VIP · *Very Important Paper*

Special Collection

Scanning Protein Surfaces with DNA-Encoded Libraries

Verena B. K. Kunig+,[a] Marco Potowski+,[a] Mateja Klika Škopić+,[a] and
Andreas Brunschweiger*[a]

Understanding the ligandability of a target protein, defined as the capability of a protein to bind drug-like compounds on any site, can give important stimuli to drug-development projects. For instance, inhibition of protein–protein interactions usually depends on the identification of protein surface binders. DNA-encoded chemical libraries (DELs) allow scanning of protein surfaces with large chemical space. Encoded library selection screens uncovered several protein–protein interaction inhibitors and compounds binding to the surface of G protein-coupled receptors (GPCRs) and kinases. The protein surface-binding chemotypes from DELs are predominantly chemically modified and cyclized peptides, and functional small-molecule peptido-mimetics. Peptoid libraries and structural peptidomimetics have been less studied in the DEL field, hinting at hitherto less populated chemical space and suggesting alternative library designs. Roughly a third of bioactive molecules evolved from smaller, target-focused libraries. They showcase the potential of encoded libraries to identify more potent molecules from weak, for example, fragment-like, starting points.

1. Introduction

Many physiological processes are regulated by direct interaction of proteins. Estimates of the size of the human interactome suggest a six-digit number of individual protein–protein interactions (PPIs, Figure 1a), and a recently published reference map of the human interactome reported 53 000 binary protein interactions.[1,2] These can be retrieved from databases such as HuRI, STRING, and BioGRID.[2–4] Several protein–protein interactions have been found to be associated with disease, providing molecular targets for intervention with therapeutic agents. A scant literature survey revealed the vast majority of well-investigated PPIs to be associated with malignant diseases, and intense drug-development efforts centered on this indication.[5–9] Yet, modulation of PPI holds promise for treatment of a much broader range of diseases, including devastating neurodegenerative disorders, and novel approaches to combat infectious diseases.[10–13] However, a highly influential opinion piece published two decades ago cast doubt on the feasibility of inhibiting PPIs with compounds that meet the requirements for peroral application and uptake into cells for cytosolic or nuclear target engagement.[14] The large binding interface, the occurrence of non-contiguous binding areas, the shallow surface of proteins involved in PPI,[15] and not the least the lack of starting points for rational drug design are formidable obstacles for inhibitor development (exemplified by

the TEAD/YAP interaction, Figure 1a).[16–24] Yet, a number of clinical stage PPI inhibitors (PPIi) and approved drugs such as the BcL-xL inhibitor Navitoclax show that this does not hold true to all disease-relevant PPIs.[16]

Current drug development efforts in this field are supported by a deeper understanding of PPIs at the molecular level. Large-scale analysis of alanine mutants revealed that in many cases protein–protein interactions critically depend on a few amino acid side chain interactions covering only a small area of the whole interface.[25–34] Such areas are called "hot spots", and they have been shown to be productive entry points for inhibitor design.[34–36] Further large scale structural analyses of proteins involved in PPIs pointed out that many of them contain cavities which could serve as locks for small organic molecule keys.[37–40] The properties of these cavities were found to differ from those formed by protein targets belonging to the "druggable genome",[37] that is, certain receptor and enzyme superfamilies such as the kinase and G protein-coupled receptor families. Likely, these cavities require chemical matter for binder/inhibitor development which is different from many of the small-molecule designs that evolved over decades of drug research on the "druggable genome".[14]

Today, the medicinal chemist has a diverse toolbox of chemical modalities and technologies available for protein binder identification, and PPIi development.[41] The modalities have been classified by Grossmann et al. They range from peptides and chemically modified, for example, cyclized peptides (class A), via alternative peptide-like oligomers such as peptoids (class B), to small molecules that are structural mimics of peptides (class C), and small molecules that mimic peptide functions (class D, Figure 1b).[42] We wish to introduce here one further class of peptidomimetics, a class E that describes small-molecule mimics of post-translationally modified peptides. For instance, orthosteric bromodomain inhibitors fall into this class. Important technologies for small-molecule protein binder identification at the disposal of medicinal chemists include structure-based (peptidomimetic) compound design exempli-

[a] V. B. K. Kunig,+ Dr. M. Potowski,+ Dr. M. Klika Škopić,+ Dr. A. Brunschweiger
Faculty of Chemistry and Chemical Biology
TU Dortmund University
Otto-Hahn-Straße 6, 44227 Dortmund (Germany)
E-mail: andreas.brunschweiger@tu-dortmund.de

[+] These authors have contributed equally.

fied by the *AnchorQuery* approach,[35,36,43,44] screening of fragment libraries by biophysical and spectroscopic methods,[45] the design of target-focused screening libraries (Figure 1c),[46,47] and, subject of this review, scanning protein surface with chemically synthesized DNA-encoded combinatorial libraries (Figure 1d). In this review, we will summarize encoded library technologies, encoded library designs, and describe successful identification of PPI inhibitors from encoded libraries. Beyond PPI targets, we will also show protease inhibitors, as these enzymes share with PPI targets the central feature of an extended binding surface, and compounds that revealed allosteric binding sites on the surface of GPCRs and kinases, that is, "druggable genome" targets. Throughout the review, we will point out where the aforementioned technologies and encoded library technology were used in a synergistic manner for encoded library design, and DEL screening hit elaboration.

2. Encoded Libraries

DNA-encoded libraries, typically abbreviated DELs, are a technology for target-based screening that relies on phenotype-genotype coupling (Figure 2a).[48–52] It is related to display technologies such as phage and RNA display, but uses organic preparative chemistry for the combinatorial synthesis of large numbers of encoded molecules.[53] For identification of bioactive molecules, encoded libraries are typically selected on tagged recombinant proteins immobilized on a surface (Figure 1d), though alternative selection formats in solution or in cells involving covalent capture of compounds or enzymatic steps have been shown.[52] Encoded one-bead/one-compound libraries even offer the opportunity to perform functional screens in miniaturized assays.[52] Over the last three decades several

encoded library formats have been introduced that shall be summarized below.

2.1. DNA-encoded solution phase combinatorial chemistry (I)

The currently most common format for DEL synthesis is the solution phase split-and-pool approach introduced by Neri and Morgan (Figure 2b).[54,55] This approach relies on concatenating short DNA oligomers containing genetic information for the chemical building blocks that are coupled to build up a DEL in a way that records library synthesis history. In the first synthesis cycle, a short, linker-modified single-stranded,[54] or, as adopted by most users in the industry,[55] a duplex DNA called "headpiece" is split and a first set of building blocks are coupled to the "headpiece" DNA followed by the ligation of the corresponding DNA codes. Afterwards, all products are pooled into a single vessel, and split for the next cycle of encoding and synthesis. Exponential library growth over 2–4 cycles and massive parallelization at each synthesis step led to numerically large encoded libraries.

2.2. DNA-templated/directed/routed chemistry (II)

DNA-directed approaches make use of DNA strands as barcodes for compound identification, for forcing reactants into proximity, and/or use them to program an encoded library synthesis. The group of David R. Liu exploited the barcoding and templating properties of DNA to introduce the DNA-templated chemical libraries (DTL, Figure 2c).[56,57] Here, the first building block is coupled to a long single-stranded template DNA that contains coding regions for programmed library synthesis. Hybridization of the template with anticodon-building block

Verena B. K. Kunig received her M.Sc. degree in chemical biology from the TU Dortmund in 2017. She is currently a PhD student under the supervision of PD Andreas Brunschweiger at the same institution. She is developing reaction methodology for DNA-encoded library synthesis and assay technology for encoded library selection.

Marco Potowski received his M.Sc. degree in chemical biology at TU Dortmund University in 2010 and his PhD under the supervision of Prof. Herbert Waldmann at the Max Planck Institute (MPI) of Molecular Physiology Dortmund in 2015. After postdoctoral research in Prof. Waldmann's group until 2017, he joined PD Andreas Brunschweiger's group at the TU Dortmund working on the development of DNA-compatible reactions for DEL synthesis.

Mateja Klika Škopić received her M.Sc. degree in chemistry from the Faculty of Chemical Engineering and Technology, University of Zagreb. After working for almost two years as a medicinal chemist in a biotech company in Zagreb, she joined the research group of PD Andreas Brunschweiger at TU Dortmund, where she earned her PhD in the field of DNA-encoded libraries. She is currently working as a postdoctoral researcher in the Brunschweiger lab on the development of micellar catalysis for DEL synthesis.

PD Andreas Brunschweiger studied pharmacy at the University of Kiel. He received his PhD in the group of Prof. Christa Müller at the University of Bonn. Following a postdoctoral stay in her group and in the group of Prof. Jonathan Hall at ETH Zürich, he has headed a research group at TU Dortmund since 2013. His group develops DNA barcoding strategies for small-molecule synthesis, designs encoded screening libraries, and uses encoded library technology to identify bioactive molecules.
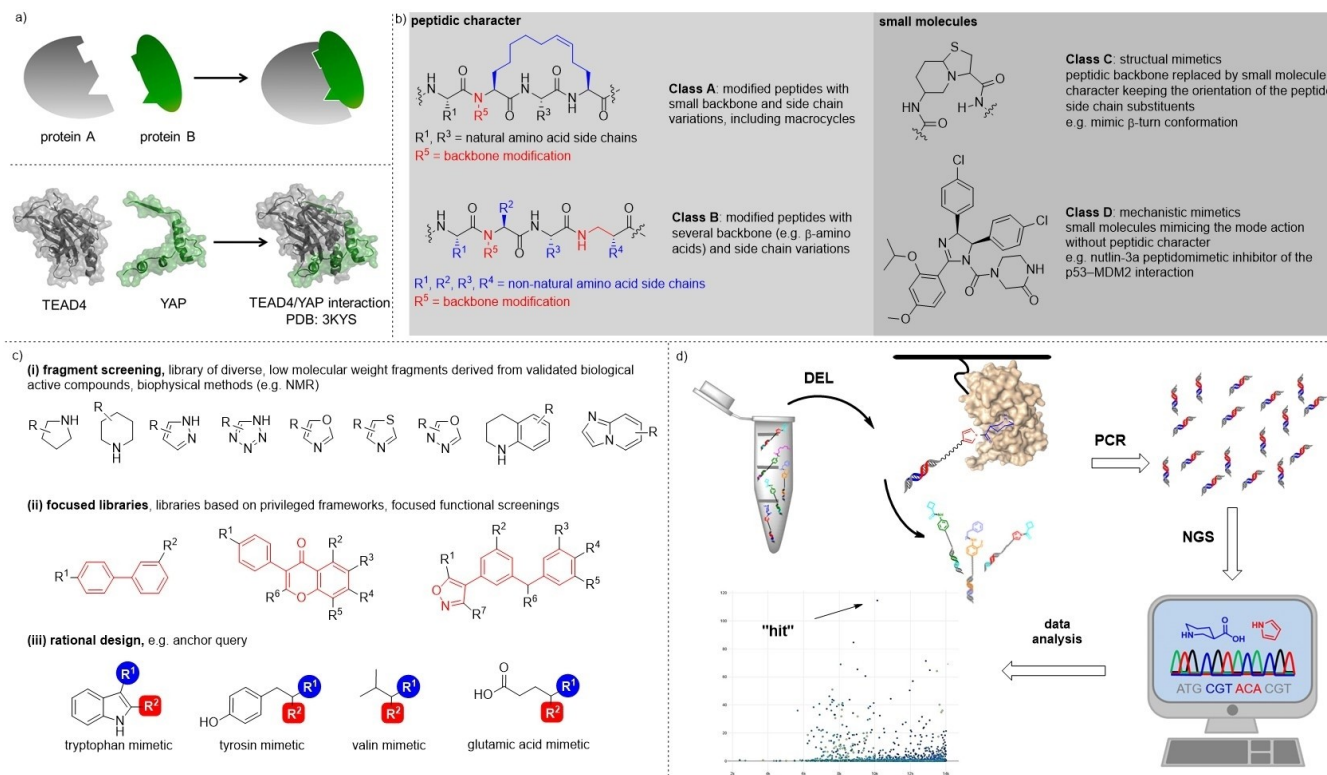
ChemMedChem

Reviews
doi.org/10.1002/cmdc.202000869

Chemistry Europe
European Chemical
Societies Publishing

Figure 1. Targeting protein–protein interactions (PPIs). a) The interaction of transcription factor TEAD and the transcriptional co-activator YAP is an exemplary protein–protein interaction. b) Classification of peptide-derived and -inspired modalities for targeting PPIs. c) Design strategies for the identification of PPI inhibitors. d) Selection of DNA-encoded libraries enables scanning protein surface with numerically large compound libraries.

conjugates, chemical reaction of the building blocks followed by cleavage of the anticodon-building block linker leads to encoded libraries. The synthesis of diverse macrocycle libraries is an impressive application of DNA-templated chemistry.[56,57] A related approach, called "yoctoreactor", was developed by Hansen and co-workers to synthesize encoded small-molecule libraries from a DNA-conjugated/encoded starting point for DEL synthesis and DNA-constructs consisting of partially comple- mentary sequences that encode bifunctional starting materials linked by a cleavable linker.[58] In an approach called "DNA- routing", Harbury used DNA anticodon strands to direct DNA- encoded libraries to vessels for programmed library synthesis.[59]

2.3. DNA-encoded solid-phase synthesis (DESPS, III)

Encoded solid-phase chemistry offers advantages such as free choice of the solvent,[53,60] and, as mentioned above, the perspective to employ different screening technologies for compound identification.[52] The Paegel and Kodadek groups established novel approaches to DNA-encoded one-bead-one- compound (OBOC, Figure 2d) libraries.[61,62] They modified TentaGel Rink-amide resin with an alkyne/amine bifunctional linker, coupled the headpiece DNA by copper(I)-catalyzed alkyne–azide cycloaddition (CuAAC) reaction to the alkyne and started encoded compound synthesis from the amine position. Unlike solution phase DELs, the DNA barcode of OBOC libraries

encodes multiple copies of an encoded compound.[61] Encoded OBOC libraries can either be screened by FACS to detect binding of labeled proteins to individual beads or in functional assays.[63,64] In the latter case, molecules are removed from the solid phase and a functional read-out is coupled to sequencing of the barcode of the active molecule.

2.4. Encoded self-assembled chemical (ESAC) libraries (IV)

Neri et al. introduced a DNA-encoded approach for fragment screening termed encoded self-assembled chemical (ESAC, Figure 2e) libraries to identify novel ligands for macromolecular targets or for affinity maturation of known protein binders. The ESAC strategy is based on the noncovalent combinatorial assembly of complementary DNA sequences from different sublibraries. The sublibraries consist of DNA oligonucleotides containing a hybridization domain and a unique DNA barcode identifying the chemical building blocks covalently attached to the 5′ or the 3′-end. Here, the combinatorial hybridization of relatively small sublibraries can lead to the formation of very large ESAC libraries.[65,66]
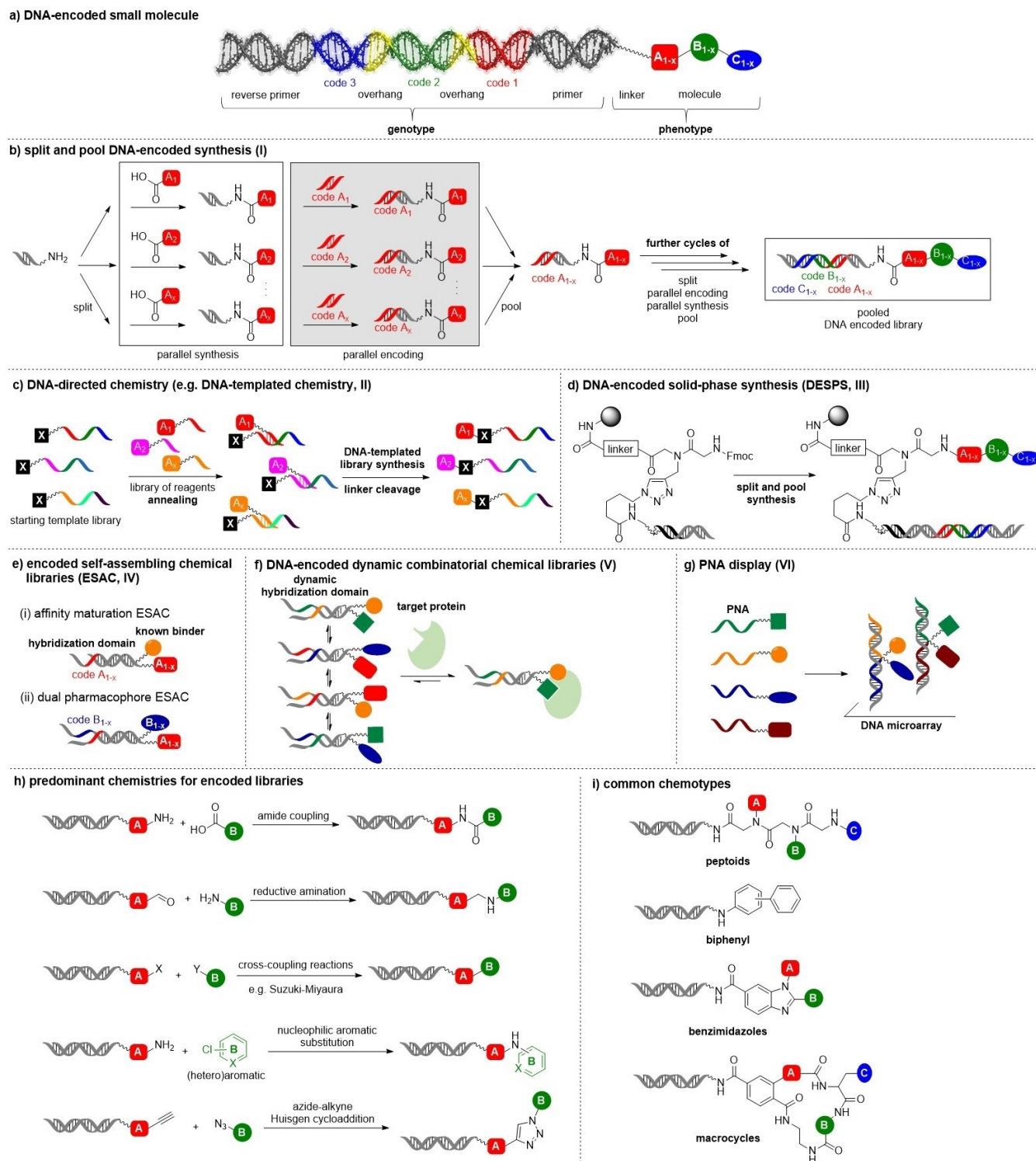
Figure 2. Encoded library technology. a) A DNA-encoded small molecule. b) Split-and-pool DNA-encoded library synthesis. c) DNA-directed chemistry. d) DNA-encoded solid-phase synthesis. e) Encoded self-assembling chemical libraries. f) DNA-encoded dynamic combinatorial chemical libraries. g) PNA display. h) Prevailing reactions for encoded library design. i) Common chemotypes found in DNA-encoded libraries.

2.5. Encoded dynamic combinatorial chemistry (V)

Dynamic combinatorial chemistry (DCC) refers to the combination of molecular building blocks through reversible reactions under thermodynamic control for the synthesis of complex small-molecule libraries. An external stimulus such as adding a biomolecule can alter the thermodynamic equilibrium of the library composition.[67] The utility of DCC to identify small-

molecule binders of target proteins was hampered by the lack of methodologies to analyze very complex small-molecule mixtures. DNA-barcoding of reactive fragments enables increasing library sizes. Encoded dynamic combinatorial chemical libraries (Figure 2f) make use of DNA-mediated hybridization of relatively "unstable" duplex DNA oligonucleotides that can be re-paired upon target addition to enrich high affinity fragment combinations.[68,69] Freezing the thermodynamic equilibrium was facilitated, for example, by photo-crosslinking or ligation of DNA oligonucleotides.[70,71]

2.6. PNA display (VI)

Winssinger and co-workers exploited the chemically much more stable peptide nucleic acid (PNA) to encode small molecules (Figure 2g). Synthesizing a PNA-encoded compound library benefits from the opportunity that the PNA tag can be co-synthesized with the organic molecule by traditional solid-phase synthesis strategy. One limitation of using PNA is that it cannot function as a template for amplification and sequencing using polymerases. Instead, DNA arrays were used to display PNA-tagged small molecules, for example, for fragment screening.[72,73]

2.7. Encoded library chemical space

The reactions for encoded library design have been extensively reviewed elsewhere.[74,75] DNA-encoded compounds mirror the linear process of DNA barcode concatenation. As the barcode grows in a linear manner, barcoded compounds are concatenated from building blocks, either in linear fashion, or coupled successively to a central scaffold displaying functional groups. Linear structures may be cyclized in the final step, yielding macrocyclic structures. Published screening hits from DEL screens have for instance validated carbonyl chemistries, C–C cross-coupling reactions, CuAAC "reaction, nucleophilic aromatic substitution reactions, and benzimidazole synthesis for library construction (Figure 2h). These reactions enrich sp^2-rich molecules and structures with peptidic character in screening libraries (Figure 2i).

3. Inhibition of Protein–Protein Interactions

3.1. Chemically modified peptides and peptide macrocycles

CBX8: Polycomb group (PcG) proteins are transcriptional repressors[76] that are part of polycomb repressive complex 1 (PRC1) and polycomb repressive complex 2 (PRC2).[77–79] CBX8 has recently emerged as a potential drug target in a variety of malignancies, such as leukemia with MLL (mixed lineage leukemia) translocations.[80] The groups of Krusemark and Dykhuizen employed a DNA-routing approach to identify potent and selective CBX8 chromodomain (ChD) inhibitors.[81,82] Selection of a peptide-DEL against a panel of CBX ChDs led to

the identification of peptide sequences with increased affinity and selectivity to CBX8 over CBX7 ChD. The authors showed then the utility of encoded combinatorial libraries by designing a focused encoded library with the aim of improving inhibitor affinity, selectivity, and cell permeability.[82] Several molecules were selected for off-DNA experiments. The K_d of fluorescently labeled compound **1** for the CBX8 ChD was ~800 nM which was similar to ~500 nM value obtained with unlabeled **1** (Figure 3) in a thermal shift assay. This compound showed high selectivity for CBX8 over CBX4 and CBX6, 20-fold selectivity over CBX7, and fivefold selectivity over CBX2. NMR spectroscopy studies indicated that inhibitor **1** can compete with histone tail binding. Biotinylated compound **1** was used to enrich CBX8 and other paralogs from mouse embryonic fibroblast (MEF) lysates and HEK293T lysates. Chromatin immunoprecipitation (ChIP) followed by quantitative PCR (ChIP-qPCR), and sequential salt extraction (SSE) validated the ability of **1** to disrupt CBX8 association with chromatin.

Compound **1** demonstrated antiproliferative activity in CBX8-dependent leukemia cells with MLL-AF9 translocations (THP1 cells), with IC$_{50}$ of 26 μM. In addition, inhibition of CBX8 ChD with **1** decreased the transcription of MLL-AF9 target genes (HOXA9, CDK6, MYB, RUNX2, and RUNX3) in THAP1 cells after 2 days of the treatment. Overall, this study highlighted the potential of small, focused DELs for targeting challenging proteins.[82]

cIAP and XIAP: Inability of cells to execute apoptosis, or programmed cell death (PCD), is associated with many malignant diseases.[83] Cellular inhibitor of apoptosis proteins (cIAPs) inhibit the extrinsic pathway of PCD by blocking the activated initiator caspase-8 protein, while X-chromosome-linked inhibitor of apoptosis proteins (XIAP) directly binds and inhibits both initiator and effector caspases associated with both PCD pathways. A 160 000-member DNA-templated library of macrocyclic pentapeptides was designed around the N-terminal alanine-residue of the native XIAP binding peptide sequence AVPI and screened for the identification of potent cIAP/XIAP antagonists.[84] The DNA-programmed peptide library was synthesized through five coupling steps with 20 different natural and unnatural amino acids and a final cyclization step via CuAAC reaction. Initial hits showed activity in the micromolar range, and their amino acid sequence was similar to the N-terminal sequence of Smac. In follow-up studies, synthesis of a small focused 1760 member DEL and structure-guided compound optimization led to compound **2** (Figure 3) which demonstrated good balanced affinity for XIAP BIR2 (IC$_{50}$ = 0.14 μM), XIAP BIR3 (IC$_{50}$ = 0.16 μM), and cIAP1 BIR3 (IC$_{50}$ = 0.02 μM). Interestingly, a dimeric molecule that formed as a side-product, showed nanomolar affinity and was used for further inhibitor design cycles.[84]

VHL: Encoded split-and-pool peptide chemistry with a final macrocyclization step by CuAAC reaction yielded a library of macrocyclic peptides through six reaction cycles leading to 2.4×10^{12} peptides of variable ring sizes ranging from four to 20 amino acids.[85] Library synthesis involved coupling of natural and non-natural amino acids, dipeptides, and tripeptides. The authors of this work included hydroxyproline, a ligand of the E3
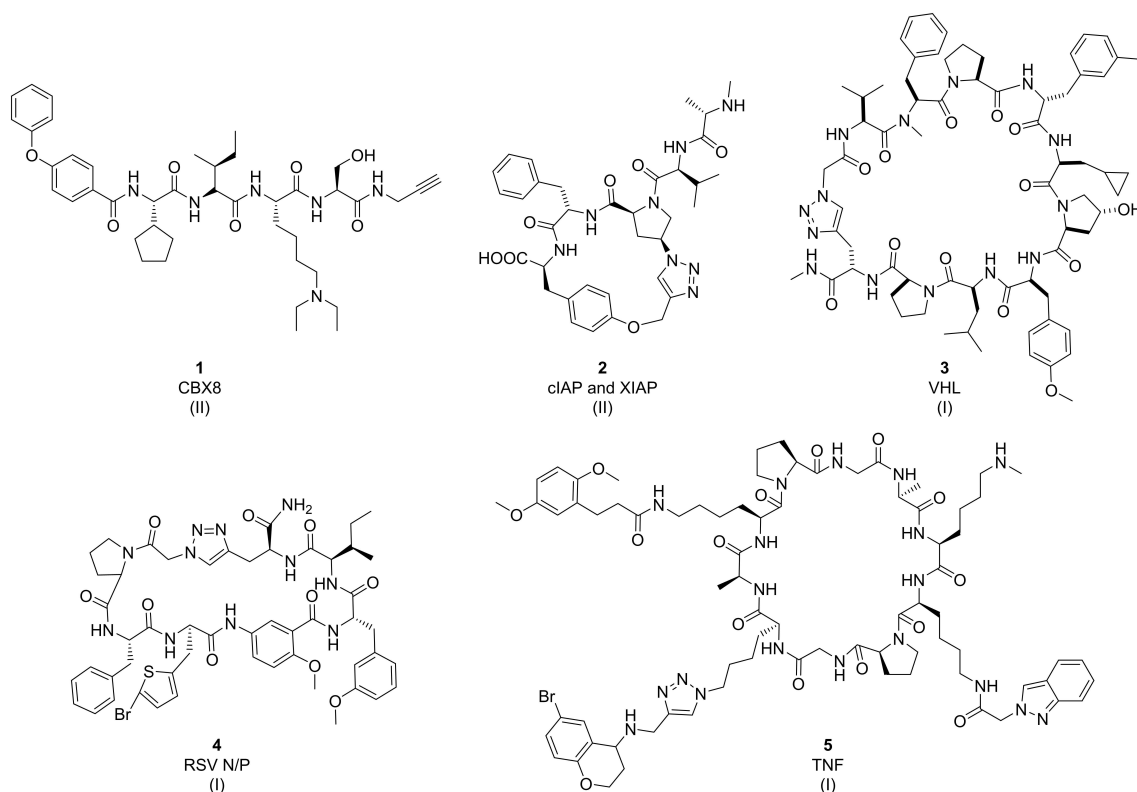
ChemMedChem

Reviews
doi.org/10.1002/cmdc.202000869

Chemistry Europe
European Chemical
Societies Publishing

Figure 3. Examples of chemically modified peptides and peptide macrocycles as PPI inhibitors (class A peptidomimetics). Roman numerals in brackets indicate the DEL technology origin of the compound.

ubiquitin ligase Von-Hippel-Lindau tumor suppressor (VHL), in the library.[86] This E3 ubiquitin ligase is involved in the ubiquitination and subsequent degradation of a hypoxia-inducible factor (HIF). HIF is a transcription factor with a critical role in the regulation of gene expression by oxygen. The selection against VHL validated the peptide macrocycle DEL, enriching macrocycles with hydroxyproline residues such as compound **3** (Figure 3). These compounds could be used as tools to probe the VHL/HIF protein–protein interaction.[85]

RSV N-protein/P-protein: The same encoded macrocycle library was selected against respiratory syncytial virus (RSV) N-protein. Interaction between RSV N-protein and P-protein is crucially important for the replication of RSV,[87] and inhibitors of this PPI hold promise for the treatment of RSV infections.[88] Macrocyclic peptides with good predicted permeability and solubility, and with at least 10-fold enrichment compared to the corresponding linear peptides were selected for functional studies. Their binding was confirmed by an affinity selection–mass spectrometry (AS-MS) assay. Functional activity of macrocyclic peptide **4** (Figure 3) was demonstrated in a time-resolved fluorescence resonance energy transfer (TR-FRET) assay that detected disruption of the interaction between RSV N-protein and P-protein (IC$_{50} \sim$ 100 nM). Comparison with linear peptides showed that peptide macrocyclization had a positive effect on affinity for the RSV N-protein.[85]

TNF: A split-and-pool DEL was designed by Neri and co-workers to mimic antibody–antigen recognition through three

diversity elements displayed on a structurally defined macrocyclic scaffold.[89] A previously reported macrocyclic scaffold with antiparallel β-sheets[90–92] was selected as a platform for the library synthesis. This scaffold was substituted on three orthogonally protected amines by combinatorial amide coupling reactions and CuAAC leading to 35 million encoded macrocycles. The library was validated by selection against carbonic anhydrase IX (CAIX), horseradish peroxidase (HRP), and tankyrase 1 (TNKS 1). In addition to the identification of novel inhibitors of human serum albumin (HSA), alpha-1 acid glycoprotein (AGP), calmodulin (CaM), and prostate-specific antigen (PSA), screening on tumor necrosis factor (TNF) also led to the identification of TNF inhibitor **5** (Figure 3). Compound **5** demonstrated activity against recombinant TNF ($K_d = 15 \mu$M) and TNF-antibody fusion L19-TNF ($K_d = 6.1 \mu$M). This study demonstrated that large libraries consisting of side-chain diversity on a constant macrocycle scaffold can deliver valuable protein binders. In addition, the structure–activity relationships (SARs) that were recognized after selection experiments suggested that design of second-generation libraries around the enriched members might lead to further potency gains.[89]

3.2. Peptoids

IgG: Distinguishing the latent infectious condition (LTB) from active infectious condition (ATB) during *Mycobacterium tuber-*

ChemMedChem

Reviews
doi.org/10.1002/cmdc.202000869

Chemistry Europe
European Chemical
Societies Publishing

culosis (Mtb) infection is important for a proper interpretation of the patient's health state.[93] Therefore, the development of highly sensitive and specific diagnostic tools is very important. The OBOC DEL technology was combined with FACS-based screening to discover ATB-specific serum ligands that bind IgG.[94] A 448 000-member OBOC DEL library was screened against LTB and ATB serum pools using fluorescently labeled IgG. The library displayed diversity at three positions and it was synthesized by peptide coupling in combination with a peptoid construction via use of halogenated acids and amines for halide substitution. Competition binding data revealed four ligands that maximally sampled the ATB patient serum samples with ligand **6** (Figure 4) as a potent and selective ATB serum IgG binder that mimics a native Ag85B epitope. The sequencing data suggested that conformational constraint is important for IgG binding.[94]

Skp2: Selection of a DNA-encoded OBOC library of cyclic peptides against oncogenic protein Skp2 (S-phase kinase-associated protein 2) identified the hit compound **7** (Figure 4, $K_d = 7.51$ µM).[95] Synthesis of the peptoid library was initiated on the amino-modified TentaGel beads. First, the Fmoc-Ahx-OH linker was coupled. Following Fmoc-deprotection, a photo-cleavable ANP linker and an azido-modified amino acid were coupled. The amino group of the linker was reacted with chloroacetic acid, and the resulting chloride was substituted with 15 different amine monomers (aliphatic, aromatic, hetero-

aromatic). The reaction sequence of chloroacylation and nucleophilic substitution with amines was repeated for the synthesis of a combinatorial hexamer peptoid library of around 11 million macrocycles. The OBOC DEL was finalized with a CuAAC reaction for the macrocycle formation. To reduce the number of initial hits identified from the affinity-based DEL selection against the Skp2-Skp1 complex for off-DNA resynthesis and validation, and in particular to avoid resynthesis of false-positive binders arising from DNA tag interaction with the target, the authors decided to synthesize a focused off-DNA sublibrary of cyclic peptides on bilayer beads and screened it against Skp2. Based on the screening results, only five molecules were selected for off-DNA hit validation. In this research work, the hit identification omitted the next-generation sequencing (NGS) but took the advantage of T-vector (TA) cloning of the selected and PCR-amplified DNA.[95]

N-Crk-SH3: The c-Crk protein, a Src homology 3 (SH3) domain containing adaptor protein, plays an important role in signal transduction. Its dysregulation is associated with malignant diseases.[96] SH3 domains are involved in numerous PPIs[97] and are considered promising drug targets.[98,99] Selection of a 100 million-member octamer peptoid PNA-encoded library against the N-terminal SH3 domain of c-Crk protein (N-Crk-SH3) led to the identification of ten peptoids selected for off-DNA validation.[100]

This library was synthesized by a reaction sequence of chloroacylation and nucleophilic substitution with amines which was repeated until octamer construction. The peptoids shared some structural features such as the tris-(2-amino-ethyl)amine side chain at the N-terminal end, bulky side chains at the first and/or second position, and three or four small side chains at the central peptoid portion. Six ligands from different ligand families showed binding affinities ($K_d = 10–100$ µM) for N-Crk-SH3, similar to those of natural SH3 peptide-based binders[101] with peptoid **8** (Figure 4) exhibiting the highest affinity for N-CrkSH3 ($K_d = 16$ µM).[100]

3.3. Structural small-molecule peptidomimetics

Mcl-1: Mcl-1 is an antiapoptotic protein from the Bcl-2 protein family. Identification of small-molecule mimetics of the BH3 domain of the pro-apoptotic inducer protein NOXA is of great interest as these molecules could potentially bind to Mcl-1 and disrupt protein–protein interaction between Mcl-1 and effector proteins BAX or BAK to initiate apoptosis.[102,103] A drug discovery program aimed at identifying novel Mcl-1 inhibitors, employed affinity-based screening of a tripeptide DEL against Mcl-1.[104] The co-crystal structure of hit compound **9** (Figure 5, $IC_{50} = 2$ µM), which comprises a dihydrobenzazepine as a β-turn mimetic core, and Mcl-1 showed that **9** binds to the BH3 binding groove of Mcl-1 where it accommodates a β-turn conformation. Structure-guided rigidification of the compound toward the bound state through macrocyclization **10** improved the potency of the initial hit by nearly three orders of magnitude ($IC_{50} = <3$ nM).[104]

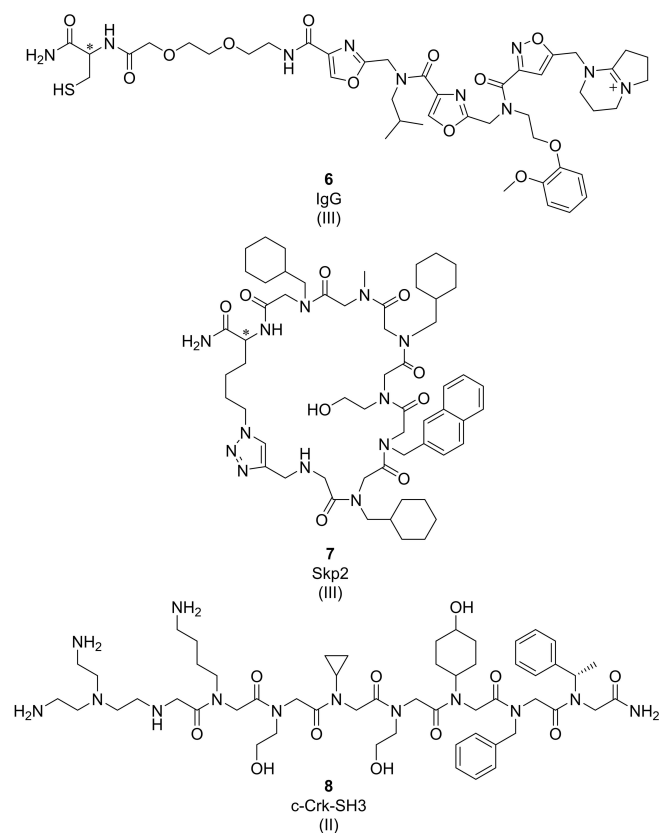**6**
IgG
(III)

7
Skp2
(III)

8
c-Crk-SH3
(II)

Figure 4. Examples of peptoids as PPI inhibitors (class B peptidomimetics). Roman numerals in brackets indicate the DEL technology origin of the compound.
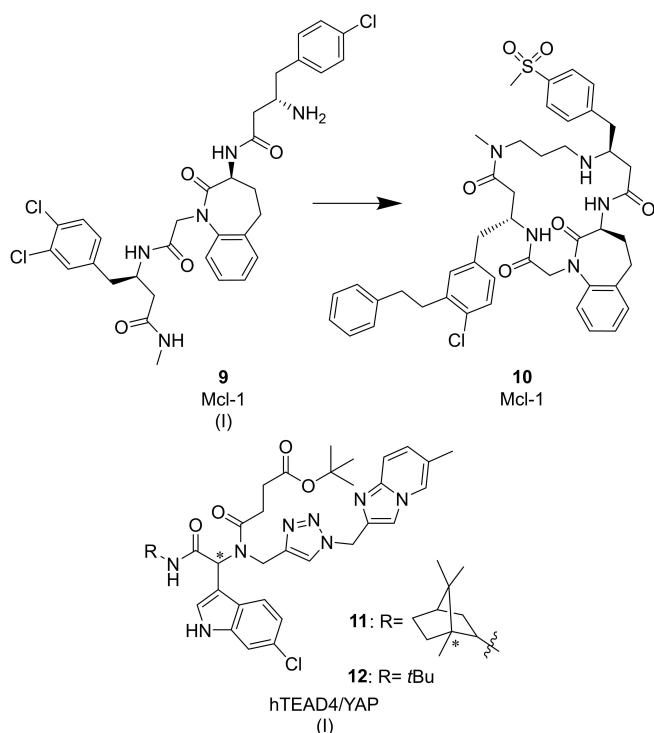
Figure 5. Examples of structural small-molecule peptidomimetics as PPI inhibitors (class C peptidomimetics). Roman numerals in brackets indicate the DEL technology origin of the compound.

hTEAD4/YAP: Dysregulation of protein–protein interactions between transcriptional enhancer factor-1 domains (TEAD1-4) and co-transcription factor Yes-associated protein (YAP), late Hippo signaling pathway effectors, is associated with oncogenic mechanisms.[105–107] Recently, screening of a small thymidine-initiated DEL (tiDEL) of peptidomimetics against YAP-interacting domain of human TEAD4 (hTEAD4) led to the identification of two PPI inhibitors **11** (Figure 5, $IC_{50} = 6.75\ \mu M$) and **12** ($IC_{50} = 5.65\ \mu M$).[108] This library was synthesized around the tryptophan side chain as an "anchor motif" with indoles introduced into the library via Ugi four-component reaction as tryptophan mimetics and CuAAC for library diversification. PPI inhibitors **11** and **12** exhibited different binding modes which are still under investigation. Both compounds showed hTEAD4/YAP interaction inhibition, however, only compound **11** exhibited inhibition of the palmitic acid-hTEAD4 interaction that takes place in the so called "central pocket" of hTEAD. Notably, compound **11** demonstrated perturbation of the expression of CTGF gene which is under control of these Hippo pathway effectors.[108]

3.4. Functional small-molecule peptidomimetics

IL-2: Interleukin-2 (IL-2) is a pro-inflammatory cytokine which is involved in high-affinity protein–protein interaction with its cognate receptor. In an early DEL project, a 30000-member split-and-pool DNA-encoded library synthesized by two combi-

natorial amide coupling reactions was screened against a panel of proteins.[109] Selection against human IL-2 revealed that many enriched sequences were coding for 2-methyl-1*H*-indole derivatives. Investigation of their binding mode in molecular docking studies revealed that the indole moiety is likely pointing towards the IL-2 surface. The most potent compound identified from the library was IL-2 inhibitor **13** (Figure 6, $K_d = 2.5\ \mu M$) which demonstrated selective IL-2 inhibition in a T-cell proliferation assay.[109]

IgG: Polyclonal human IgG binders are of great interest because they can be used to develop affinity beads for purification of monoclonal antibodies. Screening of 4000 encoded amides revealed a number of small-molecule binders of polyclonal human IgG with **14** (Figure 6) as the most enriched compound.[54]

Bcl-xL: Selection of the same library on Bcl-xL led to the identification of **15** (Figure 6) as an inhibitor of the antiapoptotic protein Bcl-xL, which is an attractive target for the development of antitumor drugs. A fluorescein-labeled derivative of compound **15** showed a higher affinity ($K_d = 0.93\ \mu M$) for Bcl-xL than the parent molecule probably due to additional interactions of the dye with the target protein.[110]

TNF Selection of a small 4000-member DEL, synthesized by Diels-Alder cycloaddition of 20 DNA-tagged dienes with 200 maleimides, on tumor necrosis factor protein (TNF) delivered hit molecule **16** (Figure 6) that was validated by fluorescence polarization assay against the trimeric EDB-TNF fusion protein ($K_d = 15\ \mu M$). This study was one of the early demonstrations of the applicability of DEL technology for the *de novo* discovery of protein–protein interaction inhibitors.[111]

3.5. Allosteric mode of action

LFA-1/ICAM-1: The integrin LFA-1 (lymphocyte function-associated antigen-1) is a leukocyte cell adhesion molecule binding to its major ligand ICAM-1 (intercellular adhesion molecule-1) on endothelial and dendritic cells.[112] Because of its key role in regulating leukocyte trafficking during inflammation and in inducing immune responses, it represents an established therapeutic target for the treatment of autoimmune and inflammatory diseases.[113,114] DEL technology was utilized for the identification of small-molecule inhibitors of LFA-1/ICAM-1 PPI.[115] The DEL was synthesized according to a seminal publication from Barry Morgan et al. by four reaction cycles yielding 4.1 billion 1,3,5-triaminotriazines.[55] Affinity selection was performed in three rounds against the soluble streptavidin-tagged LFA-1 I domain. Compound **17** (Figure 6) showed the inhibition of LFA-1/ICAM-1 interaction in the ELISA-type ligand binding assay ($IC_{50} = 23\ nM$) and demonstrated inhibition of cell adhesion to ICAM-1 in a human lymphocyte Jurkat cell-line that expresses native WT LFA-1. It exhibited the same mode of action as an established allosteric antagonist LFA703, and likely binds to the allosteric pocket below the C-terminal helix in the I domain. Finally, it was demonstrated that **17** selected on the soluble LFA-1 I domain retained affinity for native LFA-1 expressed on the cell membrane.[115]
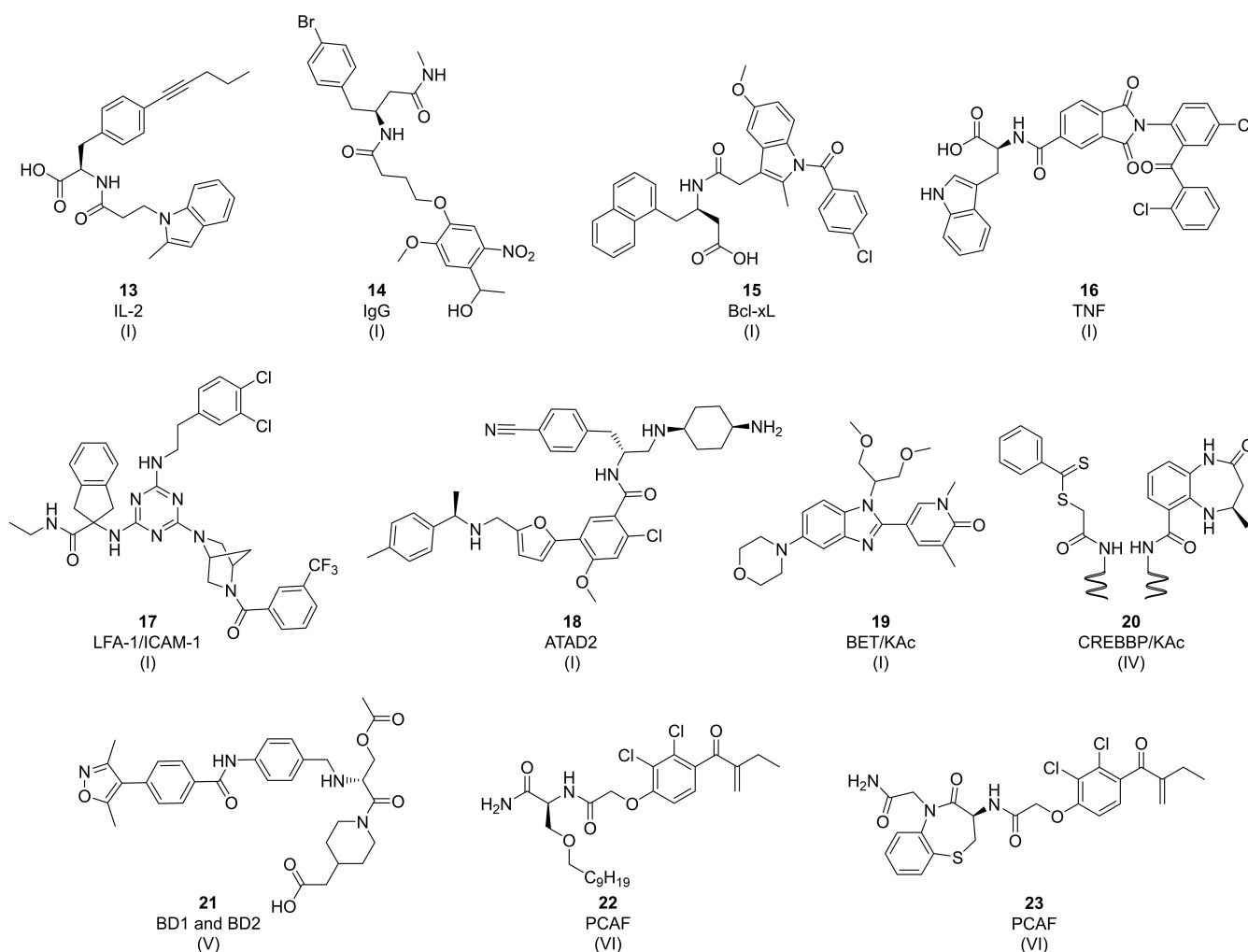
Figure 6. Examples of functional small-molecule peptidomimetics and functional mimetics of post-translationally modified peptide sequences as PPI inhibitors (class D and E peptidomimetics, see text for distinction). Roman numerals in brackets indicate the DEL technology origin of the compound.

ATAD2: ATPase family AAA-domain containing protein 2 (ATAD2 or ANCCA) is a bromodomain (BD) containing protein which acts as an epigenetic regulator and transcriptional cofactor for oncogenic transcription factors, such as ERα, AR, E2F, and Myc.[116–118] ATAD2 interacts with histone acetylation marks on newly synthesized histone H4.[119] Its overexpression is associated with different cancer types, however, its validation as a drug target is very challenging due to the lack of isoform-selective and cell active ATAD2 inhibitors. Screening a pool of 11 DELs consisting of 65 billion compounds against ATAD2 led to the identification of an isoform-selective inhibitor derived from a sublibrary.[120] This sublibrary was synthesized from three sets of building blocks: 300 different amino acids were used as the first set of building blocks (BBs), 150 different formyl acids were introduced by acylation as the second set of BBs, the third set of BBs consisted of 2341 amines introduced by reductive amination. Hit-to-lead optimization led to a selective ATAD2 inhibitor, compound **18** (Figure 6, BAY-850) which displaced the tetra-acetylated histone H4 peptide in orthogonal binding competition assays (IC$_{50}$ = 157 nM) and demonstrated activity in

cells. Interestingly, compound **18** showed an unusual mode of action as it induced protein dimer formation.[120]

3.6. Functional small-molecule mimetics of post-translationally modified peptides

BET/KAc: The bromodomain and extraterminal domain (BET) family of bromodomain containing proteins (BCPs) are epigenetic readers that recognize N-acetyl lysine (KAc) modifications on histone proteins. Dysregulation of these interactions is associated with diseases, small-molecule inhibitors of the BET/KAc interaction have entered clinical trials.[121]

A DEL screening campaign yielded a 2-(4-hydroxy-3,5-dimethylphenyl)benzimidazole series from a 117 million-member benzimidazole library.[122] The library was synthesized via three reaction cycles. First, a DNA-conjugated 4-fluoro-3-nitro benzoate was reacted by nucleophilic aromatic substitution with 65 monoprotected diamines. Following nitro reduction, the diaminoaryls were condensed with 922 aldehydes to DNA-

ChemMedChem

Reviews
doi.org/10.1002/cmdc.202000869

Chemistry
Europe
European Chemical
Societies Publishing

tagged benzimidazoles. Following the amine deprotection, the library was substituted with 1960 amine capping groups. The drug discovery program progressed by hybridizing the DEL hit with an *N*-methylpyridone fragment, a KAc mimetic obtained from fragment-based screening, demonstrating a successful combination of DEL technology with fragment screening for the development of the advanced lead molecule I-BET469 **19** (Figure 6, BRD4 BD1 $IC_{50} \sim 10$ nM).[122]

CREBBP/KAc: The cyclic-AMP response element binding protein (CREB) binding protein (CREBBP) is a transcription factor which is involved in more than 400 protein–protein interactions.[123] The CREBBP bromodomain displays two neighboring binding pockets: the KAc binding pocket and the induced-fit pocket. An encoded self-assembly chemical library was employed to identify fragments that can bind synergistically to the two adjacent CREBBP bromodomain binding sites.[124] The focused ESAC library was assembled by hybridization of partially complementary 5'-amino-modified oligonucleotides linked to 787 fragments (sublibrary A) and 3'-amino-modified oligonucleotides linked to 424 fragments (sublibrary B). The library incorporated 4,5-dihydrobenzodiazepi-none (*R*)-THBD-based ligand, a known binder of the KAc-binding pocket.[125] Affinity-based selections identified a dual fragment combination **20** (Figure 6, $K_d = 0.86$ μM) that exhibited 30-fold higher affinity for the CREBBP bromodomain than (*R*)-THBD paired with an acetyl moiety.[124]

BD1 and BD2: Selection of a focused DNA-encoded dynamic library (DEDL) against bromodomain 4 (BRD4) revealed BD1 and BD2 inhibitor **21** (Figure 6).[126] In general, design of this DEDL requires a known protein binder which can be utilized as an "anchor" with an aldehyde group and DNA-encoded library whose members are displaying a primary amine. DNA-encoded compounds compete for the anchor via imine formation, the protein target orchestrates the formation of high-affinity bind-ers, and finally reductive amination terminates the dynamic exchange. Three BD1/BD2 anchors, known binders with differ-ent affinities for BRD4, were mixed each with 67600 encoded dipeptides to form dynamic libraries that were screened against BD1 and BD2 in solution. The selection results showed that most of the compounds were more active than the anchors themselves, making this method useful for ligand optimization across a wide range of binding affinities. For an isoxazole based anchor with an IC_{50} in the mid-micromolar range, the IC_{50} improved in the case of hybrid **21** 26-fold to 1.55 μM for BD1 and 29-fold to 1.46 μM for BD2.[126]

PCAF: Selection of a library of PNA-encoded small-molecule fragments against P300-CBP associated factor (PCAF) led to the identification of 25 fragments that were used as a starting point for the synthesis of a focused PNA-encoded library of 625 com-pounds displayed on DNA microarray.[127] This small library was synthesized with two sets of building blocks. As the first set of building blocks, 25 different natural and unnatural amino acids were introduced, and then different carboxylic acids, sulfonyl chlorides, and alkynes were introduced as the second set of building blocks by acylation, sulfonylation, and CuAAC reaction, respectively. This small library yielded two PCAF binders, **22** and **23** (Figure 6), which comprised ethacrynic acid, an FDA approved drug for the treatment of high blood pressure.

Incubation of these biotin-labeled compounds with purified PCAF and PCAF in the crude lysate resulted in PCAF labeling which showed that these ethacrynic acid derivatives were engaged in covalent interactions with PCAF cysteine residues. Both compounds were used to assess the position and reactivity of different cysteine residues of 32 tested bromodo-mains and showed differential reactivity with different bromo-domains, for example, compound **22** reacted faster with BRD7, while compound **23** reacted faster with CREBBP. Both com-pounds proved useful for proteomic analysis as they were able to enrich very low concentrations of PCAF from cell lysates.[127]

4. Proteases

4.1. Chemically modified peptides and peptide macrocycles

IDE: Insulin-degrading enzyme (IDE) modulates blood glucose levels by degrading both blood glucose lowering insulin and blood glucose elevating glucagon. Selective inhibition of enzymatic insulin degradation versus glucagon degradation would open a novel treatment option for acquired diabetes. Liu et al. screened a small library of 13000 peptide macrocycles on IDE and could identify a class of molecules, exemplified by **24** (Figure 7), that blocked access of insulin to the enzyme selectively and inhibited insulin degradation with an IC_{50} of 50 nM, while the ternary complex of IDE, inhibitor, and glucagon was still catalytically active. As a consequence of the unique binding mode, these molecules reprogrammed the substrate specificity of IDE.[128,129]

4.2. Functional peptidomimetics

MMP3: Neri et al. used the ESAC technology for the identifica-tion of inhibitors of stromelysin-1 (MMP-3), a matrix metal-loproteinase, yielding novel inhibitors with micromolar potency. Matrix metalloproteinases are zinc-containing, extracellular endopeptidases that are involved in tissue remodeling proc-esses; several enzymes of this family are associated with cancer and inflammatory diseases. A DNA-encoded library with 550 compounds was selected against human MMP3 identifying a strongly enriched fragment-like compound **25** (Figure 7) which was then used as an anchor structure for the assembly of the 550-member ESAC sublibrary to identify more potent, bidentate inhibitors. Dual-pharmacophore selections against MMP3 resulted in the identification of a specific pair of compounds (structure I and II) which were then combined by different linkers for the synthesis of MMP3 inhibitors binding synergistically to the target protein among which compound **26** showed the highest inhibitory potency with an IC_{50} of 9.9 μM.[130]

Trypsin and thrombin: Focused split-and-pool DNA-encoded libraries were designed by the groups of Neri and Matzuk, respectively, around the oxyanion hole-binders benzamidine and guanidine to probe the surface around the active site of the serine proteases trypsin and thrombin, respectively.
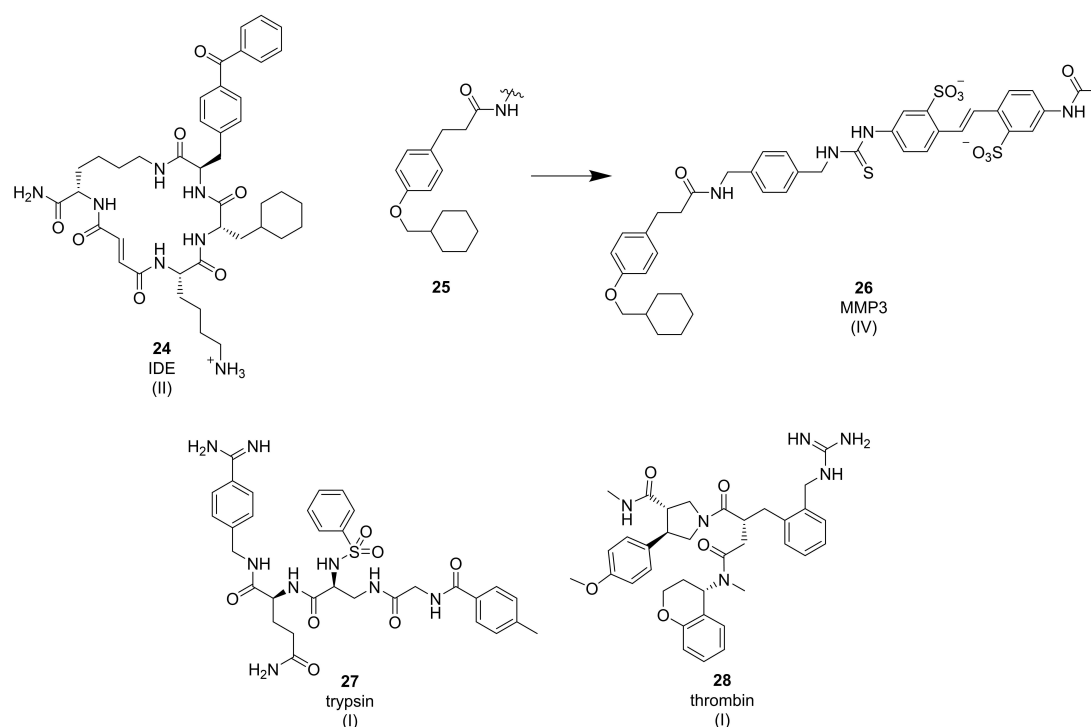
ChemMedChem

Reviews
doi.org/10.1002/cmdc.202000869

Chemistry Europe
European Chemical
Societies Publishing

Figure 7. Examples of protease inhibitors (class A (**24**) and D peptidomimetics). Roman numerals in brackets indicate the DEL technology origin of the compound.

Successful identification of nanomolar inhibitors of trypsin **27** (Figure 7), and thrombin **28** (Figure 7), showed the potential of combinatorial libraries to identify highly potent protease inhibitors by densely covering chemical space around weakly active starting points.[131,132]

5. Peptide-Competitive Ligands for Receptors and Enzymes

5.1. Chemically modified peptides and peptide macrocycles

c-Src kinase: Krusemark reported the synthesis of a 550 000-member phenol containing peptidomimetic DNA-encoded library for the identification of unnatural substrates of the tyrosine protein kinase c-Src that can serve as artificial substrates and potentially as protein substrate competitive inhibitors. The DEL contained native peptides, non-natural peptides and peptoid-inspired structures. The library was treated with ATP and Src kinase followed by affinity selection assay using a non-specific phosphotyrosine-binding antibody for the enrichment and identification of substrate molecules of c-Src. Substrate-mediated selection led to the identification of a lead compound **29** (Figure 8) that was able to serve as a substrate for phosphorylation and also to promote ATP hydrolysis. Binding of this compound to the c-Src:ATP complex was confirmed using NMR and an ester derivative of the hit compound showed inhibition of Src-dependent signaling in NME cells.[133]

c-Src kinase: The aforementioned macrocycle library from Liu et al. was selected on a set of proteins, among them several kinases. Compound **30** (Figure 8), inhibited c-Src kinase activity with an IC$_{50}$ of 960 nM and high selectivity versus a panel of 58 kinases. X-ray structures revealed a bisubstrate mode of kinase inhibition, the macrocycle occupied the ATP-binding pocket and blocked at the same time the substrate peptide-binding patch, locking the kinase in an inactive conformation.[134–136]

5.2. Functional peptidomimetics

PAR2: Protease-activated receptors are activated by enzymatic cleavage of an extracellular domain, unmasking the peptide sequence SLIGKV which binds to the receptor transmembrane domain. Excessive receptor activation is associated with inflammatory diseases. Researchers from X–Chem and AstraZeneca screened several large split-and-pool DELs that included capped diamide and benzimidazole DELs with 225 million and 7 million compounds, respectively, on a mutated, stabilized variant of the receptor in the presence and absence of a small-molecule antagonist, revealed several classes of compounds. Functional and structural studies showed that compound **31** (Figure 8) acted as functional mimic of the PAR-activating peptide sequence, while compound **32** bound to a previously unknown allosteric site and caused structural rearrangements of the receptor that precluded activation.[137,138]
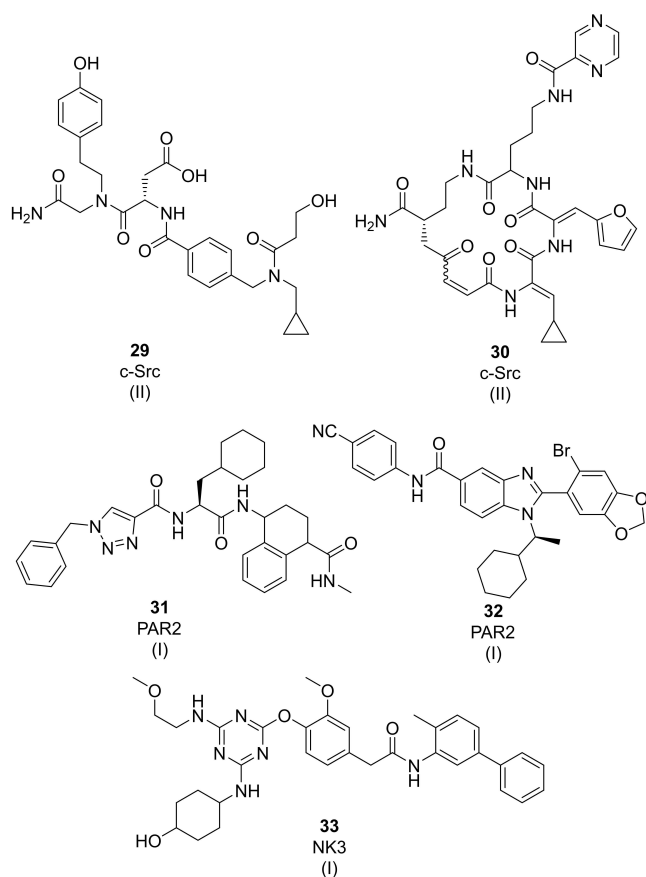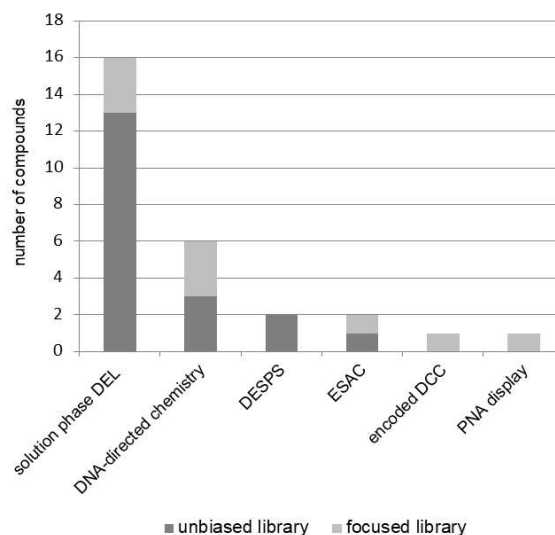
ChemMedChem

Reviews
doi.org/10.1002/cmdc.202000869

Chemistry Europe
European Chemical
Societies Publishing

Figure 8. Examples of peptide-competitive ligands for receptors and enzymes (see text for compound classification). Roman numerals in brackets indicate the DEL technology origin of the compound.
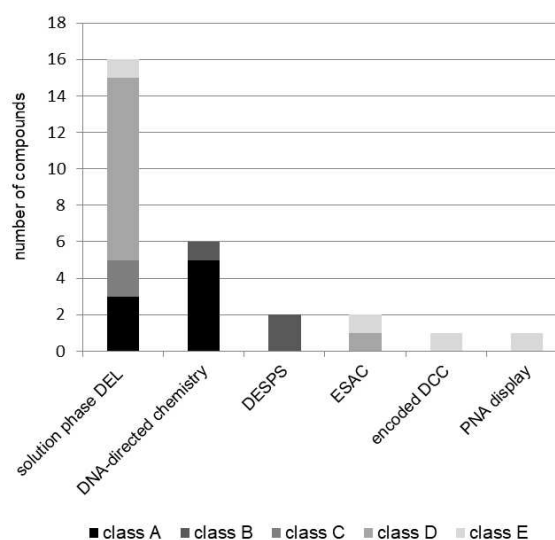

Figure 9. Statistical analysis of the compounds discussed in this review. a) Compounds identified from focused versus unbiased encoded library designs. b) Peptidomimetic compound classes identified from different DEL technologies.

NK3 receptor: The neurokinin receptor NK3, a G protein-coupled receptor, is activated by neuropeptides. It was used by a research team from GSK as a model system to establish DEL selection experiments on cell membranes. Selection of split and pool libraries on NK3 receptors that were overexpressed on HEK293 cells yielded several hit clusters, including compound **33** (Figure 8). Of note, the authors remarked that the hit clusters did not show similarity to established NK3 receptor antagonists, although such chemotypes were present in the library, furthermore DEL selections were productive on receptors for chemokines, peptide hormones and lipids, whereas they provided less hit matter on receptors for low-molecular-weight signaling molecules.[139]

6. Conclusions

Scanning of target protein surfaces with DNA-encoded libraries has delivered several compounds that inhibit protein–protein interactions. More than half of these compounds (16 out of 28) originate from solution-phase split-and-pool combinatorial libraries, a fact that reflects the widespread uptake of this library technology (Figure 9). Most of these molecules belong to the classes C–E of functional small-molecule peptidomimetics, and

only a minority are oligomeric, peptide compounds. DNA-directed chemistry on the other hand accounted for six compounds, all of which belong to the oligomeric peptides and peptoids (classes A and B peptidomimetics), and also the published hits of encoded solid-phase library screens represent bioactive peptoids. DNA-encoded fragment screens, either performed by the ESAC technology, dynamic combinatorial chemistry, or PNA display hold much promise to scan protein surface. Yet, they have to date yielded only a handful of PPI modulators. Likely this is due to the fact, that these approaches are practiced by very few research groups. Roughly a third of the compounds shown in this review were identified from smaller, target-focused libraries. They hint at the potential of

ChemMedChem

Reviews
doi.org/10.1002/cmdc.202000869

Chemistry
Europe
European Chemical
Societies Publishing

DEL technology to improve the potency of weak starting points, for example, from fragment screening. Such approaches certainly benefit from the dense coverage of chemical space by combinatorial compound synthesis.

Surprising to us was the scarce use of structural peptide mimetics for library design. These hold much promise for the identification of PPI modulators and may either be introduced by reaction methodology as for instance multicomponent reactions or by scaffolds that mimic protein secondary structure. The design of such molecule libraries may take into account the guidelines that have been published by Kihlberg et al. and adopted in a macrocycle design by the Liu group.[57,140] As a final remark, we wish to point out that DEL technology may provide excellent starting points for the development of protein-degrading or other hetero-bifunctional molecules.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

Keywords: DNA-encoded libraries · drug development · peptidomimetics · protein–protein interactions · screening

[1] K. Venkatesan, J. Rual, A. Vazquez, U. Stelzl, I. Lemmens, T. Hirozane-Kishikawa, T. Hao, M. Zenkner, X. Xin, K. Goh, *Nat. Methods* **2009**, *6*, 83–90.

[2] K. Luck, D. Kim, L. Lambourne, K. Spirohn, B. E. Begg, W. Bian, R. Brignall, T. Cafarelli, F. J. Campos-Laborie, B. Charloteaux, *Nature* **2020**, *580*, 402–408.

[3] A. Chatr-aryamontri, R. Oughtred, L. Boucher, J. Rust, C. Chang, N. K. Kolas, L. O'Donnell, S. Oster, C. Theesfeld, A. Sellam, C. Stark, B. Breitkreutz, K. Dolinski, M. Tyers, *Nucleic Acids Res.* **2017**, *45*, D369–D379.

[4] D. Szklarczyk, A. L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N.T. Doncheva, J. H. Morris, P. Bork, L. J. Jensen, C. von Mering, *Nucleic Acids Res.* **2019**, *47*, D607–D613.

[5] A. W. White, A. D. Westwell, G. Brahemi, *Expert Rev. Mol. Med.* **2008**, *10*, e8.

[6] G. Kar, A. Gursoy, O. Keskin, *PLoS Comput. Biol.* **2009**, *5*, e1000601.

[7] A. A. Ivanov, F. R. Khuri, H. Fu, *Trends Pharmacol. Sci.* **2013**, *34*, 393–400.

[8] T. L. Nero, C. J. Morton, J. K. Holien, J. Wielens, M. W. Parker, *Nat. Rev. Cancer* **2014**, *14*, 248–62.

[9] Z. Li, A. A. Ivanov, R. Su, V. Gonzalez-Pecchi, Q. Qi, S. Liu, P. Webber, E. McMillan, L. Rusnak, C. Pham, *Nat. Commun.* **2017**, *8*, 14356.

[10] T. L. Spires-Jones, J. Attems, D. R. Thal, *Acta Neuropathol.* **2017**, *134*, 187–205.

[11] A. F. Brito, J. W. Pinney, *Front. Microbiol.* **2017**, *8*, 1557.

[12] A. F. Voter, J. L. Keck, *Adv. Protein Chem. Struct. Biol.* **2018**, *111*, 197–222.

[13] D. E. Gordon, G. M. Jang, M. Bouhaddou, J. Xu, K. Obernier, K. M. White, M. J. O'Meara, V. V. Rezelj, J. Z. Guo, D. L. Swaney, *Nature* **2020**, *583*, 459–468.

[14] A. L. Hopkins, C. R. Groom, *Nature Rev. Drug Discovery* **2002**, *1*, 727–730.

[15] S. Surade, T. L. Blundell, *Chem. Biol.* **2012**, *19*, 42–50.

[16] T. Berg, *Angew. Chem. Int. Ed.* **2003**, *42*, 2462–2481; *Angew. Chem.* **2003**, *115*, 2566–2586.

[17] M. R. Arkin, J. A. Wells, *Nat. Rev. Drug Discovery* **2004**, *3*, 301–317.

[18] J. A. Wells, C. L. McClendon, *Nature* **2007**, *450*, 1001–1009.

[19] M. R. Arkin, A. Whitty, *Curr. Opin. Chem. Biol.* **2009**, *13*, 284–290.

[20] A. J. Wilson, *Chem. Soc. Rev.* **2009**, *38*, 3289–3300.

[21] Y. H. Jeon, J. Y. Lee, S. Kim, *Bioorg. Med. Chem.* **2012**, *20*, 1893–1901.

[22] L. Milroy, T. N. Grossmann, S. Hennig, L. Brunsveld, C. Ottmann, *Chem. Rev.* **2014**, *114*, 4695–4748.

[23] D. E. Scott, A. R. Bayly, C. Abell, J. Skidmore, *Nat. Rev. Drug Discovery* **2016**, *15*, 533–550.

[24] H. Lu, Q. Zhou, J. He, Z. Jiang, C. Peng, R. Tong, J. Shi, *Sig. Transduct. Target Ther.* **2020**, *5*, 213.

[25] T. Clackson, J. A. Wells, *Science* **1995**, *267*, 383–386.

[26] S. Jones, J. M. Thornton, *Proc. Natl. Acad. Sci. USA* **1996**, *93*,13–20.

[27] A. A. Bogan, K. S. Thorn, *J. Mol. Biol.* **1998**, *280*, 1–9.

[28] T. Kortemme, D. Baker, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14116–14121.

[29] B. Ma, T. Elkayam, H. Wolfson, R. Nussinov, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5772–5777.

[30] O. Keskin, B. Ma, R. Nussinov, *J. Mol. Biol.* **2005**, *345*, 1281–1294.

[31] I. S. Moreira, P. A. Fernandes, M. J. Ramos, *Proteins* **2007**, *68*, 803–812.

[32] O. Keskin, A. Gursoy, B. Ma, R. Nussinov, *Chem. Rev.* **2008**, *108*,1225–1244.

[33] D. Kozakov, L. E. Grove, D. R. Hall, T. Bohnuud, S. Mottarella, L. Luo, B. Xia, D. Beglov, S. Vajdaa, *Nat. Protoc.* **2015**, *10*, 733–755.

[34] A. A. Ibarra, G. J. Bartlett, Z. Hegedüs, S. Dutt, F. Hobor, K. A. Horner, K. Hetherington, K. Spence, A. Nelson, T. A. Edwards, D. N. Woolfson, R. B. Sessions, A. J. Wilson, *ACS Chem. Biol.* **2019**, *14*, 2252–2263.

[35] D. Koes, K. Khoury, Y Huang, W. Wang, M. Bista, G. M. Popowicz, S. Wolf, T. A. Holak, A. Dömling, C. J. Camacho, *PLoS One* **2012**, *7*, e32839.

[36] D. R. Koes, A. Dömling, C. J. Camacho, *Protein Sci.* **2018**, *27*, 229–232.

[37] F. Da Silva, G. Bret, L. Teixeira, C. F. Gonzalez, D. Rognan, *J. Med. Chem.* **2019**, *62*, 21, 9732–9742.

[38] S. Eyrisch, V. Helms, *J. Med. Chem.* **2007**, *50*, 3457–3464.

[39] G. R. Bowmana, P. L. Geissler, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 11681–11686.

[40] D. Beglov, D. R. Hall, A. E. Wakefield, L. Luo, K. N. Allen, D. Kozakov, A. Whitty, S. Vajdaa, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E3416–E3425.

[41] C. Sheng, G. Dong, Z. Miao, W. Zhang, W. Wang, *Chem. Soc. Rev.* **2015**, *44*, 8238–8259.

[42] M. Pelay-Gimeno, A. Glas, O. Koch, T. N. Grossmann, *Angew. Chem. Int. Ed.* **2015**, *54*, 8896–8927; *Angew. Chem.* **2015**, *127*, 9022–9054.

[43] D. Rognan, *MedChemComm* **2015**, *6*, 51–60.

[44] A. Ciulli, G. Williams, A. G. Smith, T. L. Blundell, C. Abell, *J. Med. Chem.* **2006**, *49*, 4992–5000.

[45] C. Reynès, H. Host, A. Camproux, G. Laconde, F. Leroux, A. Mazars, B. Deprez, R. Fahraeus, B. O. Villoutreix, O. Sperandio, *PLoS Comput. Biol.* **2010**, *6*, e1000695.

[46] N. Bosc, M. A. Kuenemann, J. Bécot, M. Vavrusa, A. H. Cerdan, O. Sperandio, *J. Chem. Inf. Model.* **2017**, *57*, 2448–2462.

[47] S. Milhas, B. Raux, S. Betzi, C. Derviaux, P. Roche, A. Restouin, M.-J. Basse, E. Rebuffet, A. Lugari, M. Badol, *ACS Chem. Biol.* **2016**, *11*, 2140–2148.

[48] D. Neri, R. A. Lerner, *Annu. Rev. Biochem.* **2018**, *87*, 479–502.

[49] J. Ottl, L. Leder, J. V. Schaefer, C. E. Dumelin, *Molecules* **2019**, *24*, 1629.

[50] D. T. Flood, C. Kingston, J. C. Vantourout, P. E. Dawson, P. S. Baran, *Isr. J. Chem.* **2020**, *60*, 268–280.

[51] M. Song, G. T. Hwang, *J. Med. Chem.* **2020**, *63*, 6578–6599.

[52] T. Kodadek, N. G. Paciaroni, M. Balzarini, P. Dickson, *Chem. Commun.* **2019**, *55*, 13330–13341.

[53] S. Brenner, R. A. Lerner, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 5381–5383.

[54] L. Mannocci, Y. Zhang, J. Scheuermann, M. Leimbacher, G. De Bellis, E. Rizzi, C. Dumelin, S. Melkko, D. Neri, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 17670–17675.

[55] M. A. Clark, R. A. Acharya, C. C. Arico-Muendel, S. L. Belyanskaya, D. R. Benjamin, N. R. Carlson, P. A. Centrella, C. H. Chiu, S. P. Creaser, J. W. Cuozzo, *Nat. Chem. Biol.* **2009**, *5*, 647–54.

[56] Z. J. Gartner, B. N. Tse, R. Grubina, J. B. Doyon, T. M. Snyder, D. R. Liu, *Science* **2004**, *305*, 1601–1605.

[57] D. L. Usanov, A. I. Chan, J. P. Maianti, D. R. Liu, *Nat. Chem.* **2018**, *10*, 704–714.

ChemMedChem

Reviews
doi.org/10.1002/cmdc.202000869

Chemistry
Europe
European Chemical
Societies Publishing

[58] M. Haahr Hansen, P. Blakskjær, L. K. Petersen, T. Heitner Hansen, J. W. Højfeldt, K. V. Gothelf, N. J. Vest Hansen, *J. Am. Chem. Soc.* **2009**, *131*, 3, 1322–1327.

[59] D. R. Halpin, P. B. Harbury, *PLoS Biol.* **2004**, *2(7)*, e173.

[60] J. Nielsen, S. Brenner, K. D. Janda, *J. Am. Chem. Soc.* **1993**, *115*, 21, 9812–9813.

[61] A. B. MacConnell, P. J. McEnaney, V. J. Cavett, B. M. Paegel, *ACS Comb. Sci.* **2015**, *17*, 518–534.

[62] N. G. Paciaroni, J. M. Ndungu, T. Kodadek, *Chem. Commun.* **2020**, *56*, 4656–4659.

[63] K. R. Mendes, M. L. Malone, J. M. Ndungu, I. Suponitsky-Kroyter, V. J. Cavett, P. J. McEnaney, A. B. MacConnell, T. M. Doran, K. Ronacher, K. Stanley, O. Utset, G. Walzl, B. M. Paegel, T. Kodadek, *ACS Chem. Biol.* **2017**, *12*, 234–343.

[64] A. B. MacConnell, A. K. Price, B. M. Paegel, *ACS Comb. Sci.* **2017**, *19*, 181–192.

[65] S. Melkko, J. Scheuermann, C. E. Dumelin, D. Neri, *Nat. Biotechnol.* **2004**, *22*, 568–574.

[66] M. Wichert, N. Krall, W. Decurtins, R. M. Franzini, F. Pretto, P. Schneider, D. Neri, J. Scheuermann, *Nat. Chem.* **2015**, *7*, 241–249.

[67] P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J. Wietor, J. K. M. Sanders, S. Otto, *Chem. Rev.* **2006**, *106*, 3652–3711.

[68] F. V. Reddavide, W. Lin, S. Lehnert, Y. Zhang, *Angew. Chem. Int. Ed.* **2015**, *54*, 7924–7928; *Angew. Chem.* **2015**, *127*, 8035–8039.

[69] G. Li, W. Zheng, Z. Chen, Y. Zhou, Y. Liu, J. Yang, Y. Huang, X. Li, *Chem. Sci.* **2015**, *6*, 7097–7104.

[70] Y. Zhou, C. Li, J. Peng, L. Xie, L. Meng, Q. Li, J. Zhang, X. D. Li, X. Li, X. Huang, X. Li, *J. Am. Chem. Soc.* **2018**, *140*, 15859–15867.

[71] F. V. Reddavide, M. Cui, W. Lin, N. Fu, S. Heiden, H. Andrade, M. Thompson, Y. Zhang, *Chem. Commun.* **2019**, *55*, 3753–3756.

[72] C. Zambaldo, S. Barluenga, N. Winssinger, *Curr. Opin. Chem. Biol.* **2015**, *26*, 8–15.

[73] J.-P. Daguer, C. Zambaldo, M. Ciobanu, P. Morieux, S. Barluenga, N. Winssinger, *Chem. Sci.* **2015**, *6*, 739–744.

[74] K. Götte, S. Chines, A. Brunschweiger, *Tetrahedron Lett.* **2020**, *61*, 151889–151899.

[75] P. R. Fitzgerald, B. M. Paegel, *Chem. Rev.* **2020**, DOI: 10.1021/acs.chemrev.0c00789.

[76] J. A. Kennison, *Annu. Rev. Genet.* **1995**, *29*, 289–303.

[77] L. Di Croce, K. Helin, *Nat. Struct. Mol. Biol.* **2013**, *20*, 1147–1155.

[78] N. J. Francis, R. E. Kingston, C. L. Woodcock, *Science* **2004**, *306*, 1574–1577.

[79] M. Lavigne, N. J. Francis, I. F. G. King, R. E. Kingston, *Mol. Cell* **2004**, *13*, 415–425.

[80] J. Tan, M. Jones, H. Koseki, M. Nakayama, A. G. Muntean, I. Maillard, J. L. Hess, *Cancer Cell* **2011**, *20*, 563–575.

[81] K. E. Denton, S. Wang, M. C. Gignac, N. Milosevich, F. Hof, E. C. Dykhuizen, C. J. Krusemark, *SLAS Discovery* **2018**, *23*, 417–428.

[82] S. Wang, K. E. Denton, K. F. Hobbs, T. Weaver, J. M. B. McFarlane, K. E. Connelly, M. C. Gignac, N. Milosevich, F. Hof, I. Paci, C. A. Musselman, E. C. Dykhuizen, C. J. Krusemark, *ACS Chem. Biol.* **2020**, *15*, 112–131.

[83] D. Hanahan, R. A. Weinberg, *Cell* **2011**, *144*, 646–674.

[84] B. A. Seigal, W. H. Connors, A. Fraley, R. M. Borzilleri, P. H. Carter, S. L. Emanuel, J. Fargnoli, K. Kim, M. Lei, J. G. Naglich, M. E. Pokross, S. L. Posy, H. Shen, N. Surti, R. Talbott, Y. Zhang, N. K. Terrett, *J. Med. Chem.* **2015**, *58*, 2855–2861.

[85] Z. Zhu, A. Shaginian, L. C. Grady, T. O'Keeffe, X. E. Shi, C. P. Davie, G. L. Simpson, J. A. Messer, G. Evindar, R. N. Bream, P. P. Thansandote, N. R. Prentice, A. M. Mason, S. Pal, *ACS Chem. Biol.* **2018**, *13*, 53–59.

[86] J. H. Min, H. Yang, M. Ivan, F. Gertler, W. G. Kaelin, N. P. Pavietich, *Science* **2002**, *296*, 1886–1889.

[87] A. P. Oliveira, F. M. Simabuco, R. E. Tamura, M. C. Guerrero, P. G. G. Ribeiro, T. A. Libermann, L. F. Zerbini, A. M. Ventura, *Virus Res.* **2013**, *177*, 108–112.

[88] G. H. Bird, S. Boyapalle, T. Wong, K. Opoku-Nsiah, R. Bedi, W. C. Crannell, A. F. Perry, H. Nguyen, V. Sampayo, A. Devareddy, S. Mohapatra, S. S. Mohapatra, L. D. Walensky, *J. Clin. Invest.* **2014**, *124*, 2113–2124.

[89] Y. Li, R. de Luca, S. Cazzamalli, F. Pretto, D. Bajic, J. Scheuermann, D. Neri, *Nat. Chem.* **2018**, *10*, 441–448.

[90] M. Mutter, R. Hersperger, K. Gubernator, K. Müller, *Proteins Struct. Funct. Bioinf.* **1989**, *5*, 13–21.

[91] P. Dumy, I. M. Eggleston, G. Esposito, S. Nicula, M. Mutter, *Biopolymers* **1996**, *39*, 297–308.

[92] S. Peluso, T. Rückle, C. Lehmann, M. Mutter, C. Peggion, M. Crisma, *ChemBioChem* **2001**, *2*, 432–437.

[93] World Health Organization, *Global Tuberculosis Report*, **2015**.

[94] K. R. Mendes, M. L. Malone, J. M. Ndungu, I. Suponitsky-Kroyter, V. J. Cavett, P. J. McEnaney, A. B. MacConnell, T. D. M. Doran, K. Ronacher, K. Stanley, O. Utset, G. Walzl, B. M. Paegel, T. Kodadek, *ACS Chem. Biol.* **2017**, *12*, 234–243.

[95] M. H. Shin, K. J. Lee, H. S. Lim, *Bioconjugate Chem.* **2019**, *30*, 2931–2938.

[96] R. B. Birge, C. Kalodimos, F. Inagaki, S. Tanaka, *Cell Commun. Signal.* **2009**, *7*, 13.

[97] G. B. Cohen, R. Ren, D. Baltimore, *Cell* **1995**, *80*, 237–248.

[98] S. Feller, M. Lewitzky, *Curr. Pharm. Des.* **2006**, *12*, 529–548.

[99] S. Hashimoto, M. Hirose, A. Hashimoto, M. Morishige, A. Yamada, H. Hosaka, K. I. Akagi, E. Ogawa, C. Oneyama, T. Agatsuma, M. Okada, H. Kobayashi, H. Wada, H. Nakano, T. Ikegami, A. Nakagawa, H. Sabe, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 7036–7041.

[100] S. J. Wrenn, R. M. Weisinger, D. R. Halpin, P. B. Harbury, *J. Am. Chem. Soc.* **2007**, *129*, 13137–13143.

[101] A. R. Viguera, A. Musacchio, M. Saraste, L. Serrano, J. L. R. Arrondo, *Biochemistry* **1994**, *33*, 10925–10933.

[102] E. Oda, R. Ohki, H. Murasawa, J. Nemoto, T. Shibue, T. Yamashita, T. Tokino, T. Taniguchi, N. Tanaka, *Science*, **2000**, *288*, 1053–1058.

[103] A. W. Roberts, M. S. Davids, J. M. Pagel, B. S. Kahl, S. D. Puvvada, J. F. Gerecitano, T. J. Kipps, M. A. Anderson, J. R. Brown, L. Gressick, S. Wong, M. Dunbar, M. Zhu, M. B. Desai, E. Cerri, S. Heitner Enschede, R. A. Humerickhouse, W. G. Wierda, J. F. Seymour, *N. Engl. J. Med.* **2016**, *374*, 311–322.

[104] J. W. Johannes, S. Bates, C. Beigie, M. A. Belmonte, J. Breen, S. Cao, P. A. Centrella, M. A. Clark, J. W. Cuozzo, C. E. Dumelin, *ACS Med. Chem. Lett.* **2017**, *8*, 239–244.

[105] F. Zanconato, M. Forcato, G. Battilana, L. Azzolin, E. Quaranta, B. Bodega, A. Rosato, S. Bicciato, M. Cordenonsi, S. Piccolo, *Nat. Cell Biol.* **2015**, *17*, 1218–1227.

[106] K. F. Harvey, X. Zhang, D. M. Thomas, *Nat. Rev. Cancer* **2013**, *13*, 246–257.

[107] P. C. Calses, J. J. Crawford, J. R. Lill, A. Dey, *Trends in Cancer* **2019**, *5*, 297–307.

[108] V. B. K. Kunig, M. Potowski, M. Akbarzadeh, M. Klika Škopić, D. dos Santos Smith, L. Arendt, I. Dormuth, H. Adihou, B. Andlovic, H. Karatas, *Angew. Chem. Int. Ed.* **2020**, *59*, 20338–20342.

[109] M. Leimbacher, Y. Zhang, L. Mannocci, M. Stravs, T. Geppert, J. Scheuermann, G. Schneider, D. Neri, *Chem. Eur. J.* **2012**, *18*, 7729–7737.

[110] S. Melkko, L. Mannocci, C. E. Dumelin, A. Villa, R. Sommavilla, Y. Zhang, M. G. Grütter, N. Keller, L. Jermutus, R. H. Jackson, J. Scheuermann, D. Neri, *ChemMedChem* **2010**, *5*, 584–590.

[111] F. Buller, Y. Zhang, J. Scheuermann, J. Schäfer, P. Bühlmann, D. Neri, *Chem. Biol.* **2009**, *16*, 1075–1086.

[112] R. Evans, I. Patzak, L. Svensson, K. De Filippo, K. Jones, A. McDowall, N. Hogg, *J. Cell Sci.* **2009**, *122*, 215–225.

[113] M. Shimaoka, T. A. Springer, *Nat. Rev. Drug Discovery* **2003**, *2*, 703–716.

[114] D. Cox, M. Brennan, N. Moran, *Nat. Rev. Drug Discovery* **2010**, *9*, 804–820.

[115] C. S. Kollmann, X. Bai, C. H. Tsai, H. Yang, K. E. Lind, S. R. Skinner, Z. Zhu, D. I. Israel, J. W. Cuozzo, B. A. Morgan, K. Yuki, C. Xie, T. A. Springer, M. Shimaoka, G. Evindar, *Bioorg. Med. Chem.* **2014**, *22*, 2353–2365.

[116] P. Filippakopoulos, S. Picaud, M. Mangos, T. Keates, J. P. Lambert, D. Barsyte-Lovejoy, I. Felletar, R. Volkmer, S. Müller, T. Pawson, A. C. Gingras, C. H. Arrowsmith, S. Knapp, *Cell* **2012**, *149*, 214–231.

[117] F. Boussouar, M. Jamshidikia, Y. Morozumi, S. Rousseaux, S. Khochbin, *Biochim. Biophys. Acta Gene Regul. Mech.* **2013**, *1829*, 1011–1014.

[118] J. T. Lloyd, K. C. Glass, *J. Cell. Physiol.* **2018**, *233*, 1877–1886.

[119] S. J. Koo, A. E. Fernández-Montalván, V. Badock, C. J. Ott, S. J. Holton, O. von Ahsen, J. Toedling, S. Vittori, J. E. Bradner, M. Gorjánácz, *Oncotarget* **2016**, *7*, 70323–70335.

[120] A. E. Fernández-Montalván, M. Berger, B. Kuropka, S. J. Koo, V. Badock, J. Weiske, V. Puetter, S. J. Holton, D. Stöckigt, A. Ter Laak, *ACS Chem. Biol.* **2017**, *12*, 2730–2736.

[121] Z. Liu, P. Wang, H. Chen, E. A. Wold, B. Tian, A. R. Brasier, J. Zhou, *J. Med. Chem.* **2017**, *60*, 4533–4558.

[122] C. R. Wellaway, D. Amans, P. Bamborough, H. Barnett, R. A. Bit, J. A. Brown, N. R. Carlson, C. W. Chung, A. W. J. Cooper, P. D. Craggs, *J. Med. Chem.* **2020**, *63*, 714–746.

[123] N. Vo, R. H. Goodman, *J. Biol. Chem.* **2001**, *276*, 13505–13508.

[124] M. Catalano, M. Moroglu, P. Balbi, F. Mazzieri, J. Clayton, K. H. Andrews, M. Bigatti, J. Scheuermann, S. J. Conway, D. Neri, *ChemMedChem* **2020**, *15*, 1752–1756.

[125] A. M. Taylor, A. Côté, M. C. Hewitt, R. Pastor, Y. Leblanc, C. G. Nasveschuk, F. A. Romero, T. D. Crawford, N. Cantone, H. Jayaram, *ACS Med. Chem. Lett.* **2016**, *7*, 531–536.

[126] Y. Deng, J. Peng, F. Xiong, Y. Song, Y. Zhou, J. Zhang, F. S. Lam, C. Xie, W. Shen, Y. Huang, L. Meng, X. Li, *Angew. Chem. Int. Ed.* **2020**, *59*, 14965–14972.

[127] J. P. Daguer, C. Zambaldo, D. Abegg, S. Barluenga, C. Tallant, S. Müller, A. Adibekian, N. Winssinger, *Angew. Chem. Int. Ed.* **2015**, *54*, 6057–6061; *Angew. Chem.* **2015**, *127*, 6155–6159.

[128] J. P. Maianti, A. McFedries, Z. H. Foda, R. E. Kleiner, X. Q. Du, M. A. Leissring, W. Tang, M. J. Charron, M. A. Seeliger, A. Saghatelian, D. R. Liu, *Nature* **2014**, *511*, 94–98.

[129] J. P. Maianti, G. A. Tan, A. Vetere, A. J. Welsh, B. K. Wagner, M. A. Seeliger, D. R. Liu, *Nat. Chem. Biol.* **2019**, *15*, 565–574.

[130] J. Scheuermann, C. E. Dumelin, S. Melkko, Y. Zhang, L. Mannocci, M. Jaggi, J. Sobek, D. Neri, *Bioconjugate Chem.* **2008**, *19*, 778–785.

[131] L. Mannocci, S. Melkko, F. Buller, I. Molnàr, J. G. Bianké, C. E. Dumelin, J. Scheuermann, D. Neri, *Bioconjugate Chem.* **2010**, *21*, 1836–1841.

[132] S. Dawadi, N. Simmons, G. Miklossy, K. M. Bohren, J. C. Faver, M. N. Ucisik, P. Nyshadham, Z. Yu, M. M. Matzuk, *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 16782–16789.

[133] D. Kim, Y. Sun, D. Xie, K. E. Denton, H. Chen, H. Lin, M. K. Wendt, C. B. Post, C. J. Krusemark, *Molecules* **2019**, *24*, 2764.

[134] B. N. Tse, T. M. Snyder, Y. Shen, D. R. Liu, *J. Am. Chem. Soc.* **2008**, *130*, 15611–15626.

[135] R. E. Kleiner, C. E. Dumelin, G. C. Tiu, K. Sakurai, D. R. Liu, *J. Am. Chem. Soc.* **2010**, *132*, 11779–11791.

[136] G. Georghiou, R. E. Kleiner, M. Pulkoski-Gross, D. R. Liu, M. A. Seeliger, *Nat. Chem. Biol.* **2012**, *8*, 366–374.

[137] D. G. Brown, G. A. Brown, P. Centrella, K. Certel, R. M. Cooke, J. W. Cuozzo, N. Dekker, C. E. Dumelin, A. Ferguson, C. Fiez-Vandal, *SLAS Discovery* **2018**, *23*, 429–436.

[138] R. K. Y. Cheng, C. Fiez-Vandal, O. Schlenker, K. Edman, B. Aggeler, D. G. Brown, G. A. Brown, R. M. Cooke, C. E. Dumelin, A. S. Doré, *Nature* **2017**, *545*, 112–115.

[139] Z. Wu, T. L. Graybill, X. Zeng, M. Platchek, J. Zhang, V. Q. Bodmer, D. D. Wisnoski, J. Deng, F. T. Coppo, G. Yao, *ACS Comb. Sci.* **2015**, *17*, 722–731.

[140] M. Tyagi, F. Begnini, V. Poongavanam, B. C. Doak, J. Kihlberg, *Chem. Eur. J.* **2020**, *26*, 49–88.