ARTICLE

OBC
www.rsc.org/obc

Self-replication *vs.* reactive binary complexes—manipulating recognition-mediated cycloadditions by simple structural modifications†

Russell J. Pearson, Eleftherios Kassianidis, Alexandra M. Z. Slawin and Douglas Philp*
Centre for Biomolecular Sciences, School of Chemistry, University of St Andrews, North Haugh, St Andrews, UK KY16 9ST. E-mail: d.philp@st-andrews.ac.uk; Fax: +44 (0)1334 463808; Tel: +44 (0)1334 467264

Received 7th May 2004, Accepted 29th September 2004
First published as an Advance Article on the web 3rd November 2004

The rate of reaction and the selectivity of a Diels–Alder cycloaddition between a furan and a maleimide can be enhanced by the introduction of complementary recognition sites on the reactant species. Subtle manipulation of other structural elements allows the generation of the observed rate enhancements and selectivities through either self-replication or formation of a pre-reactive binary complex.

Introduction

Several examples[1] of chemical systems capable of templating and catalysing their own synthesis—self-replicating systems—have appeared in the chemical literature over the last 15 years. For the biologist, these systems represent a link[2] with the origins of life—their study may be able to shed light on prebiotic chemical evolution. However, for the synthetic[3] chemist, these systems represent the ultimate synthetic machine, capable of templating the production of multiple copies of themselves from a single original molecule. Almost all of the examples of synthetic self-replicators[4] reported to date are based on the minimal model[5,6] shown in Fig. 1.

Within this minimal model for self-replication three reaction channels exist. The first is the uncatalysed bimolecular reaction between reagents **A** and **B** to afford the template **T**. However, a requirement of the minimal model is that **A** and **B** bear complementary recognition[3,7] sites. Thus, **A** and **B** can associate with each other through these complementary sites to form a binary complex, [**A**·**B**]. The presence of this complex offers a second reaction channel—the [**A**·**B**] complex[8,9] channel—in which the reaction between **A** and **B** is pseudo-intramolecular[10] and forms a closed template **T**′. The recognition used to assemble the binary complex lives on in the template, thus, although rate acceleration is achieved[1] by this mechanism, this template is catalytically inert.

The third reaction channel is the autocatalytic[11] cycle. Here, **A** and **B** bind reversibly to an open template **T** to form a catalytic ternary complex [**A**·**B**·**T**] in which the reaction between **A** and **B** is also rendered pseudo-intramolecular. 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 start of the autocatalytic cycle. Thus, the open template **T**, with unhindered recognition sites in the correct orientation, can potentially act as a self-replicator, transmitting molecular information by formation of identical template molecules.

The two recognition-mediated reaction channels shown in Fig. 1 accomplish the same goal—the acceleration and regio- and/or stereocontrol[12] of a chemical reaction. However, the [**A**·**B**] complex channel does not involve any form of recognition-mediated catalysis since the closed template **T**′ does not take any further part in the reaction scheme. By contrast, a self-replicating system is an attractive target for the synthetic chemist as it offers its non-linear kinetic behaviour as a means of amplifying, in an exponential manner, a particular structure or stereochemistry over others in the system.

We have become interested in the factors that govern the choice of reaction pathway adopted by systems whose reactivity

† Electronic supplementary information (ESI) available: details of the kinetic simulation and fitting. See http://www.rsc.org/suppdata/ob/b4/b406862a/

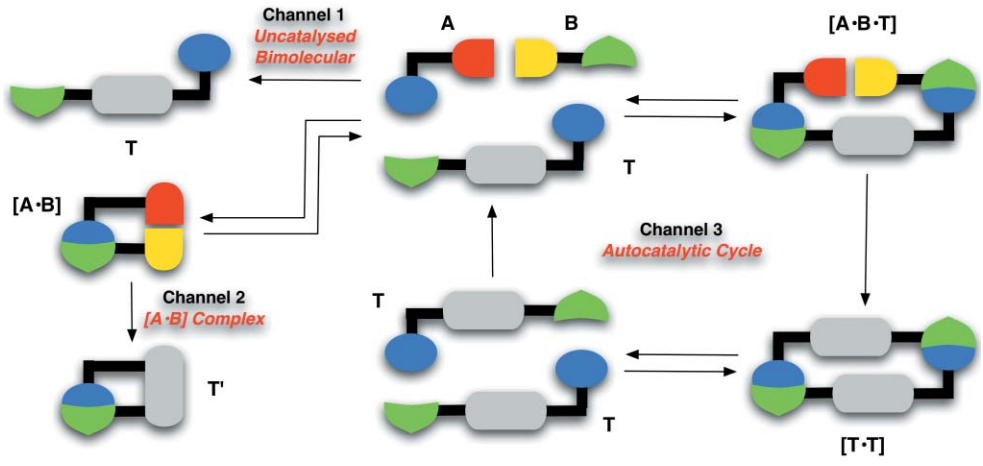
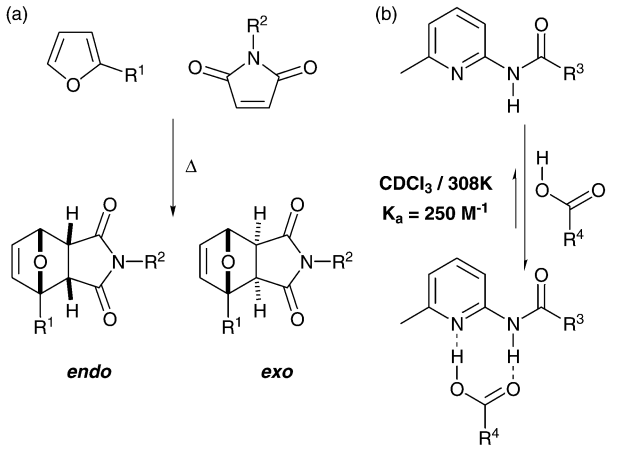


Fig. 1 The minimal model of replication. Building blocks **A** and **B** react *via* one of three pathways—the uncatalysed bimolecular pathway (Channel 1), the binary [**A**·**B**] complex channel (Channel 2) which leads to a closed template **T**′ and the autocatalytic cycle (Channel 3) which is mediated by the open template **T**.

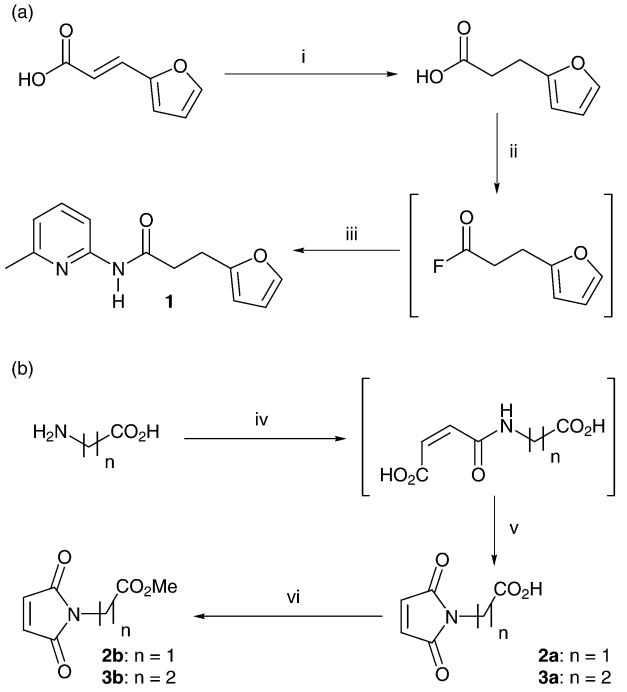
DOI: 10.1039/b406862a



Scheme 1 (a) Reaction of a 2-substituted furan derivative with an *N*-alkyl maleimide affords the diastereoisomeric *exo* and *endo* cycloadducts. (b) A carboxylic acid and an amidopyridine associate through two hydrogen bonds in non-polar solvents to form a stable complex.

is mediated by recognition processes. In particular, we wish to establish solid design criteria that will allow us to exploit replicating systems in the construction of large molecular assemblies.[13] To this end, we have identified the Diels–Alder[14] reaction between a 2-alkylfuran and a maleimide as a suitable platform upon which to conduct these studies. This reaction[12] affords two products—the *endo* cycloadduct and the *exo* cycloadduct (Scheme 1)—which differ in the relative stereochemistry at the 6–5 ring junction and these two cycloadducts have markedly different geometries. The reaction proceeds under mild conditions without the need for external reagents, making it ideal for study by NMR methods.

In this work, we describe the design and synthesis of two systems based on Diels–Alder chemistry which differ structurally by only one CH_2 group, but whose kinetic behaviours and diastereoselectivities are markedly different. These differences can be rationalised and explained in terms of the different recognition-mediated reaction pathways—facilitated by the amidopicoline–carboxylic[3,7,15] acid recognition motif (Scheme 1)—which predominate in each system.

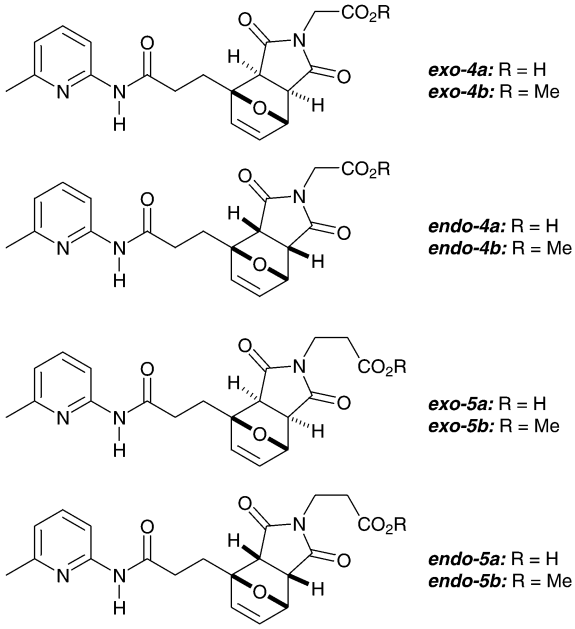


Scheme 2 *Reagents and conditions*: (i) 5% Pd/C, MeOH, 3 h, rt, 90%; (ii) cyanuric fluoride, pyridine, CH_3CN, 5 min, rt; (iii) 2-amino-6-methylpyridine, CH_2Cl_2, 16 h, rt, 58%; (iv) maleic anhydride, AcOH, 16 h; (v) AcOH, reflux, 5 h, 23%; (vi) Cs_2CO_3, MeI, DMF, 16 h, rt, 98%.

Results and discussion

The starting materials required for our recognition-mediated Diels–Alder cycloaddition reactions—compounds **1**, **2a** and **3a**—were synthesised (Scheme 2) by standard methods. Additionally, the methyl esters **2b** and **3b** were prepared to function as control dienophiles since the methyl ester is incapable of binding to the amidopicoline recognition site present in **1**.

Initially, we performed control reactions involving the methyl esters **2b** and **3b** in order to determine the inherent reactivity and selectivity within these systems. Reaction of **1** with **2b** in $CDCl_3$ at 308 K afforded the two cycloadducts *exo*-**4b** and *endo*-**4b**. When the starting concentrations of the diene and dienophile were 0.025 M, the total product concentration was around 0.003 M after 8 hours. After this time, the ratio of *exo*-**4b** : *endo*-**4b** was 2 : 1. Similarly, reaction of **1** with **3b** in $CDCl_3$ at 308 K afforded the two cycloadducts *exo*-**5b** and *endo*-**5b**. When the starting concentrations of the diene and dienophile were 0.025 M, the total product concentration was around 0.002 M after 8 hours. After this time, the ratio of *exo*-**5b** : *endo*-**5b** was 1 : 1. It is obvious from these results that neither of the two Diels–Alder reactions is particularly fast (<15% conversion in 8 hours) or particularly selective. Therefore, they are excellent candidates for the use of recognition-based strategies to enhance both their reaction rates and their diastereoselectivities.



In order to probe the effect of molecular recognition on these reactions we acquired a series of concentration–time profiles for the reaction between the diene **1** and either dienophile **2a** or dienophile **3a** in $CDCl_3$ at 308 K. The concentration–time profiles were acquired using 500 MHz ^{1}H NMR spectroscopy and the concentrations of each species present were determined by deconvolution† methods. In all of the figures depicting the results of the kinetic experiments, data for recognition-mediated reactions are always shown by filled blue circles and the corresponding control data by open blue circles.

We also performed additional experiments designed to probe the nature of the recognition-mediated processes present in each system. These experiments give a phenomenological insight into the recognition-mediated processes operating in a system.

In the first additional experiment, the so-called *inhibitor* experiment, a predefined amount of benzoic acid (4 eq) was added to the initial reaction mixture. In this type of experiment,

a reduction of the initial rate of production of cycloadduct indicates that the formation of that cycloadduct is recognition-mediated since the benzoic acid is a competitive inhibitor of the amidopicoline recognition site present in **1**. In the figures depicting the results of the kinetic experiments, data for inhibitor experiments are always shown by filled red triangles.

In the second additional experiment, the so-called *doping* experiment, a predefined amount of the appropriate presynthesised cycloadduct **4a** was added to the initial reaction mixture. In this type of experiment an enhancement of the initial rate of production of cycloadduct indicates that the formation of that cycloadduct is directed by itself (Fig. 1, Channel 3), *i.e.* an increase in the initial rate of reaction on the addition of template is a signature[1] of replication. In the figures depicting the results of the kinetic experiments, data for doping experiments are always shown by filled green diamonds.

The results of the doping and inhibitor experiments are usually conclusive in identifying the recognition-mediated reaction channels operating within a system. However, we also undertook detailed kinetic simulation and fitting of the experimental data to kinetic models describing each system. In all cases, solid lines indicate the best fit of the experimental data to the appropriate kinetic model in the concentration–time profiles.

The results of the kinetic experiments involving **1** and **2a**, and the corresponding control **2b**, are shown in Fig. 2. It is obvious immediately that the production of both *endo*-**4a** and *exo*-**4a** are enhanced by the presence of recognition. After 8 hours, the concentration of *endo*-**4a** (Fig. 2a, closed circles) is 2.2 times higher than the corresponding control *endo*-**4b** (Fig. 2a, open circles). After the same time, the concentration of *exo*-**4a** (Fig. 2b, closed circles) is 3.1 times higher than the corresponding control *exo*-**4b** (Fig. 2b, open circles). Thus, although the reaction is faster when recognition-mediated, it is, in fact, slightly less selective. In order to gain further insight into the processes operating within this system, we performed doping and inhibitor experiments. In both cases, the rates of formation of *endo*-**4a** and *exo*-**4a** are diminished (Fig. 2a and b, red triangles) in the presence of a competitive inhibitor, benzoic acid. This behaviour indicates that both the *endo* and *exo* cycloadducts are formed by recognition-mediated processes. The introduction of a relatively small amount (19 mol%) of *endo*-**4a** at the start of the reaction (Fig. 2a, green diamonds) results in a significant increase in the rate of formation of *endo*-**4a** suggesting that this cycloadduct is a template for its own formation, *i.e.* it is replicating. By contrast, the introduction of a relatively small amount (10 mol%) of *exo*-**4a** at the start of the reaction results (Fig. 2b, green diamonds) in no change in the rate of formation of *exo*-**4a** suggesting that this cycloadduct is not a template for its own formation, *i.e.* it is not replicating. In fact, the behaviour of *exo*-**4a** is more consistent with its recognition-mediated formation occurring through the [**A**·**B**] reaction channel (Fig. 1, Channel 2).

It proved possible to grow[16] single crystals of *exo*-**4a** suitable for X-ray diffraction. Surprisingly, the structure of *exo*-**4a** in the solid state (Fig. 3) reveals a head-to-tail duplex structure consistent with that shown schematically in Fig. 1 as arising from the operation of the system through the [**A**·**B**·**T**] ternary complex. Whilst this result might be attributed to crystal packing effects, we felt that this observation warranted further investigation.

In order to gain further insight into the behaviour of this system, we performed a series of molecular mechanics calculations to explore the conformational space open to both *exo*-**4a** and *endo*-**4a**. Monte Carlo conformational searches were performed from a variety of starting conformations using the AMBER* forcefield and the GB/SA salvation model for $CHCl_3$. Representative minimum energy structures are shown for *exo*-**4a** and *endo*-**4a** in Fig. 4a and b respectively. It is clear from the results of these calculations that conformational restrictions exist within both structures which prevent the formation of the classical doubly hydrogen-bonded association between the

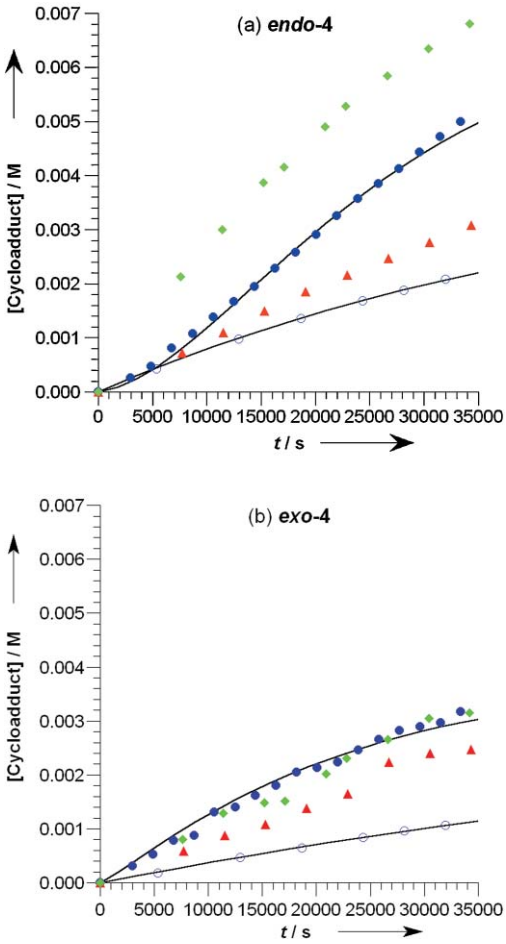


Fig. 2 Concentration–time profiles for the formation of (a) *endo*-**4** and (b) *exo*-**4** using diene **1** and dienophile **2a**. All experiments were performed with starting concentrations of 25 mM in $CDCl_3$ at 35 °C. The data for the recognition-mediated reaction between **1** and **2a** affording **4a** are shown as filled blue circles, the data for the reaction between **1** and **2a** in the presence of benzoic acid (100 mM) are shown as filled red triangles and the data for the reaction between **1** and **2a** in the presence of the appropriate template are shown as filled green diamonds. The data for the control reaction between **1** and **2b**, forming (a) *endo*-**4b** and (b) *exo*-**4b**, are shown as open blue circles. The solid lines represent the best fit of the data for the recognition-mediated and control reactions to the appropriate kinetic model (See electronic supplementary information (ESI)† for details of kinetic simulation and fitting).

amidopyridine and the carboxylic acid, shown schematically in Scheme 1b, which is particularly stabilising. In this respect, it therefore seems likely that both *exo*-**4a** and *endo*-**4a** exist in solution as open templates which are therefore capable of binding both **1** and **2a** to form the corresponding ternary complex [**1**·**2a**·**4a**] required to mediate reaction through Channel 3 in Fig. 1.

Therefore, we might expect that addition of *exo*-**4a** and *endo*-**4a** to the initial reaction mixture should enhance the rate of their own formation. However, the addition of a template which has a high dimerisation constant at concentrations well above its K_D to the reaction mixture does not, in fact, add any catalytically active *free* template to solution. Therefore, the initial rate of reaction is unaffected. In order to probe this hypothesis further, we undertook kinetic simulation†[17] and fitting of our experimental data to a kinetic model which incorporated replicating behaviour for both *exo*-**4a** and *endo*-**4a**, but incorporated different dimerisation constants for [*exo*-**4a**·*exo*-**4a**] and [*endo*-**4a**·*endo*-**4a**]. In constructing these kinetic models, we incorporated known rate constants for the bimolecular reactions which afford the two cycloadducts—determined from the reaction of **1** with **2b**—and known association constants for the complexation of a model amidopyridine with **2a**. The results of these calculations are

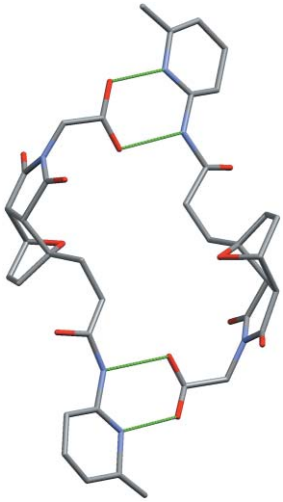


Fig. 3 Stick representation of the structure, determined by X-ray diffraction, of cycloadduct *exo*-**4a** which forms a head-to-tail duplex in the solid state. Carbon atoms are coloured grey, oxygen atoms are coloured red, nitrogen atoms are coloured blue. Hydrogen bonds are represented by green lines.

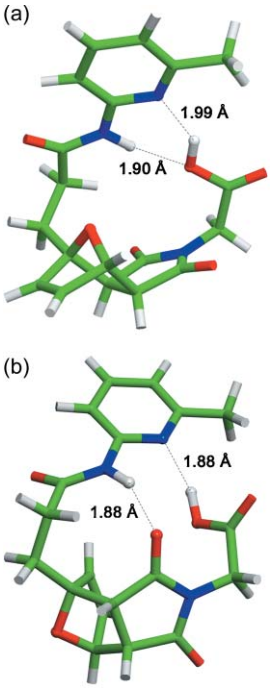
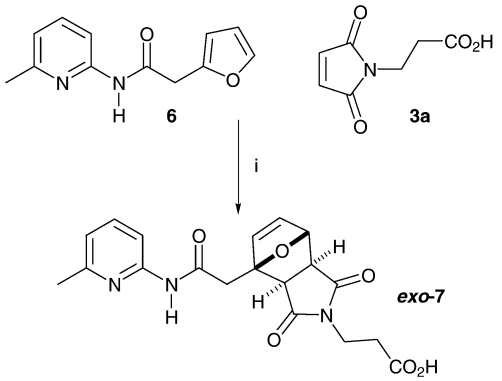


Fig. 4 Stick models of representative minimum energy conformations of *exo*-**4a** (a) and *endo*-**4a** (b) calculated using the AMBER* forcefield. Carbon atoms are coloured green, oxygen atoms are coloured red, nitrogen atoms are coloured blue, hydrogen atoms are coloured white. Hydrogen bonds are represented by dashed lines and H···acceptor atom distances are given in Å.

depicted as the solid lines in Fig. 2. Simultaneous excellent fits of the experimentally observed data to the model are achieved. Full details of the kinetic simulation and fitting are given in the electronic supplementary information (ESI).†

It is noteworthy that in order to fit the observed data successfully, several constraints are required. Firstly, in order to achieve a satisfactory fit of the observed data, it is necessary to allow reactions through *both* binary complexes (Fig. 1 Channel 2) and ternary complexes (Fig. 1 Channel 3). Secondly, the dimerisation constant for the [*exo*-**4a**·*exo*-**4a**] duplex must be $>10000\ M^{-1}$. Values of K_a less than this figure result in a uniformly poor fit of both *endo* and *exo* datasets. Thirdly, it is reassuring to note that the model is very sensitive to the value of K_a used for single amidopyridine–carboxylic acid associations. Any significant deviation (>10%) from the value determined using the model system ($K_a = 250\ M^{-1}$) results, once again, in uniformly poor fits of both *endo* and *exo* datasets. It would therefore appear from these results that the short and relatively inflexible one carbon spacer present in **2a** leads to all the recognition-mediated pathways open to the system being inefficient. Therefore, neither the [**A**·**B**] complex pathway (Channel 2, Fig. 1) nor the autocatalytic cycle (Channel 3, Fig. 1) can be utilised effectively since the system cannot easily enter the transition state for reaction from either the binary [**A**·**B**] complex or the ternary [**A**·**B**·**T**] complex. One way of viewing the inefficiency of these complexes is to regard them as being incapable of populating a significant proportion of near attack conformations[18] (NACs).



Scheme 3 *Reagents and conditions*: (i) $CDCl_3$, 25 mM, 35 °C.

Further support for the mismatching of spacers in **1** and **2a** comes from consideration of the reaction (Scheme 3) between compounds **6** and **3a**. The reaction of **6** and **3a** is fast and completely selective—forming only *exo*-**7**. This recognition-mediated process operates[8,9] through a binary [**A**·**B**] complex (Fig. 1, Channel 2), which generates a kinetic effective molarity (kEM) of 2.6 M for the cycloaddition reaction. Comparison of these two systems reveals (Fig. 5) that **1** has a CH_2CH_2 spacer between the recognition site and the diene whereas **6** has only a CH_2 spacer. By contrast, **2a** has only a CH_2 spacer between the recognition site and the dienophile whereas **3a** has a longer CH_2CH_2 spacer between the recognition site and the dienophile. In essence, in one system the lengths of the spacers on each of the components are reversed with respect to the other system. These very small changes in molecular structure, however, are apparently enough to render these two systems completely different in their kinetic behaviour.

In order to explore this relationship between spacers and recognition-mediated reactivity further, we next examined the reaction of **1** with **3a** and the corresponding control compound **3b**. In this case, there is now an additional carbon atom between the recognition site and the reactive site on the dienophile.

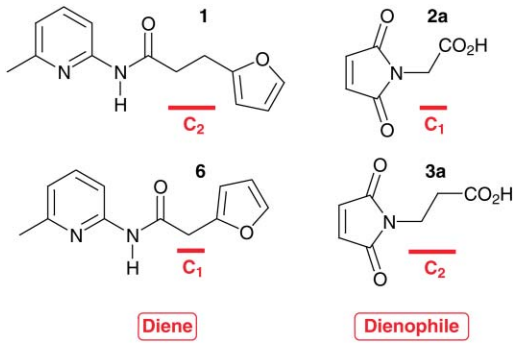


Fig. 5 A comparison of the location and length of spacers present in reactive diene and dienophile pairs: (upper) **1** and **2a** and (lower) **6** and **3a**.

The results of these kinetic experiments are shown in Fig. 6. It is obvious immediately that the production of *exo*-**5a** is enhanced significantly by the presence of recognition but the formation of *endo*-**5a** is not. After 8 hours, the concentration of *endo*-**5a** (Fig. 6a, closed circles) is almost identical to the corresponding control *endo*-**5b** (Fig. 6a, open circles). By contrast, after the same time, the concentration of *exo*-**5a** (Fig. 6b, closed circles) is eight times that for the corresponding control *exo*-**5b** (Fig. 6b, open circles). Thus, the reaction is not only faster when recognition-mediated; it is also significantly more *exo* selective—the *endo* : *exo* ratio changes from 1 : 1 in absence of recognition (Fig. 6, open circles) to 1 : 4 in the presence of recognition (Fig. 6, closed circles) after 8 hours.

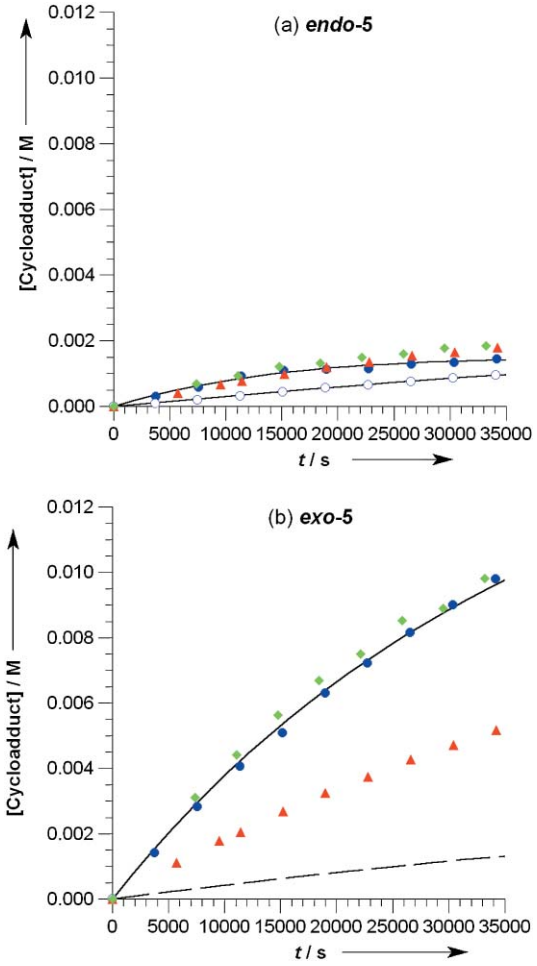


Fig. 6 Concentration–time profiles for the formation of (a) *endo*-**5** and (b) *exo*-**5** using diene **1** and dienophile **3a**. All experiments were performed at starting concentrations of 25 mM in $CDCl_3$ at 35 °C. The data for the recognition-mediated reaction between **1** and **3a** affording **5a** are shown as filled blue circles, the data for the reaction between **1** and **3a** in the presence of benzoic acid (100 mM) are shown as filled red triangles and the data for the reaction between **1** and **3a** in the presence of the appropriate template are shown as filled green diamonds. The data for the control reaction between **1** and **3b**, forming (a) *endo*-**5b** and (b) *exo*-**5b**, are shown as open blue circles. The solid lines represent the best fit of the data for the recognition-mediated and control reactions to the appropriate kinetic model.

In order to gain further insight into the processes operating within this system, we performed doping and inhibitor experiments. Neither the addition of preformed template nor the addition of benzoic acid (100 mol%) have any significant impact (Fig. 6a, red triangles and green diamonds respectively) on the formation of *endo*-**5a**. By contrast, although the rate of formation of *exo*-**5a** is diminished significantly (Fig 6b, red triangles) in the presence of a competitive inhibitor, benzoic acid (100 mol%), the addition of preformed template (29 mol%) has virtually no effect (Fig 6b, green diamonds). These results suggest that the formation of *endo*-**5a** does not occur to any significant extent through either of the two recognition-mediated pathways open to this system. By contrast, the formation of *exo*-**5a** appears to occur exclusively through the [**A**·**B**] complex reaction channel (Channel 2, Fig. 1).

In order to confirm these hypotheses, we again undertook molecular modelling studies on the structures of *exo*-**5a** and *endo*-**5a**. We performed a series of molecular mechanics calculations to explore the conformational space open to both *exo*-**5a** and *endo*-**5a**. Monte Carlo conformational searches were performed from a variety of starting conformations using the AMBER* forcefield and the GB/SA solvation model for $CHCl_3$. Representative minimum energy structures are shown for *exo*-**5a** and *endo*-**5a** in Fig. 7a, b and c respectively. It is clear from the results of these calculations that *exo*-**5a** contains several low energy conformations which possess the classical doubly

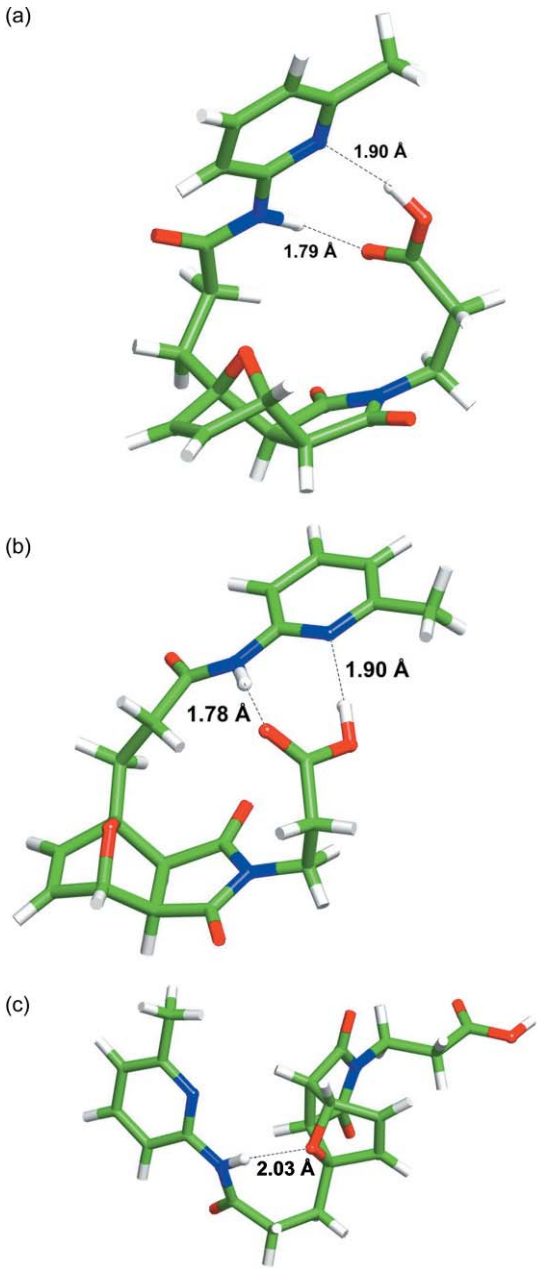


Fig. 7 Stick models of representative minimum energy conformations of *exo*-**5a** ((a) and (b)) and *endo*-**5a** (c) calculated using the AMBER* forcefield. Carbon atoms are coloured green, oxygen atoms are coloured red, nitrogen atoms are coloured blue, hydrogen atoms are coloured white. Hydrogen bonds are represented by dashed lines and H · · · Acceptor atom distances are given in Å.

hydrogen-bonded association between the amidopyridine and the carboxylic acid shown schematically in Scheme 1b. This observation is clearly consistent with the results of the kinetic experiments which suggest that the formation of *exo*-**5a** occurs through the [**A**·**B**] complex reaction channel (Channel 2, Fig. 1).

The fact that the stabilising interactions which serve to assemble the [**A**·**B**] complex prior to reaction live on in the product opens the question of whether the increase in reaction rate and extent of reaction is a result of kinetic or thermodynamic factors. Kinetic simulation and fitting of the experimental data readily provides an answer to this question. The kinetic model used[16,17] in the simulation and fitting of the kinetic data is shown in Fig. 8 and the results of these calculations are depicted as the solid lines in Fig. 6. Simultaneous excellent fits of the experimentally observed data to the model are achieved using the parameters shown in Fig. 8. Full details of the kinetic simulation and fitting are given in the Supplementary Information.

Bimolecular Reactions (Channel 1, Fig. 1)

$$\mathbf{1} + \mathbf{3a} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \textit{endo}\text{-}\mathbf{5a} \qquad k_1 = (5.040 \pm 0.081)\times10^{-5}\ M^{-1}\ s^{-1},\ k_2 = (8.560 \pm 6.300)\times10^{-8}\ M^{-1}\ s^{-1}$$

$$\mathbf{1} + \mathbf{3a} \underset{k_4}{\overset{k_3}{\rightleftharpoons}} \textit{exo}\text{-}\mathbf{5a} \qquad k_3 = (7.220 \pm 0.090)\times10^{-5}\ M^{-1}\ s^{-1},\ k_4 = (3.870 \pm 0.500)\times10^{-6}\ M^{-1}\ s^{-1}$$

[A•B] Complex Reactions (Channel 2, Fig. 1)

$$\mathbf{1} + \mathbf{3a} \overset{K_a}{\rightleftharpoons} [\mathbf{1}\bullet\mathbf{3a}] \qquad K_a = 250\ M^{-1}$$

$$[\mathbf{1}\bullet\mathbf{3a}] \underset{k_6}{\overset{k_5}{\rightleftharpoons}} \textit{endo}\text{-}\mathbf{5a} \qquad k_5 = (6.130 \pm 0.340)\times10^{-6}s^{-1},\ k_6 = (3.240 \pm 0.006)\times10^{-5}s^{-1}$$

$$[\mathbf{1}\bullet\mathbf{3a}] \underset{k_8}{\overset{k_7}{\rightleftharpoons}} \textit{exo}\text{-}\mathbf{5a} \qquad k_7 = (2.590 \pm 0.003)\times10^{-5}s^{-1},\ k_8 = (1.260 \pm 4.700)\times10^{-7}s^{-1}$$

***exo*-5a**
Kinetic EM=k_7/k_3=0.36 M
Thermodynamic EM=k_7k_4/k_8k_3=11.00 M

***endo*-5a**
Kinetic EM=k_5/k_1=0.12 M
Thermodynamic EM=k_5k_2/k_6k_1=0.0003 M

Fig. 8 Kinetic model used in the fitting of the experimental concentration–time data for the formation of *exo*-**5a** and *endo*-**5a**. For full details of the kinetic simulation and fitting protocol, see electronic supplementary information (ESI).†

It is instructive to calculate (Fig. 8) the kinetic and thermodynamic EM values for this system. These values reveal that the principal effect of recognition on the formation of *exo*-**5a** is a thermodynamic one—the thermodynamic EM (11 M) is far larger than the kinetic EM (0.36 M) for this cycloadduct. Therefore, the recognition between the carboxylic acid and the amidopyridine serves to stabilise the product ground state of *exo*-**5a** and therefore predispose[19] the reaction to form this cycloadduct in preference to the corresponding *endo* adduct. It is worth noting that the results of both the molecular modelling and the kinetic simulation and fitting suggest strongly that recognition processes are not involved to any significant extent in the formation of *endo*-**5a**. Indeed, the value of the kinetic EM (0.12 M) is barely higher than the reaction concentration itself.

Once again, a comparison of the reaction between **1** and **3a** with the reaction between **6** and **3a** (Scheme 3) is instructive. Although the outcomes of these two reactions are similar—both reactions form the *exo* cycloadduct selectively and rapidly—the mechanism by which they achieve this effect is entirely different. At thermodynamic equilibrium, the reaction between **1** and **3a** would have an *exo* : *endo* ratio of 6000 : 1, *i.e.* the reaction is effectively *exo* specific. This specificity derives from the selective stabilisation of the *exo*-**5a** ground state through two hydrogen bonds between the amidopyridine and the carboxylic acid present in the molecule. Conversely, the role played by recognition in the formation of *exo*-**7** is primarily a kinetic one—the recognition present in the [**6**·**3a**] complex stabilises the transition state leading to the *exo* cycloadduct selectively.

Conclusions

The systems discussed in this paper serve to illustrate that, whilst recognition-mediated control and acceleration of chemical reactions may be an attractive goal for the supramolecular chemist, the design and implementation of systems capable of exploiting recognition to this end is problematic—even with apparently simple compounds. We have demonstrated that, in general, for reagents bearing complementary recognition sites, given sufficient conformational freedom, reaction will occur through the binary [**A**·**B**] complex. Although this conclusion is expected on the basis of entropic considerations, the different role of recognition—kinetic (transition state stabilisation) or thermodynamic (product ground state stabilisation—in apparently similar systems (*exo*-**5a** *vs.* *exo*-**7**)) is much harder to predict. It is worth remembering that in all of the work presented here the biggest constitutional difference between any of the systems studied is a single CH_2 group. Therefore, the design and implementation of efficient and selective replication processes in structurally simple chemical systems remains an important and challenging goal.

Experimental

General procedures

All commercially available substrates, reagents, and solvents were used without further purification. ^{1}H NMR spectra were recorded at 500 MHz using either a Varian UNITYplus spectrometer or a Bruker Avance 500 spectrometer. ^{13}C NMR spectra were recorded at 125 MHz using a Bruker Avance 500 spectrometer. Coupling constants (J) are given in Hz. High-resolution mass spectra were recorded using either electron impact ionisation (EI) on a VG Autospec instrument or using electrospray ionisation (ES) on a Micromass LCT spectrometer operating in negative ion mode by the Mass Spectrometry Service at the University of St Andrews. Elemental analyses were carried out by the Microanalysis Service at the University of St Andrews. Melting points were recorded using an Electrothermal 9200 melting point apparatus and are uncorrected.

Molecular mechanics calculations were carried out using the AMBER* forcefield as implemented in Macromodel (Version 7.1, Schrodinger Inc., 2000) running on a Linux workstation.

3-Furan-2-yl-*N*-(6-methyl-pyridin-2-yl)-propionamide 1

3-(Furan-2-yl)propionic acid (3.00 g, 21.4 mmol) was dissolved in dry CH_3CN (35 cm^3) and pyridine (1.73 cm^3, 21.4 mmol) was added. This solution was placed under an atmosphere of N_2 and cooled to −15 °C. Cyanuric fluoride (0.72 g, 8.6 mmol) was

injected slowly and within 5 minutes a white precipitate formed. The reaction was stirred at −15 °C for a further 25 minutes and, after this time, the reaction mixture was diluted with CH_2Cl_2 (100 cm^3) and washed directly with brine. The organic layer was collected, dried over $MgSO_4$ and the solvent removed *in vacuo* to afford the crude acid fluoride as a yellow oil. The acid fluoride (2.53 g, 17.8 mmol) was dissolved in CH_2Cl_2 (60 cm^3) and 2-amino-6-picoline (1.90 g, 17.6 mmol) was added. The reaction mixture was stirred at room temperature for 16 hours. Standard work-up followed by column chromatography (SiO_2: EtOAc/hexane, 2 : 5) afforded **1** (2.85 g, 58% based on 3-(furan-2-yl)propionic acid) as a brown oil; CHN calculated for $C_{13}H_{14}N_2O_2$ C, 67.8; H, 6.1; N, 12.2, found C, 67.6; H, 6.3; N, 12.1%. ν_{max}(KBr)/cm^{-1} 3278, 2923, 2364, 1686, 1603, 1578, 1534, 1458, 1400, 1302, 1235, 1156, 1077, 1015; δ_H (500 MHz, $CDCl_3$) 8.07 (br s, 1H), 8.02 (d, 1H, $J = 8.3$ Hz), 7.60 (t, 1H, $J = 7.8$ Hz), 7.31 (d, 1H, $J = 1.2$ Hz), 6.89 (d, 1H, $J = 7.49$ Hz), 6.27 (dd, 1H, $J = 1.9$ and 3.1 Hz), 6.06 (dd, 1H, $J = 0.6$ and 3.1 Hz), 3.07 (t, 2H, $J = 7.52$ Hz), 2.72 (t, 2H, $J = 7.5$ Hz), 2.44 (s, 3H); δ_C (125 MHz, $CDCl_3$) 170.2, 156.6, 154.0, 150.5, 141.2, 138.8, 119.2, 110.9, 110.2, 105.5, 35.7, 23.7, 23.5; MS (EI+) *m/z* ($[M]^{+\bullet}$, 100%); HRMS (EI+) calculated for $C_{13}H_{14}N_2O_2$ $[M]^{+\bullet}$ 230.1055, found 230.1054.

(2,5-Dioxo-2,5-dihydro-pyrrol-1-yl)-acetic acid 2a

Glycine (5.00 g, 66.6 mmol) and maleic anhydride (6.53 g, 66.6 mmol) were dissolved in acetic acid (190 cm^3). The reaction mixture was stirred at room temperature for 12 hours, under an N_2 atmosphere. The resulting suspension was then refluxed for a further 8 hours to give a clear solution. The solvent was removed *in vacuo* and the residue was purified by flash column chromatography (SiO_2: CH_2Cl_2/glacial acetic acid, 95 : 5). Recrystallisation from CH_2Cl_2/hexane afforded the title compound (1.15 g, 23%) as a colourless solid: mp 112–113 °C; CHN calculated for $C_6H_5NO_4$ C, 46.5; H, 3.3; N, 9.0, found C, 46.3; H, 3.2; N, 8.9%. ν_{max}(KBr)/cm^{-1} 3101, 2726, 1755, 1446, 1397, 1193, 699, 675, 632; δ_H (300 MHz, $CDCl_3$) 8.01 (1H, br s), 6.80 (2H, s), 4.32 (2H, s); δ_C (75 MHz, $CDCl_3$) 172.6, 169.7, 134.6, 38.3; MS (ES+) *m/z* 156 ($[M]^+$, 11%) 138 (11), 110 (100), 82 (68), 54 (49); HRMS (ES+) calculated for $C_6H_5NO_4$ $[M]^+$ 155.0218, found 155.0218.

(2,5-Dioxo-2,5-dihydro-pyrrol-1-yl)acetic acid methyl ester 2b

Caesium carbonate (1.01 g, 3.2 mmol) was added to a solution of **2a** (1.03 g, 6.6 mmol) and methyl iodide (1.83 g, 12.9 mmol) in dry DMF (*ca.* 20 cm^3). The solution was stirred for 16 hours in the absence of light at room temperature, under an inert N_2 atmosphere. The reaction mixture was partitioned between water (50 cm^3) and EtOAc (3×75 cm^3). The organic extracts were combined and dried over $MgSO_4$ before being reduced to dryness *in vacuo*. The crude product was purified by flash column chromatography (SiO_2: hexane/EtOAc, 3 : 7) yielding the title compound as a pale yellow oil (1.10 g, 98%): CHN calculated for $C_7H_7NO_4$ C, 49.7; H, 4.2; N, 8.3, found C, 49.9; H, 4.1; N, 8.5%. ν_{max} (neat)/cm^{-1} 3102, 2957, 1715, 1432, 1220, 1152, 908, 830, 698; δ_H (300 MHz, $CDCl_3$) 6.78 (2H, s), 4.28 (2H, s), 3.74 (3H, s); δ_C (75 MHz, $CDCl_3$) 169.7, 167.6, 134.5, 52.7, 38.4; MS (ES+) *m/z* 169 ($[M]^+$, 42%) 110 (100), 82 (22), 54 (14), 43 (11); HRMS (ES+) calculated for $C_7H_7NO_4$ $[M]^+$ 169.0375, found 169.0382.

3-(2,5-Dioxo-2,5-dihydro-pyrrol-1-yl)propionic acid 3a

Acid **3a** (1.65 g, 35%, colourless solid) was prepared from β-alanine (2.50 g, 28.0 mmol) and maleic anhydride (2.74 g, 28.0 mmol) in the same manner as that described for **2a**: mp 92–94 °C; CHN calculated for $C_7H_7NO_4$ C, 49.7; H, 4.2; N, 8.3, found C, 49.8; H, 4.0; N, 8.4%. ν_{max} (KBr)/cm^{-1} 3455, 3096, 2949, 2646, 1714, 1694, 1453, 1412, 945, 832; δ_H (300 MHz, $CDCl_3$) 10.26 (1H, br s), 6.70 (2H, s), 3.81 (2H, t, $J = 7.0$ Hz), 2.68 (2H, t, $J = 7.0$ Hz); δ_C (75 MHz, $CDCl_3$) 176.6, 170.4, 134.3, 33.3, 32.54; MS (ES+) *m/z* 169 ($[M]^+$, 8%), 151 (30), 123 (97), 110 (100), 96 (39), 82 (92); HRMS (ES+) calculated for $C_7H_7NO_4$ $[M]^+$ 169.0375, found 169.0378.

3-(2,5-Dioxo-2,5-dihydro-pyrrol-1-yl)propionic acid methyl ester 3b

Ester **3b** (0.52 g, 96%, yellow oil) was prepared from **3a** (0.50 g, 3.0 mmol) and methyl iodide (0.37 cm^3, 5.9 mmol) in the same manner as that described for **2b**: CHN calculated for $C_8H_9NO_4$ C, 52.5; H, 5.0; N, 7.6, found C, 52.3; H, 5.2; N, 7.4%. ν_{max}(neat)/cm^{-1} 3464, 3099, 3001, 2954, 2851, 1705, 1585, 1447, 1409, 959; δ_H (300 MHz, $CDCl_3$) 6.70 (2H, s), 3.82 (2H, t, $J = 7.0$ Hz), 3.66 (3H, s,), 2.63 (2H, t, $J = 7.0$ Hz); δ_C (75 MHz, $CDCl_3$) 171.2, 170.4, 134.3, 52.0, 33.6, 32.7; MS (ES+) *m/z* 183 ($[M]^+$, 5%), 151 (29), 123 (83), 110 (100), 96 (20), 82 (73), 54 (59); HRMS (ES+) calculated for $C_8H_9NO_4$ $[M]^+$ 183.0532, found 183.0529.

exo-{1-[2-(6-Methyl-pyridin-2-ylcarbamoyl)-ethyl]-3,5-dioxo-10-oxa-4-aza-tricyclo[5.2.1.$0^{2,6}$]dec-8-en-4-yl}-acetic acid *exo*-4a

1 (0.5 cm^3 of a 50 mM solution in $CDCl_3$) and **2a** (0.5 cm^3 of a 50 mM solution in $CDCl_3$) were mixed together in an NMR tube and heated at 35 °C for six days. **4a** precipitated from the reaction mixture as a colourless crystalline solid, which was collected by filtration and washed with copious amounts of $CHCl_3$; mp: 81.3–82.4 °C; CHN calculated for $C_{19}H_{19}N_3O_6$ C, 59.2; H, 5.0; N, 10.9, found C, 59.0; H, 4.8; N, 10.7%. δ_H (500 MHz, $CDCl_3$) 11.25 (br s, 1H), 8.24 (d, 1H, $J = 8.5$ Hz), 7.72 (t, 1H, $J = 8.1$ Hz), 6.92 (d, 1H, $J = 7.7$ Hz), 6.50 (dd, 1H, $J = 1.7$ and 5.8 Hz), 6.35 (d, 1H, $J = 5.8$ Hz), 5.20 (d, 1H, $J = 1.7$ Hz), 4.24 (d, 1H, $J = 16.6$ Hz), 4.18 (d, 1H, $J = 16.6$ Hz), 3.03 (d, 1H, $J = 6.4$ Hz), 2.94 (d, 1H, $J = 6.4$ Hz), 2.79–2.35 (m, 4H), 2.48 (s, 3H); δ_C (125 MHz, $CDCl_3$) 175.6, 174.2, 172.1, 171.9, 154.5, 150.6, 141.5, 138.6, 137.6, 119.4, 112.6, 91.2, 80.3, 50.8, 49.0, 40.2, 32.3, 24.2, 21.2; MS (ES−) *m/z* 384 ($[M - H]^+$, 100%); HRMS (ES-) calculated for $C_{19}H_{19}N_3O_6$ $[M - H]^+$ 384.1196, found 384.1204.

exo-3-{1-[2-(6-Methyl-pyridin-2-ylcarbamoyl)-ethyl]-3,5-dioxo-10-oxa-4-aza-tricyclo[5.2.1.$0^{2,6}$]dec-8-en-4-yl}-propionic acid *exo*-5a

A mixture of **1** (0.5 cm^3 of a 50 mM solution in $CDCl_3$) and **3a** (0.5 cm^3 of a 50 mM solution in $CDCl_3$) in an NMR tube was heated at 35 °C for six days. *exo*-**5a** (4.33 mg, 45%) precipitated from the reaction mixture as a colourless crystalline solid which was filtered and washed with copious amounts of $CHCl_3$; mp: 166.9–167.0 °C; CHN calculated for $C_{20}H_{21}N_3O_6$ C, 60.1; H, 5.3; N, 10.5, found C, 60.3; H, 5.5; N, 10.7%. δ_H (500 MHz, $CDCl_3$) 10.47 (br s, 1H), 8.04 (d, 1H, $J = 8.3$ Hz), 7.66 (dd, 1H, $J = 7.8$ and 8.2 Hz), 6.88 (d, 1H, $J = 7.4$ Hz), 6.48 (dd, 1H, $J = 1.7$ and 5.6 Hz), 6.35 (d, 1H, $J = 5.6$ Hz), 5.23 (d, 1H, $J = 1.7$ Hz), 4.14 (ddd, 1H, $J = 4.2$, 10.5 and 14.6 Hz), 3.58 (dt, 1H, $J = 4.5$ Hz), 3.00–2.47 (m, 6H), 2.93 (d, 1H, $J = 6.1$ Hz), 2.76 (d, 1H, $J = 6.4$ Hz), 2.44 (s, 3H); δ_C (125 MHz, $CDCl_3$) 176.2, 175.75, 174.8, 172.2, 155.2, 151.0, 140.4, 139.4, 137.25, 119.3, 112.6, 90.55, 80.6, 50.1, 49.5, 35.1, 33.1, 31.9, 25.1, 22.0; MS (ES−) *m/z* 398 ($[M - H]^+$, 100%). HRMS (ES −) calculated for $C_{20}H_{21}N_3O_6$ $[M - H]^+$ 398.1352, found 398.1340.

Acknowledgements

This work was supported by the BBSRC (Grant B11855), the EPSRC (DTA award to EK) and the University of St Andrews. We thank Professor G. von Kiedrowski for providing us with

a copy of his SimFit program and Melanja Smith for technical assistance.

References

1 For reviews see: A. Robertson, A. J. Sinclair and D. Philp, *Chem. Soc. Rev.*, 2000, **29**, 141–152; E. A. Wintner and J. Rebek Jr., *Acta Chem. Scand.*, 1996, **50**, 469–485; M. Famulok, J. S. Nowick and J. Rebek Jr., *Acta Chem. Scand.*, 1992, **46**, 315–324; L. E. Orgel, *Nature*, 1992, **358**, 203–209.

2 R. Issac, Y.-W. Ham and J. A. Chmielewski, *Curr. Opin. Struct. Biol.*, 2001, **11**, 458–463; G. F. Joyce, *Nature*, 2002, **418**, 214–221; J. C. Chaput and J. W. Szostak, *J. Am. Chem. Soc.*, 2003, **125**, 9274–9275; S. Miyakawa and J. P. Ferris, *J. Am. Chem. Soc.*, 2003, **125**, 8202–8208; S. L. Miller and L. E. Orgel, *The Origins of Life on Earth*, Prentice Hall, Engelwood Cliffs, NJ, 1974.

3 L. J. Prins, D. N. Reinhoudt and P. Timmerman, *Angew. Chem., Int. Ed.*, 2001, **40**, 2382–2426; F. Diederich and P. J. Stang, *Template Directed Synthesis*, Wiley-VCH, Weinheim, 2000; D. Philp and J. F. Stoddart, *Angew. Chem., Int. Ed.*, 1996, **35**, 1154–1196. For applications see:; C. Niemeyer, *Angew. Chem., Int. Ed.*, 2001, **40**, 4128–4158; J.-M. Lehn, *Chem. Eur. J.*, 2000, **6**, 2097–2102.

4 Recent examples include: J. M. Quayle, A. M. Z. Slawin and D. Philp, *Tetrahedron Lett.*, 2002, **43**, 7229–7233; R. Isaac and J. Chmielewski, *J. Am. Chem. Soc.*, 2002, **124**, 6808–6809; X. Li and J. Chmielewski, *J. Am. Chem. Soc.*, 2003, **125**, 11820–11821; X. Li and J. Chmielewski, *Org. Biomol. Chem.*, 2003, **1**, 901–904; V. C. Allen, D. Philp and N. Spencer, *Org. Lett.*, 2001, **3**, 777–780; D. H. Lee, J. R. Granja, J. A. Martinez, K. Severin and M. R. Ghadiri, *Nature*, 1996, **382**, 525–528; S. Matsumura, T. Takahashi, A. Ueno and H. Mihara, *Chem. Eur. J.*, 2003, **9**, 4829–4837.

5 G. von Kiedrowski, *Angew. Chem., Int. Ed. Engl.*, 1986, **25**, 932–935.

6 For comprehensive kinetic analyses see: G. von Kiedrowski, *Bioorganic Chemistry Frontiers*, Vol. 3, Springer-Verlag, Berlin, Heidelberg, 1993, pp. 115–146; D. N. Reinhoudt, D. M. Rudkevich and F. de Jong, *J. Am. Chem. Soc.*, 1996, **118**, 6880–6889.

7 For recognition motifs based on hydrogen bonding see: T. Steiner, *Angew. Chem., Int. Ed.*, 2002, **41**, 48–76; G. Cooke and V. M. Rotello, *Chem. Soc. Rev.*, 2002, **31**, 275–286; D. C. Sherrington and K. Taskinen, *Chem. Soc. Rev.*, 2001, **30**, 83–93.

8 D. Philp and A. Robertson, *Chem. Commun.*, 1998, 879; C. A. Booth and D. Philp, *Tetrahedron Lett.*, 1998, 6987; A. Robertson, D. Philp and N. Spencer, *Tetrahedron*, 1999, **55**, 11365; R. M. Bennes, B. M. Kariuki, K. D. M. Harris, D. Philp and N. Spencer, *Org. Lett.*, 1999, **1**, 1087; R. M. Bennes, M. Sapro-Babiloni, W. C. Hayes and D. Philp, *Tetrahedron Lett.*, 2001, **42**, 2377; S. J. Howell, D. Philp and N. Spencer, *Tetrahedron*, 2001, **57**, 4945.

9 R. J. Pearson, E. Kassianidis and D. Philp, *Tetrahedron Lett.*, 2004, **45**, 4777–4780.

10 W. Oppolzer, *Angew. Chem., Int. Ed. Engl.*, 1977, **16**, 10–23; G. Brieger and J. N. Bennett, *Chem. Rev.*, 1980, **80**, 63–97; D. Craig, *Chem. Soc. Rev.*, 1987, **16**, 187–238; S. C. Hirst and A. D. Hamilton, *J. Am. Chem. Soc.*, 1991, **113**, 382–238.

11 K. Mikami and M. Yamanaka, *Chem. Rev.*, 2003, **103**, 3369–3400; M. H. Todd, *Chem. Soc. Rev.*, 2002, **31**, 211–222.

12 Previous efforts by other groups focused purely on rate acceleration of the binding event *via* hydrogen bonding recognition. We opted for cycloaddition reactions yielding diastereoisomeric products, since they simultaneously provide scope for selectivity investigations[4].

13 For replication of more complex entities see: J. W. Szostak, D. P. Bartel and P. L. Luisi, *Nature*, 2001, **409**, 387–390; R. Wick, P. Walde and P. L. Luisi, *J. Am. Chem. Soc.*, 1995, **117**, 1435–1436; A. Veronese and P. L. Luisi, *J. Am. Chem. Soc.*, 1998, **120**, 2662–2663.

14 J. D. Winkler, *Chem. Rev.*, 1996, **96**, 167–176; C. Cativiela, J. I. Garcia, J. A. Mayoral and L. Salvatella, *Chem. Soc. Rev.*, 1996, **25**, 209–218; H. B. Kagan and O. Riant, *Chem. Rev.*, 1992, **92**, 1007–1019.

15 F. H. Allen, W. D. S. Motherwell, P. R. Raithby, G. P. Shields and R. Taylor, *New J. Chem.*, 1999, **23**, 25–34; T. Steiner, *Acta Crystallogr., Sect. B*, 2001, **57**, 103–106.

16 Crystallographic data: Several attempts to obtain good quality crystals of ***exo*-4a** were undertaken. The best attempt yielded poor quality data, hence the limited accuracy of the refinement. However, the stereochemistry of ***exo*-4a** is unambiguous, as is the dimer-like nature of the extended three-dimensional structure. A single crystal of ***exo*-4a**·$(CHCl_3)_{0.5}$ ($0.2 \times 0.4 \times 0.8$ mm^3) suitable for X-ray analysis was obtained from $CHCl_3$ by slow evaporation. The single crystal was mounted on a quartz fiber and diffraction data were collected at 125 K on a Bruker SMART CCD diffractometer with graphite-monochromated Mo Kα radiation ($\lambda = 0.71073$ Å). An absorption correction was made using SADABS. The structure was solved by direct methods and refined by full-matrix least squares on F^2 by using SHELXTL. All non-hydrogen atoms were refined with isotropic displacement parameters. All hydrogen atoms were placed in calculated positions. Final refinement details for ***exo*-4a**·$(CHCl_3)_{0.5}$ ($C_{19.50}H_{19.50}C_{11.50}N_3O_6$, $M_w = 445.06$); space group $P2_1/c$, monoclinic; unit cell dimensions $a = 12.342(17)$ Å, $b = 28.44(4)$ Å, $c = 13.777(19)$ Å, $\beta = 107.291(18)°$, $V = 4617(11)$ Å^3; $Z = 8$ (two crystallographically independent molecules), $D_{calc} = 1.281$ Mg m^{-3}; total 17875 reflections were measured and 4770 were independent. Final R indices (435 observed reflections [$I > 2\sigma(I)$]) $R1 = 0.1620$, $wR2 = 0.2765$ and GOF = 0.818. CCDC reference number 230704. See http://www.rsc.org/suppdata/ob/b4/b406862a/ for crystallographic data in .cif or other electronic format.

17 SimFit, A program for the analysis of Kinetic Data, version 1.0, G. von Kiedrowski, 1994.

18 F. C. Lightstone and T. C. Bruice, *J. Am. Chem. Soc.*, 1996, **118**, 2595–2605; F. C. Lightstone and T. C. Bruice, *J. Am. Chem. Soc.*, 1997, **119**, 9103–9113; F. C. Lightstone and T. C. Bruice, *Bioorg. Chem.*, 1998, **26**, 193–199; T. C. Bruice and F. C. Lightstone, *Acc. Chem. Res.*, 1999, **32**, 127–136.

19 D. J. Cram, *Angew. Chem., Int. Ed. Engl.*, 1988, **27**, 1009; E. C. Constable, *Comprehensive Supramolecular Chemistry*, ed. J. L. Atwood, J. E. D. Davis, D. D. Macnicol and F. Vögtle, Elsevier, New York, 1996, vol. 9, p. 218; S. J. Rowan, D. G. Hamilton, P. A. Brady and J. K. M. Sanders, *J. Am. Chem. Soc.*, 1997, **119**, 2578; S. J. Rowan and J. K. M. Sanders, *J. Org. Chem.*, 1998, **63**, 1536.