Use of Molecular Recognition To Drive Chemical Evolution. 1. Controlling the Composition of an Equilibrating Mixture of Simple Arginine Receptors

Alexey V. Eliseev* and Marina I. Nelen

Department of Medicinal Chemistry, School of Pharmacy
State University of New York at Buffalo
Buffalo, New York 14260

Received June 5, 1996
Revised Manuscript Received December 13, 1996

Molecular recognition plays a major role in biological systems. During the last decade, considerable effort directed toward modeling of biological non-covalent binding in chemical systems resulted not only in the synthesis of numerous artificial molecular receptors but also in the development of innovative approaches to the generation of selective non-covalent binders. Besides conventional organic synthesis, combinatorial chemistry techniques have been applied to the creation of potential molecular receptors.[1,2] A merging of combinatorial chemistry and DNA cloning techniques has recently been used for the evolutionary selection/amplification of nucleic acids capable of binding to biopolymers and low-molecular targets.[3]

We introduce here a new and general approach that involves an automatic enrichment of a mixture of compounds (receptors) in the component possessing the highest affinity to the particular target ligand. This method is applicable to *any* systems in which *the components exist in a dynamic equilibrium*. To the best of our knowledge, this approach is the first application of the Darwinian principles of mutation and survival (selection) to a simple chemical system.[4] The approach is illustrated by the evolutionary-type formation of an anionic receptor for arginine.

For the generation of a simple pool of equilibrating receptors, we have designed and synthesized[5] the unsaturated dicarboxylate **1**, which is capable of existing in three isomeric forms: *trans,trans*, *cis,trans*, and *cis,cis* (Scheme 1). The three isomers can be interconverted by irradiation with UV light. According to molecular modeling, the *cis,cis* isomer of **1** possesses an ideal geometry for complexation with the guanidinium moiety of arginine *via* formation of two salt bridges (Scheme 1). The *cis,trans* and, in particular, *trans,trans* isomers were expected to show much lower affinity because of the longer distance between the carboxylates.[6]

To verify our predictions of the relative affinities, the binding constants of the disodium salts of all isomers of **1** to methylguanidinium hydrochloride were determined by the NMR titrations with the isomeric mixture in methanol and ethanol.[7] As expected from the electrostatic nature of binding, the association in ethanol is characterized both by higher absolute values of the binding constants ($K_{cis,cis}$ = 980 *vs* 170 M^{-1} in methanol) and by larger discrimination factors between the isomers ($K_{cis,cis}/K_{cis,trans}$ = 6.5, $K_{cis,cis}/K_{trans,trans}$ >100 *vs* $K_{cis,cis}/K_{cis,trans}$ = 2.4, $K_{cis,cis}/K_{trans,trans} \geq 17$ in methanol).[8]

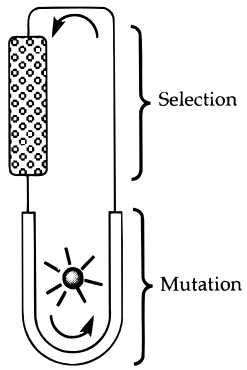
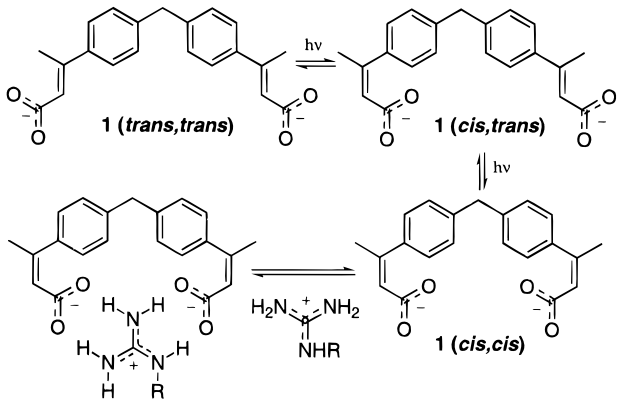


Figure 1. Schematic representation of the experimental setup.

Scheme 1



A solution of **1**·2Na in ethanol (2.6×10^{-4}M, 7 mL) was then subjected to circulation in the apparatus shown in Figure 1. Irradiation of the solution with a broad-band UV light (mercury lamps in a Rayonet reactor) in a photochemical flow cell ("mutation" chamber, Figure 1) led to the formation of a distribution of isomers[9] (*cis,cis*/*cis,trans*/*trans,trans* = 3/28/69) as shown in Figure 2a.[10] The resulting mixture was then pumped to the affinity column ("selection" chamber, Figure 1) containing 2.8 mL of 1×10^{-2} M arginine immobilized on the

(1) (a) Still, W. C. *Acc. Chem. Res.* **1996**, *29*, 155−163. (b) Borchardt, A.; Still, W. C. *J. Am. Chem. Soc.* **1994**, *116*, 373−374. (c) Boyce, R.; Li, G.; Nestler, H. P.; Suenaga, T.; Still, C. W. *J. Am. Chem. Soc.* **1994**, *116*, 7955−7956.

(2) Goodman, M. S.; Jubian, V.; Linton, B.; Hamilton, A. D. *J. Am. Chem. Soc.* **1995**, *117*, 11610−11611.

(3) (a) Famulok, M.; Szostak, J. W. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 979−988. Famulok, M. *J. Am. Chem. Soc.* **1994**, *116*, 1698−1706 and references therein.

(4) As opposed to "self-replicating systems" (for reviews, see: (a) Orgel, L. E. *Nature* **1992**, *358*, 203−209. (b) Orgel, L. E. *Acc. Chem. Res.* **1995**, *28*, 109−118), our method leads to a thermodynamic rather than kinetic preference in the *selection* of particular structure. The principle of *mutation* (variation) has not been explicitly used either in self-replicating systems or in the DNA evolution techniques (ref 3).

(5) The synthesis of **1** involved the acetylation of diphenylmethane (AcCl, $AlCl_3$ in CS_2) to give bis(4-acetylphenyl)methane in 75% yield. The latter was reacted with 2.2 equiv of methyl(diethylphosphono)acetate in a Wadsworth−Emmons reaction (NaH, DMF, 10−25 °C, 72 h) to afford the dimethyl ester of **1** (45% after column chromatography), which was then hydrolyzed by NaOH in MeOH/water and isolated as the *trans,trans* diacid of **1** (85%). ^{1}H NMR (300 MHz, DMSO-d_6): 12.1 s, 2H; 7.48 d, 4H, 7.27 d, 4H, 6.09 s, 2H; 3.98 s, 2H; 2.46 s, 2H.

(6) In the minimized structures of **1** docked with the guanidinium ion (Tripos software package SYBYL), the distances between the carboxylate oxygens and guanidinium nitrogens were in the ranges 2.5−2.6, 3.3−5.6, and 4.6−7.0 Å in the *cis,cis*, *cis,trans*, and *trans,trans* isomers, respectively.

(7) A 5×10^{-4} M solution of **1**·2Na in CD_3OD or CD_3CD_2OD was irradiated by a mercury lamp in the NMR tube for 30 min. The 400 MHz ^{1}H NMR spectrum of the resulting mixture contained sufficiently resolved signals of all isomers of 1. The solution was then titrated with methylguanidinium·HCl (5×10^{-4} to 1.3×10^{-1} M), yielding the curves of δ vs [titrant] (see Supporting Information). The individual association constants for 1:1 complexes were then calculated from nonlinear least-square curve fitting.

(8) The relatively low affinity difference between the *cis,cis* and *cis,trans* isomers is apparently due to an appreciable contribution of the second COO^- group to binding in the latter case. This contribution drops substantially in the *trans,trans* isomer, as expected from a nonlinear distance dependence of ionic interactions.

(9) The HPLC analysis was performed on a Beckman Gold chromatograph (5 cm ODS column; 70% methanol in 0.15% aqueous H_3PO_4 at 0.5 mL/min), determining each isomer of **1** at its maximum absorption wavelength (λ_{max} 262 nm (ϵ = 33 400), 275 nm (ϵ = 30 000), and 282 nm (ϵ = 26 700) for *cis,cis*, *cis,trans*, and *trans,trans* isomers, respectively). Assignment/integration of the peaks was confirmed by comparison with the ^{1}H NMR of the mixtures.

(10) Since the solution of **1** was irradiated during the limited time of passing the flow cell, the equilibrium photostationary distribution of isomers was not reached in the isomerization cycle.

S0002-7863(96)01893-8 CCC: $14.00

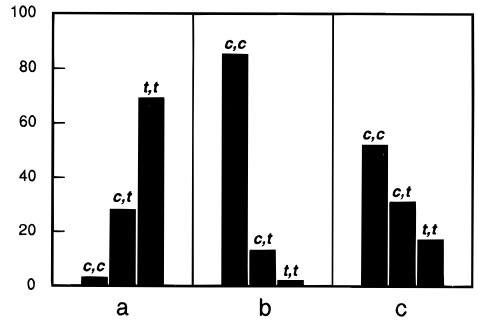


Figure 2. Distribution of the isomers of **1** (%) (a) in solution, after one irradiation cycle; (b) on the arginine-containing column, after 30 selection−isomerization cycles; and (c) in solution, in the equilibrium photostationary state.

silica gel support.[11] While passing through the column most of the high-affinity *cis,cis* isomer was retained on the silica phase as a complex with the attached arginine.

According to the experimental design, the photochemical isomerization of **1** in the subsequent cycle should regenerate the fraction of the high-affinity isomer in the mixture which is in turn transferred to the selection site. Multiple selection−isomerization ("mutation") cycles should lead to increasing the amount of the effective binder (the *cis,cis* isomer) accumulated on the solid support. This method is therefore designed to provide an automatic amplification of the high-affinity component in a system existing in dynamic equilibrium.

After *ca.* 30 isomerization−selection cycles performed within 8 h, the concentration of **1** in the circulating solution decreased to 11% of its original value and the distribution of isomers changed to *cis,cis*/*cis,trans*/*trans,trans* = 48/29/23. We then terminated the experiment and treated the silica with several portions of 1 M aqueous and aqueous/EtOH NaCl to dissociate the isomeric forms of **1** from the immobilized arginine. HPLC analysis of the resulting mixture yielded the distribution of isomers *cis,cis*/*cis,trans*/*trans,trans* = 85/13/2 (Figure 2b), the total yield of **1** on the column being 54 ± 5% of the initial amount.

For comparison, we performed a long irradiation (8 h) of the solution of **1** which led to the equilibrium photostationary distribution of isomers *cis,cis*/*cis,trans*/*trans,trans* = 52/31/17 with the total yield 66% (Figure 2c) as well as to the formation of some side products.[12] In contrast, the enriched mixture removed from the column contained no detectable amount of any side products, all of which were found to remain in the circulating solution. These results show that the circulation experiment leads to (a) amplification of the total amount of the effective binder in the system, (b) considerable selectivity of the isomerization−selection process (*e.g.*, changing the *cis,cis*/*cis,trans* ratio from 1.7 in the photostationary state to 6.5 in the selected subset), and (c) the possibility to isolate only those components of the pool which possess the desired binding affinity separating them from the side products.

A control circulation experiment performed with the sorbent unmodified with arginine (acetylated aminopropyl silica gel) led to only 12% of the isomers having accumulated on the column after 30 cycles, their ratio (*cis,cis*/*cis,trans*/*trans,trans* = 55/31/14) being close to the photostationary distribution in solution. Performing the 30 cycles with the same sorbent but without irradiation resulted in the column distribution *cis,cis*/*cis,trans*/*trans,trans* = 5/20/75 (12% total), probably due to a minor thermal isomerization. Conclusively, the result of the main circulation (evolution) experiment can only be assigned to the gradual amplification of the effective isomer in the system driven by its affinity to the immobilized target compound.

In conclusion, we have developed a new method of targeted evolution of an equilibrating mixture toward the components possessing the highest affinity to a given ligand. This method offers promise for studies of molecular recognition, in that it allows one to reveal the most effective binders among multiple structures, including complex mixtures. In addition, the method may also find synthetic utility for "one-reactor" syntheses of compounds with desired binding properties. Other future applications might include use of the technique to drive combinatorial pools of compounds toward particular subsets of structures.[13] We are currently exploring the possible use of this approach in systems containing reversibly forming *positional* isomers (*e.g.*, oligoaldehydes modified with functional amines *via* imine bonds, *etc*.).

Acknowledgment. This work was supported by the start-up grant to A.V.E. from SUNY/Buffalo and the Petroleum Research Fund (Grant No. 29948-G1). We thank Dr. D. G. Hangauer and Dr. A. J. Solo for their comments on the manuscript and their continued encouragement of this project.

Supporting Information Available: ^{1}H NMR titration curves of all isomers of **1** with methylguanidinium in ethanol, HPLC chromatograms, and the scheme of the experimental setup (8 pages). See any current masthead page for ordering and Internet access instructions.

JA961893R

(11) The concentration of immobilized arginine was chosen according to the binding constant values to provide maximum binding of **1** (*cis,cis*) (90% under given conditions) and moderate or weak binding of the other two isomers (64% and 8% for *cis,trans* and *trans,trans*, respectively). The excess of arginine over **1** was necessary to bind all portions of the effective isomer forming in the subsequent cycles. *N*-α-*t*-BOC-arginine was attached to (aminopropyl)silica (Wu, R.; Grossman, L. *Methods Enzymol*. **1987**, *154*, 299) by the DCC coupling. The residual amino groups on the support were then acylated by Ac_2O in DMF.

(12) HPLC showed formation of two peaks other than the isomers of **1** (see Supporting Information), attributable most likely to the condensation products, similar to the well-studied photodimers of cinnamic acid (see, for example: Nakamura, Y. *J. Chem. Soc., Chem. Commun*. **1982**, 477. D'Auria, M.; Vantaggi, A. *Tetrahedron* **1992**, *48*, 2523).

(13) The choice of terms "ligand" and "receptor" is rather arbitrary; the method can likewise be applied to the equilibration of potential ligands, *e.g.*, enzyme inhibitors if an immobilized enzyme is used for selection. Recent studies on the populations of effective binders in known combinatorial libraries (*e.g.*, Freier, S. M.; Konings, D. A. M.; Wyatt, J. R.; Ecker, D. J. *J. Med. Chem.* **1995**, *38*, 344−352. Wilson-Lingardo, L.; Davis, P. W.; Ecker, D. J.; Hebert, N.; Acevedo, O.; Sprankle, K. Brennan, T.; Schwarsz, L.; Freier, S. M.; Wyatt, J. R. *J. Med. Chem*. **1996**, *39*, 2720) show that small subsets of compounds exhibit considerably larger affinities than the bulk of the components. Thus, our method applied to equilibrating libraries would allow amplification of such subsets and facilitate further identification/generation of the effective components.