

Photoaffinity Labeling



Target-Templated Construction of Functional Proteomimetics Using Photo-Foldamer Libraries

Edit Wéber⁺, Péter Ábrányi-Balogh⁺, Bence Nagymihály, Dóra K. Menyhárd, Nikolett Péczka, Márton Gadanez, Gitta Schlosser, Zoltán Orgován, Ferenc Bogár, Dávid Bajusz, Gábor Kecskeméti, Zoltán Szabó, Éva Bartus, Attila Tököli, Gábor K. Tóth, Tibor V. Szalai, Tamás Takács, Elvin de Araujo, László Buday, András Perczel, Tamás A. Martinek,^{*} and György M. Keserű^{*}

Abstract: Current methods for proteomimetic engineering rely on structure-based design. Here we describe a design strategy that allows the construction of proteomimetics against challenging targets without *a priori* characterization of the target surface. Our approach employs (i) a 100-membered photoreactive foldamer library, the members of which act as local surface mimetics, and (ii) the subsequent affinity maturation of the primary hits using systems chemistry. Two surface-oriented proteinogenic side chains drove the interactions between the short helical foldamer fragments and the proteins. Diazirine-based photo-crosslinking was applied to sensitively detect and localize binding even to shallow and dynamic patches on representatively difficult targets. Photo-foldamers identified functionally relevant protein interfaces, allosteric and previously unexplored targetable regions on the surface of STAT3 and an oncogenic K-Ras variant. Target-templated dynamic linking of foldamer hits resulted in two orders of magnitude affinity improvement in a single step. The dimeric K-Ras ligand mimicked protein-like catalytic functions. The photo-foldamer approach thus enables the highly efficient mapping of protein-protein interaction sites and provides a viable starting point for proteomimetic ligand development without *a priori* structural hypotheses.

[*] E. Wéber,⁺ B. Nagymihály, F. Bogár, G. Kecskeméti, Z. Szabó, É. Bartus, A. Tököli, Prof. G. K. Tóth, Prof. T. A. Martinek
Department of Medical Chemistry,
University of Szeged
Dóm tér 8, H-6720 Szeged, Hungary
E-mail: martinek.tamas@med.u-szeged.hu

E. Wéber,⁺ F. Bogár, É. Bartus, Prof. G. K. Tóth, Prof. T. A. Martinek
HUN-REN-SZTE Biomimetic Systems Research Group
Dóm tér 8, H-6720 Szeged, Hungary

P. Ábrányi-Balogh,⁺ D. K. Menyhárd, N. Péczka, Z. Orgován,
D. Bajusz, T. V. Szalai, Prof. G. M. Keserű
Medicinal Chemistry Research Group, HUN-REN Research Centre
for Natural Sciences
Magyar Tudósok Körútja 2, H-1117 Budapest, Hungary
E-mail: keseru.gyorgy@ttk.hu

D. K. Menyhárd, M. Gadanez, Prof. A. Perczel
Laboratory of Structural Chemistry and Biology, Institute of
Chemistry,
Eötvös Loránd University
Pázmány Péter sétány 1/A, H-1117 Budapest, Hungary

D. K. Menyhárd, Prof. A. Perczel
HUN-REN-ELTE Protein Modeling Research Group, Institute of
Chemistry,
Eötvös Loránd University
Pázmány Péter sétány 1/A, H-1117 Budapest, Hungary

M. Gadanez
Hevesy György PhD School of Chemistry, Institute of Chemistry,
Eötvös Loránd University
Pázmány Péter sétány 1/A, H-1117 Budapest, Hungary

G. Schlosser
MTA-ELTE Lendület Ion Mobility Mass Spectrometry Research
Group, Institute of Chemistry,
Eötvös Loránd University
Egyetem tér 1–3, H-1053 Budapest, Hungary

T. V. Szalai
Department of Inorganic and Analytical Chemistry, Faculty of
Chemical Technology and Biotechnology,
Budapest University of Technology and Economics
Szt. Gellért tér 4, H-1111 Budapest, Hungary

T. Takács, Prof. L. Buday
Institute of Molecular Life Sciences, HUN-REN Research Centre for
Natural Sciences
Magyar Tudósok Körútja 2, H-1117 Budapest, Hungary

T. Takács
Doctoral School of Biology, Institute of Biology,
Eötvös Loránd University
Egyetem tér 1–3, H-1053 Budapest, Hungary

Prof. E. de Araujo
Centre for Medicinal Chemistry,
University of Toronto at Mississauga
Ontario L5 L 1 C6, Mississauga, Canada

Prof. G. M. Keserű
Department of Organic Chemistry and Technology, Faculty of
Chemical Technology and Biotechnology,
Budapest University of Technology and Economics
Budafoki út 8, H-1111 Budapest, Hungary

P. Ábrányi-Balogh,⁺ D. K. Menyhárd, N. Péczka, M. Gadanez,
Z. Orgován, D. Bajusz, T. V. Szalai, T. Takács, Prof. L. Buday,
Prof. A. Perczel, Prof. G. M. Keserű
National Drug Discovery and Development Laboratory
Magyar Tudósok Körútja 2, H-1117 Budapest, Hungary

[*] These authors contributed equally to this work.

© 2024 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Introduction

Engineering artificial structures that mimic complex folding and functionality of proteins remains challenging for chemists. While past studies focused on the folding properties of the synthesized molecules, now the primary goal is to create proteomimetics with functions,^[1] e.g., potent bio-inspired catalysts^[2] or compounds modulating disease-relevant protein-protein interactions (PPIs)^[3] that cannot be trivially targeted by small molecules. With the recent development of nanomolar entrapment-free cell delivery technologies for macromolecules,^[4] these larger molecular weight entities have become feasible to address even in case of intracellular targets.

The design of functional proteomimetics aimed for molecular recognition or catalytic activity is currently based on structural hypotheses – the notion that the structure of a functionally active protein should be fortified or mimicked.^[1a] While hydrophobic core packing, hydrogen bonding, salt bridging, and metal binding play critical roles in folding, covalent constraints can be introduced for further stabilization. Cross-linking was efficiently used to increase the thermal stability of enzymes,^[2a,5] and for the successful design of potent PPI modulators.^[6] A common top-down approach is incorporating non-natural amino acids into the sequence, which can improve proteolytic stability^[7] and fine-tune local folding.^[8] Using heterogeneous backbones, various tertiary motifs, such as zinc finger domain mimics,^[9] amyloid models,^[10] heterochiral knottin proteins,^[11] and foldamer-protein hybrids^[12] were constructed. Since structural and functional mimicry can also be achieved without reproducing the exact tertiary fold, some rational design strategies aim to mimic only the protein segments that are responsible for the function. For instance, antibodies and nanobodies were downsized by recapitulating their complementarity-determining loop regions.^[13]

De novo design principles proved that tertiary structures can be engineered from scratch. Topological templates were exploited for clustering secondary structural elements,^[14] various artificial helix bundles,^[15] and β -sandwich scaffolds were engineered.^[16] Computational design tools have been developed for miniproteins^[17] and tertiary structures incorporating covalently linked cores (CovCore proteins).^[18] These de novo approaches use empirical rules established for the folding motifs in question and strongly rely on *in silico* design based on a structural hypothesis.

We propose a bottom-up strategy for the design of functional proteomimetics that does not require *a priori* knowledge about the surface characteristics of the target protein or the protein-protein interaction sites. We focus on challenging PPIs with shallow or dynamic interaction surfaces that are typically less druggable by small molecules. First, we suggest that a local protein surface mimetic library can map hot spots of the targets without any *a priori* structural information. Second, we propose using a systems chemistry^[19] method to link the target-binding building blocks and select the proteomimetics according to their molecular recognition properties.

Literature data exemplify that the precise replication of the tertiary structure is not essential for functional mimicry.^[13a,b] Moreover, the characteristics of protein interaction surfaces are dominated by the nature and side chain conformation of their residues, mostly independent of the underlying specific secondary/tertiary structure.^[20] This means that the spatial extent of successful surface probes should also be on the scale of side chains. Based on these considerations, we previously constructed a focused local surface mimetic (LSM) library^[21] with short water-stable β -peptidic foldamer scaffolds and two variable proteinogenic amino acid side chains for protein surface recognition. Screening the 256-membered library against different targets demonstrated that LSMs could locally mimic protein surfaces, bind to protein hot spots with sufficient affinity, and exhibit side chain-dependent protein recognition patterns.^[21] They also discriminated between structurally highly homologous but functionally different targets.^[21–22]

In this work, we aimed to construct a minimal LSM library of proteomimetic building blocks and develop a methodology for sensitive binding detection and localization. One of the major challenges of such an approach is capturing the instantaneous proximity between the specific protein surface patches and the LSMs. Therefore, we equipped our probes with diazirine-based photoaffinity tags, which allow covalent anchoring to the surface by simple illumination.^[23] After determining their binding site, the possibility of cross-coupling or growing the LSMs allows for affinity maturation.^[24] Our goal was to simplify the iterative design-synthesis-functional test cycles by exploiting a systems chemistry approach, which can simultaneously work with networks of interacting molecules, increase the complexity of the compounds, and focus on functional properties such as target binding.^[19]

We selected one conventional and two challenging targets, the promiscuous binder calmodulin (CaM), a transcription factor (STAT3) and a major oncogene (K-Ras G12D), for proof-of-concept studies. CaM is a ubiquitous member of the calcium signal transduction pathways. It is a small, flexible protein that binds a plethora of helical substrates. STAT3 is a transcription factor containing six functionally conserved domains that is frequently activated in human solid and hematological cancers.^[25] The shallow interaction surfaces of the different domains so far restricted the development of direct STAT3 inhibitors to weak binders of its SH2 domain, intended to inhibit activating dimerization.^[26] Ras proteins are small GTPases^[27] conveying growth signals. Disease-causing mutations lock these enzymes in the GTP-loaded activated state, contributing to the emergence of nearly 20 % of all human cancers. Ras proteins are ineffective enzymes by themselves; thus, assisting proteins are required for their catalytically adequate functioning. Interaction with these and the downstream partners is focused on two short, dynamic segments, referred to as the switch regions.^[28] Since oncogenic variants are structurally very similar to the wild-type enzyme and the targetable interactions all take place on shallow and ever-changing PPIs created by the switches; it is no wonder that Ras proteins have long been considered undruggable.

Our results obtained for a range of targets demonstrate the feasibility of functional proteomimetics without *a priori* localization and characterization of the possible binding sites on the surface of the target protein. LSMs can be used to map targetable sites even for challenging hosts, and by combining LSM binders of neighboring sites, proteomimetic interaction partners of the targeted systems can be developed.

Results and Discussion

Design and Synthesis of the Photo-Foldamer Library

The LSM library was generated using an H14 helical β -peptide foldamer template^[29] equipped with two proteinogenic side chains (Figure 1a). These synthetically accessible foldamers are constructed from (1*S*,2*S*)-2-aminocyclohexanecarboxylic acids (ACHC) and β^3 -amino acid residues at the variable positions R_1 and R_2 . While the bulky and constrained ACHCs induce helical folding and serve as a solvent shield, proteinogenic side chain pairs (R_1 and R_2) pointing outward on the same face of the helix are responsible for recognizing the target.

To develop a sensitive screening method that enables the localization of the binding, we attached a UV-reactive diazirine tag to the LSMs (photo-LSM). Diazirine generates carbene with sufficient crosslinking efficiency for weak ligand-target interactions,^[23a] enabling the development of

fragment screening platforms.^[23b,30] We attached the L-photo-methionine residue bearing a flexible and medium-length side chain to the C-terminal of LSMs through a Gly spacer (Figure 1a). LSM peptides were named according to the proteinogenic side chain combinations using the one-letter codes corresponding to the standard α -amino acid notation.

PPIs are governed by the limited chemical space of proteinogenic side chains enriched at the interface.^[20] Considering the importance of the aromatic side chains at PPI interfaces,^[20a,31] we included β^3h Trp, β^3h Phe, and β^3h Tyr residues in the library. In the present work, we further sampled side chain space by selecting only one amino acid from each of the following categories: cationic (β^3h Arg), anionic (β^3h Glu), polar non-charged with hydroxyl (β^3h Ser) and with amide (β^3h Asn), short aliphatic (β^3h Val), long branched aliphatic (β^3h Leu), and non-branched aliphatic side chains (β^3h Met). Accordingly, by varying the 10 different residues at the R_1 and R_2 positions, the photo-LSM library members totaled up to 100 sequences. This is a strongly limited set compared to the size of small molecule fragment libraries.^[32] We assumed to the first approximation that these templated side chains can drive detectable binding to the target due to the similarity to the spatial arrangement of the native PPI surfaces and the solvent shielding provided by the bulky helical scaffold.

The photo-foldamer library was successfully synthesized by standard solid-phase peptide synthesis methods. The helical folding propensity of the (1*S*,2*S*)-ACHC-stabilized

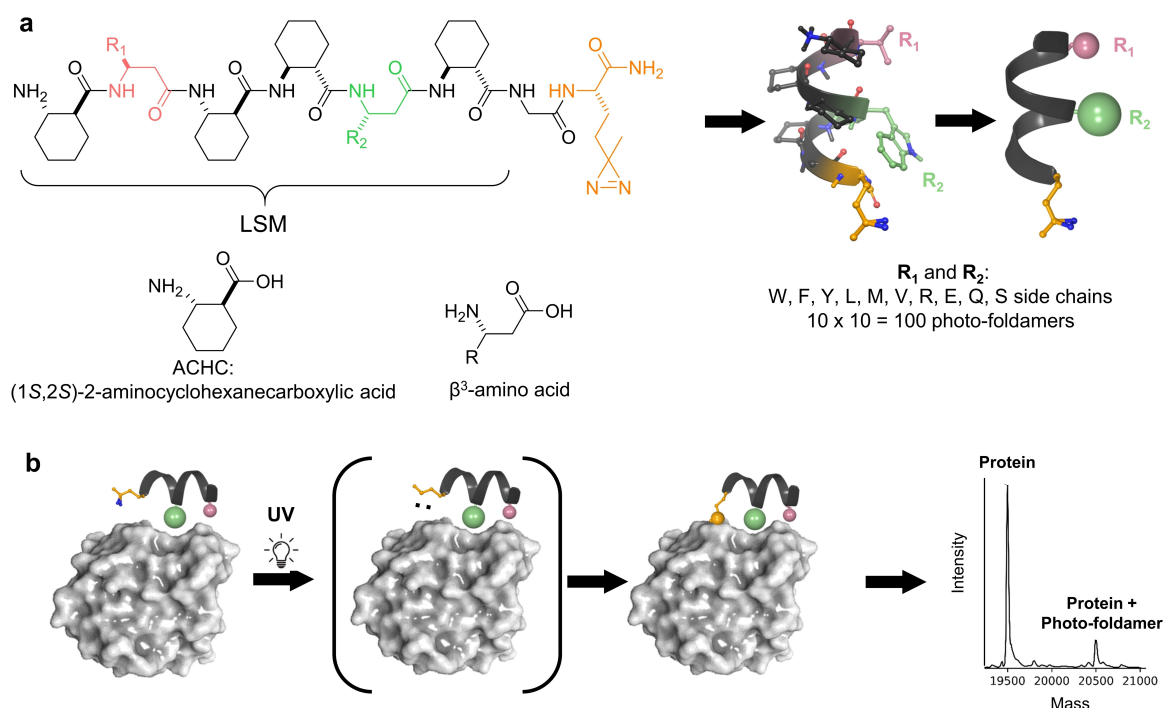


Figure 1. Photo-foldamer library design and screening workflow. a) General structure and helical fold of the local surface mimetic (LSM) β -peptide library equipped with the photoreactive diazirine moiety (photo-LSM). While ACHC residues serve as a scaffold, proteinogenic side chains in the variable R_1 and R_2 positions are responsible for target recognition. b) Proposed screening workflow: (i) during the incubation period, photo-LSM can bind to the protein surface, (ii) upon irradiation with UV light, reactive carbene radical is formed, and the bound foldamer labels the target covalently, (iii) the crosslinked foldamer-protein complex is analyzed by intact HPLC-MS.

photo-foldamers^[21,24, 33] was confirmed with 2D ¹H NMR (Figure S1) and electronic circular dichroism (ECD) measurements (Figure S2). The results of the molecular dynamics (MD) simulations (Figure S3) also supported the stability of the folded photo-LSM helices in water.

Photo-Foldamer Library Members Show Side Chain-Dependent Binding to Diverse Protein Surfaces

The screening experiments were performed according to the established photoaffinity labeling protocols.^[23b] Library members were screened individually. The purified recombinant protein was preincubated with photo-LSMs and samples were irradiated with UV light (365 nm, 10 min). Finally, each sample was analyzed with HPLC-MS (Figure 1b). Binding was detected through the mass of the crosslinked protein-photo-foldamer products. Crosslinking yields were calculated for each photo-LSM and represented as heat maps.

First, a promiscuous target was used to determine the appropriate concentration range of the diazirine-derivatized photo-LSMs. Photo-LSM concentrations in the range of 20–160 μ M resulted in complete crosslinking for CaM, and multiple labeling events, in accordance with our previous results concerning (non-derivatized) LSMs (dissociation constants in the 0.1–10 μ M range).^[24] CaM contains two very similar hydrophobic attractor regions. Using lower concentrations and equimolar conditions (2 μ M CaM, 2 μ M photo-LSM), CaM-binder photo-LSMs could still be trapped with high crosslinking yields (Figure 2a).

Next, we have evaluated the photo-LSM library against the prominent oncotarget STAT3, a multi-domain transcription factor. To date, the overwhelming majority of reported STAT3 inhibitors target the SH2 domain, which is responsible for the assembly of the active dimer (required for nuclear translocation and DNA binding),^[26] with scarce examples of targeting other areas, such as the coiled-coil or DNA-binding domains.^[34] We hypothesized that we could achieve a more thorough coverage of the possible binding surfaces of STAT3 due to the special geometric features of the photo-LSM library. Considering the extended, shallow, and often discontinuous interaction surfaces of STAT3, the screen was performed using 2 μ M STAT3 (our construct contained the 122–688 segment of STAT3 that is 64.3 kDa) and 160 μ M photo-LSM concentrations. We identified 31 hits, which could be classified into three groups according to crosslinking yields: low (0–5 %), medium (5–20 %), and high (>20 %) (Figure 2a). The binding was side chain dependent, and the best STAT3 ligands were **ph-VY**, **ph-YR**, **ph-YQ**, **ph-WR**, **ph-WQ**, **ph-EQ**, **ph-ES**, **ph-QQ**, and **ph-QS** (the prefix 'ph' denoting 'photo' was added to distinguish the photo-tagged sequences). The overall binding pattern indicated a strongly limited affinity for side chain combinations where both proteinogenic residues (R_1 and R_2) were hydrophobic. Reflecting the previous difficulties of targeting STAT3, the hit rate was considerably lower than that of the promiscuous CaM (Figure 2b), despite the higher ligand concentration and the large surface of STAT3.

We screened the photo-foldamer library against another challenging and structurally dynamic target, K-Ras G12D, an oncogenic variant of K-Ras that resulted in a remarkably side chain-dependent crosslinking efficiency. For 5 μ M K-Ras G12D, concentration-dependent responses were detected for the photo-LSMs at concentrations of 40, 80 and 160 μ M (Figure 2c-d and Figure S4). Some hydrophobic library members (**ph-LW**, **ph-MW**) did not follow the concentration-dependent trend, which could be explained by the aggregation of the compounds at higher concentrations. The best K-Ras binders were the photo-LSMs with aromatic-polar or aromatic-valine side chain combinations: **ph-VF**, **ph-VW**, **ph-RW**, **ph-RF**, **ph-QW**, **ph-QF**, **ph-SW**, **ph-SF**, **ph-FV**, and **ph-WR**. Multiple labeling was detected for **ph-QW**, **ph-SW**, and **ph-RW**. High crosslinking yields could only be obtained at higher ligand excess (160 μ M), but most hits were detectable using 80 μ M photo-LSM (Figure S4).

In order to validate the screening results, the binding of selected photo-LSMs was independently tested by ¹H saturation transfer difference (STD) NMR (Figure 2e). STD NMR confirmed the binding of the hits to STAT3 and K-Ras G12D. Non-binders of the photoaffinity tagging protocol did not exhibit STD signal (Figure S5). The signals of the proteinogenic side chains in the difference spectra indicated that these protons are in contact with the protein. This finding aligns well with the proposed proteinogenic side chain dependent binding mode and the inability of the ACHC residues to directly drive the interaction. The best hits were studied by using isothermal titration calorimetry (ITC). Titrations carried out with selected photo-LSMs (**ph-VW**, **ph-VF**, and **ph-RF**) confirmed the binding to K-Ras. Experimental data supported K_d values in the range of 200 to 700 μ M (Figure S6).

Domain/Segment-Level Proteomic Binding Site Localization and Functional Tests

For the two more challenging target proteins (STAT3 and K-Ras), we relied on the identification of tryptic peptide segments labeled by the photo-LSMs as a primary method of binding site localization. These results were further refined by MS/MS analyses (Table S1 and Table S2). Hits labeled a wide array of binding sites across the STAT3 surface, involving all the available STAT3 domains (Figure 3a and Table S1). Several photo-LSMs (e.g., **ph-YR**, **ph-RQ**, **ph-SR**, and **ph-QW**) bound to the solvent-exposed coiled-coil region of the protein, and we identified five different helical segments as labeling sites. Compounds containing hydrophilic side chains, such as **ph-QQ**, **ph-ES**, **ph-SQ**, and **ph-EQ** labeled STAT3 at the boundary of the DNA binding and the linker domain. Hydrophilic R_1 and aromatic R_2 side chain combinations (e.g., **ph-QY**, **ph-SY**, and **ph-RW**) recognized the helical and disordered part of the SH2 domain. For 10 hits, multiple binding sites were found.

We proceeded with three photo-LSM hits that either produced high-intensity signals in LC-MS/MS or labeled a unique binding site and characterized them in a STAT3

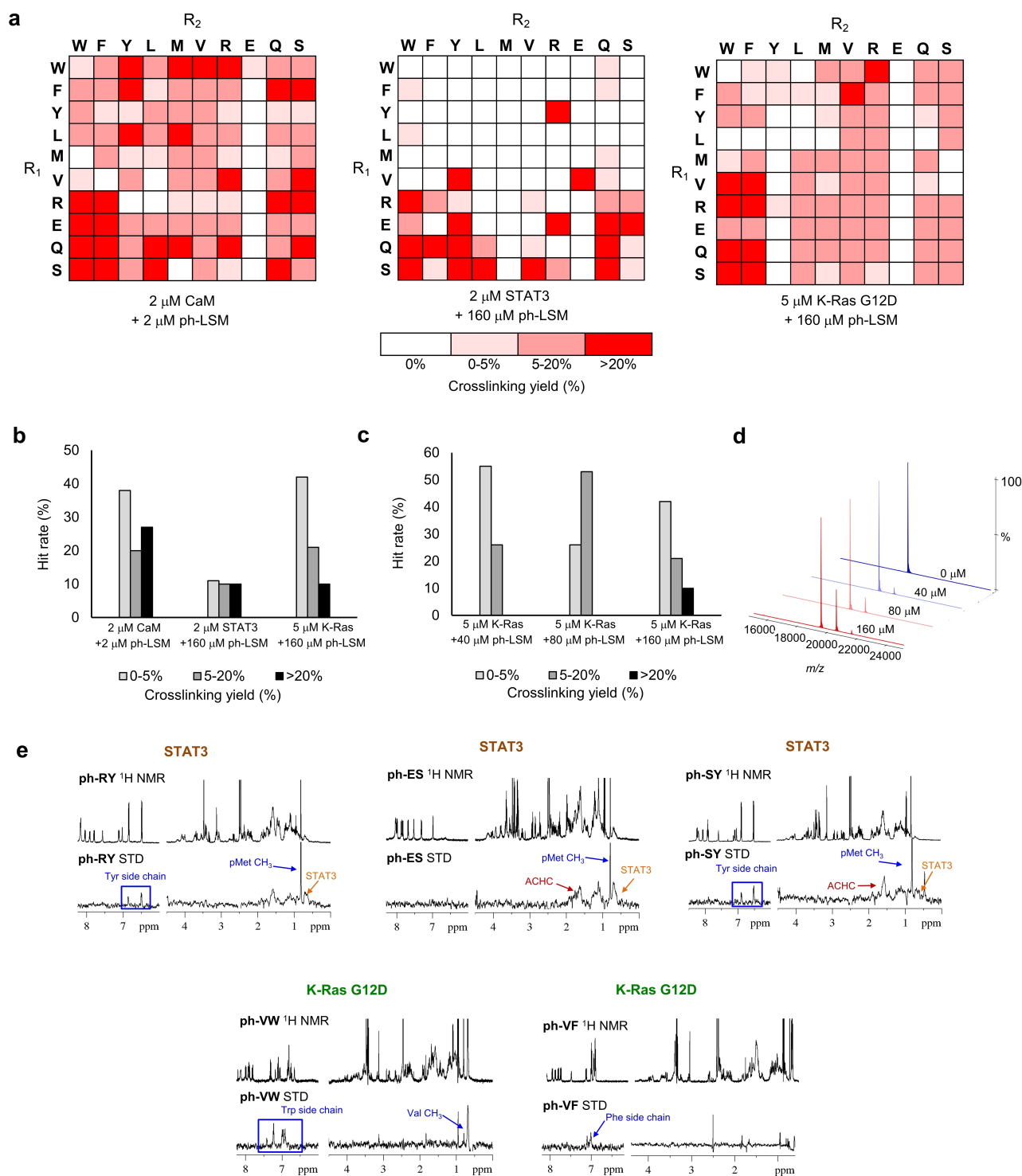


Figure 2. a) Photo-foldamer screening results for CaM, STAT3 and K-Ras G12D. Binding patterns are represented as heat maps. Side chains of the variable R_1 and R_2 positions are given at the headings of the rows and columns corresponding to the standard alpha amino acid codes. b) Hit rate statistics for the promiscuous CaM and the challenging targets as STAT3 and K-Ras G12D. c) Hit rates for K-Ras G12D using different photo-LSM concentrations. d) Intact MS spectra for K-Ras G12D with increasing concentrations of a representative photo-LSM **ph-QW** after UV irradiation. e) STD NMR experiments confirm **ph-RY**, **ph-ES**, and **ph-SY** binding to STAT3; and **ph-VW** and **ph-VF** to K-Ras G12D, respectively.

fluorescence polarization assay.^[35] Each demonstrated dose-dependent responses with IC_{50} values in the mid-to-high micromolar range after irradiation (Figure 3b). Thus, by covalently capturing members of the photo-foldamer library,

we could identify hits that bind to biologically relevant sites on STAT3.

The PPIs of K-Ras are focused on two dynamic segments, switch I and switch II, that form the primary port for

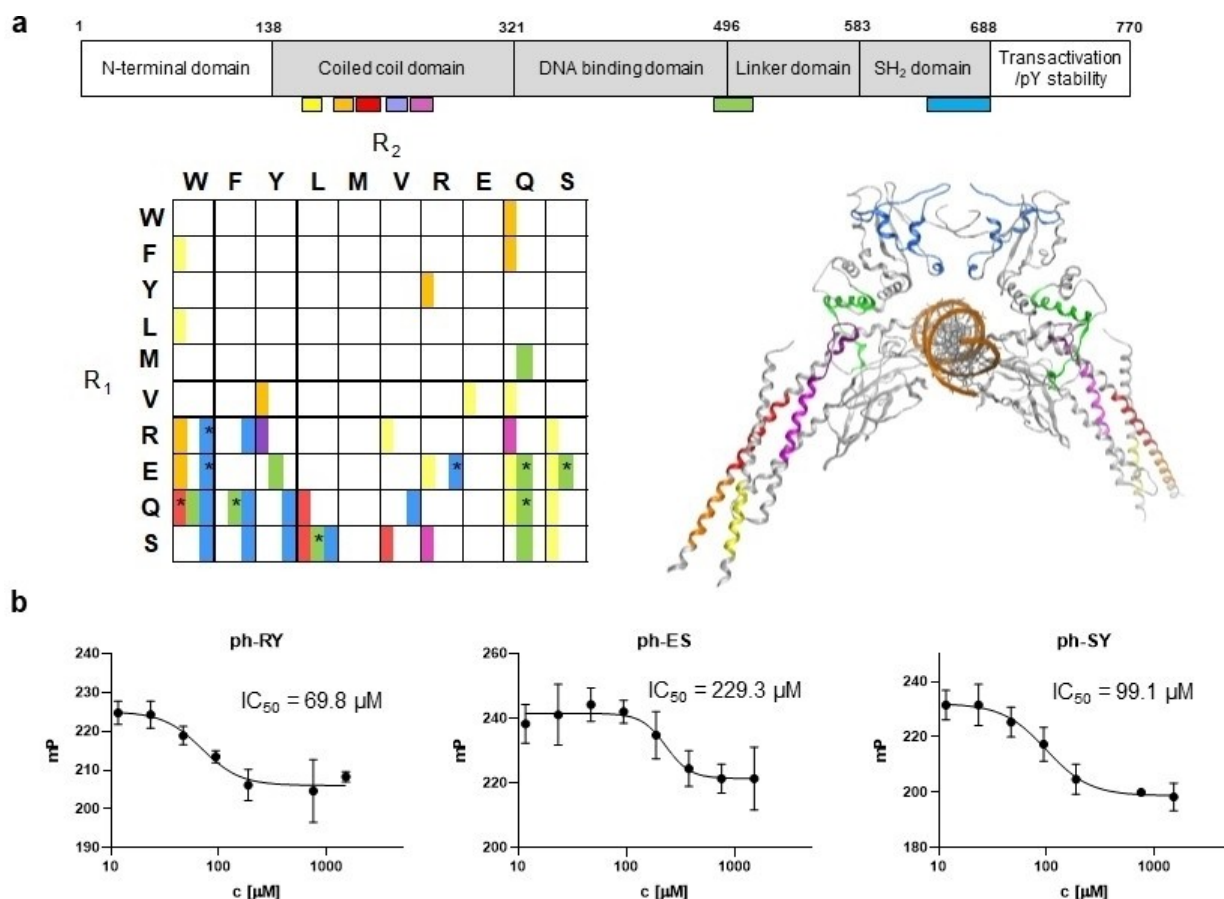


Figure 3. Segment-level LSM site localization and functional tests for STAT3. a) Binding sites of photo-LSMs on the surface of the STAT3 protein as identified by LC-MS. Labeled sites are colored according to the colors of the STAT3 segments at the ribbon representation of the protein backbone (PDB: 6QHD). To represent multiple sites, each cell corresponding to a photo-LSM compound (defined by the R₁ and R₂ groups as above) is divided into three sub-cells, and the site with the highest intensity is marked with an asterisk. b) Representative photo-LSMs bound to different domains showed STAT3 inhibitory activity in the fluorescence polarization assay.

both the effector- and activator-proteins and also participate in forming the GTPase active site. To reveal whether photo-LSMs can recognize these critical sites, complexed and irradiated samples of 33 screening hits were digested and analyzed by LC-MS/MS (Table S2). Most hits labeled the switch II region, suggesting that photo-LSMs are promising starting points for further ligand development. Besides switch II, the $\alpha 3$ and $\alpha 4$ helices, and to a lesser extent, $\beta 5$ and $\alpha 5$, were also affected, indicating that the various LSMs sampled the entire surface of the enzyme. Direct labeling within the switch I region was not found; however, covalent linkage observed at the nearby $\alpha 5$ helix suggested that the LSMs also reached the proximity of the switch I region. The localization pattern appeared side chain-dependent (Figure 4a). Hydrophobic compounds (**ph-LW**, **ph-MW**, **ph-VW**, **ph-VF**) and the photo-LSMs with Arg side chain in the R₂ position (**ph-VR**, **ph-LR**, **ph-ER**, **ph-QR**) were linked to the switch II region only. More hydrophilic photo-LSMs, typically compounds with Gln or Ser side chains (e.g., **ph-QF**, **ph-QV**, **ph-MQ**, **ph-SQ**), could be crosslinked with each of the three segments switch II region, $\alpha 3$ and $\alpha 4$ helices,

indicating that these sequences bind to multiple sites at the surface of K-Ras G12D.

Having confirmed the **ph-VW** - K-Ras G12D interaction using three different techniques (MS, STD NMR, and ITC), we tested whether the binding of the foldamer segments can exert influence over the K-Ras catalytic cycle by blocking the binding of the guanine nucleotide exchange factor SOS1 which provides considerable enhancement to the otherwise weak intrinsic exchange capacity of K-Ras (Figure 4b). The presence of non-covalently bound **ph-VW** interfered with the function of SOS1, reverting the nucleotide exchange rate toward the intrinsic level (Figure 4b, left panel). The effect was even more remarkable upon irradiation (Figure 4b, right panel). These results support the MS-based identification of the **ph-VW** binding site at switch II and the SOS1-binding interface.

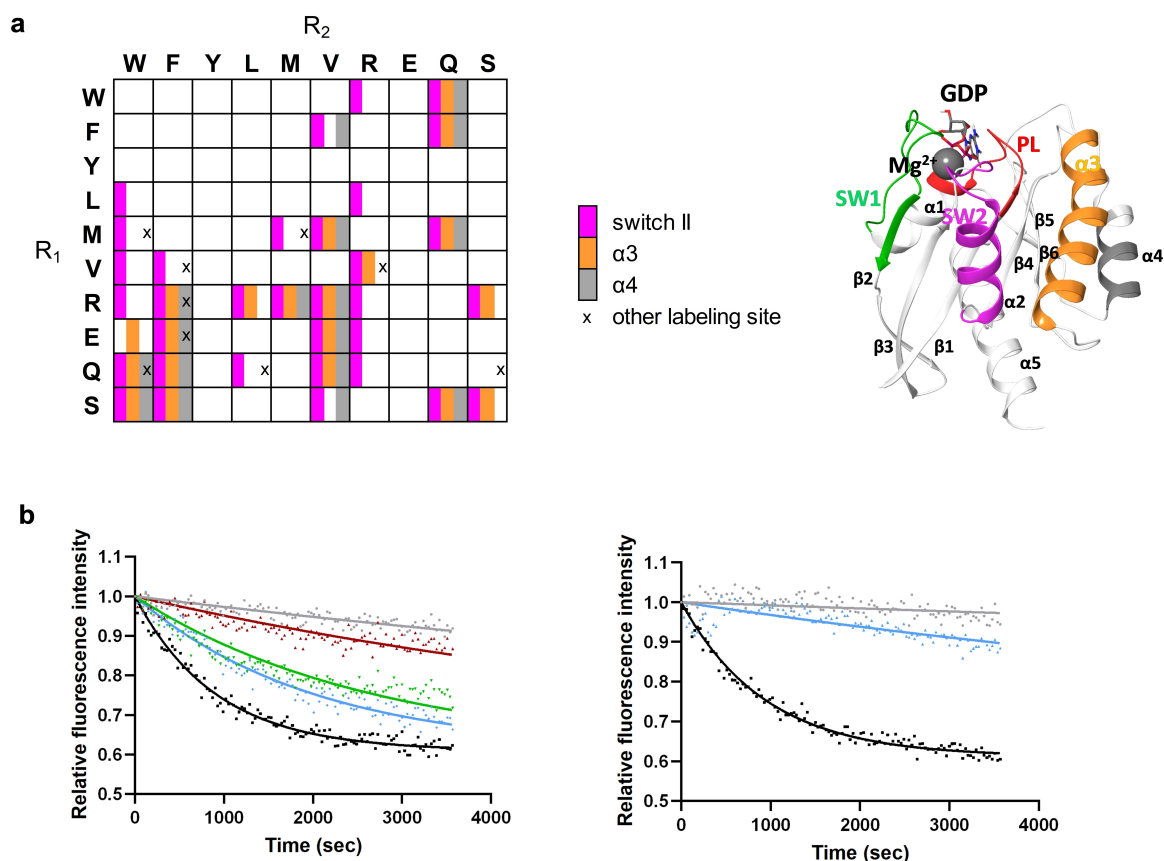


Figure 4. Segment-level LSM site localization and functional tests for K-Ras G12D. a) Photo-LSM labeled segments of K-Ras identified by LC-MS/MS. Each cell corresponding to a photo-LSM compound (defined by the R_1 and R_2 groups as above) is divided into three sub-cells representing the labeled sites colored according to the colors of the K-Ras segments at the ribbon representation of the protein. b) Foldamer fragment hits influence the functional activity of K-Ras G12D assessed in guanine nucleotide exchange experiments. In the K-Ras-SOS exchange assay, SOS1 displaces MANT-labeled GDP from K-Ras (black curve), preparing the enzyme for GTP binding and activation. Addition of **ph-VW** (blue, green and red curves for 125 μ M, 250 μ M and 500 μ M concentrations, respectively) slows the nucleotide exchange rate close to that of the intrinsic capacity of K-Ras (indicated by the grey curve) especially upon irradiation (right) - despite the presence of exchange booster SOS1 – showing that **ph-VW** effectively inhibits SOS1 binding.

NMR-Based Residue-Level Binding Site Localization for K-Ras G12D

In contrast to small molecular fragments, photo-LSMs span distances in the range of 13–15 Å over target surfaces that can be highly flexible. Thus, the spatial localization of the binding site(s) around the MS-determined anchor residues may need further measurements. In our case, LC-MS data did not identify the switch I region as a covalent anchor site, whereas the switch II region and the $\alpha 3$ and $\alpha 4$ helices were frequently labeled. To achieve a more detailed characterization of K-Ras G12D binding sites, we carried out ^{15}N HSQC NMR-based chemical shift perturbation (CSP) measurements, comparing the spectra of the covalent complexes to that of the free protein. Binding sites were defined by the spatial clustering of the affected target residues. First, we determined the center of the binding sites for each photo-foldamer using average-linkage hierarchical clustering based on the distance of the center-of-mass coordinates of the CSP-tagged residues. Second, we assumed that the target displays a limited number of consensus sites prone to bind

the photo-LSM fragments. To detect the consensus sites, we clustered all the collected binding site centers (Figure S7). Eight consensus clusters were identified (Figure 5, Table S3), spanning a large portion of the total surface of the protein. The NMR-identified sites aligned well with the anchor points revealed by MS (Figure 5 and Table S4), thus validating the photo-foldamer screening protocol. Finally, we considered those seven sites that were confirmed by MS: C0, C1, C2, C4, C5, C6, and C7. Clusters C4 and C7 explicitly include amino acids of the switch I region.

Photo-LSMs could usually be assigned to multiple sites (Figure S8 and Table S4). Hydrophobic and Trp-containing compounds (**ph-LW**, **ph-MW**, **ph-VW**, **ph-VF**, **ph-RW**, **ph-QW**) bound to C5, C6, and C7 sites only, which contain the switch regions. Besides these locations, photo-LSMs with Arg side chain in the R_2 position (**ph-VR**, **ph-LR**, **ph-ER**, **ph-QR**) could also engage the C1 site, which is a previously untargeted small pocket distant from the switch regions. Compounds with hydrophilic side chains in position R_1 (e.g., **ph-SF**, **ph-RF**, **ph-EV**, **ph-QV**, **ph-SV**) typically recognized clusters C2 and C4. C2 is a binding site of non-covalent

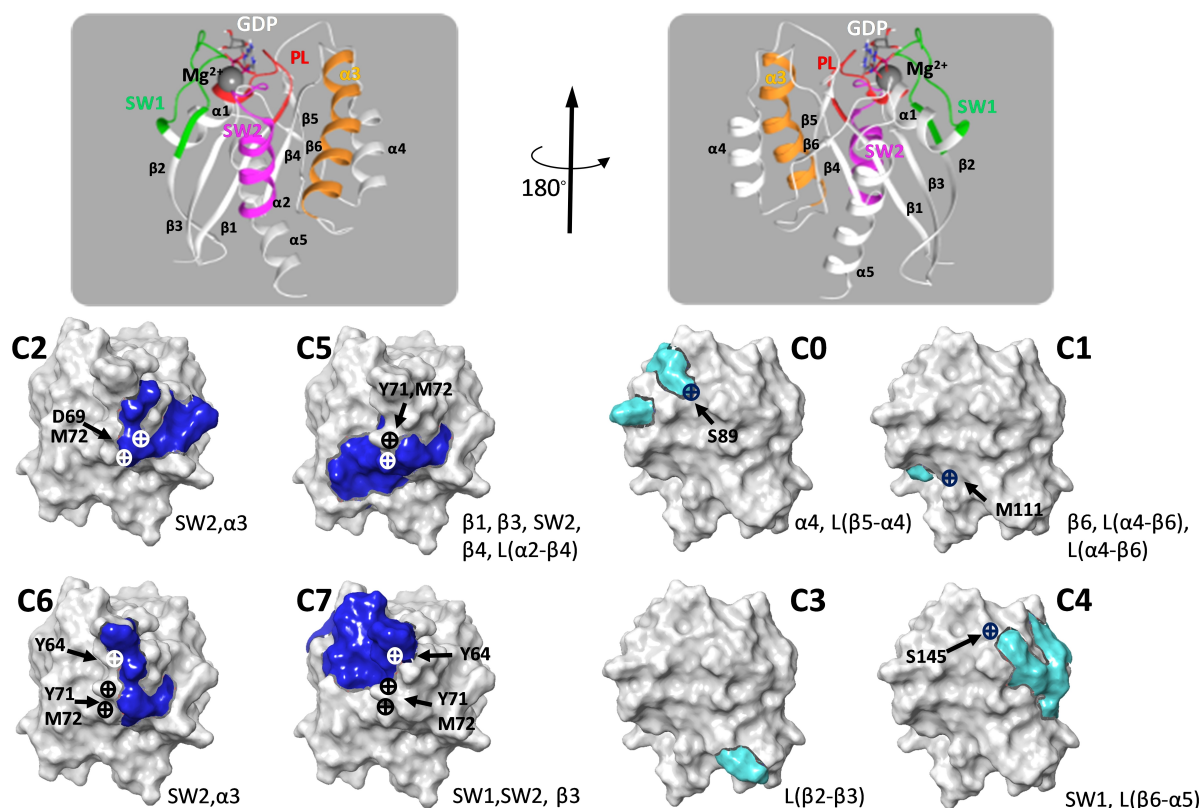


Figure 5. Consensus binding sites of photo-LSMs on the surface of the K-Ras protein. The secondary structural elements that contribute to these binding sites and the overall orientation of the protein are indicated in the top left (front side including SW2 region) and top right (back side, opposite to the nucleotide binding site) panels. Blue and cyan colors indicate front and back side positions of the consensus sites, respectively. Possible anchor residues ($\oplus$) of the covalently bound photo-LSMs are assigned to each binding site.

inhibitors located between switch II and $\alpha 3$. C4 residues flank switch I and an adjacent loop.

NMR mapping confirmed that photo-LSMs located the functionally important surface patches of K-Ras. Sites C2, C5, C6, and C7 sample the surface grooves that harbor peptides and small molecules that were shown to be effective inhibitors of the enzyme. The most extended C7 consensus site overlaps with the common PPI interface of the downstream partners, such as Raf1, PIK3CG, or RALGDS, the guanine nucleotide exchange factor SOS1, and the GTPase activating proteins (GAPs) (Figure S9). C7 contains the entirety of switch I, most of the switch II residues, and the tip of the P-loop. The short P-loop is part of the GTPase active site, where the most significant oncogenic mutations occur. Therefore, ligands occupying site C7 could not only be sensitized toward its mutation state, but they might also interfere with the binding and catalysis of the activator proteins and the downstream partners of the enzyme. All confirmed photo-LSM hits sampled the C7 site.

Our photo-foldamer-based approach identified physiologically relevant regions of K-Ras and binding sites that were previously successfully targeted (C2, C5, C6, and C7). Moreover, novel and untested niches were detected that might also exert influence over its catalytic potency (C0, C1, and C4).

Affinity Maturation by Dynamic Covalent LSM Ligation for K-Ras G12D

The MS and NMR localization revealed that the best binder photo-LSMs can be assigned to the surface clusters C7 and C5, which overlap with the biologically relevant PPI interfaces. The relative proximity of these surface patches offered the possibility of testing the ligation of the LSMs as a means of affinity maturation. We selected **VW**, **VF**, **QF**, **QW**, **WQ**, **WR**, **QR**, and **LR** as candidate surface fragments to build improved affinity dimers. The LSM dimer selection procedure was based on an adaptive self-sorting dynamic combinatorial library (DCL) approach. LSM fragments were synthesized with a Gly-Gly-Cys tag to exploit the thiol-disulfide exchange occurring in a redox buffer (GSH/GSSG mixture) under thermodynamic control. We used K-Ras G12D to template the dimer formation (Figure 6a).

Reaction mixtures obtained by incubation with and without the target were analyzed with HPLC-MS. Amplification factors (ratio of AUCs with and without K-Ras) were determined for each library member (Table S5). In the presence of K-Ras G12D, the formation of several dimers was enhanced (Figure 6b, Table S5). For **VW-WQ**, **VF-WQ**, **VW-VW**, **VF-VW**, and **VF-VF**, an over tenfold increase in concentration was observed in the protein-templated sample.

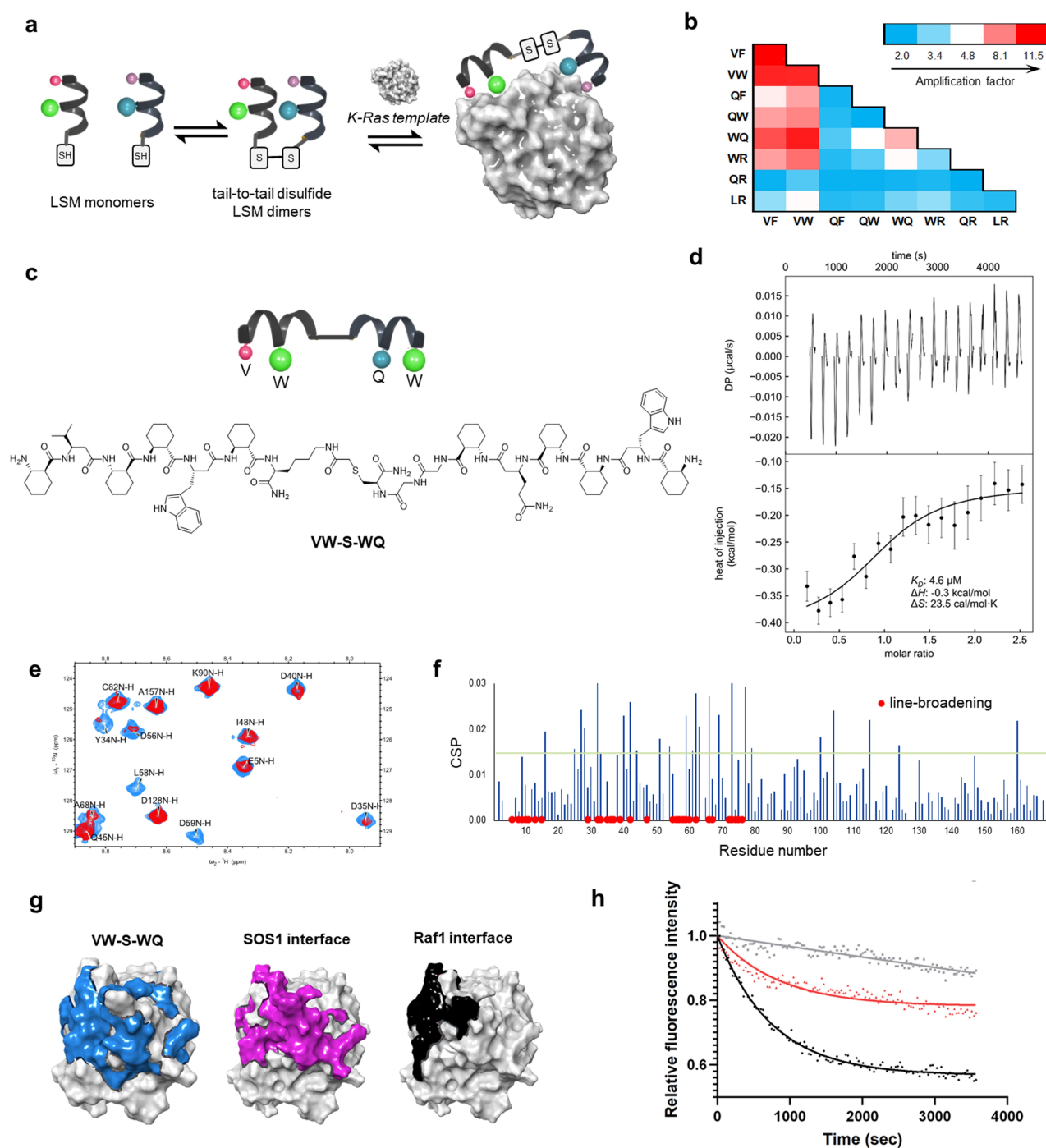


Figure 6. Affinity maturation by dynamic covalent LSM ligation for K-Ras G12D. a) General scheme for the protein-templated dynamic combinatorial library (DCL) optimization of LSM dimers using thiol–disulfide exchange driven to equilibrium. b) Amplification factors (concentration ratios of the LSM dimers with and without K-Ras) obtained in the DCL experiments represented as a heat map. Rows and columns represent the side chains of the corresponding monomeric LSM fragments. Blue and red colors indicate low and high amplifications, respectively. c) Schematic helical representation and chemical structure of the synthesized dimeric LSM with the stable thioether linkage: **VW-S-WQ**. d) ITC results obtained for **VW-S-WQ**. e) Selected region of the overlaid ^{15}N HSQC spectra for the free K-Ras (blue) and for K-Ras in complex with **VW-S-WQ** (red). Samples contained 20 μM K-Ras G12D and 40 μM **VW-S-WQ** (20 mM pH 7.4 HEPES, 50 mM NaCl, 10 mM MgCl_2 , 5% DMSO). f) CSP values of K-Ras upon complexation with **VW-S-WQ**. Residues with significant line-broadening are labeled with a red dot. g) Residues with above-average CSPs and line-broadenings mapped on the surface of K-Ras G12D (PDB: 4OBE). As a comparison, SOS1 and Raf1 interaction sites are shown (residues within 4 Å from SOS1 and Raf1 in PDB structures 1NVW and 8DGT, respectively). h) In K-Ras nucleotide exchange experiments, **VW-S-WQ** facilitates the displacement of MANT-labeled GDP from K-Ras, even without SOS1. The foldamer was applied at 100 μM (red curve) to mimic the binding site coverage of SOS1 safely. Black and grey curves correspond to SOS1 (0.5 μM) and DMSO controls, respectively.

To validate these findings, we synthesized the **VW-WQ** dimer with a redox-stable thioether linkage (**VW-S-WQ**) (Figure 6c) and subjected it to binding studies with K-Ras G12D. ITC revealed a low micromolar K_d for **VW-S-WQ** (Figure 6d), confirming that the DCL ligation approach improved the affinity by two orders of magnitude. ^{15}N HSQC NMR spectra were acquired for the K-Ras-**VW-S-WQ** complex. Compared with the control spectrum of K-Ras G12D, chemical shift perturbations (CSPs) and significant line broadenings were observed (Figure 6e-f), which we mapped to the surface of the target. (Figure 6g). The **VW-S-WQ** binding site encompassed the switch I and switch II regions. Therefore, it overlaps with the catalytically significant K-Ras interface. To test if the binding of **VW-S-WQ** to K-Ras G12D has any functional relevance, K-Ras G12D - SOS1 nucleotide exchange assay was performed in the presence of the foldamer dimer. Surprisingly, **VW-S-WQ** did not inhibit the SOS1-induced replacement of MANT-GDP as we expected based on the low micromolar K_d , the location of the binding site, and the previously confirmed inhibitory potential of **ph-VW** (Figure 4b). Thus, we carried out the assay without the assisting protein SOS1, which revealed that **VW-S-WQ** catalyzes MANT-GDP displacement from K-Ras (Figure 6h). SOS1 facilitates nucleotide release by detaching switch I from the K-Ras core via inserting its helix (residues 931–943) between the nucleotide and the switch. Our findings suggest that not only does the dimer LSM adhere to the SOS1 binding site, but its binding mode might also be similar to that of the release factor (Figure S10). In accordance with the observed effect, the signals of residues from the P-loop and switch I and II were significantly broadened in the HSQC spectra of the **VW-S-WQ** – K-Ras G12D complex. The results suggest that while the monomeric **ph-VW** LSM attaches to a functionally significant patch, the dimeric variant is capable of mimicking protein-like functions at the same site.

Discussion

We proposed that proteomimetics against challenging PPI targets with shallow, solvent-exposed, and dynamic binding sites can be developed by a bottom-up strategy, including a dynamic adaptive selection that enables the sorting of foldamer helix dimers according to their affinity.

As a first step, we combined local surface mimetic (LSM) foldamer fragments^[21] with the expedient photocrosslinking approach. Photoaffinity labeling is widely used for the sensitive detection and mapping of weak interactions, including small molecular ligand screening^[23b,30b, 36] and chemoproteomic applications.^[30d,37] We anticipated that LSMs containing two proteinogenic side chains recognize shallow sites with sufficient affinity, and the diazirine-tagged primary hits could be identified and localized by covalent capturing. As a proof of concept, screening with photo-LSMs resulted in several micromolar hits for the challenging targets STAT3 and K-Ras G12D.

Regarding crosslinking yields, labeling with photo-LSMs proved more efficient than reported for small molecular

fragments using similar conditions.^[23b] We suppose that the slower off-rates (k_{off}) of the LSM fragments due to the larger molecular weight (MW ~ 1000 Da) and the solvent shielding provided by the bulky ACHC residues resulted in longer residence times for the ligands at the protein surface leading to improved crosslinking yields. In contrast to conventional small molecular fragments, which are generally small and polar, photo-LSM hits with hydrophobic side chains were also identified. Considering that hydrophobic interactions play a key role in the stabilization of PPIs, these surface fragments can be viable starting points in the design of PPI inhibitor proteomimetics.

CaM, STAT3 and K-Ras G12D bound photo-LSM sets with different side chain combinations, proving the target selectivity of the hit fingerprints. The heat map patterns and the localization of the hits demonstrated that the interaction of the photo-foldamer probes with the target proteins is determined by the proteinogenic R_1 and R_2 side chains. This was further confirmed by the signals of proteinogenic side chains in the STD NMR spectra, indicating their proximity to the protein. Although the rigid and bulky cyclic side chains are essential for the helical fold of the scaffold and solvent shielding, direct ACHC – protein contacts do not drive the binding to the target.

Photo-LSM hits identified the functionally relevant PPI interfaces and the known ligandable sites and revealed other previously unexplored binding regions of the targets. LSMs mapped several new targetable sites on STAT3. While small molecule ligands of STAT3 typically bind to the phosphotyrosine binding pocket of the SH2 domain,^[26,38] LSMs labeled a slightly different site closer to the C-terminus of the protein. Moreover, foldamers recognized different segments in the coiled-coil domain, including the binding site of the MS3-6 monobody,^[39] and residues adjacent to that of the covalent ligand 15-keto PGE₂.^[34] As these compounds inhibit STAT3 signaling, binding photo-LSMs to similar regions might also have functional consequences. The photo-LSM sites on the surface of the GDP-loaded K-Ras G12D covered the overlapping PPI interfaces of the effectors and downstream partners (guanine nucleotide exchange factor SOS1, GAP, Raf1, PIK3CG or RALGDS) of the enzyme. Additionally, LSMs located the two shallow pockets which have proven targetable thus far: the region between switch I and II, which is the site of the nanomolar K-Ras inhibitor BI-2852;^[40] and the groove between switch II and $\alpha 3$, the binding site of the non-covalent small-molecular inhibitors,^[41] cyclic peptides,^[42] and the covalent inhibitors of K-Ras G12 C.^[43] A recent sweeping study of all possible single mutational variants of K-Ras and their altered substrate binding potential identified four allosterically active surface pockets, including the mostly unexplored ‘pocket 3’.^[44] Our photo-foldamer approach also located these sites, where ‘pocket 3’ overlaps with site C1 in our work. The most important allosteric sites (G10, G15, S17, D57, F78, P110, and V112)^[44] that influence the K-Ras interaction with all six of its protein partners are included in our sites C5 and C7. This finding indicates that photo-foldamers indeed explored the functionally relevant regions of the enzyme.

After identifying and localizing the primary hits, we linked the corresponding proteomimetic building blocks in a protein-templated DCL experiment. K-Ras G12D drove the adaptive formation of dimeric LSMs, which are compatible with its dynamic surface. Thus, the optimal combination of the four proteinogenic side chains could be selected. The resulting helix dimer **VW-S-WQ** catalyzed MANT-GDP displacement from K-Ras G12D, even without SOS1. Recently, Nanobody14^[45] and its carefully optimized loop mimetics^[13b] were shown to enhance the SOS1-catalyzed nucleotide exchange on Ras. While these activators acted via an allosteric mechanism, we propose that **VW-S-WQ** mimics SOS1 binding to K-Ras. The results demonstrate that following a bottom-up approach, a low micromolar proteomimetic compound could be generated even for such a challenging target as K-Ras. The dimeric LSMs can be viable chemical starting points for further medicinal chemistry optimizations.

In systems chemistry, peptides are used as building blocks of receptors,^[19a,46] self-templating tertiary,^[47] quaternary^[48] and supramolecular assemblies^[19a,46, 49] rather than as ligands. Our experiment using protein-guided DCL with peptidic foldamer building blocks confirms that analogous to small molecular ligand development, biomimetic components can be matured, and selecting proteomimetic ligands with advanced functions (target binding) is feasible. Thus, the presented bottom-up proteomimetics development strategy provides a chemical alternative to directed evolution methods, which can explore large sequence spaces using libraries of 10⁶ to 10¹² compounds.

The strategy can be generally applied to any protein and enables de novo development of interaction partners without any *a priori* structural hypothesis concerning the binding process. In contrast, the engineering of typical proteomimetics requires target-specific input and relies on extensive structure-based optimizations. Our approach generates viable starting points for challenging proteins with shallow and extended sites that are not targetable by conventional drug molecules. Strikingly, high hit rate, side chain selectivity, and effective surface coverage could be achieved with only a 100-membered two-side chain photo-foldamer set. This finding is in line with the limited structural diversity of the PPI interfaces,^[20b,50] a similar phenomenon to the homology of the tertiary protein structures. The evolutionarily conserved set of recognition elements is also reflected in the SpotXplorer optimized fragment pharmacophore technology, which was also developed to maximize the hit rate and pharmacophore coverage for small molecular fragments.^[51]

Conclusion

Our results show the feasibility of the bottom-up engineering of proteomimetics. The photo-foldamer approach offers advances in the chemical biology, medicinal chemistry and structural biology aspects of targeting PPI sites. The method involves the integration of foldamer chemistry with the photoaffinity-based detection of drug-like molecules to

achieve a tractable surface mapping machinery. It allows building modest-sized foldamer fragment libraries to screen against shallow, dynamic, and often hydrophobic PPI interfaces effectively. Straightforward analytical techniques, such as MS/MS analysis and HSQC NMR measurements, provide sufficient data for the identification of the binding sites on the surface of the target. Thus, the approach explores potential binding sites with high efficiency and limited synthetic efforts. Primary hits can be optimized by using the systems chemistry toolbox. The dimeric LSM variant designed against K-Ras could mimic the function of a protein interacting partner (SOS1) at a crucial target PPI. Peptidic LSMs are synthetically accessible, versatile and robustly designable; side chain and linking chemistries can be varied. We believe the proposed strategy can complement the current structure-based proteomimetics design strategies to target disease-relevant PPIs, providing a unique, protein-mimetic platform for creating surface-complementing topologies, while allowing the fine-tuning of local interactions and anchoring sites by the interchangeable proteinogenic amino acids.

Supporting Information

The authors have cited additional references within the Supporting Information.^[29a,52–71]

Acknowledgements

This research was funded by the National Research, Development and Innovation Office of Hungary, Grant Numbers GINOP-2.2.1-15-2016-00007 (T.A.M.) and NKFIH K134754. (T.A.M.), PharmaLab (RRF-2.3.1-21-2022-00015 and 2020–1.1.6-JÖVŐ-2021-00004 (G.M.K. and L.B.), NKFIH K135150 (G.M.K.), NKFIH FK146063 (D.B.), 2018–1.2.1-NKP-2018-00005 (HunProtExc) financed under the 2018–1.2.1-NKP funding Scheme. The Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund supported the KDP-2021 Program (T.T.). Project no TKP2021-EGA-32 has been implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021-EGA funding scheme (T.A.M.). Project number RRF-2.3.1-21-2022-00015 is implemented with the support of the European Union's Recovery and Resilience Instrument. G.S. was supported by the Lendület Program of the Hungarian Academy of Sciences. This research work was conducted with the support of the National Academy of Scientist Education Program of the National Biomedical Foundation under the sponsorship of the Hungarian Ministry of Culture and Innovation. The work of D.B. was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. The work was supported by the University of Szeged Open Access Fund (Grant Number: 7311).

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords: foldamer · fragment screening · photoaffinity labeling · protein design · protein-protein interaction

- [1] a) W. S. Horne, T. N. Grossmann, *Nat. Chem.* **2020**, *12*, 331–337; b) C. Durukan, T. N. Grossmann, *Synlett* **2024**.
- [2] a) M. Pelay-Gimeno, T. Bange, S. Hennig, T. N. Grossmann, *Angew. Chem. Int. Ed.* **2018**, *57*, 11164–11170; b) S. A. Kerns, A. Biswas, N. M. Minnetian, A. S. Borovik, *JACS Au* **2022**, *2*, 1252–1265; c) L. Marchetti, M. Levine, *ACS Catal.* **2011**, *1*, 1090–1118.
- [3] A. Barnard, K. Long, H. L. Martin, J. A. Miles, T. A. Edwards, D. C. Tomlinson, A. Macdonald, A. J. Wilson, *Angew. Chem. Int. Ed.* **2015**, *54*, 2960–2965.
- [4] a) N. Imre, A. Hetényi, E. Szabó, B. Bodnár, A. Szkalitsy, I. Gróf, A. Bocsik, M. A. Deli, P. Horvath, Á. Czibula, É. Monostori, T. A. Martinek, *Adv. Sci. (Weinheim, Ger.)* **2020**, *7*, 1902621; b) A. Hetényi, E. Szabo, N. Imre, K. N. Bhaumik, A. Tököl, T. Füzesi, R. Hollandi, P. Horvath, A. Czibula, É. Monostori, M. A. Deli, T. A. Martinek, *Pharmaceutics* **2022**, *14*.
- [5] G. H. Hutchins, S. Kiehstaller, P. Poc, A. H. Lewis, J. Oh, R. Sadighi, N. M. Pearce, M. Ibrahim, I. Drienovská, A. M. Rijs, S. Neubacher, S. Hennig, T. N. Grossmann, *Chem* **2024**, *10*.
- [6] a) S. H. Hong, D. Y. Yoo, L. Conway, K. C. Richards-Corke, C. G. Parker, P. S. Arora, *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118*; b) M. G. Wuo, A. B. Mahon, P. S. Arora, *J. Am. Chem. Soc.* **2015**, *137*, 11618–11621; c) S. Jeganathan, M. Wendt, S. Kiehstaller, D. Brancaccio, A. Kuepper, N. Pospiech, A. Carotenuto, E. Novellino, S. Hennig, T. N. Grossmann, *Angew. Chem. Int. Ed.* **2019**, *58*, 17351–17358; d) H. Adihou, R. Gopalakrishnan, T. Forster, S. M. Gueret, R. Gasper, S. Geschwindner, C. Carrillo Garcia, H. Karatas, A. V. Pobbati, M. Vazquez-Chantada, P. Davey, C. M. Wassvik, J. K. S. Pang, B. S. Soh, W. Hong, E. Chiarparin, D. Schade, A. T. Plowright, E. Valeur, M. Lemurell, T. N. Grossmann, H. Waldmann, *Nat. Commun.* **2020**, *11*, 5425; e) F. M. Paulussen, G. K. Schouten, C. Moertl, J. Verheul, I. Hoekstra, G. M. Koningstein, G. H. Hutchins, A. Alkir, R. A. Luirink, D. P. Geerke, P. van Ulsen, T. den Blaauwen, J. Luirink, T. N. Grossmann, *J. Am. Chem. Soc.* **2022**, *144*, 15303–15313; f) M. Wendt, R. Bellavita, A. Gerber, N. L. Efreim, T. van Ramshorst, N. M. Pearce, P. R. J. Davey, I. Everard, M. Vazquez-Chantada, E. Chiarparin, P. Grieco, S. Hennig, T. N. Grossmann, *Angew. Chem. Int. Ed.* **2021**, *60*, 13937–13944.
- [7] H. M. Werner, C. C. Cabaltea, W. S. Horne, *ChemBioChem* **2016**, *17*, 712–718.
- [8] a) W. S. Horne, S. H. Gellman, *Acc. Chem. Res.* **2008**, *41*, 1399–1408; b) W. S. Horne, M. D. Boersma, M. A. Windsor, S. H. Gellman, *Angew. Chem. Int. Ed.* **2008**, *47*, 2853–2856; c) H. S. Haase, K. J. Peterson-Kaufman, S. K. Lan Levengood, J. W. Checco, W. L. Murphy, S. H. Gellman, *J. Am. Chem. Soc.* **2012**, *134*, 7652–7655; d) U. Arnold, M. P. Hinderaker, B. L. Nilsson, B. R. Huck, S. H. Gellman, R. T. Raines, *J. Am. Chem. Soc.* **2002**, *124*, 8522–8523.
- [9] a) C. M. Lombardo, M. V. V. Kumar, C. Douat, F. Rosu, J. L. Mergny, G. F. Salgado, G. Guichard, *J. Am. Chem. Soc.* **2019**, *141*, 2516–2525; b) K. L. George, W. S. Horne, *J. Am. Chem. Soc.* **2017**, *139*, 7931–7938.
- [10] R. K. Spencer, H. Li, J. S. Nowick, *J. Am. Chem. Soc.* **2014**, *136*, 5595–5598.
- [11] S. K. Mong, F. V. Cochran, H. Yu, Z. Graziano, Y. S. Lin, J. R. Cochran, B. L. Pentelute, *Biochemistry* **2017**, *56*, 5720–5725.
- [12] G. M. Burslem, H. F. Kyle, A. L. Breeze, T. A. Edwards, A. Nelson, S. L. Warriner, A. J. Wilson, *Chem. Commun. (Cambridge, U. K.)* **2016**, *52*, 5421–5424.
- [13] a) K. Van Holsbeeck, J. C. Martins, S. Ballet, *Bioorg. Chem.* **2022**, *119*, 105563; b) K. Van Holsbeeck, B. Fischer, S. Gonzalez, C. Gadais, W. Versees, J. C. Martins, C. Martin, A. Wohlkonig, J. Steyaert, S. Ballet, *Angew. Chem. Int. Ed.* **2023**, *62*, e202219095; c) R. U. Kadam, J. Juraszek, B. Brandenburg, C. Buyck, W. B. G. Schepens, B. Kesteleyn, B. Stoops, R. J. Vreeken, J. Vermond, W. Goutier, C. Tang, R. Vogels, R. H. E. Friesen, J. Goudsmit, M. J. P. van Dongen, I. A. Wilson, *Science (New York, N. Y.)* **2017**, *358*, 496–502.
- [14] G. Tuchscherer, M. Mutter, *Pure Appl. Chem.* **1996**, *68*, 2153–2162.
- [15] a) D. S. Daniels, E. J. Petersson, J. X. Qiu, A. Schepartz, *J. Am. Chem. Soc.* **2007**, *129*, 1532–1533; b) A. G. Tebo, V. L. Pecoraro, *Curr. Opin. Chem. Biol.* **2015**, *25*, 65–70; c) E. Merljak, B. Malovrh, R. Jerala, *Nat. Commun.* **2023**, *14*; d) P. Kumar, N. G. Paterson, J. Clayden, D. N. Woolfson, *Nature* **2022**, *607*, 387–392.
- [16] a) T. M. Chidyausiku, S. R. Mendes, J. C. Klima, M. Nadal, U. Eckhard, J. Roel-Touris, S. Houlston, T. Guevara, H. K. Haddox, A. Moyer, C. H. Arrowsmith, F. X. Gomis-Rüth, D. Baker, E. Marcos, *Nat. Commun.* **2022**, *13*; b) Y. B. Yan, B. W. Erickson, *Protein Sci.* **1994**, *3*, 1069–1073; c) T. P. Quinn, N. B. Tweedy, R. W. Williams, J. S. Richardson, D. C. Richardson, *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 8747–8751.
- [17] a) K. Ozga, L. Berlicki, *ACS Bio. Med. Chem. Au* **2022**, *2*, 316–327; b) A. Ciesiolkiewicz, J. Lizandra Perez, L. Berlicki, *Bioorg. Med. Chem. Lett.* **2022**, *71*; c) J. B. Case, R. E. Chen, L. X. Cao, B. L. Ying, E. S. Winkler, M. Johnson, I. Goresnik, M. N. Pham, S. Shrihari, N. M. Kafai, A. L. Bailey, X. P. Xie, P. Y. Shi, R. Ravichandran, L. Carter, L. Stewart, D. Baker, M. S. Diamond, *Cell Host Microbe* **2021**, *29*, 1151–1161; d) S. Learte-Aymamí, C. Vidal, A. Gutiérrez-González, J. L. Mascareñas, *Angew. Chem. Int. Ed.* **2020**, *59*, 9149–9154; e) A. Roy, L. Shi, A. Chang, X. Dong, A. Fernandez, J. C. Kraft, J. Li, V. Q. Le, R. V. Winegar, G. M. Cherf, D. Slocum, P. D. Poulson, G. E. Casper, M. L. Vallecillo-Zuniga, J. C. Valdoz, M. C. Miranda, H. Bai, Y. Kipnis, A. Olshefsky, T. Priya, L. Carter, R. Ravichandran, C. M. Chow, M. R. Johnson, S. Cheng, M. Smith, C. Overed-Sayer, D. K. Finch, D. Lowe, A. K. Bera, G. Matute-Bello, T. P. Birkland, F. DiMaio, G. Raghu, J. R. Cochran, L. J. Stewart, M. G. Campbell, P. M. Van Ry, T. Springer, D. Baker, *Nat. Commun.* **2023**, *14*, 5660.
- [18] B. Dang, H. Wu, V. K. Mulligan, M. Mravic, Y. Wu, T. Lemmin, A. Ford, D. A. Silva, D. Baker, W. F. DeGrado, *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 10852–10857.
- [19] a) J. Li, P. Nowak, S. Otto, *J. Am. Chem. Soc.* **2013**, *135*, 9222–9239; b) P. T. Corbett, S. Otto, J. K. Sanders, *Chemistry* **2004**, *10*, 3139–3143; c) P. T. Corbett, S. Otto, J. K. Sanders, *Org. Lett.* **2004**, *6*, 1825–1827; d) P. T. Corbett, J. K. Sanders, S. Otto, *J. Am. Chem. Soc.* **2005**, *127*, 9390–9392.
- [20] a) A. M. Watkins, R. Bonneau, P. S. Arora, *J. Am. Chem. Soc.* **2016**, *138*, 10386–10389; b) M. Gao, J. Skolnick, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 22517–22522.
- [21] A. Tököl, B. Mag, É. Bartus, E. Wéber, G. Szakonyi, M. A. Simon, Á. Czibula, É. Monostori, L. Nyitray, T. A. Martinek, *Chem. Sci.* **2020**, *11*, 10390–10398.

- [22] M. A. Simon, É. Bartus, B. Mag, E. Boros, L. Roszjár, G. Gógl, G. Travé, T. A. Martinek, L. Nyitray, *Sci. Rep.* **2022**, *12*.
- [23] a) E. Smith, I. Collins, *Future Med. Chem.* **2015**, *7*, 159–183; b) E. K. Grant, D. J. Fallon, M. M. Hann, K. G. M. Fantom, C. Quinn, F. Zappacosta, R. S. Annan, C. W. Chung, P. Bam-borough, D. P. Dixon, P. Stacey, D. House, V. K. Patel, N. C. O. Tomkinson, J. T. Bush, *Angew. Chem. Int. Ed.* **2020**, *59*, 21096–21105.
- [24] É. Bartus, Z. Hegedüs, E. Wéber, B. Csipak, G. Szakonyi, T. A. Martinek, *ChemistryOpen* **2017**, *6*, 236–241.
- [25] B. Wingelhofer, H. A. Neubauer, P. Valent, X. Han, S. N. Constantinescu, P. T. Gunning, M. Muller, R. Moriggl, *Leuke-mia* **2018**, *32*, 1713–1726.
- [26] A. Orlova, C. Wagner, E. D. de Araujo, D. Bajusz, H. A. Neubauer, M. Herling, P. T. Gunning, G. M. Keseru, R. Moriggl, *Cancers* **2019**, *11*.
- [27] I. A. Prior, F. E. Hood, J. L. Hartley, *Cancer Res.* **2020**, *80*, 2969–2974.
- [28] G. Palfy, D. K. Menyhard, A. Perczel, *Cancer Metastasis Rev.* **2020**, *39*, 1075–1089.
- [29] a) D. H. Appella, L. A. Christianson, I. L. Karle, D. R. Powell, S. H. Gellman, *J. Am. Chem. Soc.* **1996**, *118*, 13071–13072; b) S. H. Gellman, *Acc. Chem. Res.* **1998**, *31*, 173–180.
- [30] a) Y. Liao, S. C. Chan, E. A. Welsh, B. Fang, L. Sun, E. Schonbrunn, J. M. Koomen, D. R. Duckett, E. B. Haura, A. Monastyrskyi, U. Rix, *ACS Chem. Biol.* **2023**, *18*, 251–264; b) 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, B. F. Cravatt, *Nat. Chem.* **2019**, *11*, 1113–1123; c) R. P. Thomas, R. E. Heap, F. Zappacosta, E. K. Grant, P. Pogany, S. Besley, D. J. Fallon, M. M. Hann, D. House, N. C. O. Tomkinson, J. T. Bush, *Chem. Sci.* **2021**, *12*, 12098–12106; d) F. Offensperger, G. Tin, M. Duran-Frigola, E. Hahn, S. Dobner, C. W. A. Ende, J. W. Strohbach, A. Rukavina, V. Brennstainer, K. Ogilvie, N. Marella, K. Kladnik, R. Ciuffa, J. D. Majmudar, S. D. Field, A. Bensimon, L. Ferrari, E. Ferrada, A. Ng, Z. Zhang, G. Degliesposti, A. Boeszoermyenyi, S. Martens, R. Stanton, A. C. Muller, J. T. Hannich, D. Hepworth, G. Superti-Furga, S. Kubicek, M. Schenone, G. E. Winter, *Science (New York, N. Y.)* **2024**, *384*, eadk5864.
- [31] C. Yan, F. Wu, R. L. Jernigan, D. Dobbs, V. Honavar, *Protein J.* **2008**, *27*, 59–70.
- [32] a) G. M. Keserü, D. A. Erlanson, G. G. Ferenczy, M. M. Hann, C. W. Murray, S. D. Pickett, *J. Med. Chem.* **2016**, *59*, 8189–8206; b) C. W. Murray, D. C. Rees, *Nat. Chem.* **2009**, *1*, 187–192.
- [33] L. Fülöp, I. M. Mándity, G. Juhász, V. Szegedi, A. Hetényi, E. Wéber, Z. Bozsó, D. Simon, R. Benko, Z. Király, T. A. Martinek, *PLoS One* **2012**, *7*.
- [34] E. J. Lee, S. J. Kim, Y. I. Hahn, H. J. Yoon, B. Han, K. Kim, S. Lee, K. P. Kim, Y. G. Suh, H. K. Na, Y. J. Surh, *Redox Biol.* **2019**, *23*, 101175.
- [35] J. Schust, T. Berg, *Anal. Biochem.* **2004**, *330*, 114–118.
- [36] a) Y. Liao, S. C. Chan, E. A. Welsh, B. Fang, L. X. Sun, E. Schonbrunn, J. M. Koomen, D. R. Duckett, E. B. Haura, A. Monastyrskyi, U. Rix, *ACS Chem. Biol.* **2023**, *18*, 251–264; b) A. Galmozzi, C. G. Parker, B. P. Kok, B. F. Cravatt, E. Saez, *Methods Mol. Biol.* **2018**, *1787*, 115–127; c) 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, B. F. Cravatt, *Cell* **2017**, *168*, 527–541.
- [37] a) E. Karaj, S. H. Sindi, L. M. Viranga Tillekeratne, *Bioorg. Med. Chem.* **2022**, *62*, 116721; b) M. W. Halloran, J. P. Lumb, *Chemistry* **2019**, *25*, 4885–4898; c) J. X. Gao, A. Mfuh, Y. Amako, C. M. Woo, *J. Am. Chem. Soc.* **2018**, *140*, 4259–4268; d) Y. Zhang, T. G. Kutateladze, *Nat. Chem.* **2020**, *12*, 506–508; e) J. R. Hill, A. A. B. Robertson, *J. Med. Chem.* **2018**, *61*, 6945–6963.
- [38] Y. Hua, X. Yuan, Y. H. Shen, J. Wang, W. Azeem, S. Yang, A. Gade, S. M. Lellahi, A. M. Oyan, X. Ke, W. D. Zhang, K. H. Kalland, *Front. Pharmacol.* **2022**, *13*, 836724.
- [39] G. La Sala, C. Michiels, T. Kukenshoner, T. Brandstoetter, B. Maurer, A. Koide, K. Lau, F. Pojer, S. Koide, V. Sexl, L. Dumoutier, O. Hantschel, *Nat. Commun.* **2020**, *11*, 4115.
- [40] D. Kessler, M. Gmachl, A. Mantoulidis, L. J. Martin, A. Zoephel, M. Mayer, A. Gollner, D. Covini, S. Fischer, T. Gerstberger, T. Gmaschitz, C. Goodwin, P. Greb, D. Haring, W. Hela, J. Hoffmann, J. Karolyi-Oezguer, P. Knesl, S. Kornigg, M. Koegl, R. Kousek, L. Lamarre, F. Moser, S. Munico-Martinez, C. Peinsipp, J. Phan, J. Rinnenthal, J. Sai, C. Salamon, Y. Scherbantini, K. Schipany, R. Schnitzer, A. Schrenk, B. Sharps, G. Sizler, Q. Sun, A. Waterson, B. Wolkerstorfer, M. Zeeb, M. Pearson, S. W. Fesik, D. B. McConnell, *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 15823–15829.
- [41] X. Wang, S. Allen, J. F. Blake, V. Bowcut, D. M. Briere, A. Calinisan, J. R. Dahlke, J. B. Fell, J. P. Fischer, R. J. Gunn, J. Hallin, J. Laguer, J. D. Lawson, J. Medwid, B. Newhouse, P. Nguyen, J. M. O'Leary, P. Olson, S. Pajk, L. Rahbaek, M. Rodriguez, C. R. Smith, T. P. Tang, N. C. Thomas, D. Vanderpool, G. P. Vigers, J. G. Christensen, M. A. Marx, *J. Med. Chem.* **2022**, *65*, 3123–3133.
- [42] a) A. Ohta, M. Tanada, S. Shinohara, Y. Morita, K. Nakano, Y. Yamagishi, R. Takano, S. Kariyuki, T. Iida, A. Matsuo, K. Ozeki, T. Emura, Y. Sakurai, K. Takano, A. Higashida, M. Kojima, T. Muraoka, R. Takeyama, T. Kato, K. Kimura, K. Ogawa, K. Ohara, S. Tanaka, Y. Kikuchi, N. Hisada, R. Hayashi, Y. Nishimura, K. Nomura, T. Tachibana, M. Irie, H. Kawada, T. Torizawa, N. Mura, T. Kotake, M. Tanaka, S. Ishikawa, T. Miyake, M. Tamiya, M. Arai, A. Chiyoda, S. Akai, H. Sase, S. Kuramoto, T. Ito, T. Shiraishi, T. Kojima, H. Iikura, *J. Am. Chem. Soc.* **2023**, *145*, 24035–24051; b) M. Tanada, M. Tamiya, A. Matsuo, A. Chiyoda, K. Takano, T. Ito, M. Irie, T. Kotake, R. Takeyama, H. Kawada, R. Hayashi, S. Ishikawa, K. Nomura, N. Furuichi, Y. Morita, M. Kage, S. Hashimoto, K. Nii, H. Sase, K. Ohara, A. Ohta, S. Kuramoto, Y. Nishimura, H. Iikura, T. Shiraishi, *J. Am. Chem. Soc.* **2023**, *145*, 16610–16620.
- [43] a) J. Canon, K. Rex, A. Y. Saiki, C. Mohr, K. Cooke, D. Bagal, K. Gaida, T. Holt, C. G. Knutson, N. Koppada, B. A. Lanman, J. Werner, A. S. Rapaport, T. San Miguel, R. Ortiz, T. Osgood, J. R. Sun, X. Zhu, J. D. McCarter, L. P. Volak, B. E. Houk, M. G. Fakih, B. H. O'Neil, T. J. Price, G. S. Falchook, J. Desai, J. Kuo, R. Govindan, D. S. Hong, W. Ouyang, H. Henary, T. Arvedson, V. J. Cee, J. R. Lipford, *Nature* **2019**, *575*, 217–223; b) J. B. Fell, J. P. Fischer, B. R. Baer, J. F. Blake, K. Bouhana, D. M. Briere, K. D. Brown, L. E. Burgess, A. C. Burns, M. R. Burkard, H. Chiang, M. J. Chicarella, A. W. Cook, J. J. Gaudino, J. Hallin, L. Hanson, D. P. Hartley, E. J. Hicken, G. P. Hingorani, R. J. Hinklin, M. J. Mejia, P. Olson, J. N. Otten, S. P. Rhodes, M. E. Rodriguez, P. Savechenkov, D. J. Smith, N. Sudhakar, F. X. Sullivan, T. P. Tang, G. P. Vigers, L. Wollenberg, J. G. Christensen, M. A. Marx, *J. Med. Chem.* **2020**, *63*, 6679–6693.
- [44] C. Weng, A. J. Faure, A. Escobedo, B. Lehner, *Nature* **2024**, *626*, 643–652.
- [45] B. Fischer, T. Uchański, A. Sheryazdanova, S. Gonzalez, A. N. Volkov, E. Brosens, T. Zögg, V. Kalichuk, S. Ballet, W. Versées, A. A. Sablina, E. Pardon, A. Wohlkönig, J. Steyaert, *bioRxiv* **2023**, 10.1101/2022.12.06.519258.
- [46] a) F. Sheehan, D. Sementa, A. Jain, M. Kumar, M. Tayarani-Najjaran, D. Kroiss, R. V. Ulijn, *Chem. Rev.* **2021**, *121*, 13869–13914; b) E. E. Harrison, B. A. Carpenter, L. E. St Louis, A. G.

- Mullins, M. L. Waters, *J. Am. Chem. Soc.* **2021**, *143*, 14845–14854; c) A. Chatterjee, A. Reja, S. Pal, D. Das, *Chem. Soc. Rev.* **2022**, *51*, 3047–3070.
- [47] a) B. F. Fisher, S. H. Hong, S. H. Gellman, *J. Am. Chem. Soc.* **2017**, *139*, 13292–13295; b) M. G. Woll, S. H. Gellman, *J. Am. Chem. Soc.* **2004**, *126*, 11172–11174.
- [48] J. F. Reuther, J. L. Dees, I. V. Kolesnichenko, E. T. Hernandez, D. V. Ukraintsev, R. Guduru, M. Whiteley, E. V. Anslyn, *Nat. Chem.* **2018**, *10*, 45–50.
- [49] a) H. V. Schroder, Y. Zhang, A. J. Link, *Nat. Chem.* **2021**, *13*, 850–857; b) A. Rodrigues, L. Rocard, R. Moumné, *ChemSysChem* **2023**, *5*.
- [50] a) L. Garma, S. Mukherjee, P. Mitra, Y. Zhang, *PLoS One* **2012**, *7*; b) E. Cukuroglu, A. Gursoy, R. Nussinov, O. Keskin, *PLoS One* **2014**, *9*.
- [51] D. Bajusz, W. S. Wade, G. Satala, A. J. Bojarski, J. Ilas, J. Ebner, F. Grebien, H. Papp, F. Jakab, A. Douangamath, D. Fearon, F. von Delft, M. Schuller, I. Ahel, A. Wakefield, S. Vajda, J. Gerencsér, P. Pallai, G. M. Keseru, *Nat. Commun.* **2021**, *12*.
- [52] S. B. Azatian, N. Kaur, M. P. Latham, *J. Biomol. NMR* **2019**, *73*, 11–17.
- [53] I. Vida, Z. Fazekas, G. Gyulai, D. Nagy-Fazekas, G. Pálffy, P. Stráner, É. Kiss, A. Perczel, *Microb. Biotechnol.* **2021**, *14*, 1107–1119.
- [54] T. M. M. J. Abraham, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, *Software X* **2015**, *1–2*, 19–25.
- [55] J. Huang, A. D. MacKerell, Jr., *J. Comput. Chem.* **2013**, *34*, 2135–2145.
- [56] A. Wacha, T. Beke-Somfai, T. Nagy, *ChemPlusChem* **2019**, *84*, 927–941.
- [57] S. Izadi, R. Anandakrishnan, A. V. Onufriev, *J. Phys. Chem. Lett.* **2014**, *5*, 3863–3871.
- [58] A. E. Aliev, M. Kulke, H. S. Khaneja, V. Chudasama, T. D. Sheppard, R. M. Lanigan, *Proteins* **2014**, *82*, 195–215.
- [59] M. T. Marty, A. J. Baldwin, E. G. Marklund, G. K. Hochberg, J. L. Benesch, C. V. Robinson, *Anal. Chem.* **2015**, *87*, 4370–4376.
- [60] F. C. Yu, G. C. Teo, A. T. Kong, S. E. Haynes, D. M. Avtonomov, D. J. Geiszler, A. I. Nesvizhskii, *Nat. Commun.* **2020**, *11*.
- [61] F. C. Yu, S. E. Haynes, A. I. Nesvizhskii, *Mol. Cell Proteomics* **2021**, *20*.
- [62] W. Lee, M. Rahimi, Y. Lee, A. Chiu, *Bioinformatics* **2021**, *37*, 3041–3042.
- [63] G. Pálffy, I. Vida, A. Perczel, *Biomol. NMR Assignments* **2020**, *14*, 1–7.
- [64] J. C. Hunter, D. Gurbani, S. B. Ficarro, M. A. Carrasco, S. M. Lim, H. G. Choi, T. Xie, J. A. Marto, Z. Chen, N. S. Gray, K. D. Westover, *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 8895–8900.
- [65] M. E. Welsch, A. Kaplan, J. M. Chambers, M. E. Stokes, P. H. Bos, A. Zask, Y. Zhang, M. Sanchez-Martin, M. A. Badgley, C. S. Huang, T. H. Tran, H. Akkiraju, L. M. Brown, R. Nandakumar, S. Cremers, W. S. Yang, L. Tong, K. P. Olive, A. Ferrando, B. R. Stockwell, *Cell* **2017**, *168*, 878–889.
- [66] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, J. Heming, *Information Sciences* **2023**, *622*, 178–210.
- [67] P. J. Rousseeuw, *J. Comput. Appl. Math.* **1987**, *20*, 53–65.
- [68] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, I. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, C. SciPy, *Nat. Methods* **2020**, *17*, 261–272.
- [69] T. H. Scheuermann, C. A. Brautigam, *Methods* **2015**, *76*, 87–98.
- [70] H. Y. Zhao, G. Piszczek, P. Schuck, *Methods* **2015**, *76*, 137–148.
- [71] P. A. Boriack-Sjodin, S. M. Margarit, D. Bar-Sagi, J. Kuriyan, *Nature* **1998**, *394*, 337–343.

Manuscript received: June 3, 2024

Revised manuscript received: August 23, 2024

Accepted manuscript online: September 27, 2024

Version of record online: November 6, 2024