Anal. Chem. **2009,** *81,* 490–494

Selection of Smart Small-Molecule Ligands: The Proof of Principle

Andrei P. Drabovich, Maxim V. Berezovski, Michael U. Musheev, and Sergey N. Krylov*

Department of Chemistry, York University, Toronto, Ontario M3J 1P3, Canada

The development of drugs and diagnostics with desirable characteristics requires smart small-molecule ligands—ligands with predefined binding parameters of interaction with the target. Here, we propose a general approach for selection of such ligands from highly diverse combinatorial libraries of small molecules by methods of kinetic capillary electrophoresis (KCE). We deduct three fundamental requirements for the combinatorial library to suit the KCE-based selection of smart ligands and suggest a universal design of the library for selecting smart small-molecule ligands: every small molecule in the library is tagged with DNA that encodes the structure of the molecule. Finally, we use several DNA-tagged small molecules, which represent a hypothetical library, to prove experimentally selection of smart small-molecule ligands by the proposed approach.

Small-molecule ligands (molecular weight less than 1000 Da) capable of binding therapeutic and diagnostic targets with high selectivity are used as leads to modern drug candidates and diagnostic probes.[1] Such ligands are typically selected from large combinatorial libraries of small molecules for their binding to a target using affinity methods.[2] The development of drugs with predictable pharmacokinetics and diagnostics with desirable characteristics requires smart ligands—ligands with predefined binding parameters of interaction with the target.[3] Selection of smart ligands is highly challenging and so far has been achieved only by methods of kinetic capillary electrophoresis (KCE) and only for DNA ligands (aptamers).[4,5]

KCE is a conceptual platform for the development of kinetic homogeneous affinity methods, which is defined as CE separation of species that interact during electrophoresis.[6,7] Depending on how the interaction is arranged, different KCE methods can be designed.[6,7] The spectrum of proven applications of kinetic CE methods includes (i) measurements of equilibrium and rate constants of protein−ligand interactions,[8,9] (ii) quantitative affinity analyses of proteins,[3,10] (iii) study of thermochemistry of affinity interactions,[11] and (iv) selection of smart DNA aptamers from libraries of random DNA sequences.[4,5]

In general, to facilitate the selection of smart ligands from a combinatorial library by a KCE method, the combinatorial library of molecules has to satisfy three major requirements. *First*, the library must have high structural diversity to maximize the probability of containing ligands with desirable binding parameters. *Second*, the design of molecules in the library has to be compatible with a detection method to identify structures of ligands selected in a low number of copies. *Third*, the electrophoretic mobility of the library must be different from that of the ligand−target complexes, with targets typically being proteins.

Random DNA libraries, which are used to select DNA aptamers, satisfy all three requirements. First, DNA provides virtually unlimited sequence diversity. Second, polymerase chain reaction (PCR) can be used to amplify even a single DNA molecule to facilitate its further sequence determination. Third, due to a negative charge on every nucleotide, the electrophoretic mobility of DNA is practically always different from those of DNA−protein complexes. As we mentioned above, random DNA libraries have been recently successfully used to select smart DNA ligands (smart aptamers).[4,5] Aptamers have been known for years;[12,13] however, conventional partitioning methods, such as affinity chromatography, membrane filters, and magnetic beads, were not readily suitable for selection of smart aptamers with desired binding parameters.

Libraries of small bare molecules can satisfy requirement 1; for example, a split-and-pool method[14] can generate small-molecule libraries with a practically unlimited diversity. Such libraries cannot, however, satisfy requirements 2 and 3. Indeed, mass spectrometry, the most sensitive method of small-molecule identification, requires at least 10^9 copies of the molecule.[15] Finally, the mobilities of small molecules can be similar to those of the small-molecule−protein complexes. Thus, smart ligands can hardly be selected from libraries of bare small molecules.

* To whom correspondence should be addressed. E-mail: skrylov@yorku.ca.

(1) Keseru, G. M.; Makara, G. M. *Drug Discovery Today* **2006**, *11*, 1711–1748.
(2) Comess, K. M.; Schurdak, M. E. *Curr. Opin. Drug Discovery Dev.* **2004**, *7*, 411–416.
(3) Drabovich, A. P.; Okhonin, V.; Berezovski, M.; Krylov, S. N. *J. Am. Chem. Soc.* **2007**, *129*, 7262–7261.
(4) Drabovich, A.; Berezovski, M.; Krylov, S. N. *J. Am. Chem. Soc.* **2005**, *127*, 11224–11225.
(5) Drabovich, A. P.; Berezovski, M.; Okhonin, V.; Krylov, S. N. *Anal. Chem.* **2006**, *78*, 3171–3178.
(6) Petrov, A.; Okhonin, V.; Berezovski, M.; Krylov, S. N. *J. Am. Chem. Soc.* **2005**, *127*, 17104–17110.
(7) Krylov, S. N. *Electrophoresis* **2007**, *28*, 69–88.
(8) Berezovski, M.; Krylov, S. N. *J. Am. Chem. Soc.* **2002**, *124*, 13674–13675.
(9) Okhonin, V.; Berezovski, M.; Krylov, S. N. *J. Am. Chem. Soc.* **2004**, *126*, 7166–7167.
(10) Berezovski, M.; Nutiu, R.; Li, Y.; Krylov, S. N. *Anal. Chem.* **2003**, *75*, 1382–1386.
(11) Berezovski, M.; Krylov, S. N. *Anal. Chem.* **2005**, *77*, 1526–1529.
(12) Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505–510.
(13) Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818–822.
(14) Goess, B. C.; Hannoush, R. N.; Chan, L. K.; Kirchhausen, T.; Shair, M. D. *J. Am. Chem. Soc.* **2006**, *128*, 5391–5403.
(15) Chen, G.; Pramanik, B. N.; Liu, Y. H.; Mirza, U. A. *J. Mass Spectrom.* **2007**, *42*, 279–287.

10.1021/ac8023813 CCC: $40.75
Published on Web 12/04/2008

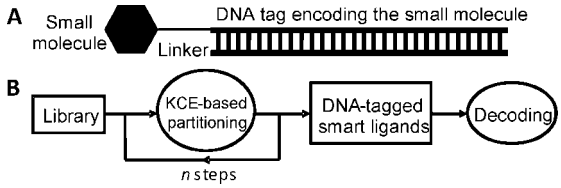

Figure 1. Concept of selection of smart small-molecule ligands: (A) the general design of molecules in the library; (B) major steps of KCE-based selection of smart ligands from such a library.

This work was inspired by the insight that requirements 2 and 3 could be satisfied if every small molecule in the library is tagged with DNA, whose sequence encodes the small-molecule structure (Figure 1A). Such a library is expected to facilitate KCE-based development of smart small-molecule ligands using a general procedure depicted in Figure 1B. A low number of copies of the DNA tag can be amplified by PCR and sequenced, thus revealing the structure of the small molecule it encodes and satisfying requirement 2. Furthermore, a required DNA tag should include a large number of bases (minimum of $\log_4 N$, where 4 is the number of possible bases and N is the number of unique small molecules in the library) and, thus, would be much heavier and have a much higher charge than a small molecule. Therefore, we predict that the influence of the small molecule on the mobility of a DNA tag will be negligible and the electrophoretic properties of such a DNA-encoded library will be similar to those of DNA, for which requirement 3 is satisfied.[3−5]

Ideologically, this work is based on recent developments in (i) selection of naive aptamers by non-SELEX (the procedure of aptamer selection that excludes aptamer amplification between the rounds of selection),[16] (ii) selection of smart aptamers by SELEX (systematic evolution of ligands by exponential enrichment),[4,5] (iii) synthesis of low-diversity DNA-encoded libraries of small molecules,[17−19] and (iv) selection of naive ligands from them by conventional separation methods.[20] Our current work proposes to advance ligand selection technologies to a new level—selection of smart small-molecule ligands. In this proof-of-principle study, we use several DNA-tagged small molecules, which represent several hypothetical libraries, to prove experimentally that the library design and the selection tactic suggested in Figure 1 satisfy the deducted requirements of smart-ligand selection. Although libraries of suitable structural diversity are not yet available, an approach to their synthesis has been already worked out.[17−19] By proving that such library design can be used to select smart small-molecule ligands, this work further stimulates research in the area of DNA-encoded small-molecule libraries and paves the way to the development and practical use of smart small-molecule ligands.

(16) Berezovski, M.; Musheev, M. U.; Drabovich, A. P.; Jitkova, J. V.; Krylov, S. N. *Nat. Protoc.* **2006**, *1*, 1359–1369.
(17) Brenner, S.; Lerner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 5381–5383.
(18) Needels, M. C.; Jones, D. G.; Tate, E. H.; Heinkel, G. L.; Kochersperger, L. M.; Dower, W. J.; Barrett, R. W.; Gallop, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 10700–10704.
(19) Melkko, S.; Dumelin, C. E.; Scheuermann, J.; Neri, D. *Drug Discovery Today* **2007**, *12*, 465–471.

EXPERIMENTAL SECTION

Chemicals and Materials. DNA-tagged cholesterol (**2′**) and desthiobiotin (**4′**) were purchased from IBA GmbH (Göttingen, Germany). Cholesterol was conjugated to the 5′ end of a 60 nt DNA sequence through a tetraethylene glycol linker. Desthiobiotin was conjugated to the 5′ end of a 60 nt DNA sequence through a C6-amino linker. DNA-tagged *N*-benzyl-4-sulfamoyl-benzamide (BSB) (**1′**) and tryptamine (**5′**) were kindly provided by Ensemble Discovery Corp. (Cambridge, MA). BSB was conjugated to the 5′ end of a 48 nt DNA sequence through the following linker [DNA]NC(=O)OCCS(=O)(=O)CCOC(=O)NC(C(=O)O)CCCC NC(=O)C(c1ccccc1)NC(=O)c2ccc(cc2)S(=O)(=O)N. Tryptamine was conjugated to the 5′ end of a 48 nt DNA sequence through the following linker: [DNA]NC(=O)C(N)CCCCNC(=O)C(O) C(O)C(=O)NCCC1=CNc2ccccc21. The bare silica capillary was purchased from Polymicro (Phoenix, AZ). Buffer components were obtained from Sigma-Aldrich (Oakville, ON, Canada). All solutions were made using Milli-Q-quality deionized water and filtered through a 0.22 μm filter (Millipore, Nepean, ON, Canada). DNA-tagged biotin (**3′**), 6-carboxyfluorescein-labeled probe, and nonlabeled primers were obtained from IDT (Coralville, IA). Biotin was conjugated to the 5′ end of a 60 nt DNA sequence through a C6-amino linker. Cholesterol, biotin, desthiobiotin, benzylamine, 4-sulfamoyl-benzoic acid, *N*-(3-dimethylaminopropyl)-*N′*-ethylcarbodiimide (EDC), carbonic anhydrase isozyme II from bovine erythrocytes, streptavidin from *Streptomyces avidinii*, β-lactoglobulin A and B from bovine milk, recombinant *Taq* DNA polymerase, and all other chemicals were from Sigma-Aldrich unless otherwise stated.

Synthesis of *N*-Benzyl-4-sulfamoyl-benzamide. A slightly modified procedure from that described elsewhere was used for synthesis.[21] A mixture of 4-sulfamoyl-benzoic acid (1.0 g; 5 mmol), benzylamine (1.1 mL; 10 mmol), and EDC (2 g; 10.4 mmol) in DMA/CH_2Cl_2 1:1 (4 mL) was stirred for 3 h. Et_2O/EtOAc 1:1 (20 mL) and 5% aqueous $KHSO_4$ (10 mL) were added. The aqueous phase was slightly acidic. After stirring for 5 min, the organic phase was washed twice with 5% aqueous $KHSO_4$ (2 × 5 mL) to convert benzylamine into the water-soluble salt. Then, the organic phase was washed three times with 5% sodium bicarbonate to extract unreacted 4-sulfamoyl-benzoic acid. Finally, the organic phase was washed three times with water (5 mL). The resulting organic phase was dried with $MgSO_4$. Evaporation of the solvent yielded BSB as a white crystalline solid (710 mg; 2.4 mmol; 49%). The relatively low yield can be explained with insufficient excess of coupling reagent (EDC) and losses during the extraction procedure. However, CE analysis requires minute amounts of BSB (5 μL of 1 mM solution, or 1.5×10^{-3} mg). Thus, the primary goal was to obtain a product with high purity, rather than a product with excellent yield.

A ^{1}H NMR spectrum of the product in DMSO-d_6 at 400 MHz was obtained (δ ppm.): 9.24 (t, 1H, NH), 8.05 (d, 2H, HCOAr), 7.92 (d, 2H, HCOAr), 7.50 (s, 2H, NH_2), 7.35 (m, 4H, HBn), 7.26 (m, 1H, HBn), 4.50 (d, 2H, CH_2). The spectrum was almost

(20) Dumelin, C. E.; Scheuermann, J.; Melkko, S.; Neri, D. *Bioconjugate Chem.* **2006**, *17*, 366–370.
(21) Rodenko, B.; Detz, R. J.; Pinas, V. A.; Lambertucci, C.; Brun, R.; Wanner, M. J.; Koomen, G. J. *Bioorg. Med. Chem.* **2006**, *14*, 1618–1629.

identical to the one published in the literature.[21] With the use of the NMR spectrum, the purity of BSB was estimated as 97%.

Electrophoretic Conditions. All experiments were conducted with an MDQ-PACE instrument (Beckman-Coulter) with either fluorescence or absorbance detectors. The length and the inner diameter of the uncoated bare silica capillary were 80 cm and 75 μm, respectively. A single buffer solution, 50 mM Tris-acetate at pH 8.2, was used for sample preparation and for capillary electrophoresis unless otherwise stated. Bare small molecules at concentrations of approximately 500 μM were detected with light absorbance at either 214 or 280 nm. Proteins (50 μM) were detected with light absorbance at 280 nm. DNA-tagged small molecules were hybridized to fluorescently labeled cDNA in the above buffer containing 50 mM NaCl. Mixtures of proteins and DNA-tagged ligands were incubated at 20 °C for 30 min to form equilibrium mixtures. A 150 nL plug of the equilibrium mixture was injected into the capillary by pressure and subjected to NECEEM or ECEEM at 37 °C and at an electric field of 350 V/cm with a positive electrode at the injection end of the capillary. All ECEEM experiments were conducted in the capillary prefilled with the buffer containing 500 nM carbonic anhydrase. A fraction was collected within a 0.7 min time window into a vial containing 5 μL of 500 nM carbonic anhydrase, and the mixture was incubated to establish a new equilibrium. The sampling of 0.03 parts (150 nL/5 μL) of the collected ligands for the second step was due to limitations of currently available commercial CE instrumentation.

Quantitative PCR Analysis of 1′ and 5′. The decoding of DNA sequences in the collected fractions was conducted by quantitative PCR using a PCR thermocycler with real-time optical detection (Bio-Rad, Hercules, CA). In addition to 5 μL of a collected fraction, the PCR mixture contained 10 μL of 2× Real Time PCR Master Mix with SYBR Green (iQ SYBR Green Supermix, Bio-Rad), 4.8 μL of ddH_20, 0.6 μL of 10 μM forward primer, and 0.6 μL of 10 μM reverse primer. The reverse primer was 5′-TGT GAG TTT GGC GTG-3′ (for analysis of **1′**) or 5′-TGT GAG TCG GTT GTG-3′ (for analysis of **5′**), while the forward primer was 5′-TAG GCT ACG ACA GAC GTC AC-3′ (for both **1′** and **5′**). The total volume of the PCR reaction mixture was 20 μL. Each PCR cycle consisted of melting at 80 °C for 10 s, detection at 57 °C for 20 s, annealing at 35 °C for 20 s, and extension at 65 °C for 20 s. Real-time PCR produced S-shaped amplification curves (product yield vs the number of cycles) that were used to build calibration curves and determine amounts of **1′** and **5′**. Calibration curves were linear in the range of 10^2-10^8 molecules of **1′** or **5′**.

RESULTS AND DISCUSSION

Choice of Experimental Model. We take a freedom to repeat the three requirements a suitable library has to satisfy. *First*, the library must have high structural diversity to maximize the probability of containing ligands with desirable binding parameters. *Second*, the design of molecules in the library has to be compatible with a detection method to identify structures of ligands selected in a low number of copies. *Third*, the electrophoretic mobility of the library must be different from that of the ligand−target complexes, with targets typically being proteins.

Although highly diverse libraries of DNA-encoded small molecules of $N \geq 10^{12}$ (smart aptamers have been selected from random-sequence DNA libraries of such diversity)[5] have not been reported yet, the approach to their synthesis is well-known.[17−19] So requirement 1 is satisfied with the suggested library design. To prove that the library design satisfies requirements 2 and 3, we used an experimental model which included several DNA-tagged small molecules and protein targets, to which these molecules could bind. The DNA-tagged molecules used here can be considered as representatives of different libraries; the differences between the molecules are greater than those between molecules within a typical library. The general design of the molecules agrees with the one shown in Figure 1A. Briefly, the small molecules used were neutral BSB (1) and cholesterol (2), negatively charged biotin (3) and desthiobiotin (4), and positively charged benzylamine (5). Different DNA tags of 48 and 60 bases were attached to the molecules through different linker arms. DNA tags were hybridized to fluorescently labeled cDNA strands to facilitate fluorescence detection and prevent possible interaction of single-strand DNA tags with protein targets.

DNA-Tagged Small Molecules. We first tested if the DNA tags linked to small molecules could be efficiently and quantitatively amplified by PCR to satisfy requirement 2 (the structure of the small molecule has to be decoded from a low number of copies of its DNA tag). No difference was found between the amplification of the DNA tag linked to a small molecule and DNA of a similar structure without the small molecule. The sensitivity of quantitative PCR was 100 copies, approximately 10^7 times better than that of mass spectrometry. Thus, the suggested library design satisfies requirement 2.

We then examined whether the suggested library design satisfies requirement 3 (the electrophoretic mobilities of the DNA-tagged small molecules are different from those of protein complexes with the DNA-tagged molecules). We deliberately included small molecules and protein targets with great ranges of mobilities (Figure 2, parts A and C) to challenge the satisfaction of requirement 3. We found that the DNA-tagged small molecules migrated as a single electrophoretic zone with mobility similar to that of DNA (Figure 2B). The mobility of the DNA-tagged molecules was considerably lower than that of the most negatively charged protein, an isoform of β-lactoglobulin with pI 5.3. Furthermore, binding of the proteins to the DNA-tagged molecules induced significant shifts in their mobility (Figure 2D). These shifts create a considerable time window for selection of smart ligands and prove that the suggested library design satisfies requirement 3. Thus, the suggested library design (Figure 1A) satisfies all three requirements for selection of smart ligands by KCE methods.

It is understood that obtaining smart ligands with a narrow range of binding parameters may require a number of consecutive steps of partitioning.[4] Here, we are answering the fundamental question: how many steps are required for completing the selection? There are two important quantitative parameters associated with requirement 3: efficiencies of collection of ligands, $k_L = L_{out}/L_{in}$, and nonligands, $k_N = N_{out}/N_{in}$ (Figure 3). L_{in} and L_{out} are amounts of ligands at the input and output of partitioning, respectively; N_{in} and N_{out} are amounts of nonligands at the input and output of partitioning, respectively. In ideal partitioning, $k_L = 1$ and $k_N = 0$; practically they vary

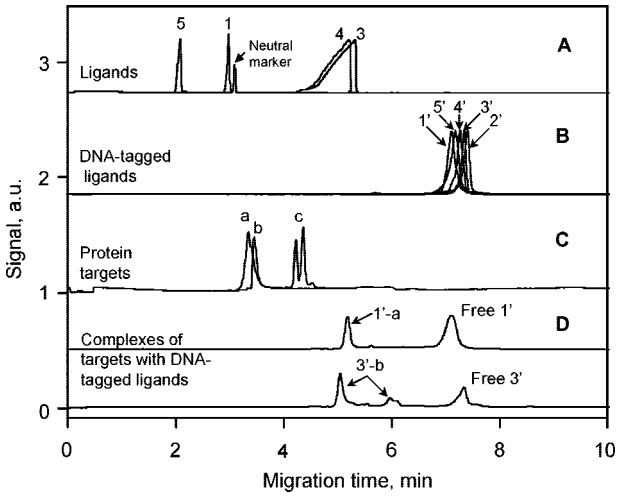

Figure 2. Electrophoresis migration patterns of small molecules (A); DNA-tagged small molecules (B); protein targets (C); complexes of proteins with DNA-tagged small molecules (D). The small molecules and proteins were *N*-benzyl-4-sulfamoyl-benzamide (1), cholesterol (2), biotin (3), desthiobiotin (4), benzylamine (5), carbonic anhydrase II (a), streptavidin (b), and two isoforms of β-lactoglobulin (c). Numbers with prime signs (**1′**−**5′**) correspond to the DNA-tagged small molecules. Primed numbers with letters (e.g., **1′**−a) correspond to the complexes. Electropherograms (D) were obtained with nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM) (ref 8).

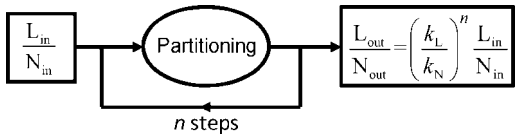

Figure 3. Schematic illustration of the ligand selection efficiency. L_{in}/N_{in} is the ratio between the amount of ligands and nonligands in the input library. L_{out}/N_{out} is the ratio between the amount of ligands and nonligands after the multistep partitioning.

between 1 and 0. Selection can be considered completed as soon as $L_{out}/N_{out} = 1$ is reached. With the use of this equality, the formula for L_{out}/N_{out} in Figure 3 can be converted to $n = \ln(N_{in}/L_{in})/\ln(k_L/k_N)$. If there is at least one molecule of a suitable smart ligand ($L_{in} \geq 1$) in a sample part of the library with a total number of molecules of $N_0 = L_{in} + N_{in}$, then $N_{in}/L_{in} = N_0/L_{in} - 1 < N_0$. The above formula for n can be rewritten as an upper estimate of a number of steps required for completing the selection:

$$n < \ln(N_0)/\ln(k_L/k_N) \quad (1)$$

The value of N_0 is typically known. The k_L/k_N ratio (the efficiency of partitioning) depends on the desired binding parameters of ligands and on the method of ligand selection. It can be determined experimentally, as we demonstrate below. Thus, n can be determined and practically used: if no ligands are selected after n consecutive steps of partitioning, the sampled library does not contain even a single copy of a smart ligand with the required binding parameters.

Selection of Smart Small-Molecule Ligands with KCE Methods. We experimentally proved the principle of selection of smart small-molecule ligands and determined n for our example

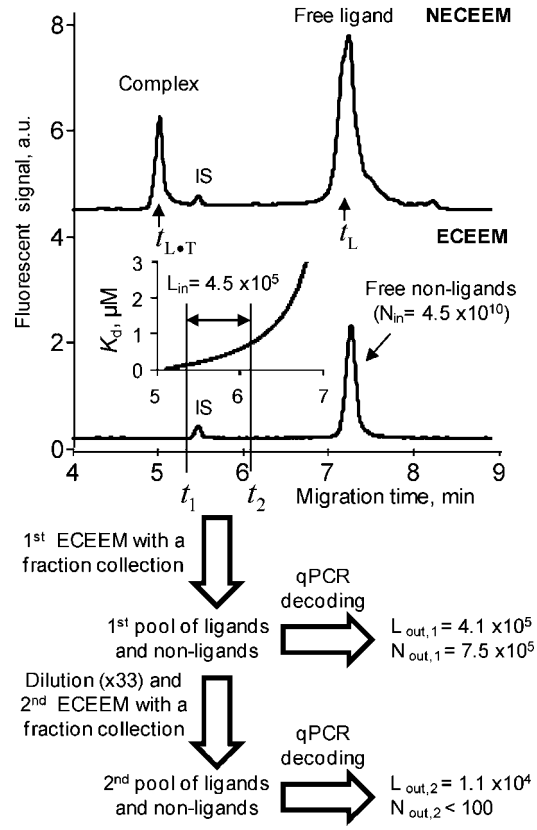

Figure 4. Selection of smart small-molecule ligands with predefined range of K_d values from a model library by a tandem of two KCE methods: NECEEM and ECEEM. Quantitative PCR was used for finding the number of DNA-tagged small molecules in the collected fractions. The top trace is a NECEEM electropherogram used to determine t_L and $t_{L\cdot T}$. The lower trace is an electropherogram of ECEEM used for partitioning of ligands from nonligands. The inset shows a theoretical dependence of K_d range of selected ligands on the time window of the collected fraction; the curve was calculated using eq 2. The bottom part of the figure shows the results of two consecutive steps of ECEEM partitioning. More experimental details can be found in the text. Peaks marked with "IS" correspond to the internal standard.

of selection. Partitioning was performed with a KCE method known as equilibrium capillary electrophoresis of equilibrium mixtures (ECEEM).[4,5] Nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM) was used to find the migration times required for selection of smart ligands by ECEEM. Briefly, both NECEEM and ECEEM start with the preparation of the equilibrium mixture of the target with the library. In NECEEM, a short plug of the equilibrium mixture is introduced into a capillary filled with a bare electrophoresis buffer and the components of the equilibrium mixture (unbound target, unbound ligand, and ligand−target complexes) are separated by capillary electrophoresis under nonequilibrium conditions. NECEEM allows for the accurate determination of migration times of the unbound ligand, t_L, and the ligand−target complex, $t_{L\cdot T}$. In ECEEM, a short plug of the equilibrium mixture is introduced into a capillary filled with electrophoresis buffer containing the target and the components of the equilibrium mixture are separated by capillary electrophoresis under quasi-equilibrium between unbound and target-bound ligands. The effective migration time, t, of ligands in ECEEM depends on the

equilibrium dissociation constant, K_d, of ligand−target interaction and the concentration of the unbound target [T]:

$$K_d(t) = [T]\frac{t_L}{t_{L\cdot T}}\frac{t - t_{L\cdot T}}{t_L - t} \quad (2)$$

ECEEM can be used to select smart ligands with a desirable range of K_d values. If a fraction is collected in time window t_1-t_2 (within a larger time window between $t_{L\cdot T}$ and t_L), ligands will be selected preferentially with K_d values in the following range:

$$[T]\frac{t_L}{t_{L\cdot T}}\frac{t_1 - t_{L\cdot T}}{t_L - t_1} < K_{d,t_1-t_2} < [T]\frac{t_L}{t_{L\cdot T}}\frac{t_2 - t_{L\cdot T}}{t_L - t_2} \quad (3)$$

The quality of ECEEM-based selection of smart ligands improves with the increasing rate of target−ligand complex re-equilibration. In general, the small molecules re-equilibrate fast which guarantees the high-quality selection.

To prove the principle of ECEEM-based selection of smart small-molecule ligands with a predefined range of K_d values, we used the following experimental model. Carbonic anhydrase II was used as a model target. The model library included a ligand (DNA-tagged BSB, K_d = 490 nM) spiked into a nonligand (DNA-tagged tryptamine, $K_d \gg 100\ \mu$M) with a ratio of $L_{in}/N_{in} = 1/10^5$. A 150 nL plug of the equilibrium mixture of the library ($L_{in} = 4.5 \times 10^5$, $N_{in} = 4.5 \times 10^{10}$, $N_0 \approx 4.5 \times 10^{10}$) with 500 nM target was injected into the capillary filled with the electrophoresis buffer containing 500 nM target and subjected to ECEEM. A fraction was collected within a time window between 5.4 and 6.1 min corresponding to a K_d range of 200−720 nM (Figure 4, lower trace). Fluorescein was added to the equilibrium mixture as an internal standard for accurate window determination. We employed quantitative PCR of the DNA tags to determine the number of copies of the ligand and nonligand in the collected fraction. We found that after step 1 of ECEEM partitioning, these numbers were $L_{out,1} = 4.1 \times 10^5$ and $N_{out,1} = 7.5 \times 10^5$. Accordingly, we could calculate $k_{L,1} = 0.9$ and $k_{N,1} = 1.7 \times 10^{-5}$. The efficiency of partitioning was $k_{L,1}/k_{N,1} = 5.3 \times 10^4$. After step 2 of ECEEM, the number of copies of the ligand and nonligand were $L_{out,2} = 1.1 \times 10^2$ and $N_{out,2} < 100$. The number of nonligands was below the limit of detection of PCR. Using $k_{L,1}/k_{N,1} = 5.3 \times 10^4$ and $N_0 = 4.5 \times 10^{10}$ in eq 1, we could calculate the upper limit for the number of partitioning steps required for completing the selection: $n < 3$. This number was confirmed experimentally by performing the second step of ECEEM partitioning, which produced more ligands than nonligands. This indicated that ligand selection could be considered completed. Our model experiment thus proved the principle of selection of smart small-molecule ligands not only qualitatively but also quantitatively.

CONCLUDING REMARKS

To conclude, we introduce a general approach for selection of smart ligands from highly diverse libraries of DNA-encoded small molecules and prove that the selection is feasible with a low number of consecutive steps of partitioning. The libraries with the desirable diversity are not available yet, but the technology for their synthesis has been already developed. To make such a selection practical, libraries with diversities of as high as 10^{12} may be needed. Two companies, Nuevolution and Praecis (acquired by GSK in 2007), synthesize highly diverse DNA-tagged libraries of small molecules (not available to outsiders). However, the maximum diversity of their libraries is 10^8 and 6.5×10^9 molecules, respectively.[22,23] The synthesis of more diverse libraries has been retarded in part by the lack of selection technologies applicable to such diverse libraries. Now, when a suitable selection technology is available, further efforts in developing highly diverse DNA-encoded libraries become fully justified. Besides, the current work will open new frontiers in selection of smart peptide ligands from libraries of mRNA-displayed peptides.[24]

We foresee that this work will stimulate selection of smart ligands and their new enabling applications in analyses and therapies that require well-defined dynamics of target−ligand interaction.

ACKNOWLEDGMENT

Funding was generously provided by the Natural Sciences and Engineering Council of Canada (a research Grant for S.N.K., an Alexander Graham Bell Canada graduate scholarship for M.U.M, and a postdoctoral fellowship for M.V.B) and the Government of Ontario (a graduate scholarship for A.P.D.). We thank Ensemble Discovery Corp. for providing two DNA-tagged small molecules.

Received for review November 10, 2008. Accepted November 13, 2008.

AC8023813

(22) Pedersen, H.; Holtmann, A.; Franch, T.; Gouliaev, A. H.; Felding, J. International Patent WO03078625, 2003.
(23) Morgan, B. Presented at the CHI 9th Annual Phage Display of Antibodies and Peptides, Boston, MA, May 14−18, 2007.
(24) Takahashi, T. T.; Austin, R. J.; Roberts, R. W. *Trends Biochem. Sci.* **2003**, *28*, 159–165.