Trio-pharmacophore DNA-encoded Chemical Library for Simultaneous Selection of Fragments and Linkers.

Meiying Cui
Technische Universität Dresden
MIchelle Patino Gaillez
B CUBE
Dzung Nguyen
DyNAbind GmbH https://orcid.org/0000-0001-7144-2525
Stephan Heiden
DyNAbind GmbH
Weilin Lin
B CUBE Center for Molecular Bioengineering, Technische Universität Dresden, 01307 Dresden,Germany
Michael Thompson
Dynabind
Francesco Reddavide
Dynabind
Qinchang Chen
School of Life Sciences and Technology, Tongji University
Yixin Zhang (yixin.zhang1@tu-dresden.de)
B CUBE Center for Molecular Bioengineering, Technische Universität Dresden https://orcid.org/0000-0002-6669-4995

Article

Keywords:

Posted Date: July 22nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1861791/v1

Abstract

The split-and-pool method has been widely used to synthesize chemical libraries of a large size, albeit without the possibility of meaningful quality control. In contrast, a self-assembled DNA-encoded chemical library (DEL) allows us to construct an m x n-member library by mixing an m-member and an n-member pre-purified sub-libraries. Herein, we report a trio-pharmacophore DEL (T-DEL) of m x l x n members through assembling three pre-purified and validated sub-libraries. The middle sub-library is synthesized using DNA-templated synthesis with different reaction mechanisms and designed as a linkage connecting the fragments displayed on the flanking two sub-libraries. In spite of assembling three fragments, the resulting compounds do not exceed the up-to-date standard of molecular weight regarding drug-likeness. The utility of T-DEL has been demonstrated in linker optimization for known binding fragments against trypsin and carbonic anhydrase II and *de novo* selections against matrix metalloprotease-2 and − 9.

Introduction

Discovering small molecular binders against protein targets of interest is important for many biochemical and pharmaceutical research fields. In recent years, DNA-encoded chemical library (DEL) technologies have emerged as a powerful combinatorial method for ligand discovery in industry and academia[1,2]. Using split-and-pool synthesis, DELs of extraordinarily large size can be synthesized, with the stepwise synthesis of each compound barcoded in the attached DNA sequence[3–7]. By using next-generation sequencing (NGS), the identities and enrichment of selected compounds can be revealed. However, these single-pharmacophore DELs (Fig. 1A), in which each DNA molecule displays one compound, have their drawbacks. As the millions and billions of different DNA-encoded compounds cannot be individually purified and characterized, the purity of a DEL decreases with the increase of reaction steps, and the final quality cannot be controlled[7].

Another combinatorial method for ligand discovery is the fragment-based approach. Fragment-based drug discovery (FBDD) identifies low-molecular-weight ligands that bind to different sites on a target protein. The structural information regarding the binding mode of these fragments is commonly determined by X-ray crystallography[8,9] or NMR spectroscopy[9,10] and is then used to design linked fragments as potent binders with drug-like properties. A variation of DEL, the self-assembled DEL, also known as dual-pharmacophore DEL (Fig. 1A), displays two compounds at the 3' and 5' ends of a DNA duplex. It resembles the FBDD approach and can facilitate the discovery of low-molecular-weight fragments. Recently, code transferring methods between two DNA strands have been developed to reveal pairing information of the enriched fragments[11–17]. Dynamic DELs[14,15,18–22] and photo-crosslinking DELs[14,19,23–28] have also been reported to improve the signal-to-noise ratio of selection processes and to allow in-solution DEL selections, respectively. Dual-pharmacophore DEL has the advantage of constructing large libraries with high purity. For example, purifying every compound in a single-pharmacophore library with 1 million members is impractical. However, assembling two 1000-compound

DNA-encoded sub-libraries can also result in a library with 1 million members, and every DNA-encoded compound can be purified by high-pressure liquid chromatography (HPLC) and characterized by mass spectrometry. However, dual-pharmacophore DELs also share the drawback of other FBDD methods, as discovering an optimal linkage between two fragments is always time-consuming and labor-intensive.

Melkko, Neri, and co-workers postulated a triplex DEL in 2007[29] (Fig. 1A). It would result in larger self-assembled DELs, in which every member can be purified and characterized. However, the construction of the triplex DEL has not been reported, as the challenges are not only associated with the library synthesis. The difficulty in finding optimal linkage between two fragments has made it intimidating to develop a general strategy to assemble three fragments with a repertoire of multi-functional scaffolds. Moreover, the resulting compounds will largely exceed the common criteria regarding drug-likeness on the aspect of molecular weight, e.g., the Lipinski rule or the up-to-date molecular mass cut-off based on the properties of orally available small molecules approved in the past decade[30].

In this work, we designed a trio-pharmacophore DEL (T-DEL), in which sub-library B (SL-B) was used as a scaffold to assemble the other two sub-libraries (SL-A and SL-C) (Fig. 1B). The SL-B will not only mediate the distances between the fragments in SL-A and SL-C but also introduce additional contacts with the protein. DNA-templated synthesis (DTS) was used to synthesize the SL-B (Fig. 1C). In conventional DEL syntheses using DTS, the organic compounds are detached from one DNA strand, presented, and encoded on the other DNA strand in the final construct[31–37]. For the SL-B of T-DEL, the small molecular compounds are flanked by two DNA strands, which will be used to assemble the SL-A and SL-C. Using this design, we synthesized a T-DEL with over 20 million members, in which every conjugate is purified via polyacrylamide gel electrophoresis (PAGE) or HPLC and characterized by mass spectrometry. After selection, the fragments revealed from SL-A and SL-C can be connected by the selected linker fragments from SL-B, resulting in potent small molecular binders against the protein target of interest.

Results

Library design and synthesis. The synthetic route of sub-library B (SL-B) is shown in Fig. 1C. We designed two 33-nt oligonucleotides that were partially complementary with 6 plus 12 base pairs. Various bi-functional building blocks were conjugated to the oligonucleotides at the 3' or 5' termini, resulting in 3' and 5' conjugates, respectively (**Figure S1**). Then, the conjugates underwent different DNA-templated reactions between matching functional groups to generate DNA-compound-DNA conjugates. The reactions were monitored by denaturing urea PAGE, and the reaction products were purified from the gel. The molecular weight of the reaction products was confirmed by LC-ESI-MS (Fig. 1D and **Supporting Information 6**). By employing a variety of reaction types, such as amide bond formation, reductive amination, azide-alkyne cycloaddition, Michael addition, and Diels-Alder reaction, we have generated 30 conjugates covering four structural categories (**Figure S2**). Each DNA-compound-DNA conjugate was then encoded by splint ligation. The encoding process was also monitored by denaturing PAGE, and only the successfully encoded conjugate was purified from the gel to ensure the high purity of the library members of SL-B (Fig. 1D).

An 883-member fragment sub-library (SL-A) and an 890-member fragment sub-library (SL-C) were synthesized to form a dynamic dual-pharmacophore DEL[15,18], which are partially complementary by 6 bp. SL-A and SL-B share a 33 bp complementary region, and SL-B and SL-C share a 13 bp complementary region. We examined whether the SL-B can assemble with the SL-A and SL-C to form a stable T-DEL using native DNA PAGE. As shown in Fig. 1E, when SL-A and SL-B (lane 4) or SL-B and SL-C (lane 6) were mixed and allowed to anneal, the bands indicative of the assembled duplexes were observed. As expected, the mixture of SL-A and SL-C did not form a larger complex (lane 5). When all three sub-libraries were mixed and allowed to anneal (lane 7), the highest band corresponding to the assembled trimeric complex was observed.

T-DEL for linker optimization. To investigate the use of T-DEL to optimize linkage between fragment pairs, we performed affinity maturation selections against the model proteins bovine carbonic anhydrase II (CAII) and bovine trypsin with their known ligand pairs. As depicted in **Figure S3A**, we utilized the reported fragment pair of CAII[15], aryl sulfonamide and 3-{5-[3-(trifluoromethyl) phenyl]-2-furyl} acrylic acid (compound A) as single-member SL-A and SL-C, respectively. After assembling with the 30-member SL-B, the T-DEL library was selected against CAII immobilized on solid support (Fig. 2A, target selection). Selection against blank solid support served as a negative control. Selection with SL-B assembled with non-modified SL-A and SL-C was also performed (**Figures S3A** and **2A**, no-ligand target selection). qPCR was used to quantify the amount of each member of SL-B in the three selections with code-specific primers (**Supporting information 7** and **Figure S49**). In Fig. 2A, the enrichment was calculated by normalizing the enrichment profile against no-target selection. As expected, the enrichment of the entire SL-B was higher in the target selection than in the no-ligand target selection, demonstrating that the ligand pair facilitates the interaction of SL-B members with the target.

We have chosen two compounds with the highest enrichment (c1 and c2), two with moderate enrichment (c3 and c4), and one compound, c5, with low enrichment for further off-DNA synthesis and validation. These selected structures from SL-B were used to connect sulfanilamide and compound A, resulting in compounds C-1 to C-5 (Fig. 2B). We also synthesized compound C-0 by connecting sulfanilamide with compound A without a linker moiety. The compounds C-0 to C-5 were validated in an enzyme inhibition assay to measure the IC_{50} values. Sulfanilamide showed an IC_{50} value of 13.36 μM, and compound A exhibited moderate inhibition at 100 μM (Fig. 2C). The compound with the highest enrichment (C-2) displayed a 20-fold improvement in the IC_{50} value (0.67 μM). Compounds with moderate and low enrichment, C-3, C-4, and C-5, exhibited lower inhibitory effects than C1 and C2, agreeing with the selection outcome. Interestingly, C-0 showed the second-highest inhibitory effect (IC_{50} 0.83 μM).

We implemented molecular docking studies to gain more insights into the compounds' binding mechanism and compared the docking poses among the compounds (Figs. 2, S4, S5, and **supporting information 5.1**). As reported previously[38,39], the sulfonamide moiety binds deeply in the catalytic site via coordinating with Zn^{2+}, and forming two hydrogen bonds with Thr198, and one with Pro200 (**Figure S4B**). Conjugation of compound A to sulfanilamide contributed predominantly to the hydrophobic interactions

with the protein, as shown in **Figure S4A**. The sulfanilamide moiety remained well-positioned in the active site in all re-synthesized compounds (C-0 to C-5) (Figs. 2 and S4). We then investigated the docking pose of each compound to understand the different inhibitory effects associated with the linker moieties. The binding pose of C-0 resembled the ligand in the reported crystal structure (PDB:6skv) (Figs. 2D and S5). C-2 adopted a compact conformation in the catalytic pocket, forming five hydrogen bonds with the surrounding residues (Fig. 2E). Also, the large hydrophobic effect and low binding energy may support the highest inhibitory effect of C-2 (**Figure S4A)**. On the contrary, the linker moiety of C-3 and C-4 protruded out of the catalytic pocket (Fig. 2F), which may explain their lower inhibitory effects.

Next, we tested the use of T-DEL for linker optimization with bovine trypsin and its ligand pair, 4-aminomethyl benzamidine and 2-iodophenyl isothiocyanate (compound B), reported by the Neri group in their DEL selection with a dual-pharmacophore library[40]. The selection and decoding strategies are identical to CAII (**Figure S3B**). We have chosen the four highly enriched linker fragments, t1 to t4, and one with low enrichment, t5 (Fig. 3A). The linkers were used to tether the fragment pair to generate small molecules T-1 to T-5 (Fig. 3B). Again, the two fragments were directly conjugated without a linker, resulting in compound T-0. The compounds were evaluated by an enzyme inhibition assay. 4-aminomethyl benzamidine showed an IC_{50} value of 147.23 μM, in agreement with the previous report[40], while compound B alone did not display any detectable inhibitory effect. T-1 to T-4 showed remarkable enhancement in the inhibitory effect, especially T-2 and T-3, displaying 70-fold and 30-fold improvement, respectively. T-0 showed an approximately 9-fold improvement (Fig. 3C).

We further studied the binding modes of compounds to trypsin by molecular docking (Figs. 3, S6, S7, and **supporting information 5.2**). As previously reported[41–43], 4-aminomethyl benzamidine binds to the substrate recognition site and forms hydrogen bonds with the key residue Asp189, the neighboring Ser190, Gly219, and Ser195 (Fig. 3D). In analogy to CAII, conjugating compound B to 4-aminomethyl benzamidine by linkers largely increased the hydrophobic contacts with trypsin, maintaining the binding mode of the benzamidine moiety (**Figure S6**). T-2 and T-4 displayed similarity in terms of the binding site and pose of both fragments, agreeing with their observed inhibitory effects (Figs. 3, S6A, and S7A). Notably, T-3 preserved the binding pose of the single compound B best in all conjugates (Figs. 3 and S7A), and T-5 displayed the worst docking score compared to other resynthesized small molecules (**Figure S6A**).

Together, the affinity maturation selections have demonstrated the capability of T-DEL to guide linker optimization for known fragment pairs.

T-DEL for de novo **selections.** A 23.576 million-member T-DEL (883 x 30 x 890) was constructed to test its utility in *de novo* selections. Matrix metalloproteinases (MMPs) are zinc-dependent endopeptidases capable of degrading and remodeling extracellular matrix components[44,45]. They are attractive therapeutic targets as high expression levels were detected in various diseases, such as inflammatory diseases, and at different stages of cancers, including metastasis, invasion, and angiogenesis[45–51]. As shown in **Figure S3C**, we performed selections against the two gelatinases (human MMP-2 and human

MMP-9) to identify binding fragments for later design and synthesis of small molecule inhibitors. After selection, the three sub-libraries were decoded, and the enrichment was calculated by dividing the post-selection fraction (count/total counts) by the pre-selection fraction (Figs. 4A and S8). Selections using a dynamic dual-pharmacophore DEL with the same members (883x890) were also performed against MMP-2 and MMP-9 to select relevant hits that enrich through different DEL formats (**Figures S3D** and **S8**).

The enrichment profiles of all three sub-libraries have shown similarities between MMP-2 and MMP-9, presumably due to the high structural homology of the two proteins[52,53] (Fig. 4B). To validate the selection outcome, we chose three fragments from SL-A (**66**, **182**, and **693**), three fragments from SL-C (**787**, **826**, and **828**), and three linker fragments with the highest enrichment (**12**, **24**, and **10**), and two linker fragments with moderate enrichment (**1** and **4**) from SL-B for further off-DNA synthesis (Fig. 4). We deployed enzyme inhibition assays of MMP-2 and MMP-9 using a fluorogenic peptide substrate for hit validation. We first measured the inhibition of two enzymes by the fragments. As shown in Figs. 5A and 5B, fragment **182** exhibited the highest inhibitory effect with IC_{50} values of 95.8 μM and 48.1 μM against MMP-2 and MMP-9, respectively. Fragments **693** and **828** displayed IC_{50} values in the high μM range against both targets, while fragment **787** showed a high μM IC_{50} value only against MMP-9. No inhibitory effects could be measured for fragments and **826**.

Since the three sub-libraries were independently decoded and analyzed, in order to identify the best combination of the selected fragments, we synthesized 45 (3x5x3) small molecules covering all possible combinations (**Figure S9**). Figure 5C shows the MW distribution of the 45 compounds. They comply with the current drug-likeness criteria regarding MW, e.g., showing an average MW of 503 Da and 90th percentile of 606 Da, similar to the analysis of all approved drug molecules in the past 20 years[30]. We then assayed the compounds against MMP-2 and MMP-9 (**Figures S10-S12**), and the resulting IC_{50} values are shown in Figs. 5D and 5E. The compounds are grouped by the fragment combinations, and each group has five compounds differing by the linker fragments. Like the enrichment profiles (Fig. 4B), the inhibitory effects of the small molecules showed similar patterns on MMP-2 and MMP-9. For both enzymes, the combinations **182 + 828**, **182 + 787**, **693 + 828**, and **693 + 787** displayed higher inhibitory effects than the other pairs, suggesting a synergistic effect from the combinations of these fragments. **66** and **826**, showing the weakest inhibition as fragments, resulted in weak binders after connecting them with various linkers. In addition to the fragment pairing, the linking moiety also impacts the inhibitory effect. When the fragments from SL-A and SL-C were linked by the linker fragments enriched from selections (**12**, **24**, and **10**), they often showed lower IC_{50} values than those with the controls (**1** and **4**). Compound **693_12_828** displayed IC_{50} values of 10 μM and 15 μM against MMP-2 and MMP-9, respectively, tens-of-fold improvements compared to the two fragments. Compounds **182_12_828** and **182_24_828** also exhibited enhanced inhibitory effects compared to the starting fragments (Fig. 5F).

Since the fragment combinations **182 + 828** and **693 + 828** displayed higher potency than the other combinations, we applied molecular docking to shed light on their binding modes. The study suggested

that the compounds bind to the catalytic domain, accommodating the hydrophobic subsite 1' (S1') of the substrate-binding cleft in both enzymes via fragment **182** or **693**. Moreover, the **828** moiety displayed interaction with the catalytic Zn^{2+} in all compounds (**Supporting information 5.3** and **Figures S13-S17**).

Conclusion

In this work, we have realized the synthesis of a trio-pharmacophore DNA-encoded chemical library (T-DEL). The resulting library has the following features:

I. Every member used in constructing the 23.576 million-member library has been purified and characterized. To our knowledge, this is the largest self-assembled DEL with purified building blocks.
II. As the SL-B can serve as both a binding and linking fragment, it represents an FBDD approach providing both fragment and linker information. In FBDD, linker optimization processes are time-consuming and labor-intensive and often lead to molecules with high MW. By directly connecting the three fragments from T-DEL selection, the compounds from off-DNA synthesis comply with the current drug-likeness criteria regarding MW.
III. With the T-DEL format, it is possible to create a focused library joining only known fragment pairs, as described for CAII and trypsin, to gain insights into linking the fragments. On the other hand, it is also possible to take full advantage of the chemical diversity and explore all the possible binding modalities, as described for the *de novo* selections against MMP-2 and MMP-9.

In summary, the T-DEL strategy has allowed us to optimize the linkers for known fragment pairs and synthesize large DEL for *de novo* identification of fragments and their linking moieties against protein targets of interest. Further molecular docking studies revealed the potential binding mode of fragment pairs tethered by different linkers. A limitation of the T-DEL design is the lack of a code-joining mechanism among the three sub-libraries. Code-transferring[11,14] and code-joining methods[12,13,15–17] between two sub-libraries in dual-pharmacophore DEL can make the pairing information readily available by sequencing the joined codes. In the future, code-transferring and code-joining methods will be investigated to develop an optimal strategy with low interference in library synthesis and selection. Further, increasing the size of linker fragment library (SL-B) is of particular interest via diversifying the scaffold structures through exploring the tool box of DNA-templated synthesis.

Methods

Affinity selections against bovine carbonic anhydrase II (CAII) and bovine trypsin

Affinity maturation of hits for CAII and trypsin were performed by mixing three sub-libraries with the 1:1:1 ratio, with each of the sublibrary B at 0.05 nM in 100 µL selection volume. Non-hit and non-target control selections were performed in parallel to eliminate promiscuous binders.

20 µL of coated and uncoated beads were washed three times with 1 mL of PBST (1x PBS+ 0.05% v/v Tween 20), then incubated with 10 µg/mL herring sperm DNA and the respective library composition according to Figure S3 in 100 µL selection volume for 1 h at RT. Next, the beads were washed three times with 1 mL PBST to remove unbound library members. The beads were suspended in 100 µL elution buffer (10 mM Tris-Cl, pH 8.5) and the bound components were eluted from the protein by heating at 95 °C for 5 min. The eluted library members were then analyzed by quantitative real-time PCR.

Quantitative real-time PCR (qPCR) for analysis of selection output

Primer pairs specifically amplifying each code sequence of the sublibrary B as well as the primer pairs quantifying the total enrichment of sublibrary A, B, and C were used to unveil the enrichment pattern of three selections for CAII and trypsin, respectively. 10 µL of reaction mixture contained 5 µL of 2x SYBR Green I master mix (Quantabio, Massachusetts, USA) 1 µL of primer pair (500 nM final concentration), 3 µL of MilliQ water, and 1 µL of selection output. qPCR was performed with the following protocol: 10 min at 95 °C, then 40 cycles of: 15 sec at 95 °C, 1 min at 53 °C, and 30 sec at 68 °C (SYBR Green I signal acquired at this step), final extension at 68 °C for 2 min, followed by melting procedure from 60 to 95 °C measuring decreasing fluorescence signal at constant interval of time points. A standard curve was generated using a series of known concentrations to correlate with the acquired Ct values. Then the Ct value corresponding to each enriched compound of sublibrary B was converted to enriched amount based on the standard curve (Tables **S2**, **S3**, and Figure **S49**).

De novo selection against MMP-2 and MMP-9 and NGS data processing

Immobilization of MMP-2/ MMP-9 on solid support.

4 µg of MMP-2/MMP-9 was dissolved in 100 µL of selection buffer containing 25 mM Tris-base, 10 mM $CaCl_2$, 150 mM NaCl, 0.05% Tween 20, pH7.5. 10 µL of APMA (100 mM) solution was added to the protein solution and incubated at 37 ºC for 1h. After activation, the solution was added on top of 40 µL pre-washed Ni-NTA resin and incubated for 30 min on a rotary shaker. The supernatant was discarded and the resin was washed three times with selection buffer.

Assembly of the chemically diverse trio-pharmacophore library.

For each independent selection, three sub-libraries were mixed in the selection buffer with final concentration of 40 nM of each sub-library in 100 µL. The mixture was heat denatured at 90 ºC for 1 min and slowly cooled down to allow the formation of the trio pharmacophore library.

Selection against MMP-2/MMP-9

The assembled library was incubated with the protein-bound solid support for 2 h at RT on a rotary shaker. In parallel, selection on bare solid support served as the negative control to exclude promiscuous binders on the resin. After panning, the supernatant was discarded and the resin was washed three times with 1 mL selection buffer. Finally, 100 µL of 50 mM Tris buffer, pH 7.4 was added to resuspend the resin and the resin was incubated at 95 ºC for 5 min to denature the protein, release and collect the bound library members.

The eluted library members were subjected to sample preparation for NGS. NGS preparation was achieved by two-step PCR. First, individual selection was indexed using primers containing unique sequence corresponding to each selection. The amplified product was purified from 2% agarose gel. Second, the purified DNA from all selections were pooled in equal amount and subjected to the 2^{nd} PCR step to attach sequences compatible with the NGS flow cell and the sequencing primers. NGS was performed by Novogene UK with Novaseq 6000. Raw data files (fastq files) obtained were decoded using a custom python script. The sequence reads corresponding to each sample was retrieved by searching for the correct index unique to each selection. Then, the code region was extracted from each sequence and assigned to the corresponding identity. The count of each code in each selection sample was obtained by looping through the total reads and counting the occurrence of each code. Enrichment fold of each compound was calculated by, first, dividing the count by total count to get the abundance, second, the abundance after selection was divided by the abundance of pre-selection to reflect the distribution change of the compound. Next, the enrichment fold was plotted against the count for each member of the library to give the scatter plots as in Figure 4.

Enzyme inhibition assays

Carbonic anhydrase II inhibition assay

CAII inhibition assay was performed to measure the inhibitory effect of resynthesized CBS-conjugates. The esterase activity of CA II with a chromogenic substrate p-nitrophenyl acetate (pNPA) was measured. The rates of hydrolysis were determined by increase of absorption at 410 nm after incubating different compounds with CA II. Resynthesized compounds were diluted in a serial spanning mM to µM concentrations. To 94 µL 1x PBS, pH 7.4 buffer containing 650 nM CAII in each well was added 1 µL of compounds of each concentration in 384-well plate and the plate was incubated for 30 min prior to the measurement. 5 µL of pNPA (50 mM in 20 % DMSO) was added to the plate right before the measurement. The absorption was measured at 37 °C each 10 sec for 15 min using a Beckman Coulter's Paradigm Detection Platform (Brea, USA). The initial V_{max} was obtained from increase of absorption and the V_{max} was plotted against the inhibitor concentrations and logistic fitting was performed to obtain IC_{50} values using Origin 2019b (OriginLab) software.

Trypsin inhibition assay

The inhibition by the benzamidine derivatives against trypsin was measured using an assay probing the enzyme proteolytic activity on the substrate Nα-Benzoyl-DL-arginine-4-nitroanilide hydrochloride (BApNA). Trypsin (2 μM, in PBS, pH 7.4) was incubated with different concentrations of benzamidine derivatives (via a two-fold serial dilution, 10 data points) for 30 min. 5 μL of BApNA solution (10 mM, dissolved in DMSO) were added immediately prior to the measurement. The final total volume was 100 μL per well. The measurements were performed in a Low Binding 384 well plate. The absorption was measured at 410 nm in intervals of 10 sec for 15 min at 37 °C using a Beckman Coulter's Paradigm Detection Platform (Brea, USA). The measurements were performed in triplicate. The activity of trypsin was calculated by the hydrolysis rate of BApNA after subtraction of the background-hydrolysis rate. The curves were plotted as a function of the concentration of the inhibitor against the relative enzyme activity. V_{max} was derived from detected absorption signal for each concentration and logistic fitting was performed to obtain IC_{50} values using Origin 2019b (OriginLab).

MMP-2 and MMP-9 inhibition assay

The inhibition by the small molecule compounds against MMP-2/MMP-9 was measured using an assay probing the enzyme proteolytic activity on the substrate DNP-Pro-Leu-Gly-Met-Trp-Ser-Arg (Enzo Life Sciences). All assays were performed in the assay buffer containing 50 mM Tris, 100 mM NaCl. 5 mM $CaCl_2$ and 0.1% Brij 35, pH 7.5. Human proMMP-2 and proMMP-9 were purchased from Sino biological (Germany). The proMMPs were activated by 1 mM APMA (4-aminophenylmercuric acetate) in the assay buffer at 37 ºC for 2h. After activation, MMP-2 was diluted to 6.5 nM of final assay concentration, and MMP-9 was diluted by buffer to reach 10 nM of final assay concentration. In each well, 49 μL of enzyme solution was incubated with 1 μL of compounds of series of concentration (nM-mM in DMSO) at RT for 30 min. 2 μL of the FRET substrate was added right before the measurement at a final concentration of 25 μM. The rate of hydrolysis was monitored by quantifying the emission at 360 nm (ex: 280 nm). The emission was measured with intervals of 30s for 30 min with Synergy H1 Plate Reader (Agilent, USA). The measurements were done in triplicate. At each concentration of compound, the rate of hydrolysis at initial stage (V_{max}) was calculated and relative enzyme activity was obtained by normalizing the data without inhibitor to 1. Then, curves were plotted as a function of the concentration of the inhibitor against the relative enzyme activity. V_{max} was derived from detected absorption signal for each concentration and logistic fitting was performed to obtain IC_{50} values using Origin 2019b (OriginLab).

Details of library synthesis, compound synthesis, and molecular docking studies are given in the supplementary information.

Declarations

Data availability

All data generated or analysed during this study are included in this published article (and its supplementary information files).

Acknowledgments

We thank Ulrike Hofmann for the technical support and Dr. Kathrin Putke from DKMS Life Science Lab (Germany), Dr. Elena Ferrari, and Meng Chen from Novogene (UK) for supporting on Next-Generation Sequencing. We thank Dr. Katerina Vafia (DyNAbind GmbH, Germany) for the support of the NGS data analysis. We thank Dr. Ekta Kumari for the technical support on the enzyme inhibitions assays.

Conflict of interest

DyNAbind GmbH is commercializing the dynamic dual-pharmacophore DNA-Encoded Library technology reported in this publication. M. Thompson, F. V. Reddavide, and Y. Zhang are shareholders of DyNAbind GmbH.

Author Contributions

M. Cui, and Y. Zhang designed the experiments. M. Cui, M. P. Gaillez, and W. Lin conducted the experiments. Q. Chen performed all molecular docking studies. Q. Chen and D. Nguyen analyzed the docking results. S. Heiden, F. V. Reddavide, and M. Thompson synthesized the 883x890-member dual-pharmacophore DNA-encoded library. All authors contributed in writing the manuscript.

References

1. Lerner, R. A.; Brenner, S. DNA-Encoded Compound Libraries as Open Source: A Powerful Pathway to New Drugs. Angew. Chemie - Int. Ed. 2017, *56* (5), 1164–1165.
2. Goodnow, R. A.; Dumelin, C. E.; Keefe, A. D. DNA-Encoded Chemistry: Enabling the Deeper Sampling of Chemical Space. Nat. Rev. Drug Discov. 2017, *16* (2), 131–147.
3. Halpin, D. R.; Harbury, P. B. DNA Display II. Genetic Manipulation of Combinatorial Chemistry Libraries for Small-Molecule Evolution. PLoS Biol. 2004, *2* (7), e174.
4. Halpin, D. R.; Lee, J. A.; Wrenn, S. J.; Harbury, P. B. DNA Display III. Solid-Phase Organic Synthesis on Unprotected DNA. PLoS Biol. 2004, *2* (7), e175.
5. Halpin, D. R.; Harbury, P. B. DNA Display I. Sequence-Encoded Routing of DNA Populations. PLoS Biol. 2004, *2* (7), e173.
6. Clark, M. A.; Acharya, R. A.; Arico-Muendel, C. C.; Belyanskaya, S. L.; Benjamin, D. R.; Carlson, N. R.; Centrella, P. A.; Chiu, C. H.; Creaser, S. P.; Cuozzo, J. W.; et al. Design, Synthesis and Selection of DNA-Encoded Small-Molecule Libraries. Nat. Chem. Biol. 2009, *5* (9), 647–654.
7. Keller, M.; Schira, K.; Scheuermann, J. Impact of DNA-Encoded Chemical Library Technology on Drug Discovery. Chimia (Aarau). 2022, *76* (5), 388.
8. Blundell, T. L.; Jhoti, H.; Abell, C. High-Throughput Crystallography for Lead Discovery in Drug Design. Nat. Rev. Drug Discov. 2002, *1* (1), 45–54.

9. Jhoti, H.; Cleasby, A.; Verdonk, M.; Williams, G. Fragment-Based Screening Using X-Ray Crystallography and NMR Spectroscopy. Curr. Opin. Chem. Biol. 2007, *11* (5), 485–493.
10. Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering High-Affinity Ligands for Proteins: SAR by NMR. Science 1996, *274* (5292), 1531–1534.
11. Wichert, M.; Krall, N.; Decurtins, W.; Franzini, R. M.; Pretto, F.; Schneider, P.; Neri, D.; Scheuermann, J. Dual-Display of Small Molecules Enables the Discovery of Ligand Pairs and Facilitates Affinity Maturation. Nat. Chem. 2015, *7* (3), 241–249.
12. Oehler, S.; Plais, L.; Bassi, G.; Neri, D.; Scheuermann, J. Modular Assembly and Encoding Strategies for Dual-Display DNA-Encoded Chemical Libraries. Chem. Commun. 2021, *57* (92), 12289–12292.
13. Plais, L.; Lessing, A.; Keller, M.; Martinelli, A.; Oehler, S.; Bassi, G.; Neri, D.; Scheuermann, J. Universal Encoding of next Generation DNA-Encoded Chemical Libraries. Chem. Sci. 2022, *13* (4), 967–974.
14. Zhou, Y.; Li, C.; Peng, J.; Xie, L.; Meng, L.; Li, Q.; Zhang, J.; Li, X. D.; Li, X.; Huang, X.; et al. DNA-Encoded Dynamic Chemical Library and Its Applications in Ligand Discovery. J. Am. Chem. Soc. 2018, *140* (46), 15859–15867.
15. Reddavide, F. V.; Cui, M.; Lin, W.; Fu, N.; Heiden, S.; Andrade, H.; Thompson, M.; Zhang, Y. Second Generation DNA-Encoded Dynamic Combinatorial Chemical Libraries. Chem. Commun. 2019, *55* (26), 3753–3756.
16. Vummidi, B. R.; Farrera-Soler, L.; Daguer, J.; Dockerill, M.; Barluenga, S.; Winssinger, N. A Mating Mechanism to Generate Diversity for the Darwinian Selection of DNA-Encoded Synthetic Molecules. Nat. Chem. 2022, *14* (2), 141–152.
17. Daguer, J.-P.; Zambaldo, C.; Ciobanu, M.; Morieux, P.; Barluenga, S.; Winssinger, N. DNA Display of Fragment Pairs as a Tool for the Discovery of Novel Biologically Active Small Molecules. Chem. Sci. 2015, *6* (1), 739–744.
18. Reddavide, F. V.; Lin, W.; Lehnert, S.; Zhang, Y. DNA-Encoded Dynamic Combinatorial Chemical Libraries. Angew. Chemie Int. Ed. 2015, *54* (27), 7924–7928.
19. Li, G.; Zheng, W.; Chen, Z.; Zhou, Y.; Liu, Y.; Yang, J.; Huang, Y.; Li, X. Design, Preparation, and Selection of DNA-Encoded Dynamic Libraries. Chem. Sci. 2015, *6* (12), 7097–7104.
20. Deng, Y.; Peng, J.; Xiong, F.; Song, Y.; Zhou, Y.; Zhang, J.; Lam, F. S.; Xie, C.; Shen, W.; Huang, Y.; et al. Selection of DNA-Encoded Dynamic Chemical Libraries for Direct Inhibitor Discovery. Angew. Chemie Int. Ed. 2020, *59* (35), 14965–14972.
21. Huang, Y.; Meng, L.; Nie, Q.; Zhou, Y.; Chen, L.; Yang, S.; Fung, Y. M. E.; Li, X.; Huang, C.; Cao, Y.; et al. Selection of DNA-Encoded Chemical Libraries against Endogenous Membrane Proteins on Live Cells. Nat. Chem. 2021, *13* (1), 77–88.
22. Farrera-Soler, L.; Daguer, J.-P.; Raunft, P.; Barluenga, S.; Imberty, A.; Winssinger, N. PNA-Based Dynamic Combinatorial Libraries (PDCL) and Screening of Lectins. Bioorg. Med. Chem. 2020, *28* (10), 115458.
23. Shi, B.; Deng, Y.; Li, X. Polymerase-Extension-Based Selection Method for DNA-Encoded Chemical Libraries against Nonimmobilized Protein Targets. ACS Comb. Sci. 2019, *21* (5), 345–349.

24. Shi, B.; Deng, Y.; Zhao, P.; Li, X. Selecting a DNA-Encoded Chemical Library against Non-Immobilized Proteins Using a "Ligate–Cross-Link–Purify" Strategy. Bioconjug. Chem. 2017, *28* (9), 2293–2301.
25. Gui, Y.; Wong, C. S.; Zhao, G.; Xie, C.; Hou, R.; Li, Y.; Li, G.; Li, X. Converting Double-Stranded DNA-Encoded Libraries (DELs) to Single-Stranded Libraries for More Versatile Selections. ACS Omega 2022, *7* (13), 11491–11500.
26. Zhao, P.; Chen, Z.; Li, Y.; Sun, D.; Gao, Y.; Huang, Y.; Li, X. Selection of DNA-Encoded Small Molecule Libraries against Unmodified and Non-Immobilized Protein Targets. Angew. Chemie - Int. Ed. 2014, *53* (38), 10056–10059.
27. Li, Y.; de Luca, R.; Cazzamalli, S.; Pretto, F.; Bajic, D.; Scheuermann, J.; Neri, D. Versatile Protein Recognition by the Encoded Display of Multiple Chemical Elements on a Constant Macrocyclic Scaffold. Nat. Chem. 2018, *10* (4), 1–8.
28. Sannino, A.; Gironda-Martínez, A.; Gorre, É. M. D.; Prati, L.; Piazzi, J.; Scheuermann, J.; Neri, D.; Donckele, E. J.; Samain, F. Critical Evaluation of Photo-Cross-Linking Parameters for the Implementation of Efficient DNA-Encoded Chemical Library Selections. ACS Comb. Sci. 2020, *22* (4), 204–212.
29. Melkko, S.; Dumelin, C. E.; Scheuermann, J.; Neri, D. Lead Discovery by DNA-Encoded Chemical Libraries. Drug Discov. Today 2007, *12* (11–12), 465–471.
30. Shultz, M. D. Two Decades under the Influence of the Rule of Five and the Changing Properties of Approved Oral Drugs. J. Med. Chem. 2019, *62* (4), 1701–1714.
31. Gartner, Z. J.; Tse, B. N.; Grubina, R.; Doyon, J. B.; Snyder, T. M.; Liu, D. R. DNA-Templated Organic Synthesis and Selection of a Library of Macrocycles. Science (80-.). 2004, *305* (5690), 1601–1605.
32. Kleiner, R. E.; Dumelin, C. E.; Tiu, G. C.; Sakurai, K.; Liu, D. R. In Vitro Selection of a DNA-Templated Small-Molecule Library Reveals a Class of Macrocyclic Kinase Inhibitors. J. Am. Chem. Soc. 2010, *132* (33), 11779–11791.
33. Hansen, M. H.; Blakskjaer, P.; Petersen, L. K.; Hansen, T. H.; Hoøjfeldt, J. W.; Gothelf, K. V.; Hansen, N. J. V. A Yoctoliter-Scale DNA Reactor for Small-Molecule Evolution. J. Am. Chem. Soc. 2009, *131* (3), 1322–1327.
34. Usanov, D. L.; Chan, A. I.; Maianti, J. P.; Liu, D. R. Second-Generation DNA-Templated Macrocycle Libraries for the Discovery of Bioactive Small Molecules. Nat. Chem. 2018, *10* (7), 704–714.
35. Li, Y.; Zhao, P.; Zhang, M.; Zhao, X.; Li, X. Multistep DNA-Templated Synthesis Using a Universal Template. J. Am. Chem. Soc. 2013, *135* (47), 17727–17730.
36. Cao, C.; Zhao, P.; Li, Z.; Chen, Z.; Huang, Y.; Bai, Y.; Li, X. A DNA-Templated Synthesis of Encoded Small Molecules by DNA Self-Assembly. Chem. Commun. 2014, *50* (75), 10997–10999.
37. Petersen, L. K.; Christensen, A. B.; Andersen, J.; Folkesson, C. G.; Kristensen, O.; Andersen, C.; Alzu, A.; Sløk, F. A.; Blakskjær, P.; Madsen, D.; et al. Screening of DNA-Encoded Small Molecule Libraries inside a Living Cell. J. Am. Chem. Soc. 2021, *143* (7), 2751–2756.
38. Babine, R. E.; Bender, S. L. Molecular Recognition of Protein-Ligand Complexes: Applications to Drug Design. Chem. Rev. 1997, *97* (5), 1359–1472.

39. Krishnamurthy, V. M.; Kaufman, G. K.; Urbach, A. R.; Gitlin, I.; Gudiksen, K. L.; Weibel, D. B.; Whitesides, G. M. Carbonic Anhydrase as a Model for Biophysical and Physical-Organic Studies of Proteins and Protein-Ligand Binding. Chem. Rev. 2008, *108* (3), 946–1051.
40. Melkko, S.; Zhang, Y.; Dumelin, C. E.; Scheuermann, J.; Neri, D. Isolation of High-Affinity Trypsin Inhibitors from a DNA-Encoded Chemical Library. Angew. Chemie Int. Ed. 2007, *46* (25), 4671–4674.
41. Plattner, N.; Noé, F. Protein Conformational Plasticity and Complex Ligand-Binding Kinetics Explored by Atomistic Simulations and Markov Models. Nat. Commun. 2015, *6* (1), 7653.
42. Rauh, D.; Klebe, G.; Stubbs, M. T. Understanding Protein-Ligand Interactions: The Price of Protein Flexibility. J. Mol. Biol. 2004, *335* (5), 1325–1341.
43. Buch, I.; Giorgino, T.; De Fabritiis, G. Complete Reconstruction of an Enzyme-Inhibitor Binding Process by Molecular Dynamics Simulations. *Proc. Natl. Acad. Sci. U. S. A.* 2011, *108* (25), 10184–10189.
44. Ratnikov, B. I.; Cieplak, P.; Gramatikoff, K.; Pierce, J.; Eroshkin, A.; Igarashi, Y.; Kazanov, M.; Sun, Q.; Godzik, A.; Osterman, A.; et al. Basis for Substrate Recognition and Distinction by Matrix Metalloproteinases. *Proc. Natl. Acad. Sci. U. S. A.* 2014, *111* (40), E4148–E4155.
45. Vandenbroucke, R. E.; Libert, C. Is There New Hope for Therapeutic Matrix Metalloproteinase Inhibition? Nat. Rev. Drug Discov. 2014, *13* (12), 904–927.
46. Winer, A.; Adams, S.; Mignatti, P. Matrix Metalloproteinase Inhibitors in Cancer Therapy: Turning Past Failures Into Future Successes. Mol. Cancer Ther. 2018, *17* (6), 1147–1155.
47. Fields, G. B. The Rebirth of Matrix Metalloproteinase Inhibitors: Moving Beyond the Dogma. Cells 2019, *8* (9), 20–23.
48. Fischer, T.; Senn, N.; Riedl, R. Design and Structural Evolution of Matrix Metalloproteinase Inhibitors. Chem. - A Eur. J. 2019, *25* (34), 7960–7980.
49. Maola, K.; Wilbs, J.; Touati, J.; Sabisz, M.; Kong, X. D.; Baumann, A.; Deyle, K.; Heinis, C. Engineered Peptide Macrocycles Can Inhibit Matrix Metalloproteinases with High Selectivity. Angew. Chemie - Int. Ed. 2019, *58* (34), 11801–11805.
50. Webb, A. H.; Gao, B. T.; Goldsmith, Z. K.; Irvine, A. S.; Saleh, N.; Lee, R. P.; Lendermon, J. B.; Bheemreddy, R.; Zhang, Q.; Brennan, R. C.; et al. Inhibition of MMP-2 and MMP-9 Decreases Cellular Migration, and Angiogenesis in in Vitro Models of Retinoblastoma. BMC Cancer 2017, *17* (1), 1–11.
51. Mondal, S.; Adhikari, N.; Banerjee, S.; Amin, S. A.; Jha, T. Matrix Metalloproteinase-9 (MMP-9) and Its Inhibitors in Cancer: A Minireview. Eur. J. Med. Chem. 2020, *194*, 112260.
52. Fabre, B.; Ramos, A.; De Pascual-Teresa, B. Targeting Matrix Metalloproteinases: Exploring the Dynamics of the S1' Pocket in the Design of Selective, Small Molecule Inhibitors. J. Med. Chem. 2014, *57* (24), 10205–10219.
53. Lenci, E.; Innocenti, R.; Di Francescantonio, T.; Menchi, G.; Bianchini, F.; Contini, A.; Trabocchi, A. Identification of Highly Potent and Selective MMP2 Inhibitors Addressing the S1' Subsite with D-Proline-Based Compounds. Bioorganic Med. Chem. 2019, *27* (9), 1891–1902.

Figures

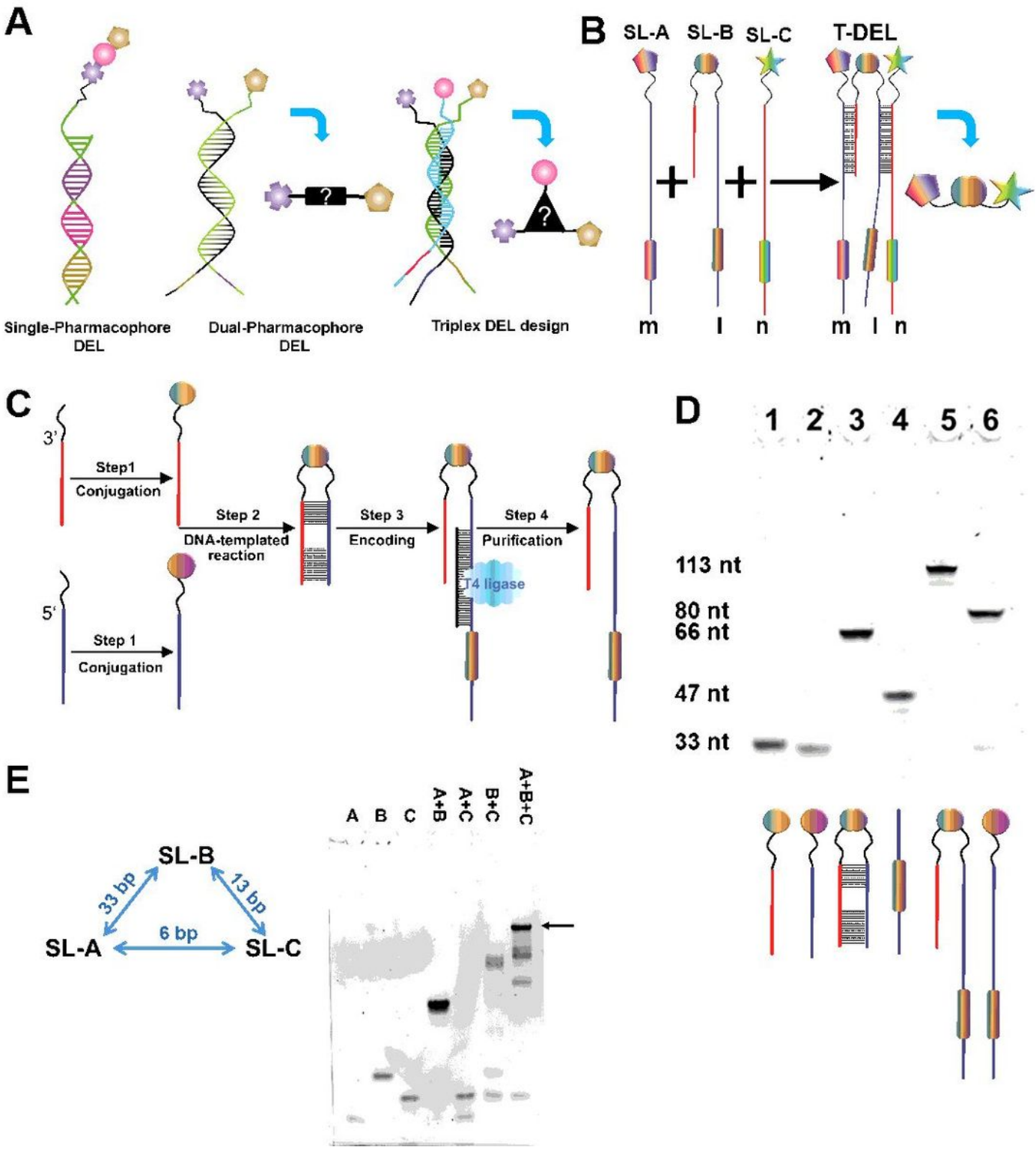

Figure 1

Trio-pharmacophore DNA-encoded chemical library (A) Schematic representation of a single-pharmacophore, a dual-pharmacophore, and a triplex DEL design proposed by Melkko et al. Figure adapted from Melkko et al. 2007. **(B)** Format of the trio-pharmacophore DEL (T-DEL) proposed in this

work. **(C)** Synthetic scheme of sub-library B. The synthesis started from two oligonucleotides with functional groups at the 3' and 5' end. The two oligonucleotides shared 6 + 12 complementary base pairs. **Step 1.** Bi-functional building blocks were conjugated to the oligonucleotides to generate single-side compound-DNA conjugate**. Step 2.** Conjugates from each side underwent DNA-templated reactions via the complementarity and the reactive functional groups from the building blocks. DNA-compound-DNA conjugates harboring both oligonucleotides were synthesized**. Step 3.** Each conjugate was then encoded with a unique barcode by using an adapter DNA (in black) and T4 DNA ligase**. Step 4.** The encoded conjugate was purified from the ligation solution to generate high purity library member of sub-library B. **(D)** Step-by-step monitoring of the library synthesis by urea-denaturing polyacrylamide gel electrophoresis (Urea PAGE). Lane 1 and lane 2 were the single-side conjugates (33 nt). Lane 3 was the DNA-templated reaction product, lane 4 was the barcode DNA (47 nt), and lane 5 was the encoded library member. Lane 6 was the control encoding product of the single-side oligonucleotide (lane 2) and the barcode DNA (lane 4), whose size should be smaller than the encoded conjugate (lane 6) and lager than the conjugate without the barcode (lane 3). The gel was stained by SyBr Green II. The construct of the DNA corresponding to each band is shown under the gel. **(E)** Assembly of three sub-libraries of the trio pharmacophore DNA-encoded chemical library. Native DNA PAGE was used to monitor the assembly of sub-libraries. Lane1-3: Sub-libraries A, B, and C were loaded separately. Lane 4: Mixture of sub-library A and B. Lane 5: Mixture of sub-libraries A and C. Lane 6: Mixture of sub-libraries B and C. Lane 7: Mixture of all three sub-libraries.

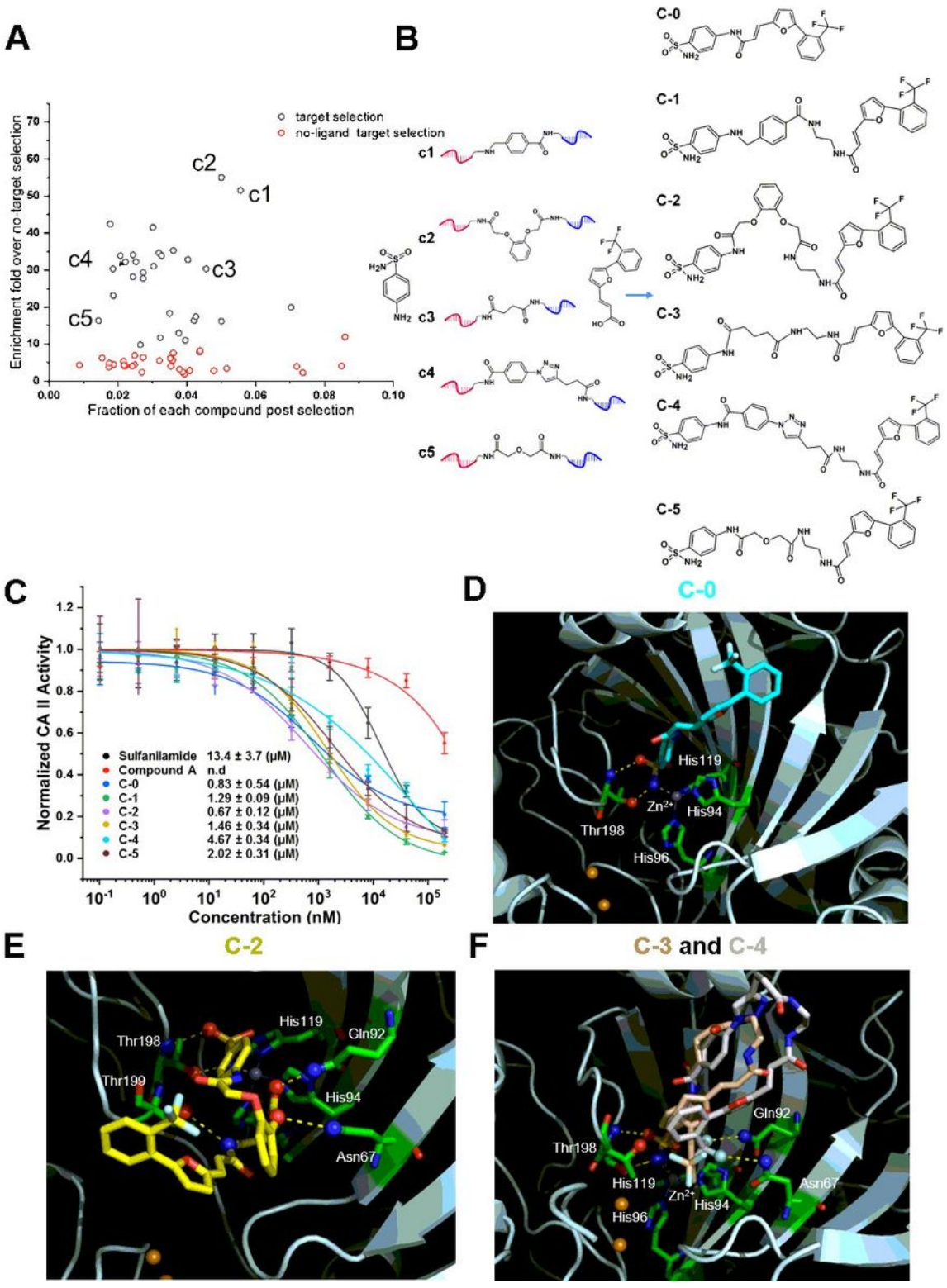

Figure 2

Affinity maturation of ligand pair against bovine carbonic anhydrase II (CAII). (A) Scatter plot of the selection outcome. **(B)** Five conjugates (c1 to c5) enriched from the selection with varying enrichment folds were chosen to link sulfanilamide and compound A, generating small molecules (C-1 to C-5) containing three moieties from the three sub-libraries. C-0 is a direct conjugation between sulfanilamide and compound A. **(C)** Hit validation by an enzyme inhibition assay. **(D, E, and F)** Docking poses of C-0, C-

2, C-3, and C-4 in complex with CAII (PDB ID: 6SKV), accordingly. The protein is in cartoon style, certain residues and the compounds are in stick representation, and the hydrogen bond-forming atoms are in ball representation. Yellow dashed lines indicate hydrogen bonds, and grey dashed lines stand for the coordination with Zn^{2+}. The gray sphere represents Zn^{2+}, and the orange spheres represent Cu^{2+}.

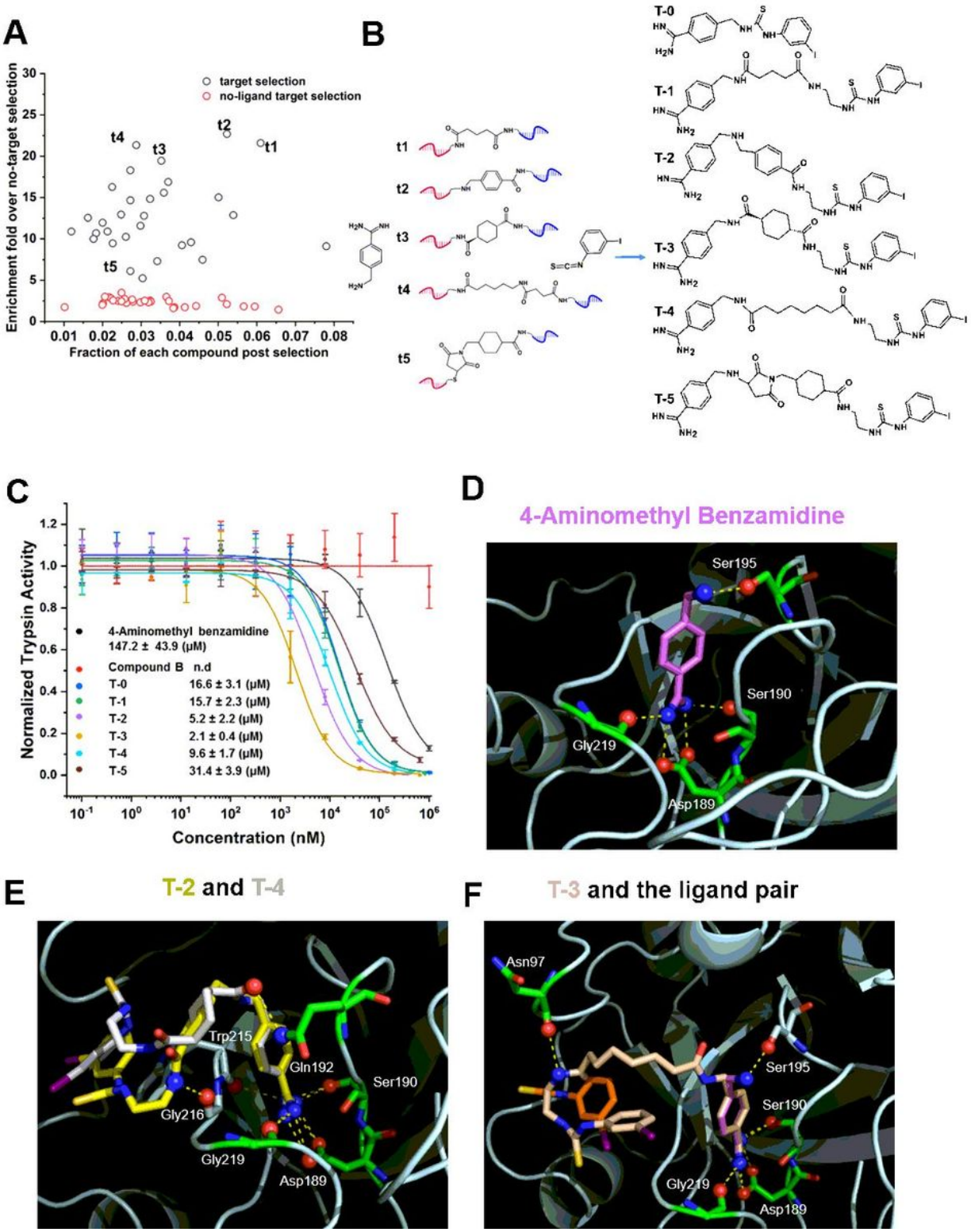

Figure 3

Affinity maturation of ligand pair against bovine trypsin. (A) Scatter plot of the selection result. **(B)** Five conjugates (t1 to t5) enriched from the selection with varying enrichment folds were chosen to link 4-amino methyl benzamidine and compound B, generating small molecules (T-1 to T-5) containing three moieties from the three sub-libraries. T-0 is a direct conjugation between 4-aminomethyl benzamidine and compound B. **(C)** Hit validation by an enzyme inhibition assay. **(D, E, and F)** Docking pose of 4-aminomethyl benzamidine **(D)**, superimposed poses of T-2 and T-4 **(E)**, and the docking pose of T-3 and the ligand pair **(F)** in complex with bovine trypsin (PDB: 1BTY). Compound B is in orange, and 4-aminomethyl benzamidine is in magenta. The protein is in cartoon style, certain residues and the compounds are in stick representation, and the hydrogen bond-forming atoms are in ball representation. Yellow dashed lines indicate hydrogen bonds.

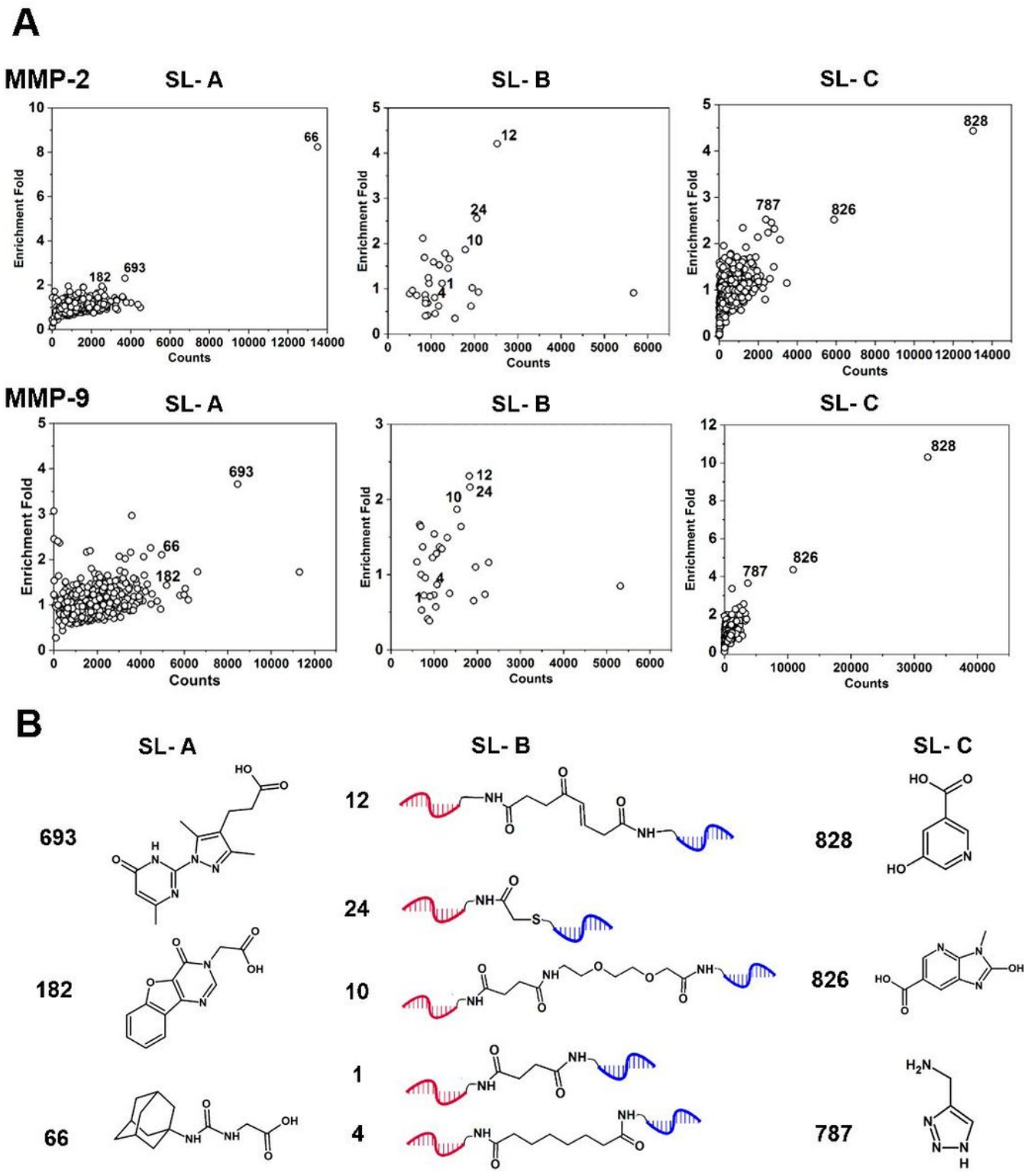

Figure 4

The outcome of the selections against MMP-2 and MMP-9. (A) Scatter plots of selection result from sub-libraries A, B, and C. **(B)** Structures of the selected building blocks from the three sub-libraries for the follow-up hit resynthesis and validation. Fragments **693**, **182**, and **66** were enriched from the SL-A, and fragments **828**, **826**, and **787** were enriched from SL-C. The linker fragments **12**, **24**, and **10** were enriched from SL-B, and **1**, and **4** were chosen to serve as negative controls.

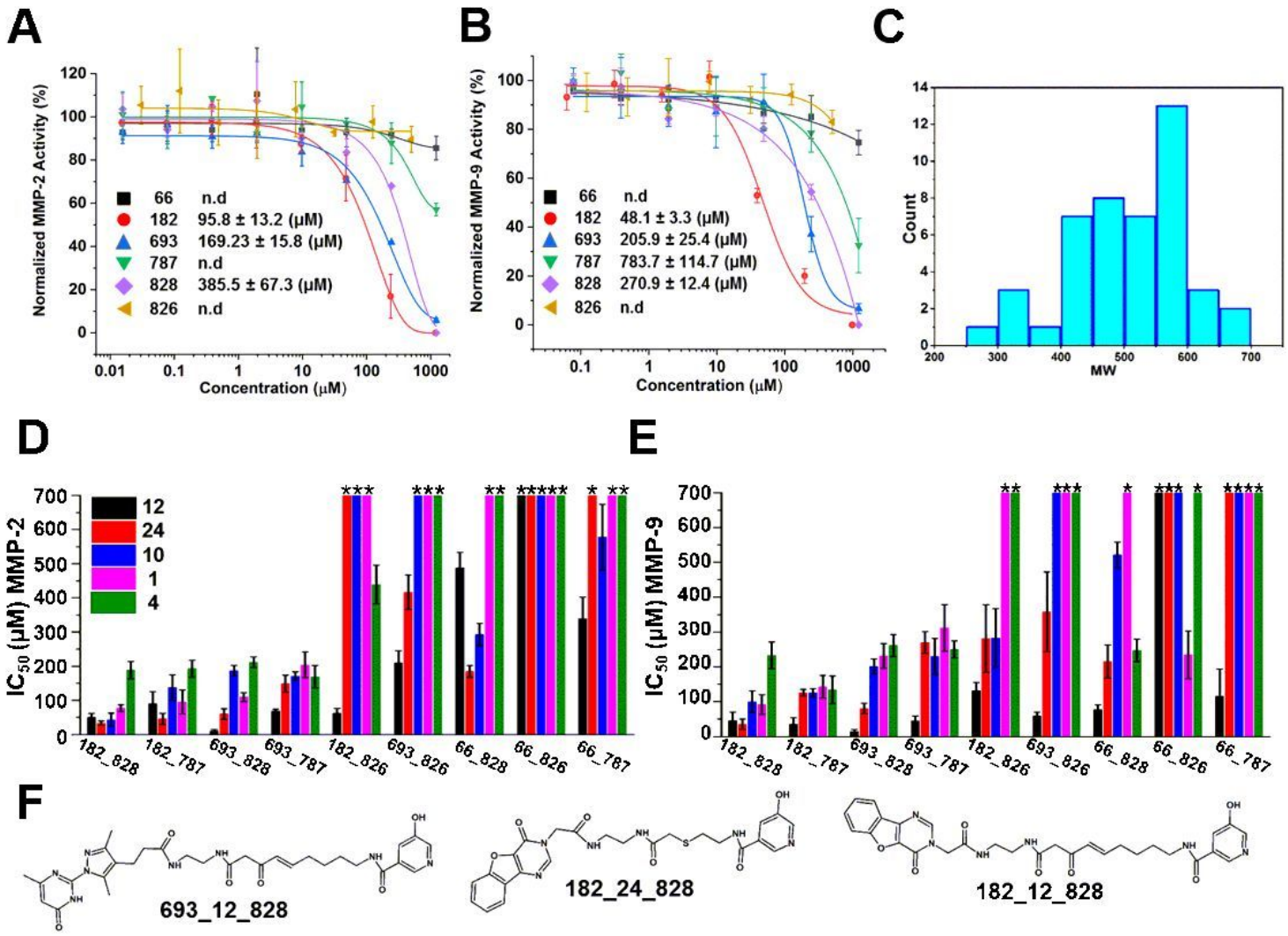

Figure 5

Hit resynthesis and validation. (A and B) Fragments enriched from SL-A (**66**, **182**, and **693**) and SL-C (**787**, **828**, and **826**) were independently validated by enzyme inhibition assay against MMP-2 and MMP-9. **(C)** Histogram of MW distribution of the 45 re-synthesized small molecule compounds. **(D and E)** Columns represent measured IC_{50} values against MMP-2 and MMP-9 of all compounds. The compounds are grouped by the nine fragment combinations. Each group contains five compounds differing by the linker moieties selected from SL-B. **(F)** The structure of compounds with the best inhibitory effect out of the 45 small molecules.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementaryinformation.pdf