

## Catalyzed Relaxation of a Metastable DNA Fuel

Georg Seelig,<sup>†</sup> Bernard Yurke,<sup>‡</sup> and Erik Winfree<sup>\*,§,||</sup>

Contribution from the Departments of Computer Science, Applied Physics, and Computation and Neural Systems, California Institute of Technology, Pasadena, California 91125, and Bell Laboratories, Lucent Technologies, 600 Mountain Avenue, Murray Hill, New Jersey 07974

Received May 22, 2006; E-mail: winfree@caltech.edu

**Abstract:** Practically all of life's molecular processes, from chemical synthesis to replication, involve enzymes that carry out their functions through the catalytic transformation of metastable fuels into waste products. Catalytic control of reaction rates will prove to be as useful and ubiquitous in nucleic-acid-based engineering as it is in biology. Here we report a metastable DNA "fuel" and a corresponding DNA "catalyst" that improve upon the original hybridization-based catalyst system (Turberfield et al. *Phys. Rev. Lett.* **90**, 118102-1–118102-4) by more than 2 orders of magnitude. This is achieved by identifying and purifying a fuel with a kinetically trapped metastable configuration consisting of a "kissing loop" stabilized by flanking helical domains; the catalyst strand acts by opening a helical domain and allowing the complex to relax to its ground state by a multistep pathway. The improved fuel/catalyst system shows a roughly 5000-fold acceleration of the uncatalyzed reaction, with each catalyst molecule capable of turning over in excess of 40 substrates. With  $k_{\text{cat}}/K_M \approx 10^7/\text{M}/\text{min}$ , comparable to many protein enzymes and ribozymes, this fuel system becomes a viable component enabling future DNA-based synthetic molecular machines and logic circuits. As an example, we designed and characterized a signal amplifier based on the fuel–catalyst system. The amplifier uses a single strand of DNA as input and releases a second strand with unrelated sequence as output. A single input strand can catalytically trigger the release of more than 10 output strands.

It is now accepted that nucleic acids play a more active role in many biological processes than they were assigned in the central dogma<sup>2–6</sup> and may have been the dominant macromolecule in primitive organisms.<sup>7</sup> This re-evaluation of the part played by nucleic acids in biology has gone hand in hand with the development of nucleic acid based systems for nano- or biotechnology applications. The relative predictability of Watson Crick base pairing and thus of inter- and intramolecular interactions allows one to design self-assembling molecules with desired folding pathways and conformational changes. This makes nucleic acids promising materials for the construction of nanoscale devices and structures. For example, DNA nanotechnology has developed methods for the self-assembly of complex two- and three-dimensional structures,<sup>8–11</sup> for control-

lable nanoscale mechanical motion,<sup>12–15</sup> for molecular walking motors,<sup>16–21</sup> and for catalysts, logic gates, and circuits using molecular input and output.<sup>1,22</sup> Similarly, nucleic acids have also proved suitable for engineering complex reactions for biological samples, both in vivo and in vitro, with applications to DNA chips,<sup>23,24</sup> gene expression,<sup>25,26</sup> and potentially even "smart" therapeutics.<sup>22,27</sup>

It would be of great interest to develop DNA-based systems that operate autonomously and control reactions to occur at specific times and under specific conditions. Biological systems

<sup>†</sup> Department of Applied Physics, California Institute of Technology.

<sup>‡</sup> Bell Laboratories, Lucent Technologies.

<sup>§</sup> Department of Computer Science, California Institute of Technology.

<sup>||</sup> Computation and Neural Systems Option, California Institute of Technology.

- (1) Turberfield, A. J.; Mitchell, J. C.; Yurke, B.; Mills, A. P., Jr.; Blakey, M. I.; Simmel, F. C. *Phys. Rev. Lett.* **2003**, *90*, 118102-1–118102-4.
- (2) Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R. *Cell* **1982**, *31*, 147–157.
- (3) Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. *Cell* **1983**, *35*, 849–857.
- (4) Lee, R. C.; Feinbaum, R. L.; Ambros, V. *Cell* **1993**, *75*, 843–854.
- (5) Bartel, D. P. *Cell* **2004**, *116*, 281–297.
- (6) Mandal, M.; Breaker, R. R. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 451–463.
- (7) Gilbert, W. *Nature* **1986**, *319*, 618–618.
- (8) Seeman, N. C. *J. Theor. Biol.* **1982**, *99*, 237–247.
- (9) Winfree, E.; Liu, F.; Wenzler, L. A.; Seeman, N. C. *Nature* **1998**, *394*, 539–544.
- (10) Shih, W. M.; Quispe, J. D.; Joyce, G. F. *Nature* **2004**, *427*, 618–621.
- (11) Rothmund P. W. K. *Nature* **2006**, *440*, 297–302.

- (12) Yurke, B.; Turberfield, A. J.; Mills, A. P., Jr.; Simmel, F. C.; Neumann, J. L. *Nature* **2000**, *406*, 605–608.
- (13) Li, J. J.; Tan, W. *Nano Lett.* **2002**, *2*, 315–318.
- (14) Feng, L.; Park, S. H.; Reif, J. H.; Yan, H. *Angew. Chem., Int. Ed.* **2003**, *42*, 4342–4346.
- (15) Chen, Y.; Wang, M. S.; Mao, C. D. *Angew. Chem., Int. Ed.* **2004**, *43*, 3554–3557.
- (16) Shin, J.-S.; Pierce, N. A. *J. Am. Chem. Soc.* **2004**, *126*, 10834–10835.
- (17) Sherman, W. B.; Seeman, N. C. *Nano Lett.* **2004**, *4*, 1203–1207.
- (18) Yin, P.; Yan, H.; Daniell, X. G.; Turberfield, A. J.; Reif, J. H. *Angew. Chem., Int. Ed.* **2004**, *43*, 4906–4911.
- (19) Pei, R.; Taylor, S. K.; Rudchenko, S.; Mitchell, T. E.; Stefanovic, D.; Stojanovic, M. N. Behavior of Polycatalytic Assemblies in Substrate-Displaying Matrix. *J. Am. Chem. Soc.*, in press.
- (20) Bath, J.; Green, S. J.; Turberfield, A. J. *Angew. Chem., Int. Ed.*, **2005**, *44*, 4358–4361.
- (21) Tian, Y.; He, Y.; Chen, Y.; Yin, P.; Mao, C. D. *Angew. Chem., Int. Ed.* **2005**, *44*, 4355–4358.
- (22) Stojanovic, M. N.; Mitchell, T. E.; Stefanovic, D. *J. Am. Chem. Soc.* **2002**, *124*, 3555–3561.
- (23) Breaker, R. R. *Curr. Opin. Biotechnol.* **2002**, *13*, 31–39.
- (24) Su, X.; Smith, L. M. *Nucleic Acids Res.* **2004**, *32*, 3115–3123.
- (25) Isaacs, F. J.; Dwyer, D. J.; Ding, C.; Pervouchine, D. D.; Cantor, C. R.; Collins, J. J. *Nat. Biotechnol.* **2004**, *22*, 841–847.
- (26) Bayer, T. S.; Smolke, C. D. *Nat. Biotechnol.* **2005**, *23*, 337–343.
- (27) Benenson, Y.; Gil, B.; Ben-Dor, U.; Adar, R.; Shapiro, E. *Nature* **2004**, *429*, 423–429.

depend on complex biochemical control circuits to make cellular processes autonomous. These circuits rely on (1) the availability of high-energy fuels (e.g., ATP) that exist in a metastable state until triggered to decay into low-energy waste products (e.g., ADP and Pi) and (2) a variety of catalytic proteins (e.g., CheA and CheY) that detect, combine, and amplify signals to control much larger molecular events in an environmentally sensitive way (e.g., flagellar movement). In a RNA world, or in nucleic-acid engineering, what might be the nucleic-acid equivalents of these fuel and catalyst molecules?

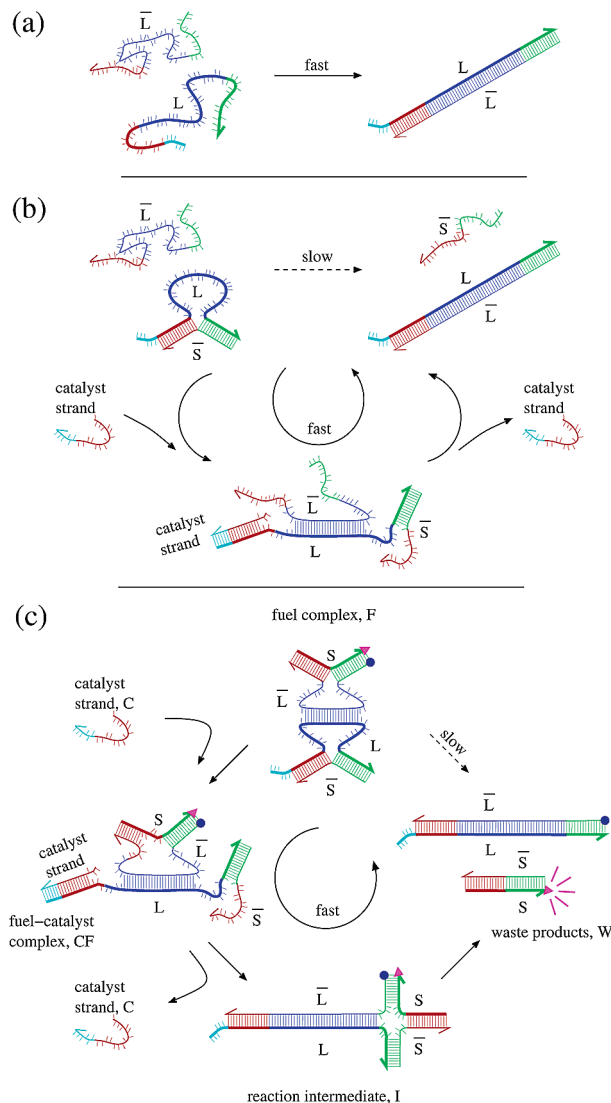
In the nucleic-acid devices that have been constructed to date, energy is primarily provided in one of two forms: either as covalent bonds in the polymer backbone that are broken by the activity of a (deoxy)ribozyme or as metastable single-stranded domains that would thermodynamically prefer to be in a low-energy double-helical state. The relative merits of these two approaches are not yet clear. To date, the latter “hybridization-based” approach has exhibited greater versatility for being incorporated into complex nanomachines but stores energy in larger molecular complexes and, most importantly, has not yet produced catalysts with rate enhancements comparable to proteins or ribozymes.

Here, we present results on a new hybridization-based substrate–catalyst system that improves previous systems by over 2 orders of magnitude, achieving catalysis that is by some measures comparable to ribozyme and protein systems. The system described here is a direct descendant of the original hybridization-based catalyst<sup>1</sup> (Figure 1a,b). Our improved results derive from increased understanding of how to engineer metastable conformations and reaction pathways. In the new catalyst–substrate system, the substrate (Figure 1c) is a four-stranded complex with a half-time for the spontaneous decay of roughly 20 days. This compares favorably with the decay half-time of roughly an hour measured for the original system. The enhanced stability may be due to a topological kinetic trap.<sup>28–30</sup> The putative reaction mechanism for catalysis is more complex, containing both three-strand and four-strand branch migration steps, but still gives rise to a catalyzed half-time of only a few minutes. A simplified model is developed that reproduces reaction kinetics under 80 experimental conditions for catalyst and fuel concentrations. Although our data implies that there is some poisoning of the catalyst, it also implies that a single-catalyst molecule can turnover at least 40 fuel substrates.

In light of our results, we also present a revised explanation for the original catalyst system.<sup>1</sup> The control over the kinetics of nucleic acid conformational changes demonstrated here is expected to find immediate application for the design of a variety of autonomous molecular systems, including molecular motors, chemical sensors, and chemical logic gates. As a specific example, we describe a modified fuel complex that can serve as an amplifier in a chemical signaling cascade.

## Materials and Methods

**System Specifications.** The sequences for strands  $L$ ,  $\bar{L}$ ,  $S$ , and  $\bar{S}$  are the same as in previous work.<sup>1</sup> Strand  $L = \tau\alpha\beta\gamma$  is a 76mer,  $\bar{L} = \bar{\gamma}\bar{\beta}\bar{\alpha}$  is a 70mer, and both  $S = \alpha\gamma$  and  $\bar{S} = \bar{\gamma}\bar{\alpha}$  are 30mers. The catalyst



**Figure 1.** Reaction schematics. The 3'-end of each strand is indicated by an arrow. Thin and thick line segments of the same color are complementary to each other (red, 15 bp; blue, 40 bp; green, 15 bp; cyan, 6 bp). (a) Hybridization reaction between two complementary strands  $L$  and  $\bar{L}$ . (b) Kinetic control over a DNA hybridization reaction, as previously proposed.<sup>1</sup> The 6 nt toe-hold (cyan) greatly enhances the rate at which the catalyst strand binds and partially displaces  $\bar{S}$  by three-strand branch migration. (c) Putative structures and reaction pathway for the catalyzed decay of a metastable fuel complex. Note that all base-paired domains must be double-helical, but by construction, the kissing loops of the fuel complex cannot be topologically linked. Fluorophore and quencher positions are indicated.

strand  $C = \bar{\alpha}\bar{\tau}$  has a length of 21 nucleotides. Greek letters identify specific subsequences, and bars indicate complementarity (i.e.,  $\alpha$  is complementary to  $\bar{\alpha}$  and  $S$  is complementary to  $\bar{S}$ ). Sequences for the toe-hold  $\tau$  (cyan in all figures), the loop region  $\beta$  (blue), and the arms  $\alpha$  (red) and  $\gamma$  (green) of the loop complexes are listed in Table 1. Writing strand sequences in terms of recurring subsequences emphasizes that all strands are constructed from the same limited set of subsequences.

Apart from the six nucleotide toe-hold,  $L$  and  $\bar{L}$  are complementary, as are  $S$  and  $\bar{S}$ . The catalyst strand  $C$  is complementary to the first 21 nucleotides of  $L$  starting from the 5'-end, including the toe-hold. Strands  $L$  and  $\bar{S}$  together form the loop complex  $L\bar{S}$ , while  $S$  and  $\bar{L}$  form the loop complex  $S\bar{L}$ . The two double-stranded segments in a loop complex are each 15 base pairs (bps) long, which ensures that the complexes form reliably and are stable at room temperature. While  $S$  and  $\bar{S}$  are

(28) Seeman, N. C. *Angew. Chem., Int. Ed.* **1998**, *37*, 3220–3238.

(29) Brunel, C.; Marquet, R.; Romby, P.; Ehresmann, C. *Biochimie* **2002**, *84*, 925–944.

(30) Bois, J. S.; Venkataraman, S.; Choi, H. M. T.; Spakowitz, A. J.; Wang, Z. G.; Pierce, N. A. *Nucleic Acids Res.* **2005**, *33*, 4090–4095.

**Table 1.** Strands and Sequences

fuel		amplifier	fuel for melt
$L = \tau\alpha\beta\gamma$		$A2 = \omega\beta\gamma$	$L1 = \delta\beta\epsilon$
$\bar{L} = \bar{\gamma}\bar{\beta}\bar{\alpha}$		$A3 = \bar{\Gamma}\bar{\beta}\bar{\Omega}$	$L2 = \bar{\eta}\bar{\beta}\bar{\zeta}$
$S = \alpha\gamma$		$A4 = \Omega\gamma\theta$	$S1 = \bar{\epsilon}\bar{\delta}$
$\bar{S} = \bar{\gamma}\bar{\alpha}$		$A1 = \Theta\bar{\gamma}\bar{\omega}\bar{\sigma}$	$S2 = \zeta\eta$
no output		Aout = $\Theta$	no output
$C = \bar{\alpha}\bar{\tau}$		mir-143 = $\sigma\omega$	no catalyst

name	use	sequence
$\tau$	fuel complex	CTGGAA (cyan, Figures 1 and 2)
$\sigma$	amplifier	TGAGAT (dark cyan, Figure 6)
$\alpha$	fuel complex	TCGTCGTTTACGGTC (red, Figures 1 and 2)
$\omega$	amplifier	GAAGCACTGTAGCTCA (dark red, Figure 6)
$\Omega$	amplifier	TGAAGCACTGTAGCTCA (dark red, Figure 6)
$\beta$	fuel complex	CACACAGTAGATCAGAATTGGCACGTTCCGCTCGCTAGGTT (blue, Figures 1, 2, and 6)
$\gamma$	fuel complex	GAAGTCACCCTCATT (green, Figures 1, 2, and 6)
$\Gamma$	amplifier	GAAGTCACCCTCATTGAT (green, Figure 6)
$\theta$	amplifier	GATGAATTGGAG (black, Figure 6)
$\Theta$	amplifier	GATGAATTGGAGGTG (black, Figure 6)
$\delta$	melt experiments	ACTGCTACACCAGCGACTTCCGTACTCAAC
$\epsilon$	melt experiments	TGTTCTGTCTCGGTCAGAGGTGTCCTACG
$\eta$	melt experiments	AGAGTGTGGACTCACGGACGTTGGTCACGA
$\zeta$	melt experiments	GCACTGGAAGCATCCGACGAAGGTAGCTCT

fully base-paired when they are part of their respective loop complexes, the central 40 nucleotides (sections  $\beta$  and  $\bar{\beta}$ , respectively) of  $L$  and  $\bar{L}$  remain unpaired (but probably have a secondary structure).

The sequences chosen for the amplifier are largely based on the sequences used in the basic catalyst experiment. However, the catalyst sequence is different (it is the DNA equivalent of the micro RNA mmu-mir-143,<sup>32</sup> 22 nt), and the sequences of the fuel strands were adjusted accordingly. Also, the amplifier contains an additional strand (the output strand), and the fuel strands were extended to accommodate this modification. The catalyst sequence is mir-143 =  $\sigma\omega$ , where  $\sigma$  is six nucleotides long and binds the toe-hold. Sequences  $\sigma$  (dark cyan) and  $\omega$  (dark red) replace  $\bar{\tau}$  and  $\bar{\alpha}$  used in the fuel and catalyst strands. The amplifier strands are  $A1 = \Theta\bar{\gamma}\bar{\omega}\bar{\sigma}$  (based on  $\bar{S}$ ),  $A2 = \omega\beta\gamma$  (based on  $L$ ),  $A3 = \bar{\Gamma}\bar{\beta}\bar{\Omega}$  (based on  $\bar{L}$ ),  $A4 = \Omega\gamma\theta$  (based on  $S$ ), and Aout =  $\Theta$  (black). Note that the toe-hold position was changed and that the toe-hold is attached to the 3'-end of A1. Here,  $\Omega$  is the same as  $\omega$ , except for a one-nucleotide extension at the 5'-end,  $\Gamma$  is the same as  $\gamma$ , with a 3-nucleotide extension at the 5'-end, and  $\theta$  is extended by three nucleotides at the 3'-end to obtain  $\Theta$ . Extensions were introduced to increase the stability of the amplifier complex against spontaneous decay. Strands A1, A2, and Aout form one loop complex (including the bound output), while A3 and A4 form a second loop complex.

For melts, modified loop complexes were designed that cannot undergo branch migration. Loop sequences ( $\beta$  and  $\bar{\beta}$ ) were left unchanged, but the sequences of the arms were modified. The sequences in all four arms were chosen to be unrelated but to have similar melting temperatures. In addition, the length of the arms was doubled from 15 to 30 bps to increase the temperature at which melting of the arms occurs and to better separate this transition from the melting of the kissing loop interaction. The sequences used, are  $L1 = \delta\beta\epsilon$ ,  $S1 = \bar{\epsilon}\bar{\delta}$ ,  $L2 = \bar{\eta}\bar{\beta}\bar{\zeta}$  and  $S2 = \zeta\eta$ . The strands form the two loop complexes  $L1S1$  and  $L2S2$ . Because the short strands  $S1$  and  $S2$  are not complementary, loop complexes  $L1S1$  and  $L2S2$  cannot decay into double-stranded waste products. Instead, upon mixing, all loop complexes will form kissing-loop structures similar to the original fuel complex.

DNA oligos were synthesized and purified by Integrated DNA Technologies, Coralville, IA. DNA stock solutions were prepared at a concentration of 50  $\mu\text{M}$  in ultrapure water. Concentrations were

determined from the measured absorbance at 260 nm and the calculated extinction coefficient (calculation based on nearest-neighbor model parameters<sup>31</sup>).

Dye labels were used to follow reaction kinetics: strand  $S$  was labeled with the fluorophore carboxytetramethylrhodamine (TAMRA) at the 3'-end, and a dark quencher (Iowa Black RQ) was attached to the 5'-end of  $\bar{L}$ . The positions of the labels within the loop are indicated in Figure 3b. Fluorescence is quenched when loop complex  $S\bar{L}$  is formed and unquenched when waste product  $S\bar{S}$  is formed.

To measure reaction kinetics of the amplifier system, the output strand Aout was labeled with a fluorophore at the 3'-end and A1 was labeled with a quencher at the 5'-end. Fluorescence is quenched when the output strand is bound to the amplifier complex and becomes unquenched when the output strand is released.

For detection of the catalyst-bound state a dye-labeled catalyst strand was used. The catalyst was labeled with Cy5 at the 3'-end with a three-base spacer (sequence AAA) inserted between the catalyst sequence  $\bar{\alpha}\bar{\tau}$  and the dye. The spacer was introduced to minimize interactions of the dye with the fuel strand, because such interactions could potentially hinder catalyst release.

TAE/Mg<sup>++</sup> buffer (0.04 M Tris acetate, 1 mM EDTA, 12.5 mM Mg acetate, pH 8.3) was used for all reactions including loop formation, fluorescence experiments, formation gels, gel elutions for purification, and melts.

**Gel Electrophoresis.** Nondenaturing polyacrylamid gels (acrylamide-bis 19:1, 8%, 2–3 h at 15 V/cm, and 4 °C) were used to verify structure formation and to analyze reaction mixtures. For imaging, gels were stained with SybrGold (Molecular Probes, Eugene, OR) for 35 min, excited at 488 nm, and imaged with 530 band-pass filter on a Bio-Rad (Hercules, CA) Molecular Imager FX Pro Plus.

**Fuel Formation and Purification.** Loop complexes were formed in a slow anneal, where the reaction mixture is heated to 90 °C and then slowly (1 °C/min) cooled to room temperature. The two complexes  $L\bar{S}$  and  $S\bar{L}$  were formed in separate reactions.

For purification, loop complexes were prepared at 20  $\mu\text{M}$  each, mixed stoichiometrically, and left to react overnight at room temperature. The reaction mix was then run on a nondenaturing gel, as specified above.

(31) Puglisi, J. D.; Tinoco, I., Jr. In *Methods in Enzymology*; Dahlberg, J. E., Abelson, J. N., Eds.; Academic Press: San Diego, 1989; Vol. 180, pp 304–325.

However, to maximize yield in the subsequent purification procedure, the amount of reactants loaded per well was increased roughly 100-fold over the amount used in a formation gel. The fuel complex formed from the two loop complexes was then purified from the gel. Gel elution (4–5 h at 5 V/cm and 4 °C) was done using the Elutrap Electroelution System (Schleicher and Schuell Bioscience, Dassel, Germany).

After purification, the concentration of the fuel complex was determined by comparing fluorescence and absorbance of test samples prepared from the purified fuel complex to known standards. For this, 50  $\mu$ L of the sample recovered after purification was annealed from 90 °C to room temperature (1 °C/min). Heating induces the metastable fuel complex to decay into the double-stranded waste products  $\bar{L}\bar{L}$  and  $\bar{S}\bar{S}$ . The concentration of fluorescent waste product  $\bar{S}\bar{S}$  was determined by comparing the sample's fluorescence to that of a series of reference samples containing  $\bar{S}\bar{S}$  at known concentrations. The concentration of  $\bar{L}\bar{L}$  and  $\bar{S}\bar{S}$  in the test sample was also determined independently by measuring UV absorption at 260 nm and comparing the result to a series of samples containing both  $\bar{L}\bar{L}$  and  $\bar{S}\bar{S}$  at known concentrations. The results for the concentration obtained from fluorescence and UV absorption measurements typically agreed to within twenty percent.

Preparation of the amplifier proceeds similarly to the preparation of the fuel complex: the two loop complexes A1A2Aout and A3A4 are formed independently, mixed, and, after overnight incubation, the band corresponding to the amplifier is purified from a gel.

**Detection of Impurity State.** Purified fuel complex was incubated with substoichiometric amounts of Cy5-labeled catalyst strand for at least 24 h at room temperature. The amount of catalyst strand ranged from 2 to 5% of the fuel concentration. The reaction mixture was loaded into a non-denaturing gel. Because the amount of material in the impurity band was expected to be not more than 1–2% of the amount of material in the fuel or waste product bands, about 10–100 times more material was used than in a standard formation gel. The gel was then stained with SybrGold and imaged. It was also excited at 635 nm and imaged with a 695-nm band-pass filter for direct detection of Cy5-labeled strands, in addition to the SybrGold imaging. For these experiments, unlabeled fuel strands were used to make sure that the signal in the Cy5 channel was not due to spurious excitation of TAMRA-labeled fuel strands.

**Kinetics Measurements.** All kinetics experiments were done in a custom-built fluorimeter. The fluorophores were excited at 532 nm, and the emitted fluorescence intensity was integrated from 550 to 600 nm. Four samples can be run in parallel, and one data point (per sample) is acquired every second. Fluorescence experiments were performed at 25 °C, with a reaction volume of 100  $\mu$ L. Fuel-formation reactions were done at a 500-nM reactant concentration. Catalysis experiments were done with purified fuel-complex concentrations varying from 20 nM to 900 nM and with catalyst concentrations varying from 1 nM to 500 nM. Reaction end points were determined by heating the sample in a water bath to 80 °C for 5 min and remeasuring fluorescence of the annealed sample. A reference sample was measured in parallel during each run, and fluorescence data was divided by this reference to correct for instrument noise.

**Selection of Reaction Kinetics Model.** The models tested were extensions of the basic enzyme–substrate interaction model  $C + F \rightleftharpoons CF \rightarrow C + W$ , where  $C$  is the catalyst,  $F$  is the fuel or substrate,  $CF$  is a complex in which catalyst and fuel are bound and  $W$  is the end or waste product. A pathway  $F \rightarrow W$  for spontaneous decay is also included. As discussed below, the rate constant for this pathway can be measured directly. In a first selection round, the number of reaction intermediates was varied systematically to correctly reproduce the overall dependence of the data on catalyst concentration and also the initial transient seen in the fluorescence traces. A total of nine models with zero, one, or two fuel–catalyst complexes,  $CF_i$ , as well as with zero, one, or two reaction intermediates,  $I_i$ , after catalyst release were generated. The most complex overall reaction pathway thus being  $C$

$+ F \rightleftharpoons CF_1 \rightarrow CF_2 \rightarrow C + I_1$ , followed by  $I_1 \rightarrow I_2 \rightarrow W$ . Several models led to similarly good fits, and the simplest of those was selected (cf. eq 2).

In a second round, different models for catalyst degradation or sequestration were compared using the model selected in the first round as a starting point. The mechanisms tested were (i) competitive inhibition by an impurity,  $C + F_i \rightleftharpoons CF_i$ , where it is assumed that the impurity concentration scales linearly with the fuel concentration, that is,  $[F_i]_0 = f_i [F]_0$ , where  $[F]_0$  is the initial fuel concentration; (ii) decay of a fuel–catalyst complex into a stuck state from which the catalyst cannot be released,  $CF \rightleftharpoons CF_{\text{stuck}}$ ; (iii) unproductive binding to a reaction intermediate,  $C + I \rightleftharpoons CI_{\text{stuck}}$ , and (iv) end-product inhibition,  $C + W \rightleftharpoons CW$ .

For every model, a parameter search was done to find a best fit to the kinetics data. Given a set of parameters, the system of ordinary differential equations was numerically integrated in MATLAB (The Mathworks, Inc., Natick, MA), and an error term was calculated for each time point.

For parameter optimization, a MATLAB algorithm for steepest gradient descent was used to minimize the mean squared error tabulated over 24 kinetic runs with fuel complexes from purifications P12, P13, and P14. Fitting parameters were the rate constants and adjustments to the estimated fuel concentrations (one fitting parameter was used per fuel purification run, bounded by  $\pm 20\%$ ). In the case of a model with competitive inhibition, the impurity fractions  $f_i = [F_i]_0/[F]_0$  were also treated as fitting parameters (one parameter per purification, bounded by 6%). It was furthermore assumed that only the waste products are fluorescent but not the possible reaction intermediates.

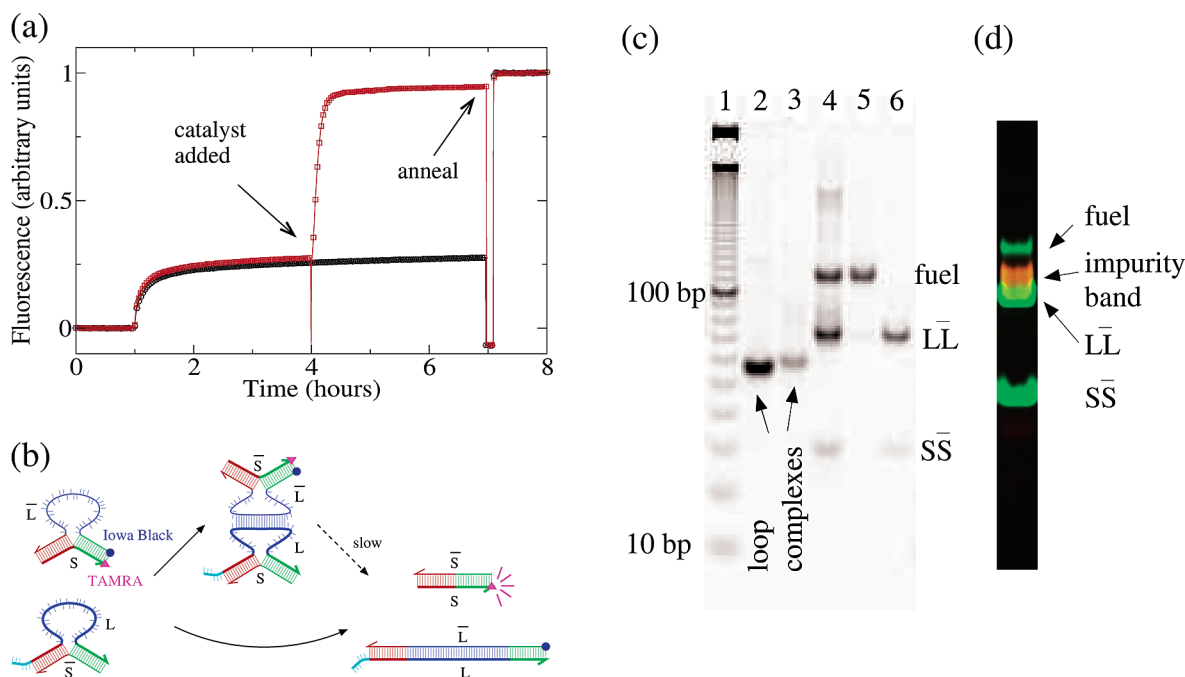
After selection of the best model (see eqs 1–4), a final parameter optimization was performed. In this last round, all 80 kinetics runs involving fuel complexes from purifications P12, P13, P14, P15, P17, and P18 were used. Of the 80 traces fit by the model, 71 have a RMS error less than 10% of their maximal value and nine are within 20%.

**Melting Experiments.** Melting experiments were used to determine the stability of fuel complexes in TAE/Mg<sup>++</sup> reaction buffer and to estimate the energy stored in the fuel. Absorbance at 260 nm was measured as a function of temperature: one data point was recorded per 0.1 °C and the temperature change rate was 0.4 °C/min. A model 14NT UV–vis spectrophotometer (AVIV Instruments Inc., Lakewood, N. J.) was used for these experiments. Melts were done at sample concentrations varying between 200 nM and 800 nM in TAE/Mg<sup>++</sup> buffer at a reaction volume of 1.5 mL. Individual loops and waste products were first annealed from 90 °C to 20 °C and then remelted from 20 °C to 90 °C. Essentially no hysteresis was observed between melt and anneal. Fuel complexes were treated slightly differently and were first melted from 20 °C to 90 °C, subsequently annealed back to 20 °C, and then melted again. Only at the onset of the first melt did the test tube contain correctly formed fuel complexes with certainty. It is at least possible that after the anneal of all four strands together fuel complexes with topologically linked loop regions can form in addition to the standard fuel complexes.<sup>30</sup>

A standard two-state approximation<sup>33</sup> was used to extract estimates for thermodynamic parameters. It is assumed that only two species are present (fully associated or fully dissociated). The temperature versus the absorbance plot is transformed into a plot of the temperature versus the fraction associated, according to  $f = (B_u - A_{260})/(B_u - B_l)$ . Here,  $f$  is the fraction associated,  $A_{260}$  is the measured absorbance, and  $B_u$  and  $B_l$  are the baseline values for the dissociated and associated species, respectively. The baselines are assumed to have a linear dependence on temperature. The melting temperature is defined as the point where the median  $(B_u + B_l)/2$  of the two baselines intersects with the data. The equilibrium constant for the association reaction  $A + B \rightleftharpoons AB$  is  $K_{\text{eq}} = C_0(1 - f)^2/f$ , where  $C_0$  is the total concentration of the species  $A$

(32) Lagos-Quintana, M.; Rauhut, R.; Yalcin, A.; Meyer, J.; Lendeckel, W.; Tuschl, T. *Curr. Biol.* **2002**, *12*, 735–739.

(33) Mergny, J.-L.; Lacroix, L. *Oligonucleotides* **2003**, *13*, 515–537.



**Figure 2.** (a) Formation of the metastable “fuel” complex. (b) Pathways for the loop–loop reaction. The location of the fluorophore (TAMRA) and quencher (Iowa Black) are indicated. (c) Formation gel. Lane 1, ten bp ladder; lane 2, unlabeled loop; lane 3, dye/quencher-labeled loop; lane 4, mixture of two loops (after  $\sim 12$  h); the three bands correspond to the fuel complex (lowest mobility), the long double-stranded waste product  $L\bar{L}$  (middle), and the short double-stranded waste product  $S\bar{S}$ ; lane 5, purified fuel complex; and lane 6, purified fuel, annealed. Dye- and quencher-labeled molecules tend to be less visible in the gel. This effect is more pronounced for the smaller molecules  $S\bar{S}$  and  $S\bar{L}$ . (d) Catalyst-impurity bound state. Images taken at two different excitation/emission wavelengths are overlaid. SybrGold-stained DNA appears green, while the Cy-5-labeled catalyst strand appears red. The catalyst is bound to an impurity state migrating slower than the long waste product  $L\bar{L}$  but faster than the fuel complex.

(or  $B$ ) when all structures are completely dissociated. Thermodynamic parameters are determined from a van’t Hoff plot of  $\ln K$  against  $1/T$ , because  $\ln K = -(\Delta H^\circ/R)(1/T) + (\Delta S^\circ/R)$ . Here  $\Delta H^\circ$  is the standard reaction enthalpy,  $\Delta S^\circ$  is the entropy, and  $R$  is the gas constant. The standard reaction free energy  $\Delta G^\circ$  at 25 °C is extrapolated from these values, which can be estimated only within the melting transition.

## Results and Discussion

**Fuel Complex Formation.** In the original hybridization-based catalyst system,<sup>1</sup> energy is provided in the form of a fuel  $L$  and an anti-fuel  $\bar{L}$ . Fuel and anti-fuel are complementary single-stranded DNA molecules that spontaneously form a low-energy duplex, the waste product (see Figure 1a). The first step toward engineering kinetic control was to slow down the spontaneous reaction, which was accomplished by partially covering the fuel strand  $L$  with a protecting strand  $\bar{S}$  that must be displaced before the fuel and anti-fuel can fully react. The reaction may then be sped-up again by introducing a catalyst strand that partially displaces  $\bar{S}$  and thus allows a fast pathway that results in the subsequent release of both the catalyst strand and the protecting group as the waste product is formed (see Figure 1b). In the original catalyst system,<sup>1</sup> a maximal catalyst-induced acceleration of roughly 30-fold was reported. Because the rate of the catalyst-induced pathway is presumably limited by the speed of branch migration, the route to an improved catalyst system appeared to be developing a fuel complex with a substantially slower spontaneous reaction rate. Preliminary results<sup>1</sup> suggested that this goal could be achieved by protecting not only  $L$  but also  $\bar{L}$  through hybridization with a short strand  $S$  complementary to  $\bar{S}$ .

Here we show that contrary to expectation, the two-loop complexes  $L\bar{S}$  and  $S\bar{L}$  do react with each other and have a half-

time comparable to that for the reaction  $L\bar{S} + \bar{L} \rightarrow L\bar{L} + \bar{S}$ . Yet, as will be shown (Figure 2), only a fraction of the loop complexes react to form the stable double-stranded DNA waste products  $S\bar{S}$  and  $L\bar{L}$ . The rest participate in a competing reaction that forms metastable two-loop complexes, which we will refer to as fuel complexes or fuel.

Figure 2a (black curve) shows a typical fluorescence trace for the reaction between the loop complexes  $L\bar{S}$  and  $S\bar{L}$ . Fluorescence is quenched at the beginning of the reaction (see Figure 2b for fluorophore and quencher locations). The increase in fluorescence is a direct measure of the decay of the loop complex  $S\bar{L}$  and the formation of the fluorescent waste product  $S\bar{S}$  as the reaction  $L\bar{S} + S\bar{L} \rightarrow S\bar{S} + L\bar{L}$  proceeds. The large fluorescence increase from the apparent reaction end point to the value reached after the anneal indicates that indeed only a fraction of the loop complexes spontaneously react to form double-stranded waste products. The fraction of loop complexes that decayed varied between experiments. Individual loop complexes were not purified, and every sample probably contained a small amount of ssDNA due to stoichiometry mismatches in loop–complex preparation. Reactions between such ssDNA complexes and loop complexes provide an additional decay pathway, and variations in the amount of ssDNA between different preparations may at least in part account for the observed variations in the fraction of loop complexes that decayed to waste product.

A gel analysis of the reaction mixture shows that three DNA species are present directly before the anneal (see Figure 2c, lane 4): the two double-stranded species  $L\bar{L}$  and  $S\bar{S}$  and a third species of higher molecular weight. Importantly, the gel also shows that at this point no unreacted loops are left. Whereas

the loop complexes migrate at a speed comparable to that of a 55 or 60 bp duplex, the high molecular weight complex migrates at the speed of a 115 bp duplex. This is consistent with a model where two loop complexes bind to each other by forming base pairs in the loop regions, as diagrammed in Figure 2b. Note that due to topological constraints, the two loops cannot form a continuous double helix.

These results also lead us to propose an alternative to the mechanism for the original hybridization-based catalyst:<sup>1</sup> namely,  $\bar{L}$  and  $L\bar{S}$  do not remain independent as previously assumed but instead react quickly to form a three-stranded complex  $\bar{L}LS$  that subsequently decays by a first-order process to  $\bar{L}\bar{L} + \bar{S}$ . The catalyst interacts with the complex  $\bar{L}LS$  and accelerates its decay.

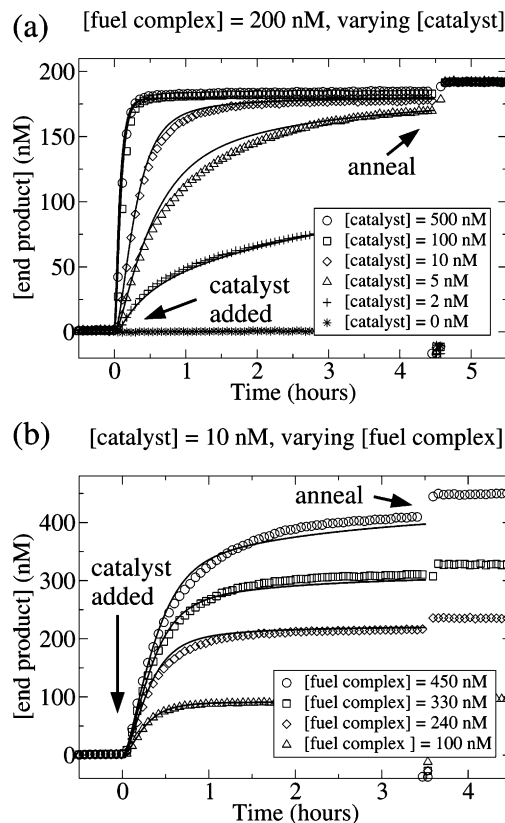
**Fuel Catalysis.** The decay of the fuel complex can not only be induced by heating but, more importantly, also by addition of catalyst strand (see red curve in Figure 2a). This suggests that the decay of the fuel complex can be triggered by binding of the catalyst to the toe-hold and subsequent opening of loop complex  $L\bar{S}$  (see, e.g., Figure 1c).

To have a well-characterized substrate for catalysis experiments, the slowly migrating fuel complex was gel-purified. The result of the purification procedure can be seen in Figure 2c, lane 5. Heating of the purified fuel complex leads to its decay into the double-stranded molecules  $\bar{L}\bar{L}$  and  $\bar{S}\bar{S}$ , as expected from the fluorescence data discussed above (see Figure 2c, lane 6).

To demonstrate catalysis in a multiple turnover situation, we measured reaction kinetics with a variety of catalyst and (purified) fuel concentrations (see Figure 3). The same reaction end point is reached in a multiple turnover situation as in a situation where catalyst strand is in excess. In the example shown, with the fuel complex concentration at 200 nM, catalyst strand at 5 nM concentration is sufficient to turn over all fuel molecules. This implies that a single catalyst strand is capable of turning over 40 or more fuel complexes.

The data for  $[C] = 0$  nM in Figure 3a shows that the spontaneous decay of the substrate proceeds exceedingly slowly. Fitting these data to a straight line allows one to estimate the decay rate as  $k_0 = 4 \times 10^{-7}$ /s. To estimate the maximal acceleration of the decay reaction due to the catalyst, without referring to any specific reaction model, we determine the initial rate for the reaction (as described below) where the catalyst concentration is  $[C] = 500$  nM. Ignoring the initial transient, we find the catalyzed decay rate to be  $1.9 \times 10^{-3}$ /s. The catalytic acceleration can then be defined as the ratio between the rates for the catalyzed and uncatalyzed reactions and is found to be roughly 4700. Two longer runs gave uncatalyzed decay rates 1.1–1.9 times slower (data not shown), so the catalytic acceleration reported here may be taken as a lower bound. The important multiple turnover regime will be discussed in more detail below.

**Catalyst Inactivation.** For the data shown in Figure 3a, the reaction does not go to completion at the lowest catalyst concentration (2 nM). In fact, in all our experiments, reactions stall at intermediate values once the fuel complex is about 50 times more concentrated than the catalyst. This suggests that there is a mechanism by which the catalyst is either captured or otherwise inactivated. A related observation is that all fluorescence traces show an increase in fluorescence from the apparent reaction end point reached just before the anneal to the value measured after the anneal. This hints at the existence

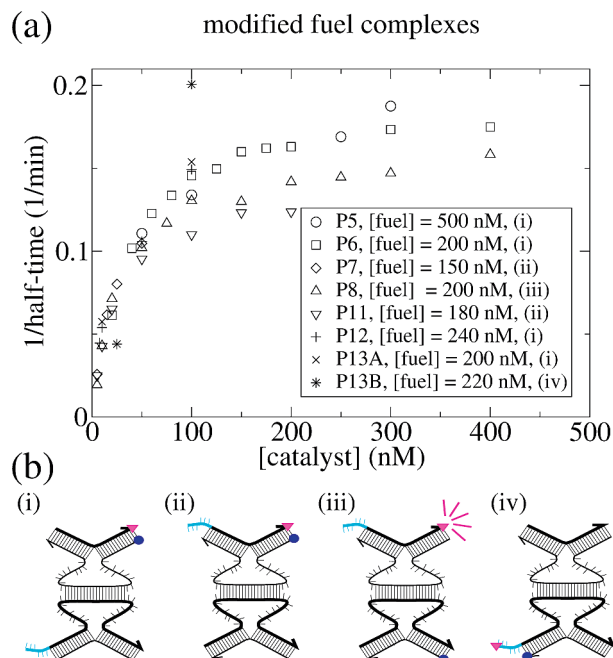


**Figure 3.** Catalysis experiments with purified fuel complex. The locations of the fluorophore and quencher in the loop are indicated in Figure 3b. (a) Decay reactions for varying catalyst concentrations (fuel complex from purification P13). (b) Decay reactions for varying fuel complex concentrations (fuel complex from purification P12). The black lines are fits using the model of eqs 1–4, with rate constants  $k_0 = 4 \times 10^{-7}$ /s,  $k_{1f} = 1.75 \times 10^5$ /M/s,  $k_{1r} = 0.006$ /s,  $k_2 = 0.050$ /s,  $k_3 = 0.004$ /s,  $k_{4f} = 1.05 \times 10^6$ /M/s,  $k_{4r} = 2.63 \times 10^{-5}$ /s,  $k_{5f} = 3.24 \times 10^3$ /M/s and  $k_{5r} = 1.44 \times 10^{-4}$ /s. The impurity concentrations are taken to be proportional to the fuel complex, that is,  $[F_i]_0 = f_i [F]_0$ , and the proportionality constants  $f_i$  vary between different purifications. Here  $f_i = 0.95 \times 10^{-2}$  for the data shown in (a) and  $f_i = 0.71 \times 10^{-2}$  for the data shown in (b).

of poorly formed fuel complexes (impurities) or reaction intermediates (stuck states) that do not decay into waste product even in the presence of catalyst. Possibly, these stuck states or impurities are responsible for catalyst inactivation.

The existence of a catalyst-binding stuck state or impurity can be tested in an electrophoresis gel experiment with a 3'-dye-labeled catalyst strand. Labeling the catalyst makes it possible to visualize it in a gel. The gel shown in Figure 2d shows that a catalyst-binding stuck state exists and is stable at least on the time scale and under the conditions of the gel experiment. The complex to which the catalyst is bound migrates at a speed intermediate to that of the purified fuel complex and the waste product  $\bar{L}\bar{L}$ . Its speed of migration is consistent with a complex that contains the two strands  $L$  and  $\bar{L}$  together with the catalyst, but our experiments did not allow us to unequivocally identify the component strands of the stuck state.

The relative concentration of impurities varied between fuel complexes from different purifications. Our model (see below) predicts impurity concentrations ranging from 0.5% of the initial fuel complex concentration in the case of purification P18 to 1.2% for P17. This may be due to small variations in the purity of the strands used for preparing different batches of fuel complex or due to some variation in the preparation protocol

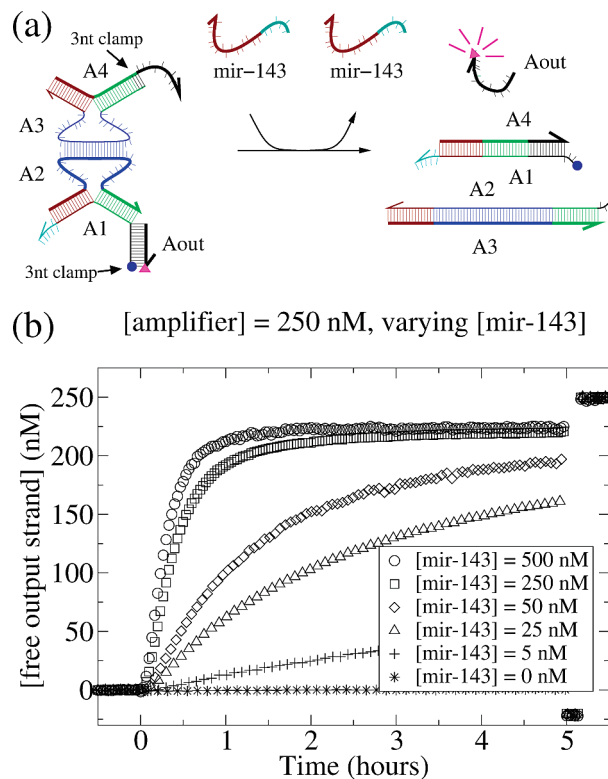


**Figure 4.** Catalysis for different purifications and modifications. We used two different toe-hold positions and three different locations for the dye and quencher in a total of four different complexes. As a model-independent measure for the speed of a reaction, we use the inverse of the reaction half-time, which can be interpreted as an overall rate constant. (a) Inverse of the reaction half-time versus catalyst concentration for the fuel complex variants (i)–(iv) shown in (b). P5–P13B indicate the different purification runs. (b) Toe-hold and dye/quencher positions used when collecting the data shown in (a).

or storage conditions (e.g., varying time intervals between preparation and use).

**Modifications and Applications.** A series of fuel complexes with toe-holds and dye/quencher pairs in a variety of different positions were also prepared. While these variations influenced the reaction kinetics quantitatively, all fuel complexes showed the same qualitative behavior, that is, they were stable at room temperature and their relaxation could be accelerated catalytically, with half-times roughly comparable to the standard molecules (see Figure 4). The reaction half-times (i.e., time to half-maximal fluorescence) were determined from the fluorescence data in Figure 3a and from similar data for the other purifications. The fluorescence value just before the anneal was used as the reaction end point when measuring reaction half-times. Half-times were only determined for reactions that went to completion. Data for very low catalyst concentrations were, therefore, not included.

The fuel complexes discussed so far all decay into chemically inert double-stranded waste products. One way to link fuel complexes to other nanomachines is to release a single-stranded output. Then, the output released in a first reaction could be used as catalyst strand in a downstream reaction, and DNA analogues to the protein enzyme based networks ubiquitous in biology could be constructed. A modified fuel complex with the ability to release a single-stranded output is shown in Figure 5. The main modifications made to the fuel strands are a 12-nucleotide extension of strand  $S$  (becomes A4) at the 3'-end and a complementary extension of  $\bar{S}$  (becomes A1) at the 5'-end. The former sequence is single-stranded in the modified fuel complex, the latter is double-stranded and hybridized to the output strand (these sequences are colored black in Figure 5).



**Figure 5.** Input signal translation and amplification. (a) Sketch of the catalytic signal amplifier. Amplifier strands are labeled Aout and A1–A4; the catalyst (dark red with dark cyan toe-hold) has the same sequence as the biological miRNA mir-143. The sequences in one arm of the loop (dark red) are adjusted accordingly. Blue and green colored regions are the same sequences as in the previous figures. The output strand and corresponding extensions of the loop strands are shown in black. Three bp clamps introduced for increased stability of the amplifier are indicated. (b) Fluorescence data for varying concentrations of input strand mir-143. Fluorescence increases as the output strand is released. Dye and quencher positions are indicated in (a).

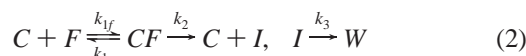
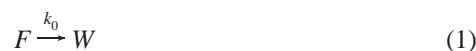
The output strand is released by strand displacement when the fuel complex decays and the double-stranded waste product forms. For fluorescence experiments, strand A1 is labeled with a quencher at the 5'-end, and the output strand Aout is labeled with a fluorophore at the 3'-end. Fluorescence is quenched when the output strand is bound, while release of the output strand leads to an increase in fluorescence. The initial and final states of a catalyzed reaction in which the output strand is released are shown in Figure 5a.

To demonstrate the modularity of our design, we used a new catalyst sequence for this experiment (with the corresponding modifications of the fuel strands). In fact, we used an oligonucleotide with a sequence homologous to that of the biological micro RNA mmu-mir-143.<sup>32</sup> At substoichiometric amounts of input strand mir-143, the modified fuel complex operates as an amplifier. A small amount of input strand catalytically releases a large amount of single-stranded output. Fluorescence data for a variety of concentrations of the input strand mir-143 are shown in Figure 5b. In a situation where the amplifier complex is about 10-fold in excess of the catalyst strand, the reaction still appears to go to completion. This implies that a single catalyst is capable of turning over on the order of 10 amplifier complexes.

For the stability of the amplifier, it is crucial that the complementary extensions on A1 and A4 cannot react spontaneously. To minimize the probability of this happening, we

introduced a three-bp clamp protecting the single-stranded extension on A4. This was done by extending strand A3 at the 5'-end by three nucleotides. Finally, to increase the binding affinity of the output strand to the extension of A1, another clamp of three-bps length was introduced before the fluorophore and quencher at the far end of the output strand.

**Modeling of Reaction Kinetics.** The kinetics data was fit to a variety of model reaction mechanisms. The simplest model having good agreement with the data (see Figure 3) included one reaction intermediate, a competitive inhibition pathway, and end-product inhibition:



Here  $C$  is the catalyst,  $F$  is the fuel complex,  $I$  is a reaction intermediate, and  $F_i$  is an impurity that competes for catalyst binding with the fuel complex. Note that the waste product  $W$  has two components, namely  $LL$  and  $SS$ . Only  $LL$ , which contains a single-stranded toe-hold, is relevant for end-product inhibition. Reactants, waste products, and intermediates for the proposed main decay pathway (eq 2) are sketched in Figure 1c.

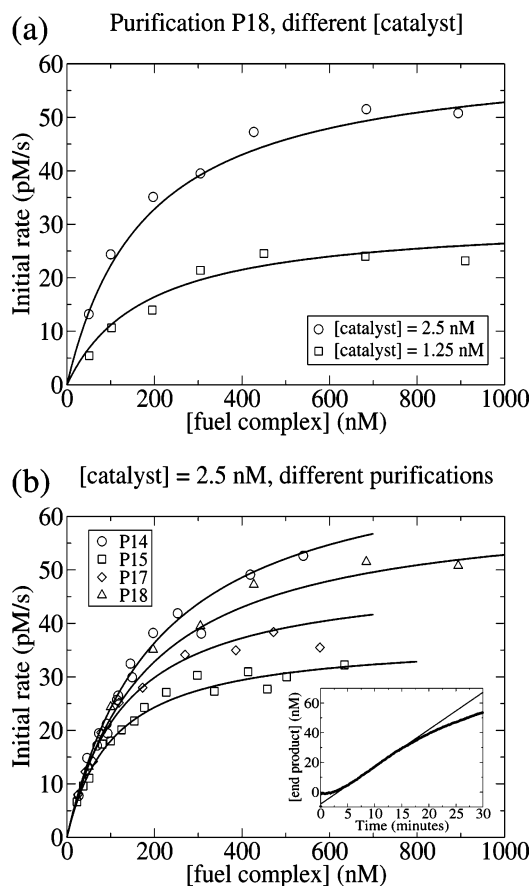
The same set of optimized parameters fits all 80 kinetic runs used for preparing Figures 3 and 6. The rate constant  $k_0$  was measured directly, as discussed above.

The rate-limiting step in the decay pathway is the decay of the intermediate species  $I$  into the waste products ( $k_3$ ); however, the rate-limiting step for catalytic turnover is release of the catalyst ( $k_2$ ) or the initial binding ( $k_{1f}$  [fuel]).

As can be seen from Figure 3, the reaction kinetics depend only weakly on catalyst concentration for concentrations within about 1 order of magnitude of the substrate concentration. Models with at least one reaction intermediate after catalyst release consistently performed better in reproducing this weak concentration dependence. The presence of at least one reaction intermediate in addition to the catalyst–fuel complex also accounts for the initial transient in the fluorescence traces (see inset of Figure 6). Adding more than one intermediate did not significantly improve the fit.

The presence of an impurity competing with the fuel complex for catalyst binding explains why reactions do not go to completion for sufficiently large differences in the concentrations of catalyst and fuel. Equation 4 accounts for end-product inhibition. As long as the equilibrium constant  $K_5 = k_{5f}/k_{5r}$  remains unchanged, the rate constants  $k_{5f}$  and  $k_{5r}$  can be varied by up to 2 orders of magnitude without deterioration to the fit.

**Kinetic Analysis.** Enzymes are often characterized in terms of their Michaelis constant  $K_M$  and catalytic constant  $k_{cat}$ . An analysis based on the Michaelis Menten equation  $v_0 = k_{cat}[C]_0[F]/(K_M + [F])$  allows one to estimate these numbers without reference to a specific reaction mechanism. Here  $v_0$  is the initial rate of waste product formation,  $[F]$  is the fuel concentration, and  $[C]_0$  is the initial catalyst concentration.



**Figure 6.** Initial rate of waste product formation for varying fuel concentrations. Data were fit to the Michaelis Menten equation  $v_0 = k_{cat}[C]_0[F]/(K_M + [F])$ . Initial rates were determined by fitting to the second five minutes after catalyst addition, thus ignoring the initial transient (see inset in (b)). Michaelis and catalytic constants obtained from the fits are listed in Table 2. (a) Initial rates measured for two different catalyst concentrations. All fuel complex used is from purification P18. (b) Initial rates measured at the same catalyst concentrations, but with fuel complexes from different purifications. The values for  $K_M$  and  $k_{cat}$  vary between different purifications, but their ratio remains relatively constant (see Table 2).

Figure 6 shows plots of measured initial rates versus substrate concentration. The figure also shows the best fits to the Michaelis Menten equation using the Eisenthal and Cornish–Bowden method.<sup>35</sup> The Michaelis Menten equation predicts that initial rates measured at different catalyst concentrations (and of course varying substrate concentration) should be fit with the same Michaelis constant  $K_M$  and catalytic constant  $k_{cat}$ . In Figure 6a, two data sets measured with fuel complex from the same purification and catalyst concentrations of 1.25 nM and 2.5 nM, respectively, are shown. In accordance with the prediction, both sets are well fit with a Michaelis Menten equation, with Michaelis constant  $K_M = 182$  nM and catalytic constant  $k_{cat} = 1.5$  /min.

Figure 6b shows four sets of initial rates data measured with substrates from different purifications. Michaelis and kinetic constants used for the fits are tabled in Table 2. The spread in the values of Michaelis and catalytic constants between different purifications is due to variations in batch purity. Interestingly,

(34) Zuker, M. *Nucleic Acids Res.* **2003**, *31*, 3406–3415.

(35) Price, N. C.; Stevens, L. *Enzymes*, 5th ed.; Horwood Publishing, Limited: Chichester, 2001; p 116.



**Table 2.** Michaelis and Catalytic Constants Obtained from the Fits Shown in Figure 6 and from a Best Fit of the Fluorescence Data to the Model of Eqs 1–4<sup>a</sup>

source	$K_M$ (nM)	$k_{cat}$ (/min)	$k_{cat}/K_M$ (/M/min)
P14 data	203	1.8	$8.7 \times 10^6$
P15 data	106	0.9	$8.4 \times 10^6$
P17 data	131	1.2	$9.0 \times 10^6$
P18 data	182	1.5	$8.3 \times 10^6$
model	319	3.0	$9.5 \times 10^6$

<sup>a</sup> From the model parameters (see Figure 3), the Michaelis and catalytic constant are calculated according to  $K_M = (k_{1r} + k_2)/k_{1f}$  and  $k_{cat} = k_2$ .

the spread in values of  $k_{cat}/K_M$  is smaller than the spread in values of either  $k_{cat}$  or  $K_M$ . All  $k_{cat}/K_M$  values are within 10% of each other, while  $k_{cat}$  and  $K_M$  vary by up to a factor of 2.

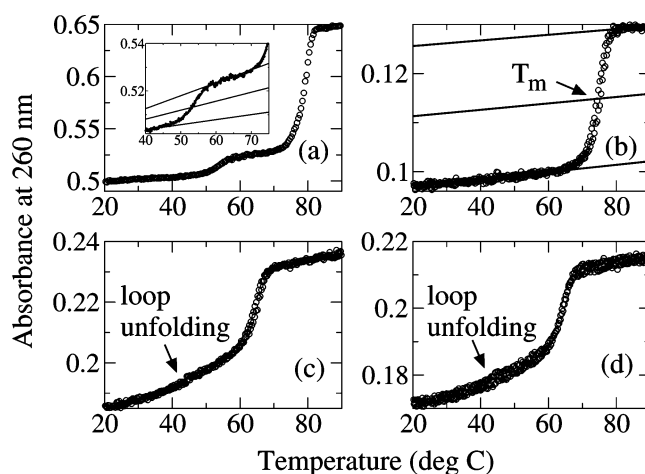
In the limiting case of no impurities and no end-product inhibition, a Michaelis Menten equation can be derived from eq 2 (assuming steady state for the reaction intermediates  $CF$  and  $D$ ). In this limit, the Michaelis and catalytic constant take the standard form  $K_M = (k_{1r} + k_2)/k_{1f}$  and  $k_{cat} = k_2$ . The ratio  $k_{cat}/K_M$ , calculated from the model parameters, is in good agreement with the independent measurements of  $k_{cat}/K_M$  discussed in this section (see Table 2).

**Fuel Energy Content.** Melting experiments were done to determine the stability of the fuel complex in TAE/Mg<sup>2+</sup> reaction buffer and to estimate the energy stored. To measure the energy stored, we need to compare the free energy released during complex formation to the energy released when the waste products  $\bar{L}\bar{L}$  and  $\bar{S}\bar{S}$  are formed instead. As described above, the formation of the fuel complex from strands  $L$ ,  $\bar{L}$ ,  $S$ , and  $\bar{S}$  is a two-step process, where first the loop complexes are formed individually. The reaction is driven by the free energy,  $\Delta G_{arms}^\circ$ , released when the 30 base pairs in the arms of the loops are formed. The total free energy released, however, is less because a free energy,  $\Delta G_{loop}^\circ$ , has to be paid for loop closing. In the second step of the fuel complex formation reaction, the two loop complexes bind to each other through their single-stranded loop regions. The energy released is  $\Delta G_{kiss}^\circ$ , where the label refers to the fact that the fuel complex is similar to a kissing hairpin complex. The total free energy associated with fuel complex formation thus is  $\Delta G_{fuel}^\circ = 2\Delta G_{arms}^\circ + 2\Delta G_{loop}^\circ + \Delta G_{kiss}^\circ$ .

The energy released when strands  $S$  and  $\bar{S}$  form a duplex is  $\Delta G_{arms}^\circ$ . Association of  $L$  to  $\bar{L}$  is driven by the free energy  $\Delta G_{arms}^\circ + \Delta G_{center}^\circ$ . Together we thus get  $\Delta G_{waste}^\circ = 2\Delta G_{arms}^\circ + \Delta G_{center}^\circ$ . The free energy stored in the fuel can now be written as the difference of the free energy released in forming the fuel and the waste products and we obtain

$$\Delta G_{stored}^\circ = \Delta G_{center}^\circ - 2\Delta G_{loop}^\circ - \Delta G_{kiss}^\circ$$

Approximate numerical values for the first two terms in this expression can be obtained using DNA mfold.<sup>34</sup> For the sequences used in our experiments and with parameters that best approximate our experimental conditions (12.5 mM Mg<sup>2+</sup> and 25 °C), we obtain  $\Delta G_{center}^\circ = -63$  kcal/mol and  $2\Delta G_{loop}^\circ = +15$  kcal/mol. An estimate for  $\Delta G_{kiss}^\circ$  was obtained experimentally (DNA mfold cannot be used because the fuel complex is pseudoknotted. In addition, the two loop regions cannot form an extended double helix). From a melting experiment (see Figure 7a) and using a two-state approximation to analyze the data, we obtained  $\Delta G_{kiss}^\circ = -23$  kcal/mol. We have deter-



**Figure 7.** Melting experiments. Concentrations of all samples are 200 nM. (a) Fuel melt. Data for a melt from 20 to 90 °C is shown. The melting of the kissing loop complex can be seen as a step around 54 °C. Inset: complex melting transition. Upper and lower baselines and their median are indicated. The melting temperature is  $T_m = 54$  °C. (b)–(d) In all figures, two data traces corresponding to annealing from 90 to 20 °C and subsequent remelting from 20 to 90 °C are shown. (b) Melt of the short waste product  $\bar{S}\bar{S}$ ,  $T_m = 75$  °C. Upper and lower baseline and their median are shown. (c) Melt of loop  $\bar{L}\bar{S}$ . Melting of the arms ( $T_m = 65$  °C) is preceded by a slow increase in absorbance due to melting of the secondary structure and unstacking of single-stranded regions in the loop. (d) Melt of loop  $\bar{S}\bar{L}$  (arms melt at  $T_m = 65$  °C).

mined  $\Delta G_{kiss}^\circ$  from several data sets measured at three different concentrations and from both annealing and melting data. The number cited here was the most negative thus obtained (and, therefore, the least favorable in terms of the energy stored). With these numbers, we are now able to estimate the total energy stored in the fuel complex as  $\Delta G_{stored}^\circ = -55$  kcal/mol. This number should be considered an order of magnitude estimate. Still, it is interesting to compare it with the free energy of ATP hydrolysis, which is  $\Delta G_{ATP}^\circ = -7.3$  kcal/mol and, thus, almost an order of magnitude smaller (for a 2 orders of magnitude smaller molecule, however).

Several caveats apply: First, the two-state approximation is generally assumed to break down for melting of DNA molecules longer than 14 bps or so and the assumption of a temperature independent  $\Delta H^\circ$  may not always hold. Second, the exact choice of baselines for the fuel-melting transition is subjective. Third, the equilibrium constant can be determined experimentally only in a relatively narrow temperature interval around the melting temperature, and small deviations from linearity in the  $\ln K$  vs  $1/T$  plot can potentially lead to large variation in  $\Delta S^\circ/R$  obtained from the y-axis intercept. Finally, the pH of Tris buffers is strongly temperature-dependent, and Tris buffers are, therefore, not ideal for melting experiments. However, these experiments still allow one to observe the main features of different melting transitions. Also, we felt that it was important to study the thermodynamic stability of the fuel complexes in the same buffer used for kinetics experiments.

To evaluate the validity of our approach, we also melted the waste products  $\bar{S}\bar{S}$  and  $\bar{L}\bar{L}$  as well as the loop complexes  $\bar{L}\bar{S}$  and  $\bar{S}\bar{L}$ . For melting of the 30mer  $\bar{S}\bar{S}$  (see Figure 7b), the melting transition was well-approximated by a two-state model, and thermodynamic parameters ( $\Delta G^\circ = -44$  kcal/mol,  $\Delta H^\circ = -242$  kcal/mol,  $\Delta S^\circ = -662$  cal/mol/K) were within 5% of the parameters calculated with DNA mfold. Not surprisingly, the melting transition for the 70mer  $\bar{L}\bar{L}$  was not well-approximated

by a two-state model.<sup>33</sup> The free energy obtained from the data was 30–40% lower than that predicted by DNA mfold. The melting transition for the loop complexes (see Figure 7c,d) is clearly not two-state. The melting of the arms is preceded by a gradual signal increase due to melting of secondary structure and unstacking in the single-stranded loop regions. Note that none of the melting data for loop complexes or waste products shows a transition similar to the transition that we associate with melting of the fuel complex.

**Conclusions.** An improved hybridization-based fuel–catalyst system was demonstrated and characterized. The stability of the fuel complex was about 2 orders of magnitude better than for a previous hybridization-based scheme, while the reaction in the presence of the catalyst was comparably fast. While the best DNA enzyme, deriving its driving energy from the cleavage of covalent bonds, has a specificity constant  $k_{\text{cat}}/K_M$  on the order of  $4.5 \times 10^9$  /M/min and a turnover rate of  $k_{\text{cat}} = 3.4$  /min at 37 °C,<sup>36</sup> the system discussed here with a specificity constant of  $8.7 \times 10^6$  /M/min and a turnover rate of  $k_{\text{cat}} = 1.8$  /min at 25 °C (from P14 data, see Table 2) is comparable to many ribozymes and deoxyribozymes.

Furthermore, the acceleration of reaction rates in our fuel/catalyst system, while not as large as many enzyme catalysts, is substantial: in a single-turnover regime, the ratio of initial rates for reactions with and without an excess of catalyst was measured at 4700, while in a multiturnover regime, the catalytic rate enhancement  $k_{\text{cat}}/k_0$  was measured to be roughly 75 000 (using P14 data from Table 2). Note that the corresponding ratios of model parameters,  $k_3/k_0$  and  $k_2/k_0$ , are twice and two-thirds larger, respectively.

Although our data and modeling imply that there is some poisoning of the catalyst, they also imply that a single-catalyst molecule can turnover on the order of 100 fuel substrates in roughly 30 min. These results establish that altering covalent bonds is not necessary in order to have a viable chemical catalyst and fuel for driving nucleic-acid systems. A two-strand fuel/catalyst system, described previously, but not fully characterized,<sup>30</sup> also looks promising for similar applications and appears to operate on similar principles, further supporting this conclusion.

In fact, hybridization-based systems have the advantage that the same secondary structures and pathways can be embodied

using a wide variety of (almost arbitrary) base sequences, usually with similar performance. For example, given a sequence of interest,  $X$ , one can design a new fuel complex for which  $X$  is the catalyst strand. Because of the specificity inherent in the initial toe-hold binding and subsequent three-strand branch migration, we expect that multiple fuels and catalysts can coexist in the same reaction solution with little cross-talk. This provides great flexibility for coupling the catalyst system to other nucleic-acid systems of interest, making it likely that our demonstrated control over reaction kinetics and conformational changes will prove useful for many applications in DNA nanotechnology.

The basic fuel system discussed here could potentially be useful as a fluorescence amplifier with a 50-fold gain. A DNA fuel complex, with the appropriate sequence, could thus be used as a highly specific detector for a miRNA, for example, in the context of in situ imaging. By coupling an aptamer or aptazyme to the catalyst, small molecules could also be sensed. We have also demonstrated a modified fuel complex where the output is not a fluorescence signal but a single-stranded DNA molecule. This allows for a more general type of signal amplification and for coupling of the fuel complex to other nanomachines. If the activating inputs come from other DNA-based logic elements, the catalyst could, for example, participate as an amplifier in a molecular information processing circuit.

Interestingly, the same catalyst system can also serve as a basic component for seemingly different applications. As a molecular motor, the catalyst strand undergoes a work cycle of binding to the fuel complex and assuming double-helical form, followed by release and return to a random coil; analogously to previous work,<sup>19–21</sup> constructing a molecule containing several catalysts could produce a molecular motor that walks on a surface displaying fuel molecules.

**Acknowledgment.** We thank Richard W. Roberts for use of his Elutrap. G.S. was supported by the Swiss National Science Foundation and the Center for Biological Circuit design at Caltech. E.W. acknowledges NSF Awards #0093846, #0533064, and #0506468. B.Y. thanks the Moore Foundation for a Moore Distinguished Scholar Fellowship. Preliminary results for this work were previously reported.<sup>37</sup>

JA0635635

(36) Santoro, W. S.; Joyce, G. F. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 4262–4266.

(37) Seelig, G.; Yurke, B.; Winfree E. *Lect. Notes Comput. Sci.* **2005**, *3384*, 329–343.