PRELIMINARY COMMUNICATION

Open Access

Unravelling a fulvene based Replicator: Experiment and Theory in Interplay

Arne Dieckmann[1*], Sabrina Beniken[2], Christian Lorenz[3], Nikos L Doltsinis[2], Günter von Kiedrowski[1]

Abstract

A self-replicating system based on a cycloaddition of a fulvene derivative and a maleinimide is investigated using a two-pronged approach combining NMR spectroscopy with computer simulations. In the course of the reaction, two diastereomers are formed with identical rates in the absence of replication. When replication is enabled, a network emerges in which one diastereomer takes over the resources as a "selfish" autocatalyst while exploiting the competitor as a weak "altruist". The structure and dynamics of the reaction network is studied using 1 D and 2 D NMR techniques supported by dynamically averaged *ab initio* chemical shifts and *ab initio* molecular dynamics simulations. It is shown that this combination is a powerful means to understand the observed experimental behaviour in great detail.

The field of self-replication chemistry [1-23] currently aims to understand and to control the energy landscape determining the type of autocatalytic growth, which distinguishes the chemistry of parabolic coexistence[24] from the biological physics building on the Darwinian principle [25]. Small organic replicators, especially those which are the design offspring of the system introduced by Wang and Sutherland [16], offer particular advantages in the quest for a more detailed understanding of replicator dynamics: They are large enough to exhibit autocatalysis coupled to information transfer and they are small enough to be treatable from first principles. Here we introduce a replicator which utilizes a fulvene-based Diels-Alder reaction (see Fig. 1) and show that the conjunction of ab initio molecular dynamics (AIMD) as developed by Car and Parrinello[26] with NMR kinetics supported by computed chemical shifts and 2D-NMR methods allows us to decipher the structural and energetic rationale behind the observed behaviour, while static computational methods currently used in the field did not reproduce the experimental data.
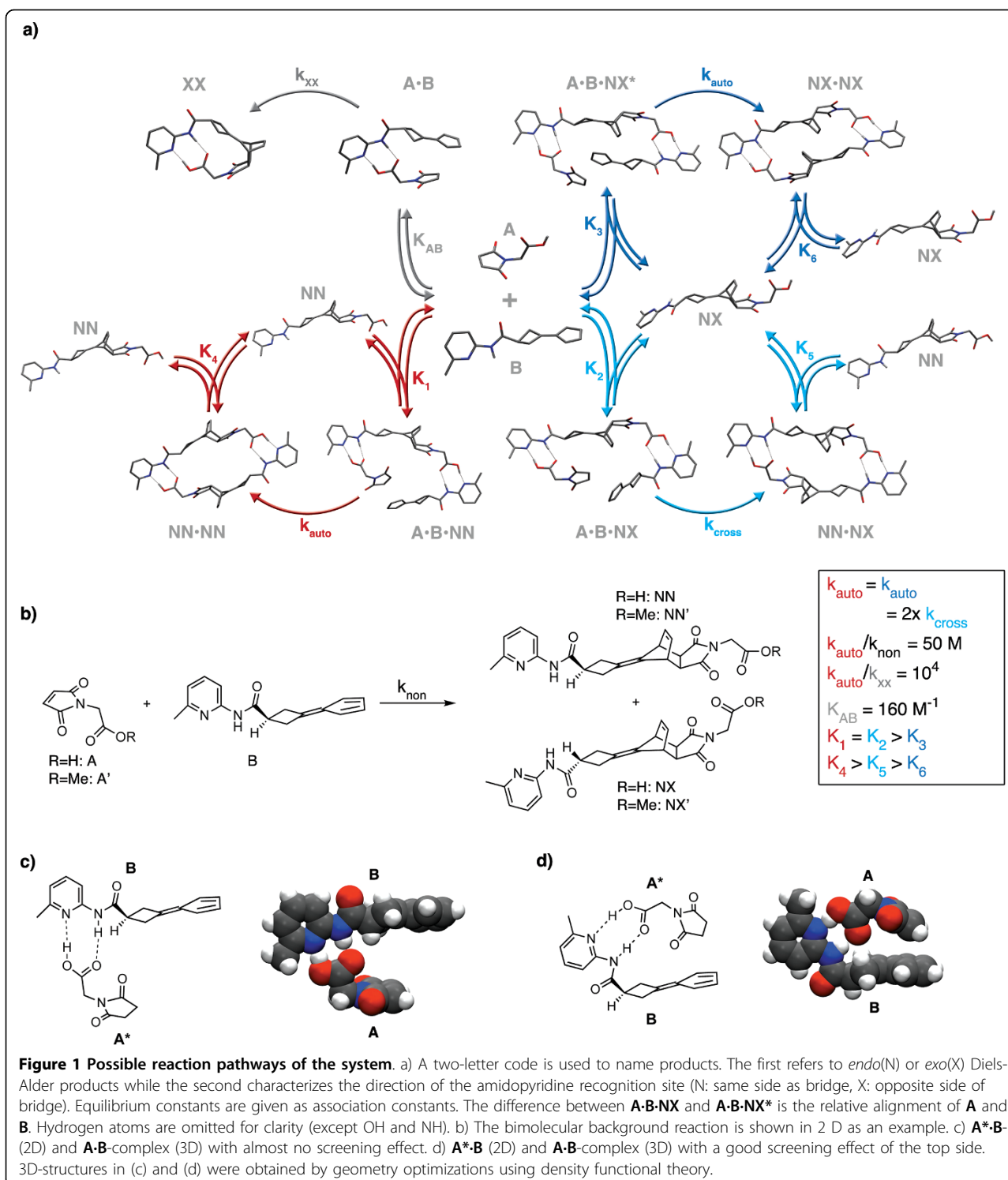
We selected fulvene chemistry for replicator construction for two reasons: First, fulvene chemistry allows facile variation of the diene part. Second, with respect to

an earlier described system [17], we aimed to arrive at a nonchiral reaction product in order to simplify the kinetic analysis. In principle, there are four possible products – two *endo* and two *exo* diastereomers – from the reaction of **A** and **B**, three of them being observed by NMR (Fig. 2). Only *endo*-products (**NN**, **NX**) are able to replicate via a termolecular complex. The kinetics of the reaction of **A** with **B** were studied by time-resolved ^{1}H-NMR (600 MHz) in CDCl$_3$ at 293 K (Fig. 2a). One main product is apparent whose curve exhibits a sigmoidal shape caused by an induction period, which is typical of autocatalytic reactions. To prove the specificity of the catalysis 2 eq benzoic acid was added as a competitive inhibitor, resulting in a deceleration of the reaction. On the other hand, when a 10% product mixture – corresponding to almost pure main product – was added as a catalyst at the start, the production of the main product was accelerated and the induction period almost vanished [see Additional file 1]. We observed that product precipitation is kinetically hindered even for days. Once precipitated however (e.g. by solvent evaporation) the products could not be dissolved. The latter prevented chromatographic separation and as a consequence discrete addition of each product. Another measurement was carried out using the methylester **A'** to disable molecular recognition in the system, allowing independent measurement of the background reaction. As expected, the resulting concentration-time data does not show an induction period and the reaction proceeds

* Correspondence: arne.dieckmann@rub.de
[1]Lehrstuhl für Organische Chemie I, Bioorganische Chemie, Ruhr-Universität Bochum, 44780, Bochum, Germany
Full list of author information is available at the end of the article

Figure 1 Possible reaction pathways of the system. a) A two-letter code is used to name products. The first refers to *endo*(N) or *exo*(X) Diels-Alder products while the second characterizes the direction of the amidopyridine recognition site (N: same side as bridge, X: opposite side of bridge). Equilibrium constants are given as association constants. The difference between **A·B·NX** and **A·B·NX*** is the relative alignment of **A** and **B**. Hydrogen atoms are omitted for clarity (except OH and NH). b) The bimolecular background reaction is shown in 2 D as an example. c) **A*·B**-(2D) and **A·B**-complex (3D) with almost no screening effect. d) **A*·B** (2D) and **A·B**-complex (3D) with a good screening effect of the top side. 3D-structures in (c) and (d) were obtained by geometry optimizations using density functional theory.

much more slowly (Fig. 2a). Chromatographic separation by high performance liquid chromatography (HPLC) and identification of the background reaction products by rotating frame nuclear Overhauser enhancement spectroscopy (ROESY) revealed that only **NN** and

NX are formed. There is a slight diastereoselectivity of 2:3 for **NX**, which must be due to subtle electronic effects, as there are no steric interactions disfavouring the formation of **NN**. However, if the formation of an unreactive **A*·B**-complex as a model for the real
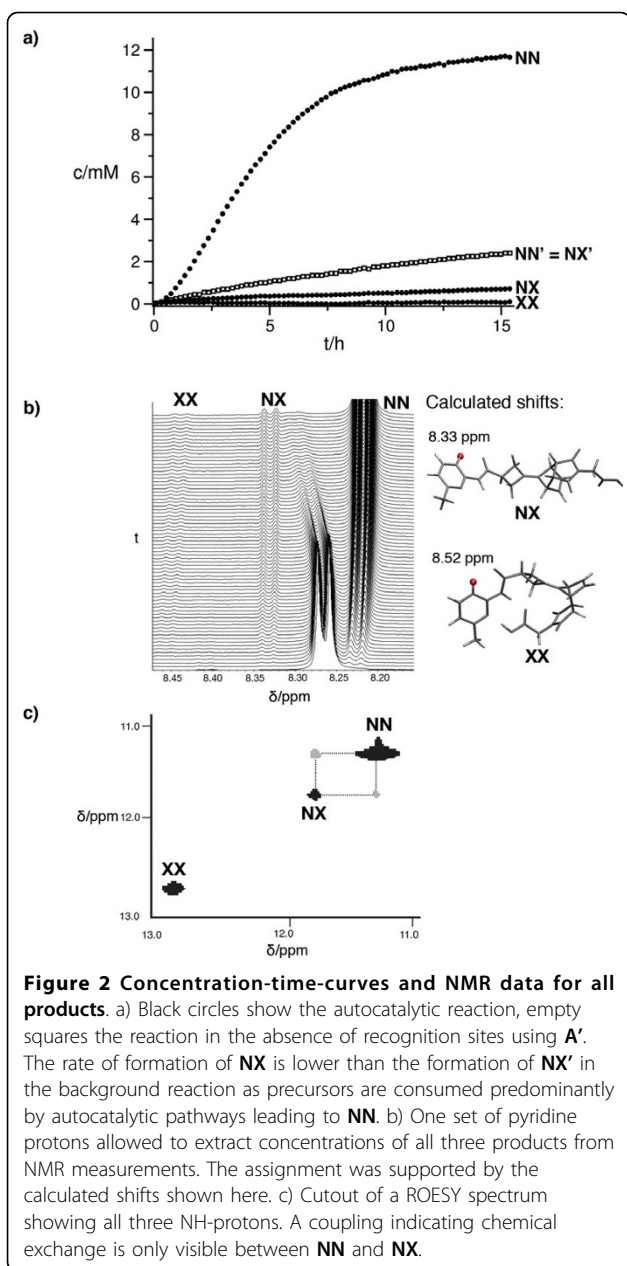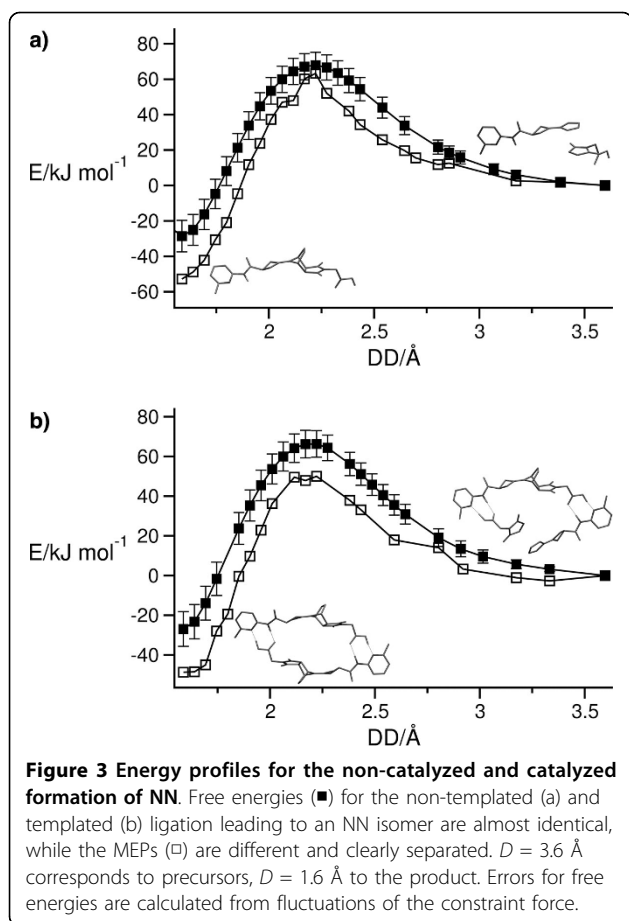
Figure 2 Concentration-time-curves and NMR data for all products. a) Black circles show the autocatalytic reaction, empty squares the reaction in the absence of recognition sites using **A'**. The rate of formation of **NX** is lower than the formation of **NX'** in the background reaction as precursors are consumed predominantly by autocatalytic pathways leading to **NN**. b) One set of pyridine protons allowed to extract concentrations of all three products from NMR measurements. The assignment was supported by the calculated shifts shown here. c) Cutout of a ROESY spectrum showing all three NH-protons. A coupling indicating chemical exchange is only visible between **NN** and **NX**.

A·B-complex was allowed by adding one equivalent of succinimide **A*** (Fig. 1c/d) both isomers were formed with identical rates. This can be explained by a more effective screening of one side of the fulvene in the **A*·B**-complex rotamer depicted in Fig. 1d. It is evident from the respective three-dimensional models of the real **A·B** complex that the same argument holds for the autocatalyic system. As the background reaction dominates the system at early reaction times when **A·B** complexes are the predominant species, it can be assumed that both templates are formed with identical rates via this pathway.

As most signals of different product isomers in the autocatalytic reaction overlapped due to structural similarities, it was impossible to determine the composition of the product mixture from 1D-NMR. Again, we applied ROESY after the reaction was completed (see Fig. 2c). The obtained spectra clearly showed that the amide NH-proton of one isomer does not undergo chemical exchange, while it was observable for the remaining two isomers. Only **XX** is expected to exhibit intramolecular hydrogen bonding and therefore no chemical exchange; its existence in the mixture was further supported by cross peaks indicating an *exo*-Diels-Alder product. The presence of **XN** could be ruled out by the following reasoning: ROESY spectra and HPLC plots of the background reaction using methylester **A'** did not show any *exo*-products. The occurence of an *exo*-product in the presence of recognition sites then means that a recognition-mediated reaction pathway is exploited. Such a pathway is only available for the **XX** isomer via an **A·B**-complex. The remaining two products were identified as the **NN** (main product) and **NX** isomers (side product).

Since the catalytic properties of the indivual products could not be characterized by experiment, we performed AIMD free energy calculations of the **A**+**B** [4+2]-cycloaddition in the presence of different template isomers using the recently proposed *dynamic distance* constraint[27-29], which controls the root mean square distance, D, between the two pairs of carbon atoms involved in the cycloaddition. This 'dynamic' distance D was varied incrementally from 3.6 Å (precursors) to 1.6 Å (product) and free energy profiles were obtained by thermodynamic integration. These are compared to minimum energy paths (MEPs) obtained from constrained geometry optimizations along this reaction coordinate. Fig. 3 clearly shows that there are significant entropic contributions separating the free energy curves from the MEPs. Moreover, the MEPs display a substantial ruggedness due to the fact that at each value of D the highly flexible structures **B** and **NN** or **NX** possess a large number of local minima, further demonstrating the need to carry out molecular dynamics studies. Hence interpreting replicators purely on the basis of 0 K energy barriers obtained from optimized transition state structures bears considerable uncertainties [17,19-21,23]. In the case of the **NN** product, for instance, the reaction barrier of the MEP appears to be lower for the templated than for the non-templated reaction by 13 kJ/mol. Crucially, however, the free energy profiles are nearly identical. Thus the catalytic effect of the template is not the result of a change in electronic structure of the **A**+**B** transition state, but rather of a change in molecularity of the ligation reaction from second-order
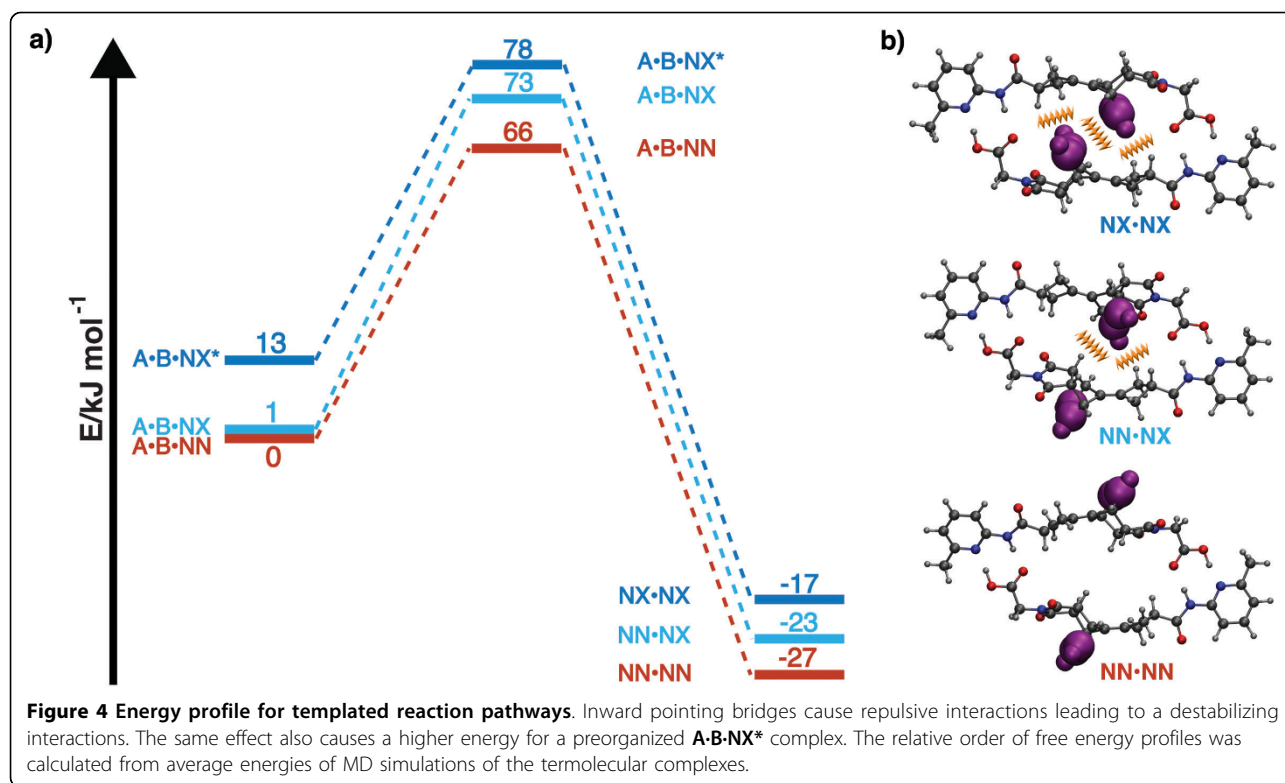
Figure 3 Energy profiles for the non-catalyzed and catalyzed formation of NN. Free energies (■) for the non-templated (a) and templated (b) ligation leading to an NN isomer are almost identical, while the MEPs (□) are different and clearly separated. *D* = 3.6 Å corresponds to precursors, *D* = 1.6 Å to the product. Errors for free energies are calculated from fluctuations of the constraint force.

to pseudo first-order without being energetically penalized.

AIMD simulations for different reaction paths of the system revealed that the **NN** template is only able to catalyze its own formation. On the other hand, the **NX** template cannot only reproduce itself, but also catalyze the formation of an **NN** template. Both autocatalytic reactions feature similar free energy barriers – 65 and 66 kJ/mol, respectively – while the cross-catalytic pathway leading to a new **NN** product templated by **NX** has a slightly higher barrier of 72 kJ/mol. Calculated free energy profiles ordered by averaged absolute energies of template duplexes are displayed in Fig. 4. Obviously, an inward-pointing bridge of an **NX** template destabilizes complexes by repulsive steric interactions, promoting **NX·NX** to an even higher energy than **NN·NX** due to a closer spatial proximity of both bridges. The same destabilizing effect is present in **A·B·NX***, as **B** is very close to the bridge of **NX**. This finding combined with the "selfish" behaviour of the **NN** isomer means that it will outcompete the **NX** isomer, which remains a side product. The reason why the **XX** isomer is observed in the recognition-mediated reaction as a side-product is that its formation is entropically favoured by proceeding in a

pseudo-intramolecular fashion and therefore accelerated (Fig. 1).

Although the composition of the experimental product mixture could be elucidated by 2D-NMR and calculated free energy profiles, an assignment of isomers to 1D-NMR signals was still necessary for a kinetic modelling of the system. The assignment of the **NN** isomer is straightforward, as it is the main product, but it is difficult to distinguish between the **NX** and **XX** isomer on the basis of the available data. For a direct assignment of the experimental NMR spectra we calculated thermally averaged ab initio chemical shifts. A comparison of calculated and experimental shifts for the set of non-overlapping protons used to extract time-dependent concentrations (see Fig. 2a/b) shows a remarkable agreement for both isomers with a deviation of just 0.05 ppm for **XX** and 0.03 ppm for **NX**, respectively. Our final assignment was supported by these shifts and corroborated by the fact that an inverse assignment did not allow for a good fit of the experimental kinetic data to models that were in accordance with results from our calculations. A 16:1 diastereoselectivity for **NN** was determined by integration of the respective NMR peaks, which is a true emergent property, as it results exclusively from the interactions between templates and precursors. It even reverses the slight selectivity for **NX** in the background reaction.

Our kinetic model was constructed based on information about possible reaction channels from AIMD simulations (see Figs. 1 and 4). Complex equilibria of **A·B·NN** and **A·B·NX** complexes were modeled with the same association constant, while **A·B·NX*** was modeled with a separate association constant to account for different relative complex energies. For the same reason all three duplex equilibria were attributed different association constants. Different rate constants were assigned to autocatalytic and crosscatalytic ligations. The rate constant for uncatalyzed reactions to **NN** and **NX** was known from separate measurements of the background reaction. Complex associations were assumed to be limited only by diffusion. In order to quantify the rate constants for these processes separate classical MD simulations of **A**, **B** and **NN** in chloroform were performed and the diffusion constant – which is proportional to the rate constant in this scenario – was determined from the center-of-mass mean square displacement via the Einstein relation. Thus we arrived at rate constants of the order of 10^{10} M^{-1}s^{-1} for all diffusion limited processes. Kinetic data was fitted to the model using Simfit [30] to obtain rate and equilibrium constants. According to the model, the cycloaddition of **A**+**B** is rather efficiently catalyzed in the presence of **NN** or **NX**, the rate constant k_{auto} being about 50 times larger than k_{non} for a non-catalyzed background reaction

Figure 4 Energy profile for templated reaction pathways. Inward pointing bridges cause repulsive interactions leading to a destabilizing interactions. The same effect also causes a higher energy for a preorganized **A·B·NX*** complex. The relative order of free energy profiles was calculated from average energies of MD simulations of the termolecular complexes.

(corresponding to an effective kinetic molarity of 50 M). The crosscatalytic mechanism is less efficient, its rate constant k_{cross} is predicted to be approximately one half of k_{auto}, which is in agreement with our calculated free energy profiles. Furthermore, k_{auto} is four orders of magnitude larger than the rate constant k_{XX} of the **XX** formation via an **AB** channel. This means, although still present, this undesirable pathway is sufficiently suppressed. Template duplexes are predicted to be more stable than termolecular complexes, suggesting that the system suffers from product inhibition. Interestingly, association constants for different termolecular complexes and duplexes reflect the relative order of calculated complex energies. All in all, the model is able to describe the dynamic behaviour of the system very well. Nevertheless, one has to keep in mind the system's complexity and very limited amount of accessible observables. As a consequence, kinetic and thermodynamic parameters obtained by kinetic fitting cannot be expected to be highly accurate. On the other hand, our method allowed us to construct a meaningful model in the first place, which would have been impossible without access to free energy profiles of all major reaction paths.

Complex reaction networks with interesting dynamic signatures in which obstacles like chemical lability or similarity lead to an incomplete base of solid chemical knowledge are expected to challenge chemistry in the future. Our approach of merging experimental NMR kinetics with *ab initio* dynamical chemical shifts and free energy landscapes enabled us to comprehend a dynamic puzzle which otherwise would have had to remain unsolved.

Additional material

Additional file 1: Supporting Information. Description of the synthesis of all compounds, kinetic modelling and setup of calculations.

Author details
[1]Lehrstuhl für Organische Chemie I, Bioorganische Chemie, Ruhr-Universität Bochum, 44780, Bochum, Germany. [2]Department of Physics King's College London, WC2R 2LS, London, UK. [3]Department of Engineering, King's College London, WC2R 2LS, London, UK.

Authors' contributions
SB calculated ab initio NMR shifts. CL performed classical MD simulations to determine diffusion coefficients. AD synthesized all compounds, conducted kinetic measurements as well as data analysis/fitting and performed AIMD simulations. GvK. suggested to use fulvenes for replicator construction. AD, NLD and GvK wrote the paper. NLD and GvK supervised the project.

Competing interests
The authors declare that they have no competing interests.

References

1. von Kiedrowski G: **Minimal replicator theory I: Parabolic versus exponential growth.** *Bioorg Chem Front* 1993, **3**:113-146.
2. von Kiedrowski G: **A Self-Replicating Hexadeoxynucleotide.** *Angew Chem Int Ed* 1986, **25**:932-935.
3. Sievers D, von Kiedrowski G: **Self-replication of complementary nucleotide-based oligomers.** *Nature* 1994, **369**:221-224.
4. Zielinski W, Orgel L: **Autocatalytic synthesis of a tetranucleotide analogue.** *Nature* 1987, **327**:346-347.
5. Tjivikua T, Ballester P, Rebek J: **A self-replicating system.** *J Am Chem Soc* 1990, **112**:1249-1250.
6. Wintner E, Conn M, Rebek J: **Self-replicating Molecules: A second generation.** *J Am Chem Soc* 1994, **116**:8877-8884.
7. Menger F, Eliseev A, Khanjin N: **A Self-Replicating System: New Experimental Data and a New Mechanistic Interpretation.** *J Am Chem Soc* 1994, **116**:3613-3614.
8. Reinhoudt D, Rudkevich D, de Jong F: **Kinetic Analysis of the Rebek Self-Replicating System: Is There a Controversy?** *J Am Chem Soc* 1996, **118**:6880-6889.
9. Lee D, Granja J, Martinez J, Severin K, Ghadiri M: **A self-replicating peptide.** *Nature* 1996, **382**:525-527.
10. Saghatelian A, Yokobayashi Y, Soltani K, Ghadiri M: **A chiroselective peptide replicator.** *Nature* 2001, **409**:797-801.
11. Ashkenasy G, Ghadiri M: **Boolean Logic Functions of a Synthetic Peptide Network.** *J Am Chem Soc* 2004, **126**:11140-11141.
12. Wagner N, Ashkenasy G: **Symmetry and order in systems chemistry.** *J Chem Phys* 2009, **130**:164907.
13. Rubinov B, Wagner N, Rapaport H, Ashkenasy G: **Self-Replicating Amphiphilic b-Sheet Peptides.** *Angew Chem Int Ed* 2009, **48**:6683-6686.
14. Xu S, Giuseppone N: **Self-Duplicating Amplification in a Dynamic Combinatorial Library.** *J Am Chem Soc* 2008, **130**:1826-1827.
15. Nguyen R, Allouche L, Buhler E, Giuseppone N: **Dynamic Combinatorial Evolution within Self-Replicating Supramolecular Assemblies.** *Angew Chem Int Ed* 2009, **48**:1093-1096.
16. Wang B, Sutherland I: **Self-replication in a Diels-Alder reaction.** *Chem Comm* 1997, **16**:1495-1496.
17. Kindermann M, Stahl I, Reimold M, Pankau W, von Kiedrowski G: **Systems Chemistry: Kinetic and Computational Analysis of a Nearly Exponential Organic Replicator.** *Angew Chem Int Ed* 2005, **44**:6750-6755.
18. Stahl I, von Kiedrowski G: **Kinetic NMR Titration: Including Chemical Shift Information in the Kinetic Analysis of Supramolecular Reaction Systems such as Organic Replicators.** *J Am Chem Soc* 2006, **128**:14014-14015.
19. Kassianidis E, Philp D: **Design and Implementation of a Highly Selective Minimal Self-Replicating System.** *Angew Chem Int Ed* 2006, **45**:6344-6348.
20. Kassianidis E, Pearson R, Philp D: **Specific Autocatalysis in Diastereoisomeric Replicators.** *Org Lett* 2005, **7**:3833-3836.
21. Kassianidis E, Pearson R, Philp D: **Probing Structural Effects on Replication Efficiency through Comparative Analyses of Families of Potential Self-Replicators.** *Chem Eur J* 2006, **12**:8798-8812.
22. del Amo V, Philp D: **Making Imines Without Making Water–Exploiting a Recognition-Mediated Aza-Wittig Reaction.** *Org Lett* 2009, **11**:301-304.
23. Pearson R, Kassianidis E, Slawin A, Philp D: **Comparative Analyses of a Family of Potential Self-Replicators: The Subtle Interplay between Molecular Structure and the Efficiency of Self-Replication.** *Chem Eur J* 2006, **12**:6829-6840.
24. Szathmary E, Gladkih I: **Sub-Exponential Growth and Coexistence of Non-Enzymatically Replicating Templates.** *J Theor Biol* 1989, **138**:55-58.
25. Eigen M: **Selforganization of Matter and the Evolution of Biological Macromolecules.** *Naturwis-senschaften* 1971, **58**:465-523.
26. Car R, Parrinello M: **Unified Approach for Molecular Dynamics and Density-Functional Theory.** *Phys Rev Lett* 1985, **55**:2471-2474.
27. Doltsinis N: **Free Energy and Rare Events in Molecular Dynamics.** *NIC Series, ISBN 3-00-017350-1* John von Neumann Institute for Computing, JülichGrotendorst J, Blügel S, Marx D 2006, **31**:375-387.
28. Markwick R, Doltsinis N, Schlitter J: **Probing irradition induced DNA damage mechanisms using excited state Car-Parinello molecular dynamics.** *J Chem Phys* 2007, **126**:045104.
29. Burisch C, Markwick P, Doltsinis N, Schlitter J: **Dynamic Distance Reaction Coordinate for Competing Bonds: Applications in Classical and Ab Initio Simulations.** *J Chem Theory Comp* 2008, **4**:164-172.
30. von Kiedrowski G: *SimFit 32* Bochum, Germany 1990.