Pergamon

Tetrahedron 55 (1999) 11687–11697

TETRAHEDRON

Selection of DNA-Binding Compounds *via* Multistage Molecular Evolution

Bryan Klekota and Benjamin L. Miller*
Department of Chemistry, University of Rochester
Rochester, New York 14627

Received 21 September 1998; accepted 20 October 1998

Abstract: Combinatorial libraries incorporating multiple equilibria offer opportunities to study molecular evolution, and are a novel method of identifying ligands for biological receptors. We describe the construction and evaluation of a multi-equilibrium combinatorial library, in which structural diversity and structural mutation are accomplished via reversible imine formation and transition-metal complexation. We demonstrate that oligo d(A•T)-cellulose resin can select subsets of this libray, in accord with measured solution-phase affinities. © 1999 Elsevier Science Ltd. All rights reserved.

The Darwinian evolution of biomolecules via a continuous process of mutation, selection, and amplification is one of the keystone concepts of biology. Laboratory applications of this process, including the polymerase chain reaction (PCR)[1] and the selection and amplification of RNA aptamers (for example, the SELEX method)[2] have proven to be invaluable tools for the generation of biopolymers with specific constitutions or activities. Recently, our group[3] and others[4,5,6] have been exploring methods for applying evolution, mutation, and selection principles to the generation of non-biopolymeric compounds with specific properties. In essence, these methods take advantage of the thermodynamics of binding events to influence the equilibrium concentration of a mixture. Such methods of generating molecular diversity via dynamic processes hold particular promise as new methods for the generation of small-molecule ligands for receptors, since the amount of a molecule in the mixture should be influenced by its affinity for the receptor. In addition to potentially providing a rapid means of identifying novel probes of macromolecular structure and function, this technique could increase the speed with which therapeutic lead compounds are discovered.

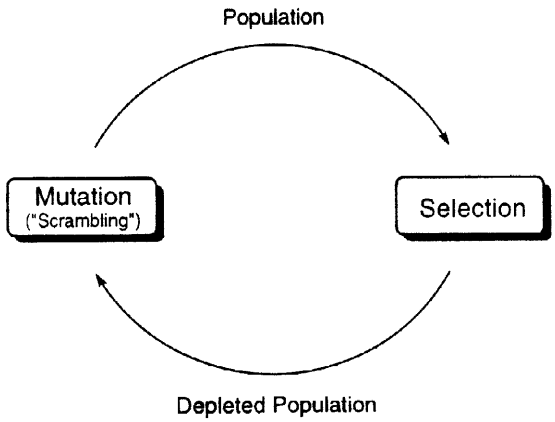
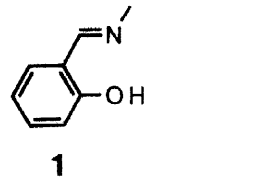


Figure 1: Combinatorial Molecular Evolution

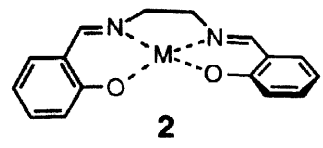
There are several requirements for the construction of an effective system for molecular evolution. First and foremost, the mechanism of library "scrambling" (i.e., bond formation, coordination/complexation, or isomerization) must be rapid and reversible. Second, "scrambling" of the library must be carried out in such a way that the receptor is unaffected. Third, either the products of the scrambling reaction or the components that combine with one another in the scrambling process must be separable, and identifiable by some spectroscopic or spectrometric technique. Finally, in an ideal molecular evolution system all hypothetical library constituents will have equivalent heats of formation, as well as similar solubilities. This ensures that selection is based solely on differences in binding affinity for the receptor, and not on the relative stability or solubility of a component of the mixture. In practice, this "ideal" is unapproachable, but should not prevent the experiment from being a success.

PII: S0040-4020(99)00687-0

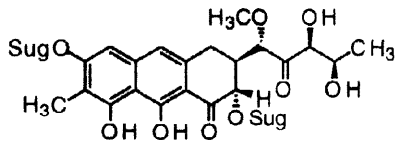
We have been exploring the utility of labile organotransition metal coordination complexes as scaffolding from which to append functionality potentially capable of binding to oligonucleotides. In particular, coordination complexes formed between salicylaldimines (**1**) and $Zn^{(2+)}$ seemed to fulfill all of the requirements we had stipulated; ligand substitution in complexes of this type is known to be rapid, and substantial structural characterization of such complexes (in the solid state) is available.[7] Furthermore, the closely related metal-salen complexes **2** have been used extensively as DNA- and RNA-binding reagents.[8] Although the goal of this study was not necessarily to generate pharmaceutically viable compounds, it should be noted that several biologically active compounds, including chromomycin A3[9] (**3**) and a series of recently discovered serine protease inhibitors (i.e., **4**),[10] gain their activity via coordination to a metal ion *in situ*.



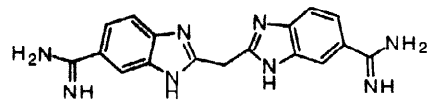
1



2

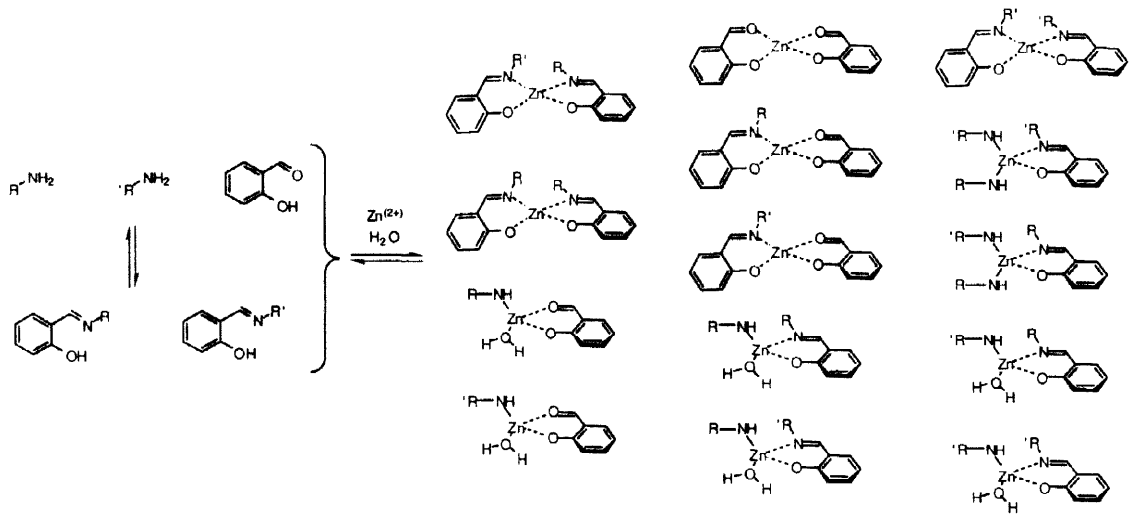


3



4

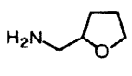
Our initial feasibility studies tested the ability of cellulose-supported DNA to select and amplify compounds from an equilibrating combinatorial library of what we presumed to be bis(salicylaldiminato)zinc complexes formed by mixing pure salicylaldimines with zinc salts in aqueous solution. Several observations made during the course of that investigation merited further study. First, as was to be expected based on the known hydrolytic instability of imines in water, we observed significant amounts of hydrolysis for pre-formed imines on standing in aqueous solution. Conversely, mixing an amine with salicylaldehyde in water was found to provide a significant amount of imine. Indeed, the kinetics of imine formation between various amines and salicylaldehyde in the presence of a series of divalent metal salts, as studied extensively by Leussing and



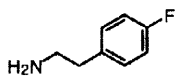
Scheme 1: A subset of the complexes that might be formed in a multistage equilibration

coworkers some time ago,[11] support the contention that transition metals can template the reversible formation of salicylaldimines, and mono- or bis-salicylaldiminato(metal) complexes. Furthermore, during the course of our initial study, the Lehn group reported observing differences in product ratios for libraries of imines formed in the presence or absence of a receptor.[12] This suggested to us that a multi-stage equilibration (imine formation, complexation, and receptor binding) might in fact be operational in our selection experiments, rather than the two-stage equilibration (complexation, receptor binding) we had envisioned. In such a multi-stage equilibrium process, the presence of amines, salicylaldehyde, imines, and buffer components provides an extremely complex mixture, since all of these materials are in equilibrium with one another as well as with various complexed forms. For example, considering only four-coordinate zinc in complex with an equilibrating mixture of two amines, salicylaldehyde, and water produces a mixture of at least 1,521 components (Scheme 1), all of which can potentially participate in DNA binding.[13]

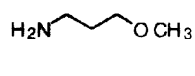
In addition to explicitly examining receptor-mediated selection of compounds from a library assembled under multistage equilibrium conditions, we wished to address the potential participation of the buffer in the selection experiments. The selection results obtained in our previous study, conducted in Tris buffer, reflected the relative solution phase affinities of various complexes for DNA (measured in Tris), but whether the buffer itself had played a role in the selection was unclear. Tris is itself capable of nonspecific binding to DNA, and is known to be a strongly metal-coordinating buffer.[14]. Furthermore, since Tris is itself a primary amine, it would form imines with salicylaldehyde and complex zinc, potentially obscuring our ability to observe differential DNA binding among other zinc-salicylaldimine complexes. Therefore, we sought alternatives.



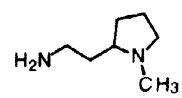
5



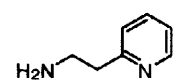
6



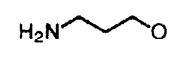
7



8



9



10

RESULTS

Since data regarding the rates of imine formation, complexation, and ligand exchange in buffered solutions were lacking, we chose to first examine the time-dependence of imine formation and equilibration in the presence and absence of zinc. In phosphate-buffered saline solution (PBS) containing 100 mM KCl (pH 7.3), 157 μM salicylaldehyde, and 157 μM each of **5**-**10**, the absorbance band corresponding to the imine at 380 nm essentially reaches a maximum after two hours in the absence of zinc. In the presence of 40 μM $Zn^{(2+)}$, the rate of imine formation is significantly faster (Figure 1), in accordance with Leussing's observations. However, after reaching a maximum value, the absorbance falls off over time. At the end of the run a significant amount of precipitate could be observed in the spectrophotometric cuvette. Most likely, this precipitate is composed of insoluble zinc-phosphate complexes, by analogy to known reactions of Zn(II) salts with phosphate solutions.[15] The same experiment, when run in the much less coordinating buffer N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid) (HEPES; 10 mM pH 6.8, 100 mM KCl), showed approximately the same rate of imine formation, but the rate of decrease in absorbance was markedly slower. Therefore, this suggested that HEPES might be a more effective buffer for selection and amplification experiments, since precipitation or other mechanisms of imine loss would be less problematic. That the rate of absorbance decrease in HEPES is nonzero suggests that at least some of the complexes formed *in situ* are of low solubility, although we observed no precipitate.

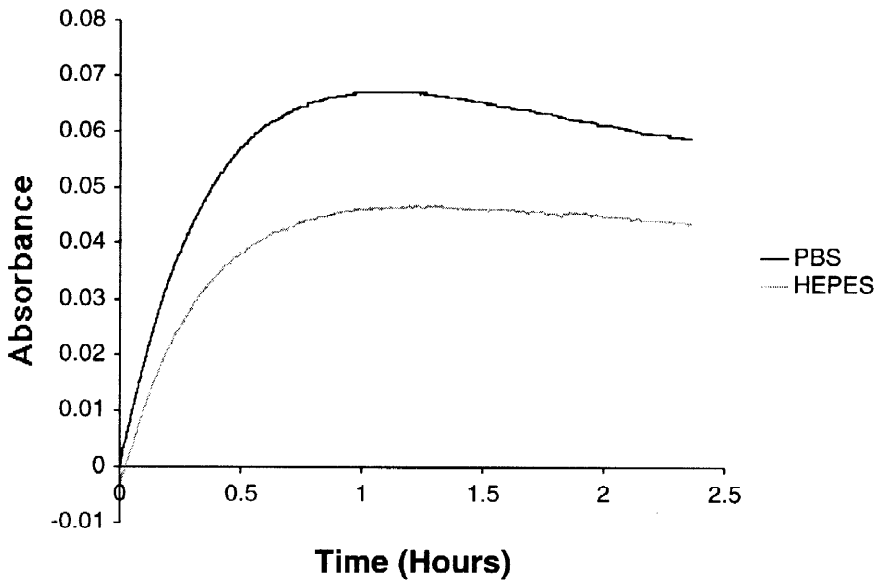


Figure 1: Imine formation in the presence of $Zn^{(2+)}$ in PBS or HEPES buffers

A second interesting observation was that the addition of poly(dA•dT) to an equilibrated mixture of salicylaldehyde, amines, and zinc causes a significant increase in the rate of decay of the absorbance at 380 nm (Figure 2). This could be due to one of at least two factors. First, free amines and zinc might bind nonspecifically to DNA, reducing the concentration of amines in solution. This would shift the salicylaldehyde-salicylaldimine equilibrium in favor of salicylaldehyde, reducing the 380 nm absorbance. Second, binding of imines and/or complexes to DNA could lead to quenching of the absorbance at 380 nm via an electronic interaction. Experiments designed to test the effect of increasing DNA concentration on the salicylaldehyde - imine - complex equilibrium are in progress. Again, this rate of decay is much faster in PBS/KCl than in HEPES/KCl.[16]

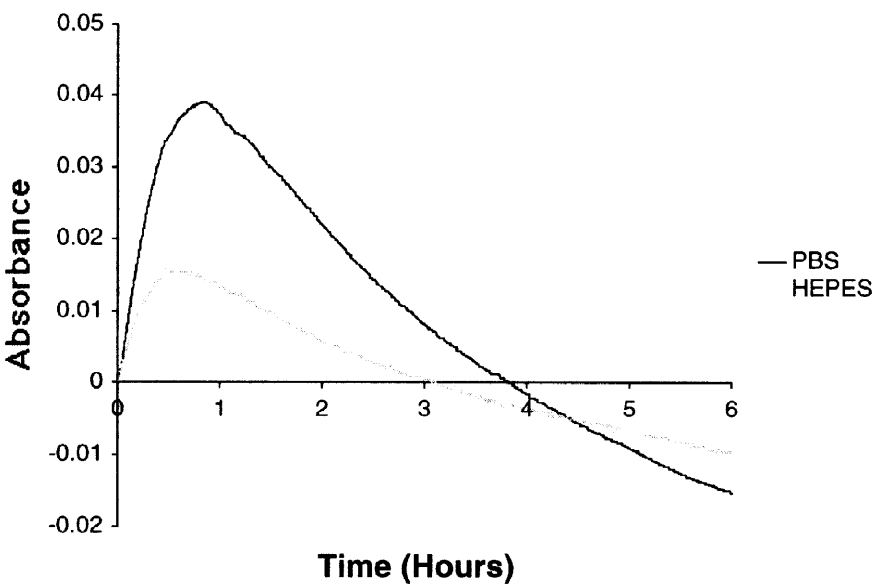


Figure 2: Imine formation in the presence of $Zn^{(2+)}$ and poly(dA•dT) in PBS or HEPES buffers

SELECTION EXPERIMENTS

As a first attempt at carrying out a multistage selection experiment in HEPES buffer, we incubated libraries derived from six amines (**5-10**), salicylaldehyde, and $Zn^{(2+)}$ with double-stranded oligo(dA•dT)-cellulose resin. Initial experiments employed concentrations of 82 μM in each amine, zero or 82 μM salicylaldehyde, and zero, 41, or 82 μM $Zn^{(2+)}$. This provided a total initial ratio of 6:1 amine:salicylaldehyde, which was anticipated to be a sufficient excess of amine over aldehyde to promote imine formation. The UV kinetics experiments described above suggested that the half-life for equilibration of the library in the presence of $Zn^{(2+)}$ was on the order of 30 minutes; therefore, libraries were allowed to pre-incubate for one hour to allow them to reach an initial equilibrium prior to their addition to the DNA resin.[17] Samples were incubated on the resin for three hours prior to being eluted, hydrolyzed, and derivatized with 2-naphthoyl chloride. Derivatized samples were analyzed by HPLC, the identity of specific derivatives determined by comparison with the retention times observed for standard solutions, and peak areas recorded.

Unfortunately, we found that the above set of reagents and conditions produced data that were not acceptably reproducible; standard deviations among identical sets of experiments were too high to permit any reasonable analysis. Thinking that perhaps increasing the concentrations of materials involved in the experiment might improve reproducibility, we carried out another series of experiments in which the amount of each amine was doubled (providing a total amine:salicylaldehyde ratio of 12:1). In this case, the reproducibility of measurements was significantly improved, allowing us to draw conclusions about the results of the experiment.

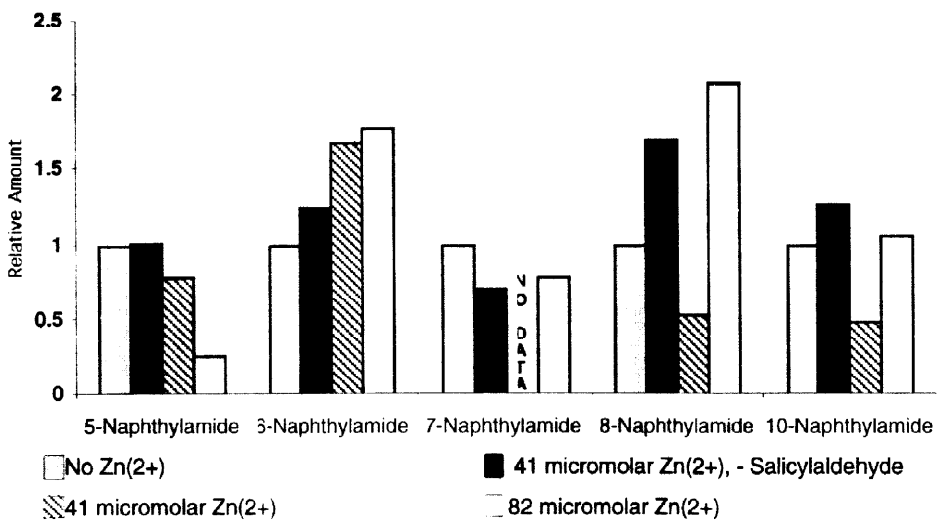


Figure 3

Average results for derivatized amines from initial elutions are shown in Figure 3. In each case, areas for eluted peaks in the set of zinc-free samples (164 μM each amine, 82 μM salicylaldehyde, no $Zn^{(2+)}$) were used to normalize all other values. In accord with the results of our previous experiments, a significant reduction in the amount of derivatized N-methyl 2-aminoethyl pyrrolidine (**8**) was observed from the three replicate experiments containing both salicylaldehyde and 41 μM $Zn^{(2+)}$. Similarly, a large (≥ 50%) reduction in the amount of 3-hydroxypropylamine (**10**) was observed under these conditions. An unidentified impurity unfortunately obscured the peak corresponding to **9**-naphthylamide, preventing us from evaluating the role of **9** in library selection. Comparison with the initial elutions from unfunctionalized cellulose resin (Figure 4) indicates similar levels of retention of **10**, but not **8**. A slight (*ca.* 20%) reduction in the amount of 2-

aminomethylfuran (**5**) in salicylaldehyde and 41 μM $Zn^{(2+)}$-containing libraries eluted from DNA-cellulose resin was observed; this may indicate a slight affinity of complexes incorporating **5** for DNA. Similarly to our previously reported results for libraries incorporating pre-formed salicylaldimines, these data would lead us to the conclusion that zinc complexes incorporating salicylaldehyde and N-methyl 2-aminoethylpyrrolidine (**8**) have the highest affinity for oligo (dA•dT) at the relative concentrations tested.

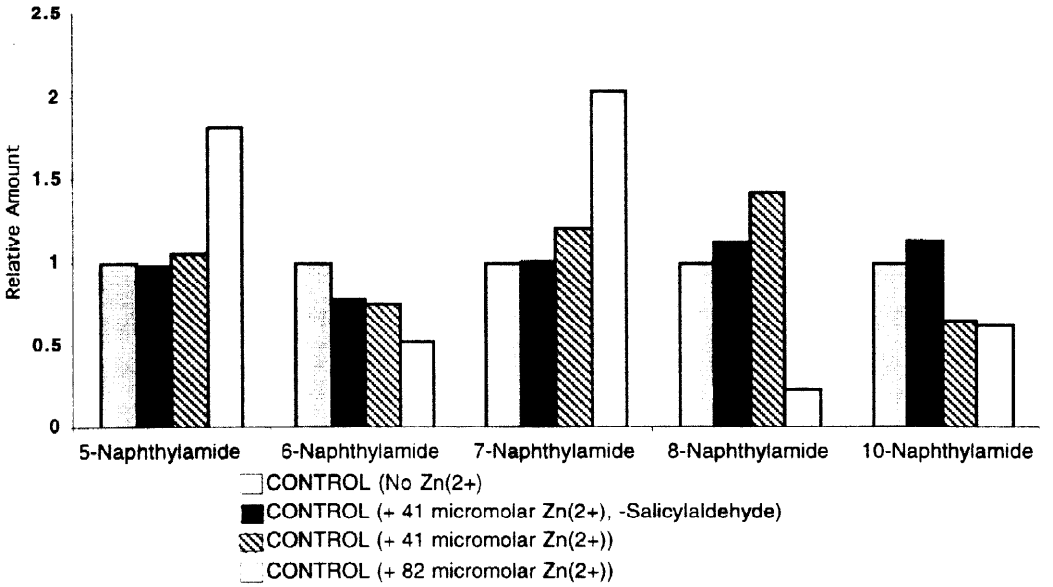


Figure 4

Curiously, libraries incorporating higher (82 μM) concentrations of $Zn^{(2+)}$ gave markedly different results from those incorporating 41 μM $Zn^{(2+)}$. Most notably, we observed a significant reduction in the amount of **8** retained on the column. The most likely explanation for this is that the added zinc changes the population of complexes present in the library. However, it is not known at present what the precise structural reasons for this might be. We have observed a strong dependence of the UV spectra of mixtures of salicylaldehyde and N-methyl 2-aminoethyl pyrrolidine (**8**) in HEPES buffer on the concentration of $Zn^{(2+)}$ (Figure 5). However, the concentrations of $Zn^{(2+)}$ evaluated in these UV titrations which caused the inverse parabolic variations observed were 10-fold higher than those employed in our affinity experiments.

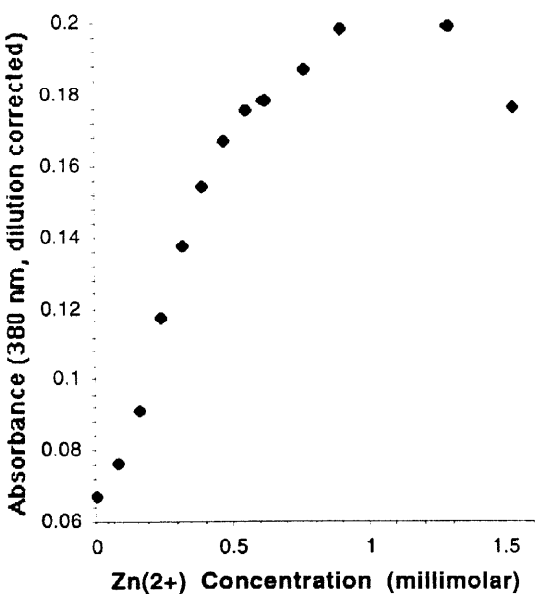


Figure 5

To verify that the presence of zinc was not producing artifacts in our ability to derivatize samples, we prepared control libraries 164 μM in each amine, 82 μM in salicylaldehyde, and 0 or 41 μM $Zn^{(2+)}$. After allowing these libraries to incubate in solution (i.e., without contact with any sort of resin) for 6 hours, they were lyophilized, derivatized, and analyzed in an identical manner to samples from selection experiments. As shown in Figure 6, differences in peak volumes for libraries prepared in the presence or absence of $Zn^{(2+)}$ are less than 10%, with the exception of 3-methoxypropylamine (**7**). In each case, the differences

could be due to experimental error, participation of $Zn^{(2+)}$ in the acylation reaction, or by $Zn^{(2+)}$-accelerated partial reversion of hydrolyzed complexes to imines prior to acylation. These results also suggest that differences in peak volumes for derivatized amines that are greater than 10% are probably significant, with the exception of 3-methoxy propylamine.

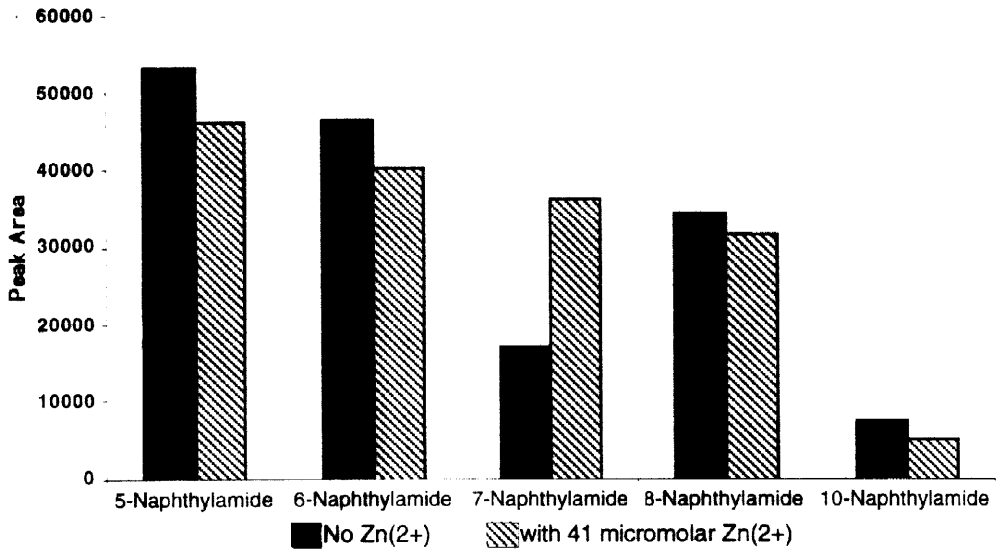


Figure 6

A potentially more desirable method of identifying the library constituents with the highest affinity for DNA would be to remove them from the receptor in some manner following the initial elution of "nonbinding" components for subsequent derivatization and analysis. We assumed that simply washing the DNA resin with water would be sufficient to denature the receptor, since low-salt conditions are known to cause DNA melting. However, we observed only small amounts of derivatized amines from wash solutions, with no experimentally significant differences among experiments or between DNA-cellulose (Figure 7) and control cellulose (Figure 8) resins. One potential explanation for this is that the affinity of N-methyl 2-aminoethyl pyrrolidine-containing complexes is similar for either single or double-stranded DNA; experiments to test this hypothesis are in progress.

CONCLUSIONS

We have demonstrated that the receptor-mediated selection of subsets of libraries prepared by a multistage equilibrium process is possible. The observation that complexes incorporating **8** were selected by this method is consistent with measured solution-phase affinities for DNA.[3,18] Furthermore, as would be expected for a process under thermodynamic control, the results of the selection experiments reported herein are similar to the results we reported earlier which used pre-formed salicylaldimines. That they are not *identical* to our previous results (i.e., there is some variance with previously reported amounts of observed elution products) is most likely the result of other experimental changes, including the overall concentration of material as well as the buffer in which the experiments were conducted. Efforts to fully characterize the *structure* (as opposed to the constitution) of the high-affinity complexes identified using this procedure are in progress. In order to improve the speed and accuracy with which future multistage selection experiments may be conducted, it will be necessary to design libraries in which library scrambling is performed by some external agent which may be removed from the mixture. Initial exploratory experiments in this direction, as well as

experiments which incorporate starting materials with significantly greater structural complexity, are currently underway in our laboratories.

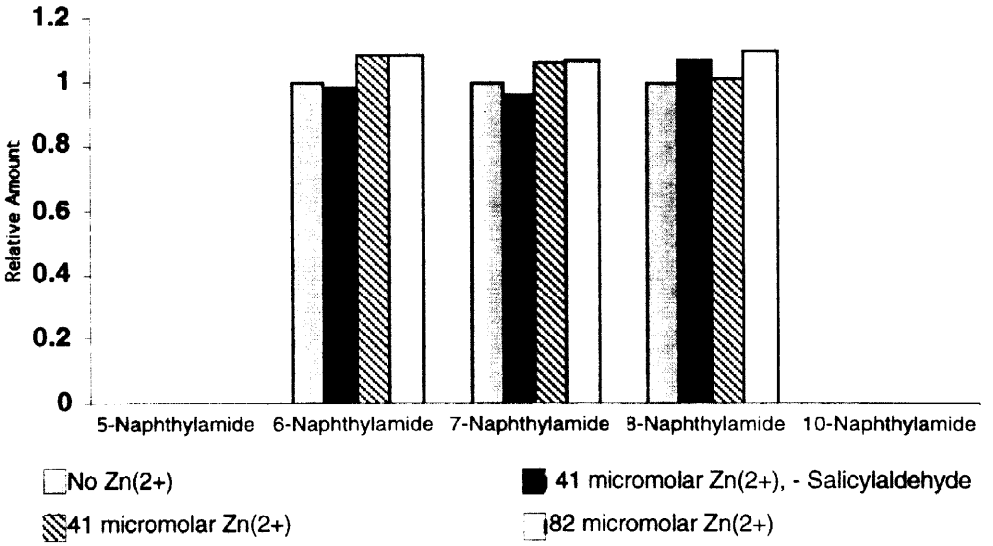


Figure 7

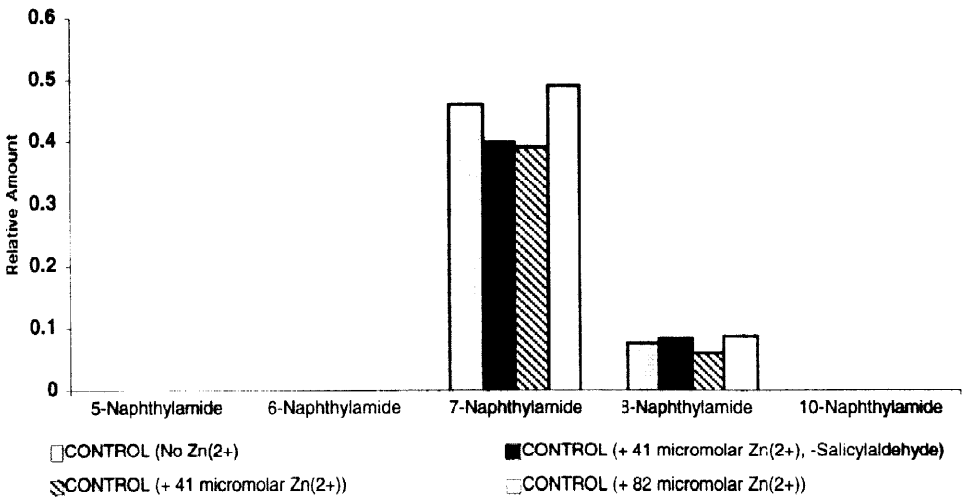


Figure 8

EXPERIMENTAL SECTION

General Materials and Methods: Salicylaldehyde, 3-methoxypropylamine, 3-hydroxypropylamine, 2-aminoethylpyridine, N-methyl 2-aminoethyl pyrrolidine, 2-aminomethylfuran, (4-fluoro)phenethylamine, trifluoroacetic acid, 2-naphthoyl chloride, and methylene chloride were obtained from Aldrich chemical company and used without further purification. Triethylamine was obtained from J. T. Baker corp. and distilled over calcium hydride under nitrogen atmosphere prior to use. HEPES and PBS buffers were prepared according to standard protocols from materials purchased from Sigma chemical company. Deionized water

was subjected to redistillation in an all-glass apparatus prior to use. Zinc chloride and zinc acetate wre used as obtained from J. T. Baker Corp. Oligo(dT)-cellulose resin, oligo(dA), and poly(dA•dT) were obtained from Pharmacia Biochem. Unfunctionalized cellulose for control experiments was used as obtained from Sigma chemical company. HPLC data on derivatized amines were obtained using either a Beckman 112 ternary gradient HPLC system equipped with an Alcott fixed-fill autosampler, a Whatman EQ-C18 reverse-phase column (4.6 x 300 mm), and UV detection, or a Hewlett-Packard 1050 quaternary HPLC equipped with an autosampler, Hamilton RP-1 reverse-phase column (4.6 x 150 mm), and UV detection. LC-MS was carried out using a Hewlett-Packard Series 1100 MSD, using electrospray ionization in positive ion mode and a Whatman EQ-C18 reverse-phase column. UV-Vis spectra and kinetics experiments were carried out on a Shimadzu 1601-PC spectrophotometer at ambient temperature (23 °C +/- 1 °C).

Selection Experiments:

(1) Preparation of DNA affinity resins

For each sample, 15 mg oligo(dT) cellulose was weighed into a 2 ml Bio-Rad BioSpin® polypropylene chromatography column equipped with a glass frit. The resin was then taken up in 750 ul HEPES buffer (10 mM HEPES, 100 mM KCl, pH 6.8). 0.5 $A_{(260)}$ units of oligo(dA) were then added in 50 µl H_2O, and allowed to incubate on the resin for one hour at ambient temperature prior to the addition of library samples. During this preincubation period, as well as throughout the remainder of the selection experiments, gentle agitation was provided by rotating the samples on a Barnstead-Thermolyne LabQuake (tm) rotary shaker. In a similar manner, 15 mg of unfunctionalized cellulose for each control experiment was allowed to soak in 800 µl HEPES buffer at ambient temperature for one hour prior to the addition of library samples.

(2) Preparation of libraries

Samples were prepared in triplicate from stock solutions as indicated in Tables 1 and 2 to provide a total volume of 450 µl. An identical set of samples was prepared for control experiments. In each case, libraries were allowed to incubate at ambient temperature for one hour prior to addition to the DNA-cellulose or control cellulose resin. Additional sets of library samples were prepared according to lines C and D of Table 2 for control experiments conducted in the absence of DNA-cellulose or unfunctionalized cellulose resin.

Table 1: Library samples, "low" amine concentration:

Sample Set	Concentration of each amine	$Zn^{(2+)}$ Concentration	Salicylaldehyde Concentration
A	82	0	82
B	82	41	0
C	82	41	82
D	82	82	82

All samples were prepared in triplicate, as described in the text. Concentrations are in micromolar, and reflect the final concentration obtained following addition of the library mixture to the DNA-cellulose or control resin.

Table 2: Library samples, "high" amine concentration

Sample Set	Concentration of each amine	$Zn^{(2+)}$ Concentration	Salicylaldehyde Concentration
A	164	0	82
B	164	41	0
C	164	41	82
D	164	82	82

All samples were prepared in triplicate, as described in the text. Concentrations are in micromolar, and reflect the final concentration obtained following addition of the library mixture to the DNA-cellulose or control resin.

(3) Incubation of libraries with affinity resins

Following the pre-incubation period, each sample was added to either a DNA-cellulose or unfunctionalized cellulose column as appropriate. These samples were then allowed to incubate on the resin with gentle agitation at ambient temperature for either three (Table 1) or six (Table 2) hours. Solutions were then eluted from the resin, frozen, and lyophilized prior to derivatization and analyisis as described below. Each resin sample was resuspended in 1 ml H_2O and agitated for 24 hours; these wash solutions were then eluted from the resin, frozen, and lyophilized prior to derivatization and analyisis as described below.

(4) Derivatization of samples

Each sample was resuspended in 1.0 ml of a 50% trifluoroacetic acid solution in dichloromethane, vortexed, and allowed to react for one hour. Solvent was removed under reduced pressure. To each sample of dried material was then added a solution of 0.017 mmol 2-naphthoyl chloride and 200 µl triethyl amine in 0.8 ml methylene chloride. After allowing the reaction to proceed for 10 hours, solvent was removed under reduced pressure. Each crude sample was prepared for HPLC analysis by dissolution in 700 µl of a 20% H_2O - 80% CH_3CN solution.

(5) HPLC analysis of samples

In order to separate and identify as many constituents of each sample as possible, a minimum of two HPLC elutions were performed for each sample. In the first, a 30 minute linear gradient of 30% to 100% CH_3CN in H_2O-0.1% trifluoroacetic acid was applied. In the second analysis, a linear gradient of 30% to 100% CH_3CN in 10 mM NH_4OAc (pH 4.5) was applied. Elutions were monitored at 230 nm. HPLC peaks were assigned by comparison with retention times obtained for standards using both solvent systems. Further confirmation of the identity of HPLC peaks was obtained by analyzing selected samples by LC-MS, using the identical HPLC conditions.

Acknowledgement: We thank the University of Rochester for generous financial support.

REFERENCES

1. Mullis, K. B.; Faloona, F. A. *Methods Enzymol.* **1987**, *155*, 335-350.
2. (a) Szostak, J. W. *Trends Biochem. Sci.* **1992**, *17*, 89-93; (b) Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505-510.
3. Klekota, B.; Hammond, M. H.; Miller, B. L. *Tetrahedron Lett.* **1997**, *38*, 8639-8642.

4. (a) Rowan, S. J.; Hamilton, D. G.; Brady, P. A.; Sanders, J. K. M. *J. Am. Chem. Soc.* **1997,** *119*, 2578-2579; (b) Browan, S. J.; Sanders, J. K. M. *J. Org. Chem.* **1998,** *63*, 1536-1546.
5. Hioki, H., Still, W.C. *J. Org. Chem.* **1998**, *63*, 904-905.
6. Eliseev, A. V.; Nelen, M. I. *J. Am. Chem. Soc.* **1997,** *119*, 1147-1148.
7. Garnovskii, A. D.; Nivorozhkin, A. L.; Minkin, V. I. *Coord. Chem. Rev.* **1993**, *126*, 1-69.
8. (a) Gravert, D. J.; Griffin, J. H. *J. Org. Chem.* **1993**, *58*, 820-822. (b) Muller, J. G.; Paikoff, S. J.; Rokita, S. E.; Burrows, C. J. *J. Inorg. Biochem.* **1994**, *54*, 199-206.
9. (a) Gao, X.; Patel, D. J. *Biochemistry* **1989**, *28*, 751-762. (b) Gao, X.; Mirau, P.; Patel, D. J. *J. Mol. Biol.* **1992**, *223*, 259-279. (c) Silva, D. J.; Kahne, D. *J. Am. Chem. Soc.* **1994**, *116*, 2641-2642.
10. Katz, B. A.; Clark, J. M.; Finer-Moore, J. S.; Jenkins, T. E.; Johnson, C. R.; Ross, M. J.; Luong, C.; Moore, W. R.; Stroud, R. M. *Nature* **1998**, *391*, 608-612.
11. (a) Leach, B. E.; Leussing, D. L. *J. Am. Chem. Soc.* **1971**, *93*, 3371-3384; (b) McQuate, R. S.; Leussing, D. L. *J. Am. Chem. Soc.* **1975**, *97*, 5117-5124; (c) Hopgood, D.; Leussing, D. L. *J. Am. Chem. Soc.* **1969**, *91*, 3740-3750.
12. Huc, I., Lehn, J.M. *Proc. Natl. Acad. Sci* **1997**, *94*, 2106-2110.
13. This value was derived as follows: considering only tetrahedral coordination, and ignoring symmetry, we have 6 major components (2 amines, salicylaldehyde, 2 imines, and water) to permute. Assuming salicylaldehyde and the two salicylaldimines are the only materials capable of divalent coordination, we have:

    (1) 2 species coordinated: 6^4 = 1296
    (2) 3 species coordinated: 6^3 = 216
    (3) 2 species coordinated: 3^2 = 9

    which provides a total of 1521 distinct complexes.
14. Good, N. E.; Winget, G. D.; Winter, W.; Connolly, T. N.; Izawa, S.; Singh, R. M. M. *Biochemistry* **1966**, *5*, 467-477.
15. Bailar, J. C., Jr.; Emeléus, H. J.; Nyholm, R.; Trotman-Dickernson, A. F., Eds. *Comprehensive Inorganic Chemistry* **1973**, Pergamon Press, Oxford, 1973, vol. 3 p. 238.
16. It is unclear whether the differences in the maximum amount of absorbance observed are significant, or simply a result of experimental error.
17. Of course, since equilibration continues once the library is added to the resin, it is probable that this pre-equilibration step could be eliminated without altering the result.
18. In addition to the measured affinities described in reference (3), we have measured the affinity of the mixture composed of **6**, salicylaldehyde, and Zn(2+) for oligo d(A•T) in HEPES/KCl, and found it to be 11 micromolar.