DNA-Encoded Library

DOI: 10.1002/ange.200700654

Isolation of High-Affinity Trypsin Inhibitors from a DNA-Encoded Chemical Library**

*Samu Melkko, Yixin Zhang, Christoph E. Dumelin, Jörg Scheuermann, and Dario Neri**

The isolation of small organic binding molecules specific to proteins of interest is a central problem in chemistry, biology, and medicine. The use of DNA fragments as "barcodes" for the identification of chemical compounds in a library[1] represents an attractive option for the synthesis and screening of large combinatorial libraries. Herein, we describe the use of a DNA-encoded chemical library for the isolation of a family of potent trypsin inhibitors with IC_{50} values in the nanomolar range. The best inhibitor *N*-(4-carbamimidoylbenzyl)-2-(3-(3-(3-iodophenyl)thioureido)phenyl)acetamide displayed a remarkable selectivity among closely related serine proteases, exhibiting a 40- and 6500-fold lower potency towards thrombin and factor Xa, respectively, compared with the inhibition of trypsin.

Biochemical display technologies (such as phage display[2] and ribosome display[3]) routinely allow the isolation of polypeptides with high affinity to virtually any protein target of interest, starting from judiciously designed, large libraries of protein mutants. The success of these technologies is due not only to plasticity of certain proteins (e.g., antibodies) for the recognition of cognate antigens, but also to the fact that libraries of large size can be subjected to affinity-capture procedures, followed by the amplification and identification of enriched library members by virtue of the physical linkage between phenotype (i.e., the antigen binding property) and genotype (i.e., the gene coding for the antibody).

In principle, large collections of chemical compounds, individually coupled to unique oligonucleotides, may allow us to perform similar affinity-capture procedures for small molecular entities followed by PCR-based amplification and identification of preferred binding specificities.[4] Such DNA-encoded chemical libraries may be constructed as single-pharmacophore libraries (in which suitable chemical moieties are added in a stepwise fashion to molecular scaffolds[5–8]) or as dual-pharmacophore chemical libraries (also termed encoded self-assembling chemical (ESAC) libraries[9,10]). This second methodology is particularly suited for the improvement of binding affinity and specificity, starting from lead compounds of modest potency.

In this work, we used ESAC technology for the isolation of potent inhibitors of bovine trypsin, starting from benzamidine, a trypsin inhibitor with an IC_{50} value in approximately the 100 μM range (Figure 1a). Trypsin was chosen as a model protease in view of the wealth of structural and functional information available for this enzyme.[11,12] Furthermore, several trypsin-like serine proteases represent targets of considerable biological and pharmaceutical relevance.[13,14]

We conjugated 5-(4-carbamimidoylbenzylamino)-5-oxopentanoic acid to the 3′ end of an amino-modified 24-mer oligonucleotide (see Scheme 1 in the Supporting Information). The resulting conjugate was allowed to anneal to the cognate radiolabeled oligonucleotide and used in affinity-capture assays on trypsin resins at different coating densities. Figure 1b shows that the benzamidine–DNA conjugate was well retained on affinity resins coated at 2.5 mg mL^{-1}, whereas retention efficiency decreased at lower coating densities, which is similar to what we have previously reported for other protein binders.[15,16] On the basis of these results, we decided to pair the benzamidine–oligonucleotide conjugate to a DNA-encoded library of 620 chemical compounds[15] (each individually coupled to a distinctive oligonucleotide; Figure 1a) and to perform affinity-capture selections on trypsin resins coated at 0.1 and 0.02 mg mL^{-1} density, which corresponds to more stringent selection conditions.

Figure 2a shows representative results of affinity selections performed with the 620 benzamidine derivatives, which were obtained by PCR amplification of the code-containing DNA moieties with a fluorescently labeled primer, followed by hybridization to microarrays spotted in quadruplicate with cognate oligonucleotides corresponding to the 620 different codes. Analysis of the ratios of fluorescence signal intensities for the library compounds after selection on trypsin and on a resin without antigen (Figure 2b) revealed that compounds **73**, **77**, **79**, **437**, and **585** were preferentially enriched after selections performed at both 0.1 and 0.02 mg mL^{-1} coating densities (Figure 2c). Interestingly, three of the five most promising compounds contained a phenylthiourea moiety, which suggests the presence of an exosite compatible with

[*] Dr. Y. Zhang, C. E. Dumelin, Dr. J. Scheuermann, Prof. D. Neri
Department Chemie und Angewandte Biowissenschaften
ETH Zürich
Wolfgang-Pauli-Strasse 10, 8093 Zürich (Switzerland)
Fax: (+41) 44-633-1358
E-mail: neri@pharma.ethz.ch
Dr. S. Melkko
Philochem AG
c/o ETH Zürich
Wolfgang-Pauli-Strasse 10, 8093 Zürich (Switzerland)

[**] We are grateful to Dr. Jens Sobek and Dr. Ralph Schlapbach (Functional Genomics Center Zürich) for help with the spotting of oligonucleotide chips, and we acknowledge financial support from the ETH Zürich, the Bundesamt für Bildung und Wissenschaft for the EU Project "STROMA", the Swiss National Science Foundation, Philogen SpA, and Philochem AG. The authors declare competing financial interests. ESAC technology is covered by a patent application, which was licensed from the ETH Zürich to Philogen under a share of revenues agreement. D.N. owns shares of Philogen and consults for this company. Since November 2006, S.M. is employed by Philochem, a daughter company of Philogen.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

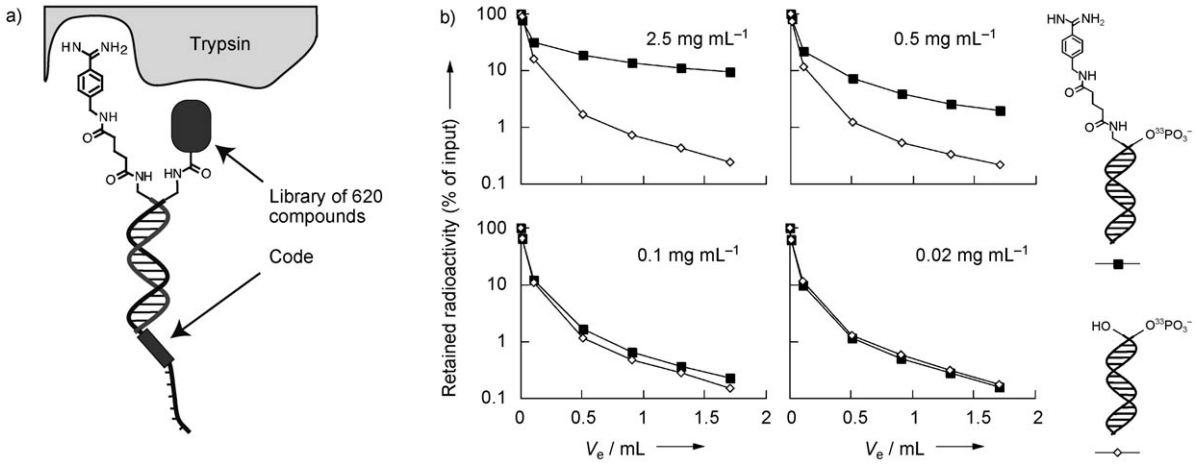


Figure 1. a) Schematic representation of a 24-mer oligonucleotide (covalently conjugated to the weak trypsin inhibitor benzamidine) paired with a library of 620 DNA-encoded chemical compounds, constituting a library of 620 DNA-encoded benzamidine derivatives. b) Benzamidine–oligonucleotide conjugate and the corresponding negative-control oligonucleotide were paired with a cognate radiolabeled oligonucleotide and subjected to an affinity-capture assay by using trypsin-coated sepharose beads of four different coating densities. The radioactivity of aliquots of the input, the flow-through fractions of different washing steps, and of the remaining slurry were determined. V_e = elution volume.

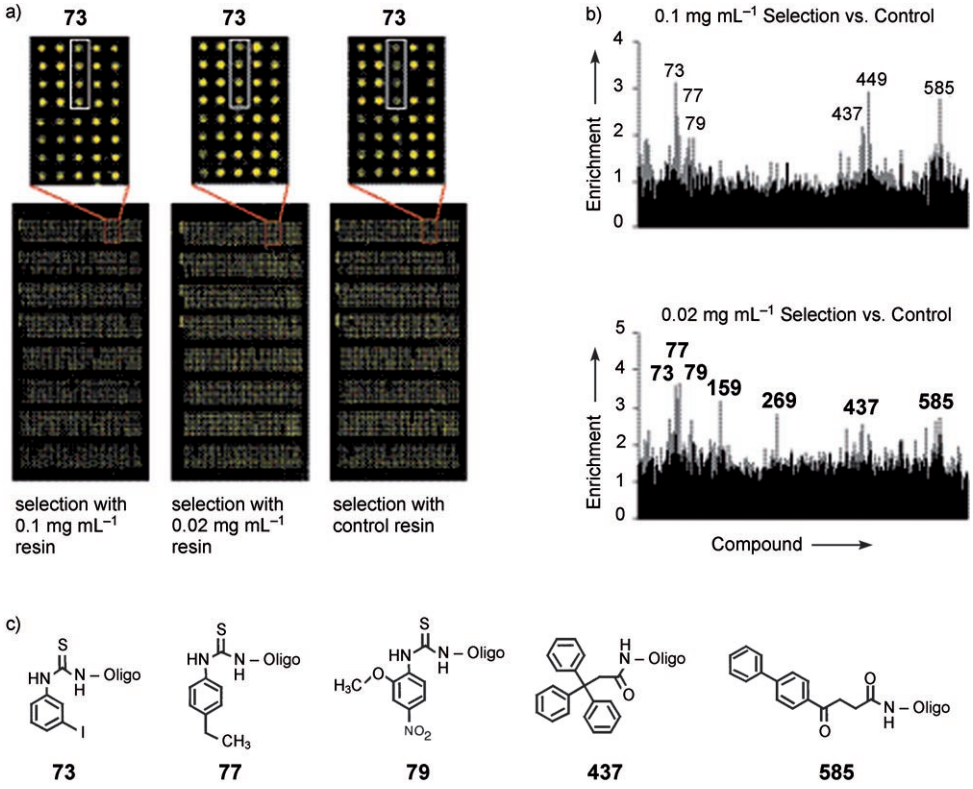
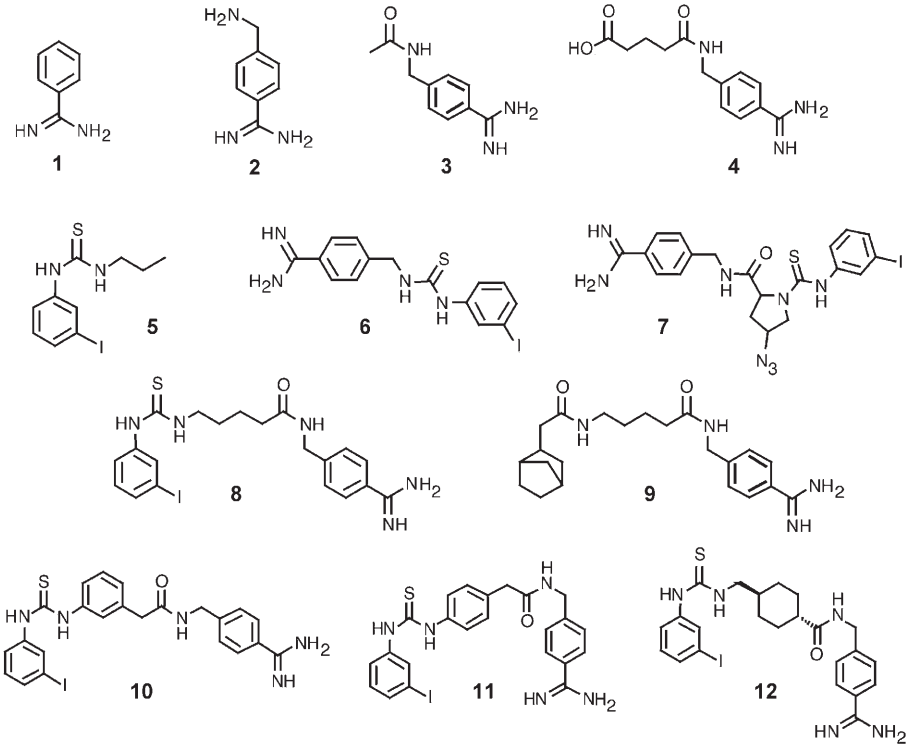


Figure 2. a) Microarray decoding of ESAC affinity maturation selections against trypsin. The microarray slides for selections with trypsin resin of two different coating densities and with negative-control resin are shown. A section of the microarray is magnified, showing the preferential enrichment of the signal for compound **73** when the selections were performed in the presence of trypsin. b) The graphs show ratios of the average microarray signal intensities for individual compounds of selections to trypsin against the negative-control selection. c) Structures of library compounds **73**, **77**, **79**, **437**, and **585**. The signals of these compounds were enriched in the selections to trypsin.

phenylthiourea binding adjacent to the S1 (and benzamidine) binding pocket of trypsin. Compound **73** was chosen for further characterization as it yielded the best results on affinity-capture radioactive assays similar to the ones of Figure 1b (data not shown). Bidentate derivatives consisting of benzamidine and of compound **73** connected by different linkers were synthesized (see Scheme 2 in the Supporting Information) and tested in trypsin inhibition assays with a fluorogenic substrate (see Figure 1 in the Supporting Information).

Scheme 1 shows the chemical structures and the trypsin IC_{50} values for a series of four parental benzamidine derivatives (**1**–**4**), a derivate of the selected compound **73** (**5**), six bidentate conjugates of benzamidine and compound **73** (**6**–**8**, **10**–**12**), and, as a negative control compound, a bidentate conjugate of benzamidine and compound **322** (**9**). The most active inhibitor **10** (*N*-(4-carbamimidoylbenzyl)-2-(3-(3-(3-iodophenyl)thioureido)-phenyl)acetamide) contained a phenyl moiety in the linker and exhibited an IC_{50} value of 98 nM. However, bidentate ligands (e.g., **8**) with flexible alkyl linkers also exhibited a dramatically improved affinity (600 nM), compared with a set of parental benzamidine derivatives (**1**–**4**), whose IC_{50} values were in the 11–220 μM



Compound	IC_{50} / μM	Compound	IC_{50} / μM
1	35	**7**	35
2	90	**8**	0.60
3	11	**9**	12
4	220	**10**	0.098
5	n.d.*	**11**	0.54
6	5.2	**12**	7.7

Scheme 1. Structures of chemical compounds that were tested for their potency to inhibit trypsin. The determination of the IC_{50} values was performed as described in the Supporting Information. * For compound **5**, no inhibition was detected at a compound concentration of 250 μM.

range. Bidentate conjugates of benzamidine and library compounds, which had not been enriched in the affinity selections, displayed IC_{50} values comparable with the one of benzamidine (e.g., 12 μM for **9**, a conjugate of benzamidine and library compound **322**).

The IC_{50} values of compound **10** towards factor Xa and thrombin, two trypsin-like serine proteases, were 640 μM and 4 μM, which corresponds to a 6500- and 40-fold selectivity, respectively, in favor of trypsin (data not shown).

The results presented herein show that 620 DNA-encoded benzamidine derivatives could be rapidly prepared and screened for improved trypsin binders. In principle, the same methodology can be applied to the maturation of any lead compound that can be coupled to an oligonucleotide, for virtually any protein target of interest, without the need to resynthesize the chemical library. After selection decoding and a modular synthesis of bidentate ligands, selective inhibitors with nanomolar potency to trypsin could be rapidly identified. In further studies, it will be interesting to elucidate the structural basis for the improvement of binding affinity conferred by the selected phenylthiourea moiety. As for phage-display technology, a judicious choice of suitable coating density on affinity-capture support is crucial to the success of selection methodologies with DNA-encoded chemical libraries. Such investigations are facilitated by the possibility of performing model selections with radiolabeled DNA derivatives (Figure 1b).

We used oligonucleotide-based microarray technology for the decoding of the selection experiments (Figure 2a), however, in principle other methodologies such as high-throughput sequencing can be considered.[1,17,18] The site-specific immobilization of amino-tagged oligonucleotides on epoxysilane-modified glass slides resulted in highly reproducible fluorescence signal intensities (Figure 2a), which are required for efficient decoding procedures. The essentially flat profile for the ratios of fluorescence signal intensities corresponding to selections performed in the presence or absence of target antigen (Figure 2b) constitutes additional evidence for the robustness of the decoding procedure, facilitating the identification of bidentate binders, which were preferentially enriched with trypsin resin.

In summary, the results discussed here provide evidence that ESAC technology can be used for the rapid affinity maturation of lead protein binders[9] and stimulates research efforts for the development of improved synthetic and decoding methods related to DNA-encoded chemical libraries.

Received: February 12, 2007
Published online: May 11, 2007

Keywords: benzamidines · compound libraries · self-assembly · supramolecular chemistry · trypsin

[1] J. Scheuermann, C. E. Dumelin, S. Melkko, D. Neri, *J. Biotechnol.* **2006**, *126*, 568.
[2] G. Winter, A. D. Griffiths, R. E. Hawkins, H. R. Hoogenboom, *Annu. Rev. Immunol.* **1994**, *12*, 433.
[3] C. Schaffitzel, J. Hanes, L. Jermutus, A. Pluckthun, *J. Immunol. Methods* **1999**, *231*, 119.
[4] S. Brenner, R. A. Lerner, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 5381.
[5] Z. J. Gartner, B. N. Tse, R. Grubina, J. B. Doyon, T. M. Snyder, D. R. Liu, *Science* **2004**, *305*, 1601.
[6] D. R. Halpin, P. B. Harbury, *PLoS Biol.* **2004**, *2*, e174.
[7] D. R. Halpin, P. B. Harbury, *PLoS Biol.* **2004**, *2*, e173.
[8] D. R. Halpin, J. A. Lee, S. J. Wrenn, P. B. Harbury, *PLoS Biol.* **2004**, *2*, e175.
[9] S. Melkko, J. Scheuermann, C. E. Dumelin, D. Neri, *Nat. Biotechnol.* **2004**, *22*, 568.
[10] K. I. Sprinz, D. M. Tagore, A. D. Hamilton, *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3908.
[11] J. J. Perona, C. S. Craik, *J. Biol. Chem.* **1997**, *272*, 29987.
[12] T. A. Steitz, R. G. Shulman, *Annu. Rev. Biophys. Bioeng.* **1982**, *11*, 419.
[13] A. Schweinitz et al., *J. Biol. Chem.* **2004**, *279*, 33613; see the Supporting Information.
[14] B. Turk, *Nat. Rev. Drug Discovery* **2006**, *5*, 785.
[15] C. E. Dumelin, J. Scheuermann, S. Melkko, D. Neri, *Bioconjugate Chem.* **2006**, *17*, 366.
[16] S. Melkko, C. E. Dumelin, J. Scheuermann, D. Neri, *Chem. Biol.* **2006**, *13*, 225.
[17] S. Melkko, J. Sobek, G. Guarda, J. Scheuermann, C. E. Dumelin, D. Neri, *Chimia* **2005**, *59*, 798.
[18] M. Margulies et al., *Nature* **2005**, *437*, 376; see the Supporting Information.

