



Discovery of potent thrombin inhibitors from a protease-focused DNA-encoded chemical library

Surendra Dawadi^{a,1}, Nicholas Simmons^a, Gabriella Miklossy^a, Kurt M. Bohren^a, John C. Faver^a, Melek Nihan Ucsik^a, Pranavanand Nyshadham^a, Zhifeng Yu^a, and Martin M. Matzuk^{a,1}

^aCenter for Drug Discovery, Department of Pathology & Immunology, Baylor College of Medicine, Houston, TX 77030

Contributed by Martin M. Matzuk, May 27, 2020 (sent for review March 24, 2020; reviewed by Philip E. Dawson and Stanton F. McHardy)

DNA-encoded chemical libraries are collections of compounds individually coupled to unique DNA tags serving as amplifiable identification barcodes. By bridging split-and-pool combinatorial synthesis with the ligation of unique encoding DNA oligomers, million- to billion-member libraries can be synthesized for use in hundreds of healthcare target screens. Although structural diversity and desirable molecular property ranges generally guide DNA-encoded chemical library design, recent reports have highlighted the utility of focused DNA-encoded chemical libraries that are structurally biased for a class of protein targets. Herein, a protease-focused DNA-encoded chemical library was designed that utilizes chemotypes known to engage conserved catalytic protease residues. The three-cycle library features functional moieties such as guanidine, which interacts strongly with aspartate of the protease catalytic triad, as well as mild electrophiles such as sulfonamide, urea, and carbamate. We developed a DNA-compatible method for guanidinylation of amines and reduction of nitriles. Employing these optimized reactions, we constructed a 9.8-million-membered DNA-encoded chemical library. Affinity selection of the library with thrombin, a common protease, revealed a number of enriched features which ultimately led to the discovery of a 1 nM inhibitor of thrombin. Thus, structurally focused DNA-encoded chemical libraries have tremendous potential to find clinically useful high-affinity hits for the rapid discovery of drugs for targets (e.g., proteases) with essential functions in infectious diseases (e.g., severe acute respiratory syndrome coronavirus 2) and relevant healthcare conditions (e.g., male contraception).

DNA-encoded chemical library | focused DECL | protease inhibitor | COVID-19 | SARS-CoV-2

DNA-encoded chemical libraries (DECLs) are a powerful tool for target-directed hit discovery (1–4). By bridging split-and-pool combinatorial synthesis approaches with the ligation of unique encoding DNA oligomers, million- to billion-member libraries may be synthesized at micromole scales for use in hundreds of target screens. Although DECLs are often built as a general screening resource and the significant cost and reagent savings theoretically enable routine screens of an organization's entire DECL collection for any compatible target, in certain cases it is desirable to design and use focused DECLs for a specific target. These focused libraries are particularly suitable for targets in which mechanistic details, structural information, or relevant chemical matter are well-known (5, 6). Use of focused DECLs may bolster chances for robust hit generation as well as enable deeper and broader exploration of structure-activity features if combined with curated building block selection and deep sequencing. Focused DECLs may also empower designs with motifs not normally included in general DECLs such as DECLs with covalent modifiers (7) and DECLs incorporating macrocycles (8) and complex natural products (9).

Proteases, enzymes which catalyze the hydrolysis of peptidic bonds in proteins, have been an intensely studied class of targets for drug discovery due to their involvement in various pathophysiological processes (e.g., protein degradation, protein processing, and signal transduction pathways). Mechanistically,

proteases generally fall into five classes: cysteine proteases, serine proteases, metalloproteases, threonine proteases, and aspartic proteases (10). After activation of the amide with key residues, three of the classes—cysteine, serine, and threonine—utilize the namesake residue to attack the amide carbonyl group, where metalloproteases and aspartic proteases use an activated water. Strategies to directly target proteases with small molecules have generally focused on engaging the active-site catalytic triad, leading to peptidomimetic inhibitors that either react with a nucleophilic residue in a covalent fashion or interact with the critical catalytic residues (11). This strategy has been particularly successful for inhibitors of serine proteases, which are designed to reveal a mild electrophile such as sulfonamide to interact with the serine hydroxyl group and a positively charged guanidine/benzimidazole to interact with a nearby aspartic acid carboxylate (12). Clinically, there have been numerous successes including angiotensin-converting enzyme inhibitors for cardiovascular disorders (13), thrombin inhibitors for thromboembolism and bleeding disorders (14, 15), and HIV protease inhibitors in the treatment of HIV and AIDS (16), among others (17, 18). Recently, viral main protease (M^{pro}, a cysteine protease) has presented as an attractive drug target for the development of novel therapeutics in the COVID-19 (coronavirus disease 2019) pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), and a number of broad-spectrum inhibitors have been developed that target this main protease (19). Another target in the COVID-19 pandemic is human transmembrane serine protease 2 (TMPRSS2), which functions in the processing of the SARS-CoV-2 spike (S) protein

Significance

To rapidly identify small-molecule lead compounds to target healthcare-associated proteases, we constructed a unique 9.8-million-membered protease-focused DNA-encoded chemical library. Affinity selection of this library with a healthcare-relevant protease (i.e., thrombin, a key protein necessary for blood coagulation) revealed potent inhibitors in the first screening attempt. Our results emphasize the utility of a structurally focused DNA-encoded chemical library approach to rapidly uncover hits for healthcare targets (e.g., proteases) where no drug exists (e.g., male contraception) and for emerging diseases (e.g., coronavirus disease 2019).

Author contributions: S.D., N.S., and M.M.M. designed research; S.D., G.M., K.M.B., J.C.F., M.N.U., P.N., and Z.Y. performed research; S.D., N.S., and J.C.F. analyzed data; and S.D., N.S., and M.M.M. wrote the paper.

Reviewers: P.E.D., The Scripps Research Institute; and S.F.M., The University of Texas at San Antonio.

The authors declare no competing interest.

This open access article is distributed under [Creative Commons Attribution License 4.0 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).

¹To whom correspondence may be addressed. Email: dawadi@bcm.edu or mmatzuk@bcm.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2005447117/-DCSupplemental>.

First published July 8, 2020.

(20). Proteases also play important roles in reproduction, and several serine proteases and serine protease-like proteins are essential for normal fertility and therefore could be unique targets for nonhormonal male contraception (i.e., a male pill). For example, among the proteins essential for male fertility are the following: TMPRSS12, a testis-specific transmembrane serine protease required for sperm motility and uterotubal junction migration (21) and structurally similar to TMPRSS2; serine protease 55 (PRSS55), a glycosylphosphatidylinositol-anchored testis-specific serine protease that is necessary for sperm uterotubal junction migration and sperm-ovocyte binding (22, 23); serine protease 37 (PRSS37), a testis-specific inactive serine protease-like protein required for sperm uterotubal junction migration and fertilization (24); and ovoidysmase 2 (OVCH2), an epididymis-specific secreted serine protease required for maturation of sperm (25). Thus, there are multiple healthcare areas that could benefit from advances in the production of unique protease-biased chemical matter for inhibiting this important target class.

Here, we report on a protease-focused DECL and DECL screening strategies designed to engage the protease catalytic triad. Since many of the known protease inhibitors contain guanidine, sulfonamide, urea, and carbamate moieties, we incorporated these functional groups into the design of this library (26, 27). We utilized the structure of argatroban (Fig. 1A), an arginine containing peptidomimetic thrombin inhibitor, as inspiration for the design of this library since it bears a positively charged guanidinium group flanked by two hydrophobic groups (28). Retrosynthetically, argatroban could be partitioned into sections that could be connected in a three-cycle DECL. The piperidine carboxylic acid can be viewed as building block 1 (BB1) which can be covalently linked to DNA via an amide linker. Interestingly, this carboxylate is situated in the direction to open space of the thrombin binding pocket (29), which makes it correctly positioned for a DNA linkage site. The arginine moiety of argatroban can be visualized as BB2 and the aryl sulfonamide as BB3. Thus, the proposed protease DECL (Fig. 1B) could feature a BB1 set of various $-NBoc$ amino acids and $-NBoc$ diamines connected by amide bonds to a DNA-attached amino or

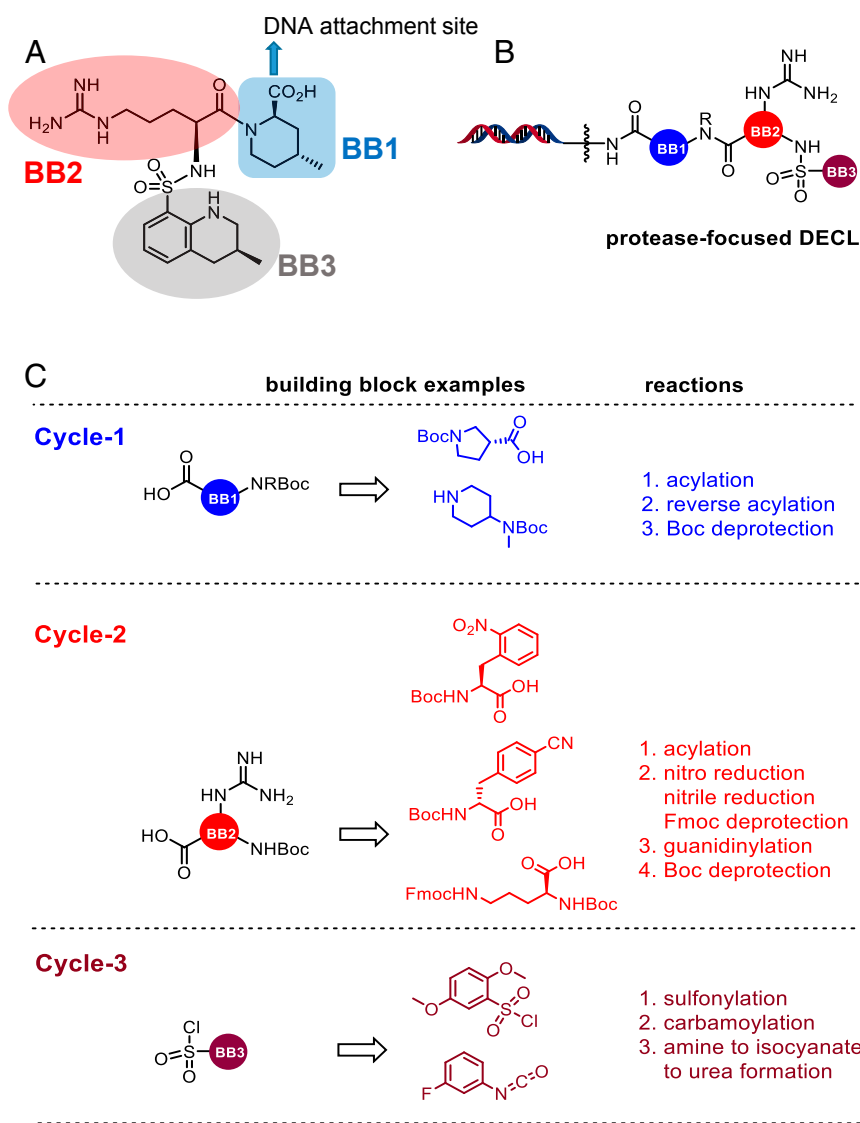


Fig. 1. Protease-focused DECL design. (A) Structure of argatroban, a thrombin inhibitor in clinical use. Argatroban served as a prototype for the design of the protease-focused DECL described in this study. (B) Design of a guanidine containing three-cycle protease-focused DECL. The library features guanidine, as well as a sulfonamide, urea, and carbamate moiety, which are commonly found in protease inhibitors. (C) Examples of various types of building blocks in each cycle of the library build and a list of reactions required to accomplish the synthesis of this library.

carboxyl terminated linker, respectively. After $-NBoc$ deprotection, a variety of different trifunctional cores containing a free carboxyl group, primary $-NHBoc$ amine and orthogonally masked amine ($-Nfmoc$, nitro, or nitrile group) could be connected by amidation to serve as BB2. After deprotection of the orthogonally masked Fmoc group, and reduction of nitro or nitrile groups and subsequent guanidinylation, the remaining $-NHBoc$ amine could be deprotected and functionalized with BB3 through sulfonylation, carbamoylation, and carbamylation (Fig. 1C). Critical to this design is both the synthesis and maintenance of the guanidine

moiety through multistep DECL synthesis. However, to the best of our knowledge there have been no reports of guanidines within published DECLs, although a benzamidine-containing dual pharmacophore DECL has been reported (30).

Results and Discussion

To evaluate the synthetic feasibility of this DECL design, we looked at the availability of DNA-compatible chemical reactions for the construction of this library (Fig. 1C). The acylation reaction, where an amine functionalized DNA is coupled with a

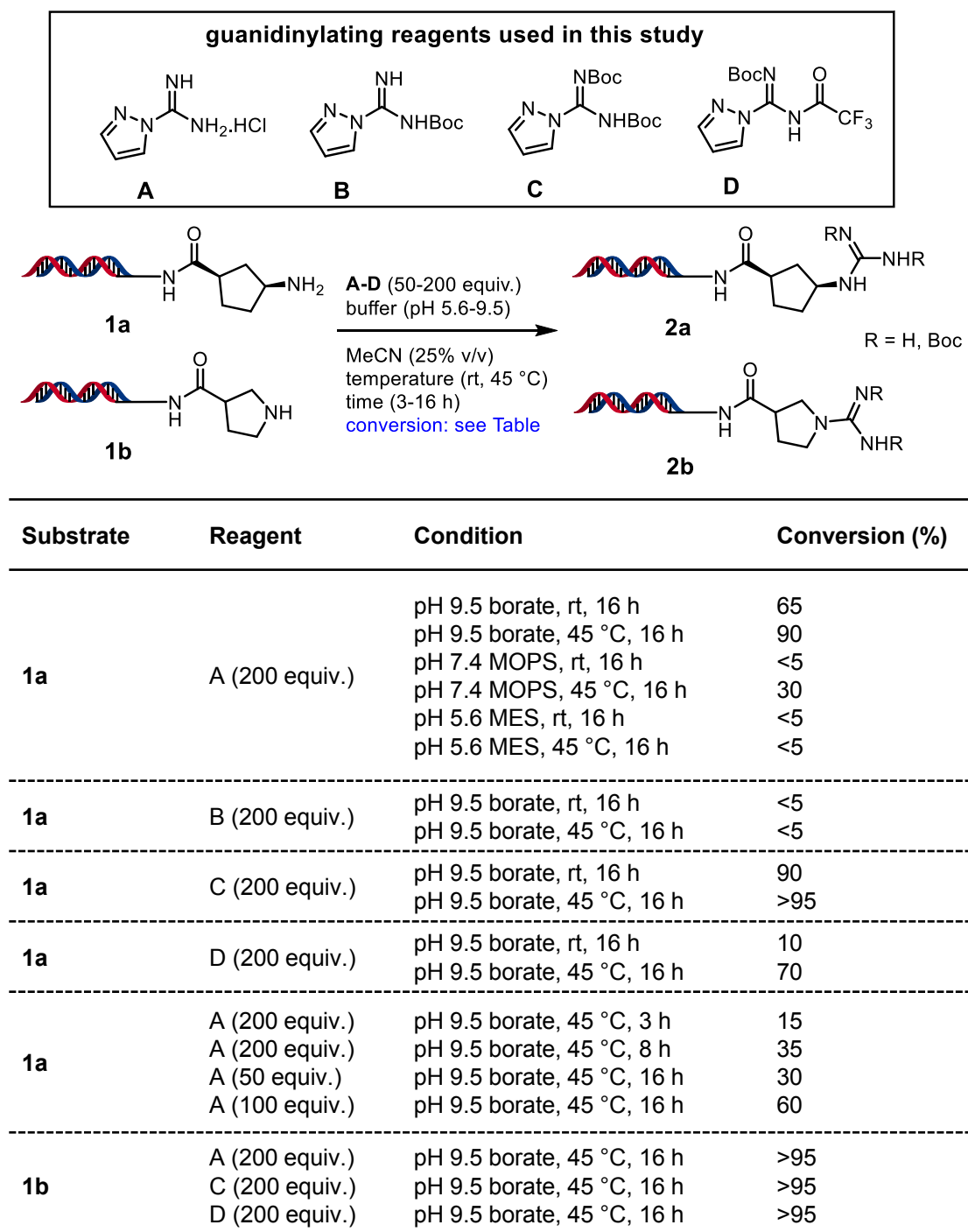


Fig. 2. Optimization of reaction conditions for guanidinylation of amines.

carboxylic acid building block, or vice versa, forming an amide bond is one of most common reactions in DECL technology. Similarly, reports of deprotection conditions of $-NBoc$ and $-Nfmoc$ groups are also common. The reactions of an amine with electrophiles such as sulfonyl chlorides and isocyanates to produce sulfonamides and ureas, respectively, are also known (31). Several conditions for the reduction of aromatic nitro groups to the amine have also been reported (32, 33) Although the conversion of an amine to an isocyanate followed by reaction with nucleophiles such as amine and phenol to form urea and carbamate (34) is known in the presence of DNA, it has not been utilized in the construction of a reported DECL. Overall two of our planned reactions, the conversion of amine to guanidine and the reduction of nitriles to amines, would require optimization to find a general, DNA-compatible process.

Optimization of DNA-Compatible Guanidinylation of Amines and Reduction of Nitriles. To explore the DNA-compatible guanidine formation, we studied four commercially available 1*H*-pyrazole-1-carboxamide-based guanidinylation reagents A through D

(Fig. 2) (35). The DNA linked primary and secondary amine substrates **1a** and **1b** were prepared from the DNA starting material DNA headpiece (DNA-HP **3**; see *SI Appendix, Scheme S1*) by acylation with corresponding $-NBoc$ amino acid followed by the $-NBoc$ deprotection (*SI Appendix, Fig. S2*). When **1a** and **1b** were treated with reagents A through D in various reaction conditions (Fig. 2 and see *SI Appendix* for detailed optimization conditions), we observed formation of guanidine products **2a** and **2b**. Among different buffer systems (pH range from 5.6 to 9.9) studied, pH 9.5 borate buffer gave better conversions. Warming the reaction to 45 °C gave higher conversions in a shorter time. Overall, we determined that 45 °C for 16 h (overnight) in pH 9.5 borate buffer (25% vol/vol MeCN) with either reagent A or C gave excellent conversions with substrate **1a**, whereas substrate **1b** provided excellent conversions with reagents A, C, and D. We employed these optimal conditions to study the substrate scope of guanidinylation. The amines **1c–n** and anilines **1o–u** are obtained from acylation of DNA-HP **3** with commercially available $-NBoc$ amino acids followed by $-NBoc$ deprotection (*SI Appendix, Fig. S2*). We performed the guanidinylation

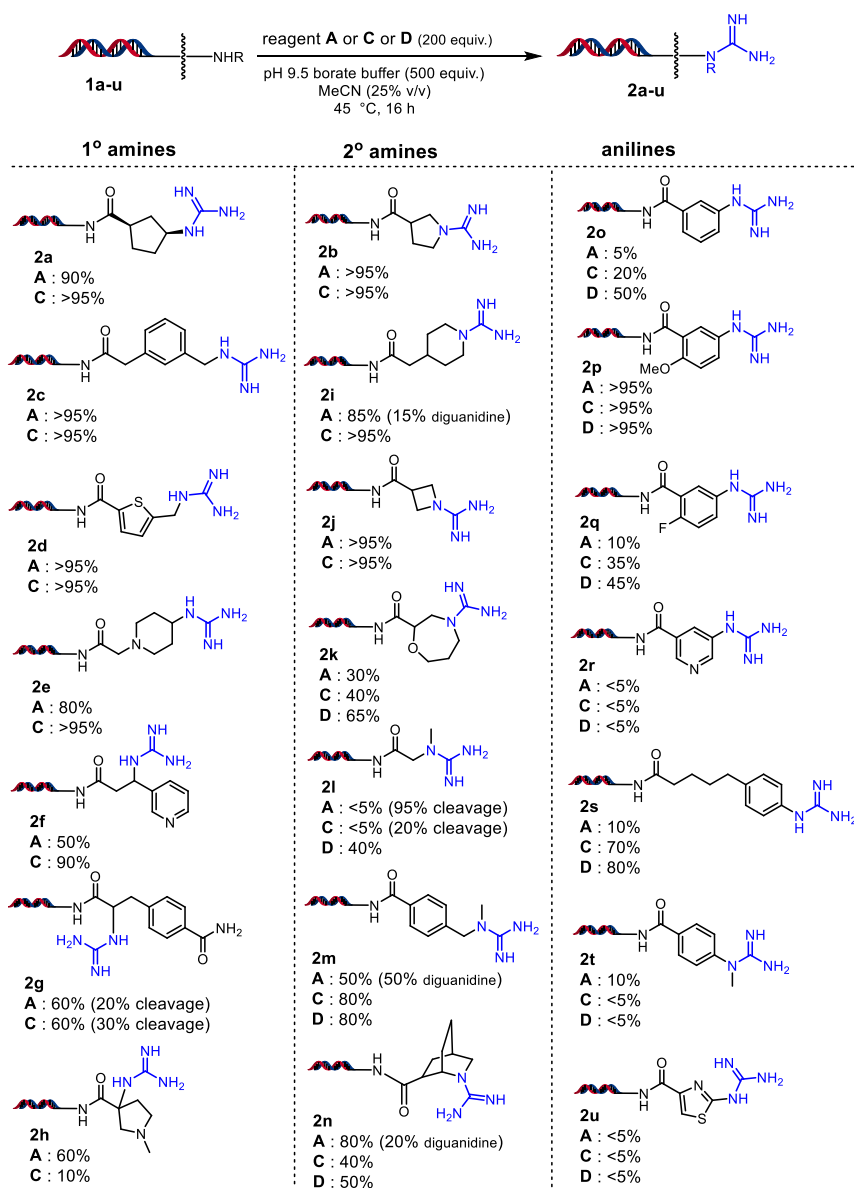


Fig. 3. Scope of DNA-compatible guanidinylation of amines and anilines.

of these substrates and the products **2a–u** were obtained in various conversions (Fig. 3). Primary and unhindered secondary amines readily undergo guanidinylation, whereas anilines have a wide range of reactivity depending upon electronics and sterics. With some secondary amines such as **1k** and most anilines, reagent D gave significantly better conversions. Regardless of the reagent and conditions used, the electron-deficient anilines such as **1g**, **1r**, and **1u** are poor substrates for guanidine formation, which is also observed with the secondary aniline **2t**. Importantly, in the library synthesis setting, we avoided using such building blocks containing electron-deficient anilines in cycle-2 chemistry. Overall, this result suggested that we should use reagent D for guanidinylation of anilines and some secondary amines whereas reagents A and C are more preferable for the guanidinylation of reactive amines.

Next, we studied the reduction of nitriles (see *SI Appendix* for detail optimization condition). Among several conditions studied, the best outcomes were observed while utilizing NaBH_4 as hydride donor in the presence of following additives: Raney Ni and NiCl_2 (*SI Appendix*, Table S2). A mixture of Raney Ni and NaBH_4 in 40% MeOH in pH 5.6 MES [2-(*N*-morpholino)ethanesulfonic acid] buffer reduced both aromatic and aliphatic nitrile groups with excellent conversions within 1 h at room temperature. Also, Raney Ni is known to be compatible with DNA and has been previously employed for nitro reduction (34). One important consideration provided while using Raney Ni was the removal of metal particles by centrifugation before EtOH precipitation of DNA material to prevent DNA damage in subsequent step especially if it is the -NBoc deprotection.

Chemistry Validation and Library Construction. After optimization of the guanidinylation and nitrile reduction, we performed the validation of these two new reactions by the synthesis and analysis of three authentic on-DNA compounds **41–43** (*SI Appendix*, Scheme S2). A more detailed three-cycle chemistry validation process that mimics the library design was performed to prove the feasibility and versatility of our synthetic methods (*SI Appendix*, Scheme S3). Importantly, we need to confirm that the guanidine moiety formed in cycle-2 does not impact cycle-3 amine capping reactions. The validation process started with compound **1b**, which was acylated with six cycle-2 BBs that were

representative of the three kinds of trifunctional groups (i.e., two nitros, two nitriles, and two -NFmoc groups) to afford **44a–f**. The subsequent nitro and nitrile reduction and Fmoc deprotection proceeded with excellent conversion to provide **45a–f**. The guanidinylation of newly formed amines and anilines went smoothly with reagent A or D and the subsequent -NBoc deprotection provided compounds **47a–f**. The newly exposed primary amine underwent *N*-capping reactions with a sulfonyl chloride, isocyanate, or anhydride cycle-3 BB to produce a sulfonamide, urea, or amide bond. In addition to those reactions, the primary amine also underwent isocyanate formation using dipyrindin-2-yl carbonate followed by reaction with amines and phenols to produce the corresponding urea and carbamate. These reactions collectively afforded the final products **48a–i**. The chemistry validation also confirmed that the guanidine has no reactivity toward cycle-3 BBs. Additionally, we confirmed that the reaction conditions have negligible impact on DNA recovery and integrity, which were evaluated with optical density measurement (ultraviolet absorbance at 260 nm) and ligation efficiency.

After establishing all optimized reactions work well in a DECL synthesis setting, the next step was to select the building blocks. Due to two -NBoc deprotection reactions that require heating of the DNA material at 80 °C for >24 h, there are some limitations on cycle-1 BB selection. If we use alpha -NHboc amino acids, the amide bond is cleaved by BB2 amino group during the -NBoc deprotection during cycle 2 (*SI Appendix*, Fig. S4A). Similarly, if we use primary -NHboc amino acids, then the amide -NH reacts with isocyanate intramolecularly during cycle 3 to produce a stable hydantoin (*SI Appendix*, Fig. S4B). These results suggest only nonalpha secondary -NBoc amino acids can be used in cycle 1. For cycle 2, the BB has to be a primary -NHboc amino acid, preferably containing a nitro, nitrile, or -NFmoc group. We also decided to include BBs that do not have this third functional group and do not undergo guanidine formation but preferably have some guanidine bioisosteres. For cycle-3 *N*-capping reactions, a huge collection of BBs is available and they were selected based on diversity calculations and reactivity.

We proceeded to construct the library (Fig. 4 and see *SI Appendix* for detailed methods). The synthesis began with two DNA starting materials, amine **3** and carboxylic acid **4**. They were split

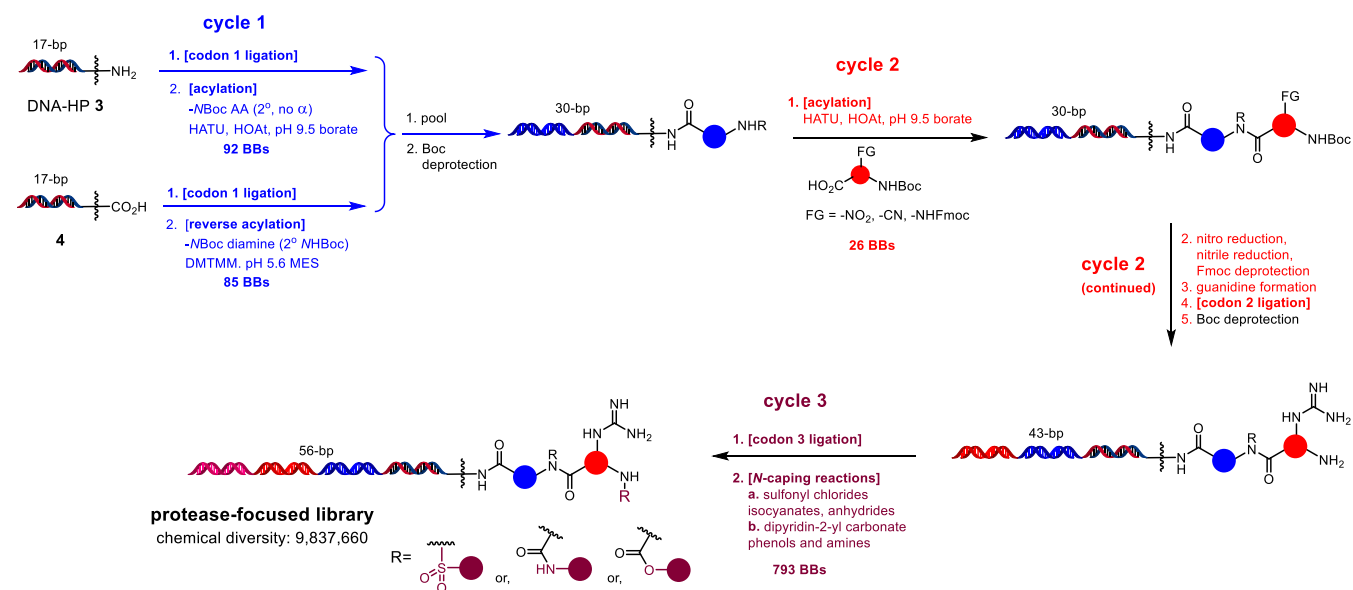


Fig. 4. Construction of a 9.8-million-member protease-focused DECL.

into a desired number of wells, ligated with cycle-1 codons, and reacted with *N*-Boc amino acids and *N*-Boc diamines. Both individual ligations and acylations were monitored by liquid chromatography–mass spectrometry (LC-MS). After the completion of cycle 1, the DNA materials were pooled and *N*-Boc deprotection was performed. The DNA material was split based on the number of cycle-2 BBs and acylation reactions were carried out. Based on the nature of the BBs, nitro reduction, nitrile reduction, Fmoc deprotection, or guanidine formation reactions were performed. After the completion of these cycle-2 reactions, we performed codon-2 ligations which were monitored by gel electrophoresis. After the completion of ligation, DNA materials were pooled and *N*-Boc deprotection was performed. The DNA material was split and codon-3 ligations were carried out followed by the *N*-capping reactions with cycle-3 BBs. Importantly, each cycle pooled material is used as a blank in the next cycle to encode the unreacted materials in those cycles. After completion of the main library build, the entire library material was ligated with a duplexed pair of 12-mer oligonucleotides to encode the library structure/design. Small portions of this tagged library material were further ligated with two oligonucleotides called closing primers to enable later amplification. At this point, the library material was ready for selection experiments.

Affinity Selection with Thrombin. After completion of the build of this 9.8-million-member library and prior to the selection experiment with proteins, the quality of the DECL was evaluated with a naïve sequencing codon distribution report (SI Appendix, Fig. S8A). Further analysis of the relative codon distributions between each of the cycle-2 subpools confirmed that the codons representing all building blocks were observed (SI Appendix, Fig. S8B). The *N*Fmoc codon distribution served as a control for DNA integrity when compared to the nitrile reduction method. The observed codon counts were comparable between the building blocks undergoing mild *N*Fmoc deprotection and the ones undergoing nitrile reduction, which was additional evidence of DNA stability under newly developed reaction conditions. After these evaluations of DNA integrity, we performed a selection experiment to find small-molecule binders of thrombin within the library. Selection assays were performed on streptavidin-coated magnetic beads, which were pretreated with biotinylated human thrombin. The binders were then eluted and the DNA was amplified by PCR and analyzed by next-generation Illumina sequencing in comparison with no-target control (NTC). The enrichment of *n*-synthons (i.e., the encoded association of a chemical building block and unique DNA sequence) for thrombin in comparison with NTC were plotted (Fig. 5A) (36). The enrichment of each *n*-synthon was measured using a normalized z-score metric. The resulting enrichment values (AC-zscore-n) were compared by plotting the enrichment in the target sample (*x* axis) against enrichment in the NTC sample (*y* axis). We observed a strong enrichment of a number of three-cycle trisynthons displaying structure–enrichment relationships. The same trisynthons encoded by different DNA tags were similarly enriched, which confirmed the robustness of the selection result. The enriched features enabled us to identify the structure of the best binders **5** and **6** (Fig. 5B). In the library, these molecules are covalently attached to DNA at the methyl amide site.

Hit Resynthesis and Evaluation of Thrombin Inhibition. The identified hits were first synthesized on-DNA. The successful completion of on-DNA synthesis confirmed that the hits are indeed produced in the library. Also, by performing this we were able to identify any byproducts that might have formed during the library build. Thus, we prepared DNA-**5** and DNA-**6** on DNA-HP using the synthetic sequence exactly as in the library synthesis (Fig. 6A). DNA-HP **3** on acylation with (\pm)-**7** followed by *N*-Boc

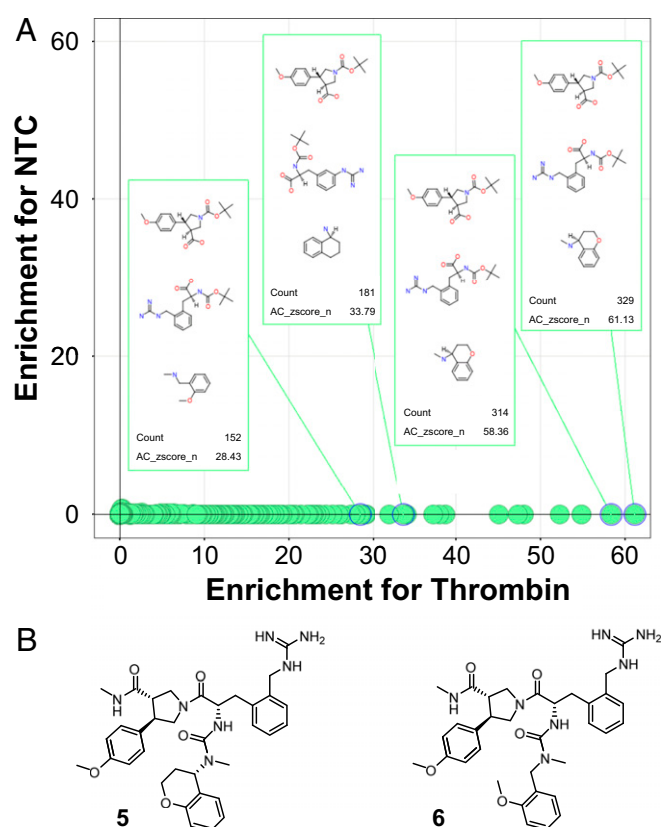


Fig. 5. Discovery of thrombin inhibitors from the protease-focused DECL. (A) Analysis of enrichment of building blocks in the selection of thrombin with the library. The resulting enrichment values (AC-zscore-n) (**36**) were compared by plotting the enrichment in the target sample (*x* axis) against enrichment in the NTC sample (*y* axis). The higher the number in the *x* axis, the stronger the enrichment of synthons. (B) Structure of the top two thrombin hits **5** and **6** identified from the selection experiment. In the library pool, these molecules are attached to the DNA via the linker at the methyl amide site.

deprotection condition afforded **8**. We anticipate the formation of two diastereomers of **8**. Compound **8** underwent amidation with **9** to provide **10**. Nitrile reduction of **10** afforded **11** which underwent guanidinylation followed by *N*-Boc deprotection to provide **12**. The primary amine **12** produced an isocyanate intermediate in the presence of the reagent dipyrindin-2-yl carbonate, which was treated with two amines **13** and **14** to afford urea DNA-**5** and DNA-**6**, respectively. We did not observe formation of any significant by-products along the synthesis and the final products were purified by high-performance LC (HPLC). We studied the thrombin inhibitory potency of DNA-**5** and DNA-**6** (Fig. 6C) and observed that DNA-**5** (58 nM), which is the top hit, is about 10-fold more potent than DNA-**6** (589 nM) (Fig. 6E).

Our analysis of the on-DNA potency of DNA-**5** and DNA-**6** led to the decision to synthesize isomerically pure **5** off-DNA (Fig. 6B). The homochiral 3*S*,4*R*-enantiomer of **7** was the most enriched cycle-1 building block in the selection experiment and we desired to use this pure enantiomer for off-DNA hit resynthesis. The starting material **7** was commercially available as a racemic mixture and the enantiomers of **7** were separated by chiral HPLC with retention times of 9 and 11 min. The absolute configuration of each enantiomer was determined by chemical derivatization to a known compound with reported optical rotation (**37**). Once we established that the enantiomer with retention time of 9 min was the desired 3*S*,4*R* isomer of **7**, this enantiomer was coupled with methylamine using HATU to

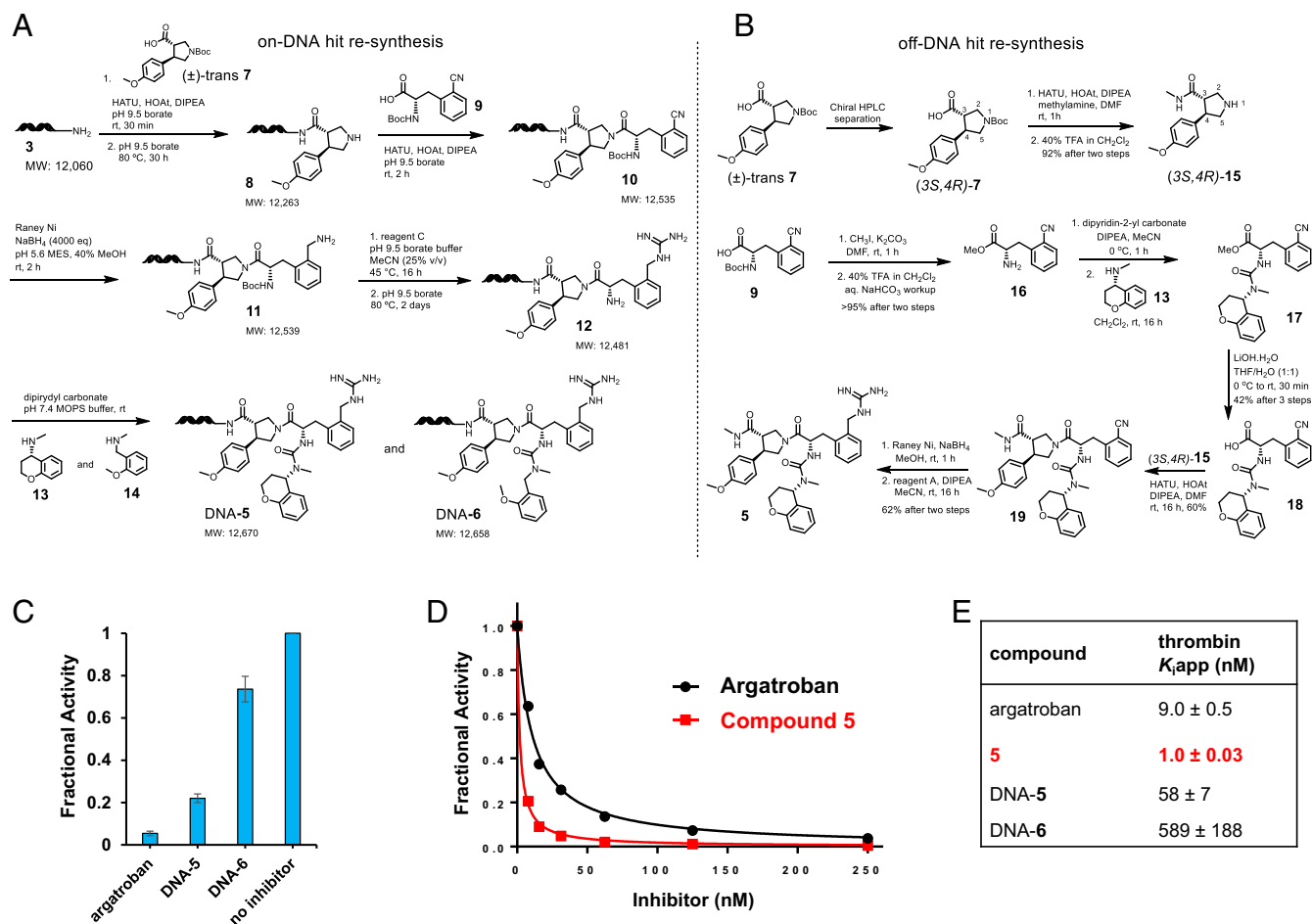


Fig. 6. Synthesis and biological evaluation of thrombin hits. (A) On-DNA resynthesis of two hits to obtain small-molecule DNA-conjugates DNA-5 and DNA-6. (B) Off-DNA resynthesis of the top hit 5. (C) Inhibition of thrombin by DNA-5 and DNA-6 compared with argatroban. (D) Dose–response curves of thrombin inhibition by compound 5 and argatroban. (E) Inhibition constant (apparent K_i) values of synthesized hits. DNA-5 is 10-fold more potent than DNA-6. Compound 5 is 10-fold more potent than argatroban against thrombin.

provide methyl amide, which under $-NBoc$ deprotection condition provided isomerically pure (3*S*,4*R*)-15. The synthesis of the other part of the molecule began with compound **9**, which when treated with iodomethane and K_2CO_3 afforded the methyl ester which in turn underwent $-NBoc$ deprotection to provide free amine **16**. When **16** was treated with dipyridin-2-yl carbonate at 0 °C, the amine was converted to isocyanate, which when treated with freshly prepared secondary amine **13** afforded the urea **17**. The primary amine **16** is prone to react with the isocyanate obtained from itself and produce dimeric urea in 50% yield. Unfortunately, various purification methods were unable to separate the desired product **17** from this by-product. Both **17** and the by-product were taken forward for ester hydrolysis with LiOH at 0 °C. Purification by silica gel chromatography afforded the acid **18** in 42% yield in three steps. We confirmed that compound **18** is a single stereoisomer by optical rotation and NMR, which suggested racemization did not occur during ester hydrolysis. The acid **18** was coupled with amine **15** using HATU to afford **19**, which upon nitrile reduction and guanidinylation with reagent A provided the final product **5**. With purified compound **5** in hand, we examined the inhibitory potency with thrombin. Compound **5** inhibited thrombin very strongly in a dose–response manner (Fig. 6D) with an inhibition constant of 1 nM. Compound **5** is about 10-fold more potent thrombin inhibitor than the clinical drug argatroban (9 nM) (Fig. 6E).

In summary, starting with known protease pharmacophores, we designed and constructed a protease-focused DECL. Selection of this library with thrombin revealed a number of potent thrombin binders. Our best thrombin inhibitor, urea- and guanidine-functionalized compound **5**, showed a remarkable inhibitory potency ($K_i = 1$ nM). The discovery of **5** has led to the initiation of a medicinal chemistry project to further optimize this small molecule thrombin inhibitor with a goal of identifying novel anticoagulants for the treatment of blood clotting complications. In vitro metabolism studies as well as in vivo clotting time studies such as diluted thrombin time and activated partial thromboplastin time are currently ongoing. Additionally, a number of viral, bacterial, and contraception-relevant proteases are undergoing affinity selection with our validated DECL. We anticipate that this approach of utilizing structurally focused DECLs for protease drug discovery can be applied to a wide range of protein targets of therapeutic interest.

Materials and Methods

Procedure for reactions on DNA, protocols of library synthesis, protocol of affinity selection, procedures of hit resynthesis, ultra-performance LC–MS chromatograms of on-DNA compounds, and NMR characterization data of synthetic compounds are available in *SI Appendix*.

Data Availability. The paper and *SI Appendix* contain all datasets generated during this study.

ACKNOWLEDGMENTS. This work is supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (P01HD087157), the Bill and Melinda Gates Foundation (INV-001902), the

Welch Foundation (H-Q-0042), a Core Facility Support Award from the Cancer Prevention Research Institute of Texas (RP160805), and a COVID-19-relevant exploratory grant from Baylor College of Medicine.

1. R. M. Franzini, D. Neri, J. Scheuermann, DNA-encoded chemical libraries: Advancing beyond conventional small-molecule libraries. *Acc. Chem. Res.* **47**, 1247–1255 (2014).
2. D. Neri, Twenty-five years of DNA-encoded chemical libraries. *ChemBioChem* **18**, 827–828 (2017).
3. R. A. Lerner, S. Brenner, DNA-encoded compound libraries as open source: A powerful pathway to new drugs. *Angew. Chem. Int. Ed. Engl.* **56**, 1164–1165 (2017).
4. R. A. Goodnow Jr., C. E. Dumelin, A. D. Keefe, DNA-encoded chemistry: Enabling the deeper sampling of chemical space. *Nat. Rev. Drug Discov.* **16**, 131–147 (2017).
5. L. H. Yuen *et al.*, A focused DNA-encoded chemical library for the discovery of inhibitors of NAD⁺-dependent enzymes. *J. Am. Chem. Soc.* **141**, 5169–5181 (2019).
6. S. Wang *et al.*, Optimization of ligands using focused DNA-encoded libraries to develop a selective, cell-permeable CBX8 chromodomain inhibitor. *ACS Chem. Biol.* **15**, 112–131 (2020).
7. C. Zambaldo, J.-P. Dagher, J. Saarbach, S. Barluengaa, N. Winssinger, Screening for covalent inhibitors using DNA-display of small molecule libraries functionalized with cysteine reactive moieties. *MedChemComm* **7**, 1340–1351 (2016).
8. C. J. Stress, B. Sauter, L. A. Schneider, T. Sharpe, D. Gillingham, A DNA-encoded chemical library incorporating elements of natural macrocycles. *Angew. Chem. Int. Ed. Engl.* **58**, 9570–9574 (2019).
9. P. Ma *et al.*, Functionality-independent DNA encoding of complex natural products. *Angew. Chem. Int. Ed. Engl.* **58**, 9254–9261 (2019).
10. T. Klein, U. Eckhard, A. Dufour, N. Solis, C. M. Overall, Proteolytic cleavage—Mechanisms, function, and “omic” approaches for a near-ubiquitous post-translational modification. *Chem. Rev.* **118**, 1137–1168 (2018).
11. G. Fear, S. Komarnytsky, I. Raskin, Protease inhibitors and their peptidomimetic derivatives as potential drugs. *Pharmacol. Ther.* **113**, 354–368 (2007).
12. F. M. Franco *et al.*, Structure-based discovery of small molecule hepsin and HGFA protease inhibitors: Evaluation of potency and selectivity derived from distinct binding pockets. *Bioorg. Med. Chem.* **23**, 2328–2343 (2015).
13. K. R. Acharya, E. D. Sturrock, J. F. Riordan, M. R. W. Ehlers, Ace revisited: A new target for structure-based drug design. *Nat. Rev. Drug Discov.* **2**, 891–902 (2003).
14. A. Straub, S. Roehrig, A. Hillisch, Oral, direct thrombin and factor Xa inhibitors: The replacement for warfarin, leeches, and pig intestines? *Angew. Chem. Int. Ed. Engl.* **50**, 4574–4590 (2011).
15. E. A. Nutescu, A. K. Wittkowsky, Direct thrombin inhibitors for anticoagulation. *Ann. Pharmacother.* **38**, 99–109 (2004).
16. S. G. Deeks, M. Smith, M. Holodniy, J. O. Kahn, HIV-1 protease inhibitors. A review for clinicians. *JAMA* **277**, 145–153 (1997).
17. A. A. Agbowuro, W. M. Huston, A. B. Gamble, J. D. A. Tyndall, Proteases and protease inhibitors in infectious diseases. *Med. Res. Rev.* **38**, 1295–1331 (2018).
18. A. Eatemadi *et al.*, Role of protease and protease inhibitors in cancer pathogenesis and treatment. *Biomed. Pharmacother.* **86**, 221–231 (2017).
19. L. Zhang *et al.*, Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α -ketoamide inhibitors. *Science* **368**, 409–412 (2020).
20. M. Hoffmann *et al.*, SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280.e8 (2020).
21. T. Larasati *et al.*, TMPRSS12 is required for sperm motility and uterotubal junction migration in mice. *Biol. Reprod.*, 10.1093/biolre/foaa060 (2020).
22. K. Kobayashi *et al.*, Prss55 but not Prss51 is required for male fertility in mice. *Biol. Reprod.*, 10.1093/biolre/foaa041 (2020).
23. X. Shang *et al.*, Serine protease PRSS55 is crucial for male mouse fertility via affecting sperm migration and sperm-egg binding. *Cell. Mol. Life Sci.* **75**, 4371–4384 (2018).
24. C. Shen *et al.*, Prss37 is required for male fertility in the mouse. *Biol. Reprod.* **88**, 123 (2013).
25. D. Kiyozumi *et al.*, NELL2-mediated lumicrine signaling through OVCH2 is required for male fertility. *Science* **368**, 1132–1135 (2020).
26. D. A. Bachovchin, B. F. Cravatt, The pharmacological landscape and therapeutic potential of serine hydrolases. *Nat. Rev. Drug Discov.* **11**, 52–68 (2012).
27. J. Wagner, J. Kallen, C. Ehrhardt, J.-P. Evenou, D. Wagner, Rational design, synthesis, and X-ray structure of selective noncovalent thrombin inhibitors. *J. Med. Chem.* **41**, 3664–3674 (1998).
28. S. Okamoto *et al.*, Potent inhibition of thrombin by the newly synthesized arginine derivative No. 805. The importance of stereo-structure of its hydrophobic carboxamide portion. *Biochem. Biophys. Res. Commun.* **101**, 440–446 (1981).
29. R. Kikumoto *et al.*, Selective inhibition of thrombin by (2*R*,4*R*)-4-methyl-1-[*N*2-[1,2,3,4-tetrahydro-8-quinoliny]sulfonyl]-L-arginyl]-2-piperidinecarboxylic acid. *Biochemistry* **23**, 85–90 (1984).
30. L. Mannocci *et al.*, Isolation of potent and specific trypsin inhibitors from a DNA-encoded chemical library. *Bioconjug. Chem.* **21**, 1836–1841 (2010).
31. H. Deng *et al.*, Discovery of highly potent and selective small molecule ADAMTS-5 inhibitors that inhibit human cartilage degradation via encoded library technology (ELT). *J. Med. Chem.* **55**, 7061–7079 (2012).
32. H.-C. Du *et al.*, DNA-compatible nitro reduction using B₂(OH)₄. *Org. Lett.* **21**, 2194–2199 (2019).
33. H.-C. Du, H. Huang, DNA-compatible nitro reduction and synthesis of benzimidazoles. *Bioconjug. Chem.* **28**, 2575–2580 (2017).
34. A. L. Satz *et al.*, DNA compatible multistep synthesis and applications to dna encoded libraries. *Bioconjug. Chem.* **26**, 1623–1632 (2015).
35. M. S. Bernatowicz, Y. Wu, G. R. Matsueda, 1*H*-Pyrazole-1-carboxamide hydrochloride an attractive reagent for guanylation of amines and its application to peptide synthesis. *J. Org. Chem.* **57**, 2497–2502 (1992).
36. J. C. Faver *et al.*, Quantitative comparison of enrichment from DNA-encoded chemical library selections. *ACS Comb. Sci.* **21**, 75–82 (2019).
37. D. Antermite, D. P. Affron, J. A. Bull, Regio- and stereoselective palladium-catalyzed C(sp³)-H arylation of pyrrolidines and piperidines with C(3) directing groups. *Org. Lett.* **20**, 3948–3952 (2018).