the first bromodomain of BRD4, a protein-protein interaction represented by the oncogenic KRas^{G12D} protein, and the yet unliganded *N*-terminal domain of the STAT5B transcription factor. We have discovered several fragment hits against all three targets and identified their binding sites via enzymatic digestion, structural studies and modeling. Our results revealed that this protocol outperforms screening traditional fully functionalized and photoaffinity fragments in better exploration of the available binding sites and higher hit rates observed for even difficult targets.

Fragment-based drug discovery (FBDD) has been a definitive trend of the past thirty years of drug discovery¹, and has come of age with the approval of vemurafenib² and further drugs that originated from FBDD programs^{3–5}. The main rationale in generating chemical starting points by screening moderate-sized libraries of less complex, small and polar compounds comes from the realization that screening a few thousand fragments (≤ 16 heavy atoms) provides a better sampling of the respective chemical space⁶ than screening millions of larger compounds (≤ 36 heavy atoms), translating to substantially higher hit rates, typically at a smaller cost¹. FBDD comes with its own challenges though, particularly in the experimental detection of the targeted binding site(s) of the respective hit(s). Biochemical screens are usually less sensitive to weak fragment binders and do not provide structural information on their binding site. Biophysical approaches such as the thermal shift assay, ligand-observed nuclear magnetic resonance (NMR) spectroscopy, surface plasmon resonance are typically used in primary fragment screening. Other techniques might offer more detailed results, but are more resource-intensive, e.g., X-ray screening needs the crystals of the

target protein available for high-concentration soaking experiments mostly at synchrotrons or protein observed NMR that requires labeled protein samples and access to high-field NMR facilities.

For the simple and efficient detection of fragment hits, the concept of photoaffinity labeling has gained significant ground in the past years. Photoaffinity probes usually contain a pharmacophore pattern recognized by the target protein that forms non-covalent interactions, and a photo-reactive group for the light-induced anchorage of the probe after the binding event^{7,8}. The captured compound can be identified directly on the labeled protein by mass spectrometry (MS)⁹, or the probe might be equipped with a biorthogonal handle to enable downstream workflows^{10,11}. This latter concept was pioneered at the Cravatt group introducing fully-functionalized fragments (FFFs)^{12,13} that were successfully utilized in diverse screening campaigns¹⁴, as well as chemical proteomics^{15,16}. Notably, the concept has since been endorsed by several major chemical vendors that distribute commercial libraries of FFFs. Recently, a GSK team developed a screening platform (PhotoAffinity Bits, or PhABits), that captures fragment-protein

A full list of affiliations appears at the end of the paper. ✉ e-mail: keseru.gyorgy@ttk.hu

interactions by photoaffinity labeling^{17,18}. In our setup, members of the fragment library constitute the variable, pharmacophore-optimized non-covalent cores for target binding, and are equipped with a diazirine photoaffinity tag to enable anchorage and MS-based detection¹⁸.

Efficient library design is a central phenomenon in fragment screening since library composition impacts the chemical space and the protein interactome covered. Recently we developed a design protocol, dubbed SpotXplorer, that provides fragment libraries with maximal coverage and diversity in their representation of the unique pharmacophore patterns that were detected in over 3000 experimental protein-fragment complexes¹⁹. Analyzing an FFF library of a major vendor (640 compounds) reveals an uneven distribution of experimentally validated and evolutionary conserved fragment-binding pharmacophores (Fig. 1a)¹⁹. By contrast, a pilot library of 96 SpotXplorer fragments provides balanced coverage of the unique fragment-binding pharmacophores (Fig. 1b). The pilot library has been successfully validated against both conventional (protease and GPCRs) and challenging targets (histone methyltransferase SETD2 and SARS-CoV-2 viral targets) using biochemical and X-ray screening.

The specificity of fragment binding has been addressed in multiple studies. An analysis of 35 fragment screening campaigns of Novartis concluded that 63% of the screened fragments had never been observed as hits, while a smaller number of privileged fragments were found to be active on more than one target²⁰. Similarly, many fragments that appear in crystallized complexes in the Protein Data Bank (PDB) have multiple, seemingly unrelated protein targets²¹. Interestingly, the analysis of the available X-ray structures suggested these fragments form specific binding interactions²². In line with these observations, we found that most fragments represent more than one unique pharmacophore and furthermore, multiple sites might recognize the same pharmacophore(s). Consequently, individual fragments can bind to diverse binding sites having the key recognition elements represented by the actual fragment pharmacophore. Considering the key objective of fragments screening as the identification of a potential binding site and a viable chemical starting point, specificity at the fragment level is not a strict requirement. In fact, in FBDD settings, binding specificity is typically ensured by the interactions formed upon the optimization of the initial fragment.

Maximizing the coverage of the experimentally validated set of binding pharmacophores at the fragment level, here we report a pharmacophore-optimized photoaffinity library (PhP) containing 100 fragments with diverse binding pharmacophores, equipped with a diazirine-type photoaffinity tag. We have characterized the library by screening against three benchmark proteins including carbonic anhydrase II, myoglobin and lysozyme. Next, we employed the library against three oncological targets with decreasing tractability: the first bromodomain of the BRD4 enzyme

(BRD4-BD1) with multiple known ligands, the oncogenic KRas^{G12D} GTPase with a few structurally related inhibitors, and the N-terminal domain of the STAT5B transcription factor, which has no reported ligands. Binding events were first confirmed by mass spectrometry and the exact site of labeling was identified after enzymatic digestion. The identified binding sites were characterized by X-ray crystallography, HSQC NMR experiments and modeling. Our efforts revealed that pharmacophore-optimized photoaffinity fragments map the available binding sites effectively, identifying yet unexploited, tractable sites on challenging drug targets and provided viable chemical starting points for small molecule drug discovery efforts.

Results

Library design, synthesis, and screening

The PhP library was compiled by the SpotXplorer technology¹⁹ providing the optimal coverage of the unique fragment-protein binding pharmacophores as its main objective. We have recently found that experimental fragment-binding modes are represented by a limited set of 425 unique binding pharmacophores with up to four features, including H-bond donors/acceptors, positively/negatively charged groups, aromatic rings and hydrophobic groups. After suitable ligand preparation, commercial fragment collections can be utilized for assembling small fragment libraries that provide optimal coverage against this set of experimentally validated and evolutionary-conserved unique pharmacophores and can thus be used in target-agnostic screening campaigns.

To compile the PhP library, we have utilized the Enamine primary amine collection, using the primary amine group as the attachment point for the photoreactive diazirine unit. Briefly, protomers and conformers of the available amines were screened against the unique pharmacophores, and selected with a diversity picker based on the pharmacophore sets that they represent. Since fragments typically have small sizes and limited pharmacophore elements²³ that form only a few characteristic interactions with their targets²⁴, we selected a minimal set of fragments covering the largest possible fraction of experimentally validated 2- and 3-point binding pharmacophores. During library preparation, we disregarded the primary amine function as a potential pharmacophore feature (as it served the purpose of the attachment point for the diazirine unit), and we discarded any molecules with multiple primary amine groups, as well as pan-assay interference compounds (PAINS)²⁵. A set of 160 compounds was assembled to provide optimal pharmacophore coverage (88% and 83%) and diversity over the 117 unique 2- and 3-point pharmacophores, respectively (Fig. 2a).

The 160 amines representing the set of experimentally validated fragment pharmacophores were coupled with the free carboxyl group of the photoaffinity tag using HATU coupling in a plate-based parallel synthesis setup. Products were subjected to purification and were analyzed by plate-

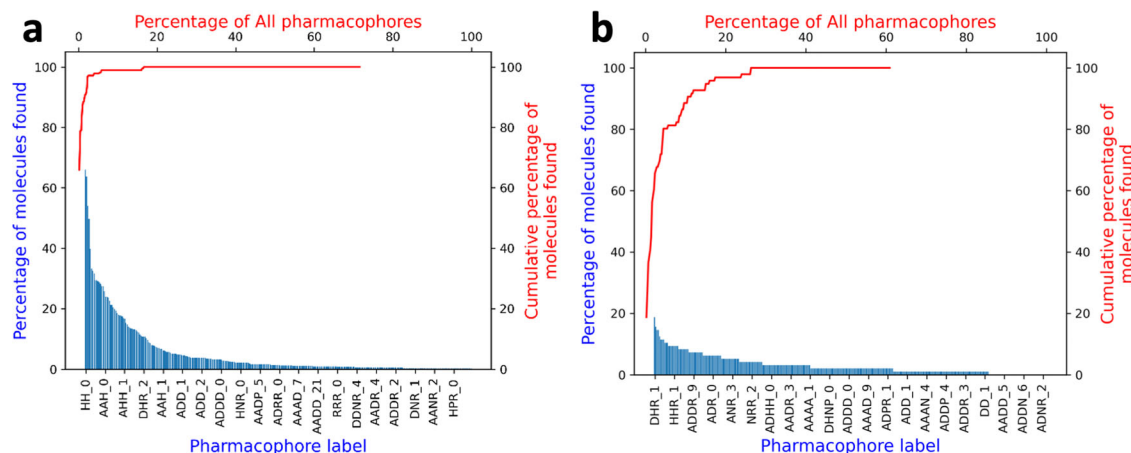


Fig. 1 | Pharmacophore coverage of fragment libraries. Relative occurrences (blue) of unique binding pharmacophores (lower x axis) in a commercial FFF library (a) vs. the SpotXplorer pilot library (b). The cumulative coverage of library members

(red line) further highlights that almost all members of the commercial library are represented by the few most privileged pharmacophores (vs. a more even distribution for the SpotXplorer library that is about six times smaller).

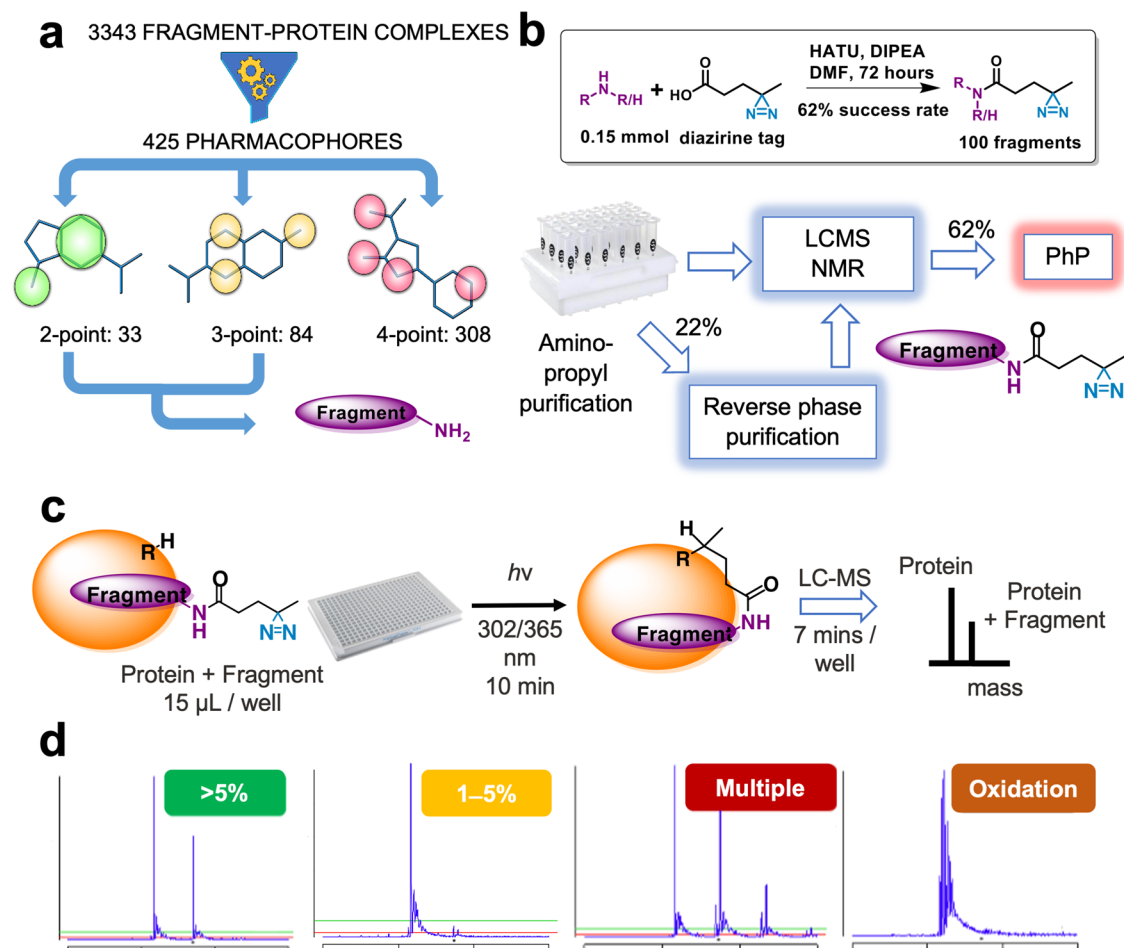


Fig. 2 | Concept, synthesis, and general screening workflow of the pharmacophore-optimized photoaffinity (PhP) library. **a** Primary amines from the Enamine collection were compiled with the SpotXplorer workflow to optimize diversity and coverage against 117 unique 2- and 3-point fragment binding pharmacophores. **b** HATU coupling in plate-based parallel synthesis and subsequent purification yielded 100 (out of 160) fragments equipped with a diazirine tag. **c** The

screening paradigm employs 10 min of irradiation with the protein target and subsequent MS detection of the anchored protein-fragment complex. **d** Screening results were categorized by the analysis of their mass spectra as strong binders (>5% labeling), weak binders (1–5% labeling), multiple binders (multiple labeling), oxidized samples (having only protein +16 or multiplied peaks) and non-binders (<1% labeling) samples showing no change in the MS spectra.

based LC-MS and NMR that confirmed 100 photoaffinity labeled fragments, representing a 62% synthesis success rate (Fig. 2b). The final PhP library covers 88% and 75% of the experimentally validated 2- and 3-point pharmacophores, respectively.

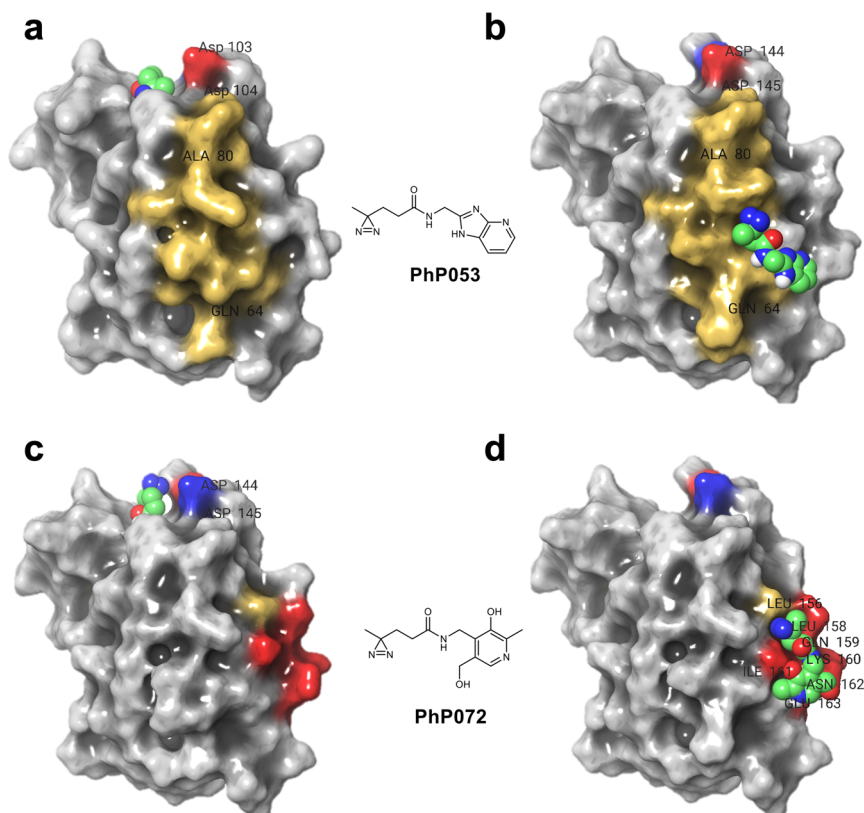
Using a plate-based format with one fragment per well, each fragment was incubated with the target proteins for 15 min to allow for the formation of non-covalent interactions. The plate was then irradiated at 302 or 365 nm for 10 min to enable the bound fragments to crosslink to the protein²⁶. Finally, the results were directly analyzed from the plate using intact protein mass spectrometry by observing the mass additions that correspond to successful labeling (Fig. 2c, for labeling % values see Supplementary data). The library was first characterized against three benchmark proteins like carbonic anhydrase (CA), lysozyme (Lyo), and myoglobin (Myo). These proteins were labeled by 17, 16, and 23 fragments, respectively. This result showed a consistent efficiency of the library compared to the larger set (556 PhaBits) of Grant et al. that showed 47% labeling efficiency for Lyo, 18% for CA and <3% for Myo¹⁷.

Next, we screened the library against three oncology targets with decreasing tractability. Among the therapeutically relevant targets, BRD4-BD1 represents a conventional target having 6000+ ligands available in the ChEMBL database. We considered the oncogenic KRas^{G12D} mutant as a challenging target with a few ligands reported recently in the literature. Finally, we screened the PhP library against the N-terminal domain (NTD) of the transcription factor STAT5B that has no known ligands reported to

date. Hits were detected by two different MS setups: for BRD1-BD4 and KRas^{G12D} we used a high-end LC-MS-TOF spectrometer, while STAT5B-NTD was screened on a readily available standard UHPLC-MS system. Hits were classified by the analysis of the spectra following criteria on the labeling pattern (Fig. 2d).

From all targets, BRD4-BD1 was the most vulnerable (30 fragments labeled >1%) followed by STAT5B-NTD and KRas G12D (26 and 25, respectively). There were 44 fragments in the library that did not label any of the proteins more than 1% and further 19 did not reach 2%. This is in line with Grant et al. reporting that majority of the 556-membered library did not crosslink to any targets. In their investigation 10 PhaBits out of 556 (1.8%) labeled six proteins, while 58 (10.4%) labeled at least 4 targets¹⁷. Among our most privileged fragments, two have labeled four proteins each (PhP018 and PhP060), while four have labeled two targets in higher than 10% amount (PhP038, PhP034, PhP071, PhP087). These might represent privileged scaffolds that could fit into several protein pockets, or be due to nonspecific crosslinking through lipophilic interactions with the protein or formation of long lifetime carbenes¹⁷. Importantly, all proteins had privileged pharmacophores that labeled their targets in a significantly higher amount than others (see Supplementary Table S1 and the corresponding discussion). CA was mostly preferred by PhP003 (4.7%) that slightly labeled the others (<1.1% for each). BRD4-BD1 was mostly preferred by PhP053 (6.7% vs. <2%). In the case of KRas, PhP048 and PhP012 were the most selective compounds (29.4% and 15.8% vs. <3%, respectively). PhP092,

Fig. 3 | Binding sites identified on the BRD4-BD1 surface revealed by screening the PhP library. Experimental (a) and modeled (b–d) binding poses of two verified hits against the primary binding site (left column) vs. other locations (right column) on BRD4-BD1. In the newly identified, surface-exposed sites, **PhP053** binds via multiple H-bond and π - π contacts to the Y65 and K160 sidechains, while **PhP072** occupies a neighboring site by multiple H-bonds towards K160 and E163. Ligands are represented as spheres, labeling sites (as determined by peptide mapping) are colored red (exact site found) or yellow (approximate site found along a longer sequence) on the protein surface. The binding mode of **PhP053** vs. the primary site was solved by X-ray crystallography and has been deposited into the PDB database with the accession code 8Q34.



PhP001 and **PhP088** preferred Lyo (35.6%, 19.2% and 4.8% vs. <1%), while Myo was targeted selectively only by **PhP082** (52.9% vs. <0.5%). STAT5B-NTD was most selectively labeled by **PhP040**, **PhP077**, **PhP065** and **PhP097** (75.0%, 20.0%, 6.5% and 5.7% vs. <3%, respectively).

Usually, >20% labeling efficiencies are accompanied with oxidation, protein degradation or other side reactions. We identified six fragments that showed labeling efficiencies in the 1–13% range consistently over all protein targets (**PhP099**: $5.2 \pm 4\%$, **PhP072**: $4.4 \pm 4.3\%$, **PhP080**: $3.6 \pm 2.2\%$, **PhP037**: $3.4 \pm 3.8\%$, **PhP098**: $3.3 \pm 3.1\%$, **PhP035**: $1.5 \pm 1.2\%$). From this set, **PhP037** and **PhP098** differ only in a substituent on the phenyl ring connected to imidazole and the labeling pattern is very similar e.g., KRas G12D is labeled in 11.2% and 9.6%, BRD4-BD1 in 4.8% and 4.3%, while STAT5B-NTD in 1.4% and 2.3%. **PhP072** might also be highlighted as a consistently labeling fragment with a pyridoxamine core structure. In the following subsections, we present our results on the three oncological targets in more detail.

Fragment hits against BRD4, a tractable target

As a tractable protein target with thousands of reported ligands, we have first screened the PhP library against the first bromodomain of BRD4 (BRD4-BD1). BRD4 is a member of the Bromodomain and Extraterminal protein family, and has important roles in the expression regulation of oncogenes, as well as the maintenance of genome stability²⁷. As such, BRD4 is an important oncotarget for small-molecule intervention, with an accessible and druggable acetyl-lysine (AcK) binding site²⁸.

Intact MS screening against BRD4-BD1 has identified five fragment hits with >1% target labeling, and we could verify the binding sites of two of them by peptide mapping performed by LC-MS/MS following tryptic digestion (Fig. 3). One hit, **PhP053** was successfully co-crystallized with the protein, providing further evidence of binding in the primary (AcK) binding site. Other binding modes, including one in a secondary site for **PhP053**, as well as the AcK site and a third site (residues 156–163) for **PhP072** were modeled with induced fit docking, considering the sites of labeling, as verified by peptide mapping. These efforts revealed that the

PhP library could identify hits against the main binding site of a tractable target, and in addition, some hits could identify yet unreported orthogonal, surface-exposed binding sites, facilitating the design of allosteric ligands.

Fragment binders for KRas^{G12D}, a challenging target

KRas^{G12D} represents a challenging drug target previously considered as non-druggable²⁹. Oncogenic KRas mutants attracted considerable interest in recent years and drug discovery efforts identified multiple fragments and drug-like compounds having affinity against KRas^{G12D}. In fact, several distinct binding sites were successfully targeted by recent efforts³⁰, including the Switch II pocket³¹, the SOS interface^{32,33} and the interface of the Switch I/II regions^{34,35}. To our current knowledge though, there are no reported inhibitors targeting the RBD and dimerization interfaces, or the Switch I region³⁶. Given the fact that KRas has relatively shallow, solvent-exposed pockets, we hoped that any labeling would be an indication of a high-affinity fragment binding site of the target.

Intact MS screening of the PhP library resulted in crosslinking for 11 compounds. Their sites of labeling were investigated by mass spectrometry after tryptic digestion, while the contacting amino acids were identified by HSQC NMR. Combining these orthogonal experiments, we identified the respective binding sites of the three confirmed hits with induced fit docking using experimental information available from MS-based peptide mapping and HSQC NMR. This analysis revealed that PhP fragments have identified multiple allosteric pockets or protein-protein interaction contact surfaces that were not yet targeted by small molecules (Fig. 4).

Following up on the three fragment hits, we have proven by a KRas-SOS exchange assay that **PhP072** fragment binding at the Switch I, II and SOS interfaces influences KRas function (Fig. 4c). Consistently with the generally low affinity of fragment-sized ligand upon non-covalent binding, **PhP072** could restore KRas function only upon irradiation (Fig. 4c, bottom panel). In contrast, **PhP071** and **PhP060** fragments that bind at the RBD and dimerization interfaces have no effect on the nucleotide exchange, before or after irradiation.

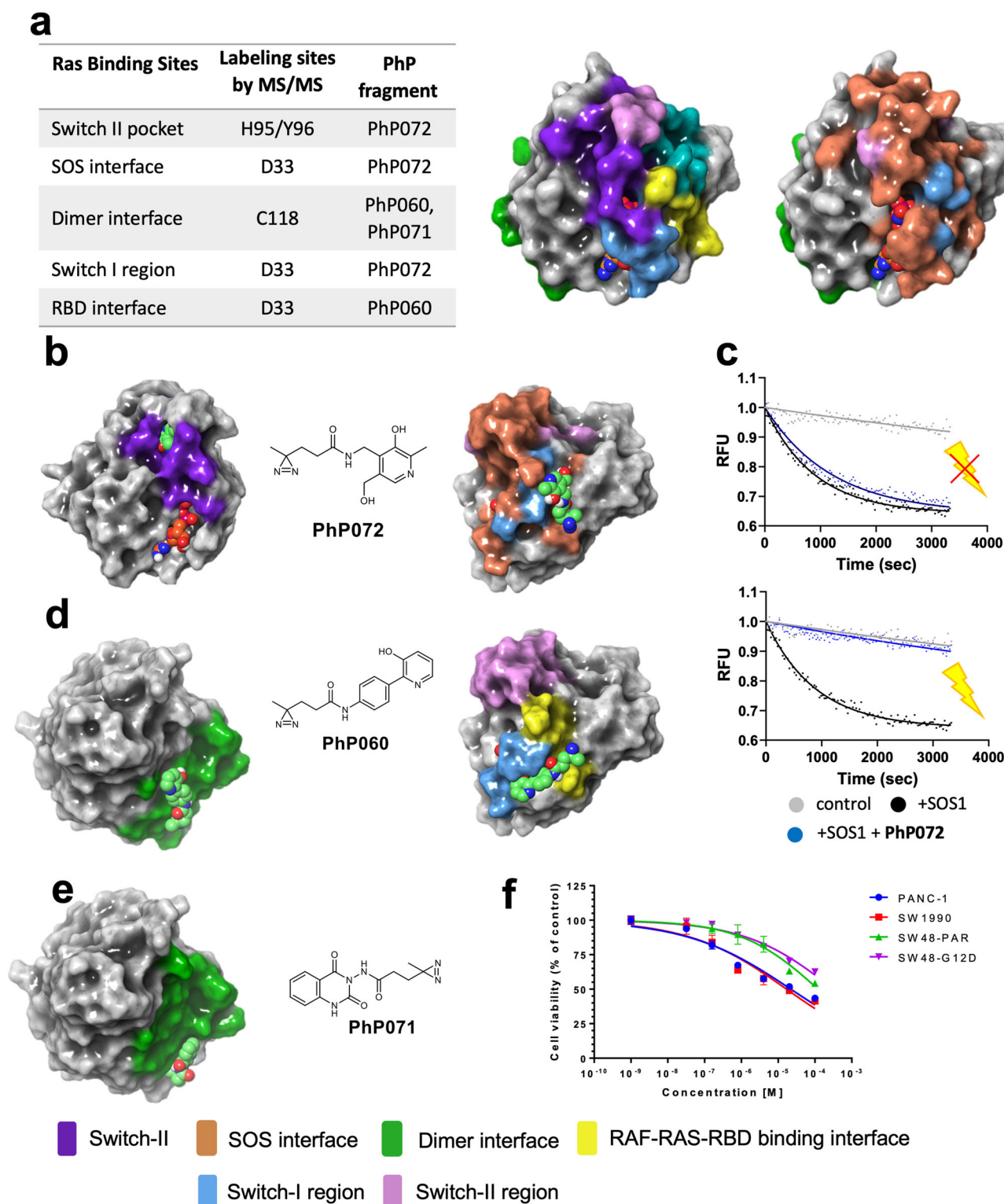


Fig. 4 | Binding sites identified on the KRas^{G12D} surface revealed by screening the PhP library. a Experimentally validated labeling sites of fragment hits on the surface of KRas^{G12D}. **b** PhP072 binds to distinct sites in the Switch-II pocket and SOS interface. PhP072 forms H-bond interactions with D69 and E63 in the Switch-II pocket, and with S39, T35, E37. **c** In the KRas-SOS exchange assay, SOS1 displaces MANT-labeled GDP from KRas (black curve). Addition of PhP072 (blue curve) has little effect prior to irradiation (top), but fully restores KRas function to the control level (gray curve) upon irradiation (bottom). **d** PhP060 binds to distinct sites on the dimerization and RBD interfaces via H-bonds to D33, H27, T35, and a π - π

interaction to H27. **e** PhP071 binds to the dimerization interface via multiple H-bond interactions with R123 and T127. **f** PhP072 shows antiproliferative effect on KRas-dependent pancreatic and colon cancer cells (blue: PANC-1, red: SW1990, green: SW48-PAR, purple: SW48-G12D). Small molecules are shown as spheres: fragments with green carbons, GDP with orange carbons. In (c), each data point and the fitted kinetic curves are shown. In (f), data are presented as mean values $\pm$ SD, calculated from two biologically independent samples, each with three repetitions. Source data are provided in the Source Data file.

Encouraged by the positive labeling and functional results, we have undertaken further efforts to assess the utility of the fragment hits in living cells. The dependency of the applied cell lines for the protein targets was determined by obtaining CRISPR knockout sensitivity data from <https://depmap.org>, and the combined CRISPR sensitivity values (DepMap 22Q2 Public+Score, Chronos) were downloaded (Supplementary Fig. S1). First, we have tested the effect of **PhP072** against a range of KRas-dependent cancer cells including PANC-1 (human pancreatic cancer, KRas^{G12D}), SW1990 (human pancreatic cancer, KRas-G12D), SW48-PAR (human colon cancer; parental cell line KRas^{wt}), and SW48-G12D (human colon cancer; heterozygous knockin of the KRas^{G12D} activating mutation) cancer cell lines. **PhP072** showed antiproliferative activity against all the KRas-dependent cell lines with IC₅₀ values of 17.2 and 22.6 μ M against SW1990 and PANC-1, respectively.

Fragment hits against STAT5B NTD, an unliganded target

Finally, we aimed to identify viable fragment starting points against signal transducer and activator of transcription 5B (STAT5B), an important transcription factor mediating or even driving cancer progression through hyperactivation or gain-of-function mutations³⁷. While there is great interest in its small molecule targeting, no direct STAT inhibitor has reached the market yet. Furthermore, the few ligands of STAT5B reported so far are all targeting its highly conserved SH2 domain³⁸. Therefore, targeting the yet unliganded N-terminal domain (NTD) represents an approach for the direct inhibition of STAT5B function that would open a possibility for pharmacological intervention to cancer progression. Recent works to characterize and target the N-terminal domain of STAT3, including its published X-ray structure³⁹ and a high-throughput virtual screen⁴⁰, provide a solid basis for this effort.

After successful validation against BRD4-BD1 and KRas^{G12D}, we introduced two methodological improvements in the screening process applied against STAT5B-NTD. First, we replaced the high-end LC-MSMS-TOF platform by a readily available UPLC-MS system that might broaden the user community. Second, we attempted to increase the sensitivity of screening by improving the labeling efficiency of PhP fragments. We have noticed that the labeling efficiency of the diazirines is usually around or less than 5% that limits the sensitivity of detecting bound fragments. Therefore, we synthesized an iridium-based photocatalyst (Fig. 5d) and designed a methodology to conjugate the catalyst to the target protein in water using the activating agent BOP, or ((1H-benzotriazol-1-yl)oxy)tris(dimethylamino)phosphonium hexafluorophosphate⁴¹. During the screening of photoreactive fragments, the catalyst can activate the diazirine moiety by Dexter-energy transfer to result in increased labeling efficiency (Fig. 5c)^{41–43}. (To note, the small size of the N-terminal domain and the long-range effect of Dexter-energy transfer allow for the activation of the photoaffinity tag practically anywhere on the STAT5B-NTD surface, see Supplementary Fig. S2.)

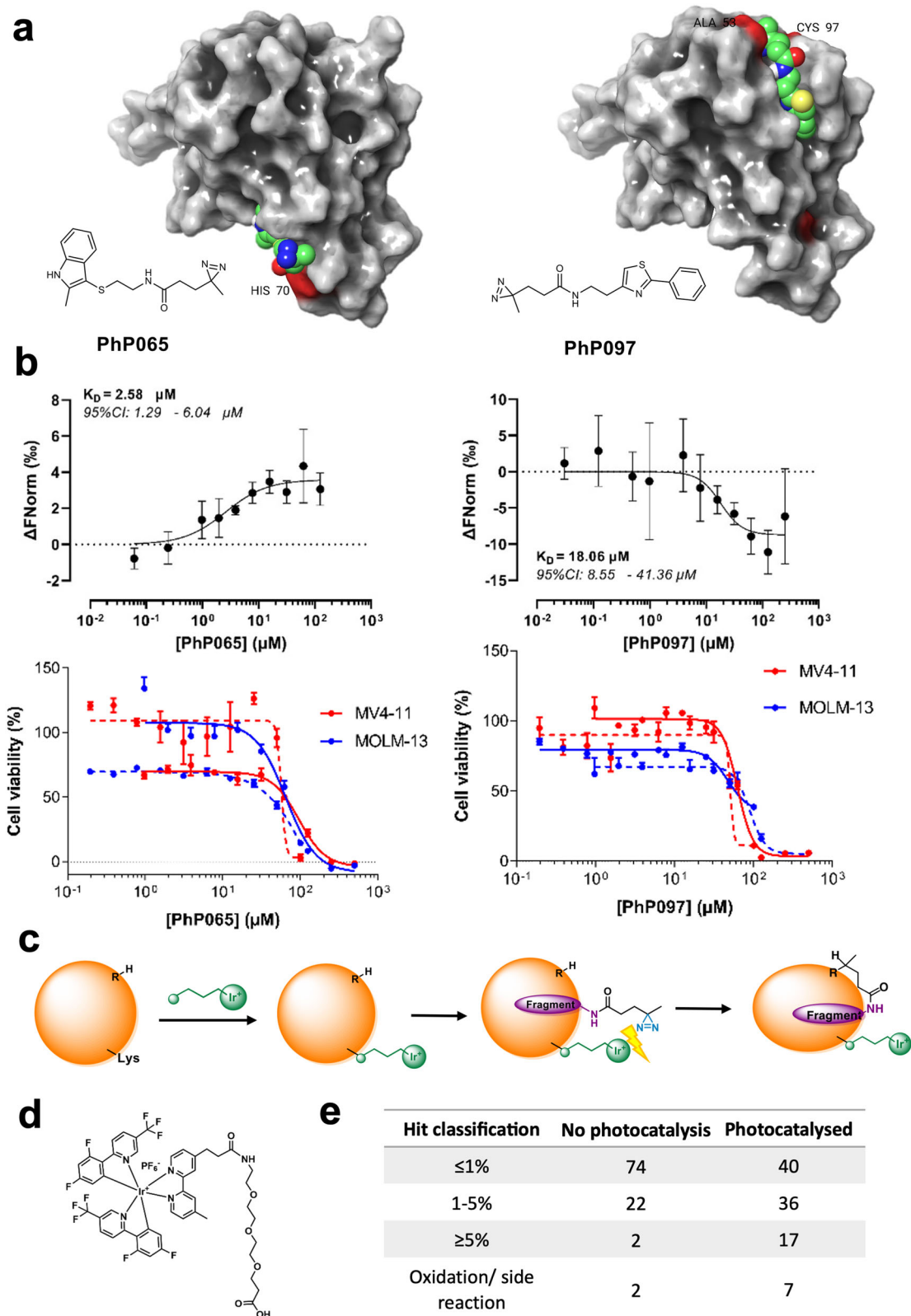
Intact MS screening has identified 24 hits against STAT5B-NTD. As follow-up, the labeling site and efficiency of the hits were analyzed by mass spectrometry, and their prospective binding modes were predicted by ligand docking. Since the labeling sites are close to the two known dimerization interfaces of STAT N-terminal domains, we presume that the hits can inhibit the formation of the “handshake dimer” or the Ni²⁺-mediated tetramer of the N-terminal domain (Fig. 5a)³⁹. This was further verified on the two best fragment hits for the two mentioned sites, **PhP065** and **PhP097** respectively, which provided over 5% labeling and no side reactions, even without the use of the photocatalyst. Here, microscale thermophoresis measurements that revealed direct binding of the two hits to STAT5B-NTD even without irradiation, with K_d values in the low micromolar range (Fig. 5b, black curves). To investigate the translational potential of the resulting fragment hits, we have tested the effects of **PhP065** and **PhP097** on the viability of two established leukemia cell lines, MV4-11 and MOLM-13 (see CRISPR knockout sensitivities in Supplementary Fig. S1). Here, the hit fragments exhibited cellular IC₅₀ values in the mid-micromolar range (Fig. 5b, red and blue curves), further verifying them as viable starting points

for hit elaboration. Finally, we have shown significantly increased labeling efficiencies and more primary hits in case of the photocatalyzed intact MS screening of PhP-fragments, confirming the usefulness of the improved protocol (Fig. 5e).

Discussion

We introduced a fragment screening concept using pharmacophore-optimized photoaffinity fragments (PhPs), by combining the advantages of two recent approaches reported by our groups^{17,19}. The advantages of PhP are apparent in comparison to SpotXplorer, as the PhP library readily offers the capacity to experimentally detect the relevant binding sites of the discovered hits, in a platform that is more accessible to a wider community. By comparison, in validating the SpotXplorer technology against targets that were new at the time, specifically the SARS-CoV-2 viral targets 3CLPro and NSP3, we needed a state-of-the-art synchrotron facility to obtain high-resolution crystal structures of the confirmed hits. Additionally, our current results have illustrated that there are distinct cases regarding the available binding site(s). In some cases, we can infer analogies based on previously reported ligands (BRD4-BD1). However, for the more elusive target KRas^{G12D}, there are numerous binding sites available, so the direct experimental mapping provides invaluable and readily available guidance for hit-to-lead follow-up at an early stage. Finally, for the unexplored STAT5B-NTD target, we have gathered first evidence for targeting the dimerization interface, as verified by MS screening. We should highlight that the orthogonal, secondary binding assays (X-ray crystallography for BRD4-BD1, HSQC NMR for KRas^{G12D} and MST for STAT5B-NTD) were carried out without irradiation, verifying the affinity of the fragments towards the respective targets via a non-covalent mechanism of action. Additionally, the detected affinities against KRas^{G12D} and STAT5B-NTD (IC₅₀ and K_d values in the 2–20 μ M range) exceeded our expectations toward fragment hits. These promising on-target affinities prompted us to test the best fragments against KRas^{G12D} and STAT5B-dependent cell lines, respectively. Although their observed cellular activity supports that the fragments interact with the target in cells, but further studies will be required to ascribes its functional effects to on-target binding given their promiscuity. In fact, during the revision process, Offensperger et al. published a large-scale chemoproteomics study, using 407 diazirine-tagged fragments⁴⁴. From this set, we identified compounds with notable structural similarities to our fragment hits (Supplementary Table S2) that labeled 10–300 proteins in cells.

In comparison to the original PhABits screening platform, the main advantages of the PhP library are the higher hit rates and, by design, its better coverage of the unique fragment-binding pharmacophore patterns. For the highly tractable BRD4-BD1 target, both libraries provided fragment hits against its primary binding site (nine primary hits for PhABits and five for PhP). Notably though, the smaller number of hits provided about the same coverage of the pharmacophore space of the previously reported, fragment-sized BRD4-BD1 inhibitors (58 fragments downloaded from the ChEMBL database^{45,46} with 10–16 heavy atoms and a pChEMBL value ≥ 3 , representing millimolar or better activity). Specifically, 63 and 59% of the unique pharmacophores represented by at least one literature inhibitor were recovered by the nine PhABit and five PhP hits, respectively (vs. a priori expectations of 16 and 9%, assuming an even distribution of pharmacophores). The smaller number of PhP hits can be rationalized by their optimized pharmacophore coverage as compared to PhABits. The similar coverage of known pharmacophores, however, was achieved by significantly fewer PhP fragments as PhABits (100 vs. 556 compounds). Comparing the screening efficiency of PhABit and PhP libraries revealed three times higher hit rate for the latter platform (1.6% vs 5%, respectively). In addition, each of the two hits that were further verified by peptide mapping, **PhP053** and **PhP072**, have labeled an alternative binding site in addition to the main one. The difference is even more apparent for KRas^{G12D}, where the three verified PhP hits have labeled five different binding sites, in comparison to one site labeled by the four hits resulting from the larger PhABits library. This is a direct consequence of the SpotXplorer design, that optimized the library



towards the best possible coverage of the diverse pharmacophore patterns observed in experimental protein-fragment complexes. Furthermore, screening the PhP library provided a more than seven times higher hit rate (5% vs. 0.7%) against this challenging target. Therefore, the PhP hits explored a higher diversity of potential binding sites with higher screening efficiency as compared to the PhABit hits.

Finally, we highlight two important methodological developments. First, we adapted the screening technology to conventional LC-MS platforms, in line with recent efforts to democratize MS-based fragment screening approaches for academic labs^{47–49}. Second, we improved its sensitivity by a recently reported bioconjugated photocatalyst that improves the efficiency of photoaffinity labeling. This approach can eliminate the

Fig. 5 | Results of screening the PhP library against the STAT5B N-terminal domain. **a** Structures and predicted binding modes of **PhP065** and **PhP097** against the dimerization interfaces of the STAT5B N-terminal domain. **PhP065** binds close to the cavity formed in the handshake dimer interface by H-bonds with the Q36 and K70 residues. **PhP097** binds at the tip of the $\alpha 6$ and $\alpha 7$ helices next to the Ni^{2+} -mediated tetramer interface with an H-bond to the T58 sidechain and further hydrophobic contacts. **b** In microscale thermophoresis, **PhP065** and **PhP097** bind directly to STAT5B-NTD, with K_d values of 2.58 and 18.06 μM , respectively (black curves). Furthermore, the fragments inhibit the viability of the established leukemia cell lines MV4-11 (red) and MOLM-13 (blue) with IC_{50} values of 76.15 and 100.1 μM (**PhP065**), and 56.62 and 68.22 μM (**PhP097**), respectively. **c** In our

bottleneck of limited labeling efficiency in screening low-affinity fragments, and to our knowledge is the first successful attempt at attaching this photoaffinity tag to a full protein. Electrophiles targeting a specific sidechain, such as cysteine, typically have a high labeling efficiency, but a limited proteome coverage, as the targeted sidechain must be present in the binding site. By comparison, photoaffinity tags offer a far superior proteome coverage by design, but this comes at a cost of lower modification yields, cf. our classification of primary hits with >5% labeling rates (Fig. 2d). Ultimately, this is a key factor regarding the detection sensitivity of the platform. One possibility for improving modification yields is using different photoreactive functionalities. However, a higher reactivity might come at a cost in other properties, such as a lower aqueous stability⁵⁰. In the present work, we have followed a different train of thought by boosting the labeling efficiency of our photoaffinity-tagged fragments with a bioconjugated iridium-based photocatalyst. This approach doubled the number of hits identified against the previously unliganded and challenging transcription factor target, STAT5B-NTD (Fig. 5c). It should be noted that in addition to the much-improved labeling efficiency we observed slightly higher number of side reactions such as oxidation or water elimination. These reactions might limit the use of certain moieties; however, the covered pharmacophore information can be preserved by choosing more photoresistant scaffolds. These developments in the screening technology improved the detection sensitivity of the platform especially on low-end LC-MS platforms.

In conclusion, combining the advantages of two fragment screening platforms developed at our labs we introduced pharmacophore-optimized photoaffinity fragments that provides higher hit rates and better exploration of available protein binding sites even for challenging or unliganded targets. Improving the sensitivity of hit detection by a bioconjugated photocatalyst and transforming the screening platform to readily available LC-MS systems, we widened the scope of potential users. Our work can also be considered as an example of academia-industry collaborations that facilitate drug discovery efforts and therapeutic innovations via new tools available for the wider community.

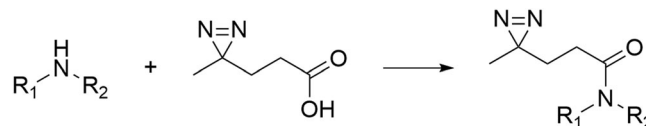
Methods

Library design

The library was designed as reported in our recent work¹⁹. Briefly, pharmacophore models were extracted from PDB structures with the ePharmacophore module of the Schrödinger software suite (version 2017-4), and clustered with scipy (version 1.0.1). Ligand preparation, conformer generation and pharmacophore screening were carried out with the Epik, MacroModel and Phase modules of the Schrödinger software suite, respectively (version 2022-4).

Fragments in the Enamine primary amine collection were filtered for size (10–16 heavy atoms), primary amine group count of exactly one, and the absence of PAINS moieties²⁵. Candidate fragments were screened against, and annotated with the full set of 2-, and 3-point pharmacophore models and stored as fingerprints with bits for pharmacophores present set to 1. We have used a diversity picker to select 160 fragments with maximized coverage and diversity across the 117 bit positions (pharmacophores). The final set of successfully synthesized 100 PhP fragments provided coverages of 88% and 75% for the 2- and 3-point pharmacophores, respectively. SMILES codes, images and

modified workflow, the iridium-based photocatalyst Ir-G2-PEG3-COOH is first attached to the target protein. After initial non-covalent binding, the crosslinking of the photoaffinity-labeled fragments is enhanced by Dexter-energy transfer from the photocatalyst. **d** Structure of the photocatalyst Ir-G2-PEG3-COOH⁴¹. **e** Application of the photocatalyst results in higher labeling efficiency and, by transition, a higher number of primary hits (>1% labeling). In the K_d plots, data are presented as mean values $\pm$ SD, calculated from two biologically independent samples, each with three repetitions. In the cellular IC_{50} plots, data are presented as mean values $\pm$ SEM, separately for two biologically independent samples, each with three repetitions. Source data are provided in the Source Data file.



Scheme 1 | General procedure for attaching the diazirine tag onto amine-containing fragments. Amide coupling with the HATU (Hexafluorophosphate Azabenzotriazole Tetramethyl Uronium) coupling reagent was utilized to award the photoaffinity-tagged fragments of the PhP library.

pharmacophore fingerprints of the PhP fragments are included in Supplementary Data 1.

General procedures for synthesis and compound characterization

PhP fragments were synthesized according to the general procedure outlined in Scheme 1. 3-(3-Methyl-3H-diazirin-3-yl)propanoic acid (0.025 g, 0.195 mmol) was dissolved in *N,N*-dimethylformamide (DMF), 1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate (HATU, 0.071 g, 0.188 mmol) and DIPEA (0.105 mL, 0.600 mmol) were added. The reaction mixture was stirred for 15 min at room temperature. 0.15 mmol amine was added, then the mixture was stirred at RT for 16–72 h covered in tin foil, and progress was assessed by LCMS.

After 16–72 h the reactions were processed as follows:

- DMF in the reaction mixtures was removed in a Genevac. The crude product was dissolved in chloroform (0.4 mL) and loaded onto a 1.0 g NH_2 -Isolute SPE column, pre-equilibrated with chloroform for 2 column volumes (CV's). The product was eluted off the column using 2.5 mL of 10% methanol in ethyl acetate. This was repeated with another 2.5 mL of 10% methanol in ethyl acetate. The eluent was collected for each reaction in pre-weighed T-vials and dried under a stream of nitrogen in the Radleys blowdown apparatus to obtain a product.
- The reaction mixture was loaded to a prep-HPLC column and purified with a linear gradient from 5 to 100 vol% MeCN in water containing 0.1 vol% formic acid over 15 min. The solvent was then evaporated.

^1H NMR was recorded in $\text{DMSO}-d_6$, CDCl_3 , CD_3CN or D_2O solution at room temperature, on a Varian Unity Inova 500 (Varian, Palo Alto, CA, USA) (500 MHz) and on a Varian 300 spectrometer (300 MHz), with the deuterium signal of the solvent as the lock. Chemical shifts (δ) and coupling constants (J) are given in ppm and Hz, respectively, and the spectra are collated in Supplementary Data 2.

HPLC-MS measurements were performed using a Shimadzu LC-MS-2020 device equipped with a Reprospher-100 C18 (5 μm ; 100 $\times$ 3 mm) column and positive-negative double ion source (DUIS $\pm$) with a quadrupole MS analyzer in a range of m/z 50–1000. Samples were analyzed with gradient elution using eluent A (0.1% formic acid in water) and eluent B (0.1% formic acid in acetonitrile). Flow rate was set to 1 mL/min. The initial condition was 5% B eluent, followed by a linear gradient to 100% B eluent by 1 min, from 1 to 3.5 min 100% B eluent was retained; and from 3.5 to 4.5 min back to initial

condition with 5% B eluent and retained to 5 min. The column temperature was kept at room temperature and the injection volume was 1–10 μL . Purity of compounds was assessed by HPLC with UV detection at 254 nm; all tested compounds were >95% pure.

High-resolution mass spectrometric measurements were performed using a Q-TOF Premier mass spectrometer (Waters Corporation, Milford, MA, USA) in positive or negative electrospray ionization mode. Reactions were monitored with Merck silica gel 60 F²⁵⁴ TLC plates (Darmstadt, Germany). All chemicals and solvents were used as purchased from commercial suppliers. The column chromatography purifications were performed using Teledyne ISCO CombiFlash Lumen+ R_f. For buffer media exchange, a GE Healthcare PD SpinTrap™ G-25 desalting column was used.

Detailed analytical results are included in Supplementary Note 1 for each compound.

Sample preparation for STAT5B-NTD illumination and photocatalysis experiments

STAT5B protein in 7 μM concentration in pH 7.4 PBS buffer was pre-incubated with 0.17 μL PhP (in 100 mM DMSO solution) for 1 h at room temperature in dark. Irradiation was carried out at 4 °C for 10 min using 365 nm UV lamp.

The synthesis of short- and long-linker Ir-G2 photocatalysts was started from dMebppy⁴¹. First, we synthesized Gen 2 Iridium catalyst (cpd. 13 in ref. 41) that was not effective for labeling STAT5B-NTD, presumably due to the short propionic acid linker. However, the carboxylic acid analog with a longer PEG3 linker (Ir-G2-PEG3-COOH, or cpd. 17 in ref. 41) successfully labeled the protein target in the presence of BOP (but not of other activating agents like HATU, HCTU, TSTU or PyAOP) with a 27% yield (Supplementary Fig. S3).

For the photocatalysis experiments, 1.4 μL photocatalyst (in 1.4 mM DMSO solution, 10 eq) was pre-incubated with 1.5 μL coupling reagent (0.75 μL , 5.6 mM EDC (20 eq) + 0.75 μL 4.2 mM NHS and 1.5 μL 2.1 mM (15 eq) BOP for 30 min in dark. 30 μL STAT5B-NTD (in 7 μM concentration) was added and incubated further at room temperature overnight or at 37 °C for 1 h in dark. After addition of 0.14 μL PhP (in 100 mM DMSO) the solution was further incubated in dark for 1 h. Irradiation was carried out using 450 nm LED lamp (7.6Vx0A) for 10 min at 4 °C.

Intact mass spectrometry (MS) screening

The PhP library was screened against BRD4-BD1 and KRas^{G12D} as follows. Using a Labcyte Echo® 555 Liquid Handler, 150 nL of fragment solution (20 mM) was transferred into a Greiner 384 low-volume plate (white) to prepare the probe plate. The probe plate was placed on ice. 15 μL of protein stock solution (1 μM BRD4-BD1 (GSK/GenScript) or KRas^{G12D} (GSK/GenScript) in PBS buffer) was dispensed into wells containing fragments. The plate was left on ice for 15 min for incubation, then irradiated at 302 nm for 10 min, and centrifuged (1000 rpm, 1 min) to remove any bubbles. The plate was then analyzed by LCMS-TOF mass spectrometry (Agilent 1200 series liquid chromatography with Agilent Bio-HPLC PLRP-S (1000 Å, 5 μm × 50 mm × 1.0 mm, PL1312-1502) reverse phase HPLC column at 70 °C equipped with an Agilent G6224 time-of-flight (ToF), see Supplementary Note 2 for the full protocols). The deconvoluted spectra were analyzed using R Studio software. Spectra of the hits are reported in Supplementary Note 2.

For screening against STAT5B-NTD, we have used a UHPLC-MS system that consisted of a Waters ACQUITY UPLC I-Class setup coupled with a Waters ACQUITY UPLC Peptide BEH C18 Column (130 Å, 1.7 μm , 2.1 mm × 100 mm), connected to a Waters Xevo G2-XS QT-ToF instrument equipped with a Waters Z-spray ESI source. During the analysis, the column temperature was maintained at a constant 60 °C, and a sample volume of 3 μL was injected for each analysis. Data acquisition was conducted in positive ion mode within the m/z 100–2000 (mass-to-charge ratio) range. For the full protocol and spectra of the hits, see Supplementary Note 2.

Binding site identification by LC-MS/MS peptide mapping—BRD4-BD1 and STAT5B-NTD

In follow-up of the intact MS screening, the fragment hits were further analyzed by a Triple TOF 5600+ hybrid Quadrupole-TOF LC/MS/MS system, after digesting the resulting fragment-protein complexes by Trypsin/Lys C mix. Data acquisition and processing were performed using Analyst TF software version 1.7.1 (AB Sciex Instruments, CA, USA). Chromatographic separation was achieved on the Discovery® BIO Wide Pore C-18-5 (250 mm × 2.1 mm, 5 μm , 300 Å) HPLC column. MS/MS spectra were obtained on the 8 most abundant parent ions present in the TOF survey scan with the Information Dependent Acquisition (IDA) mode, and peaks were evaluated with PeakView (version 2.2, Sciex) and Biologics Explorer (version 3.0.3, Sciex). The full sample preparation and data acquisition protocols, as well as the spectra of the reported hits, are available in Supplementary Note 3.

Modeling the binding poses of fragment hits by docking

Docking calculations were performed on BRD4-BD1 (PDB ID: 7A9U¹⁷), KRas^{G12D} (PDB ID: 4OBE⁵¹) and STAT5B-NTD with the same protocol. Briefly, the relevant X-ray structures were downloaded from PDB and were prepared using Protein Preparation Wizard⁵². The structure of STAT5B-NTD was homology modeled based on the published structures of STAT3-NTD (PDB ID: 4ZIA³⁹). Ligands were prepared with Ligprep⁵², and docking was performed with the Induced Fit Docking protocol of Schrödinger. For the grid box generation, the experimental results (MS labeling data and NMR shift perturbations, where available) were used. At most 20 possible binding conformations were generated in the first docking step⁵³. Redocking was done into structures within a 30 kcal/mol energy window from the best structure, and within the top 20 structures overall, using the single precision (SP) method.

Cell viability measurements

Pancreatic cancer cell lines PANC-1 (cat. no. CRL-1469), SW1990 (cat. no. CRL-2172) were obtained from ATCC (American Type Culture Collection). Genetically modified isogenic colon cancer cell lines SW48-PAR and SW48-G12D were obtained from Horizon Discovery Ltd. PANC-1 (human pancreatic cancer, KRas^{G12D}), SW1990 (human pancreatic cancer, KRas^{G12D}), SW48-PAR (human colon cancer; parental cell line KRas^{wt}), and SW48-G12D (human colon cancer; heterozygous knockin of the KRas^{G12D} activating mutation) cancer cell lines were cultured in Roswell Park Memorial Institute Medium (RPMI; Biosera, Nuaille, France), supplemented with 10% heat-inactivated Fetal Bovine Serum (FBS; Biosera), and with 1% Penicillin/Streptomycin (Biosera). Cells were cultured in sterile T75 flasks with ventilation cap (Sarstedt, Nümbrecht, Germany) at 37 °C in a humidified atmosphere with 5% CO₂ in ESCO CelCulture Incubator (ESCO, Friedberg, Germany). Manipulations with the cells were performed in biosafety cabinet (laminar) ESCO Sentinel Gold class II model AC2-4E8 (ESCO).

For the evaluation of the in vitro antiproliferative activity of fragments, cell viability was determined with the MTT assay (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazolium bromide from Sigma Aldrich). After standard harvesting of the cells by trypsin-EDTA (Biosera), 7 × 10³ cells per well depending on the cell line, were seeded in serum-containing growth medium to 96-well plates and incubated. After 24 h, cells were treated with various concentrations of fragments (32 nM–100 μM), dissolved in serum-containing medium, and incubated under standard conditions. Control wells were treated with medium. Final concentration of serum was 2.5%, final concentration of DMSO was 0.2%. Treatment was for 72 h continuously.

Afterwards, MTT assay was performed in order to determine cell viability, by adding 20 μL of MTT solution (5 mg/mL in PBS) to each well and after 2 h of incubation at 37 °C, the supernatant was removed. The formazan crystals were dissolved in 100 μL of a 1:1 solution of DMSO (Sigma-Aldrich):EtOH (Molar Chemicals) and the absorbance was measured after 15 min at λ = 570 nm by using a microplate reader (BioTek

Table 1 | Refinement statistics for the X-ray structure 8Q34

PDB ID	8Q34 [*]
Data collection	
Space group	P1
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	37.24 44.09 78.48
<i>α</i> , <i>β</i> , <i>γ</i> (°)	90.03 90.00 90.03
Resolution (Å)	78.48–1.48 (1.51–1.48)
<i>R</i> _{merge}	0.14 (1.31)
<i>I</i> / <i>σI</i>	5.7 (0.8)
Completeness (%)	90.9 (51.3)
Redundancy	3.4 (3.0)
Refinement	
Resolution (Å)	1.48
No. reflections	259312
<i>R</i> _{work} / <i>R</i> _{free}	0.184/0.216
No. atoms	9429
Protein	8457
Ligand	168
Water	804
<i>B</i>-factors	
Protein	16
Ligand	14
Water	26.41
R.m.s. deviations	
Bond lengths (Å)	0.0068
Bond angles (°)	2.905

^{*}Diffraction data from single crystal was used to determine the structure. Values in parentheses are for highest-resolution shell. Section headings are highlighted in bold.

800TS, Agilent, Santa Clara, CA, USA). The IC₅₀ values of the fragments were calculated using GraphPad Prism 6 (GraphPad Software, San Diego, CA, USA). The experiments were done in triplicate, and each experiment was repeated two times.

BRD4-BD1 expression, purification, crystallization, and structure

BRD4 (residues 44–171, cloned into pNIC28-Bsa4 using LIC cloning, SGC ID: BRD4A-c001) was produced from *E. coli* (BL21) RR in TB media (Formedia) with expression induced with 1 mM IPTG when OD₆₀₀ = 0.6 was reached at 18 °C overnight. The cells were lysed (50 mM HEPES, 500 mM NaCl, 5% glycerol) by sonication, purified by IMAC using Talon resin (50 mM HEPES, 500 mM NaCl, 5% glycerol, 300 mM imidazole), tag removed with TEV cleavage overnight at 4 °C and polished using gel filtration with Superdex 75 column. Purified protein in (50 mM HEPES, 500 mM NaCl, 5% glycerol) was concentrated to 5.6 mg/mL and crystallized in 150 nL drop in 1:2 protein: reservoir solution ratio in 20% PEG6000, 10% ethylene glycol, 0.1 M HEPES pH 7.0 and 0.2 M sodium chloride at 277 K using sitting drop vapor diffusion. Crystals formed after 28 days. **PhP053** (100 mM, dissolved in MeOH, not DMSO) was diluted 10 times with reservoir solution added to the crystals and incubated for a further 1 h at 4 °C. Diffraction data were collected at i04 at Diamond Light source as part of BAG allocation mx28172, autoprocessed using the autoPROC pipeline⁵⁴ and phased using BUSTER⁵⁵ using 4MEN. Manual model rebuilding was done using the CCP4 cloud⁵⁶ alternated with structure refinement was performed in COOT⁵⁷ and REFMAC5⁵⁸. The resulting structure has been deposited in the PDB database (<https://rcsb.org>), with the accession code **8Q34**. Refinement statistics are provided in Table 1, as well as in Supplementary Data 1.

Binding site identification of KRas^{G12D} hits by LC-MS/MS peptide mapping

Modification sites were determined by proteolysis and reversed-phase LC-MS peptide mapping. Briefly, proteins were enzymatically digested in 25 mM NH₄HCO₃ solution after buffer exchange using Amicon Ultra-0.5 mL Centrifugal Filter units (10 kDa, Merck Millipore). Trypsin-LysC mixture and ProAlanaase (Promega Corporation, Madison, USA) were used for the enzymatic digestion. Protein samples were reduced by dithiothreitol at 37 °C for 30 min. After reduction, proteins were digested using 1:50 enzyme:protein ratio or 4 h at 37 °C, followed by an additional short incubation with dithiothreitol (5 min, 37 °C). Overnight digestion was performed by Trypsin-LysC mixture in 50 mM NH₄HCO₃ solution at 37 °C. Tryptic digestion was stopped by adding formic acid in a final concentration of 0.2% (V/V). ProAlanaase digestion was performed for 4 h at 37 °C in 50 mM HCl and was stopped by heating at 90 °C for 10 min. After digestions an additional short incubation with dithiothreitol was repeated (5 min, 37 °C).

Mass spectrometric experiments were performed on a high-resolution hybrid quadrupole-time-of-flight mass spectrometer equipped with a cyclic ion mobility separator (Waters Select Series Cyclic IMS, Waters Corp., Wilmslow, U.K.). Chromatographic separation was performed using a Waters Acquity I-Class UPLC system, coupled directly to the mass spectrometer. Waters Acquity CSH Peptide C18 UPLC column (2.1 × 150 mm, 1.7 μm) was used for chromatography. Gradient elution was performed under the following parameters: eluent A: 0.1% formic acid in water, eluent B: 0.1% formic acid in acetonitrile; flow rate: 300 μL/min; column temperature: 60 °C; gradient: 2 min: 2%B, 80 min: 45%B, 81 min: 85%B. HDMS^E experiments were performed using a single-pass cyclic ion mobility separation and fragmentation in the transfer cell with collision voltage ramping. MS data acquisition was performed with the following parameters: *m/z* 50–2000, V-mode, scan time: 0.3 s, single Lock Mass: leucine enkephalin; low energy: 6 V, high energy: ramping 19–45 V. BiopharmaLynx 1.3.5 software (Waters Corp., Wilmslow, U.K.) was used to for data analysis. Spectra of the hits are reported in Supplementary Note 3.

Binding site identification of KRas^{G12D} hits by NMR spectroscopy

NMR samples contained 70 μM ¹⁵N-labeled GDP-bound KRas^{G12D} protein, 100–600 μM binding partner, 10 mM MgCl₂, 3 mM NaN₃, in PBS buffer, 5% DMSO-d₆, 7% D₂O, 0.5% DSS and the pH was set to 7.4. ¹H,¹⁵N-SOFAST-HMQC (fast version of ¹H,¹⁵N-HSQC) NMR spectra were acquired at 298 K on a Bruker AVANCE III spectrometer (Bruker Biospin, Rheinstetten, Germany) operating at 700.05 MHz for ¹H and 70.94 MHz for ¹⁵N, equipped with a 5 mm Prodigy TCI H&F-C/N-D, z-gradient probe head. Temperature was calibrated by standard methanol solution. The chemical shifts were referenced with respect to the ¹H-resonance of an internal DSS standard, while ¹⁵N chemical shifts were referenced indirectly via the gyromagnetic ratios according to the IUPAC conventions. All NMR data were processed with Bruker TOPSPIN 3.6 and analyzed in POKY software⁵⁹. The shifted crosspeaks were compared to the free KRas^{G12D} chemical shifts⁶⁰. Spectra of the hits are reported in Supplementary Note 4.

Kras-SOS exchange assay

MANT-GDP loading assay: Kras^{G12D} protein was first buffer exchanged into low magnesium buffer (20 mM HEPES-NaOH (pH 7.5), 50 mM NaCl, 0.5 mM MgCl₂) using a NAP5 column (catalog no.: 17-0583-1, Cytiva). The proteins were then incubated with 20-fold molar excess of *N*-methylanthraniloyl (MANT)-GDP (catalog no.: 69244, Sigma-Aldrich) in loading buffer (50 mM NaCl, 20 mM HEPES-NaOH [pH 7.5], 0.5 mM MgCl₂, 10 mM EDTA, and 1 mM Dithiothreitol (DTT)) in a total volume of 200 μL at 20 °C for 90 min. The reaction was stopped by adding MgCl₂ to a final concentration of 10 mM, then incubated at 20 °C for 30 min. The unbound MANT-GDP was removed using the NAP-5 column equilibrated with nucleotide exchange buffer (40 mM HEPES-NaOH (pH 7.5), 50 mM NaCl, 10 mM MgCl₂, 2 mM DTT).

MANT-GDP exchange assay: First the MANT-GDP bound Kras G12D protein (in a final 1 μ M molar concentration) was preincubated with the inhibitor molecules for 60 min. After the preincubation the samples were exposed to UV light for 10 min with a wavelength of 366 nm. The UV-treated MANT-GDP Kras-inhibitor mix was loaded into a black 384 well microplate. The nucleotide exchange reaction was initiated by adding 100 fold molar excess of GppNHp (catalog no.: G0635, Sigma-Aldrich), a non-hydrolyzable GTP analog and the SOS1 exchange domain (catalog no.: GE02, Cytoskeleton, Inc.) protein in 0.5 μ M final concentration. The change in fluorescence intensity was measured every 30 s in room temperature for 60 min on an EnSpire plate reader (PerkinElmer, Inc.). The measured fluorescence values were fitted to a single exponential function by using GraphPad Prism 8 software.

STAT5B-NTD expression, purification, and MST measurements

Recombinant protein expression was performed similarly to our recently published protocol for expressing full-length STAT5B⁶¹. STAT5B N-terminal domain (1-123, NCBI Accession Number NP_036580.2) was codon optimized and synthesized by Genscript and cloned into pET28b+ plasmid using NheI and XhoI cloning sites with an N-terminal His-SUMO tag. The plasmid was used to transform BL21 RILP cells and single colonies were selected and used to inoculate 3 mL cultures in Super Broth (with 34 μ g/mL chloramphenicol and 50 μ g/mL kanamycin). Once cultures reached an OD₆₀₀ of ~1.0, they were transferred to 1 L Super Broth (supplemented with 10 mM MgSO₄, 0.1% [v/v] glucose, 34 μ g/mL chloramphenicol and 50 μ g/mL kanamycin). At OD₆₀₀ = 1.5, the temperature was reduced to 16 °C and the media was supplemented with 1 mM IPTG and 3% (v/v) ethanol. The cells were harvested following 20 h induction and frozen at -80 °C.

For protein purification, the cell pellets were thawed in lysis buffer and ruptured through sonication. The cell lysate was cleared by centrifugation and loaded onto a 3 mL Ni²⁺-NTA resin (GE Healthcare). The column was washed and the protein was eluted and loaded onto a Superdex 200 Increase 10/300 GL column (GE Healthcare). The fractions were treated with His-Ulp1 protease to cleave the His-SUMO tag and passed through a 1 mL Ni²⁺-NTA resin to remove any residual tag. The flow through was collected and assessed for purity via SDS-PAGE and the protein was dialyzed into 100 mM HEPES pH 7.4, 2% (v/v) glycerol. Protein concentration was determined by BCA assay (Thermo Fisher Scientific) and aliquots of N-terminal domain were flash-frozen in liquid nitrogen and stored at -80 °C. The compositions of all purification buffers are listed in additional detail in ref. 61.

For the microscale thermophoresis (MST) studies, we prepared 16 two-fold serial dilutions of compounds starting from 500 μ M. Titration series were prepared that contained 10 μ L of compounds' solutions of varying concentrations and 10 μ L RED-NHS 2nd Generation labeled STAT5B NTD with a concentration of 168 nM for compound **PhP097**, and a concentration of 84 nM for compound **PhP065**. Final buffer composition included 1X PBS containing 0.5% DMSO. All measurements were taken in Premium Coated Capillaries on a Monolith NT.115 instrument (Nano-Temper Technologies, Munich, Germany) using 80% infrared laser power for compound **PhP097**, 60% infrared laser power for compound **PhP065**, and an LED excitation source with λ = 650 nm at a temperature of 25 °C. Results were expressed as the mean of two separate experiments, with three technical replicates each. GraphPad Prism 9.5.1 software was used to fit the data and to determine the K_D values.

MV4-11 and MOLM-13 cell lines were purchased from DSMZ (Braunschweig, Germany), and grown at 37 °C and 5% CO₂ in RPMI 1640 medium (Gibco™, Thermo Fisher Scientific). Media were supplemented with 10% fetal calf serum (FCS), 10 U/mL penicillin, 10 μ g/mL streptomycin and 2 mM L-glutamine (all Gibco™, Thermo Fisher Scientific).

To determine the IC₅₀ of the selected compounds on the cell lines, CellTiter-Blue® cell viability assay (Promega) was performed. For this, cells were seeded in 96-well flat bottom plates at a cell density of 10,000 cells/well. Cells were treated in triplicates with the compound of interest at various concentrations or with 10 μ M Bortezomib (S1013; Selleck Chemicals,

Houston, TX, USA) as a positive control. Cell viability of treated cell lines was measured using CellTiter-Blue® after 72 h incubation. Plates were measured using a GloMax® plate reader (Promega) and IC₅₀ values were determined by non-linear regression using the GraphPad Prism 9.1.1 (GraphPad Software, Inc.) and the data are reported as mean values $\pm$ SEM.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Structure data that support the findings of this study have been deposited in the PDB database (<https://rcsb.org>), with the accession code 8Q34. Data generated during the computational and experimental screening of the described fragment library are reported in Supplementary Data 1. BRD4-BD1 ligands and bioactivity data were downloaded from the ChEMBL database (<https://www.ebi.ac.uk/chembl/>). Source data are provided with this paper.

Received: 1 June 2024; Accepted: 23 July 2024;

Published online: 31 July 2024

References

- Murray, C. W. & Rees, D. C. The rise of fragment-based drug discovery. *Nat. Chem.* **1**, 187–192 (2009).
- Bollag, G. et al. Vemurafenib: the first drug approved for BRAF-mutant cancer. *Nat. Rev. Drug Discov.* **11**, 873–886 (2012).
- Souers, A. J. et al. ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nat. Med.* **19**, 202–208 (2013).
- Tap, W. D. et al. Structure-guided blockade of CSF1R kinase in tenosynovial giant-cell tumor. *N. Engl. J. Med.* **373**, 428–437 (2015).
- Perera, T. P. S. et al. Discovery and pharmacological characterization of JNJ-42756493 (Erdafitinib), a functionally selective small-molecule FGFR family inhibitor. *Mol. Cancer Ther.* **16**, 1010–1020 (2017).
- Dobson, C. M. Chemical space and biology. *Nature* **432**, 824–828 (2004).
- Kotzyba-Hibert, F., Kapfer, I. & Goeldner, M. Recent trends in photoaffinity labeling. *Angew. Chem. Int. Ed. Engl.* **34**, 1296–1312 (1995).
- Singh, A., Thornton, E. R. & Westheimer, F. H. The photolysis of diazoacetylchymotrypsin. *J. Biol. Chem.* **237**, PC3006–PC3008 (1962).
- Grant, E. K. et al. A photoaffinity displacement assay and probes to study the cyclin-dependent kinase family. *Angew. Chem. Int. Ed.* **58**, 17322–17327 (2019).
- Guo, H. & Li, Z. Developments of bioorthogonal handle-containing photo-crosslinkers for photoaffinity labeling. *Med. Chem. Commun.* **8**, 1585–1591 (2017).
- Lapinsky, D. J. & Johnson, D. S. Recent developments and applications of clickable photoprobes in medicinal chemistry and chemical biology. *Future Med. Chem.* **7**, 2143–2171 (2015).
- Parker, C. G. et al. Ligand and target discovery by fragment-based screening in human cells. *Cell* **168**, 527 (2017).
- Forrest, I. & Parker, C. G. Proteome-wide fragment-based ligand and target discovery. *Isr. J. Chem.* **63**, e202200098 (2023).
- Galmazzi, A., Parker, C. G., Kok, B. P., Cravatt, B. F. & Saez, E. Discovery of modulators of adipocyte physiology using fully functionalized fragments. In *Phenotypic Screening: Methods and Protocols* (ed. Wagner, B.) 115–127 (Springer, New York, NY, 2018). https://doi.org/10.1007/978-1-4939-7847-2_9.
- Wang, Y. et al. Expedited mapping of the ligandable proteome using fully functionalized enantiomeric probe pairs. *Nat. Chem.* **11**, 1113–1123 (2019).

16. Liao, Y. et al. Chemical proteomics with novel fully functionalized fragments and stringent target prioritization identifies the glutathione-dependent isomerase GSTZ1 as a LUNG CANCER Target. *ACS Chem. Biol.* **18**, 251–264 (2023).
17. Grant, E. K. et al. A photoaffinity-based fragment-screening platform for efficient identification of protein ligands. *Angew. Chem. Int. Ed.* **59**, 21096–21105 (2020).
18. Mullard, A. Fragment-based screening sees the light. *Nat. Rev. Drug Discov.* **19**, 742–743 (2020).
19. Bajusz, D. et al. Exploring protein hotspots by optimized fragment pharmacophores. *Nat. Commun.* **12**, 3201 (2021).
20. Kutchukian, P. S. et al. Large scale meta-analysis of fragment-based screening campaigns: privileged fragments and complementary technologies. *SLAS Discov.* **20**, 588–596 (2015).
21. Drwal, M. N., Bret, G. & Kellenberger, E. Multi-target fragments display versatile binding modes. *Mol. Inform.* **36**, 1700042 (2017).
22. Swain, C. Fragment-based screening published hits. *Cambridge MedChem Consulting* https://www.cambridgemedchemconsulting.com/resources/hit_identification/fragment_based_screening_reported_hits.html.
23. Keserü, G. M. et al. Design principles for fragment libraries: maximizing the value of learnings from pharma fragment-based drug discovery (FBDD) programs for use in academia. *J. Med. Chem.* **59**, 8189–8206 (2016).
24. Ferenczy, G. G. & Keserü, G. M. Thermodynamics of fragment binding. *J. Chem. Inf. Modeling* **52**, 1039–1045 (2012).
25. Baell, J. B. & Holloway, G. A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J. Med. Chem.* **53**, 2719–2740 (2010).
26. Conway, L. P. et al. Evaluation of fully-functionalized diazirine tags for chemical proteomic applications. *Chem. Sci.* **12**, 7839–7847 (2021).
27. Donati, B., Lorenzini, E. & Ciarrocchi, A. BRD4 and Cancer: going beyond transcriptional regulation. *Mol. Cancer* **17**, 164 (2018).
28. Zeng, L. & Zhou, M.-M. Bromodomain: an acetyl-lysine binding domain. *FEBS Lett.* **513**, 124–128 (2002).
29. Huang, L., Guo, Z., Wang, F. & Fu, L. KRAS mutation: from undruggable to druggable in cancer. *Sig Transduct. Target Ther.* **6**, 1–20 (2021).
30. Weng, C., Faure, A. J., Escobedo, A. & Lehner, B. The energetic and allosteric landscape for KRAS inhibition. *Nature* 1–10 <https://doi.org/10.1038/s41586-023-06954-0> (2023).
31. Bröker, J. et al. Fragment optimization of reversible binding to the switch II pocket on KRAS leads to a potent, in vivo active KRASG12C inhibitor. *J. Med. Chem.* **65**, 14614–14629 (2022).
32. Sun, Q. et al. Discovery of Small Molecules that Bind to K-Ras and Inhibit Sos-Mediated Activation. *Angew. Chem. Int. Ed.* **51**, 6140–6143 (2012).
33. Maurer, T. et al. Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated nucleotide exchange activity. *Proc. Natl Acad. Sci.* **109**, 5299–5304 (2012).
34. Kessler, D. et al. Drugging an undruggable pocket on KRAS. *Proc. Natl Acad. Sci.* **116**, 15823–15829 (2019).
35. Mathieu, M. et al. KRAS G12C fragment screening renders new binding pockets. *Small GTPases* **13**, 225–238 (2022).
36. Moore, A. R., Rosenberg, S. C., McCormick, F. & Malek, S. RAS-targeted therapies: is the undruggable drugged? *Nat. Rev. Drug Discov.* **19**, 533–552 (2020).
37. Orlova, A. et al. Direct targeting options for STAT3 and STAT5 in cancer. *Cancers* **11**, 1930 (2019).
38. De Araujo, E. D. et al. Structural implications of STAT3 and STAT5 SH2 domain mutations. *Cancers* **11**, 1757 (2019).
39. Hu, T. et al. Impact of the N-terminal domain of STAT3 in STAT3-dependent transcriptional activity. *Mol. Cell. Biol.* **35**, 3284–3300 (2015).
40. Bonilla, P. A. et al. Virtual screening of ultra-large chemical libraries identifies cell-permeable small-molecule inhibitors of a “non-druggable” target, STAT3 N-terminal domain. *Front. Oncol.* **13**, 1144153 (2023).
41. Trowbridge, A. D. et al. Small molecule photocatalysis enables drug target identification via energy transfer. *Proc. Natl Acad. Sci.* **119**, e2208077119 (2022).
42. Geri, J. B. et al. Microenvironment mapping via Dexter energy transfer on immune cells. *Science* **367**, 1091–1097 (2020).
43. Seath, C. P. et al. Tracking chromatin state changes using nanoscale photo-proximity labelling. *Nature* **616**, 574–580 (2023).
44. Offensperger, F. et al. Large-scale chemoproteomics expedites ligand discovery and predicts ligand behavior in cells. *Science* **384**, eadk5864 (2024).
45. Mendez, D. et al. ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids Res.* **47**, D930–D940 (2019).
46. Davies, M. et al. ChEMBL web services: streamlining access to drug discovery data and utilities. *Nucleic Acids Res.* **43**, W612–W620 (2015).
47. Hallenbeck, K. K. et al. A liquid chromatography/mass spectrometry method for screening disulfide tethering fragments. *SLAS Discov.* **23**, 183–192 (2018).
48. Resnick, E. et al. Rapid covalent-probe discovery by electrophile-fragment screening. *J. Am. Chem. Soc.* **141**, 8951–8968 (2019).
49. Nonomiya, J., Li, K. S., Babin, B. M. & Mulvihill, M. M. Covalent library screening by targeted mass spectrometry for rapid binding site identification. *Anal. Chem.* **95**, 3779–3788 (2023).
50. Gilbert, K. E. et al. Profiling sulfur(VI) fluorides as reactive functionalities for chemical biology tools and expansion of the ligandable proteome. *ACS Chem. Biol.* **18**, 285–295 (2023).
51. Hunter, J. C. et al. In situ selectivity profiling and crystal structure of SML-8-73-1, an active site inhibitor of oncogenic K-Ras G12C. *Proc. Natl Acad. Sci. USA* **111**, 8895–8900 (2014).
52. Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput-Aided Mol. Des.* **27**, 221–234 (2013).
53. Sherman, W., Day, T., Jacobson, M. P., Friesner, R. A. & Farid, R. Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* **49**, 534–553 (2006).
54. Vonrhein, C. et al. Data processing and analysis with the autoPROC toolbox. *Acta Cryst. D* **67**, 293–302 (2011).
55. Smart, O. S. et al. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Cryst. D* **68**, 368–380 (2012).
56. Krissinel, E., Uski, V., Lebedev, A., Winn, M. & Ballard, C. Distributed computing for macromolecular crystallography. *Acta Cryst. D* **74**, 143–151 (2018).
57. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Cryst. D* **66**, 486–501 (2010).
58. Vagin, A. A. et al. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Cryst. D* **60**, 2184–2195 (2004).
59. Lee, W., Rahimi, M., Lee, Y. & Chiu, A. POKY: a software suite for multidimensional NMR and 3D structure calculation of biomolecules. *Bioinformatics* **37**, 3041–3042 (2021).
60. Pálffy, G., Vida, I. & Perczel, A. 1H, 15N backbone assignment and comparative analysis of the wild type and G12C, G12D, G12V mutants of K-Ras bound to GDP at physiological pH. *Biomol. NMR Assign.* **14**, 1–7 (2020).
61. de Araujo, E. D., Geletu, M. & Gunning, P. T. Strategies for over-expression and purification of recombinant full length STAT5B in *Escherichia coli*. *Protein Expr. Purif.* **129**, 1–8 (2017).

Acknowledgements

The authors are grateful to Mike Hann of GSK for initiating and managing this collaboration within the frame of FRAGNET. The authors acknowledge the contribution of Lilla Fajka to the synthesis of the library and of Elizabeth MacLean and Lizbe Koekemoer of University of Oxford to the BRD4 crystallization., as well as Dániel Pölöske and Richard Moriggi at the University of Veterinary Medicine, Vienna to the cell viability measurements against the MV4-11 and MOLM-13 cell lines. The authors thank the Bio-molecular Self-Assembly Research Group (and in particular, Tünde Juhász) at the HUN-REN Research Centre for Natural Sciences for providing access to their MST device. The authors received funding from the MSCA ITN 'FRAGNET' [grant no. 675899 to G.M.K.] and 'ALLODD' [grant no. 956314 to G.M.K.]; National Research Development and Innovation Office of Hungary [2020-1.1.2-PIACI-KFI-2020-00039 and K135150 to G.M.K., PharmaLab (RRF-2.3.1-21-2022-00015) to G.M.K., L.B. and A.P., FK146063 to D.B.] and from the National Brain Research Program [NAP 3.0 to G.M.K.] of the Hungarian Academy of Sciences. The work of P.Á-B. and D.B. was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and the ÚNKP-22-5, ÚNKP-23-5 New National Excellence Program of the Ministry for Innovation and Technology. The authors also acknowledge financial support from the National Laboratories Excellence program [under the National Tumor Biology Laboratory Project (NLP-17, 2022-2.1.1-NL-2022-00010)] and the Hungarian Thematic Excellence Program [TKP2021-EGA-44] to I.R. and J.T. This work was completed in the ELTE Thematic Excellence Program supported by the Hungarian Ministry for Innovation and Technology. Project no. 2018-1.2.1-NKP-2018-00005 has been implemented with the support provided by the National Research, Development and Innovation Fund of Hungary, financed under the 2018-1.2.1-NKP funding scheme. Project number RRF-2.3.1-21-2022-00015 is implemented with the support of the European Union's Recovery and Resilience Instrument. Supported by the Ministry for Innovation and Technology from the Hungarian NRD Fund (2020-1.1.6-JÖVŐ-2021-00004, 2020-1.1.6-JÖVŐ-2021-00010). The work of T.T. was supported by the KDP-2021 program of the Ministry of Innovation and Technology from the source of the National Research, Development and Innovation Fund.

Author contributions

G.M.K. and J.B. conceived the study. P.Á.-B. and G.M.K. supervised the project. D.B. and G.M.K. developed the SpotXplorer approach. D.B. performed theoretical calculations for compiling the PHP library. P.Á.-B., E.d.A., L.P., L.B., A.P., J.T. and T.K. contributed to experimental design and methodology. P.Á.-B., L.P., T.V.S., Q.A., Z.O., N.P., I.T., G.S., I.R., M.B., G.P., M.G., T.T., A.M., A.B.K., and E.K.G. performed experimental and

computational work, including synthesis, compound characterization, analytics and biological measurements. All authors contributed to writing and reviewing the manuscript, with leading contributions from P.Á.-B., D.B., Z.O., and G.M.K.

Funding

Open access funding provided by HUN-REN Research Centre for Natural Sciences.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42004-024-01252-w>.

Correspondence and requests for materials should be addressed to György M. Keserű.

Peer review information *Communications Chemistry* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

¹Medicinal Chemistry Research Group, HUN-REN Research Centre for Natural Sciences, Budapest, Hungary. ²National Drug Research and Development Laboratory, HUN-REN Research Centre for Natural Sciences, Budapest, Hungary. ³Department of Organic Chemistry and Technology, Faculty of Chemical Technology and Biotechnology, Budapest University of Technology and Economics, Budapest, Hungary. ⁴Department of Inorganic and Analytical Chemistry, Faculty of Chemical Technology and Biotechnology, Budapest University of Technology and Economics, Budapest, Hungary. ⁵Laboratory of Structural Chemistry and Biology & HUN-REN-ELTE Protein Modelling Research Group, Eötvös Loránd University, Budapest, Hungary. ⁶Hevesy György PhD School of Chemistry, Eötvös Loránd University, Budapest, Hungary. ⁷GlaxoSmithKline, Hertfordshire, UK. ⁸MS Metabolomics Research Group, HUN-REN Research Centre for Natural Sciences, Budapest, Hungary. ⁹Signal Transduction and Functional Genomics Research Group, HUN-REN Research Centre for Natural Sciences, Budapest, Hungary. ¹⁰Doctoral School of Biology, Institute of Biology, ELTE Eötvös Loránd University, Budapest, Hungary. ¹¹National Tumor Biology Laboratory and Department of Experimental Pharmacology, National Institute of Oncology, Budapest, Hungary. ¹²KINETO Lab Ltd, Budapest, Hungary. ¹³Department of Pathology, Forensic and Insurance Medicine, Semmelweis University, Budapest, Hungary. ¹⁴Department of Chemical and Environmental Process Engineering, Faculty of Chemical Technology and Biotechnology, Budapest University of Technology and Economics, Budapest, Hungary. ¹⁵Waters Research Center, Budapest, Hungary. ¹⁶MTA-ELTE Lendület Ion Mobility Mass Spectrometry Research Group, Eötvös Loránd University, Budapest, Hungary. ¹⁷Department of Chemical & Physical Sciences, University of Toronto Mississauga, Mississauga, ON, Canada. ¹⁸Centre for Medicinal Chemistry, University of Toronto at Mississauga, Mississauga, ON, Canada. ¹⁹These authors contributed equally: Péter Ábrányi-Balogh, Dávid Bajusz, Zoltán Orgován. ✉ e-mail: keseru.gyorgy@ttk.hu