Combinatorial Chemistry

DOI: 10.1002/anie.200702460

# Systems Chemistry: Pattern Formation in Random Dynamic Combinatorial Libraries**

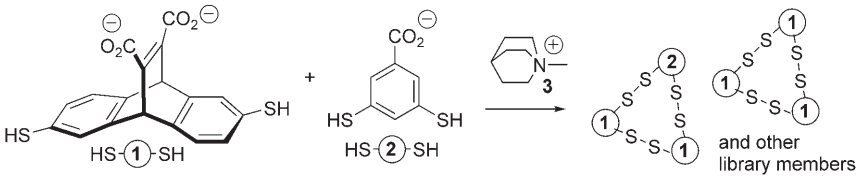
*Peter T. Corbett, Jeremy K. M. Sanders, and Sijbren Otto**

Complex systems have received much attention in mathematics, physics, and biology, but are underdeveloped in chemistry. The study of complex networks of interacting molecules is however of fundamental importance in understanding the organizational principles of biological systems,[1] and may well be the key to unraveling the origin of life.[1,2] One of the central challenges in complex-systems research is establishing how the interactions between the components in a system give rise to emergent properties; that is, characteristics of the system that cannot be ascribed to any of its components acting in isolation. Examples of such emergent behavior in chemical systems include oscillating reactions and the related formation of Turing patterns from a set of interconnected kinetically controlled reactions.[3] Other important recent contributions to systems chemistry[4] include work on self-replicators[5] and self-sorting in multicomponent systems,[6] chemical models for blood clotting,[7] and replicating vesicles.[8]

Dynamic combinatorial libraries (DCLs)[9] present an opportunity for exploring the behavior of mixtures of interacting molecules that are under thermodynamic control. Herein we report a computational study of the global behavior of a DCL in response to molecular-recognition events induced by exposing the complex equilibrium mixture to a template. We show how patterns can emerge in DCLs whose components have random affinities for the template. This unexpected behavior can be traced back to a competition between library members for specific building blocks, in which library members that are not dependent on those building blocks have a competitive advantage.

A DCL is an equilibrium mixture of molecules resulting from exchange between the building blocks from which the library members are constructed. Molecular-recognition events that selectively stabilize particular library members are likely to cause a shift in the library distribution towards these species. The relationship between the amplification factors[10] of the various library members and their affinities for the template has received considerable attention.[11] The emphasis of these studies has been on identifying conditions under which the amplification of library members is selective for the best binder, which ensures that these species are correctly identified in dynamic combinatorial selection experiments. More detailed studies which focus on the global behavior of DCLs as a goal in its own right have only recently started to emerge. These include the use of dynamic combinatorial libraries (DCLs) as sensors,[12] the exploration of different network topologies, based on self-sorting building blocks,[13] the interplay between complementary reversible chemistries,[14] and strategies to simultaneously extract the various host–guest binding constants from library compositions.[15]

We have recently shown how the amplification of the strongly binding receptor $\mathbf{1}_3$ induced by template **3** (Scheme 1) is reduced by binding of the same template to




***Scheme 1.*** Combination of library members **1** and **2** with template **3** to form $\mathbf{1}_2\mathbf{2}$, $\mathbf{1}_3$, and other members.

$\mathbf{1}_2\mathbf{2}$, even though the latter is a weaker binder than $\mathbf{1}_3$.[11e] While it appears likely that such effects not only affect $\mathbf{1}_3$, but propagate through the entire DCL, it is not feasible to study this experimentally owing to the low concentrations of most of the other library members. Thus, we resorted to computation for studying the global response of DCLs to competition scenarios, such as that outlined above, using our recently developed software package (DCLSim) that allows the product distribution of a DCL to be simulated.[11c,d] The input consists of the equilibrium constants that describe the formation of all library members from the building blocks and the free energies of binding of each of the members to the template. No other chemical information, such as structure, is considered. We were particularly interested in the behavior of DCLs in which the template binding energies are drawn randomly from a log-normal distribution (mean $\log K_{\text{template}} = 2$; standard deviation $= 1$).[11c] This situation is likely to reflect that encountered in experimental systems, in which most library members would have a modest affinity for the

[*] Dr. P. T. Corbett, Prof. J. K. M. Sanders, Dr. S. Otto
Department of Chemistry, University of Cambridge
Lensfield Road, Cambridge CB2 1EW (UK)
Fax: (+44) 1223-336017
E-mail: so230@cam.ac.uk
Homepage: http://www-otto.ch.cam.ac.uk

[**] We are grateful for financial support from the Royal Society and EPSRC.

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

template, and only a few would bind particularly strongly (or particularly weakly). Introduction of a template to such a library results in a change in the concentrations of most library members. A typical result of such a theoretical experiment is shown in Figure 1a, where the amplification factor[10] for each library member is plotted against the Gibbs energy of binding to the template. The library was made from seven building blocks A–G (total concentration of 10 mM) which were allowed to combine into 322 different oligomeric species ranging from dimers to tetramers. As intuitively expected, introducing a template (1 mM) causes a net flow of building blocks from the weaker binders to the better binders (Figure 2a).

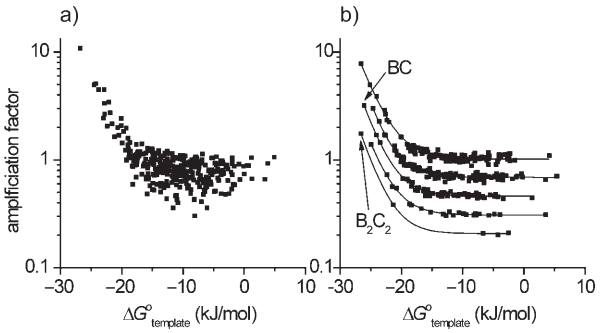



***Figure 1.*** a) Typical correlation between the amplification factor[10] and the affinity for the template for a 322-component DCL with randomly assigned template binding constants. b) Emergence of a pattern in the same library with different set of randomly assigned binding constants. Two members, BC and $B_2C_2$, are indicated, see text for details.

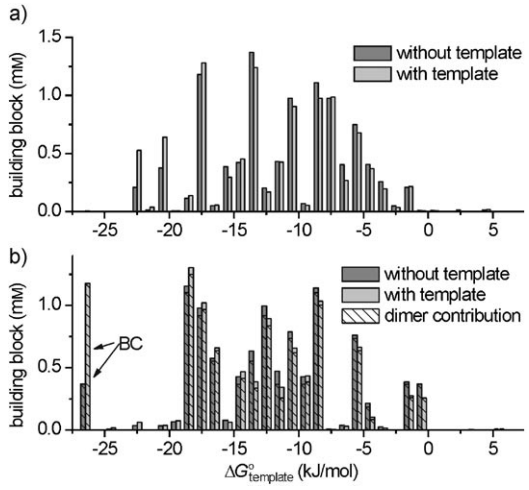



***Figure 2.*** Distribution of building blocks over different library members grouped in template binding affinity classes for the library in a) Figure 1a and b) Figure 1b, in the absence (dark grey) and presence (light grey) of template. The contribution of the dimers in (b) is hatched.

Repeating the experiment in Figure 1a using a different set of random binding affinities, but otherwise identical conditions, can give drastically different results. Under particular circumstances (see below) unexpected patterns emerge, where data points appear clustered in lines. Figure 1b shows an example of such behavior. Inspection of the individual data points reveals that all the library members on the top line are devoid of building blocks B and C. The line below consists exclusively of all the library members into which one copy of B or C has been incorporated. Each of the compounds on the third line contains two units of B or two of C or one of each. The fourth line features the compounds that contain three units of B and/or C, while the bottom line unites the tetramers built up exclusively from B and/or C. Thus, it appears that all compounds in this mixture that require B or C are penalized: with a given template binding energy, their concentration is increased less than that of their counterparts which do not have any of these two subunits in their structures. In the presence of the template, the incorporation of building block B or C requires more free energy than the recruitment of any of the other building blocks. The penalty for incorporating larger numbers of B or C building blocks is correspondingly larger and perfectly additive. This behavior can be traced to the influence of a single library member BC, which binds the template with a high (but not the highest) affinity. Repeating the simulation with weaker binding of BC to the template, but with all other equilibrium constants unchanged, causes the pattern to disappear rapidly: a reduction of BC binding affinity from $-26$ to $-20$ kJ mol$^{-1}$ causes the pattern to disappear completely (Figure 3).

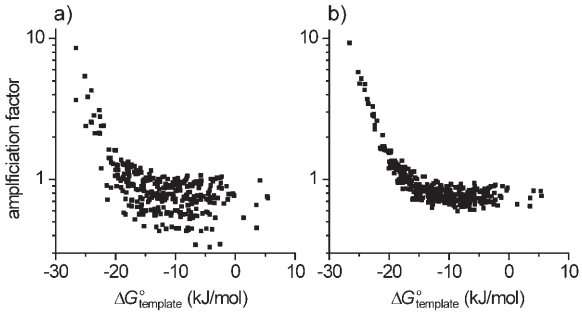



***Figure 3.*** Correlation between amplification factor[10] and template binding affinity for the DCL of Figure 1b but with a reduced $\Delta G^0$ for BC of a) $-24$ kJ mol$^{-1}$ and b) $-20$ kJ mol$^{-1}$. See the Supporting Information for the behavior across a wider range of $\Delta G^0$ values.

It is now well established that, in the presence of excess template, library members that are made from a relatively small number of building blocks have a competitive advantage over those that contain more building blocks.[11a–c] It is important to realize that the library will maximize the template binding energy of the entire system: Given that the amount of available building block is limited, producing many copies of BC ($\Delta G_{template} = -26.1$ kJ mol$^{-1}$) realizes more binding energy than, for example, generating a smaller number of the somewhat better binder $B_2C_2$ ($\Delta G_{template} = -26.6$ kJ mol$^{-1}$; Figure 1b). This effect is similar to that observed experimentally in the competition between $\mathbf{1}_2\mathbf{2}$ and $\mathbf{1}_3$ (see above), which bind template **3** with affinities of $-26.8$ and $-28.0$ kJ mol$^{-1}$, respectively.[11e]

There is a second factor which is of critical importance for emergence of the pattern in Figure 1b: the concentration of BC in the mixture in the absence of template. At moderate building-block concentrations, the formation of dimers such as BC is favored over larger oligomers as it is entropically less demanding to make dimers than trimers or tetramers. This situation is evident from Figure 2b, which shows how the building blocks are distributed over the different affinity classes in the absence and presence of the template. The hatched parts of the bars indicate the contribution of the various dimers, which between them account for 90% of the library material. The leftmost set of bars captures the behavior of dimer BC, which, in the presence of the template, accounts for 41% of the amount of B and C in the library. Repeating the simulation with the original set of binding constants but starting from a tenfold-reduced concentration of dimer BC (in the absence of the template) caused the pattern of Figure 1b to disappear (producing a graph very similar to Figure 3b).

In general, for an individual library member to exert a dominant influence on the global behavior of the mixture, four criteria need to be satisfied: 1) the compound in question needs to have a high affinity for the template, 2) the compound needs to be among the smaller oligomers, 3) the compound needs to be abundant in the absence of the template, and 4) not more than a modest excess of template needs to be available (representative examples of the behavior of the library in Figure 1b at various template concentrations are provided in the Supporting Information). Patterns such as that in Figure 1b emerge when there is only one such high-affinity small oligomer in the mixture. Given that the pattern is caused by a library member with an unexceptional affinity for the template of $-26.1\ \mathrm{kJ\,mol^{-1}}$, similar effects can in principle occur in many of the experimental libraries, as these frequently feature library members with affinities of $-25$ to $-30\ \mathrm{kJ\,mol^{-1}}$.

The results given above have important implications: information on the amount of the template in the mixture can be transmitted to molecules that need not be able to directly interact with this template. This constitutes an indirect form of signaling, which relies on sharing of building blocks rather than on direct molecular interactions. This principle can be used to engineer dynamic mixtures that produce specific responses to the presence of chemical entities. While we have described how an individual library member can influence the behavior of all other library members (potentially inducing patterns), it is also possible to go in the other direction and construct equilibrium mixtures where the concentration of an individual library member is controlled by the properties of the other species in the mixture. Two examples below illustrate how such signaling systems may be constructed.

Consider a small DCL made from equimolar amounts of building blocks A–E, which can all combine with each other and in which compound AB binds to effector molecule T1 ($K_a = 10^6\ \mathrm{M^{-1}}$) while CD binds to another effector species T2 ($K_a = 10^6\ \mathrm{M^{-1}}$). None of the other library members binds T1 or T2. The composition of this dynamic mixture will respond to the presence of T1 by producing more AB and to the presence of T2 by amplifying the amount of CD. Each of these events will indirectly result in an increase in the amount of reporter molecule EE because the alternative partners for building block E are tied up in AB or CD. The largest increase in EE will occur when both T1 and T2 are present (Figure 4a). Representative examples of the behavior of the system at other values of $K_a$ are shown in the Supporting Information.

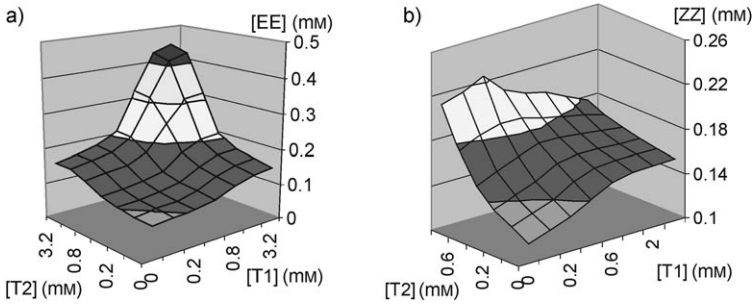



***Figure 4.*** a) Concentration of library member EE as a function of the concentration of effector molecules T1 and T2 in a dynamic library made from building blocks A–E in which T1 binds to AB and T2 to BC, both with $K_a = 10^6\ \mathrm{M^{-1}}$. b) Concentration of library member ZZ as a function of the concentration of effector molecules T1 and T2 in a dynamic library made from building blocks W–Z in which T1 binds to WW and T2 to WXY.

More complex responses to the presence of T1 and T2 can also be engineered. Consider, for example, another dynamic mixture based on equimolar amounts of building blocks W–Z, in which T1 is recognised by WW and T2 by WXY. In this case, ZZ reports on the presence of the effector molecules in a complex fashion: An increase in the amount of T2 will lead to an increase in the concentration of ZZ, but addition of T1 in the presence of T2 will moderate the response. High levels of T1 may cause a reduction of the amount of ZZ (Figure 4b) as T1 will recruit building block W from WXY that was formed under the influence of T2. This then liberates X and Y that can pair up with Z, reducing the amount of ZZ.

In conclusion, dynamic combinatorial libraries constitute a useful new entry point into the as-yet underdeveloped discipline of systems chemistry. Computer simulations represent an efficient way of rapidly exploring the behavior and of improving the understanding of complex equilibrium systems. Starting from unexpected patterns that were observed in libraries with randomly generated affinities between template and library members, the rules that govern such behavior were uncovered. On the basis of these insights, new dynamic mixtures may be engineered that respond to the presence of effector molecules. Such systems provide new opportunities for transmitting information about the composition of mixtures of molecules. They are chemically simple, encompassing one (or more) reversible reaction(s) and a few selective molecular recognition events. It is therefore tempting to speculate that such systems may have played a role in processing molecular information during the early stages of the development of life.[2]

## Experimental Section

The compositions of the dynamic combinatorial libraries were simulated using our DCLSim software package (version 1.1).[11c,d] The equilibrium constants describing the formation of the library members from building blocks were chosen to generate a statistical distribution (i.e., at equal concentrations of building blocks A and B in the absence of template the concentration of AB is twice that of AA). The host–guest binding constants for the libraries of Figures 1 and 2 were assigned randomly from a log-normal distribution ($\mu = 2$ and $\sigma = 1$). These libraries were generated from equimolar amounts of seven building blocks A–G (at a total concentration of 10 mM) which were combined to give all possible library members ranging from dimers to tetramers. No distinction was made between sequence isomers. The amplification factors were calculated by dividing the concentration of the specific library members in the presence of the template (1 mM) by their concentration in the absence of the template. The libraries in Figure 4a and b were generated from equimolar mixtures of five (A–E) and four (W–Z) building blocks (1 mM each), respectively. These were allowed to combine into all possible dimeric and trimeric library members. Equilibrium constants for binding T1 and T2 ($K_a = 10^6\,M^{-1}$) were assigned only to AB, CD, WW and WXY; the other library members did not bind to T1 or T2. The various equilibrium constants and the resulting concentrations of all library members for the DCLs in Figure 1a and b are provided in the supporting information.

Received: June 6, 2007
Revised: July 25, 2007
Published online: October 17, 2007

**Keywords:** combinatorial chemistry · host–guest systems · molecular recognition · supramolecular chemistry · systems chemistry

[1] P. Schuster, P. F. Stadler, *Complexity* **2002**, *8*, 34 – 42.
[2] See: S. A. Kauffman, *The Origins of Order*, Oxford University Press, Oxford, **1993** and references therein.
[3] a) I. R. Epstein, K. Showalter, *J. Phys. Chem.* **1996**, *100*, 13132 – 13147; b) I. R. Epstein, J. A. Pojman, O. Steinbock, *Chaos* **2006**, *16*, 037101; c) G. Biosa, S. Bastianoni, M. Rustici, *Chem. Eur. J.* **2006**, *12*, 3430 – 3437.
[4] For reviews, see: a) G. M. Whitesides, R. F. Ismagilov, *Science* **1999**, *284*, 89 – 92; b) R. F. Ludlow, S. Otto, *Chem. Soc. Rev.*, DOI: 10.1039/B611921M.
[5] a) A. Robertson, A. J. Sinclair, D. Philp, *Chem. Soc. Rev.* **2000**, *29*, 141 – 152; b) N. Paul, G. F. Joyce, *Curr. Opin. Chem. Biol.* **2004**, *8*, 634 – 639; c) G. Ashkenasy, M. R. Ghadiri, *J. Am. Chem. Soc.* **2004**, *126*, 11140 – 11141; d) G. Ashkenasy, R. Jagasia, M. Yadav, M. R. Ghadiri, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 10872 – 10877; e) M. Kindermann, I. Stahl, M. Reimold, W. M. Pankau, G. von Kiedrowski, *Angew. Chem.* **2005**, *117*, 6908 – 6913; *Angew. Chem. Int. Ed.* **2005**, *44*, 6750 – 6755.
[6] a) A. X. Wu, L. Isaacs, *J. Am. Chem. Soc.* **2003**, *125*, 4831 – 4835; b) P. Mukhopadhyay, P. Y. Zavalij, L. Isaacs, *J. Am. Chem. Soc.* **2006**, *128*, 14093 – 14102.
[7] a) M. K. Runyon, B. L. Johnson-Kerner, R. F. Ismagilov, *Angew. Chem.* **2004**, *116*, 1557 – 1562; *Angew. Chem. Int. Ed.* **2004**, *43*, 1531 – 1536; b) C. J. Kastrup, M. K. Runyon, F. Shen, R. F. Ismagilov, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 15747 – 15752.
[8] M. M. Hanczyc, J. W. Szostak, *Curr. Opin. Chem. Biol.* **2004**, *8*, 660 – 664.
[9] P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders, S. Otto, *Chem. Rev.* **2006**, *106*, 3652 – 3711.
[10] The amplification factor is the ratio between the concentration of a specific library member in the presence of the template and its concentration in the absence of template.
[11] a) Z. Grote, R. Scopelliti, K. Severin, *Angew. Chem.* **2003**, *115*, 3951 – 3955; *Angew. Chem. Int. Ed.* **2003**, *42*, 3821 – 3825; b) K. Severin, *Chem. Eur. J.* **2004**, *10*, 2565 – 2580; c) P. T. Corbett, S. Otto, J. K. M. Sanders, *Chem. Eur. J.* **2004**, *10*, 3139 – 3143; d) P. T. Corbett, S. Otto, J. K. M. Sanders, *Org. Lett.* **2004**, *6*, 1825 – 1827; e) P. T. Corbett, J. K. M. Sanders, S. Otto, *J. Am. Chem. Soc.* **2005**, *127*, 9390 – 9392; f) I. Saur, K. Severin, *Chem. Commun.* **2005**, 1471 – 1473; g) G. Gasparini, M. Martin, L. J. Prins, P. Scrimin, *Chem. Commun.* **2007**, 1340 – 1342.
[12] a) A. Buryak, K. Severin, *Angew. Chem.* **2005**, *117*, 8149; *Angew. Chem. Int. Ed.* **2005**, *44*, 7935 – 7938; b) A. Buryak, K. Severin, *J. Comb. Chem.* **2006**, *8*, 540 – 543.
[13] B. Saur, R. Scopelliti, K. Severin, *Chem. Eur. J.* **2006**, *12*, 1058 – 1066.
[14] J. Leclaire, L. Vial, S. Otto, J. K. M. Sanders, *Chem. Commun.* **2005**, 1959 – 1961.
[15] R. F. Ludlow, J. Liu, H. Li, S. L. Roberts, J. K. M. Sanders, S. Otto, *Angew. Chem.* **2007**, *119*, 5864 – 5866; *Angew. Chem. Int. Ed.* **2007**, *46*, 5762 – 5764.