Autocatalysis

DOI: 10.1002/anie.200802177

Systems Chemistry on Ribozyme Self-Construction: Evidence for Anabolic Autocatalysis in a Recombination Network

*Eric J. Hayden, Günter von Kiedrowski, and Niles Lehman**

Abiogenesis requires a transition from simple chemical precursors to macromolecules that can catalyze their own production. This development necessarily requires two types of reaction pathways: anabolic (the formation of macromolecules) and autocatalytic (the feedback by certain products to enhance their own production). Studies of autocatalytic feedback are central to containment-based, metabolic, and genetic theories on the origins of life.[1–6] In principle, metabolic autocatalysis can run in an anabolic or catabolic direction. Catabolic autocatalysis is understood in the sense that larger or more complex molecules are autocatalytically transformed into smaller or simpler products. Implementations are found in two recent schemes involving hybridization networks.[7,8] Genetic autocatalysis is usually attributed to self-replication (the autocatalytic transfer of information from templates to copies), and examples based on nucleic acids,[9–14] peptides,[15,16] and organic molecules[17] as templates have been studied. This approach was greatly influenced by Eigen's description of the self-organization of matter that requires autocatalysis to evolve from a prebiotic "chemical" phase to a self-organization of replicating "individuals" phase, which must have occurred contemporaneously.[18] Therefore, the origin of life requires reaction networks that define the very boundaries of these phases, involving both anabolic reactions and autocatalytic feedback.

The significance of autocatalysis within complex reaction networks is underlined by recent emphasis on systems chemistry, the main objectives of which are to investigate autocatalytic reaction systems within supramolecular, prebiotic, and other fields of chemistry.[19] We report herein on a case of anabolic autocatalysis within a reaction network that assembles a functional ribozyme from smaller oligonucleotide precursors.

We previously described this system,[20,21] which starts from four RNA fragments of the Azoarcus group I intron, termed **W**, **X**, **Y**, and **Z**, that self-assemble into an active complex that then catalyzes the recombination of these fragments to produce a covalently contiguous ribozyme, **W·X·Y·Z** (Scheme 1). This set of reactions demonstrates a spontaneous path to the formation of a single catalytic molecule from four individually inactive oligonucleotides. Once formed, this covalently contiguous ribozyme has the opportunity to act as an autocatalyst by enhancing the rates of each reaction leading to its own production (Scheme 1a, curved arrows). Previously, autocatalysis was inferred by an increase in the initial rate of **W·X·Y·Z** production when the reaction was seeded with this product.[20] Herein, we confirm the autocatalytic capacity of this system and describe in detail how autocatalysis contributes to the reaction network and directs intermediates to the production of **W·X·Y·Z**.

We first set out to isolate and study the three component reactions in which **W·X·Y·Z** has the potential for direct (proximal) autocatalysis by catalyzing its own production in a single step (Scheme 1a, solid curved arrows). To do this, we synthesized each pair of RNA fragments necessary for the reactions in Equations (1–3).

$$\mathbf{W} + \mathbf{X}\cdot\mathbf{Y}\cdot\mathbf{Z} \rightleftharpoons \mathbf{W}\cdot\mathbf{X}\cdot\mathbf{Y}\cdot\mathbf{Z} \quad (1)$$

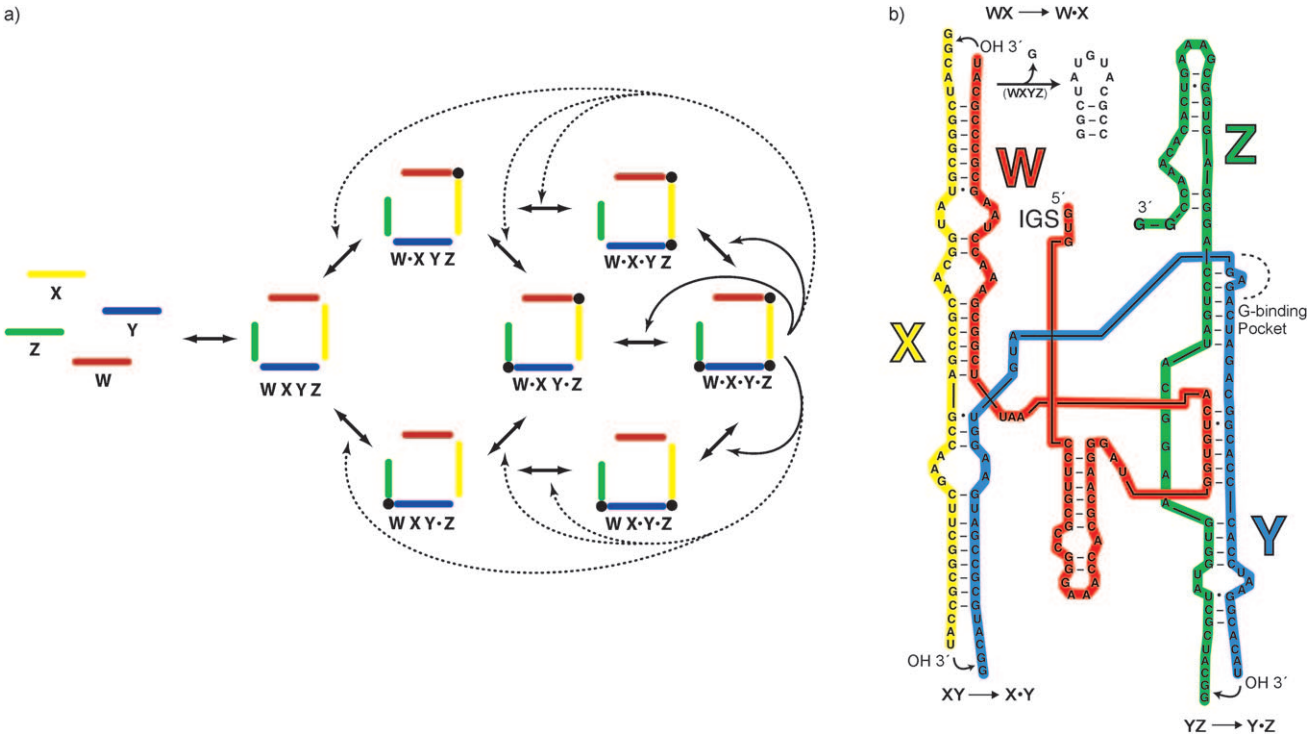
$$\mathbf{W}\cdot\mathbf{X} + \mathbf{Y}\cdot\mathbf{Z} \rightleftharpoons \mathbf{W}\cdot\mathbf{X}\cdot\mathbf{Y}\cdot\mathbf{Z} \quad (2)$$

$$\mathbf{W}\cdot\mathbf{X}\cdot\mathbf{Y} + \mathbf{Z} \rightleftharpoons \mathbf{W}\cdot\mathbf{X}\cdot\mathbf{Y}\cdot\mathbf{Z} \quad (3)$$

We then studied the effect of doping **W·X·Y·Z** on the initial rate of its own production in each reaction (Figure 1). In all cases, when 1 μM of each fragment was reacted, the initial rate of **W·X·Y·Z** formation increased linearly with the initial concentration of **W·X·Y·Z** up to 2 μM. The positive slopes of the plots demonstrate autocatalysis in each reaction. In fact, the linear increase in the initial rate of **W·X·Y·Z** formation indicates a potential for exponential growth.[22] The data can be fit to a line of the form $k_i = k_a[\mathbf{W}\cdot\mathbf{X}\cdot\mathbf{Y}\cdot\mathbf{Z}] + k_b$. For this system, k_b (y intercept) is the rate constant exhibited in the absence of covalent **W·X·Y·Z**, and k_a (slope) is the empirical autocatalytic rate constant; k_i is the initial rate of the overall reaction. Equation (3) demonstrates the highest rate contributions for both the non-autocatalytic and autocatalytic pathways. However, by taking the ratio k_a/k_b for each reaction, we can compare the autocatalytic efficiencies ($\varepsilon = k_a/k_b$), which quantify the effect the autocatalyst has relative to the non-autocatalytic rate.[23] It should be noted that this treatment of the data is a simplification that ignores the concentration of noncovalent complexes and binding efficiencies of each catalyst, but it provides a convenient way of comparing each autocatalytic path. When juxtaposed in this way, Equation (2) demonstrates the highest efficiency, while Equation (3) actually has the lowest. The lower autocatalytic efficiency of Equation (3) is a result of it having the highest

[*] Dr. E. J. Hayden, Prof. N. Lehman
Department of Chemistry, Portland State University
PO Box 751, Portland, OR 97207 (USA)
Fax: (+1) 503-725-8769
E-mail: niles@pdx.edu
Homepage: http://www.chem.pdx.edu/~niles/lehman_lab_at_psu/home.html
Prof. Dr. G. von Kiedrowski
Lehrstuhl für Organische Chemie I, Bioorganische Chemie
Ruhr-Universität Bochum (Germany)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200802177.



Scheme 1. The anabolic autocatalytic reaction network formed from the four fragments of the Azoarcus ribozyme. a) A scheme of covalent-bond-forming reactions (· represents a covalent bond) from the four fragments through a series of noncovalent intermediates to the fully covalent ribozyme **W·X·Y·Z**. Double-sided arrows indicate reversible phosphodiester bond formation, which is catalyzed initially by the noncovalent complex. Curved arrows indicate feedback by **W·X·Y·Z** to catalyze its own production in a single step (solid), or by catalyzing the production of an intermediate (dashed). Intermediates can also catalyze their own production, as indicated by the gray curved arrow. b) Proposed secondary structure of the noncovalent intermediate **WXYZ**, and the tRNA-like splicing mechanism. IGS = internal guide sequence.[25]

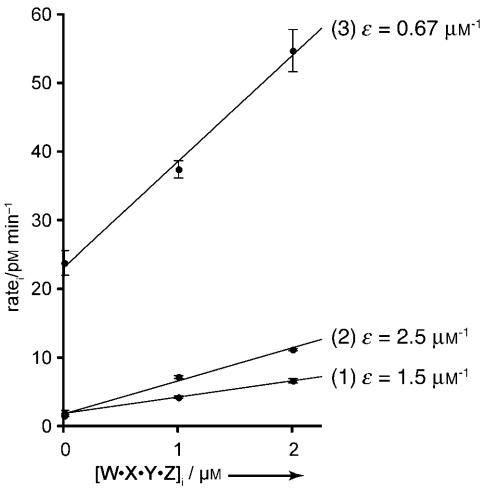


Figure 1. Initial rate of **W·X·Y·Z** formation as a function of the initial concentration of **W·X·Y·Z** for the three component reactions of the network in which **W·X·Y·Z** is formed in a single step. Reactions containing 1 μM oligonucleotides **W** + **X·Y·Z** (1), **W·X** + **Y·Z** (2), or **W·X·Y** + **Z** (3) were incubated with 0–2 μM **W·X·Y·Z**. Data points represent the average of three independent trials, and lines are linear fits to the data. The autocatalytic efficiencies ε are the ratio of the slope to the y intercept for each line.

rate in the non-autocatalytic channel. The noncovalent catalyst in Equation (3) is broken only at the **Y**–**Z** junction, retaining in a single molecule (**W·X·Y**) the bulk of the structural integrity of the native ribozyme, as determined by folding assays.[24,25] All three ε values lie toward the higher end of the spectrum of efficiencies reported for other autocatalytic systems.[26] Strong autocatalytic enhancement was also detected in a peptide system that can self-replicate through bipartite ligation reactions, despite reliance on the competition of nontemplated background reactions for initiation.[16]

Many other feedback reactions are imaginable within this system (Scheme 1 a). In addition to the proximal autocatalytic steps described above, there are distal autocatalytic steps, in which **W·X·Y·Z** catalyzes a reaction that produces an intermediate of the network. To evaluate the total autocatalysis of the system in the context of the entire reaction network, we developed a kinetic model and fit it to our data using SimFit[27] (Table 1).

Reactions a–c (Table 1) account for the formation of the noncovalent complex **WXYZ**, the system's initial catalyst. The model also allows for five (reversible) covalent-bond-forming reactions each to be catalyzed by the noncovalent complex (reactions d–h in Table 1) and the fully covalent product **W·X·Y·Z** (reactions i–m in Table 1). Reactions d–f and i–k account for a pathway to **W·X·Y·Z** through intermediates **W·X** and **W·X·Y**, which can be observed on a denaturing gel when **W** is radiolabeled at the 5′ terminus.[25] However, because the concentration of **W·X·Y·Z** quickly surpasses that of **W·X·Y**, and because the concentration of **W·X·Y** does not pass through a maximum, there is at least one

Table 1: The reaction model and calculated rate parameters. Catalysis is represented by the catalyst appearing on both sides of the reaction equation.

	Type[a]	Reaction	$k_{forward}$	$k_{reverse}$
a	SA	**W** + **X** ⇄ **WX**	1.0×10^{6} (fixed)	$(9.7\pm1.5)\times10^{-3}$
b	SA	**WX** + **Y** ⇄ **WXY**	1.0×10^{6} (fixed)	$(9.7\pm1.5)\times10^{-3}$
c	SA	**WXY** + **Z** ⇄ **WXYZ**	1.0×10^{6} (fixed)	$(9.7\pm1.5)\times10^{-3}$
d	NC	**WXYZ** + **WXYZ** ⇄ **W·XYZ** + **WXYZ**	$(1.9\pm0.41)\times10^{2}$	$(1.1\pm1.7)\times10^{3}$
e	NC	**W·XYZ** + **WXYZ** ⇄ **W·X·YZ** + **WXYZ**	$(7.6\pm2.0)\times10^{2}$	$(6.6\pm7.8)\times10^{3}$
f	NC	**W·X·YZ** + **WXYZ** ⇄ **W·X·Y·Z** + **WXYZ**	$(2.3\pm0.63)\times10^{2}$	$(8.2\pm4.7)\times10^{2}$
g	NC	**W·XYZ** + **WXYZ** ⇄ **W·XY·Z** + **WXYZ**	$(2.3\pm0.63)\times10^{2}$	$(8.2\pm4.7)\times10^{2}$
h	NC	**W·XY·Z** + **WXYZ** ⇄ **W·X·Y·Z** + **WXYZ**	$(7.6\pm2.0)\times10^{2}$	$(6.6\pm7.8)\times10^{2}$
i	AC	**WXYZ** + **W·X·Y·Z** ⇄ **W·XYZ** + **W·X·Y·Z**	$(2.2\pm1.1)\times10^{4}$	$(3.5\pm1.7)\times10^{4}$
j	AC	**W·XYZ** + **W·X·Y·Z** ⇄ **W·X·YZ** + **W·X·Y·Z**	$(5.9\pm4.6)\times10^{2}$	$(1.2\pm0.34)\times10^{2}$
k	AC	**W·X·YZ** + **W·X·Y·Z** ⇄ **W·X·Y·Z** + **W·X·Y·Z**	$(1.6\pm1.5)\times10^{3}$	$(6.8\pm5.8)\times10^{3}$
l	AC	**W·XYZ** + **W·X·Y·Z** ⇄ **W·XY·Z** + **W·X·Y·Z**	$(1.6\pm1.5)\times10^{3}$	$(6.8\pm5.8)\times10^{3}$
m	AC	**W·XY·Z** + **W·X·Y·Z** ⇄ **W·X·Y·Z** + **W·X·Y·Z**	$(5.9\pm4.6)\times10^{2}$	$(1.2\pm0.34)\times10^{2}$

[a] SA = self-assembly, NC = non-autocatalytic catalysis, AC = autocatalysis.

other significant pathway to **W·X·Y·Z**. From previous results, we saw that **X·Y** is not a populated intermediate.[20] Thus reactions g, h, l, and m were chosen to account for a path to **W·X·Y·Z** that does not go through **W·X·Y**.

This model attempts to explain the data through the following simplifying assumptions: 1) Formation of covalent bonds is truly reversible; 2) the products reach equilibrium through catalysis; 3) catalysis occurs by either a termolecular noncovalent complex (substrate) or a completely covalent unimolecular product; and 4) catalytic steps are rate-limiting.[28] Consequently, the model sacrifices mechanistic detail to highlight the difference between initial catalysis and product catalysis.[27]

The model was fitted to time-course data of the observables **W·X**, **W·X·Y**, and **W·X·Y·Z** obtained from denaturing PAGE of a reaction starting with 1 μM each **W**, **X**, **Y**, and **Z** (Figure 2a). The model fits well to the experimental data, with a root mean square (RMS) value of 3.5%. Models that did not include autocatalysis, but only noncovalent catalysis, did not give a satisfactory fit to the data, with an RMS of no less than 13%.[25] Removing the autocatalytic rate parameters after fitting gives a visual image of the extent of autocatalytic feedback in the reaction network (Figure 2b). From this model, noncovalent catalysis can only account for approximately 33% of the **W·X·Y·Z** formed after 240 min. Autocatalytic feedback is responsible for the remaining approximately 66%.

The model predicts that the sigmoidal shape of **W·X** formation (Figure 3a, dashed curve) results from strong distal autocatalytic feedback by **W·X·Y·Z**. We thus hypothesized that seeding the reaction with **W·X·Y·Z** should increase the initial rate of **W·X** formation and decrease the lag phase of the reaction profile. When we repeated the reaction but seeded it with 1 μM **W·X·Y·Z**, the experimental results confirmed the hypothesis formed from the model (Figure 3a, upper solid curve).

Other intermediates within the system also have the potential to act autocatalytically to enhance their own production through participation in noncovalent complexes such as **W·XYZ** and **W·X·YZ** (e.g. Scheme 1a, gray arrow). To characterize this aspect of autocatalysis, we repeated the reaction with 1 μM each **W**, **X**, **Y**, and **Z**, but seeded the reaction with either 1 μM **W·X** or 1 μM **W·X·Y**. The initial rate of **W·X** formation is only slightly altered by the addition of **W·X**, indicating that this intermediate has very limited autocatalytic potential (Figure 3a). The intermediate **W·X·Y**, however, shows a noticeable increase in the rate of its own production (Figure 3b). In fact, this intermediate seems specifically to catalyze its own production over other intermediates, because it does not dramatically increase the rates of either **W·X** production or **W** consumption.[25] This finding is consistent with the data in Figure 1, which shows a large noncatalytic rate constant (*y* intercept) for Equation (3), confirming the catalytic prowess of the noncovalent complex **W·X·YZ**, as noted above. These data rank the efficiency of autocatalysis as **W·X** < **W·X·Y** < **W·X·Y·Z**, meaning the total autocatalysis of the network increases as longer molecules are produced. Hence, **W·X·Y·Z** is the most effective autocatalyst in the system, further increasing our confidence in the parameterization of the model.

In summary, this system demonstrates the emergence of autocatalysis from a self-organizing anabolic reaction net-

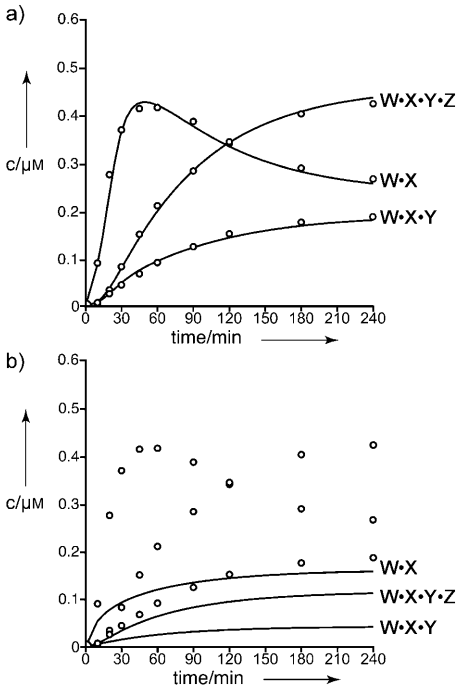


Figure 2. Kinetic modeling of the system. a) Concentrations of observables **W·X**, **W·X·Y**, and **W·X·Y·Z** as a function of time. ○: experimental data obtained from denaturing polyacrylamide gel electrophoresis (PAGE) and autoradiography;[25] ——: best fit of the model to the kinetic data using the program SimFit. b) Result of removing the autocatalytic rate constants from the model after fitting.

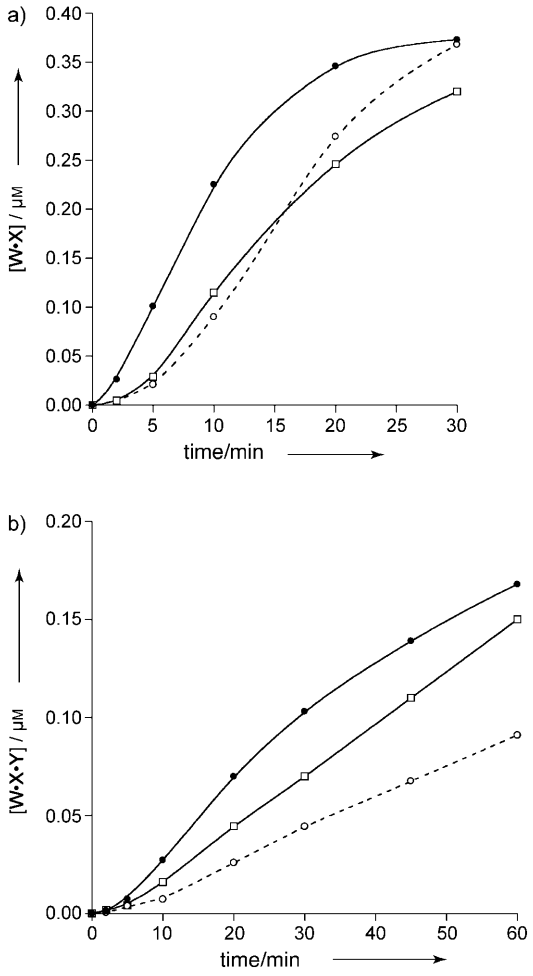


Figure 3. Time course showing the effect of **W·X** and **W·X·Y** on their own production. All reactions initially contained 1 μM each **W**, **X**, **Y**, **Z**. a) Concentration of **W·X** as a function of time with no products added (----, ○) or in the presence of either 1 μM **W·X** (□) or 1 μM **W·X·Y·Z** (●). b) Concentration of **W·X·Y** as a function of time with no products added (----, ○) or in the presence of either of 1 μM **W·X·Y** (□) or 1 μM **W·X·Y·Z** (●). Curves are for visual aid only.

work involving four individually noncatalytic RNA molecules. Despite the simplifications made, the parameters found through kinetic modeling are in qualitative agreement with the results from our studies of individual initial-rate measurements of subsystems. The model supports our basic claim that autocatalysis is a dominant component of the reaction network. This is true despite a strong non-autocatalytic route, which is necessary for the initial formation of the autocatalytic species. Although much of the autocatalysis by **W·X·Y·Z** is nonspecific, the incremental increase in autocatalysis by larger molecules provides a way for the system to bootstrap itself toward greater complexity. Therefore, the entire network could be viewed as an "individual", in which specific sequences are required for collaboration to produce an autocatalytic network. This system provides an example of predicted missing links[9] between the first self-replicating nucleic acid templates and a more complex RNA world, in line with Eigen's demarcation. Future work will examine whether collaborating molecules can display a selective advantage through autocatalysis and if a mutated population derived from these sequences can evolve over time.

Experimental Section

RNA fragments were synthesized by in vitro transcription from synthetic DNA templates. Reactions were initiated by the addition of 30 mM 3-[4-(2-hydroxyethyl)-1-piperazinyl]propanesulfonic acid buffer (EPPS, pH 7.5) containing 100 mM $MgCl_2$ and incubation at 48 °C. Samples (5 μL) were removed at desired time points, immediately quenched with 125 mM ethylenediaminetetraacetic acid (EDTA) and urea-containing loading dye, and then separated on 8% polyacrylamide gels containing 8 M urea. Product molecules used in seeding experiments were obtained by gel excision of actual reaction products and subsequent reverse-transcription PCR and transcription. The **W**-containing fragments were labeled with ^{32}P at the 5′ terminus prior to reaction. Product concentrations were calculated from a percent of reacted material in each lane. Initial rates were determined from time points at which the extent of reaction was less than 10%. Theoretical kinetic data was calculated by integration of the set of stiff rate equations derived from our reaction model using SimFit.[27] Nonlinear optimization was achieved through a simplex method followed by Newton–Raphson optimization.

Received: May 9, 2008
Revised: June 27, 2008
Published online: September 9, 2009

Keywords: autocatalysis · origin of life · ribozymes · RNA · self-assembly

[1] E. Szathmáry, *Philos. Trans. R. Soc. Lond. B Biol Sci.* **2006**, *361*, 1761–1766.
[2] T. Gánti, *The Principles of Life*, Oxford University Press, Oxford, **2003**.
[3] S. A. Kauffman, *The Origins of Order*, Oxford University Press, New York, **1993**.
[4] P. L. Luisi, *The Emergence of Life*, New York, Cambridge University Press, **2006**.
[5] L. E. Orgel, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 12503–12507.
[6] A. Eschenmoser, *Tetrahedron* **2007**, *63*, 12821–12844.
[7] P. Yin, H. M. T. Choi, C. R. Calvert, N. A. Pierce, *Nature* **2008**, *451*, 318–322.
[8] D. Y. Zhang, A. T. Turberfield, B. Yurke, E. Winfree, *Science* **2007**, *318*, 1121–1125.
[9] M. Levy, A. D. Ellington, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 6416–6421.
[10] W. S. Zielinski, L. E. Orgel, *Nature* **1987**, *327*, 346–347.
[11] G. von Kiedrowski, *Angew. Chem.* **1986**, *98*, 932–934; *Angew. Chem. Int. Ed. Engl.* **1986**, *25*, 932–935.
[12] T. Achilles, G. von Kiedrowski, *Angew. Chem.* **1993**, *105*, 1225–1228; *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 1198–1201.
[13] D. Sievers, G. von Kiedrowski, *Nature* **1994**, *369*, 221–224.
[14] N. Paul, G. F. Joyce, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12733–12740. Reported an autocatalytic efficiency of $3 \times 10^8\,M^{-1}$.
[15] D. H. Lee, J. R. Granja, J. A. Martinez, K. Severin, M. R. Ghadiri, *Nature* **1996**, *382*, 525–528.
[16] K. Severin, D. H. Lee, J. A. Martinez, M. Vieth, M. R. Ghadiri, *Angew. Chem.* **1998**, *110*, 133–135; *Angew. Chem. Int. Ed.* **1998**, *37*, 126–128.
[17] T. Tjivikua, P. Ballester, J. Rebek, *J. Am. Chem. Soc.* **1990**, *112*, 1249–1250.
[18] M. Eigen, *Naturwissenschaften* **1971**, *58*, 465–528. Quotation marks are Eigen's.

[19] 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.
[20] E. J. Hayden, N. Lehman, *Chem. Biol.* **2006**, *13*, 909 – 918.
[21] W. E. Draper, E. J. Hayden, N. Lehman, *Nucleic Acids Res.* **2008**, *36*, 520 – 531.
[22] G. von Kiedrowski, *Bioorg. Chem. Front.* **1993**, *3*, 113 – 146.
[23] G. von Kiedrowski, B. Wlotzka, J. Helbing, M. Mazten, M. S. Jordan, *Angew. Chem.* **1991**, *103*, 456 – 459; *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 423 – 426.
[24] B. Rangan, P. Masquida, E. Westhof, S. A. Woodson, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1574 – 1579.
[25] These data or discussion can be found in the Supporting Information.
[26] N. Paul, G. F. Joyce, *Curr. Opin. Chem. Biol.* **2004**, *8*, 634 – 639.
[27] I. Stahl, G. von Kiedrowski, *J. Am. Chem. Soc.* **2006**, *128*, 14014 – 14015.
[28] L. Kuo, L. A. Davidson, S. Pico, *Biochim. Biophys. Acta Gene Struct. Expression* **1999**, *1489*, 281 – 292.

