Pergamon

Tetrahedron Letters 43 (2002) 7229–7233

TETRAHEDRON LETTERS

A structurally simple minimal self-replicating system

Julie M. Quayle, Alexandra M. Z. Slawin and Douglas Philp*

Centre for Biomolecular Sciences, School of Chemistry, University of St Andrews, North Haugh, St Andrews, Fife KY16 9ST, UK

Received 27 June 2002; accepted 1 August 2002

Abstract—Molecular recognition between an amidopyridine and a carboxylic acid through two complementary hydrogen bonding sites renders the reaction between an azide and a maleimide self-replicating, but not autocatalytic.

Examples of chemical systems capable of templating and catalyzing their own synthesis—so-called self-replicating systems—have begun to appear[1] in the chemical literature over the last 15 years. For the biologist, these systems represent a putative link with the origin of life and their study may shed light[2] on prebiotic chemical evolution. However, for the synthetic chemist, the autocatalytic properties[3] of such systems are enticing. The concept of a chemical template that is capable of making billions of exact copies of itself, given appropriate starting materials, is a highly attractive one for the synthetic chemist. In addition, the growing field of nanotechnology could benefit immensely from systems that are capable of self-synthesis (i.e. replication).

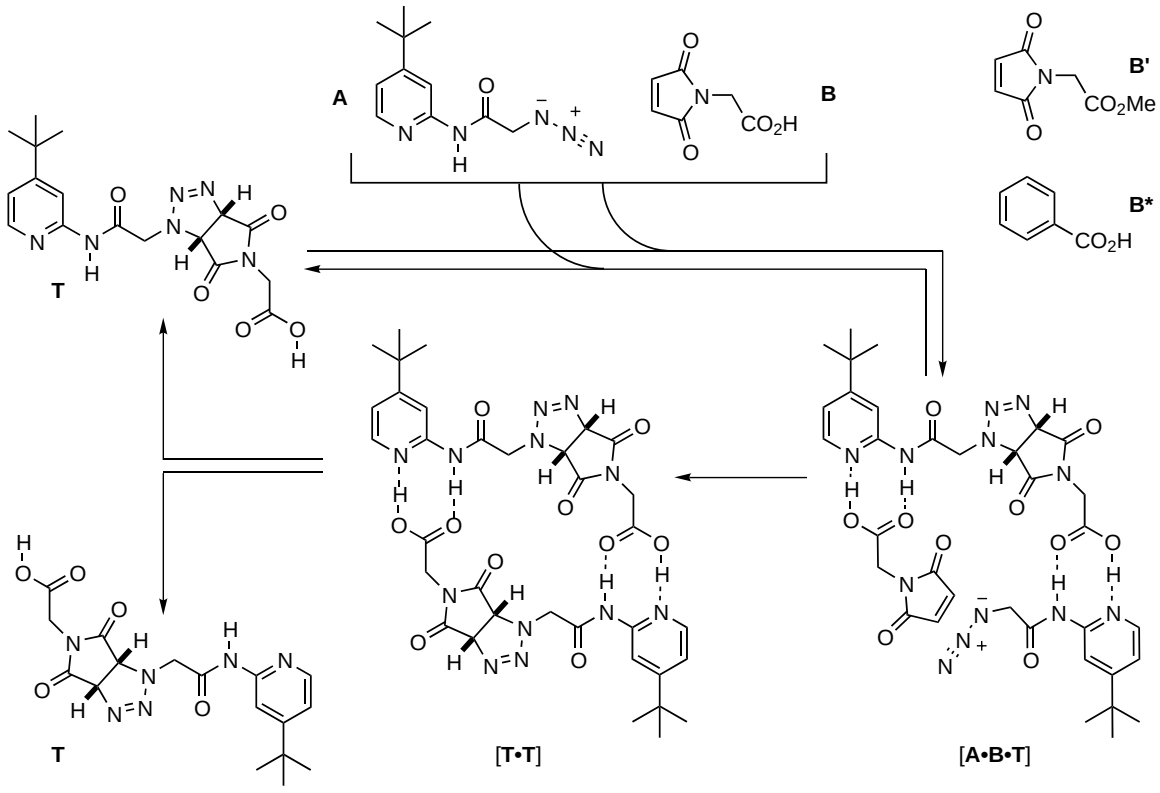
In principle, all one requires to implement a self-replicating system is two reactive groups that are each attached to self-complementary recognition sites. Thus, the development of synthetic replicating systems is an ideal challenge for supramolecular chemistry. In this work, we set out to develop self-replicating systems based on cycloaddition chemistry that were very simple in a structural sense. The reasons for this approach were two-fold. Firstly, if replication is to be of use synthetically, we must demonstrate that the addition of two self-complementary recognition sites to our reactive groups is all that is necessary to introduce autocatalytic and self-replicating behavior. Secondly, replicating systems are notorious[4] for their complex kinetic behavior. Therefore, the development of structurally simple systems should, in principle, allow easier investigation of their properties.

Accordingly, we designed a synthetic self-replicator[5] based on the minimal model[6] shown in Scheme 1. In this model, two reaction channels exist. The first is the uncatalyzed bimolecular reaction of **A** and **B** to give **T**. The second is the *autocatalytic cycle* in which **A** and **B** associate with **T** to form the catalytic ternary complex [**A**·**B**·**T**]. Bond formation occurs between **A** and **B** to give the product duplex [**T**·**T**] which then dissociates to return *two* molecules of **T** to the autocatalytic cycle. The dissociation of the [**T**·**T**] duplex is a key consideration in the design of replicating systems and for efficient autocatalytic turnover—its stability must be lower or, at worst, comparable to that of the ternary complex [**A**·**B**·**T**]. An additional reaction channel[7] involves the formation of a binary complex [**A**·**B**] in which the reaction between **A** and **B** is now pseudo-intramolecular. This pathway also accelerates the formation of **T**, but is not catalytic—the template remains closed at the end of the reaction since the recognition utilized to assemble [**A**·**B**] lives on in the final product. This pathway cannot be observed in the system shown in Scheme 1 as a result of the short spacer lengths between the recognition and reactive sites in **A** and **B**.

Here, we report the synthesis properties of the structurally simple system (Scheme 1) that exploits[8] molecular recognition to achieve modest acceleration of a cycloaddition reaction between an azide and a maleimide. The weakly replicating behavior of this system is demonstrated through a series of control experiments and preliminary kinetic modeling. The relative inefficiency of this system is traced to a product duplex [**T**·**T**] that is stable enough to be characterized in the solid state.

* Corresponding author.

0040-4039/02/$ - see front matter
PII: S0040-4039(02)01613-1



Scheme 1.

Azide **A** and maleimide **B** were synthesized[9] by standard methods. In order to make valid comparisons between the recognition-mediated and bimolecular reaction pathways, a control compound was required which possessed that same chemical functionality as the normal reagents **A** or **B**, but which was incapable of participating in a recognition mediated reaction. We have found[9] that the methyl ester of **B**, **B**′ (Scheme 1), is a suitable control compound, the carboxymethyl group in **B**′ being completely incapable of binding to the amidopyridine recognition site in **A**.

The reaction between azide **A** and maleimide ester **B**′, performed at 30°C at a concentration of 25 mM in $CDCl_3$, is very slow leading to <10% overall conversion to the corresponding triazoline over a period of 16 h. Kinetic simulation and fitting of the rate profile obtained for this reaction (Fig. 1) gave an excellent fit to a bimolecular reaction model.

When the recognition-mediated reaction between azide **A** and maleimide ester **B** is performed under identical conditions, it is immediately obvious (Fig. 1) that the rate of reaction is significantly higher in the case of the recognition-mediated reaction. Less obvious is the fact that the concentration–time profile exhibits a slight upward curve. In other words, the rate of reaction is increasing with time as opposed to decreasing with time. This behavior is indicative of an autocatalytic process in operation.

The efficient operation of a self-replicating system hinges on the reversible binding events that occur during the autocatalytic cycle. If the crucial recognition processes are interfered with, rendering the binding events inefficient or non-existent, then the autocatalytic cycle is unable to operate, or may only do so with low efficiency. Initially, we wished to prove that the reaction between **A** and **B** was recognition-mediated. In

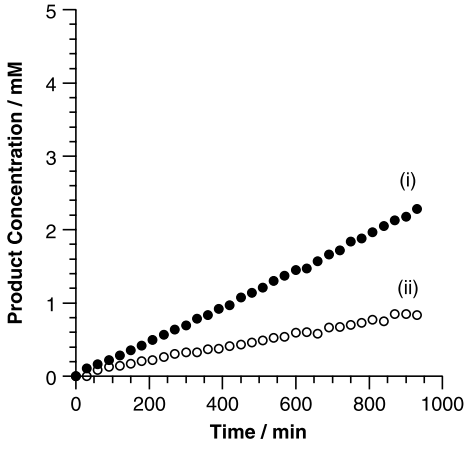


Figure 1. (i) Rate profile obtained for the reaction of **A** and **B** at 25 mM at 30°C, in $CDCl_3$. The formation of **T** is shown as filled circles. (ii) Rate profile obtained for the reaction of **A** and **B**′ at 25 mM at 30°C, in $CDCl_3$. The formation of the product triazoline is shown as open circles. For clarity, error bars are omitted from the graphs; however, errors in concentration are estimated to be ±4%.

order to achieve this, we performed a competition experiment in which we added an unreactive molecule **B***, in this case benzoic acid, which is capable of binding in an unproductive manner to **A** and thereby decreasing the rate of reaction. When the recognition-mediated reaction between azide **A** and maleimide ester **B** is performed under identical conditions to those described before, but in the presence of 100 mM **B***, the reaction rate is diminished almost to that of the control (Fig. 1(i)), thus demonstrating the reliance on molecular recognition in the rate enhancement observed in the reaction between **A** and **B**.

During the progress of a self-replicating reaction, the formation of product, **T**, during the initial stages of the reaction, is primarily via a simple uncatalyzed bimolecular pathway. However, once the product in solution reaches a critical concentration, then the autocatalytic cycle may begin to operate. Therefore, the presence of presynthesized template, **T**, at the beginning of the reaction, ($t=0$), should result in the experimentally observable loss of the initial lag period in the rate profile for the reaction. In order to demonstrate this effect, azide **A** and maleimide **B**, were reacted under identical conditions to those described previously, but with the addition of either 10 mol% or 50 mol% of added template **T**. The results of this experiment (Fig. 2) clearly show that the rate of production of **T** is enhanced by the addition of template **T**.

The efficient operation of the autocatalytic cycle hinges on the effective dissociation of the product duplex [**T**·**T**] in the autocatalytic cycle. If this duplex is too stable, no new template is returned to solution and the autocatalytic cycle fails. Unfortunately, the template **T** was not soluble enough in $CDCl_3$ to perform a direct determination of the association constant for the [**T**·**T**] duplex by NMR dilution methodology. However, an approximate value for the dimerization constant was obtained by fitting the chemical shift changes observed in the resonance arising from the amide NH proton in **T** during the reaction to the appropriate binding isotherm. The dimerization constant obtained by this methodology, 1000±100 M^{-1}, indicates that the [**T**·**T**] duplex is very stable.

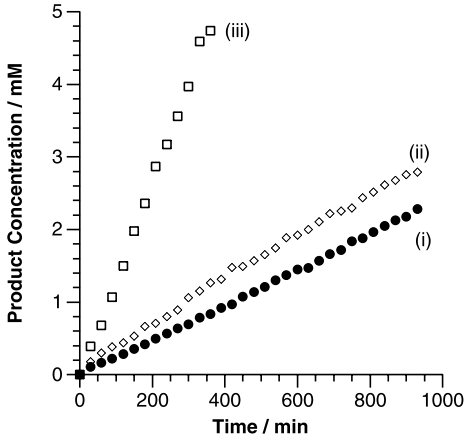


Figure 2. (i) Rate profile obtained for the reaction of **A** and **B** at 25 mM at 30°C, in $CDCl_3$. The formation of **T** is shown as filled circles. (ii) Rate profile obtained for the reaction of **A** and **B** at 25 mM at 30°C, in $CDCl_3$ in the presence of 10 mol% preformed **T**. The formation of **T** is shown as open diamonds. (iii) Rate profile obtained for the reaction of **A** and **B** at 25 mM at 30°C, in $CDCl_3$ in the presence of 50 mol% preformed **T**. The formation of **T** is shown as open diamonds. For clarity, error bars are omitted from the graphs; however, errors in concentration are estimated to be ±4%.

Further evidence for the stability of the [**T**·**T**] duplex comes from the solid state structure of **T** determined by single-crystal X-ray crystallography. Very small single crystals[10] of **T** were obtained by slow evaporation of a solution of **T** in a mixture of CH_2Cl_2 and hexane. The solid state structure reveals that **T** exists as a homodimer (Fig. 3), i.e. as the [**T**·**T**] duplex, in the solid state. This behavior in the solid state indicates that the [**T**·**T**] duplex is relatively stable since molecules that are capable of self-association often crystallize in extended structures as opposed to homodimers.

Therefore, in order to gain some insight into the kinetic behavior of the system we turned to kinetic simulation. Von Kiedrowski has introduced a simple model to describe this behavior (Fig. 4).

In this minimal model, the parameter p describes the autocatalytic behavior of the system. A value of p that is 0.5 denotes the fact that the replicating system obeys the square root law,[11] indicative of a stable [**T**·**T**] duplex. However, if the [**T**·**T**] duplex is relatively unstable, the value of p will tend to 1. An additional parameter ε is defined in this minimal model which

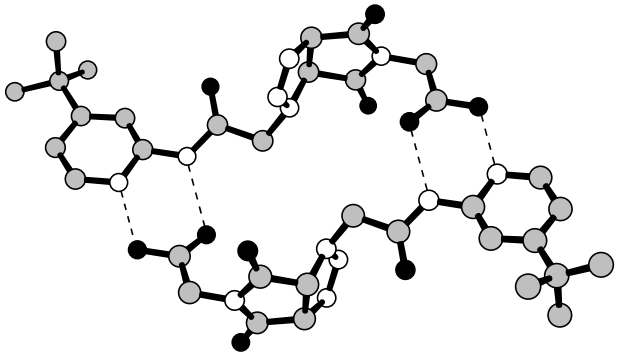


Figure 3. Structure of [**T**·**T**] duplex as determined by single crystal X-ray diffraction. Carbon atoms are gray, oxygen atoms are black and nitrogen atoms are white. Hydrogen atoms have been omitted for clarity. The hydrogen bonds stabilizing the duplex are shown as dashed lines.

$$\mathbf{A} + \mathbf{B} \xrightarrow{k_1} \mathbf{T}$$

$$\mathbf{A} + \mathbf{B} + p\mathbf{T} \xrightarrow{k_2} (1+p)\mathbf{T}$$

$$\frac{d[\mathrm{T}]}{dt} = k_1[\mathrm{A}][\mathrm{B}] + k_2[\mathrm{A}][\mathrm{B}][\mathrm{T}]^p \qquad \varepsilon = \frac{k_2}{k_1}$$

Figure 4. Minimal kinetic model for replication. For a detailed explanation of this model, see Ref. 11.

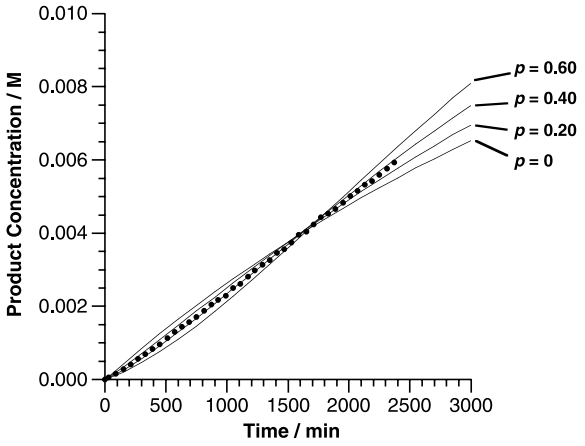


Figure 5. Results of kinetic simulation and fitting of the kinetic data for the reaction between **A** and **B** to the appropriate model. Solid lines represent the best fits of the data to the model at various values of p. The simulation was carried out using the data from both the undoped and the doped experiments. However, for clarity, only the fit to the undoped data is shown.

describes the relative importance[12] of the bimolecular pathway with respect to the autocatalytic channel. Fitting of our experimental data to this minimal model proceeded smoothly[13] (Fig. 5) and afforded best-fit values of p and ε of 0.40 and 20, respectively.[14] These results suggest strongly that the [**T**·**T**] duplex is very stable in our system and this leads to almost no turnover in the autocatalytic cycle. In this case, the system is self-replicating—the template **T** is capable of organizing **A** and **B** and accelerating the reaction between them—but not autocatalytic—the stability of the [**T**·**T**] duplex prevents turnover.

In summary, we have demonstrated that the reaction between azide **A** and maleimide **B** at 30°C in $CDCl_3$ at a concentration of 25 mM forms a template **T** that is capable of self-replication. We have demonstrated directly, through X-ray crystallography, that this template forms a strong duplex [**T**·**T**] in the solid state and indirectly, through kinetic simulation, that this template duplex is stable in solution. The stability of this duplex precludes efficient autocatalysis in this system. The challenge is now to deduce the structural basis for the fact that some minimal replicators show efficient autocatalysis[15] and others do not. This understanding is a requirement for the successful application of this technology in synthetic chemistry and beyond.

Acknowledgements

This work was supported by the Engineering and Physical Sciences Research Council (Quota Award to J.M.Q.), the Biotechnology and Biological Sciences Research Council (Grant B6/11855) and the University of St Andrews. We thank Professor G. von Kiedrowski for providing us with a copy of his SimFit program.

References

1. (a) Robertson, A.; Sinclair, A. J.; Philp, D. *Chem. Soc. Rev.* **2000**, *29*, 141; (b) Lee, D. H.; Severin, K.; Ghadiri, M. R. *Curr. Opin. Chem. Biol.* **1997**, *1*, 491; (c) Philp, D.; Stoddart, J. F. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 1154; (d) Famulok, M.; Nowick, J. S.; Rebek, J., Jr. *Acta Chem. Scand.* **1992**, *46*, 315.
2. (a) Bridson, P. K.; Orgel, L. E. *J. Mol. Biol.* **1980**, *144*, 567; (b) Kuhn, H.; Waser, J. *Angew. Chem., Int. Ed. Engl.* **1981**, *20*, 500; (c) Inoue, T.; Orgel, L. E. *Science* **1983**, *219*, 859; (d) Cech, T. R. *Sci. Am.* **1986**, *255*, 76; (e) Joyce, G. F. *Nature* **1989**, *338*, 217; (f) Cairns-Smith, A. G. *Genetic Takeover*; CUP: Cambridge, 1982; (g) *Clay Minerals and the Origin of Life*; Cairns-Smith, A. G.; Hartman, H., Eds.; CUP: Cambridge, 1986.
3. Drexler, K. E. *Nanosystems: Molecular Machinery, Manufacturing, and Computations*; Wiley: New York, 1992; Chapter 8. See also: Blackmond, D. G. *Acc. Chem. Res.* **2000**, *33*, 402.
4. Menger, F. M.; Eliseev, A. V.; Khanjin, N. A.; Sherrod, M. J. *J. Org. Chem.* **1995**, *60*, 2870.
5. (a) Severin, K.; Lee, D. H.; Martinez, J. A.; Vieth, M.; Ghadiri, M. R. *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 126; (b) Lee, D. H.; Granja, J. R.; Martinez, J. A.; Severin, K.; Ghadiri, M. R. *Nature* **1996**, *382*, 525; (c) Lee, D. H.; Severin, K.; Yokobayashi, Y.; Ghadiri, M. R. *Nature* **1997**, *390*, 591; (d) Severin, K.; Lee, D. H.; Martinez, J. A.; Vieth, M.; Ghadiri, M. R. *Chem. Eur. J.* **1997**, *3*, 1017.
6. von Kiedrowski, G. *Angew. Chem., Int. Ed. Engl.* **1986**, *25*, 932.
7. (a) Howell, S. J.; Philp, D.; Spencer, N. *Tetrahedron* **2001**, *57*, 4945; (b) Bennes, R. M.; Sapro-Babiloni, M.; Hayes, W. C.; Philp, D. *Tetrahedron Lett.* **2001**, *42*, 2377; (c) Bennes, R. M.; Philp, D.; Spencer, N.; Kariuki, B. M.; Harris, K. D. M. *Org. Lett.* **1999**, *1*, 1087.
8. Soai, K.; Shibata, T.; Sato, I. *Acc. Chem. Res.* **2000**, *33*, 382.
9. Booth, C. A. Ph.D. Thesis, University of Birmingham, 1998.
10. X-Ray diffraction studies on crystals of **T** were performed at 293 K using a Bruker SMART diffractometer with graphite-monochromated Mo-Kα radiation (λ= 0.71073 Å). The crystals were very small and it required many attempts before a crystal that gave any diffraction at all was found. As the diffraction was very weak, all the non-hydrogen atoms were refined isotropically in order to maximize the data/parameter ratio. The structure was solved by direct methods. The hydrogen atoms on N(7) and O(20) were located by a ΔF map and allowed to refine isotropically subject to a distance constraint (X–H = 0.98 Å). The remaining hydrogen atoms bound to carbon were refined in idealized geometries. Structural refinements were by the full-matrix least-squares method on F^2 using the program SHELXTL. $C_{17}H_{20}N_6O_5$, M= 388.39, monoclinic, space group $P2_1/c$, a = 13.013(4), b = 17.121(6), c = 8.522(3) Å, β = 98.104(7)°, U = 1879.8(11) Å^3, Z = 4, D_c = 1.372 Mg/m^3, μ = 0.104 mm^{-1}, F (000) = 816, crystal size = 0.1×0.1×0.01 mm. Of 9453 measured data, 2720 were unique (R_{int} 0.6151), and 550 observed ([$I>2\sigma(I)$]) to give R_1 = 0.1166 and wR_2 = 0.1948.

11. For a complete discussion, see: von Kiedrowski, G. *Bioorg. Chem. Front.* **1993**, *3*, 151. There has been considerable discussion in the literature concerning the benefits and drawbacks of using a minimal model compared to full kinetic modeling (Reinhoudt, D. N; Rudekevich, D. M.; de Jong, F. *J. Am. Chem. Soc.* **1996**, *118*, 6880). We have performed both minimal and full kinetic modeling on this system. The data from the full kinetic modeling is entirely consistent with the conclusions drawn here; however, its description is beyond the scope of this paper and will be presented elsewhere. The minimal modeling is presented here in order to set our system in the context of other replicating systems published previously.
12. Only two of the three reaction channels are open to this system as reaction is impossible within the [**A**·**B**] complex as a result of the short methylene spacer between the maleimide and the carboxylic acid in **B**.
13. Kinetic simulation and fitting were performed using the program SimFit: Sievers, D.; von Kiedrowski, G. *Chem. Eur. J.* **1998**, *4*, 629.
14. Systems based on nucleic acids have p values of around 0.5. Peptide-based systems (Ref. 5) have p values between 0.6 and 0.7. The Diels–Alder based system of Sutherland (Ref. 15) has a p value of 0.80 and our nitrone based system (Ref. 15) has a p value of 0.91.
15. (a) Wang, B.; Sutherland, I. O. *Chem. Commun.* **1997**, 1495; (b) Allen, V. C.; Philp, D.; Spencer, N. *Org. Lett.* **2001**, *3*, 777.