

# Phase Transitions and Control in Self Assembly

Yuliy Baryshnikov<sup>1</sup>, Ed Coffman<sup>2</sup>, and Petar Momčilović<sup>3</sup>

<sup>1</sup> Bell Labs, Lucent Technologies, Murray Hill, NJ 07974

<sup>2</sup> Department of Electrical Engineering, Columbia University, New York, NY 10027

<sup>3</sup> IBM T.J. Watson Research Center, Yorktown Heights, NY 10598

**Abstract.** We introduce a general mathematical model of processes by which self-assembled objects are built from physical (e.g., molecular) units and clusters of these units. We operate within the paradigm of dynamical systems theory, thus representing the self assembly process by a system of ODEs. Our primary interest is in *yield*, i.e., the fraction of the absorbing state containing the desired product, and how it can be controlled. We show that when reaction rates are varied, unexpected phase transitions can occur where yields become 100%.

## 1 Introduction

Self assembly of physical systems brings order out of disorder; it can be described as the basic bottom-up paradigm of numerous constructive and computational processes that have applications spanning nanoscale electronics, biotechnology, medicine, and sensor networks. The scale of the process varies, as do the applications, from the molecular to the astronomical; however, the work proposed here focuses on self assembly at nano scale. Such processes grow complexes from the distributed and autonomous interaction of vast numbers of particles subject to simple bonding rules. Particles are not controlled individually (hence the term “self”), but in the aggregate, e.g., by changes in temperature or relative concentrations of particle types in the initial state.

Research continues apace in molecular self assemblies [1–3] (e.g., polymerization) by which novel functional materials are produced, and nanowire self assembly for producing elementary nanodevices such as switches, gates, and memory elements [4] which can be further assembled into nanocomputers. On the computational side, self assembly plays a fundamental role in amorphous computing [5] and DNA-based computing [6, 7].

The connections with computing remain exciting and promising areas of research, both from the experimental and theoretical points of view. Note the two orthogonal ways that self assembly has influenced research in computer science and engineering: First, the process itself is a computation in DNA-based computing, and second, it is the (nano) fabrication of devices that do the computation. Indeed, the efficiency and robustness of self assembly in these two applications are critical to the success of the underlying concepts.

The dramatic advances in self assembly, and in nanoscience more generally, belong primarily to the experimental scientists; as in other applications,

mathematical modeling and the development of reference theories tend to follow experimental proofs of concept. There have been a number of combinatorial studies [6, 8], but self assembly typically involves random phenomena, and is intrinsically a dynamic process, evolving in continuous time, which puts it in striking contrast with classical technologies and computing paradigms. And although insightful stochastic models and reference theories have been virtually ubiquitous in the physical sciences, they remain a fertile ground for self assembly research, where stochastic analysis has only scarcely begun. The early work of Adleman and colleagues is the springboard in this area, and serves as our point of departure.

In this paper, we study a general model of self assembly not tied to specific applications (although we often adopt the “tile” terminology of DNA models rather than alternatives such as particles). In particular, we study processes in which self-assembled objects are built from physical units (e.g., monomers) and clusters of these units.

We chose to work in the general paradigm of dynamical systems theory, as it reflects, in our view, certain components lacking in available studies of self-assembling systems, yet reflects crucial features of these systems. In particular, we represent the dynamics of self assembly by a system of ordinary differential equations resembling those describing the processes of chemical kinetics. This reduction abstracts the problem of the underlying “physical content” yet allows us to analyze the system quite precisely.

Again, following the standard methodology of bifurcation theory, we introduce the reaction rates as parameters of the system. These reaction rates, one for each distinct cluster pair, influence qualitative and quantitative behavior of the resulting dynamical system, and we concentrate on this dependence. In particular, we are interested in the *yield*, i.e., the fraction of the absorbing state that contains the desired product, or self-assembled objects. We show that when reaction rates are varied, unexpected phase transitions can occur where yields become 100%.

A need for high yields motivated our work in [9]: the analysis of a control mechanism (modeling temperature variation for example) that implements *reversible* self-assembly, where disassembly of objects can be exploited to improve the overall assembly process. We review these results here.

## 2 A baseline self assembly model

In the DNA-computing terms of [10] elementary particles called *tiles* combine to form progressively larger geometric shapes, or complexes, called *supertiles*. Hereafter, we say *i-tile*,  $i \geq 1$ , when referring to a supertile containing  $i$  unit tiles.

In the general tile-system model, there is a set of tile types and a collection of bonding (sticking or gluing) rules governing how tiles and supertiles combine into larger supertiles. In the model here, all elementary, unit tiles are subject to the same bonding rules, or equivalently, the system is homogeneous with just

a single tile type. Supertiles of size  $i$  combine with supertiles of size  $j$  to form supertiles of size  $i + j$  subject to a reaction rate (or sticking probability) of  $p_{ij} \geq 0$ . These rates, semantically, reflect the energetic reaction threshold (and temperature) via the familiar formula

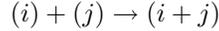
$$p_{ij} \approx \exp\left(-\frac{U_{ij}}{kT}\right),$$

where  $U_{ij}$  is the energetic barrier for tiles of sizes  $i$  and  $j$  to combine. Notice that we assume that once combined, the tiles are locked, i.e., the process is irreversible.

The crucial assumptions about the system are the following:

- tiles of size  $N$  precipitate, where  $N$  is a given parameter of the tile system,
- no combination resulting in tiles heavier than  $N$  is allowed.

We assume **ergodicity**: the rate of collision of particles of the same type  $i$  is proportional to  $x_i^2$ , while the rate of collision of different types  $i$  and  $j$  is proportional to  $2x_i x_j$ . Hence, the total rate of reaction



is the product of the rate of collisions of  $i$ - and  $j$ -tiles, and of  $p_{ij}$ .

Combining all possible interactions, we arrive at the following **master system** of ODEs:

$$\begin{aligned} \frac{dx_1}{dt} &= -2p_{1,1}x_1^2 - 2p_{1,2}x_1x_2 - \dots - 2p_{1,N-1}x_1x_{N-1} \\ \frac{dx_2}{dt} &= p_{1,1}x_1^2 - 2p_{1,2}x_1x_2 - 2p_{2,2}x_2^2 \dots \\ &\vdots \\ \frac{dx_N}{dt} &= 2p_{1,N-1}x_1x_{N-1} + 2p_{2,N-2}x_2x_{N-2} + \dots \end{aligned}$$

Note that there is no intrinsic limitation to spatial linearity in this model. Also, since we are exploring behavior in the fluid limit, our results also apply to situations in which there are different tile types, as long as they have the same initial concentrations and symmetric behavior. For example, suppose the final supertiles are 3-tiles, there are 3 tile types, no two tiles of the same type can appear in the same supertile, and the concentrations of the 3 tile types are equal in the initial state.

### 3 $N = \infty$

As often happens, the case where the size of the precipitating supertile  $N$  is very large can be analyzed explicitly, if one assumes all  $p_{ij} = 1$ .

Indeed, we have

$$\frac{dx_l}{dt} = - \sum_{i=1}^{\infty} 2x_l x_i + \sum_{m+n=l} x_m x_n.$$

Passing to the generating function  $f(z, t) = \sum_{i=1}^{\infty} x_i(t) z^i$  yields

$$\frac{\partial f(z, t)}{\partial t} = -2f(z, t)f(1, t) + f(z, t)^2.$$

Solving first for  $\phi(t) = f(1, t)$  gives  $\phi(t) = 1/(t+1)$ , whence

$$f(z, t) = (t+1)^2 \sum_{k=0}^{\infty} \left( \frac{t}{t+1} \right)^k f_0(z)^k,$$

where  $f_0(z) = f(z, 0)$ .

In particular, if  $f_0(z) = z$  (which corresponds to the case of interest for us, when all the tiles initially are of size 1), then at any time  $t$ , the generating function  $f(z, t)$  describes a geometric distribution,

$$x_k(t) = (t+1)^2 \left( \frac{t}{t+1} \right)^k.$$

The profile of concentrations is geometric, spreading further and further as  $t$  grows. One can easily adjust this method of solution to the situation where the rate matrix  $(p_{ij})$  is of rank 1, i.e. when

$$p_{ij} = p_i p_j$$

for some sequence  $\{p_i\}_{i \geq 1}$ ,  $p_i \geq 0$ .

## 4 Assembly of triangles

For  $N = 3$  the master system is

$$\begin{aligned} \frac{dx_1}{dt} &= -2p_{11}x_1^2 - 2p_{12}x_1x_2, \\ \frac{dx_2}{dt} &= p_{11}x_1^2 - 2p_{12}x_1x_2. \end{aligned}$$

(The special case with  $p_{11} = 1$  and  $p_{12} = 1/2$  was presented in our earlier paper [9].) The *trajectories* of this system satisfy the following ODE (we use the notation  $y := x_2(x_1)$ ,  $x := x_1$  for brevity):

$$\frac{dy}{dx} = \frac{y - (p/2)x}{y + px}, \quad y(1) = 0,$$

where  $p = p_{11}/p_{12}$  is the sole parameter determining the limiting concentration  $y(0) = x_2(\infty)$ . Let  $z_1, z_2$  be the two roots of

$$z^2 + (p-1)z + p/2 = 0$$

Computations give the following.

1. If  $p \geq 2 + \sqrt{3}$ , then  $z_{1,2}$  are both real and negative, and

$$x_2(\infty) = (-z_1)^{\frac{1-z_2}{z_2-z_1}} (-z_2)^{\frac{z_1-1}{z_2-z_1}},$$

where

$$z_1 = \frac{1-p - \sqrt{p^2 - 4p + 1}}{2}$$

$$z_2 = \frac{1-p + \sqrt{p^2 - 4p + 1}}{2}.$$

In particular, as  $p \rightarrow \infty$ ,

$$z_1 \approx p, \quad z_2 \rightarrow 1/2,$$

and

$$x_2(\infty) \approx 1/2 - \ln p/p.$$

2. If  $2 - \sqrt{3} < p < 2 + \sqrt{3}$ , then

$$x_2(\infty) = \sqrt{\frac{p}{2}} \exp \left( \frac{p+1}{\sqrt{-p^2 + 4p - 1}} \left( \arccos \frac{1-p}{\sqrt{2p}} - \pi \right) \right).$$

3. The most interesting behavior occurs when  $p = 2 - \sqrt{3}$ , for if  $0 \leq p \leq 2 - \sqrt{3}$ , then  $x_2(t) \rightarrow 0$  as  $t \rightarrow \infty$ . In fact,  $x_2(t)/x_1(t) \rightarrow z_2$ , where

$$z_2 = \frac{(1-p + \sqrt{p^2 - 4p + 1})}{2}$$

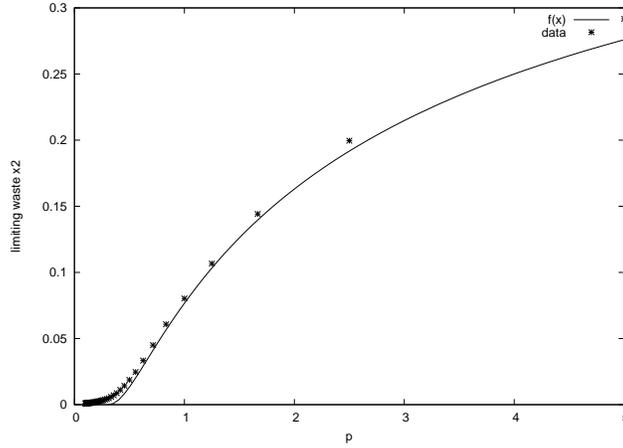
is the larger of the two positive roots of  $z^2 + (p-1)z + p/2$ .

The plot in Figure 1 gives the waste as a function of  $p$  (together with simulation results).

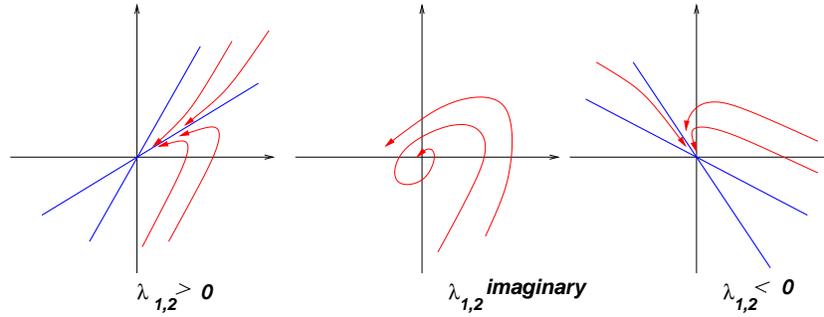
In other words, this dynamical system exhibits a *phase transition* from zero waste (for  $p < p_* = 2 - \sqrt{3}$ ) to positive waste ( $p > p_* = 2 - \sqrt{3}$ ).

At first glance, this phase transition might seem surprising; indeed, one would expect that the asymptotic yield increases as  $p$  decreases (one releases the 2-tiles more slowly, and they become easily 3-tiles absorbing relatively abundant 1-tiles), but in principle some positive amount of wasted material would seem natural for any system of this kind. Yet, there is some threshold, where 100% of efficiency is achieved, still for a positive ratio  $p$ .

A closer look shows what happens here. When  $p$  is large, the linear system describing the trajectories has real eigenvalues and the eigenvectors are pointing in the 2nd and 4th quadrants. Then the eigenvalues become imaginary. At  $p = 2 - \sqrt{3}$  one more bifurcation occurs, but the eigenvectors are now in the 1st and 3rd quadrants. See Figure 2 for details.



**Fig. 1.** Phase transition in waste as a function of  $p$



**Fig. 2.** Eigenvectors of the linear system

## 5 Phase transition for general systems

In general, the phase transition described in the preceding section for  $N = 3$  can be observed for any  $N$ . That is

**Theorem 1** *The space of parameters  $\mathcal{P} = \{(p_{ij}), p_{ij} \geq 0, j < N\}$  contains open nonempty sets  $\mathcal{P}_0$  and  $\mathcal{P}_+$  such that for parameters in the closure of  $\mathcal{P}_0$  the limiting yield  $Nx_N(\infty) = 1$ , while for parameters in  $\mathcal{P}_+$  the limiting waste is positive,  $Nx_N(\infty) < 1$ .*

We will just sketch a proof here; a complete exposition will appear elsewhere.

To prove the theorem, we first exhibit, for any  $N$ , some systems that have positive waste. Clearly, this will prove that  $\mathcal{P}_+$  is nonempty; its openness would follow from general results on the continuous dependence on parameters.

## 5.1 Some special systems, for arbitrary $N$

Let us set  $p_{ij} = 0$  when both  $i, j > 1$ , i.e. we assume that the rates are non-vanishing only when at least one of the tiles being attached has size 1, a common assumption as noted in [11]. We call this *incremental self assembly*. Further, set  $p_{1,1} = 1$ ,  $p_{1,i} = 1/2$ ,  $j > 1$ . (In the original set-up this would result if tile *pairs* were chosen at a rate proportional to their concentrations.) We covered this case recently in [9]. There, we arrived promptly at

$$\begin{aligned}\frac{dx_1}{dt} &= -x_1(2x_1 - \sum_{k=2}^{N-2} x_k) \\ \frac{dx_k}{dt} &= x_1(x_{k-1} - x_k), \quad k = 2, \dots, N-1.\end{aligned}$$

or, with the substitution  $ds := x_1 dt$ , the system

$$\begin{aligned}\frac{dx_1}{ds} &= -2x_1 - \sum_{k=2}^{N-2} x_k \\ \frac{dx_k}{ds} &= x_{k-1} - x_k, \quad k = 2, \dots, N-1.\end{aligned}$$

Surprisingly, this system can be solved explicitly, as we showed in [9]:

$$x_k(s) = \frac{1}{N} \sum_{\ell=1}^{N-1} (\varepsilon^{-\ell} - 1) \varepsilon^{\ell k} e^{\lambda_\ell s}$$

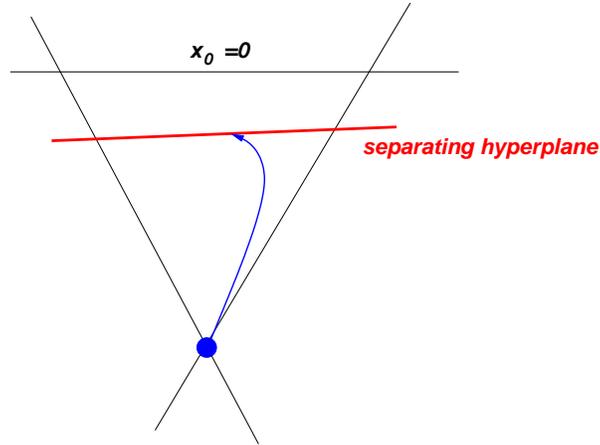
where  $\lambda_\ell = e^{i2\pi\ell/N} - 1$ ,  $\ell = 1, \dots, N-1$ . In particular, one can deduce that the waste (the residual mass not precipitated as the product ( $N$ ) in the limit  $x_1 \rightarrow 0$ ) is always positive for these values of  $p_{ij}$  (and in fact, the total wasted mass rapidly converges to 1 as  $N$  grows).

## 5.2 Phase transitions for arbitrary $N$

On the other hand, consider the incremental system with  $p_{1,1}$  close to 0. We claim that the yield in such systems is 100%, for  $p = p_{1,1} > 0$  small enough. Indeed, one can show that for small  $p$ , the corresponding linear system described by the matrix

$$\begin{bmatrix} -2p & -2p_{1,2} & 0 & \dots & -2p_{1,N-1} \\ p & -2p_{1,2} & 0 & \dots & 0 \\ 0 & 2p_{1,2} & -2p_{1,3} & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & -2p_{1,N-1} \end{bmatrix}$$

has a *left* eigenvector  $(c_1, \dots, c_{N-1})$  with  $c_1 > 0$ ,  $c_i < 0$ ,  $i > 1$ . This eigen(co)vector defines a hyperplane which separates the initial point  $x_0 = (1, 0, \dots, 0)$  from the face of the waste  $\{x_1 = 0, \sum_{i>1} x_i > 0\}$  on the phase space simplex.



**Fig. 3.** Existence of phase transitions for general  $N$

Standard arguments (homogeneity of the system and a certain structural stability) prove that a separating hypersurface survives small perturbations of the coefficients.

### 5.3 An example

A small illustration is given in Figure 4 and shows the critical curve, i.e., plot of  $(p, \lambda)$  for which the matrix describing the *linearized* system for  $N = 4$ , with  $p_{1,1} = p$  has  $\lambda$  as eigenvalue. The prediction is that the phase transition happens at the critical value of  $p(\lambda)$  closest to the origin, i.e., at  $p \approx 0.012$ .

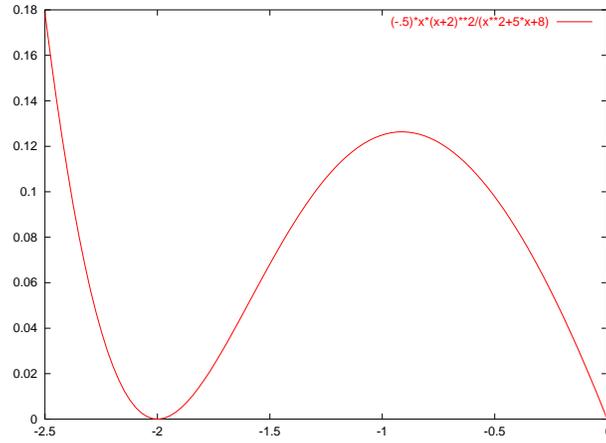
The results of experiments are shown in Figure 5.

## 6 The reversible process

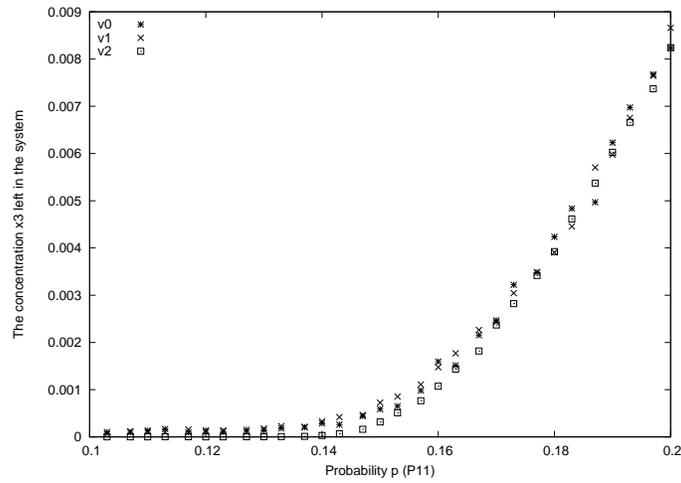
Higher, indeed maximum, yields are achieved by the reversible process in which  $i$ -tiles,  $i > 1$ , can decompose, returning two smaller tiles to the system. Reversibility is controlled by a time-varying parameter  $u(t)$  and is reflected in “control terms” appended to the rate equations. For example, in [9], we considered the triangle ( $N = 3$ ) case in the incremental model. We have

$$\begin{aligned}\frac{dx_1}{dt} &= -2x_1^2 - x_1x_2 + 2ux_2, \\ \frac{dx_2}{dt} &= x_1^2 - x_1x_2 - ux_2\end{aligned}$$

One may think of  $u(t)$  as time-varying temperature.



**Fig. 4.**  $N = 4$  example



**Fig. 5.** Experimental results showing phase transitions

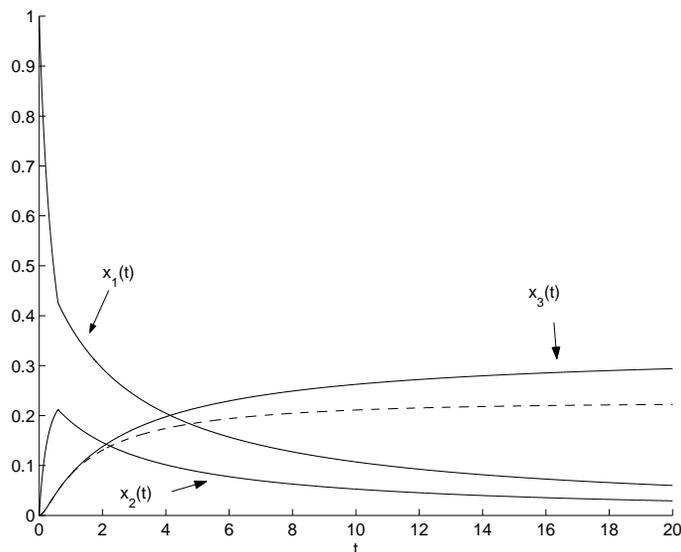
The goal now is to set up the control so that, for a given time horizon  $T$ , the concentrations converge to a state at time  $T$  in which the yield is maximized. We showed in [9] that the optimal strategy is defined by

$$u(t) = \frac{7/2 \cdot 1\{t > \tau\}}{3(t - \tau)/2 + 1/\alpha}$$

where  $\tau := \inf\{t > 0 : x_1 = 2x_2\}$  and  $x_2(\tau) = \alpha$  must be determined numerically. In words, the system is allowed to evolve (absent control) from the initial state

$x_1 = 1$  until such time as the concentration of 1-tiles first becomes exactly twice that of 2-tiles (at time  $\tau$ ). At this point, control is exerted as above so as to preserve the relative concentrations of 1- and 2-tiles.

The numerical solutions for  $x_1$ ,  $x_2$  and  $x_3$  as functions of time for the reversible process in the optimal-control case are shown in Figure 6 taken from [9]. As can be seen in the figure, the reversible set-up results in a higher yield than the irreversible one.



**Fig. 6.** Concentrations of 1, 2 and 3-tiles as functions of time for the reversible process under the optimal control. The assembly starts at time  $t = 0$ . After time  $t = \tau$  the ratio of  $x_1/x_2$  is fixed at 2. The dashed line indicates the concentration of 3-tiles in the corresponding irreversible process.

## References

1. Lopinski, G., Wayner, D., Wolkow, R.: Self-directed growth of molecular nanostructures. *Nature* **406** (2000)
2. Gomez-Lopez, M., Preece, J., Stoddart, J.: The art and science of self-assembling molecular machines. *Nanotechnology* **7** (1996) 183–192
3. Whitesides, G., Mathias, J., Seto, C.: Molecular self-assembly and nanochemistry: A chemical strategy for the synthesis of nanostructures. *Science* **254** (1991) 1312–1319
4. Gracias, D., Tien, J., Breen, T., Hsu, C., Whitesides, G.: Forming electrical networks in three dimensions by self-assembly. *Science* **289** (2000) 1170–1173

5. Abelson, H., Allen, D., Coore, D., Hanson, C., Homsy, G., Knight, T., Nagpal, R., Rauch, E., Sussman, G., Weiss, R.: Amorphous computing. *Communications of the ACM* **43** (2000) 74–82
6. Adleman, L., Cheng, Q., Goel, A., Huang, M.D.: Running time and program size for self-assembled squares. In: *Proc. ACM Symp. Th. Comput.* (2001) 740–748
7. Carbone, A., Seeman, N.: Circuits and programmable self-assembling DNA structures. *Proc. Natl. Acad. Sci. USA* **99** (2002) 12577–12582
8. Adleman, L., Cheng, Q., Goel, A., Huang, M.D., Kempe, D., de Espanés, P.M., Rothmund, P.: Combinatorial optimization problems in self-assembly. In: *Proc. ACM Symp. Th. Comput., Montreal, Canada* (2002) 23–32
9. Baryshnikov, Y., Coffman, E., Momčilović, P.: Incremental self-assembly in the fluid limit. In: *Proc. 38th Ann. Conf. Inf. Sys. Sci., Princeton, NJ* (2004)
10. Adleman, L., Cheng, Q., Goel, A., Huang, M.D., Wasserman, H.: Linear self-assemblies: Equilibria, entropy, and convergence rates. In Elaydi, Ladas, Aulbach, eds.: *New progress in difference equations*. (Taylor and Francis, London (to appear))
11. Rothmund, P., Winfree, E.: The program-size complexity of self-assembled squares. In: *Proc. ACM Symp. Th. Comput.* (2001) 459–468