Intercalation-Mediated Ligation

Enzymatic Behavior by Intercalating Molecules in a Template-Directed Ligation Reaction**

*Swapan S. Jain, Frank A. L. Anet, Christopher J. Stahle, and Nicholas V. Hud**

Since the discovery of catalytic RNA two decades ago,[1] much attention has focused on the hypothesis that an early form of life used nucleic acids for both information storage and catalysis before the advent of proteins.[2] However, it is still a mystery how the first nucleic acid polymers assembled and replicated, as these tasks are carried out by protein enzymes in contemporary life. Decades of research have led to the inescapable conclusion that Watson–Crick base pairing alone does not sufficiently stabilize the assembly of mononucleo-

[*] S. S. Jain, C. J. Stahle, Dr. N. V. Hud
School of Chemistry and Biochemistry
Parker H. Petit Institute for Bioengineering and Bioscience
Georgia Institute of Technology, Atlanta, GA 30332 (USA)
Fax: (+1) 404-894-2295
E-mail: hud@chemistry.gatech.edu

Dr. F. A. L. Anet
Department of Chemistry and Biochemistry
University of California, Los Angeles, CA 90095 (USA)

[**] This work was supported by the Georgia Institute of Technology Office of the Vice Provost for Research. We thank Dr. A. Ellington (University of Texas, Austin) for helpful discussions.

  DOI: 10.1002/ange.200353155

tides on a template strand in aqueous solution to allow spontaneous self-replication.[3] Investigations of nonnatural mononucleotide-coupling chemistries and chemical activation have proven more successful than attempts to condense the natural mononucleotide triphosphates on single-stranded DNA or RNA templates.[4] Nevertheless, a prebiotically plausible method to bridge the gap from small molecules to self-replicating RNA-like polymers has not been found. Herein, we report that a small molecule that intercalates the bases of DNA and RNA can increase the template-directed coupling rate of short oligonucleotides by three orders of magnitude. Several of these molecules work together in a cooperative manner to function, in essence, as a concentration-dependent multimolecular "enzyme". These results support the recently made hypothesis that an intercalating molecule could have acted as a "molecular midwife" that facilitated the replication of information-containing polymers before the existence of the RNA world, as well as the replication of RNA itself, at least in the early stages of the RNA world.[5]

We have conducted a series of experiments to test whether intercalation in present-day nucleic acids can facilitate the template-directed synthesis of nucleic acids. Our experimental test system involves suitably modified forms of the short oligonucleotides, $(dT)_3$ and $(dT)_4$ (dT = deoxythymidylate). The chemistry used to couple these oligonucleotides makes use of an iodine atom as a leaving group on 5′-iodo-$(dT)_4$ and leads to formation of a covalent bond with the sulfur atom of 3′-phosphorothioate-$(dT)_3$.[6] A graphical representation of this ligation test system is shown in Figure 1. The intercalator used is a planar tricyclic cationic molecule commonly known as proflavine (Figure 1 b, c), which closely matches the shape of a Watson–Crick base pair (Figure 1 d, e). By labeling the 3′-phosphorothioate-$(dT)_3$ substrate on its 5′ end with a ^{32}P-phosphate group, we are able to follow product formation as a function of intercalator concentration by quantification of the $(dT)_7$ product after polyacrylamide gel electrophoresis. An image of a gel for a set of ligation experiments with increasing concentrations of proflavine is presented in Figure 2. This gel shows that the $(dT)_7$ product is virtually undetectable for reactions containing only the substrates $(dT)_3$ and $(dT)_4$ with the $(dA)_{16}$ template strand (dA = deoxyadenylate). The addition of proflavine to the reaction mixture of $(dT)_3$ and $(dT)_4$ produces a detectable increase in the yield of the $(dT)_7$ ligation product, even in the complete absence of the template. The significance of this result will be discussed later. A far more dramatic increase in the yield of $(dT)_7$ occurs when both the $(dA)_{16}$ template and proflavine are present (Figure 2). Quantification of gel band intensities (Table 1) shows that proflavine catalyzes the ligation rate of $(dT)_3$ and $(dT)_4$ by three orders of magnitude over reactions relying on only the $(dA)_{16}$ template strand to organize the substrates. These results are consistent with proflavine promoting the formation of a $(dT)_3$,$(dT)_4$·$(dA)_{16}$ duplex that acts as a ligation complex in which the reactive ends of the $(dT)_3$ and $(dT)_4$ substrates can meet. The ligation product is a phosphorothioate-linked analogue of $(dT)_7$ (Figure 1 a). The importance of Watson–Crick base pairing in the proflavine-catalyzed ligation reaction is illustrated by the fact that product yield

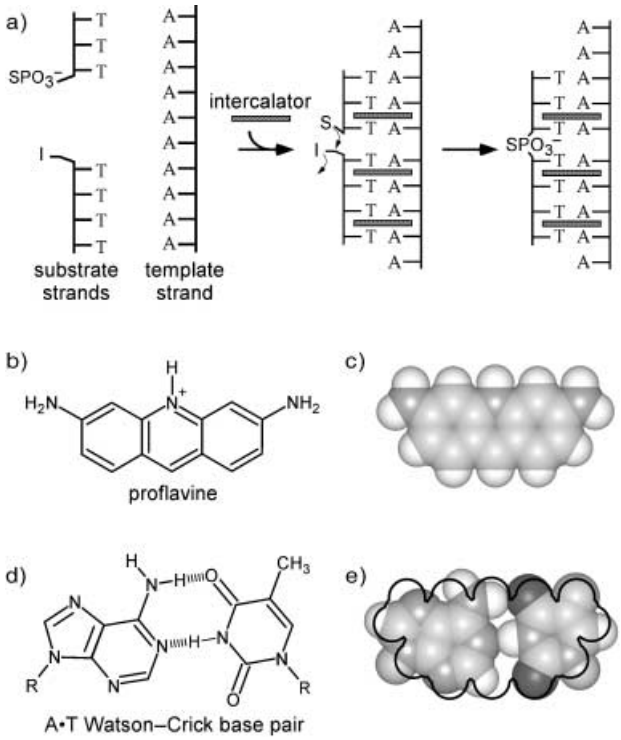

Figure 1. A schematic representation of the test system for investigating intercalation-mediated template-directed synthesis, as well as the applicable molecular structures. a) A template strand in solution with substrate strands. The substrate strands are sufficiently short that the equilibrium amount of substrate strands bound to the template strand is extremely small. The addition of an intercalating molecule to the solution facilitates the formation of a duplex between the template strand and the substrate strands with Watson–Crick complementary sequences. Chemical ligation is used to join the backbones of substrate strands aligned along the template strand. b) and c) The chemical structure and space-filling model of proflavine. d) and e) The chemical structure and space-filling model of the Watson–Crick A·T base pair. A black outline of the proflavine van der Waals surface is superimposed on the space-filling model of the A·T base pair to illustrate the close match between the shapes of these molecular structures.

drops significantly when DNA templates with sequences other than $(dA)_{16}$ are used with the $(dT)_3$ and $(dT)_4$ substrates (Table 1).

A plot of the rate of $(dT)_7$ ligation on a $(dA)_{16}$ template demonstrates that the rate of reaction is enhanced with increasing proflavine concentrations, up to approximately 100 μM proflavine (Figure 3). A least-squares fit of this data by the Hill equation indicates that at least three proflavine molecules bind cooperatively to the substrate and template strands, each with a binding constant of around 60 μM, to create the active ligation complex. According to the nearest-neighbor exclusion principle the bases of nucleic acid duplexes can only bind one intercalating molecule per two base pairs.[7] Thus, the substrates $(dT)_3$ and $(dT)_4$, when forming a duplex with a $(dA)_n$ template strand, would be expected to bind one and two proflavine molecules, respectively (Figure 1 a), for a total of three molecules, which is in agreement with our experimental data.

The 1000-fold increase in the rate of formation of the $(dT)_7$ ligation product in a solution containing 140 μM

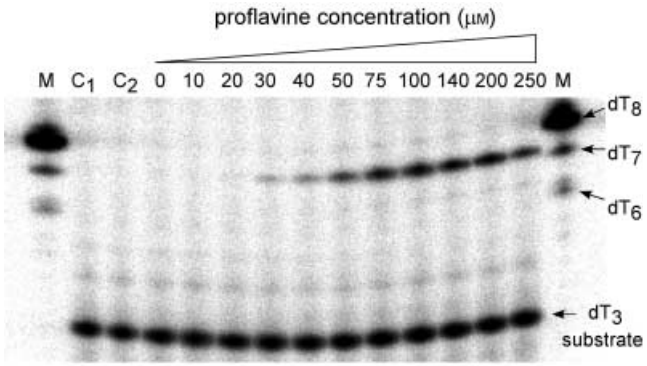

Figure 2. Denaturing polyacrylamide gel after electrophoresis analysis that illustrates the effect of proflavine on the ligation of 3′-phosphorothioate-$(dT)_3$ and 5′-iodo-$(dT)_4$ with $(dA)_{16}$ as a template strand. Lane C1: Only ^{32}P-labeled 3′-phosphorothioate-$(dT)_3$. Lane C2: Substrates ^{32}P-labeled 3′-phosphorothioate-$(dT)_3$ and 5′-iodo-$(dT)_4$. Lanes labeled 0–250: ^{32}P-labeled 3′-phosphorothioate-$(dT)_3$, 5′-iodo-$(dT)_4$, template strand $(dA)_{16}$, and proflavine at a concentration corresponding to the number above the lane, in units of μM. All reaction mixtures were incubated for 24 h at 277 K. Lane M: Molecular-weight marker bands of $(dT)_8$, $(dT)_7$, and $(dT)_6$.

Table 1: Quantitative analysis of ligation test-system results.

Template[a]	Intercalator[b]	Ligation rate[c]	Half max[d]
–	–	< 15	NA
$(dA)_{16}$	–	< 15	NA
$(dA)_{16}$	proflavine	10000 ± 1000	51 μM
$(dAATA)_4$	proflavine	5100 ± 600	ND
$(dN)_{16}$	proflavine	ca. 70	ND
$(dA)_{16}$ (298 K)	proflavine	6300 ± 700	ND
$(rA)_{16}$	proflavine	7100 ± 800	87 μM
$(dA)_{16}$	ethidium	ca. 35	> 500 μM
–	proflavine	ca. 70	ND

[a] Template $(dAATA)_4$ is d(AATAAATAAATAAATA); template $(dN)_{16}$ is d(GATCCGAATTCACGTG), where dG = deoxyguanylate and dC = deoxycytidylate. [b] Intercalator concentrations were 140 μM, where an intercalator is listed. [c] Ligation rates were determined based upon radioactive decay counts from gel bands and have been normalized with respect to the highest ligation rate experiment, which has been scaled to 10000. [d] Half max = the concentration of intercalator at which the rate of product yield is one half of the maximum ligation rate, NA = not applicable, ND = not determined. All reactions contained both the $(dT)_3$ and $(dT)_4$ substrates as described in the text. All experiments were carried out at 277 K, unless otherwise indicated. The ligation reaction time was 24 h for all experiments.

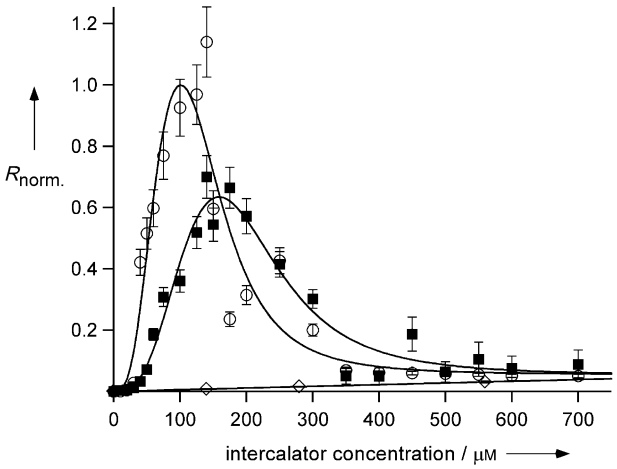

Figure 3. Plots of the relative ligation rates (R) for formation of the $(dT)_7$ product as a function of template strand, intercalator species, and intercalator concentration: ○: template = $(dA)_{16}$ and intercalator = proflavine; ■: template = $(rA)_{16}$ and intercalator = proflavine; ◇: template = $(dA)_{16}$ and intercalator = ethidium. Rates shown have been normalized to the maximum of the fit of the data for proflavine with the DNA template. Substrates and template strands were in concentrations of 1.0 μM for all reactions. The reaction mixtures were incubated for 24 h at 277 K. The error bars show known sources of error only. A few data points, for unknown reasons, show unexpectedly large deviations from the fitted curves; omitting these points does not change the fits appreciably.

proflavine implies that proflavine reduces the overall free-energy barrier for ligation by approximately 3.8 kcal mol^{-1} at 277 K. When the same reaction was carried out at 298 K, the rate of product formation was reduced by a factor of 0.63 with respect to the rate at 277 K (Table 1). Xu and Kool have shown that, for a stable nucleic acid assembly, the rate of the phosphorothioate ligation reaction increases with temperature over this range.[6] Thus, the reduction in the rate of product formation observed in our system with increased temperature must be the result of a reduction in the concentration of the duplex structure containing three intercalated proflavine molecules, a conclusion that is consistent with the expected negative entropy for the formation of such a complex. However, the analysis is complicated by the presence of a second intercalation complex that is evident at still higher proflavine concentrations.

The maximum rate of proflavine-catalyzed $(dT)_7$ ligation on the DNA template $(dA)_{16}$ is achieved at around 100 μM proflavine under the conditions used, but the rate does not remain on a plateau at higher concentrations. Instead, it decreases smoothly after the maximum point to reach a much lower constant rate at approximately 600 μM proflavine (Figure 3). The shape of the curve for concentrations of 0–700 μM proflavine is consistent with four additional proflavine molecules binding to the reaction complex with a weaker binding constant (ca. 160 μM) than the three proflavine molecules that assemble the catalytically active complex. The decrease in the ligation rate at high proflavine concentrations indicates that the assembly bound with more than three proflavine molecules is a much less catalytically active complex than the three-proflavine complex. It is possible that binding the secondary set of proflavine molecules arranges the $(dT)_3$ and $(dT)_4$ oligonucleotides such that their reactive groups are too far away from each other for bond formation, or high proflavine concentrations may induce the $(dA)_{16}$ template to dimerize.[8] In any case, the significant decline in reaction rate upon the binding of more than three proflavine molecules fits a cooperative phenomenon.

In Figure 3 we also present results from proflavine-catalyzed ligation of $(dT)_3$ and $(dT)_4$ on the RNA template $(rA)_{16}$ (rA = adenylate). The overall results are similar to those with the analogous DNA template, except that the curve is shifted to higher proflavine concentrations, a result

indicating that the intercalation complex is somewhat less favorable with the RNA than with the DNA template. This result shows the interplay that exists between a small-molecule intercalator and the backbone structure, even though an intercalator such as proflavine is expected to have minimal direct contact with the backbone (Figure 1e).

Ethidium, a common fluorescent intercalator, was also investigated in our ligation test system. Far less $(dT)_7$ ligation product was observed in comparison to the yield from the same reaction with proflavine (Figure 3). The binding constants of proflavine and ethidium for a DNA duplex are very similar.[9] Thus, the ability for an intercalating molecule to act as a midwife must also depend on the shape of the molecule, rather than simply on its binding constant. Proflavine has three linearly fused aromatic rings, whereas ethidium has its three aromatic rings angularly fused and it also has a pendant phenyl group that is not present in proflavine. This hydrophobic phenyl group would tend to increase the binding of ethidium in an aqueous medium, but it might well be detrimental to the ligation reaction itself.

As noted above, a small but distinct increase in the rate of formation of the $(dT)_7$ ligation product over the background rate was detected in a reaction mixture containing 140 μM proflavine but no $(dA)_{16}$ template strand (Table 1). A small increase in the template-free ligation rate of $(dT)_3$ and $(dT)_4$ by proflavine is of interest because it shows that an intercalator, perhaps through nonspecific stacking interactions with the terminal bases of $(dT)_3$ and $(dT)_4$, can create a small equilibrium amount of a ligation-active complex. This means that DNA- and RNA-like polymers could have been synthesized de novo by intercalators at low rates without the requirement for preexisting templates. Once this occurs, the system could become autocatalytic if complementary Watson–Crick bases were both present as activated monomers, since the spontaneous emergence of template strands would greatly enhance the production of complementary strands in the presence of the proper intercalator.

The plots of $(dT)_7$ ligation rates shown in Figure 3 can also be viewed as plots of the rates of enzyme-catalyzed reactions as functions of the enzyme concentrations. The rate of an enzyme-catalyzed reaction typically increases linearly with enzyme concentration (that is, first order with respect to enzyme concentration). In contrast, the cooperative increase in the rate of formation of the $(dT)_7$ ligation product with proflavine concentration indicates that the three proflavine molecules of the active complex are working together. Thus, the small-molecule proflavine can be viewed as a cooperative, concentration-dependent multimolecular enzyme. This fact has significant implications regarding the possible utility of small planar molecules and the role of intercalation in the early stages of life.[5]

In conclusion, our demonstration that an intercalating molecule can greatly increase the efficiency of a template-directed ligation reaction has important implications for contemporary nucleic acid chemistry, as well as potential implications concerning the mechanism of nucleic acid synthesis in early life. For over thirty years researchers have sought to improve the yield of protein-free template-directed nucleic acid ligation reactions. Past efforts have included careful sequence design, exhaustive exploration of solution conditions, the use of templates with nonnatural backbones, and the development of novel substrate-linkage chemistries.[3] Our results demonstrate that the simple act of adding an intercalating molecule to a ligation reaction can have a huge effect on improving the coupling efficiency. There has also been much speculation concerning the possible role of inorganic surfaces in the origin of life,[10] as the collection of materials on surfaces could serve as a means to concentrate and spatially organize the molecular components necessary for life. However, as we have illustrated here, a relatively simple molecule with a flat surface could have accomplished these tasks in a much more versatile way than a solid macroscopic surface. Molecules that intercalate DNA and RNA duplexes do so in part because their shapes match those of the Watson–Crick base pairs. In the same way, molecules that could have acted as molecular midwives in the assembly and replication of the first informational polymers may have played a significant role in selecting the nucleotide bases as a consequence of their ability to form structures that matched the structure of the midwife's surface.

Experimental Section

Sample preparation: Substrate oligodeoxynucleotides were synthesized on an automated synthesizer by using the phosphoramidite coupling chemistry. Synthesis of 3′-phosphorothioate-$(dT)_3$ was accomplished by using a 3′-phosphate controlled-pore glass support (Glen Research), with the oxidation reagent normally added during the first nucleotide coupling cycle replaced by a sulfurizing reagent (Glen Research). The 5′-iodo-$(dT)_4$ substrate oligonucleotide was synthesized by using the commercially available 5′-iodothymidine phosphoramidite reagent (Glen Research). Following deprotection, substrate oligonucleotides were purified by reversed-phase HPLC on a C_{18} semipreparative column. Template-strand oligonucleotides were purified on a 1-m G-15 column (Pharmacia). Stock solutions of oligonucleotides were prepared by resuspending freeze-dried purified samples in deionized H_2O. Oligonucleotide concentrations were determined spectrophotometrically.

The 3′-phosphorothioate-$(dT)_3$ substrate was radioactively labeled with a ^{32}P-phosphate group at the 5′ end by diluting 3′-phosphorothioate-$(dT)_3$ from a stock solution to a concentration of 50 μM in T4 polynucleotide kinase buffer (100 μL; New England Biolabs). T4 polynucleotide kinase (30 units; New England Biolabs) was added to the buffered DNA solution. γ-^{32}P-adenosine triphosphate (3 μL; 100 μCi μL^{-1}; ICN) was then added to the solution and the mixture was incubated at 37 °C for 30 min.

Proflavine (hemisulfate salt) was purchased from Sigma. Stock solutions of proflavine were prepared by dissolving the solid proflavine salt in deionized H_2O. Stock-solution concentrations were determined spectrophotometrically by using the extinction coefficient $\varepsilon_{444} = 38\,900\,\text{M}^{-1}\,\text{cm}^{-1}$.

Ligation reactions: Reactions were carried out in 100-μL volumes in a solution containing 10 mM tris(hydroxymethyl)aminomethane buffer (pH 8.2), 10 mM NaCl, and 100 mM 2-thioethanol. The substrate 5′-iodo-$(dT)_4$, the substrate ^{32}P-labeled 3′-phosphorothioate-$(dT)_3$, and the template $(dA)_{16}$ were each added to the reaction buffer to a final concentration of 1.0 μM. The presence of 2-thioethanol in the reaction buffer was necessary to prevent dimerization of the 3′-phosphorothioate-$(dT)_3$ substrate. Ligation reactions were stopped by plunging the reaction test tubes into liquid nitrogen and freeze-drying.

Product analysis: Freeze-dried reaction samples were resuspended in 8 M urea solution (10 μL) and loaded onto a denaturing

30% polyacrylamide gel (acrylamide:bisacrylamide (19:1)). Gels were subjected to electrophoresis at a constant power of 65 W for 6 h. The relative yield of the $(dT)_7$ product for each reaction was determined by imaging the gel on a Fuji Phosphor Imager and quantifying for each lane the integrated intensity of gel bands that corresponded to the $(dT)_7$ ligation product by using the software package Image Gauge V3.12. Background correction was accomplished by subtracting the integrated intensity from all reaction samples of an area in a control lane run with only ^{32}P-labeled $(dT)_3$.

Received: October 27, 2003 [Z53155]
Published Online: February 16, 2004

Keywords: chemical ligation · intercalations · nucleic acids · self-replication · template synthesis

[1] a) C. Guerrier-Takada, K. Gardiner, T. Marsh, N. Pace, S. Altman, *Cell* **1983**, *35*, 849–857; b) K. Kruger, P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling, T. R. Cech, *Cell* **1982**, *31*, 147–157.

[2] *The RNA World: The Nature of Modern RNA Suggests a Prebiotic RNA World, 2nd ed.* (Eds.: R. Gesteland, J. F. Atkins), Cold Spring Harbor Laboratory Press, New York, **1999**.

[3] a) G. F. Joyce, *Cold Spring Harbor Symp. Quant. Biol.* **1987**, *52*, 41–51; b) G. F. Joyce, L. E. Orgel in *The RNA World: The Nature of Modern RNA Suggests a Prebiotic RNA World, 2nd ed.* (Eds.: R. F. Gesteland, J. F. Atkins), Cold Spring Harbor Laboratory Press, New York, **1999**.

[4] a) T. Inoue, L. E. Orgel, *J. Mol. Biol.* **1982**, *162*, 201–217; b) E. Kanaya, H. Yanagawa, *Biochemistry* **1986**, *25*, 7423–7430; c) J. P. Ferris, C.-H. Huang, W. J. Hagan, Jr., *Nucleosides Nucleotides* **1989**, *8*, 407–414; d) J. T. Goodwin, D. G. Lynn, *J. Am. Chem. Soc.* **1992**, *114*, 9197–9198; e) Y. Gat, D. G. Lynn, *Biopolymers* **1998**, *48*, 19–28; f) Z. Y. Li, Z. Y. J. Zhang, R. Knipe, D. G. Lynn, *J. Am. Chem. Soc.* **2002**, *124*, 746–747; g) Z. T. Gartner, R. Grubina, C. T. Calderone, D. R. Liu, *Angew. Chem.* **2003**, *115*, 1408–1413; *Angew. Chem. Int. Ed.* **2003**, *42*, 1370–1375.

[5] N. V. Hud, F. A. L. Anet, *J. Theor. Biol.* **2000**, *205*, 543–562.

[6] Y. Z. Xu, E. T. Kool, *Nucleic Acids Res.* **1999**, *27*, 875–881.

[7] a) L. S. Lerman, *J. Mol. Biol.* **1961**, *3*, 18–30; b) P. H. von Hippel, J. D. McGhee, *Annu. Rev. Biochem.* **1972**, *41*, 231–300.

[8] a) E. Westhof, M. Sundaralingam, *Proc. Natl. Acad. Sci. USA* **1980**, *77*, 1852–1856; b) M. Polak, N. V. Hud, *Nucleic Acids Res.* **2002**, *30*, 983–992; c) S. S. Jain, M. Polak, N. V. Hud, *Nucleic Acids Res.* **2003**, *31*, 4608–4615.

[9] a) J. B. LePecq, C. Paoletti, *J. Mol. Biol.* **1967**, *27*, 87–106; b) G. Lober, *Photochem. Photobiol.* **1968**, *8*, 23–30; c) X. G. Qu, J. B. Chaires, *J. Am. Chem. Soc.* **2001**, *123*, 1–7.

[10] a) J. D. Bernal, *The Physical Basis of Life*, Routledge & Kegan Paul, London, **1951**; b) A. G. Cairns-Smith, *Genetic Takeover: And the Mineral Origins of Life*, Cambridge University Press, Cambridge, **1982**; c) J. P. Ferris, A. R. Hill, R. H. Liu, L. E. Orgel, *Nature* **1996**, *381*, 59–61; d) A. W. Schwartz, *Chem. Biol.* **1996**, *3*, 515–518.

