Synthetic Molecular Machines for Transportation and Pumping, and Theoretical Analysis for Them

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Science and Engineering

2021

Shuntaro Amano

Department of Chemistry

List of Contents

Total word count (excluding References, Declaration and copyright statement, Acknowledgements and Appendix) = 53,332

List of Figures, Tables and Boxes

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of Abbreviations

Ac	acetate
ADMP	2-azido-1,3-dimethylimidazolinium hexafluorophosphate
ADP	adenosine diphosphate
ami	amide
AmPhos	[4-(*N*,*N*-Dimethylamino)phenyl]di-*tert*-butyl phosphine
APCI	atmospheric-pressure chemical ionization
aq.	aqueous solution
ASAP	atmospheric pressure solids analysis probe
ATP	adenosine triphosphate
ax	axial
Bn	benzyl
Boc	*tert*-butyloxycarbonyl
calcd.	calculated
car	carbamate
Cat	catalyst
CBS	Corey-Bakshi-Shibata
CCW	counterclockwise
chem	chemical
COSY	correlation spectroscopy
CRN	chemical reaction network
CuAAC	copper(I)-catalysed alkyne-azide cycloaddition
CW	clockwise
Cy	cyclohexyl
d	days
δ	chemical shift
dba	dibenzylideneacetone
DBA	dibenzyl ammonium
DBF	dibenzofulvene
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DIBAL-D	diisobutylaluminum deuteride
DIPEA	diisopropylethylamine
DMAP	4-(dimethylamino)pyridine
DMEAD	di-2-methoxyethyl azodicarboxylate
DMEDA	dimethylethylenediamine
DMF	*N*,*N*-dimethylformamide
DMP	Dess-Martin periodinane
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dpp	diphenylphenanthroline
dppf	1,1'-bis(diphenylphosphino)ferrocene
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
eq	equatorial
eq., equiv.	equivalent
ESI	electrospray ionization
Et	ethyl
ET	energy transfer
Fmoc	9-fluorenylmethyloxycarbonyl

fum, Fum	fumaramide
GTP	guanosine triphosphate
h	hours
HESI	heated electrospray ionization
Hex	hexane
HMBC	heteronuclear multiple-bond correlation spectroscopy
HRMS	high-resolution mass spectrometry
HSQC	heteronuclear single-quantum correlation spectroscopy
i-Pr	*iso*-propyl
mal	maleamide
MBA	monobenzyl ammonium
Me	methyl
mech	mechanical
min	minutes
Ms	mesyl
MS	mass spectrometry
NBS	*N*-bromosuccinimide
NMR	nuclear magnetic resonance
PG	protecting group
Ph	phenyl
Pi	inorganic phosphate
pin	pinacolato
PMB	*p*-methoxybenzyl
PMP	*p*-methoxyphenyl
ppm	parts per million
PTLC	preparative thin layer chromatography
quant.	quantitative
ROESY	rotating frame nuclear overhauser effect spectroscopy
r.t., RT	room temperature
s, sec	seconds
SEC	size exclusion chromatography
SS	stationary state
succ	succinamide
TBAF	tetrabutylammonium fluoride
TBDMS	*tertiary*-butyldimethylsilyl
TBTA	tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine
t-Bu	*tertiary*-butyl
TCEP	tris(2-carboxyethyl)phosphine
terpy	terpyridine
Tf	triflyl
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TIPS	triisopropylsilyl
TLC	thin layer chromatography
TMS	trimethylsilyl
Tr	trityl
tri	triazole
Ts	tosyl
UV	ultraviolet
Vis	visible

v/*v*	volume per volume
w/*w*	weight per weight

Conventional abbreviations for units, constants, physical quantities and chemical functional groups are used.

Abstract and Structure of the Thesis

Biology utilises molecular machines for various purposes, such as transportation, synthesis, signal transduction and information processing. Synthetic chemists have strived to design and synthesise molecular machines that are reminiscent of their biological counterparts. Here we present two molecular machines for transportation with different mechanisms. One is a directional molecular transporter utilising multiple hydrazone-type switches, which coordinates the operation of multiple movable parts. The other is a catalysis-driven small-molecule pump, which autonomously and progressively pumps multiple macrocycles as long as the energy source is present. We also analysed the mechanism of a chemically driven autonomous synthetic molecular motor with information thermodynamics, a framework that treats information as a physical variable and quantitatively relates it with other thermodynamic parameters such as entropy and free energy. These attempts to create synthetic molecular machines that are capable of sophisticated tasks and to theoretically analyse their mechanisms will pave the way to future nanotechnology with life-like properties, and advance the understanding of biological molecular machines.

Chapter 1 contains an introduction to molecular machines: definition and general principles of molecular machines, overview of synthetic molecular machines for motility and transportation, and theoretical analyses of their mechanisms.

Chapter 2 presents an ongoing work of a unidirectional molecular transporter utilising multiple hydrazone-type switches. This molecular machine is aimed at transporting a molecular cargo from one end of the machine to the other by the coordinated operation of two hydrazone-type switches.

Chapter 3 describes an autonomous chemically fueled small-molecule pump, which pumps up to three macrocycles from bulk solution to the thread via an information ratchet mechanism.

Chapter 4 details the analysis of a chemically driven synthetic molecular motor operating via an information ratchet mechanism with the framework of information thermodynamics.

This thesis is presented in a journal format. Chapter 3 has been published as a peer-reviewed article and Chapter 4 has been submitted to a peer-reviewed journal for publication. Both articles have been edited only to ensure consistency within this thesis. A copy of published/submitted manuscripts is attached in the Appendix of the thesis. The Appendix also includes another manuscript that has been accepted by a journal for publication but is not included in the thesis. The author's contribution to each project is detailed at the beginning of each chapter.

Declaration and Copyright Statement

Unless otherwise stated at the beginning of the relevant chapter, no portion of the work referred to in the dissertation has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in The University's policy on Presentation of Theses.

Acknowledgements

First, I would like to thank Prof. David A. Leigh for letting me work in his group and supervising my thesis. He gave me a large degree of freedom and resources that I needed to pursue my goals. Through discussing him, I learned a lot about how to brush up ideas, how to manage projects, how to write a good manuscript and how to deal with editors and referees, so basically everything I needed to know as a researcher. He showed me how to balance between big dreams and practical considerations. People tend to focus on one of them and forget about the other, but I believe considering both of these aspects are really important. Also, the atmosphere in the Leigh group has been exceptionally good. Every time I heard about horrible situations in other groups, I thanked Dave for maintaining this friendly and collaborative environment. If I were to become a PI, I will surely try to create such a working environment.

I would like to thank Sau Yin Sek and Valérie Bruyr for their efforts for maintaining a great working environment. Because of their tireless efforts, we can focus on research without worrying about other stuffs. I believe this is the best possible situation we can ask for as a researcher.

I also thank Dr. Vanesa Marcos Algaba, Dr. Salma Kassem, Dr. Marcel Dommaschk, Dr. Shaofei Zhang, Dr. Javier Echavarren and Dr. Albano Galan for helping me work on the project of a molecular transporter utilising multiple hydrazone-type switches, so-called “Bucket Brigade project”. Due to its complicated structure, it has been a tough project and synthesis has been difficult. Without you, the project would not have reached to the current stage. Through working with you, I learned a lot about organic synthesis, especially how to move forward when there seems to be no way forward. In particular, I thank Salma, who taught me basic skills of organic synthesis and mentored me when I joined the group. It has

been a basic foundation as a researcher for me, and everything I have done and will do will be built on this solid foundation.

I thank Dr. Stephen D. P. Fielden for tackling together the project of a chemically driven autonomous pump. This project was exciting and led to an excellent outcome. It was a highlight in my PhD and working with Steve was a great pleasure. His broad knowledge and excellent skills show his passion for research, and I am glad that I could share it with you.

I thank Prof. Massimiliano Esposito, Dr. Elisabeth Kreidt, Emanuele Penocchio and Benjamin M. W. Roberts for collaborating in the information thermodynamic analysis of a molecular motor project. It was an 'abnormal' project in that we initially had very little common language and we needed to develop it in the course of collaboration. This required a lot of patience for both sides. I am glad that I could share this invaluable experience with you, and I am looking forward to the future opportunities of working with you again.

I thank Dr. Stefan Borsley and Dr. Zhanhu Sun for writing a Chemical Engine Perspective together. Discussing with you guys have always been fun. I had an experience of merging different opinions and perspectives to produce something much better than my personal thought. Previously I found it difficult, but now I can enjoy it because of the experience I had with you.

I also thank everyone who worked together in other projects, namely Dean Thomas, Dr. Daniel J. Tetlow, Martin Power, Dr. Chong Tian and Hayley Frost. It has been inspiring to work with you and I learned a lot from each of you.

In general, big thanks to all the lab members. It has been a great joy to work with you. I have not had opportunity to collaborate with everyone in the research project, but still talking to you during the lunch time or in the pub has been a great pleasure and comfort. I will miss it when I leave the group. Special thanks go to my office mates and bay mates. I do not list everyone's name here because I am so forgetful that I might miss someone. But I

truly enjoyed sharing the same space with you and talking about various topics ranging from weekend activities to Japanese culture.

The President's Doctoral Scholar Award from the University of Manchester has been indispensable for my pursuit of PhD and I thank the university for offering this scholarship.

Moving away from the lab, I also express my gratitude to my father, Shiro, my mother, Junko, my sister, Mai and our dog, Basil, for supporting me from Japan through Skype, email and sending Japanese foods and items. Thanks to your support, I did not become homesick. I hope I can travel to Japan soon and meet you again.

Finally, I would like to thank my girlfriend, Huifeng Li, for always being beside me. Although we have not seen each other for almost two years due to the pandemic, I did not feel lonely because I could always talk to you. Research is sometimes stressful, but communicating with you has made it bearable. Hopefully we can meet soon, when the situation gets better.

To E., who first supported my pursuit of chemistry.

Chapter 1.

Introduction: Molecular Machines

Blank Page

Declaration

An earlier version of this chapter has been submitted to the University of Manchester to support the author's application for the degree of Master of Science by Research in 2018.

Acknowledgements

This chapter was prepared solely by the author. Dr. Daniel J. Tetlow is gratefully thanked for proofreading this chapter.

Synopsis

This chapter presents the state of the art in the field of synthetic molecular machines. First, definition and general principles of molecular machines common to both biological and synthetic molecular machines are introduced. Then, different mechanisms of molecular machines for motility and transportation, and their synthetic examples are reviewed. These mechanisms are a switch, an energy ratchet and an information ratchet. Finally, theoretical analyses of molecular machines, especially chemically driven information ratchets, are reviewed. The chapter concludes with describing challenges in the field and how the projects in the following chapters addressed them.

1.1. Aim and structure of the literature review

In a recent review[1], molecular machines are defined as a "subset of "molecular devices" (functional molecular systems) in which some stimulus triggers the controlled, large amplitude or directional mechanical motion of one component relative to another (or of a substrate relative to the machine) which results in a net task being performed." Such molecular machines are widely utilised in biological systems, carrying out crucial tasks. These biological molecular machines include dynein[2], kinesin[3,4] (for the transportation along microtubules), myosin[4] (for the contraction of muscles), ion pumps[5,6], ribosomes[7] (for the synthesis of proteins) and ATP synthase[8] (for the synthesis of ATP and generating the proton concentration gradient across the membrane). Inspired by these sophisticated biological molecular machines, chemists set out to develop artificial molecular machines. Such synthetic molecular machines reminiscent of biology will be crucial for future nanotechnology that can perform various tasks such as controlled transportation, motility, synthesis, and adaptation to the environment.

The aim of this introduction part is to review the development of the field of synthetic molecular machines and to clarify the context of the author's PhD research. The review mainly focuses on the molecular machines operated in solution. There are molecular machines that operate in other conditions, such as on solid surface[9,10], however these machines are omitted in this review. This chapter will first discuss general principles of molecular machines (Section 1.2.). Then, representative examples of synthetic molecular machines that operate with different mechanisms will be reviewed. These mechanisms are molecular switches (Section 1.3.), energy ratchets (Section 1.4.1.) and information ratchets (Section 1.4.2.). Next, theoretical analysis of chemically driven molecular motors will be

discussed (Section 1.5.). Finally, challenges in the field of synthetic molecular machines and aims of the projects described in this thesis will be explained (Section 1.6.).

1.2. General principles of molecular machines

Studies in physics and biology have revealed some of the general principles of molecular machines that apply both to biological and synthetic systems[1,11,12]. They are significantly different from the principles of macroscopic machines, and understanding these principles is important for designing new synthetic molecular machines. These principles are as follows:

1. The effect of scale: On a molecular scale, the significance of the physical phenomena that affect the operation of machines, such as Brownian motion (random motion of an object caused by the collision of solvent molecules to it), heat dissipation, solvation, momentum, inertia, gravity, differ fundamentally from a macroscopic scale. On a large scale, inertial terms and gravity play an important role in determining the motion of the objects. However, in a scale smaller than micrometre, momentum and gravity lose their relevance while viscous forces and Brownian motion become dominant. Therefore, in the design of molecular machines, these forces should be taken into consideration. The consequence of the relevance of viscous forces is that, unlike in macroscopic machines, inertia cannot be effectively utilised for causing directional movement: even when some directional mechanical force is applied to a microscopic object in solution, its effect will be immediately dampened by strong viscous forces and does not result in significant directional movement. Instead, the movement of such a microscopic object is governed entirely by the forces acting on it at every instant.

2. Utilisation of Brownian motion: One of the dominant forces that act to a microscopic object in solution is Brownian motion. The studies of biological molecular machines suggest that it is more effective to utilise Brownian motion, rather than to counteract it. This is because the energy scale of Brownian motion at room temperature is comparable to or larger than the scale of energy that molecular machines can utilise from an energy source. For example, a typical motor protein consumes 100 to 1,000 ATP fuel molecules per second, which enables maximum possible power output in the region of 10^{-16} to 10^{-17} W per molecule[13]. This magnitude of energy is much smaller than that of Brownian motion that molecules experience in solution at room temperature, which is approximately 10^{-8} W.

3. Breaking detailed balance: The principle of detailed balance states that in an equilibrium state, the rates of the transitions between any two states in opposite directions are the same. Therefore, no net flux can be generated. This principle implies that in order for molecular machines to perform net tasks, they have to be driven out of equilibrium and break the detailed balance.

In addition to the physical principles discussed above, biological molecular machines also provide practical guidelines for the design of synthetic molecular machines[1]. Biological molecular machines are made of intrinsically flexible molecules and utilise chemical energy, either released from the conversion of chemical fuel (e.g. ATP) to waste (e.g. ADP and inorganic phosphate) or in the form of concentration gradient across a membrane. Additionally, they utilise architectures that restrict most of the degrees of freedom of the machine components and/or the substrates they interact with. Such architectures include tracks for the walker molecular machines and membranes where ion pumps are embedded.

These architectures enable molecular machines to perform tasks progressively (e.g. walking along a track step by step, pumping ions to generate concentration gradient across a membrane). Understanding design principles of biological molecular machines will facilitate design of their synthetic counterparts with similar properties and, ideally, comparable sophistication.

Another important concept to be aware of in the field of molecular machines is the fundamental distinction between "switch" and "motor", which exhibit different behaviors when the system is returned to an original state[1]. If a molecular switch operates with some mechanical effect on an external system, and then returned to its original state, the mechanical task it performed is undone. This is because the process of returning is exactly the reverse of the process of performing a task. Conversely, if a molecular motor operates with some mechanical effect on an external system, and then returned to its original state, the mechanical task it performed is not necessarily undone. This is because the molecular motor can return to its original state *via* a different process to the one it underwent to perform a task. For example, when a molecular rotary motor undergoes 360º directional rotation, it returns to the original state but the mechanical task it performed is not reset. The key reason why this distinction between the "switch" and the "motor" is important is that the former cannot perform tasks progressively so that they accumulate, while the latter can. In the following sections, general mechanisms and representative examples of molecular switches and motors will be reviewed.

1.3. Molecular switches

Molecular switches generally switch the thermodynamically most stable state by the change of the operating environment. Immediately after this change, the system is in an out-

of-equilibrium state. Then, Brownian motion drives majority of the molecular machines into the new thermodynamic sink, relaxing the system to a new equilibrium. This process results in a net directional motion. However, when the operating environment is reset and concomitantly the system returns to its original equilibrium, the machine undergoes the reverse process of the initial switching, undoing the mechanical task it performed. This is because there is no mechanism to prevent the reverse backward motion. This property exhibits a stark contrast to the ratchet mechanisms required for molecular motors (Section 1.4.), which have a mechanism to prevent the backward motion.

The 'switch' type of molecular machines include azobenzene tweezers[14], hydrazone switches[15,16] and a photo-responsive molecule that twists the guest[17]. These machines utilise the change of configuration induced by external stimuli to perform mechanical tasks. For example, the photo-responsive switch by Aida[17] (**1** in **Figure 1**) changes the configuration upon the irradiation of UV light and visible light. This change is caused by *cis-trans* isomerisation of the azobenzene moiety in the molecule and results in the twisting of the guest molecule (**2** in **Figure 1**). Another example of a molecular switch utilising configurational change is Aprahamian's hydrazone switch[15,16] (**Figure 2**). This switch changes its configuration and conformation depending on the acidity of the operating environment. Initially, the switch has *E*-configuration (*E*-**3**). The protonation of the pyridine ring isomerises the switch into *Z*-configuration (*Z*-**3**·H^+). Further protonation flips the quinoline part, but the configuration of the C-N double bond remains the same (*Z*-**3**·$2H^+$). Deprotonation of the switch flips back the quinoline part (*Z*-**3**), which isomerises into more stable *E*-configuration (*E*-**3**).

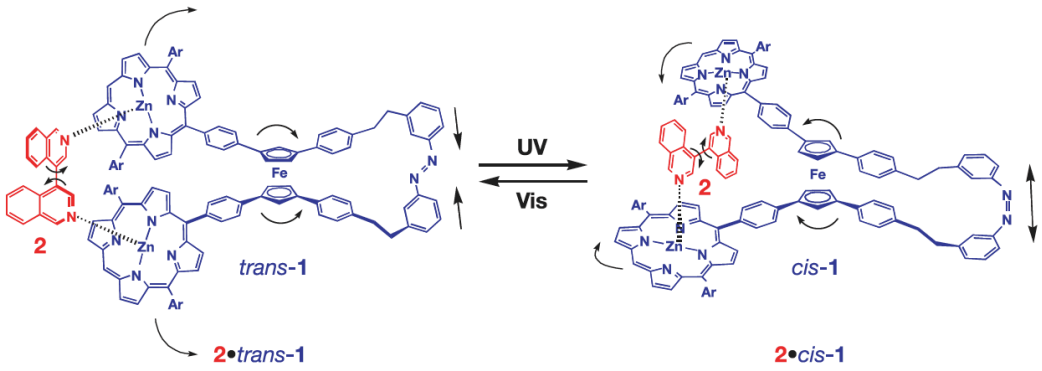


Figure 1. Photo-responsive switch that twists a guest molecule by Aida[17]. Reprinted by permission from Muraoka, T., Kinbara, K. & Aida, T. Mechanical twisting of a guest by a photoresponsive host. *Nature* **440**, 512–515 (2006), Springer Nature.

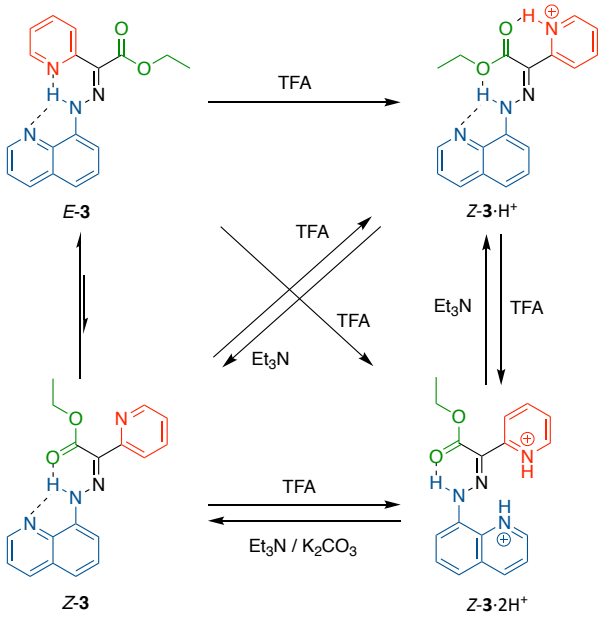


Figure 2. Aprahamian's hydrazone switch[15,16].

Other examples of molecular switches have interlocked structures such as catenanes[18] and rotaxanes[19]. A catenane consists of multiple interlocked macrocycles, whereas a rotaxane is composed of a macrocycle part and an axle part that threads through the macrocycle. In molecular switches with interlocked structures, multiple binding sites for a macrocycle are introduced in the other macrocycle (in the case of a [2]catenane) or in the axle part (in the case of a rotaxane). The binding affinity of these sites can be modulated by external stimuli, causing large amplitude motion of a macrocycle against the other component (a macrocycle or an axle). For example, Sauvage's [2]catenate[18] consists of one

macrocycle that only has a diphenylphenanthroline (dpp) unit, the other macrocycle that has a dpp unit and a terpyridine (terpy) unit, and a copper ion that is chelated by these units (**Figure 3**). When the copper ion takes the oxidation state of +1 (Cu(I)), it favors four-coordinate tetrahedral complexes, and two dpp units of each macrocycle coordinate to the Cu(I) ion (dpp-[Cu^{I}(**4**)]). If the copper ion is oxidized to Cu(II), it favors five- or six-coordinate complexes, and the macrocycle with two binding sites rotates against the other macrocycle so that one dpp unit and one terpy unit coordinate to the Cu(II) ion (terpy-[Cu^{II}(**4**)]). When the copper ion is reduced back to Cu(I), the catenate returns to the original state of dpp-[Cu^{I}(**4**)].

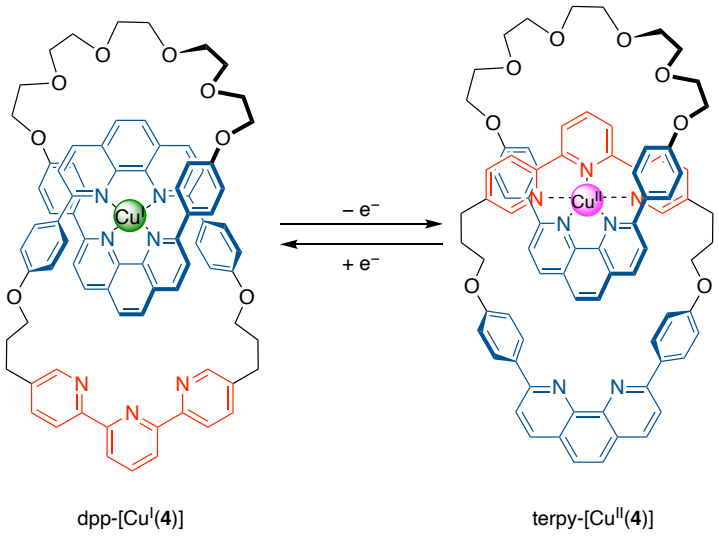


Figure 3. Sauvage's switchable catenane[18].

An example of a rotaxane-based switch is Stoddart's rotaxane shuttle[19] (**Figure 4**), which contains two binding sites for a positively-charged cyclophane macrocycle, a benzidine station and a biphenol station. Under neutral conditions, the macrocycle binds predominantly at the benzidine station due to the stronger stabilization by donor-acceptor interactions (**5**). When the operating environment becomes acidic, the benzidine station protonates and causes electronic repulsion against the macrocycle. This repulsion causes the

macrocycle to move to the biphenol station that now binds the macrocycle more strongly (**5**·2H$^+$).

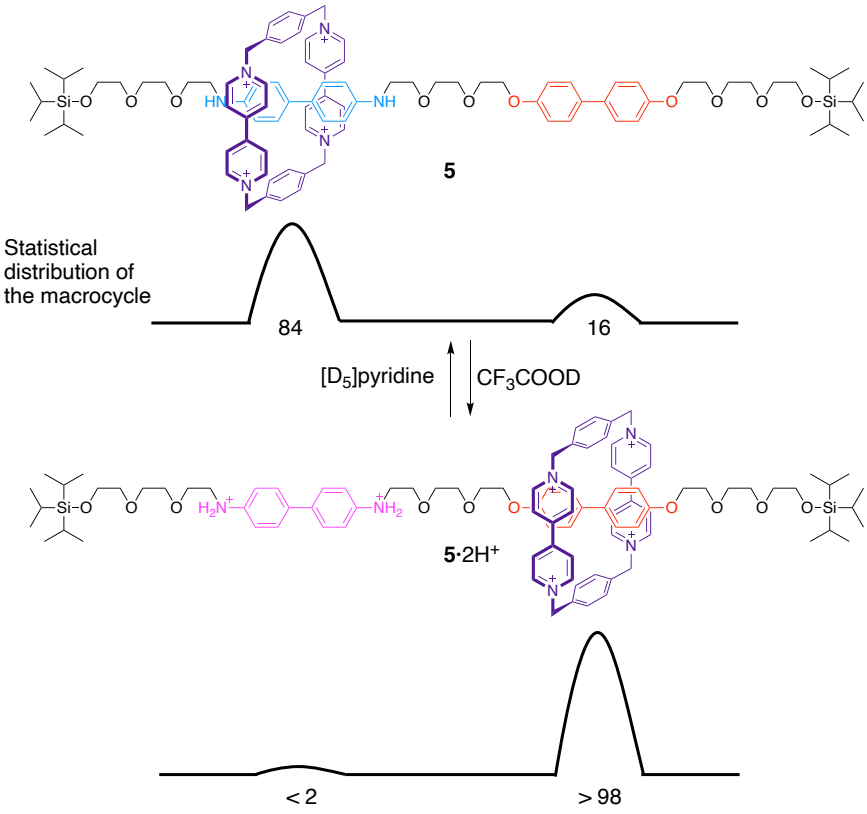


Figure 4. Stoddart's rotaxane shuttle with two binding sites[19].

Directional motion caused by molecular switches can be used for performing tasks. For example, the molecular transporter reported by Feringa[20] transports a cargo molecule from one position in the machine to another (**Figure 5**). The design of this machine is based on the light-driven molecular motor developed by the Feringa group[21]. This machine achieves the transportation of an acetyl group from one end of the machine to the other by photoisomerisation of the central, crowded alkene. First, the thiol group on the arm attacks the active ester group under basic conditions and picks up the acetyl group (*Z*-**6**-OAc→ *Z*-**6**-SAc). Next, *Z*→*E* photoisomerisation of the alkene transports the cargo to the opposite side of the molecule (*Z*-**6**-SAc→ *E*-**6**-SAc). Finally, the amine group present attacks the thioester group and completes the transportation process (*E*-**6**-SAc→ *E*-**6**-NHAc).

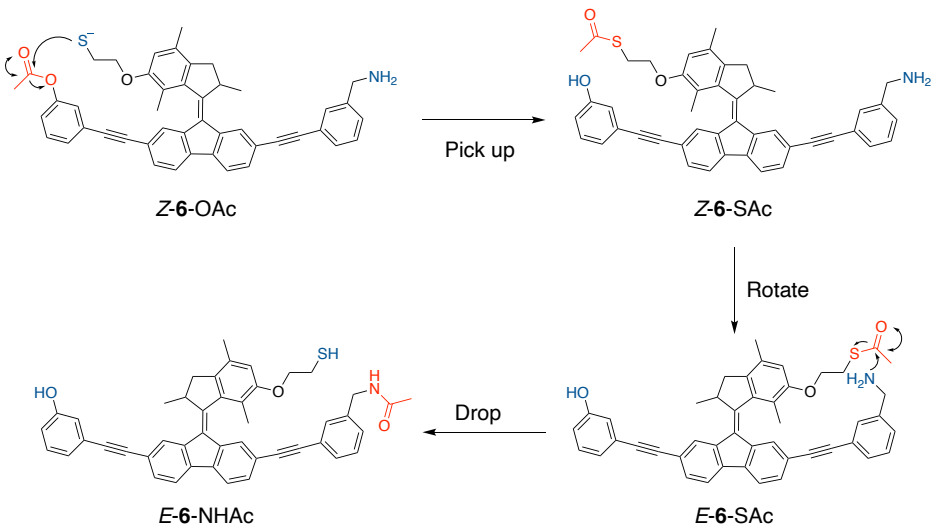


Figure 5. Light-driven molecular transporter by Feringa[20].

Another example of a molecular transporter was realised by the Leigh group[22] (**Figure 6**) which utilises Aprahamian's hydrazone switch[15,16] (**Figure 2**). This molecular transporter can transport a cargo molecule between two aldehyde stations. Initially, the cargo molecule (red part of *EZ*-**7**-left) is bound to the deuterated aldehyde station in the left half of the molecule (blue part in *EZ*-**7**-left). The central switch is then isomerised to its *Z*-configuration by acid (*Z*-**7**-left). Oxidation of the thiols connects the cargo molecule to the arm part by a disulfide bond (*Z*-**8**-left). Subsequent addition of acid triggers the isomerisation of the switch to its *E*-configuration, while at the same time promoting hydrazone exchange. As a result the cargo molecule is transported to the non-deuterated aldehyde station (*E*-**8**-right). Finally, reductive cleavage of the disulfide bond completes the process (*EZ*-**7**-right). This process is reversible: the cargo molecule can also be transported from the non-deuterated aldehyde station to the deuterated aldehyde station.

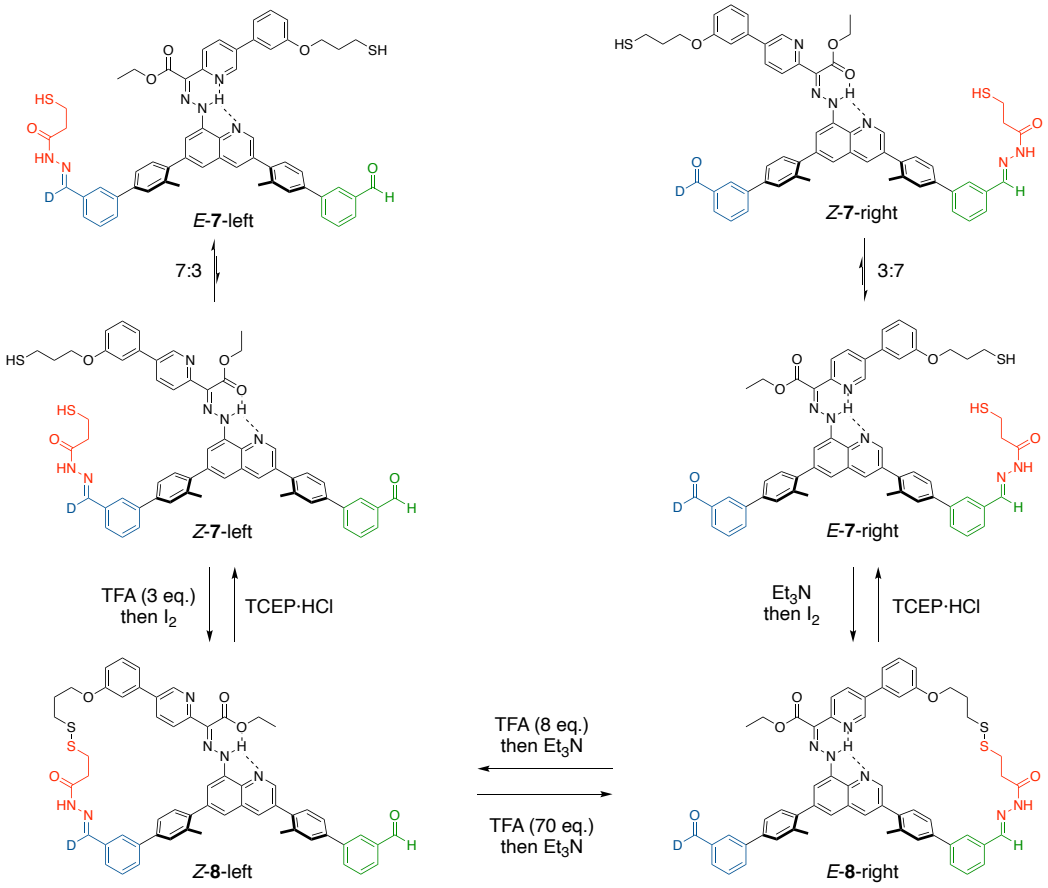


Figure 6. Molecular transporter utilising a hydrazone switch[22].

1.4. Molecular motors

In the molecular switches reviewed in the previous section, the task they performed is undone once the machine returns to its original state. However, the tasks performed by biological molecular machines are not undone even when the machine returns to its original state. This means that biological molecular machines are "motors", which require Brownian ratchet mechanisms[12,23]. Brownian ratchets rectify random Brownian motion into net directional motion, and consist of three components[24]: (1) a randomising element (Brownian motion of the machine components or substrates); (2) an energy input to break the detailed balance under the restriction of the second law of thermodynamics; (3) asymmetry in the energy potential to bias Brownian motion and generate net directional motion. The last two

requirements (energy input and asymmetry) can be provided in different ways depending on the type of the Brownian ratchet: energy ratchets and information ratchets.

1.4.1. Energy ratchets

The energy ratchet is one type of Brownian ratchet mechanism where potential-energy surface or operating conditions (e.g. temperature) are varied in a periodic or stochastic fashion, independent of the position of the particle (moving component of the molecular machine) on the potential-energy surface[1,25]. Among various subcategories of energy ratchet mechanisms, the most commonly observed mechanism in molecular machines is the flashing ratchet mechanism, which involves switching between two different potential-energy surfaces (**Figure 7**). Typically, it consists of a set of asymmetric potential-energy surfaces, each of them having a periodic series of two different energy minima and two different energy maxima. When these energy surfaces are alternately applied, a Brownian particle is directionally transported along these surfaces: consider that the Brownian particle starts in a green well in **Figure 7a**. When the energy surface is switched (**Figure 7b**), the yellow well becomes thermodynamically more stable than the blue well, due to the switching of the

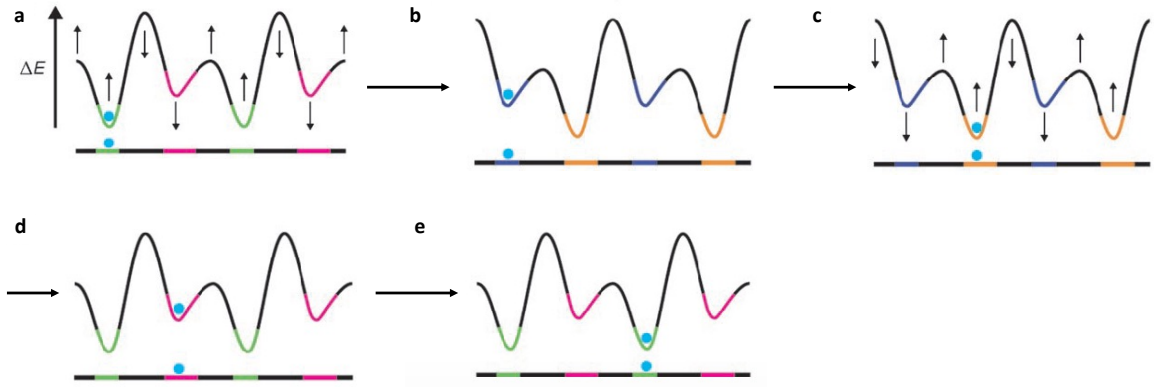


Figure 7. Typical mechanism of a flashing ratchet[1]. The horizontal axis, vertical axis, curve line and a blue circle represent reaction coordinate, free energy level, potential-energy surface and a Brownian particle, respectively. Adapted by permission from Kay, E. R., Leigh, D. A. & Zerbetto, F. Synthetic molecular motors and mechanical machines. *Angew. Chem. Int. Ed.* **46**, 72–191 (2007), John Wiley and Sons.

energies of the wells. Also, the kinetic barrier between these wells is lowered. Therefore, the Brownian particle will be transported from the blue well to the yellow well (**Figure 7c**). Next, the potential-energy surface is switched back to the initial one (**Figure 7d**), which makes the green well more stable than the red well. Importantly, the kinetic barriers are also switched, which makes the transportation of the Brownian particle from the red well to the righthand green well more kinetically favorable than to the lefthand green well (**Figure 7e**). Such concomitant switching of the energy minima (energies of the wells) and energy maxima (kinetic barriers) prevents backward motion and enables net directional motion. Note that the switching between two energy surfaces occurs independent of the position of the Brownian particle, which exhibits stark contrast to the information ratchet mechanism (Section 1.4.2.).

The energy ratchet mechanism has been the basis of various molecular machines, such as a molecular walker driven by photoisomerisation and dynamic covalent chemistry[26], molecular motors that rotate around a C−C single bond[27-29], catenane-type molecular motors[30-32] and a rotaxane-type molecular pump[33]. One of the earliest examples is a catenane-type molecular rotary motor reported by Leigh, which rotates one macrocycle around the other by the sequential addition of stimuli[31] (**Figure 8**). The initial state of the motor, *fum*-*E*-**9**, has a macrocycle on the fumaramide binding site. There is another binding site on the track (succinamide group), but the shuttling between two binding sites is sterically prevented by a TBDMS group and a trityl group. On irradiation of UV light, the fumaramide undergoes photoisomerisation to be converted to a maleamide (*mal*-*Z*-**9**), which has lower binding affinity to the macrocycle than the fumaramide station. Therefore, after the de-silylation and re-silylation, the macrocycle shuttles to the succinamide station because it has higher binding affinity than the maleamide (*succ*-*Z*-**9**). Then, addition of piperidine

isomerises the maleamide back to a fumaramide (*succ-E*-**9**). Following de-tritylation and re-tritylation facilitates the shuttling of the macrocycle back to the fumaramide station (*fum-E*-**9**), because the fumaramide has higher binding affinity to the macrocycle than the succinamide. This sequential control of the binding site affinities and kinetic barriers results in a 360° rotation of the macrocycle along the track. The sense of rotation can be inverted by changing the order of adding reagents (photoisomerisation of the fumaramide, de-tritylation/re-tritylation, reisomerisation of the maleamide, de-silylation/re-silylation).

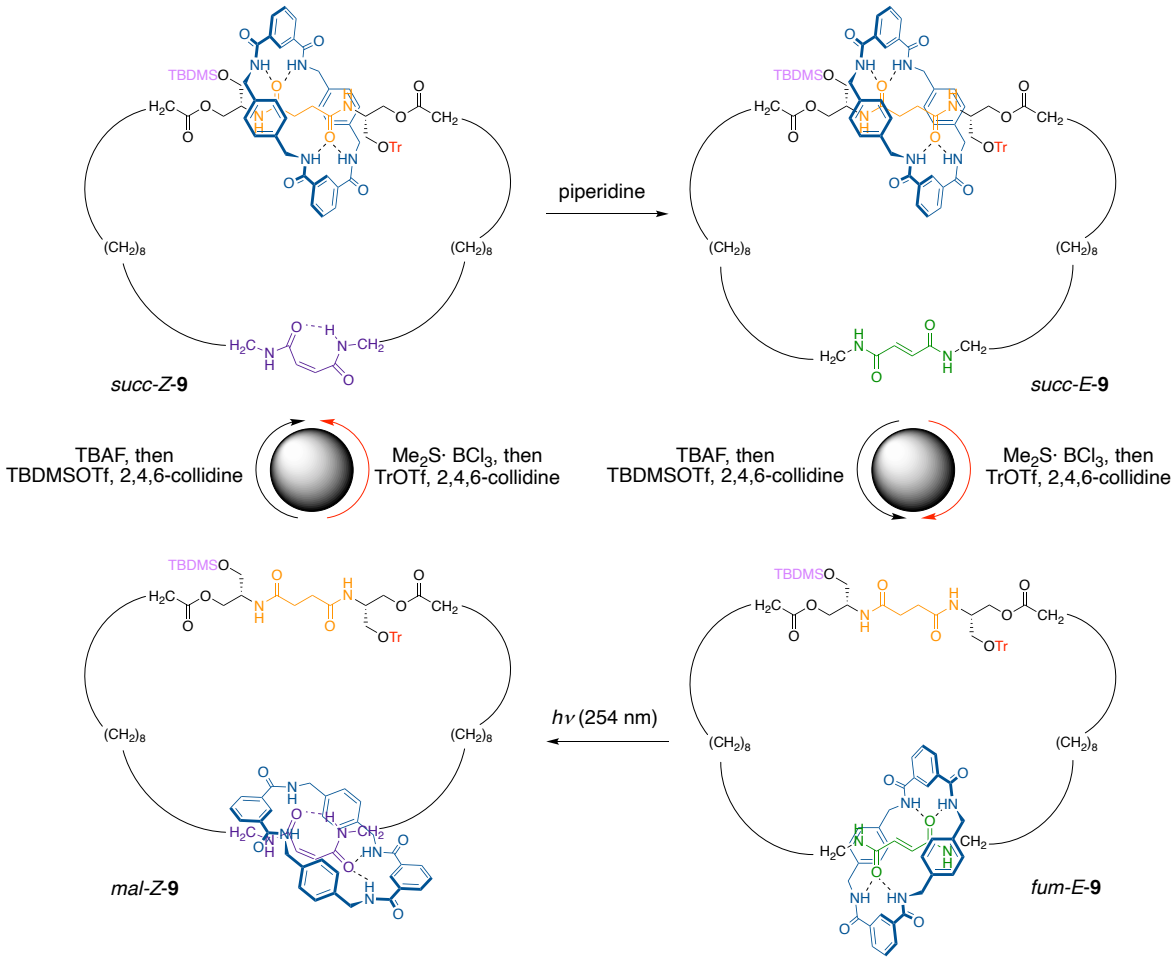


Figure 8. Leigh's catenane-type molecular rotary motor[31].

An energy ratchet mechanism has also been utilised for realising directional linear motion. For example, a synthetic walker reported by Leigh[26] achieved a directional linear transport of a walker molecule along the track (**Figure 9**). First, *E*→*Z* isomerisation of the

stilbene linkage by UV light and disulfide exchange under basic conditions transport the walker (red part in **Figure 9**) from the (1, 2) position to the (2, 3) position (*Z*-**10**-(2,3)). Next, *Z*→*E* isomerisation of the stilbene by visible light with iodine causes strain in the macrocycle in the (2, 3) isomer. Finally, hydrazone exchange under acidic conditions releases this strain, leading to biased formation of the (3, 4) isomer (*E*-**10**-(3,4)). The direction of transportation can be reversed by changing the order of operation: the walker can be transported from the (3, 4) position to the (1, 2) position by *E*→*Z* photoisomerisation, hydrazone exchange, *Z*→*E* photoisomerisation and disulfide exchange.

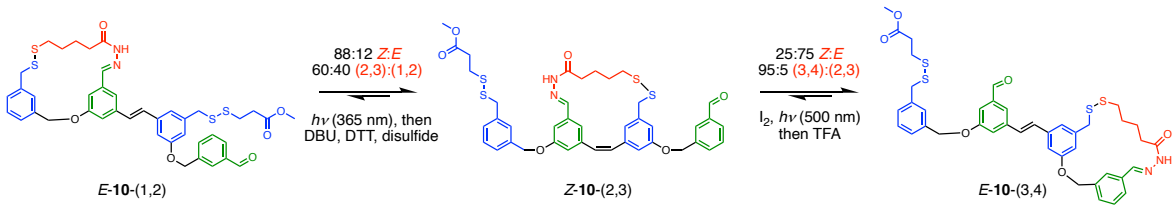


Figure 9. Unidirectional and reversible molecular walker by Leigh[26].

The two systems described above control the thermodynamic stability of states and kinetic barriers between the states by different stimuli. However, this approach can make the operation of such molecules complicated and therefore is not practical for further application. To address this problem, energy ratchets that control the thermodynamic stability of states and kinetic barriers with the same stimuli have been developed. One prominent example is a molecular pump reported by Stoddart[33], which pumps multiple macrocycles from bulk solution onto a catchment thread (**Figure 10**). Initially, the macrocycle (**11**) and the pump molecule (**12**) repel one another as they are both positively charged (**Figure 10a**). However, once the macrocycle and the viologen unit are reduced to radicals, an attractive interaction between them emerges due to stabilising radical–radical pimerisation. Also, the reduction decreases the repulsion between the macrocycle and the pyridinium unit, which lowers the

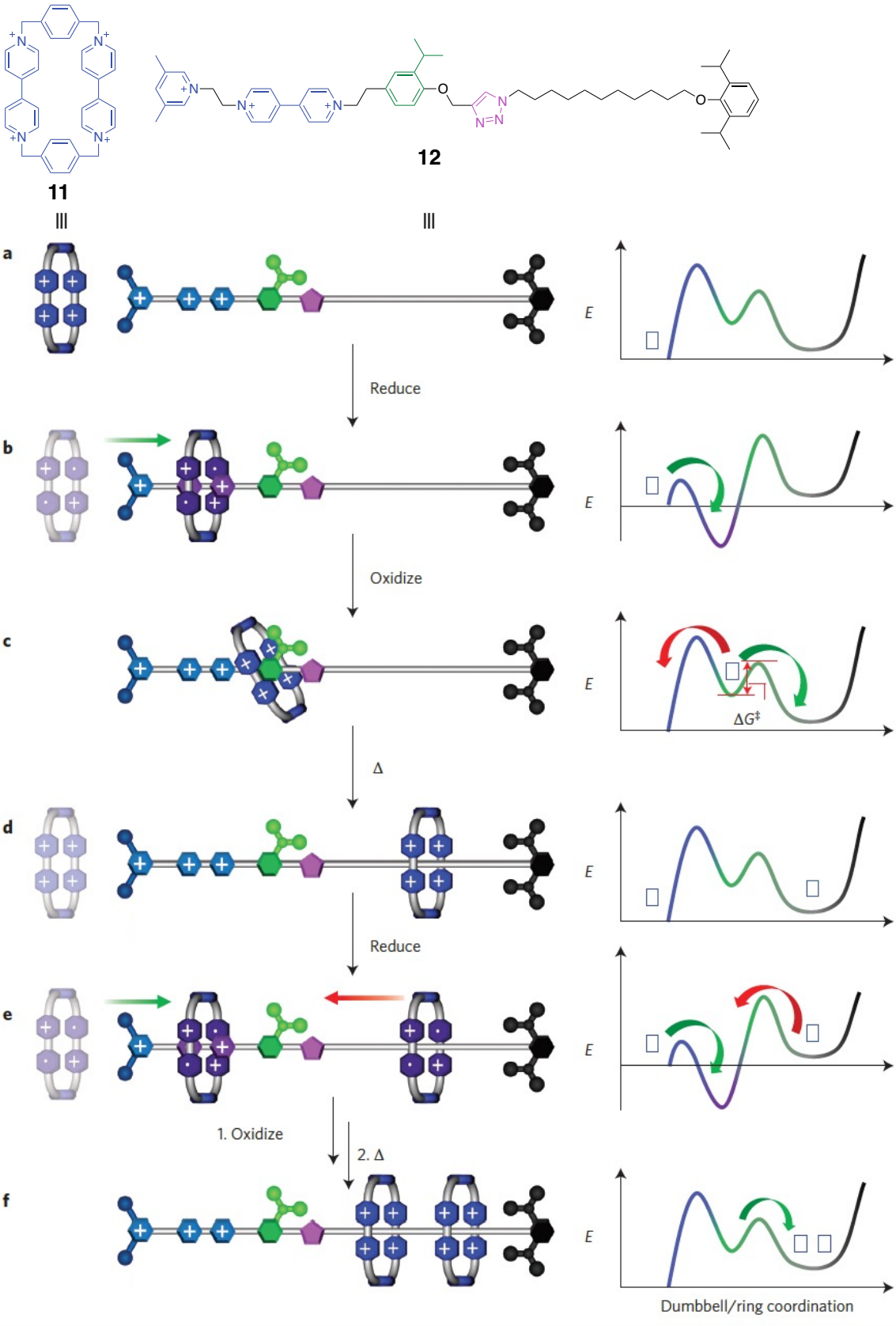


Figure 10. Molecular pump operated with redox reactions by Stoddart[33]. The operation cycle is shown by cartoon illustrations of the macrocycle (**11**) and the pump (**12**) with a simplified potential-energy surface at each step (horizontal axis: reaction coordinate along the pump molecule, vertical

axis: free energy level). Adapted by permission from Cheng, C. *et al.* An artificial molecular pump. *Nat. Nanotechnol.* **10**, 547–553 (2015), Springer Nature.

kinetic barrier for the threading of the macrocycle onto the pump. Consequently, the macrocycle gets threaded on the viologen unit to form a thermodynamically stable complex (**Figure 10b**). Subsequent oxidation of the macrocycle and the viologen unit restores the repulsion between them (**Figure 10c**). The macrocycle would become thermodynamically most stable if it were to dethread back into the bulk solution, but this process is kinetically prevented due to insurmountable electrostatic barrier (red arrow). Alternatively, the macrocycle undergoes a more kinetically favoured process, which is the translocation over the isopropylphenyl unit to the alkyl catchment thread (green arrow, **Figure 10d**). When the macrocycle and the viologen unit are reduced again, the threading of the second macrocycle (green arrow in **Figure 10e**) occurs faster than the relocation of the first macrocycle to the viologen unit (red arrow), due to the steric hinderance caused by the isopropylphenyl unit. Further oxidation and translocation of the second macrocycle results in the pumping of the second macrocycle onto the catchment thread (**Figure 10f**).

An example of utilising a simpler energy ratchet mechanism for directional rotation is Feringa's molecular motor based on a palladium redox cycle that rotates around a C−C single bond[28] (**Figure 11**). The motor consists of a biaryl unit connected to a sulfoxide group ((*S*,*M*)-**13**) and has two sources of chirality: one from the chirality over the biaryl axis and the other from the central chirality at sulfur. Under ambient conditions, both of the atropisomers and the chiral sulfoxide are stable and do not undergo epimerisation. Treatment of (*S*,*M*)-**13** with palladium(II) complex leads to formation of a palladacycle via C−H activation, where the sulfoxide acts as a directing group ({Pd[(*R*,*P*)-**14**]Cl}$_2$). This step rotates the upper aryl ring (blue part) in a clockwise sense with respect to the lower aryl ring (red part) and prevents backward motion. The formation of a six-membered palladacycle

considerably lowers the energy barrier between two atropisomers, allowing the conversion of {Pd[(*R*,*P*)-**14**]Cl}$_2$ to a more stable {Pd[(*R*,*M*)-**14**]Cl}$_2$. Subsequent reintroduction of the C−H bond yields (*S*,*P*)-**13**, completing a directional clockwise 180° rotation. This process also reduces palladium(II) to palladium(0), which can react with the C−Br bond *via* oxidative addition on addition of a phosphine ligand. This step leads to the formation of a new palladacycle due to the coordination of the sulfoxide to the palladium(II) centre (Pd[(*R,P*)-**15**]BrPCy$_3$). As in the first half of the rotation, this six-membered palladacycle significantly lowers the energy barrier between two atropisomers, allowing the conversion of Pd[(*R,P*)-**15**]BrPCy$_3$ to a more stable Pd[(*R,M*)-**15**]BrPCy$_3$. Subsequent reintroduction of the C−Br bond yields (*S,M*)-**13**, completing the second clockwise 180° rotation. The key element in this rotation mechanism is the different selectivity of palladium(II) and palladium(0) toward the reaction with the C−H bond and the C−Br bond: palladium(II) selectively reacts with the C−H bond, while palladium(0) with the C−Br bond, allowing formation of desired palladacycles at each step. In 2020, the Feringa group reported a molecular rotary motor that operates with a similar energy ratchet mechanism but utilises different chemistry of reversible lactone formation[29].

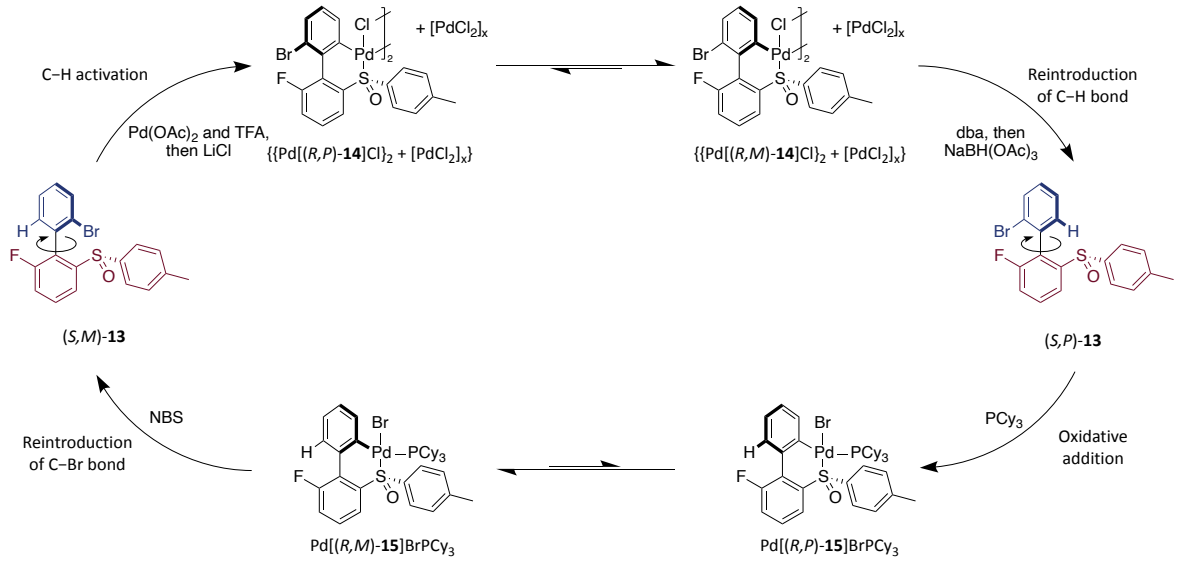


Figure 11. Molecular rotary motor based on a palladium redox cycle by Feringa[28]. *S* and *R* show the fixed central chirality at sulfur, while *M* and *P* show the chirality over the biaryl axis.

Recently, a system that utilises the same chemical processes for realising different types of motion has been realised, which is a rotary motor and a molecular pump by Leigh[32] (**Figure 12**). This system employs trichloroacetic acid as a chemical fuel to change the operating conditions[34,35]. In the case of a molecular motor (**Figure 12a**, compound **16**), the addition of trichloroacetic acid facilitates hydrazone exchange and protonates the dibenzylamine station. This simultaneously lowers the kinetic barrier for the shuttling of the macrocycle between the triazolium and dibenzylamine station, and increases the binding affinity of the macrocycle to the dibenzylamine station (now dibenzylammonium station). Consequently, the macrocycle is translocated from the triazolium to the dibenzylammonium station through the left-hand pathway (**16**·H^+). The trichloroacetic acid fuel gets decarboxylated and the media becomes basic, locking the hydrazone exchange while facilitating disulfide exchange. Also, the dibenzylammonium station gets deprotonated, providing thermodynamic driving force for the macrocycle to shuttle back to the triazolium station *via* the right-hand pathway. One pulse of the trichloroacetic acid fuel therefore results in a 360° clockwise rotation of the macrocycle along the track. The same operating mechanism was utilised for a molecular pump (**Figure 12b**, compound **20**). The first pulse of the trichloroacetic acid fuel leads to the pumping of two macrocycles onto the triazolium stations (**21**), and further pulses of the fuel results in pumping of two more macrocycles on the dibenzylamine stations (**22**).

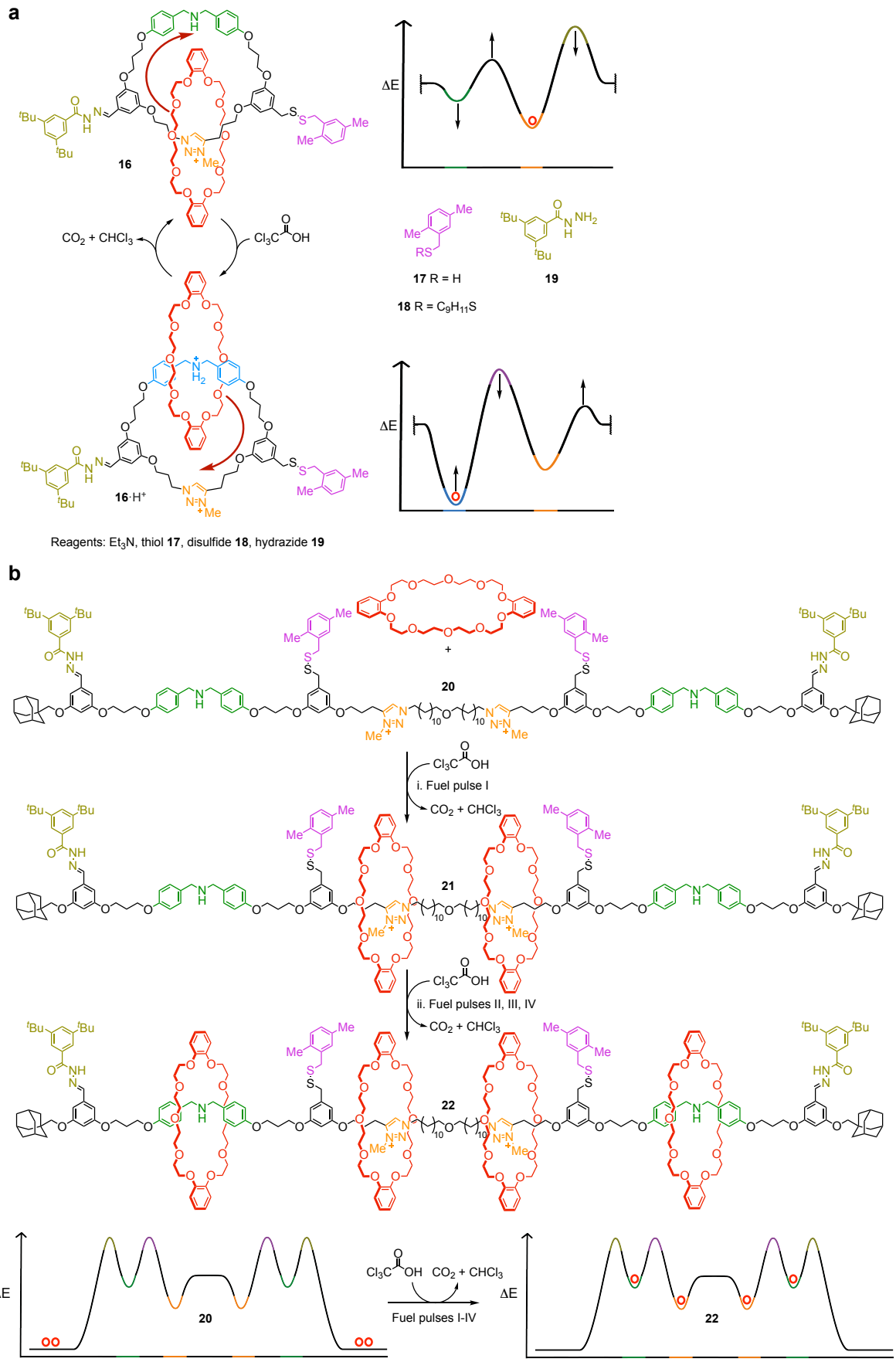

a
16
CO2 + CHCl3
Cl3C OH
ΔE
17 R = H
18 R = C9H11S
19
16·H+
Reagents: Et3N, thiol 17, disulfide 18, hydrazide 19
b
20
i. Fuel pulse I
21
ii. Fuel pulses II, III, IV
22
Fuel pulses I-IV

Figure 12. Leigh's molecular rotary motor and linear motor driven by pulses of trichloroacetic acid fuel[32]. **a**, molecular rotary motor. **b**, molecular pump. The two-dimensional diagrams represent potential-energy surface of each state (horizontal axis: reaction coordinate along the track, vertical axis: free energy level).

1.4.2. Information ratchets

An information ratchet is another type of Brownian ratchet mechanism where the potential-energy surface is switched depending on the position of the particle (moving component of the molecular machine) on the energy surface[1,36,37]. This mechanism involves information transfer from the particle to the switching process, and resembles the mechanism of a well-known thought-machine, Maxwell's demon[38]. This entity was devised by Scottish physicist James Clerk Maxwell in an attempt to illustrate the statistical feature of the Second Law of thermodynamics. Maxwell imagined a system consisting of a gas in a container, which does not allow energy or matter exchange with the external environment (**Figure 13**). No temperature or pressure gradient can spontaneously emerge because of the Second Law of thermodynamics. The container is separated into two parts by a partition with a hole. The hole can be opened or closed by a small "demon", depending on the velocity of the molecule approaching the hole. The "demon" opens the hole only when molecules moving slower than the average (blue dots) approach the hole from left, or when molecules moving faster than the average (red dots) approach the hole from right. By repeating this process, "hot" and "cold" molecules will be confined in the left and right compartment, respectively. This "demon" seemingly contradicts the Second Law, because it seems to spontaneously generate a temperature gradient in an isolated system. The apparent paradox of Maxwell's demon stimulated the investigation of the relationship between thermodynamics and information processing[38], because the demon's ability to generate temperature gradient originates from its intelligence; its capability of measuring the molecule's velocity and open the hole accordingly. Recently, the paradox has finally been resolved by explicitly considering

mutual information during the demon's operation, founding the new field of information thermodynamics[39] (see Chapter 4, Section 4.1.).

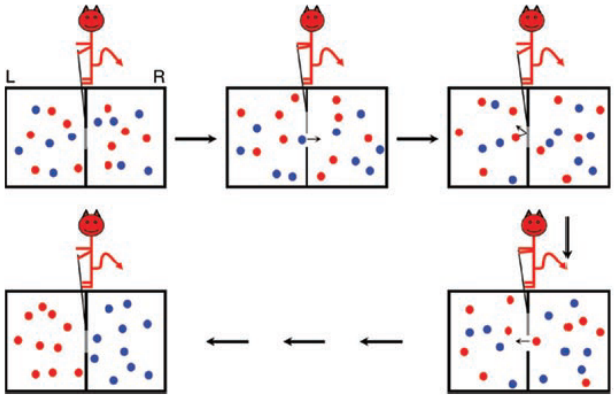


Figure 13. Maxwell's "temperature demon", in which a gas at a certain uniform temperature is sorted into "hot" and "cold" molecules, depending on their velocities[1]. Particles with velocity higher than the average are illustrated by red dots, while particles with velocity lower than the average by blue dots. Adapted by permission from Kay, E. R., Leigh, D. A. & Zerbetto, F. Synthetic molecular motors and mechanical machines. *Angew. Chem. Int. Ed.* **46**, 72–191 (2007), John Wiley and Sons.

The mechanism of information ratchets is similar to Maxwell's demon, although it requires the input of free-energy and therefore evidently consistent with the Second Law. A typical information ratchet consists of a coupling between energy input processes (chemical or light-driven processes) and mechanical processes (movement of one part of the molecular machine relative to another; **Figure 14a**)[40,41]. This figure specifically illustrates a chemically driven information ratchet, but similar illustration will be possible for light-driven information ratchets. Consider that the machine is in the state $\mathbf{M_b}$. It can undergo a mechanical process to be converted to the state $\mathbf{M_a}$. Both of these two states can react with a fuel molecule (**F**) which supplies free energy to the system, but with different rates: $\mathbf{M_a}$ reacts with **F** faster than $\mathbf{M_b}$ does (**Figure 14b**, left). Therefore, if $\mathbf{M_a}$ and $\mathbf{M_b}$ have the same standard chemical potential, formation of $\mathbf{M_a'}$ is kinetically preferred to $\mathbf{M_b'}$. After the formation of $\mathbf{M_a'}$, it can undergo another mechanical process to be converted to the state $\mathbf{M_b'}$. Both of these two states react to emit waste species (**W**) but with different rates: $\mathbf{M_b'}$ reacts faster than $\mathbf{M_a'}$ does (**Figure 14b**, right). Therefore, if $\mathbf{M_a'}$ and $\mathbf{M_b'}$ have the same

standard chemical potential, formation of $\mathbf{M_b}$ is kinetically preferred to $\mathbf{M_a}$. Under the conditions where all chemical and mechanical processes can occur, these kinetic selections in fuel-consuming and waste-forming reactions will result in net clockwise directional motion of the machine ($\mathbf{M_b \rightarrow M_a \rightarrow M_a' \rightarrow M_b' \rightarrow M_b}$). In this mechanism, the switching between two potential-energy surfaces (**Figure 14b**, left and right) takes place depending on the mechanical state of the machine: the switching from the left to right surface occurs faster when the machine is in the mechanical state **a**, whereas the opposite switching occurs faster when the machine is in the mechanical state **b**. Note that not all information ratchets have every element illustrated in **Figure 14**, but still they can be regarded as information ratchets as long as they incorporate kinetic selection depending on the mechanical state as a key mechanism for generating directional motion. In the rest of this section, representative chemically driven synthetic information ratchets, and then light-driven information ratchets, will be reviewed.

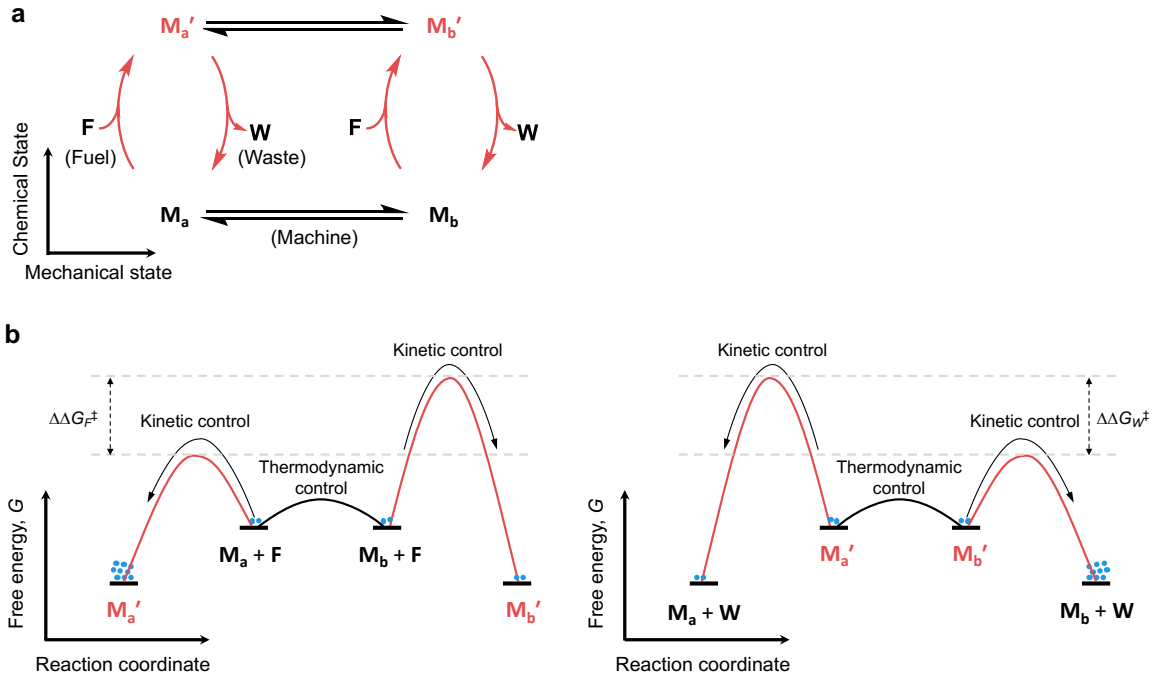


Figure 14. Typical information ratchet mechanism[41]. This figure specifically illustrates a chemically driven information ratchet, but similar illustration will be possible for light-driven information ratchets. **a**, an information ratchet mechanism represented as a coupling between chemical processes (fueled reaction cycles, **M**→[+**F**]→**M′**→[-**W**]→**M**) and mechanical processes ($\mathbf{M_a \rightleftharpoons M_b}$). **M** stands for a molecular machine, and the prime denotes a state of the molecular machine bound to or reacted

with the fuel (chemical states), whereas subscript letters indicate the dynamic mechanical state of the machine. **b**, potential-energy surfaces relevant to each chemical state of the molecular machine. $\Delta\Delta G_F^{\ddagger}$ and $\Delta\Delta G_W^{\ddagger}$ are difference between the standard chemical potential of the transition states of two different paths for the fuel-addition reaction and waste-forming reaction, respectively.

One of the earliest examples of chemically driven information ratchets is a molecular rotary motor based on a lactone formation by Feringa[42] (**Figure 15**). The motor has a biaryl structure, where the phenyl ring has two hydroxy groups protected by orthogonal protecting groups and the naphthyl moiety has a carboxy group. When the hydroxy groups and the carboxy group are disconnected, the atropisomerisation around the biaryl axis is sterically prevented ((*M*)-**23** and (*P*)-**23**). However, when one of the hydroxy groups and the carboxy group form a lactone, the energy barrier for the atropisomerisation is significantly lowered, allowing interconversion between two enantiomers ((*P*)-**24** and (*M*)-**24**). Asymmetric ring opening using a chiral catalyst ((*S*)-CBS catalyst) and following allyl protection and oxidation lead to selective formation of (*P*)-**23** over (*M*)-**23**. Subsequent selective deprotection of the PMB group spontaneously results in a cyclisation, rotating the rotor part (phenyl ring) with respect to the stator part (naphthyl moiety) by 180° ((*P*)-**25** and (*M*)-**25**). The second asymmetric ring opening and following PMB protection and oxidation lead to selective formation of (*M*)-**23** over (*P*)-**23**. Selective deprotection of the allyl group and addition of a carbodiimide reagent enable regeneration of (*P*)-**24**, which completes the 360° rotation.

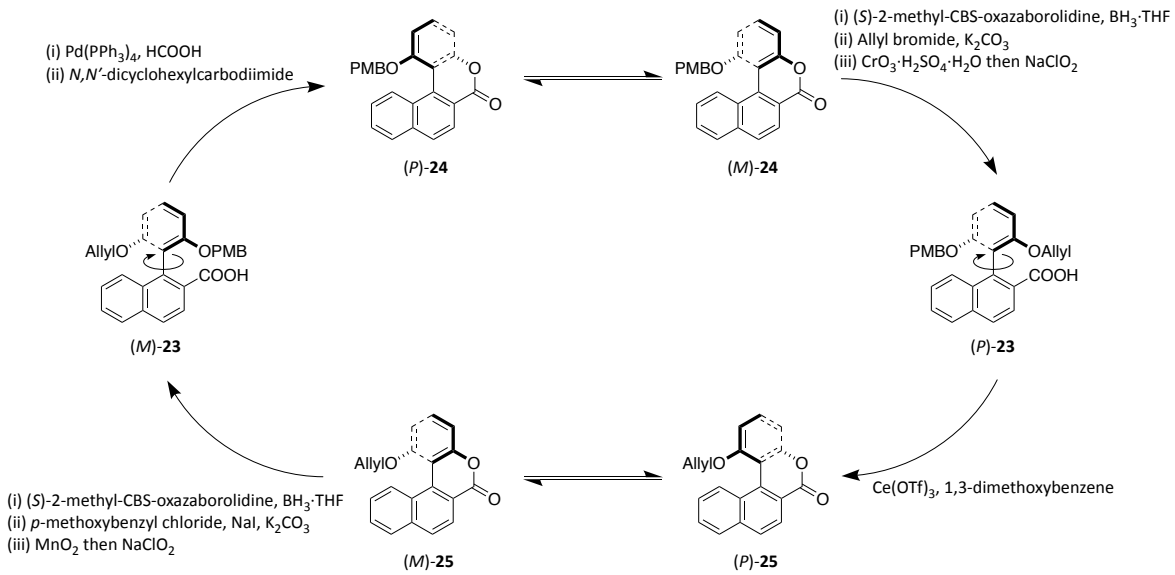


Figure 15. Molecular rotary motor based on a lactone formation by Feringa[42]. *P* and *M* show the chirality over the biaryl axis.

The mechanism of this rotary motor apparently looks similar to the biaryl rotary motor based on a palladium redox cycle by Feringa[28] (**Figure 11**), but the key mechanism for generating directional motion is fundamentally different. In the motor based on a palladium redox cycle (**Figure 11**), the energy of the two interconverting atropisomers (e.g., {Pd[(*R,P*)-**14**]Cl}$_2$ and {Pd[(*R,M*)-**14**]Cl}$_2$) are different, because they are diastereomers. This is why (*S,P*)-**13** preferentially forms over (*S,M*)-**13** on the reintroduction of the C−H bond. In other words, the selectivity is controlled thermodynamically. In contrast, the energy of the two interconverting atropisomers in the case of lactone-based motor (e.g., (*P*)-**24** and (*M*)-**24** in **Figure 15**) are identical, because they are enantiomers. Still, selective formation of (*P*)-**23** over (*M*)-**23** can be achieved due to asymmetric ring opening with a chiral catalyst. In other words, the selectivity is controlled kinetically, and such kinetic selection is a distinct characteristic of information ratchets. In 2005, Branchaud reported another lactone-based biaryl information ratchet with asymmetric ring opening process[43].

A similar kinetic selection mechanism was utilised in Leigh's rotaxane-type information ratchet[44] (**Figure 16**). The two co-conformers (*FumH*$_2$-**26** and *FumD*$_2$-**26**) are interconvertible in the initial state, and the short distance between the central hydroxy group and the fumaramide stations ensures that the presence of the macrocycle in either station produces a chiral environment at the hydroxy group. Introduction of the barrier to the alcohol traps the macrocycle on either side of the thread (*FumH*$_2$-**27** or *FumD*$_2$-**27**). If a chiral electrophile is used for this barrier formation, one of these co-conformers preferentially forms over the other, generating a biased, non-equilibrium distribution. This is because the conversion from *FumH*$_2$-**26** to *FumH*$_2$-**27** and from *FumD*$_2$-**26** to *FumD*$_2$-**27** occur *via* diastereomeric transition states with different energies, which results in different rates in these two reactions. This system was further extended to a rotaxane information ratchet with three fumaramide stations and two hydroxy groups[45].

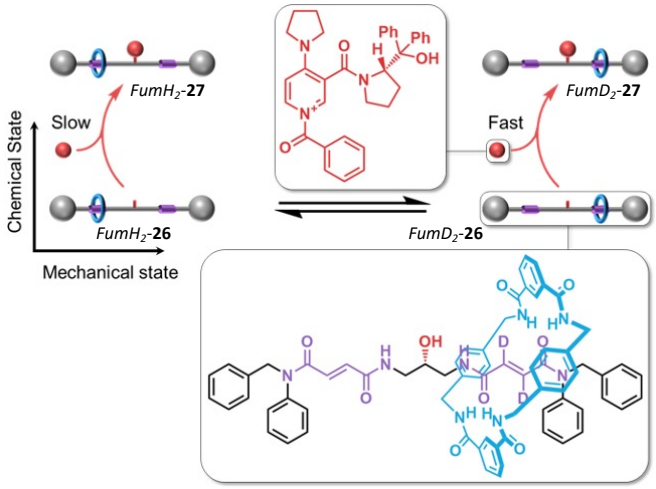


Figure 16. Rotaxane-type information ratchet by Leigh[44]. *FumH*$_2$ and *FumD*$_2$ denote that the macrocycle is on the non-deuterated fumaramide group or on the deuterated fumaramide group, respectively.

Kinetic selection for the information ratchet mechanism can be induced by means other than an asymmetric reaction to enantiomeric reactants as well. A chemically driven autonomous molecular rotary motor by Leigh utilises steric hinderance caused by the rotating component for inducing kinetic selection[46] (**Figure 17**). The system consists of a

[2]catenane structure with two hydroxy groups, where an Fmoc group can attach and detach. The macrocycle (purple) can shuttle between the two fumaramide sites (yellow and green) along the unblocked pathway between *FumH*$_2$-**28** and *FumD*$_2$-**28** or between *FumH*$_2$-**30** and *FumD*$_2$-**30**. When an Fmoc-Cl chemical fuel is present, Fmoc attachment occurs faster with *FumD*$_2$-**28** and *FumH*$_2$-**30** than *FumH*$_2$-**28** and *FumD*$_2$-**30**, whereas any Fmoc removal from *FumH*$_2$-**29** and *FumD*$_2$-**29** occurs at the same rate. The kinetic selection in the Fmoc attachment emerges due to the steric hinderance caused by the macrocycle in *FumH*$_2$-**28** and *FumD*$_2$-**30**, and leads to net clockwise rotation of the macrocycle. The unique characteristic of this system is its autonomy: unlike other information ratchets discussed previously, it continues to rotate progressively as long as the energy source, Fmoc-Cl, is present.

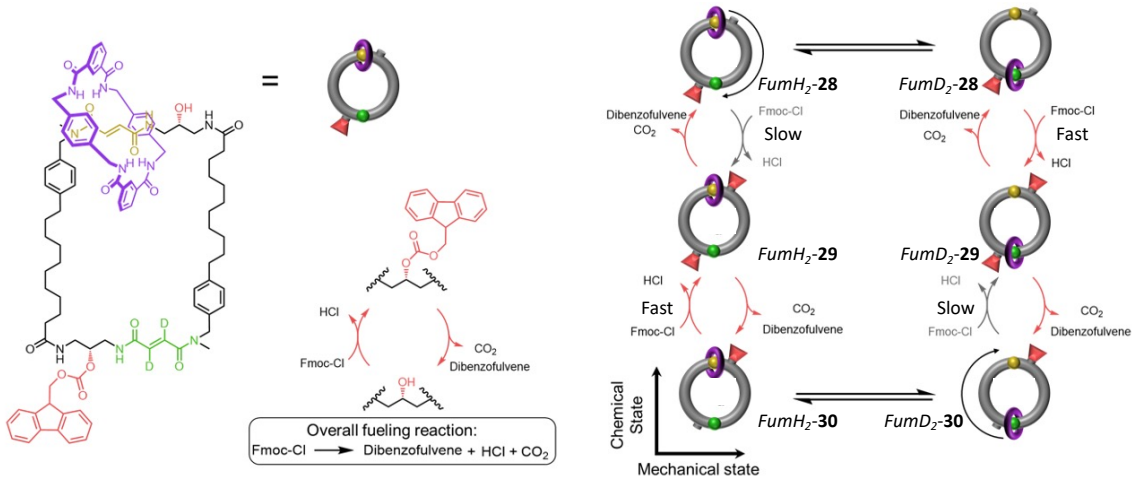


Figure 17. Chemically driven autonomous molecular rotary motor by Leigh[46]. *FumH*$_2$ and *FumD*$_2$ denote that the macrocycle is on the non-deuterated fumaramide group or on the deuterated fumaramide group, respectively.

One significant drawback of the information ratchet fueled by the decomposition of Fmoc-Cl is that the efficiency in terms of the fuel usage is low. Substantial amount of the Fmoc-Cl fuel is decomposed in a background reaction without reacting with the rotary motor to power directional motion. Recently, the Leigh group addressed this problem by utilising a hydrolysis of carbodiimide to drive directional motion[47]. In this system, more than 80% of the carbodiimide fuel was consumed by the machine, thereby significantly improving the

efficiency. The examples of information ratchets described so far all employ chemical reactions involving small molecules only, but an information ratchet that utilises an enzymatic reaction to achieve kinetic selection has also been reported[48].

There are also various examples of synthetic light-driven information ratchets, one of which is Feringa's molecular motor that utilises photoisomerisation of alkene for causing rotatory motion[21] (**Figure 18**). In this system, irradiation with ≥280 nm UV light isomerises (*P*,*P*)-*trans*-**31** into (*M*,*M*)-*cis*-**31**. Then, thermal helix inversion gives rise to the more stable (*P*,*P*)-*cis*-**31** isomer. After another sequence of photoisomerisation and helix inversion, 360° directional rotation can be achieved and the molecule returns to the initial state ((*P*,*P*)-*trans*-**31**). This system can be regarded as an information ratchet because there is an element of kinetic selection in the photoisomerisation step: when we consider the two mechanically interconvertible species, (*M*,*M*)-*cis*-**31** and (*P*,*P*)-*cis*-**31**, under the irradiation with ≥280 nm UV light, photoisomerisation occurs preferentially in the latter than in the former. Similarly, comparing (*M*,*M*)-*trans*-**31** and (*P*,*P*)-*trans*-**31**, the photoisomerisation occurs preferentially in the latter than in the former. Note that, however, chemically driven molecular machines and light-driven molecular machines obey fundamentally different constraints, microscopic reversibility for the former and the Einstein relations between absorption and stimulated and spontaneous emission for the latter[49]. Therefore, in-depth physical analysis of their operating mechanisms must take into account this difference.

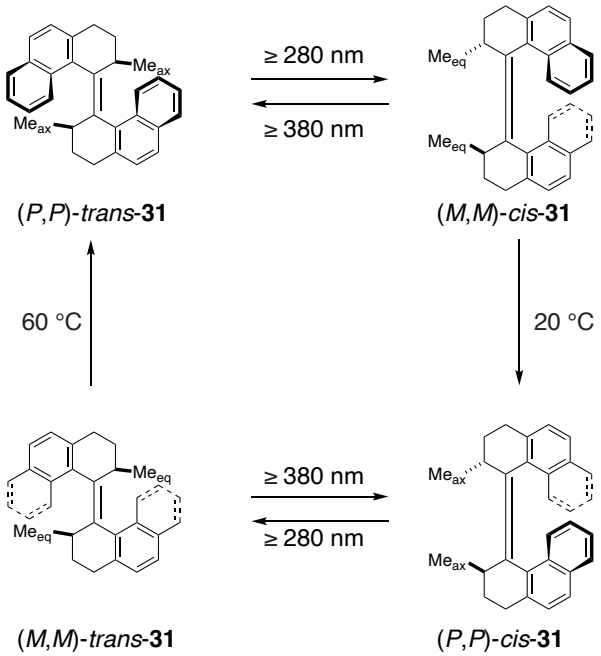


Figure 18. Light-driven molecular rotary motor by Feringa[21]. *P* and *M* show the helicity in each half of the structure.

Another prominent example of a synthetic light-driven information ratchet is a rotaxane-based system by Leigh[50] (**Figure 19**). This system consists of a dibenzo-24-crown-8-based macrocycle and a linear track that has two binding sites for the macrocycle (dibenzyl ammonium (DBA) and monobenzyl ammonium (MBA) stations) and an α-methyl stilbene unit. Under the irradiation at 350 nm wavelength and in the presence of benzil (PhCOCOPh), the macrocycle distribution between the two stations is driven out of equilibrium, thus 'ratcheted' onto one of the stations. When the stilbene is in the *E*-configuration, the macrocycle can shuttle between two binding sites, but when the stilbene is in the *Z*-configuration, the macrocycle shuttling is blocked. There are two major processes that mediate photoisomerisation of the stilbene under the operating conditions: intermolecular energy transfer from benzil and intramolecular energy transfer from benzophenone (PhCOPh) which is incorporated in the macrocycle. The benzil-sensitised photostationary state of α-methyl stilbene is typically *Z*:*E* = 82:18, whereas the benzophenone-sensitised photostationary state is *Z*:*E* = 55:45. This difference in the ratio is key for driving the

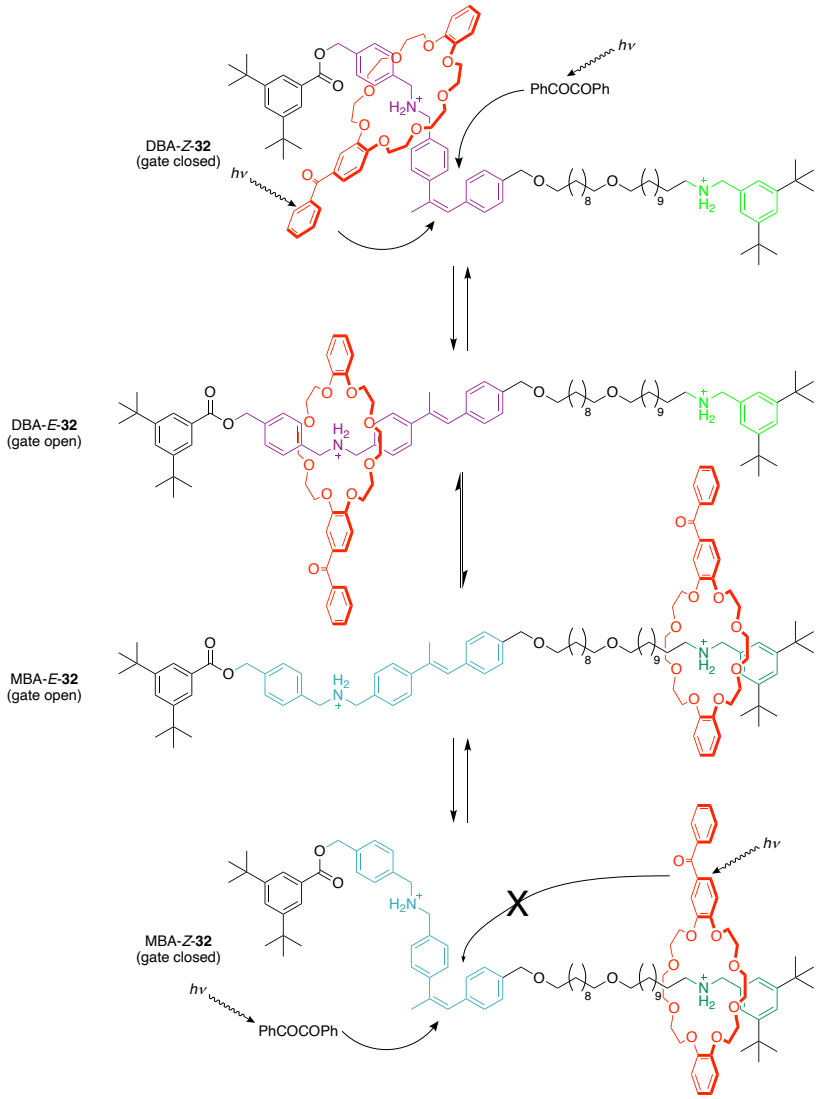


Figure 19. Rotaxane-based light-driven information ratchet by Leigh[50]. For clarity, the photoinduced energy transfer (ET) processes are only shown for the *Z*- forms of the rotaxane, but the same processes occur for the equivalent *E*- translational isomers.

macrocycle distribution out of equilibrium. When the macrocycle is at the MBA station (MBA-*E*-**32** and MBA-*Z*-**32**), only the energy transfer from the benzil can occur due to the long distance between the MBA station and the stilbene unit. Therefore, the stilbene unit remains mostly in the *Z*-configuration. However, when the macrocycle is at the DBA station (DBA-*Z*-**32** and DBA-*E*-**32**), the energy transfer from the benzophenone dominates and the stilbene unit is more likely to be in the *E*-configuration than when the macrocycle is at the MBA station. Consequently, the stilbene unit tends to isomerise from *E*- to *Z*-configuration when the macrocycle shuttles from the DBA station to the MBA station, effectively

preventing the backward shuttling from the MBA station to the DBA station. When this information ratchet is operated in the presence of benzil (5 eq.), the distribution of the macrocycle shifts from the equilibrium distribution of DBA:MBA = 65:35 to the out-of-equilibrium distribution of DBA:MBA = 45:55, which testifies the design principle.

Light-driven information ratchets have enabled not only biasing of the distribution of isomers, but also directional linear motion. Credi developed a system where a macrocycle directionally and autonomously travels over a non-symmetric molecular axle upon irradiation with light[51] (**Figure 20**). The macrocycle is dinaphtho-24-crown-8 (**33**) and the axle consists of a photoswitchable azobenzene unit, an ammonium binding site for the macrocycle and a methylcyclopentyl pseudo-stopper (*EZ*-**34**). When the axle is in the *E*-configuration (*E*-**34**), the macrocycle threads preferentially from the azobenzene side of the axle to form **33**·*E*-**34** due to kinetic reasons (**Figure 20b**, left). If the system is operated with irradiation at 365 nm wavelength, the ratio of concentrations of the products and the reactants for photoisomerisation is larger for the interlocked species than for the non-interlocked axle (i.e., $K_{2,h\nu} > K_{4,h\nu}$). Therefore, the photoisomerisation of **33**·*E*-**34** to **33**·*Z*-**34** occurs preferentially compared to the photoisomerisation of *E*-**34** to *Z*-**34**. The *E*-to-*Z* photoisomerisation both raises the kinetic barrier for dethreading of the macrocycle at the azobenzene side and destabilises the macrocycle on the ammonium station (**Figure 20b**, right, $K_1 > K_3$). Consequently, some macrocycles of **33**·*Z*-**34** dethread and this dethreading preferentially occurs from the methylcyclopentyl side to form *Z*-**34**. The light irradiation isomerises *Z*-**34** back to *E*-**34**, which completes the operating cycle. During this cycle, there is a net directional linear motion of the macrocycle over the axle, where the macrocycle enters from the azobenzene side and leaves from the methylcyclopentyl side.

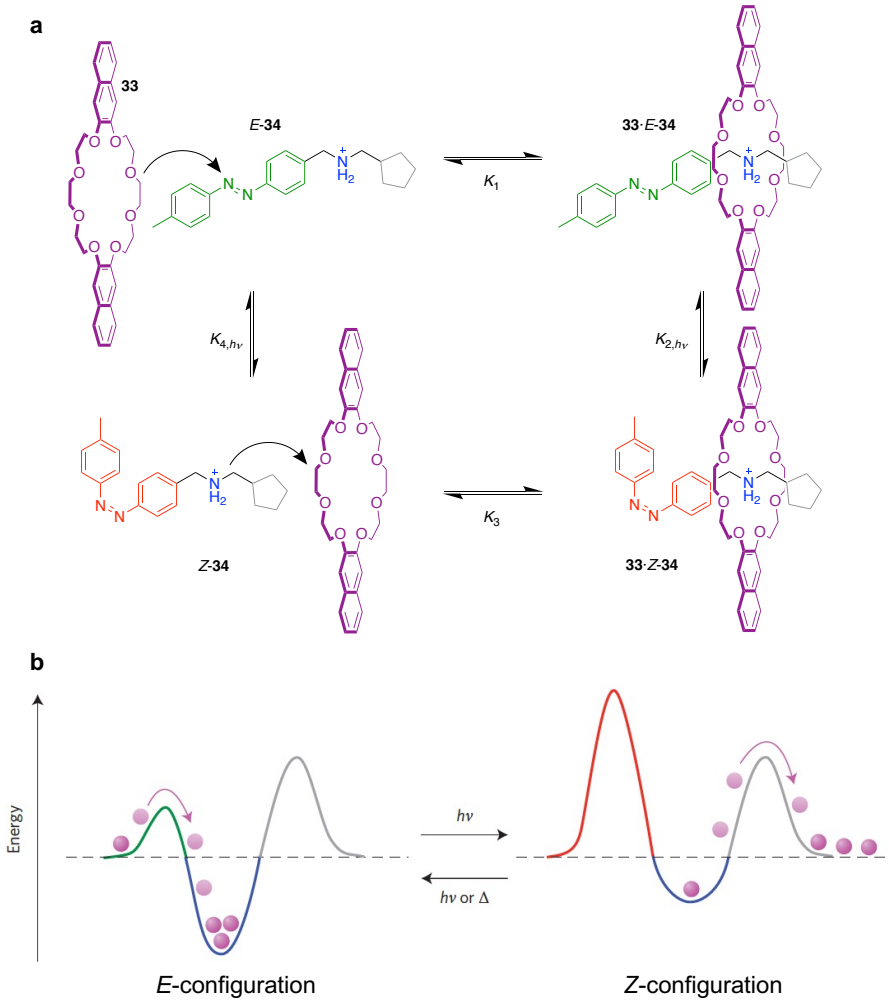


Figure 20. Light-powered autonomous and directional motion of a macrocycle over an axle by Credi[51]. **a**, reaction network of the system. K_1 and K_3 are equilibrium constants of the self-assembly processes, while $K_{2,h\nu}$ and $K_{4,h\nu}$ are the ratios of concentrations of the products and the reactants for respective photoisomerisation under the photostationary state. The indicated parameters refer to the reactions read from left to right and from top to bottom. **b**, simplified potential-energy surface for the macrocycle **33** along the molecular axle *EZ*-**34**. Adapted by permission from Ragazzon, G., Baroncini, M., Silvi, S., Venturi, M. & Credi, A. Light-powered autonomous and directional molecular motion of a dissipative self-assembling system. *Nat. Nanotechnol.* **10**, 70–75 (2015), Springer Nature.

1.5. Theoretical analysis of chemically driven molecular motors

This section reviews theoretical understanding of chemically driven molecular motors, specifically biological motor proteins. The aim is to set the context of the author's PhD project related to theory (Chapter 4). To describe the function of motor proteins, two mechanisms have been proposed: a power stroke mechanism and a Brownian ratchet mechanism[52]. These two mechanisms will be first reviewed individually, and then their

relationship will be discussed. Note that the description in the preceding sections of this chapter has been based on the Brownian ratchet mechanism (Section 1.4.), because it is generally applicable to various examples of synthetic molecular machines. However, this does not mean that the power stroke mechanism is completely irrelevant (see Section 1.5.3. for the relationship between the two mechanisms).

1.5.1. Power stroke mechanism

A power stroke is "a viscoelastic, free-energy releasing, large-amplitude conformational change"[49,53] which results in a nearly irreversible directional motion of a protein motor. In the examples of synthetic molecular machines, the power stroke refers to energetically downhill mechanical processes, such as the displacement of a crown ether from a triazolium station to a dibenzylammonium station[32] (**Figure 12**). Researchers who support the power stroke mechanism think that the power stroke is an essential element of a stepping mechanism of protein motors[52].

Two views have been proposed regarding how the power stroke is realised. One explanation is that the molecular motor is strained like a spring and release of the elastic energy stored in the strained state through a mechanical forward step results in a power stroke[53,54]. However, this idea is unlikely to be applicable to molecular motors. It has been shown that the magnitude of elastic deformation in a motor protein to store the free energy that can be released in the power stroke is comparable to or smaller than the amplitude of thermal fluctuation[52]. This suggests that it is unrealistic to utilise elastic strain for generating a power stroke. Additionally, this explanation does not apply to some of the power strokes observed in synthetic molecular machines. For example, displacement of a crown ether from

a triazolium station to a dibenzylammonium station (**Figure 12**) is a power stroke, but no release of elastic energy is involved.

The other explanation is that a power stroke results from a transition from a high-energy conformation to a low-energy conformation[54]. This description is applicable both to biological and synthetic molecular machines, although in some cases of synthetic molecular machines, it is a co-conformational change, not a conformational change, that is involved in a power stroke.

1.5.2. Brownian ratchet mechanism

A Brownian ratchet mechanism refers to biasing or rectifying the random Brownian motion of the motor to generate directional motion[12,23-25]. As was discussed in Section 1.4., there are two types of Brownian ratchet mechanisms depending on how energy input and asymmetry in the potential-energy surfaces are provided: energy ratchets (Section 1.4.1.) and information ratchets (Section 1.4.2.). Non-autonomous chemically driven molecular motors can be realised by an energy ratchet mechanism, but autonomous chemically driven molecular motors require an information ratchet mechanism. Here, autonomy is a property of continuing to operate over multiple cycles under constant operating conditions, as long as the energy source is present. Synthetic examples such as Stoddart's molecular pump[33] (**Figure 10**) is non-autonomous, whereas examples such as Leigh's catenane rotary motor fueled by Fmoc-Cl[46] (**Figure 17**) is autonomous. Because energy ratchets necessitate alternation of operating conditions to switch between multiple potential-energy surfaces, they cannot realise chemically driven fully autonomous molecular motors.

Analysis of the mechanism of biological molecular machines has so far been carried out based on the concept of the information ratchet, because they are chemically driven autonomous molecular machines. Significant contributions in this direction have been made by Astumian. He argues that a power stroke is irrelevant in determining directionality, stopping force and optimal efficiency of chemically driven molecular motors[49]. Instead, he has demonstrated that these properties are determined by the relative heights of the energy barriers between states, which mainly determines the parameter called kinetic asymmetry[40]. In the case of **Figure 14b**, the kinetic asymmetry can be related to $\Delta\Delta G_F^{\ddagger} + \Delta\Delta G_W^{\ddagger}$. Astumian further proposed intrinsic directionality[55], or a ratcheting constant[56], as a quantity to determine the directionality of chemically driven molecular motors (see Chapter 4, Section 4.1., **Box I** for its mathematical expression).

1.5.3. Relationship between the power stroke mechanism and the Brownian ratchet mechanism

Although the power stroke and the Brownian ratchet mechanism are sometimes thought to be contradictory to one another, they are actually not mutually exclusive[57,58]. In the mechanism of energy ratchets (Section 1.4.1., **Figure 7**), power strokes (displacement of the Brownian particle from the blue well to the yellow well, and from the red well to the green well) are crucial for generating net directional motion. In the mechanism of information ratchets as well (Section 1.4.2., **Figure 14**), a power stroke can be present if the standard chemical potential of state $\mathbf{M_a}$ is lower than $\mathbf{M_b}$, or if the standard chemical potential of $\mathbf{M_b'}$ is lower than $\mathbf{M_a'}$. However, a power stroke is not crucial for information ratchets to generate net directional motion, as is evident from the example such as Leigh's autonomous catenane rotary motor[46] (**Figure 17**). In terms of theory, there is an attempt to discuss the power stroke and the Brownian ratchet mechanism from a comprehensive framework of a transition-state

model and stochastic thermodynamics by Wagoner and Dill[57]. They found that, under all of conditions they studied, power stroke motors are faster, more powerful, and more efficient at constant velocity than Brownian ratchet motors without power strokes.

1.6. Challenges in the field and aims of the projects in this thesis

Decades of efforts by synthetic chemists, physicists and biophysicists have led to significant advance in the development of synthetic molecular machines and the understanding of principles common to all molecular machines for motility and transportation. However, there are still fundamental challenges that need to be addressed to develop molecular machines with more sophisticated, life-like functions and to understand their operating mechanisms deeply. One of these challenges is that most of the synthetic molecular machines reported to date have only one movable part, such as a switching part or a rotating part. For artificial molecular machines to achieve complex and sophisticated tasks, coordination of multiple units is necessary. Another challenge is that there are limited number of examples of chemically fueled autonomous molecular machines. The fact that most of biological molecular machines fall into this category, implies that development of this type of synthetic molecular machines will be crucial for future nanotechnology with life-like properties. Finally, theoretical understanding of autonomous molecular machines is not complete. Current understanding is based on kinetics (e.g., the kinetic asymmetry principle, Section 1.5.2.) and has limited thermodynamic perspective, which makes it difficult to understand the thermodynamics of molecular machines (e.g., how the free energy supplied from the chemical fuel is converted and utilised in molecular machines). Theoretical understanding based on thermodynamics will enable us to define efficiency of each process in molecular machines, which will be a useful quantity to compare different design strategies. Such a framework will also provide practical implications for designing and

improving molecular machines. In the following chapters, the author describes our attempts to address these challenges. Chapter 2 details a molecular machine that coordinates operation of two hydrazone-type switches to transport a cargo molecule. This project is built on the molecular transporter previously developed by the Leigh group[22] (**Figure 6**). Chapter 3 presents a chemically fuelled autonomous molecular pump that operates *via* an information ratchet mechanism[59]. This machine utilises Fmoc nitrophenol as a chemical fuel to drive the system out of equilibrium, which is similar to the autonomous catenane rotary motor reported by Leigh[46] (**Figure 17**), and progressively pumps multiple macrocycles onto the machine from a bulk solution. Chapter 4 describes an analysis of the synthetic molecular rotary motor that operates *via* an information ratchet mechanism[46] with a framework of information thermodynamics[60]. Information thermodynamics allows us to quantitatively relate information and other thermodynamic variables such as entropy and free energy, and has been established through the efforts to resolve the paradox of Maxwell's demon[38,39] (**Figure 13**). Inspired by the similarity between the mechanisms of Maxwell's demon and information ratchets, we applied this framework to a chemically driven information ratchet. It offered in-depth thermodynamic insights, as well as clarification of the role of power strokes in the Brownian ratchet mechanism (Section 1.5.3.). Finally, the conclusion and outlook summarise each of the chapters and discuss future prospects of the field.

1.7. References

1 Kay, E. R., Leigh, D. A. & Zerbetto, F. Synthetic molecular motors and mechanical machines. *Angew. Chem. Int. Ed.* **46**, 72–191 (2007).

2 Schliwa, M. & Woehlke, G. Molecular motors. *Nature* **422**, 759–765 (2003).

3 Amos, L. A. Molecular motors: not quite like clockwork. *Cell. Mol. Life Sci.* **65**, 509–515 (2008).

4 Vale, R. D. & Milligan, R. A. The way things move: looking under the hood of molecular motor proteins. *Science* **288**, 88–95 (2000).

5 Gouaux, E. & MacKinnon, R. Principles of selective ion transport in channels and pumps. *Science* **310**, 1461–1465 (2005).

6 Jorgensen, P. L., Håkansson, K. O. & Karlish, S. J. Structure and mechanism of Na, K-ATPase: functional sites and their interactions. *Annu. Rev. Physiol.* **65**, 817–849 (2003).

7 Yonath, A. Hibernating bears, antibiotics, and the evolving ribosome (Nobel Lecture). *Angew. Chem. Int. Ed.* **49**, 4340–4354 (2010).

8 Von Ballmoos, C., Wiedenmann, A. & Dimroth, P. Essentials for ATP synthesis by F_1F_0 ATP synthases. *Annu. Rev. Biochem.* **78**, 649–672 (2009).

9 Nguyen, T. D. *et al.* A reversible molecular valve. *Proc. Natl. Acad. Sci. USA* **102**, 10029–10034 (2005).

10 Chen, K.-Y. *et al.* Control of surface wettability using tripodal light-activated molecular motors. *J. Am. Chem. Soc.* **136**, 3219–3224 (2014).

11 Erbas-Cakmak, S., Leigh, D. A., McTernan, C. T. & Nussbaumer, A. L. Artificial molecular machines. *Chem. Rev.* **115**, 10081–10206 (2015).

12 Astumian, R. D. Design principles for Brownian molecular machines: how to swim in molasses and walk in a hurricane. *Phys. Chem. Chem. Phys.* **9**, 5067–5083 (2007).

13 *Molecular Motors*. (eds. Schliwa, M., Wiley-VCH, 2003).

14 Shinkai, S., Nakaji, T., Ogawa, T., Shigematsu, K. & Manabe, O. Photoresponsive crown ethers. 2. Photocontrol of ion extraction and ion transport by a bis (crown ether) with a butterfly-like motion. *J. Am. Chem. Soc.* **103**, 111–115 (1981).

15 Landge, S. M. & Aprahamian, I. A pH activated configurational rotary switch: controlling the *E*/*Z* isomerization in hydrazones. *J. Am. Chem. Soc.* **131**, 18269–18271 (2009).

16 Su, X. & Aprahamian, I. Switching around two axles: controlling the configuration and conformation of a hydrazone-based switch. *Org. Lett.* **13**, 30–33 (2010).

17 Muraoka, T., Kinbara, K. & Aida, T. Mechanical twisting of a guest by a photoresponsive host. *Nature* **440**, 512–515 (2006).

18 Livoreil, A., Dietrich-Buchecker, C. O. & Sauvage, J.-P. Electrochemically triggered swinging of a [2]-catenate. *J. Am. Chem. Soc.* **116**, 9399–9400 (1994).

19 Bissell, R. A., Córdova, E., Kaifer, A. E. & Stoddart, J. F. A chemically and electrochemically switchable molecular shuttle. *Nature* **369**, 133–137 (1994).

20 Chen, J., Wezenberg, S. J. & Feringa, B. L. Intramolecular transport of small-molecule cargo in a nanoscale device operated by light. *Chem. Commun.* **52**, 6765–6768 (2016).

21 Koumura, N., Zijlstra, R. W., van Delden, R. A., Harada, N. & Feringa, B. L. Light-driven monodirectional molecular rotor. *Nature* **401**, 152–155 (1999).

22 Kassem, S., Lee, A. T., Leigh, D. A., Markevicius, A. & Solà, J. Pick-up, transport and release of a molecular cargo using a small-molecule robotic arm. *Nat. Chem.* **8**, 138–143 (2016).

23 Chatterjee, M. N., Kay, E. R. & Leigh, D. A. Beyond switches: ratcheting a particle energetically uphill with a compartmentalized molecular machine. *J. Am. Chem. Soc.* **128**, 4058–4073 (2006).
24 Astumian, R. D. Thermodynamics and kinetics of a Brownian motor. *Science* **276**, 917–922 (1997).
25 Reimann, P. Brownian motors: noisy transport far from equilibrium. *Phys. Rep.* **361**, 57–265 (2002).
26 Barrell, M. J., Campaña, A. G., von Delius, M., Geertsema, E. M. & Leigh, D. A. Light-driven transport of a molecular walker in either direction along a molecular track. *Angew. Chem. Int. Ed.* **50**, 285–290 (2011).
27 Kelly, T. R., De Silva, H. & Silva, R. A. Unidirectional rotary motion in a molecular system. *Nature* **401**, 150–152 (1999).
28 Collins, B. S., Kistemaker, J. C., Otten, E. & Feringa, B. L. A chemically powered unidirectional rotary molecular motor based on a palladium redox cycle. *Nat. Chem.* **8**, 860–866 (2016).
29 Zhang, Y. *et al.* A chemically driven rotary molecular motor based on reversible lactone formation with perfect unidirectionality. *Chem* **6**, 2420–2429 (2020).
30 Leigh, D. A., Wong, J. K., Dehez, F. & Zerbetto, F. Unidirectional rotation in a mechanically interlocked molecular rotor. *Nature* **424**, 174–179 (2003).
31 Hernández, J. V., Kay, E. R. & Leigh, D. A. A reversible synthetic rotary molecular motor. *Science* **306**, 1532–1537 (2004).
32 Erbas-Cakmak, S. *et al.* Rotary and linear molecular motors driven by pulses of a chemical fuel. *Science* **358**, 340–343 (2017).
33 Cheng, C. *et al.* An artificial molecular pump. *Nat. Nanotechnol.* **10**, 547–553 (2015).
34 Berrocal, J. A., Biagini, C., Mandolini, L. & Di Stefano, S. Coupling of the decarboxylation of 2-cyano-2-phenylpropanoic acid to large-amplitude motions: a convenient fuel for an acid-base-operated molecular switch. *Angew. Chem. Int. Ed.* **55**, 6997–7001 (2016).
35 Brown, B. R. The mechanism of thermal decarboxylation. *Q. Rev. Chem. Soc.* **5**, 131–146 (1951).
36 Zhou, H. & Chen, Y. Chemically driven motility of Brownian particles. *Phys. Rev. Lett.* **77**, 194–197 (1996).
37 Astumian, R. D. & Derényi, I. Fluctuation driven transport and models of molecular motors and pumps. *Eur. Biophys. J.* **27**, 474–489 (1998).
38 *Maxwell's Demon 2. Entropy, Classical and Quantum Information, Computing*. (eds. Leff, H. S., Rex, A. F., Institute of Physics Publishing, 2003).
39 Parrondo, J. M. R., Horowitz, J. M. & Sagawa, T. Thermodynamics of information. *Nat. Phys.* **11**, 131–139 (2015).
40 Astumian, R. D. Kinetic asymmetry allows macromolecular catalysts to drive an information ratchet. *Nat. Commun.* **10**, 3837 (2019).
41 Amano, S., Borsley, S., Leigh, D. A. & Sun, Z. Chemical engines: driving systems away from equilibrium through catalyst reaction cycles. *Nat. Nanotechnol.* **16**, 1057–1067 (2021).
42 Fletcher, S. P., Dumur, F., Pollard, M. M. & Feringa, B. L. A reversible, unidirectional molecular rotary motor driven by chemical energy. *Science* **310**, 80–82 (2005).
43 Lin, Y., Dahl, B. J. & Branchaud, B. P. Net directed 180° aryl–aryl bond rotation in a prototypical achiral biaryl lactone synthetic molecular motor. *Tetrahedron Lett.* **46**, 8359–8362 (2005).

44 Alvarez-Perez, M., Goldup, S. M., Leigh, D. A. & Slawin, A. M. A chemically-driven molecular information ratchet. *J. Am. Chem. Soc.* **130**, 1836–1838 (2008).
45 Carlone, A., Goldup, S. M., Lebrasseur, N., Leigh, D. A. & Wilson, A. A three-compartment chemically-driven molecular information ratchet. *J. Am. Chem. Soc.* **134**, 8321–8323 (2012).
46 Wilson, M. R. *et al.* An autonomous chemically fuelled small-molecule motor. *Nature* **534**, 235–240 (2016).
47 Borsley, S., Leigh, D. A. & Roberts, B. M. W. A doubly kinetically-gated information ratchet autonomously driven by carbodiimide hydration. *J. Am. Chem. Soc.* **143**, 4414–4420 (2021).
48 Martin, C. J., Lee, A. T. L., Adams, R. W. & Leigh, D. A. Enzyme-mediated directional transport of a small-molecule walker with chemically identical feet. *J. Am. Chem. Soc.* **139**, 11998–12002 (2017).
49 Astumian, R. D. Irrelevance of the power stroke for the directionality, stopping force, and optimal efficiency of chemically driven molecular machines. *Biophys. J.* **108**, 291–303 (2015).
50 Serreli, V., Lee, C.-F., Kay, E. R. & Leigh, D. A. A molecular information ratchet. *Nature* **445**, 523–527 (2007).
51 Ragazzon, G., Baroncini, M., Silvi, S., Venturi, M. & Credi, A. Light-powered autonomous and directional molecular motion of a dissipative self-assembling system. *Nat. Nanotechnol.* **10**, 70–75 (2015).
52 Hwang, W. & Karplus, M. Structural basis for power stroke vs. Brownian ratchet mechanisms of motor proteins. *Proc. Natl. Acad. Sci. USA* **116**, 19777–19785 (2019).
53 Howard, J. Protein power strokes. *Curr. Biol.* **16**, R517–R519 (2006).
54 Leibler, S. & Huse, D. A. Porters versus rowers: a unified stochastic model of motor proteins. *J. Cell. Biol.* **121**, 1357–1368 (1993).
55 Astumian, R. D. Stochastic conformational pumping: a mechanism for free-energy transduction by molecules. *Annu. Rev. Biophys.* **40**, 289–313 (2011).
56 Ragazzon, G. & Prins, L. J. Energy consumption in chemical fuel-driven self-assembly. *Nat. Nanotechnol.* **13**, 882–889 (2018).
57 Wagoner, J. A. & Dill, K. A. Molecular motors: power strokes outperform Brownian ratchets. *J. Phys. Chem. B* **120**, 6327–6336 (2016).
58 Wang, H. & Oster, G. Ratchets, power strokes, and molecular motors. *Appl. Phys. A* **75**, 315–323 (2002).
59 Amano, S., Fielden, S. D. P. & Leigh, D. A. A catalysis-driven artificial molecular pump. *Nature* **594**, 529–534 (2021).
60 Amano, S. *et al.* Insights from an information thermodynamics analysis of a synthetic molecular motor. *ChemRxiv*, doi:10.33774/chemrxiv-2021-60k1r (2021).

Chapter 2.

A Unidirectional Molecular Transporter Utilising Multiple Hydrazone-type Switches

Blank Page

Declaration

Preliminary work in this chapter has been submitted to the University of Manchester to support the author's application for the degree of Master of Science by Research in 2018.

Acknowledgements

The author's contribution to this work includes the design of the system, model studies of cargo transportation, synthesis and characterisation of the intermediates, and writing of the main text and the experimental section. Dr. Vanesa Marcos Algaba and Dr. Salma Kassem contributed to the design of the system. Dr. Salma Kassem, Dr. Marcel Dommaschk, Dr. Shaofei Zhang, Dr. Javier Echavarren and Dr. Albano Galan helped the synthesis and characterisation. Prof. David A. Leigh directed the research. Everyone involved in this project are gratefully thanked for their contribution. Dr. Stefan Borsley is gratefully thanked for proofreading this chapter.

Synopsis

This chapter details the molecular machine that directionally transports cargo by the coordinated operation of two switches. Most synthetic molecular machines to date have only one movable part, such as a switching or rotating part. For artificial molecular machines to achieve complex and sophisticated tasks, the coordination of multiple units in one machine is necessary. Therefore, this project aims to create a molecular machine that directionally transports a cargo molecule, 3-mercaptopropanehydrazide, between two aldehyde stations by the coordinated motion of two hydrazone switches. These hydrazone switches change their configuration and conformation depending on the protonation state, allowing them to pick up and transport a cargo molecule between them. The cargo molecule is picked up from

an initial binding site by one arm, transferred to the other and eventually deposited at the final site. This transportation mechanism allows the cargo to be passed between the sites and arms without dissociation from the machine. So far, model studies of the cargo transfer between the two arms have been conducted and the synthesis of the machine molecule is ongoing. The success of this project will pave the way to molecular machines with coordinated dynamics that perform sophisticated tasks reminiscent of biological systems.

2.1. Introduction

Biology utilises mechanical motion of molecules to achieve various tasks such as synthesis[1,2], intracellular transportation[3,4] and motility[5]. For example, metazoan fatty acid synthase passes a growing fatty acid chain between enzyme domains in the protein superstructure in order to synthesise fatty acid[6-8]. The possibility of mechanical manipulation of matter at a nanoscale by artificial systems was first proposed by Feynman in his lecture 'There's a plenty of room at the bottom[9]'. Another conceptual origin of mechanical manipulation of matter at a nanoscale is Drexler's suggestion of nanomachines, which should be capable of positioning atoms, or clusters of atoms, and could potentially be used for assembling almost any substance or material[10,11]. However, Smalley[12] and others[13,14] have argued that such molecular 'Drexlerian assemblers' are impractical. This is because the manipulator hand of the assembler itself is made of atoms, which makes it impossible to accommodate in a nanoscale reaction region all manipulators necessary for complete control of the assembling process (so-called 'fat fingers' problems). The manipulator hand is also 'sticky' in that the atoms that constitute the manipulator hand will adhere to the atoms that are being transported. This makes it difficult to release the atoms at the right position (so-called 'sticky fingers' problems). Still, these conceptual discussions have been helpful for elucidating practical pathways to realise synthetic systems for mechanical manipulation.

To date, numerous synthetic molecular machines for mechanical manipulation have been reported, such as a photo-responsive host molecule that mechanically twists a guest molecule[15] (Chapter 1, **Figure 1**), small-molecule machines that perform tasks at the macroscopic level[16-19], a DNA-based robotic arm[20] and a dynamic system that can transport gold nanoparticles[21]. However, most of these systems have only one movable part per machine, such as a switching or rotating part. For synthetic molecular machines to achieve complex and sophisticated tasks, the coordinated motion of multiple units within one

machine is necessary. Therefore, the project in this chapter aims to develop a molecular machine that transports a cargo molecule by the coordinated motion of two hydrazone-type switches (**1** in **Figure 1**).

The hydrazone-type switch utilised in the project described in this chapter was developed by Aprahamian, and changes their configuration and conformation depending on the protonation state[22,23] (Chapter 1, **Figure 2**). This hydrazone switch was utilised in a molecular transporter by Leigh[24], which transports a cargo molecule between the stations by the isomerisation of the hydrazone switch (Chapter 1, **Figure 6**). The project in this chapter is built upon this previous transporter, and aims at transporting a cargo molecule from one side of the machine to the other by coordinating the switching of two hydrazone-type switches. The success of this project will pave the way to molecular machines with coordinated dynamics that perform sophisticated tasks reminiscent of biological systems. In this chapter, the design and the mechanism of the transporter machine will be first explained (Section 2.2.), followed by the initial model study of cargo transfer between two arms and cargo reloading (Section 2.3.). Then, the synthesis of the first design of the machine will be detailed (Section 2.4.). A cargo transfer model study was conducted with a left part of the whole machine (Section 2.5.). Based on its result, we decided to change the design of the machine (Section 2.6. and Section 2.7.). Finally, model studies to prepare for the final step of the synthesis of the machine will be explained (Section 2.8.).

2.2. Design and mechanism

The whole design of the molecular machine is shown in **Figure 1**. The cargo molecule is shown in red. The machine has two aldehyde stations (green and blue), two hydrazone switches (orange), a biphenyl linker between the switches, a mercaptopyridine arm (pink), and a thiol arm (brown). The key component of this machine is the hydrazone switch

developed by Aprahamian[22,23] (Chapter 1, **Figure 2**) and recently utilised in a molecular transporter by Leigh[24] (Chapter 1, **Figure 6**). Compared to the previous molecular transporter[24], there are two key differences in this machine. First, this design has two hydrazone switches oriented in opposite directions. Second, there are two types of the arm: a mercaptopyridine arm and a thiol arm. This is to enable the directional transfer of the cargo molecule from the mercaptopyridine arm to the thiol arm. Due to the stability of the resulting thiopyridone moiety after the transfer, this process is practically irreversible, which generates unidirectionality.

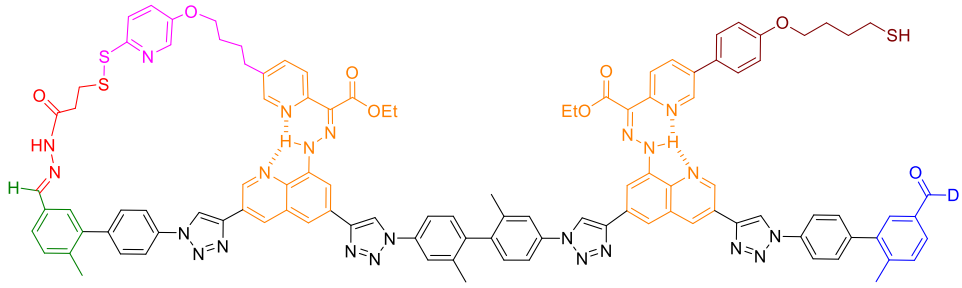


Figure 1. Structure of the molecular transporter with two hydrazone switches (**1**).

The proposed mechanism of this molecular machine is shown in **Figure 2**. Initially, the cargo is bound to the green station *via* a hydrazone bond (**Figure 2a**). The addition of acid and *p*-anisaldehyde facilitates hydrazone exchange, cleaving the cargo molecule from the station. It also changes the configuration of the two hydrazone switches (**Figure 2b**). As the two arms are brought within close proximity of one another, disulfide exchange transfers the cargo molecule from the mercaptopyridine arm to the thiol arm irreversibly (**Figure 2c**). Unlike common disulfide exchange, this process occurs in acidic conditions because the mercaptopyridine is protonated to be a good leaving group. The feasibility of this cargo transfer process was confirmed in the model study described in Section 2.3. After the addition of a large excess of acid, the configuration of the two hydrazone switches returns to the original state and the cargo molecule is connected to the blue station (**Figure 2d**).

Subsequent addition of base stops hydrazone exchange, kinetically locking the cargo at the final site.

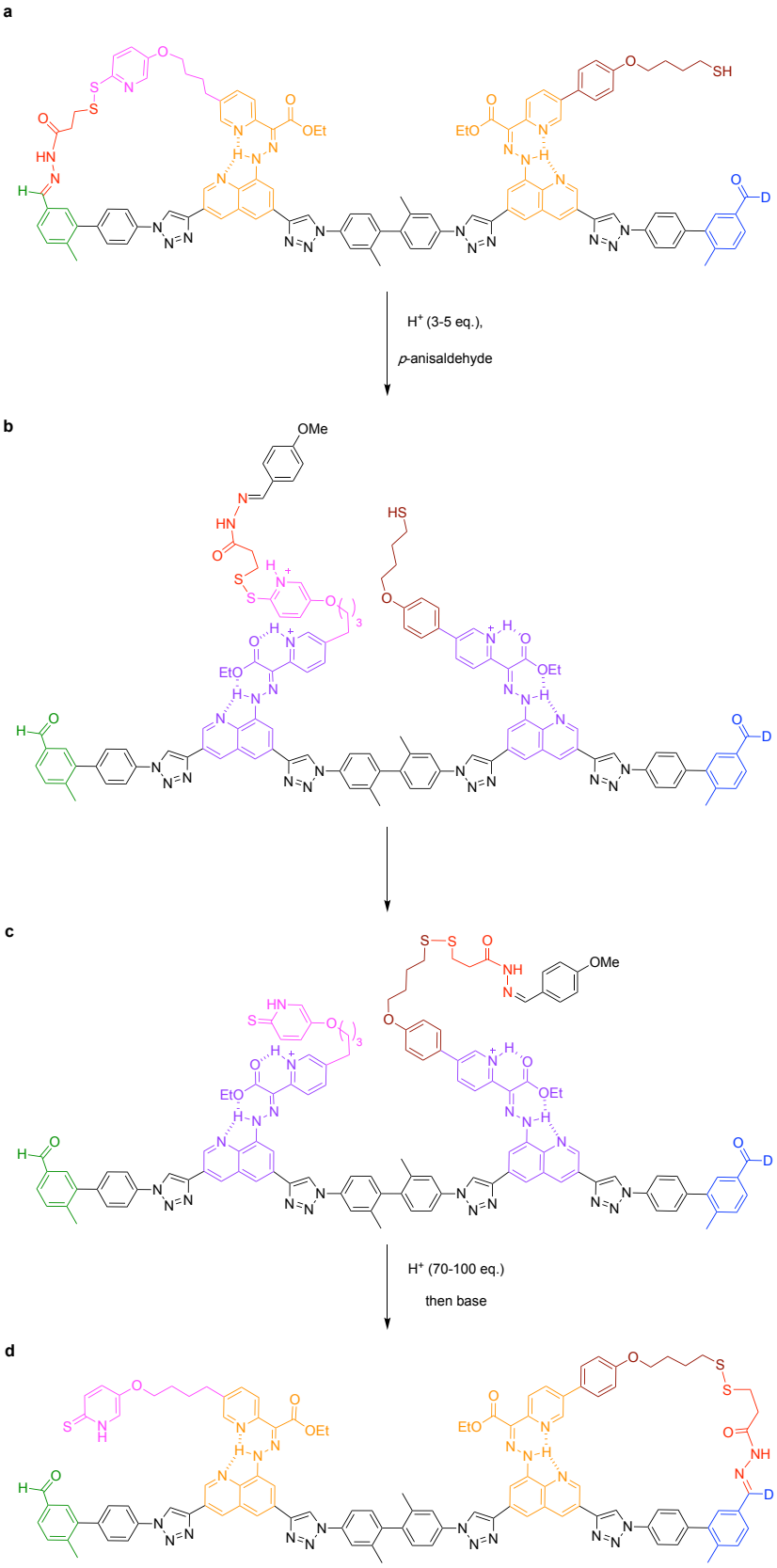


Figure 2. Proposed mechanism of operation of a molecular transporter with two hydrazone switches.

2.3. Model study of cargo transfer and reloading

Initially, a model study was conducted to check if the transfer of the cargo molecule from the mercaptopyridine arm to the thiol arm was possible, and if the mercaptopyridine arm was reloadable with another cargo molecule. The latter property will ensure that the machine is, in principle, reusable for the transportation of multiple cargo molecules. For the model study, compound **5** and **8** were synthesised to represent the two arms (**Figure 3**). In compound **8**, the mercaptopyridine arm was already connected with a cargo molecule by a disulfide bond, which is similar to the initial state of the designed molecular machine (**Figure 2a**).

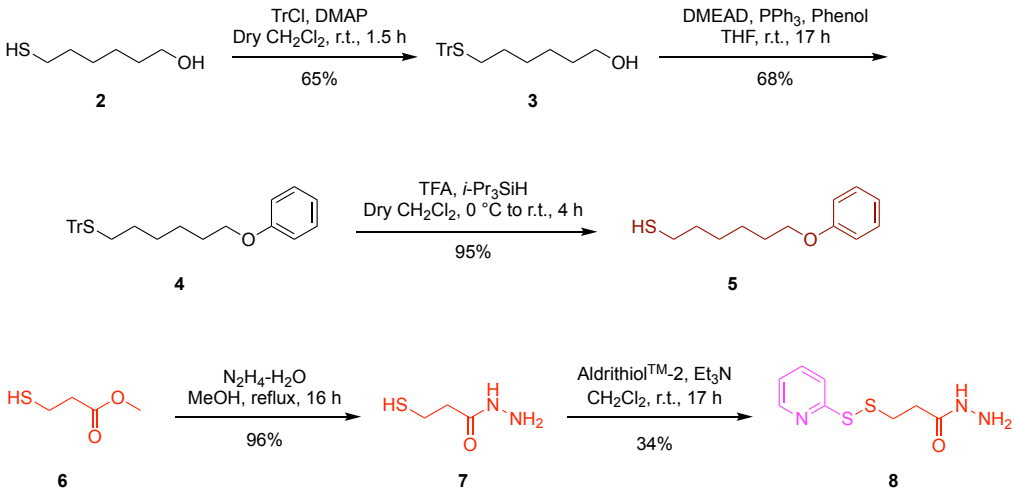
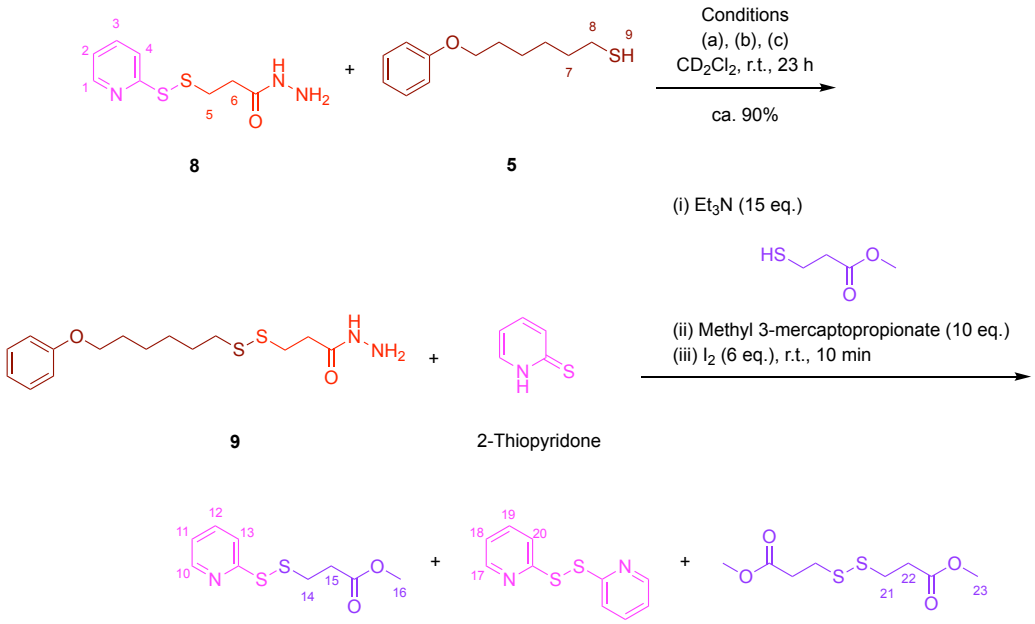


Figure 3. Synthesis of model compounds.

When these two compounds were mixed under acidic and basic conditions (condition (a), (b) and (c) in **Figure 4**), compound **9** and 2-thiopyridone were formed in about 90% yield. This result was confirmed by ^{1}H NMR spectroscopy (**Figure 5**). In all conditions, all signals of the pyridine ring of compound **8** (H_1, H_2, H_3 and H_4) shifted upfield, which indicates the formation of 2-thiopyridone. Furthermore, the decrease of the signal of the thiol in compound **5** (H_9) and the shifts of signal H_5, H_6, H_7 and H_8 indicates the formation of a disulfide bond. This result confirmed that the cargo transfer from the mercaptopyridine arm to the thiol arm was possible in both acidic and basic conditions. This ensured the validity

of the mechanism for the irreversible cargo transfer between the arms in the designed molecular machine.



Conditions	**10**	Aldrithiol^{TM}-2	**11**
(a) AcOH (0.2 eq.)	1	0.25	7.5
(b) TFA (0.2 eq.)	1	0.25	7
(c) Et_3N (0.2 eq.)	1	0.25	8.5

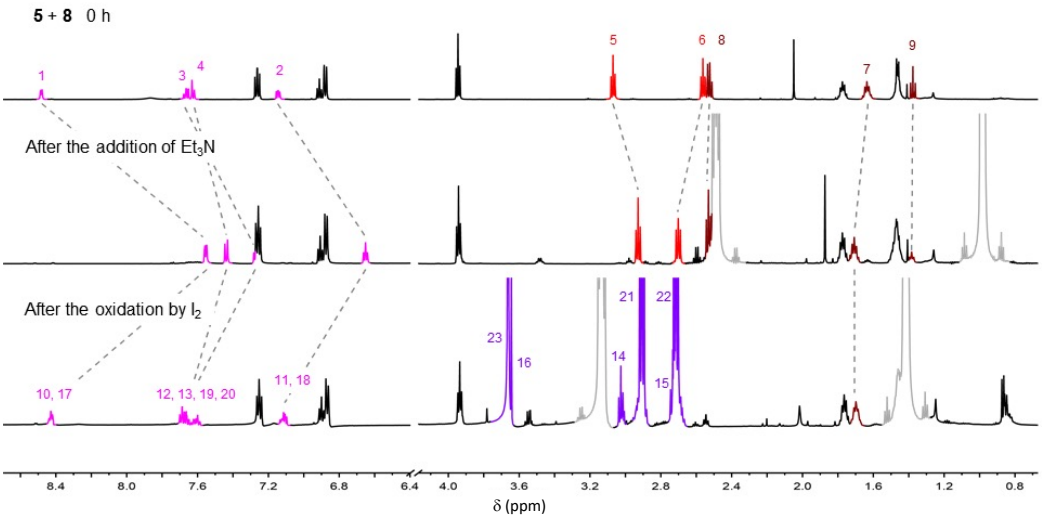
Compound **9** was omitted in the table for clarity.

Figure 4. Results of the model study. The numbers in the table indicate the relative amount of substance.

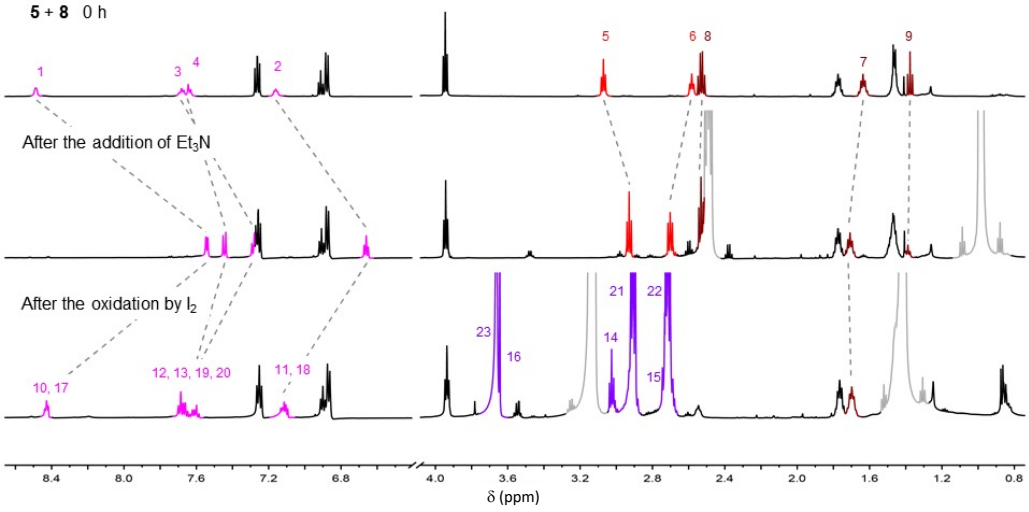
To the resulting mixture after the first step were added excess amount of triethylamine, 10 equivalents of methyl 3-mercaptopropionate as the new cargo molecule, and iodine as oxidant to reload the arm. Triethylamine was added to sequester protons that would be formed after oxidation. A large excess of methyl 3-mercaptopropionate was added to prevent the formation of a disulfide bond between two 2-thiopyridone molecules. The results of the second step were analysed by ^{1}H NMR spectroscopy (**Figure 5**). All signals of the pyridine ring of 2-thiopyridone shifted downfield to the similar chemical shifts with compound **8** (H_{10}, H_{11}, H_{12}, H_{13}, H_{17}, H_{18}, H_{19}, H_{20}). Close investigation revealed that there were two sets of peaks, which indicates the formation of compound **10** and Aldrithiol^{TM}-2. Three large

signals (H_{21}, H_{22} and H_{23}) correspond to the dimer of the cargo molecule (**11**). Based on the integration values of the signals, the relative ratio of the amount of the products was calculated. The results are shown in the bottom table of **Figure 4**. In all conditions of the first step (i.e. 0.2 eq. AcOH, 0.2 eq. TFA, 0.2 eq. Et_3N), the ratio of the three products was almost the same. Pleasingly, compound **10** formed preferentially over Aldrithiol^{TM}-2. Compound **11** formed inevitably through the disulfide formation between two cargo molecules. This result indicates that, after the transfer of the cargo between two arms, the mercaptopyridine arm can be reloaded with another cargo molecule, although dimerisation between two arms occurs as a side reaction. Therefore, the model study could confirm two intended properties of the molecular machine, the ability of irreversible transfer of the cargo between two arms and the reloadability of the first arm. Note that the model study was done with 30 mM of compound **8** and **5**, which is a significantly higher concentration than the concentration of the full machine for the actual operation (ca. 1 mM). This difference may cast a doubt whether the result of this model study can be used to predict the result of the cargo transfer process in the actual operation. However, the cargo transfer in this model study was an intermolecular process, while that in the actual operation is an intramolecular process. Therefore, in the actual operation of the full machine, the effective concentration of the cargo part for the alkyl thiol arm (and vice versa) should be much higher than the concentration of the machine itself (1 mM), which justifies the model study with higher concentrations of the corresponding model compounds (30 mM).

(a) 0.2 eq AcOH



(b) 0.2 eq TFA



(c) 0.2 eq Et_3N

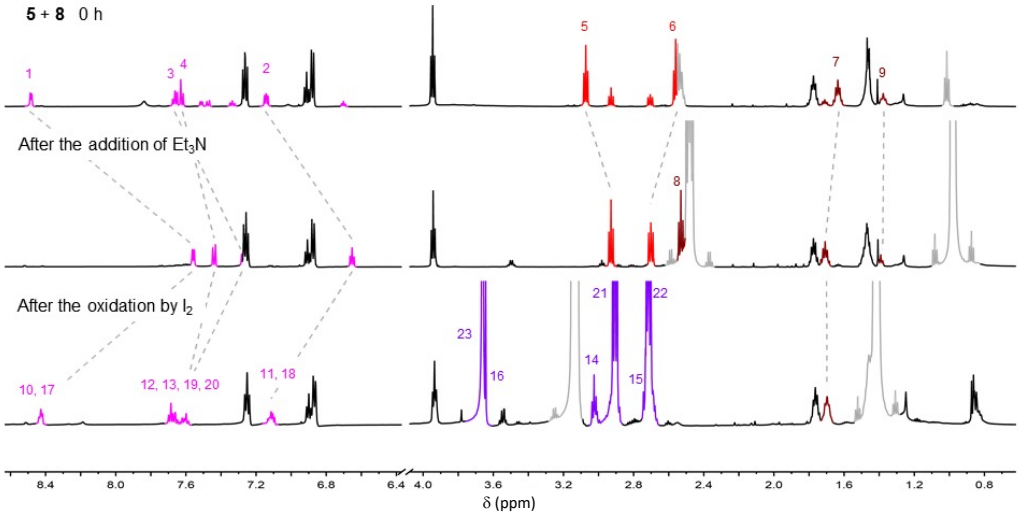


Figure 5. ^{1}H NMR spectra during the model study (600 MHz, CD_2Cl_2, 295 K). **(a)** 0.2 eq. AcOH, **(b)** 0.2 eq. TFA, **(c)** 0.2 eq. Et_3N. In each stack plot, '0 h' is immediately after the mixing of **5**, **8** and the reagent, 'After the addition of Et_3N' is after adding Et_3N after the first step finished, and 'After the oxidation by I_2' is after reacting the mixture with I_2 for 10 min.

2.4. Synthesis for the first design of the machine

After the promising results from the model study, the synthesis of the molecular machine was undertaken. In the retrosynthesis, the target molecule was divided into eight parts (**Figure 6**): cargo **12**, non-deuterated station **13** (the alcohol was oxidized to an aldehyde at a later step), thiopyridone (mercaptopyridine) arm and the upper part of the hydrazone switch **14**, lower parts of the switch **15** and **17**, linker **16**, thiol arm and upper part of the switch **18** and deuterated station **19**. Thiopyridone is a more stable tautomeric form of mercaptopyridine. Since the two hydrazone switches were to be synthesised in different ways (via a hydrazone condensation and the Japp-Klingemann reaction, respectively), two different lower parts of the switch were prepared. The following parts of this section will explain the synthesis of each component and their connection in detail.

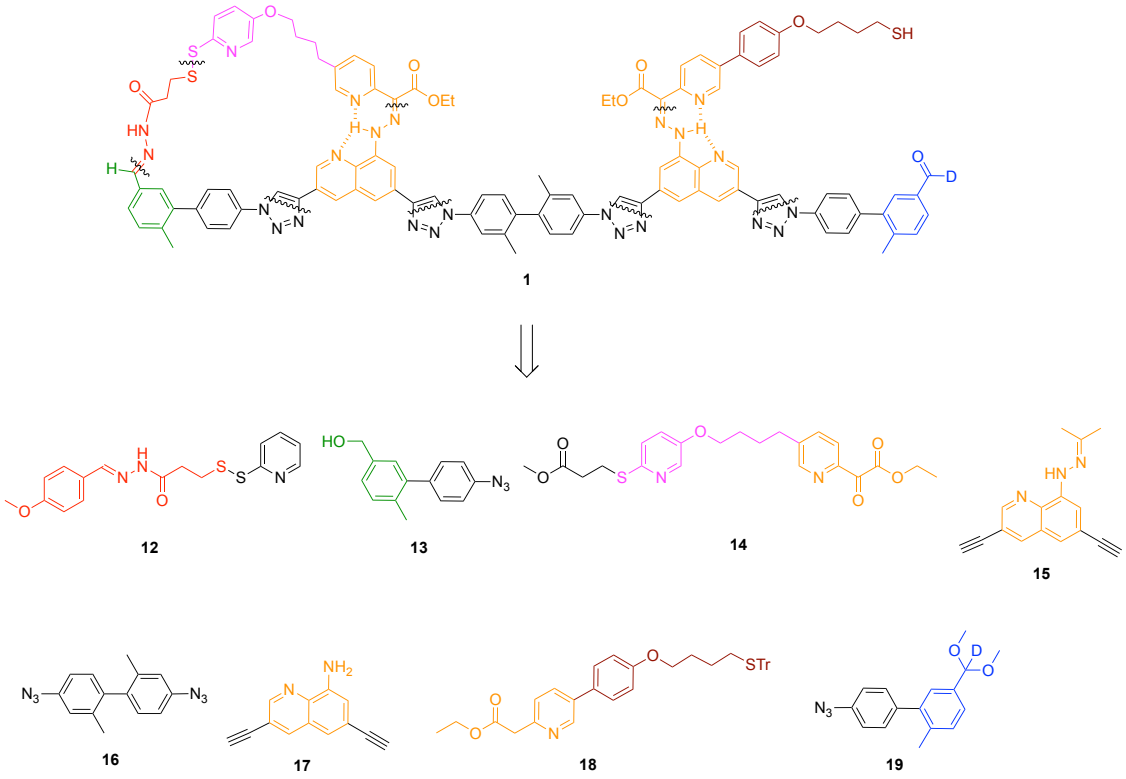


Figure 6. Retrosynthesis.

Cargo molecule **12** was synthesized from a known compound **20**[24] in one step (**Figure 7a**). **Figure 7b** shows the route of the synthesis of the non-deuterated station **13**. Starting from commercially available compound **21** and **22**, **13** was synthesised in three steps. **Figure 7c** shows the route of the synthesis of deuterated station **19** starting from methyl 3-bromo-4-methylbenzoate **21**. Borylation and Suzuki coupling gave ester **26**. Next, DIBAL-D was used as the reducing agent to convert ester **26** to the deuterated benzyl alcohol. The resulting benzyl alcohol was treated with Dess-Martin periodinane to afford the desired deuterium-labelled aldehyde **27** in 81% yield. Acetal **28** was prepared from the corresponding aldehyde in 82% yield, which was finally converted to the deuterium-labelled building block **19** in 95% yield.

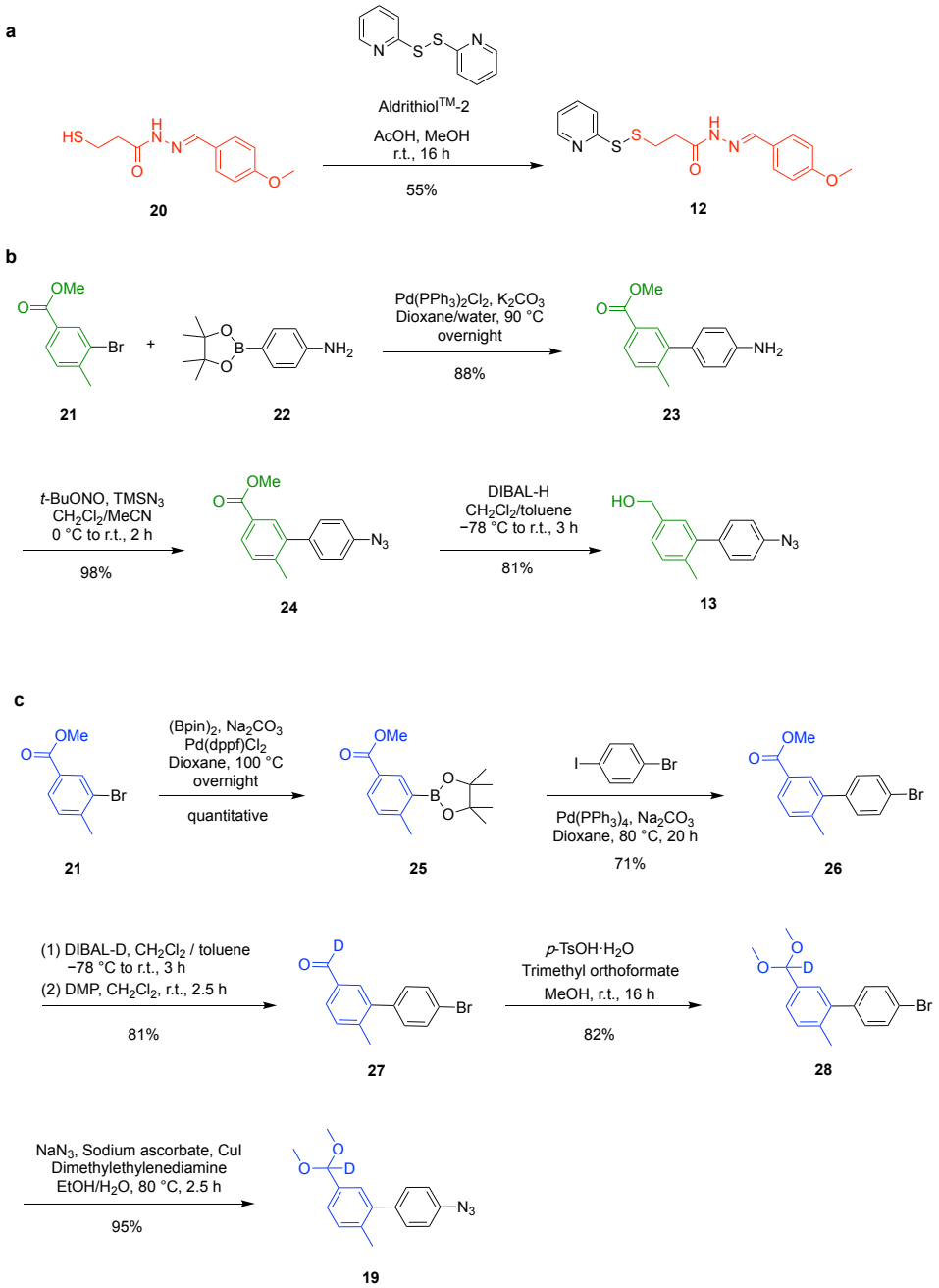


Figure 7. Synthesis of cargo **12** (**a**), left station **13** (**b**), and right station **19** (**c**).

Figure 8 shows the synthesis route of compound **14**. First, the precursor of thiopyridone arm **30** was synthesised from 2-bromo-5-hydroxypyridine **29** in one step (**Figure 8a**). Subsequently, the top half of hydrazone switch **34** was prepared from known compound **31** in three steps (**Figure 8b**). Finally, the top half of switch **34** and the precursor of thiopyridone arm **30** were linked by an S_N2 reaction and a further 2-step modification afforded compound **14** (**Figure 8c**).

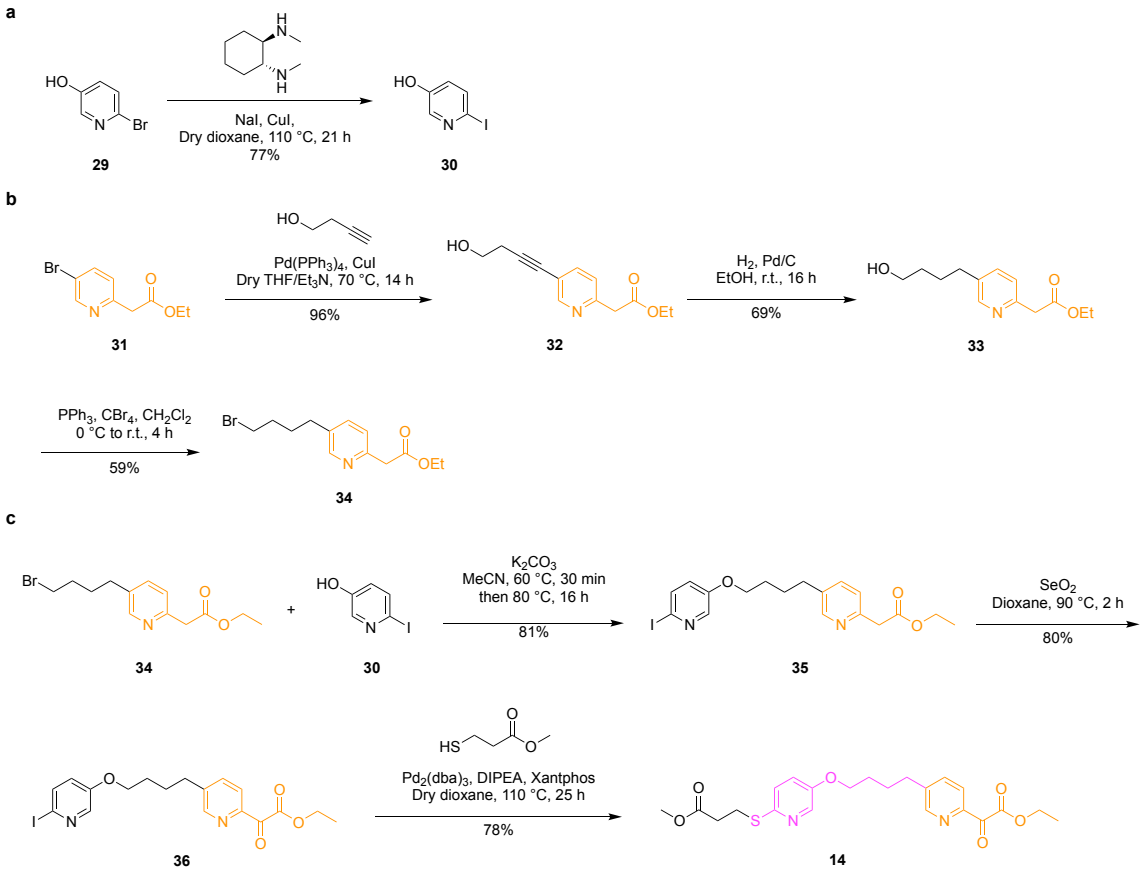


Figure 8. Synthesis of precursor of the thiopyridone arm **30** (**a**), top half of the hydrazone switch **34** (**b**), compound **14** (**c**).

Compound **15**, which forms the core of the lower half of the hydrazone switches, is already known.[25] Linker **16** was prepared from diamine **37** by treatment with trimethylsilyl azide in the presence of *tert*-butyl nitrite (**Figure 9a**). The reaction required low temperatures (−20 °C) and the crude product should be concentrated at room temperature to prevent thermal decomposition of the product. After purification by flash column chromatography, linker **16** was obtained as a colourless oil in 70% yield.

The bottom part of the hydrazone switch **17** was prepared from 8-nitroquinoline **38** (**Figure 9b**). Bromination and reduction proceeded in a moderate to high yield to give **40**. The subsequent double Sonogashira coupling with TMS-acetylene and the deprotection of the trimethylsilyl group gave compound **17** in good yield.

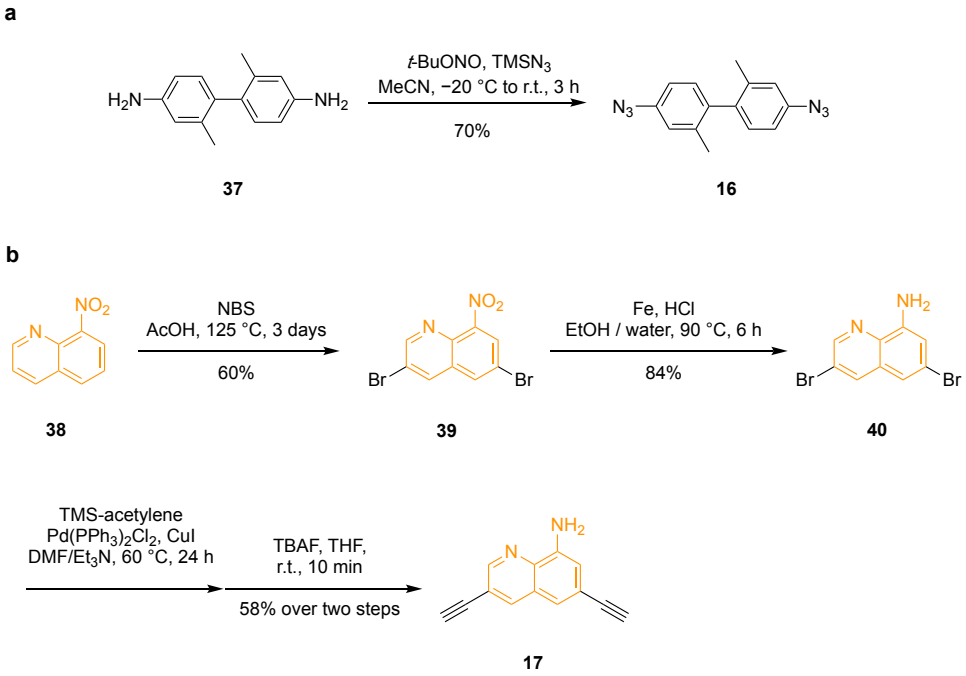


Figure 9. Synthesis of **16** (**a**) and **17** (**b**).

Compound **18** consists of two parts: protected thiol arm **45** and the upper part of the hydrazone switch **46**. The synthesis of compound **45** required 4-mercapto-1-butanol **43** as starting material. First, the preparation of this compound was attempted by the mesylation of one of the hydroxyl groups in 1,4-dibutanol **41** and subsequent nucleophilic substitution with NaSH (**Figure 10a**). However, the second reaction did not proceed and the starting material was largely recovered. Thus, a more reliable route to prepare **45** via the ring-opening of a thiolactone was devised (**Figure 10b**). First, the reduction of γ-thiobutyrolactone **44** with $NaBH_4$ gave compound **43**. Crude **43** was subsequently treated with tritylchloride and DMAP to give alcohol **45** in 57% yield. Compound **46** was synthesised by the Suzuki coupling between ester **31** and 4-hydroxyphenylboronic acid, catalysed by $Pd(AmPhos)_2Cl_2$ in 94% yield. Finally, alcohol **45** and phenol **46** were connected by the Mitsunobu reaction in 81% yield to give compound **18**.

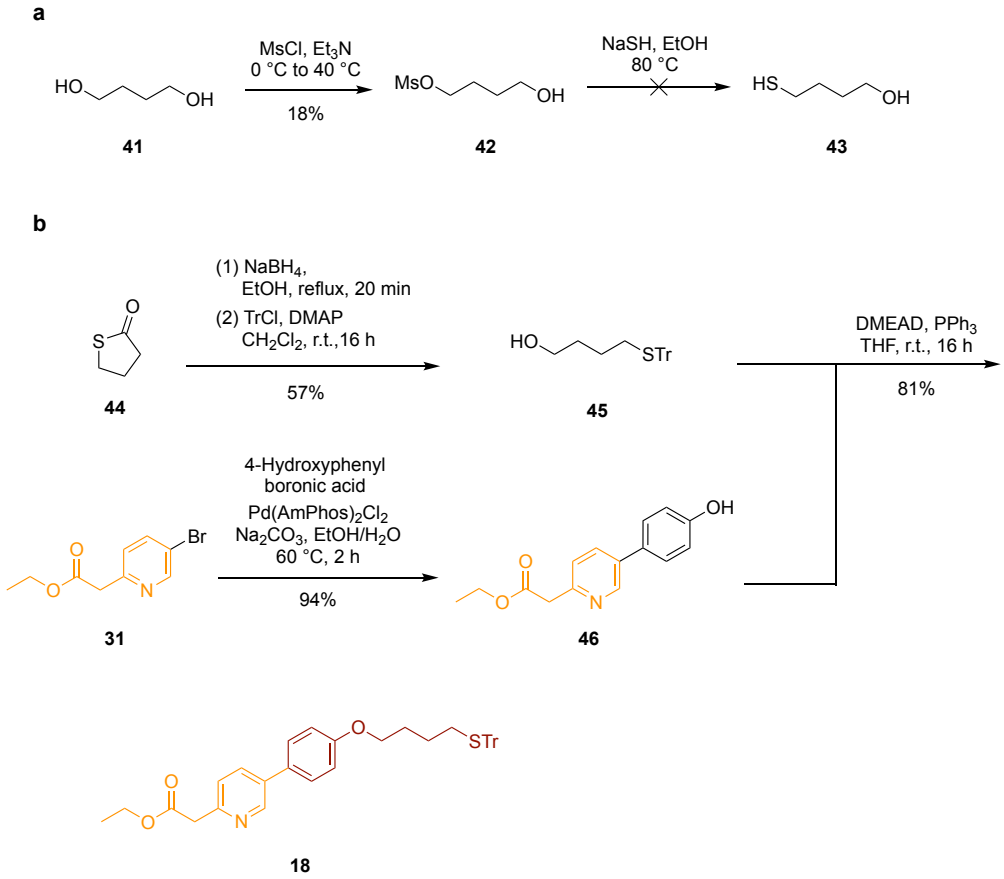


Figure 10. Synthesis of **18**. **a**, attempt by mesylation and the S_N2 reaction, **b**, attempt by reduction of γ-thiobutyrolactone.

The left half of the designed molecular machine was synthesised as a macrocycle **53** (**Figure 11**) with a cargo molecule (red part) connected to the initial aldehyde station (green part) and the thiopyridone arm (pink part). With this design, the cargo has already been picked up by the first arm when the cargo transportation process starts. The left macrocycle **53** was synthesised in seven steps from the components described so far (**Figure 11a**). The synthesis of the right half of the machine is currently ongoing. To date, the right station **19** and the bottom part of hydrazone switch **17** have been connected by a copper(I)-catalysed alkyne-azide cycloaddition (CuAAC) to yield compound **54** (**Figure 11b**).

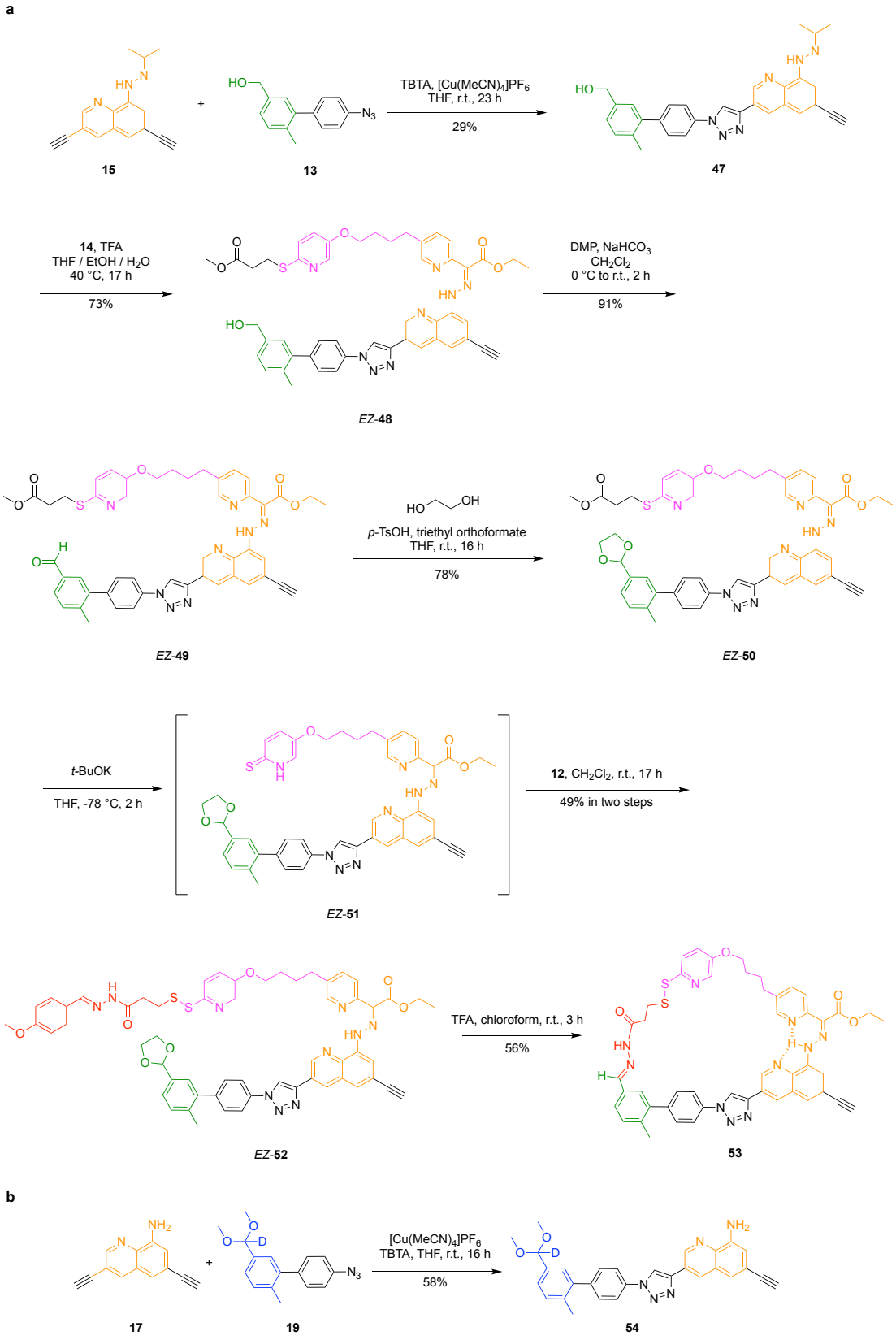


Figure 11. Synthesis of (**a**) the left macrocycle and (**b**) the initial synthesis of the right half of the machine.

2.5. Model study with the first design of the left macrocycle

After synthesising left macrocycle **53**, a model study of the ring opening, switching and cargo transfer process was attempted (**Figure 12**). When compound **53** (0.15 mM) was mixed with TFA (2.6 mM) and *p*-nitrobenzaldehyde (1.8 mM), hydrazone exchange ring-opened the macrocycle and isomerisation of the hydrazone switch occurred to generate **55**. This was confirmed by ^{1}H NMR spectroscopy and mass spectrometry (**Figure 13b** and **Figure 14a**). In the ^{1}H NMR spectrum, a characteristic signal of the protonated hydrazone switch in the *Z*-configuration (14.23 ppm, signal **a**) and the free aldehyde station (10.04 ppm, signal **b**) could be observed. In the mass spectrum, the main signal corresponded to compound **55** (*m/z* = 1038.33). Subsequently, 1-hexanethiol (0.9 mM) was added to model the right arm and study the cargo transfer process. ^{1}H NMR spectra revealed some inconclusive changes (**Figure 13c**). The most significant changes were the emergence of new peaks at 9.5–8.0 ppm, indicating that some changes might have occurred at the hydrazone switch and/or the thiopyridone moiety. The expected compound after the transfer of the cargo molecule (**56**, *m/z* = 787.28) was not observed by mass spectrometry. Instead, the most significant change observed was the emergence of a peak at *m/z* = 903.42 (**Figure 14b**), which corresponds to the mass of disulfide **58**. The emergence of disulfide **58** instead of **56** indicates that the 2-thiopyridone moiety (pink) is a worse leaving group than the cargo molecule (red) for the disulfide exchange, and the intended cargo transfer from the thiopyridone arm to the alkyl thiol arm (brown) is less favoured.

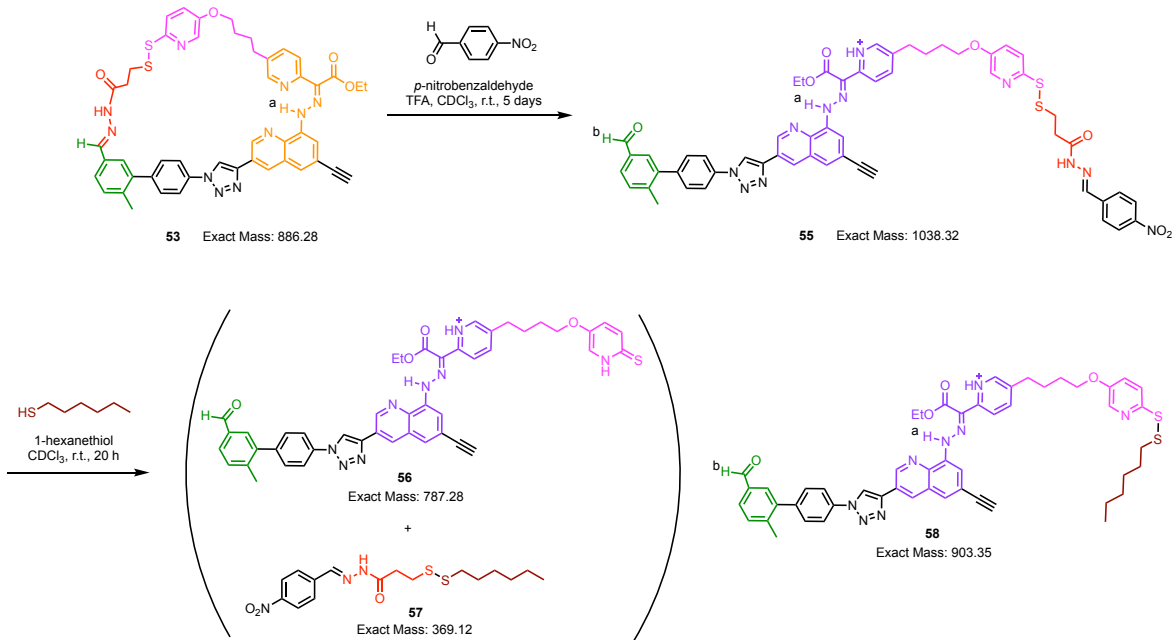


Figure 12. Ring opening and cargo transfer study with the first design of the left macrocycle.

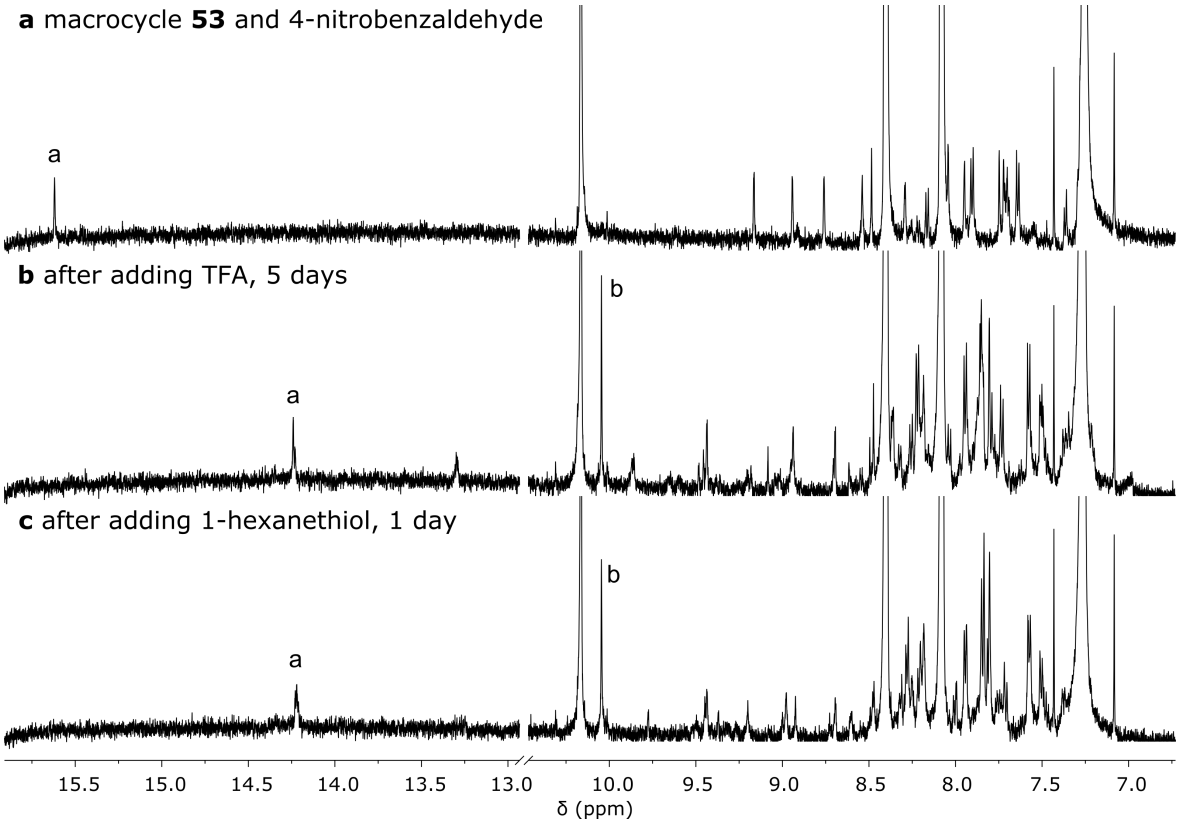


Figure 13. Partial ^{1}H NMR spectra during the model study (600 MHz, $CDCl_3$, 295 K). Lettering relates protons in the chemical structures in **Figure 12** to the corresponding signals in the ^{1}H NMR spectra. **a**, macrocycle **53** (0.15 mM) and 4-nitrobenzaldehyde (1.8 mM). **b**, after adding TFA (2.6 mM), 5 days. **c**, after adding 1-hexanethiol (0.9 mM), 1 day.

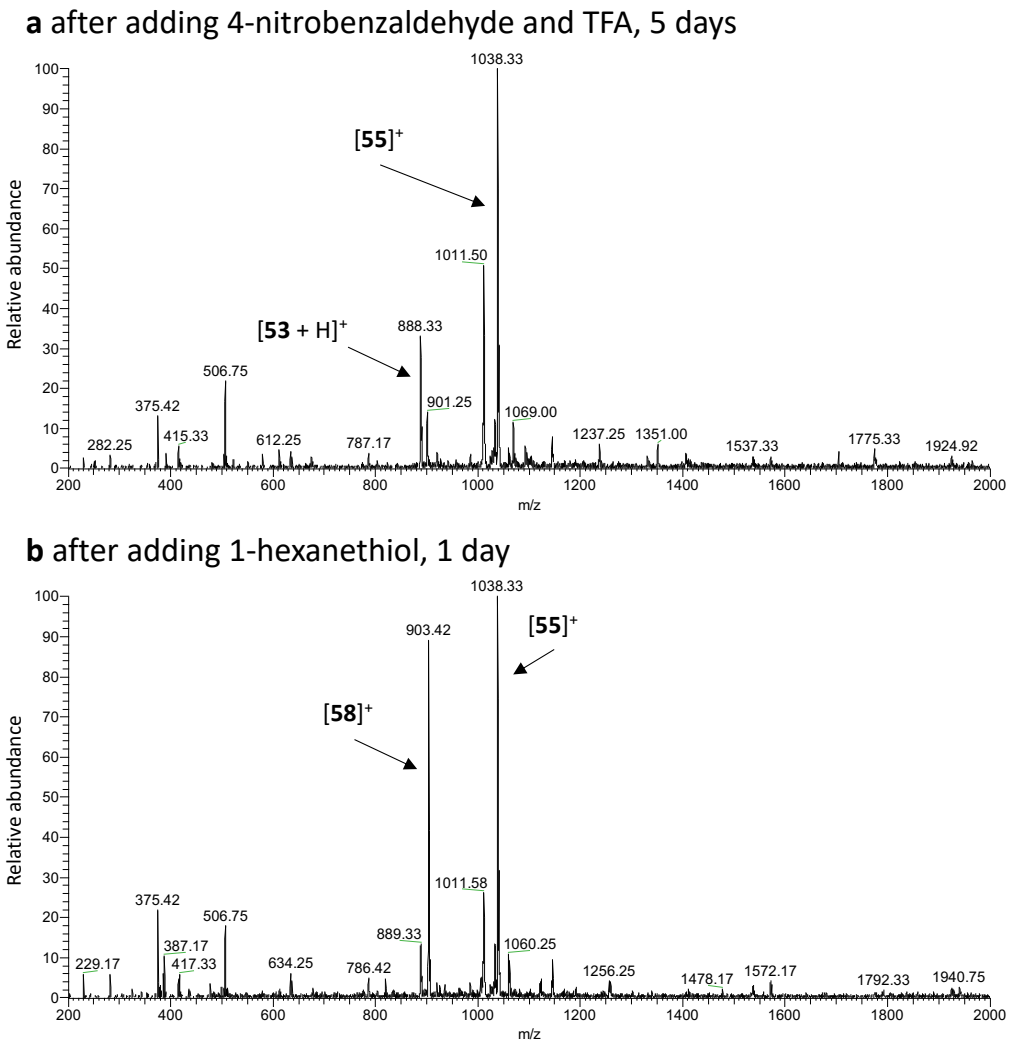


Figure 14. ESI-MS spectra during the model study. **a**, after adding 4-nitrobenzaldehyde and TFA, 5 days. **b**, after adding 1-hexanethiol, 1 day.

2.6. Model study of cargo transfer to change the design of the left arm

Based on the observation of the incorrect disulfide exchange in the model study (Section 2.5), the design of the thiopyridone arm was changed. **Figure 15** shows old design **53** and the new design of the left macrocycle. In the new design, the oxygen atom connected to the mercaptopyridine (thiopyridone) moiety is replaced with a methylene group, which has less electron-donating character than the oxygen atom, so that the thiopyridone moiety becomes a better leaving group in the disulfide exchange for the cargo transfer.

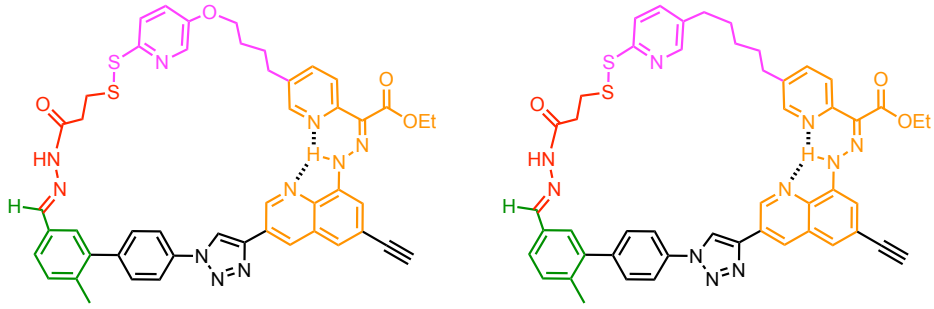


Figure 15. Old design **53** (left) and new design (right) of the left macrocycle.

A model study was initially performed to confirm that the cargo transfer from the new design of the thiopyridone arm to the alkyl thiol arm occurs preferentially (**Figure 17**). A new cargo molecule (**62** in **Figure 16a**) and a model compound of the cargo connected to the new thiopyridone arm (**65** in **Figure 16b**) were synthesised. The difference between the old cargo molecule (**12** in **Figure 7**) and the new one **62** is the presence of the nitro group. This strongly electron withdrawing nitro group serves to make the mercaptopyridine moiety in cargo **62** a better leaving group for disulfide exchange, thus facilitating the loading of the cargo to the left arm (i.e. cargo loading to model left arm **64** to yield compound **65**).

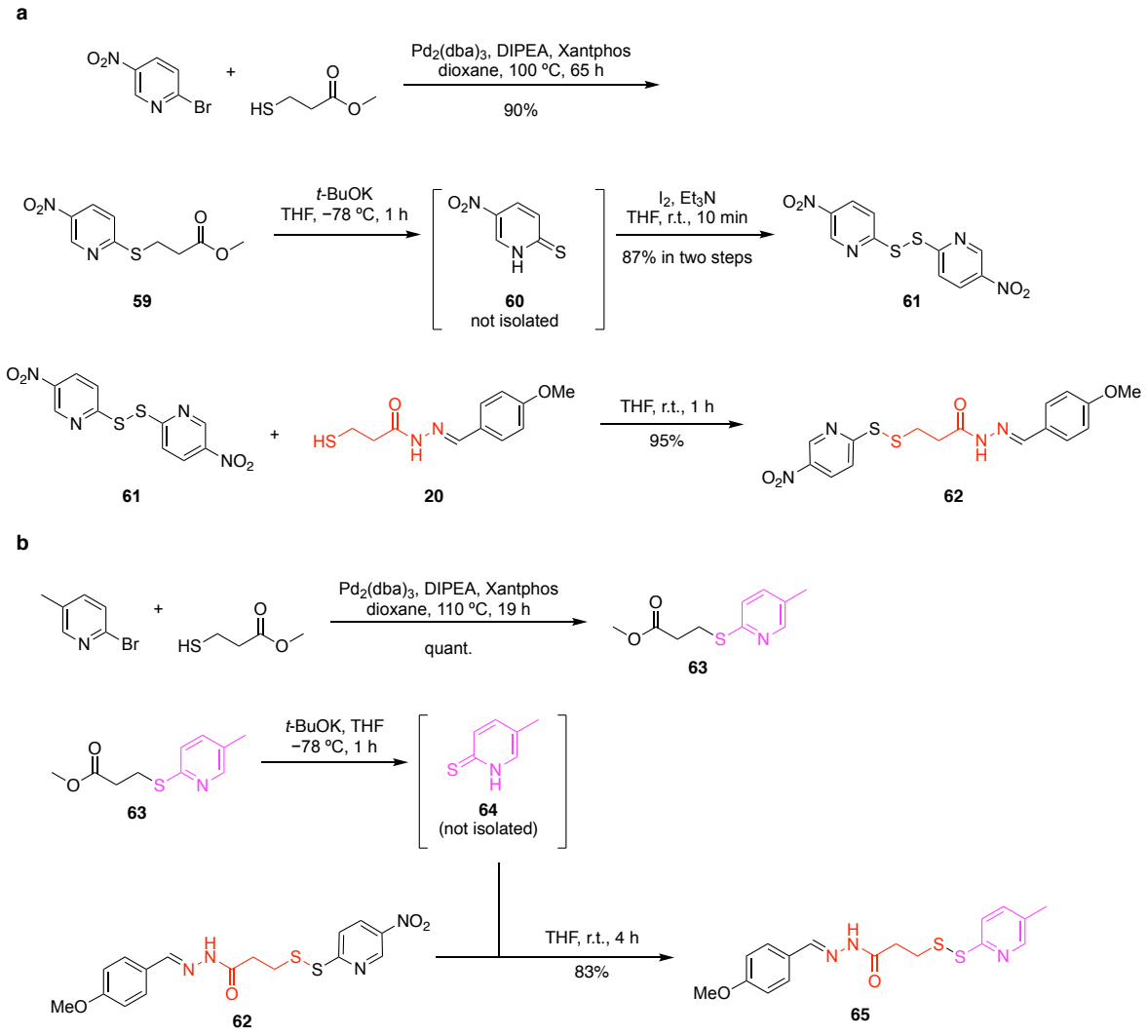


Figure 16. Synthesis of (**a**) the new cargo molecule and (**b**) the model compound of the cargo connected to the new thiopyridone arm.

Cargo transfer model study was conducted by adding 1-hexanethiol (model compound of the right arm, 0.9 mM) to the solution of compound **65** (0.9 mM) in $CDCl_3$ (**Figure 17**). Samples with and without TFA (2.7 mM) were prepared to examine the effect of TFA. ^{1}H NMR spectra after 24 h were inconclusive due to overlapping peaks. However, by mass spectrometry, the formation of compound **66** could be observed in both samples (**Figure 18a** and **b**) and compound **64** in the sample with TFA (**Figure 18c**), while neither of compound **20** or **67** was detected. Therefore, it is likely that the disulfide exchange occurred with desired selectivity to facilitate cargo transfer from the mercaptopyridine arm (pink) to the alkyl thiol arm (brown).

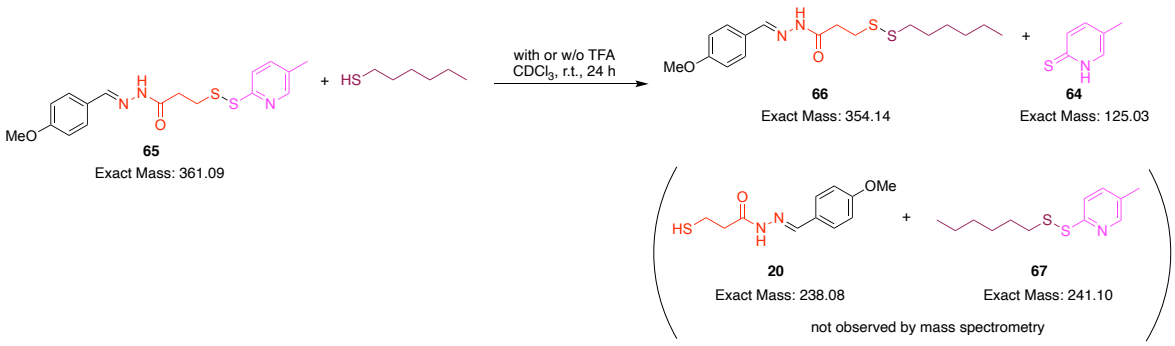


Figure 17. Model study of cargo transfer to change the design of the left arm using compound **65** (0.9 mM) and 1-hexanethiol (0.9 mM) with or without TFA (2.7 mM).

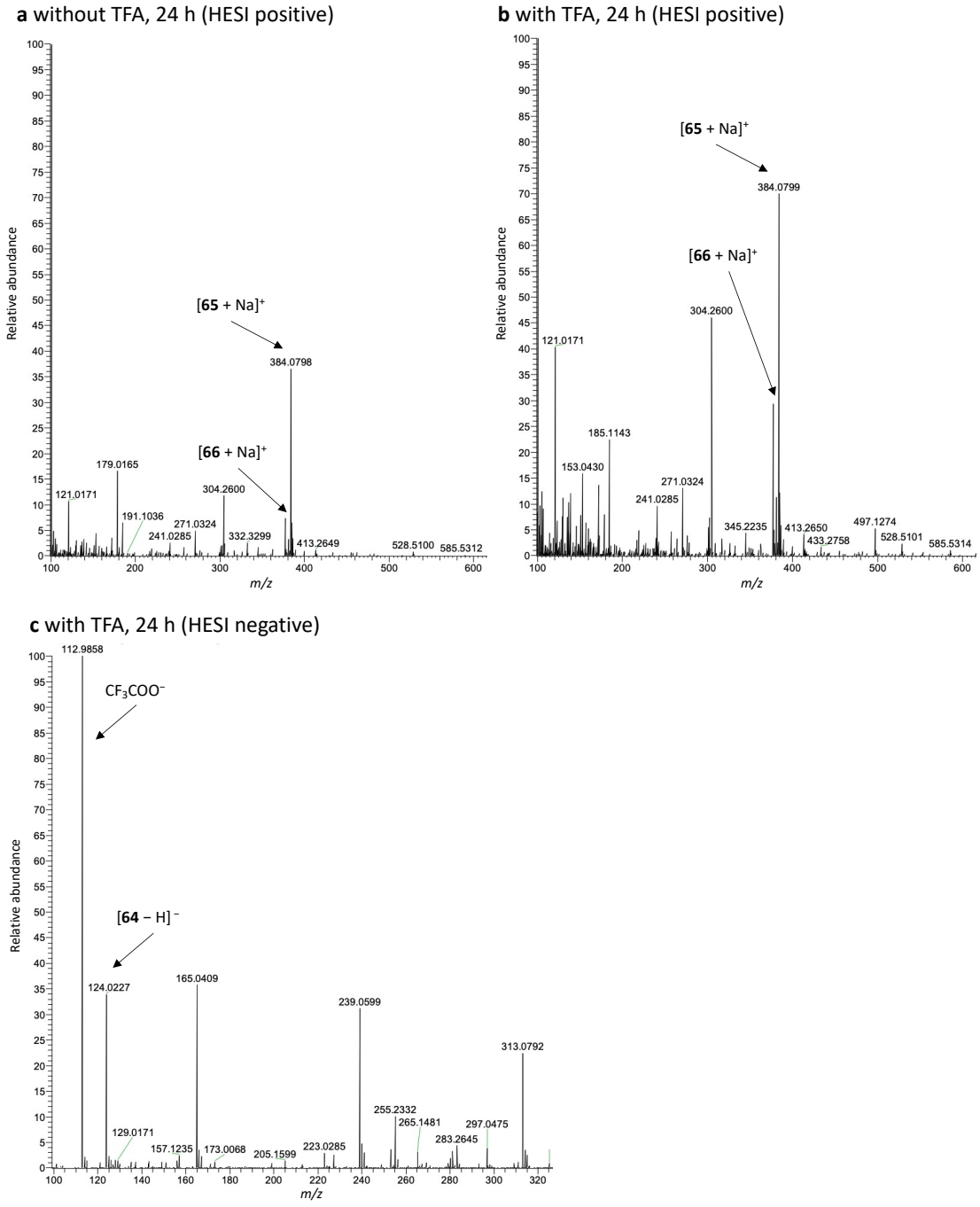


Figure 18. Mass spectra of (**a**) the sample without TFA, HESI positive, (**b**) the sample with TFA, HESI positive and (**c**) the sample with TFA, HESI negative.

2.7. Synthesis for the second design of the left macrocycle

Encouraged by the result of the cargo transfer model study, synthesis of the new design of the left macrocycle (**Figure 19**) was commenced. 2-Chloropyridine derivative **68** was selected as a precursor for the mercaptopyridine arm, because 2-bromopyridine derivative underwent side reactions during the hydrogenation of the alkyne. After converting alcohol **69** to alkyne **71** in two steps, the left arm was connected to the upper half of the hydrazone switch **31** by the Sonogashira coupling. After reduction of alkyne intermediate **72**, compound **73** was obtained. While synthesis has not yet progressed further, the plan for the remainder of the synthesis is outlined here briefly. The Riley oxidation to form 1-ketone-2-ester species will be followed by introduction of a protected mercaptopyridine arm by a nucleophilic aromatic substitution between a thiol and the 2-chloropyridine moiety following a literature procedure[26].

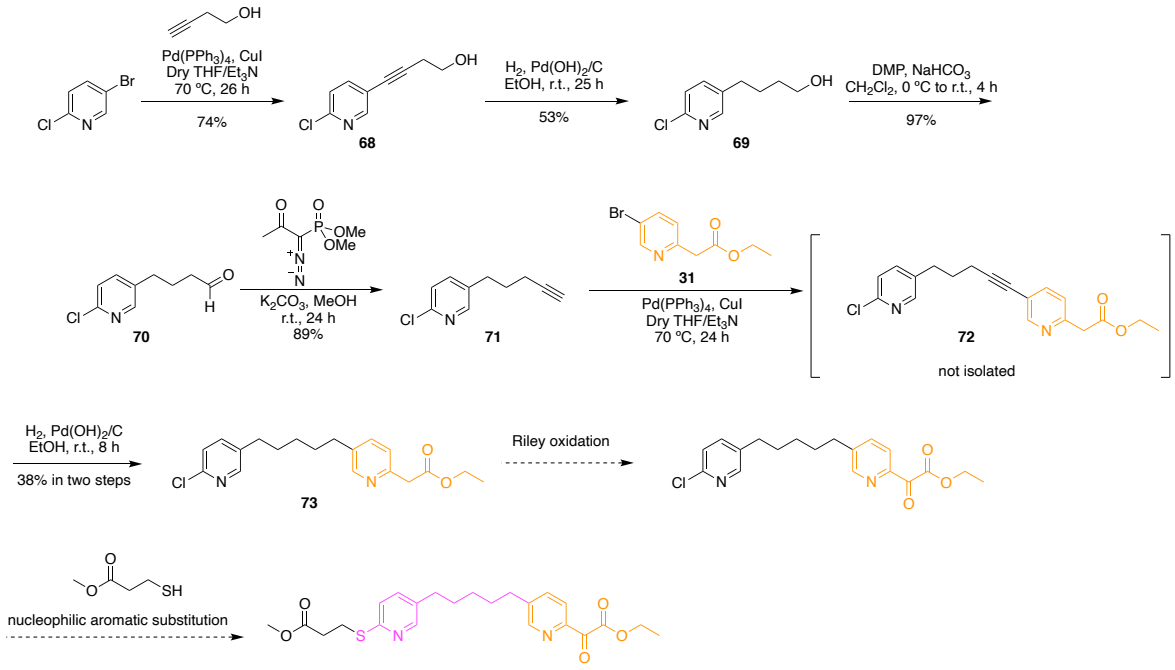


Figure 19. Ongoing synthesis for the second design of the left macrocycle.

2.8. Model studies for CuAAC in the presence of a hydrazone switch

While synthesising the new design of the left macrocycle, model studies for the CuAAC in the presence of a hydrazone switch were also undertaken. This is because the intended final step of the synthesis of the whole machine is a CuAAC (**Figure 20**). This final reaction might be difficult due to the presence of hydrazone switches: these moieties strongly bind to the copper catalysts for CuAAC, which is likely to inhibit the reaction. Therefore, this model study is crucial to determine optimal conditions to perform the final CuAAC in the presence of a hydrazone switch.

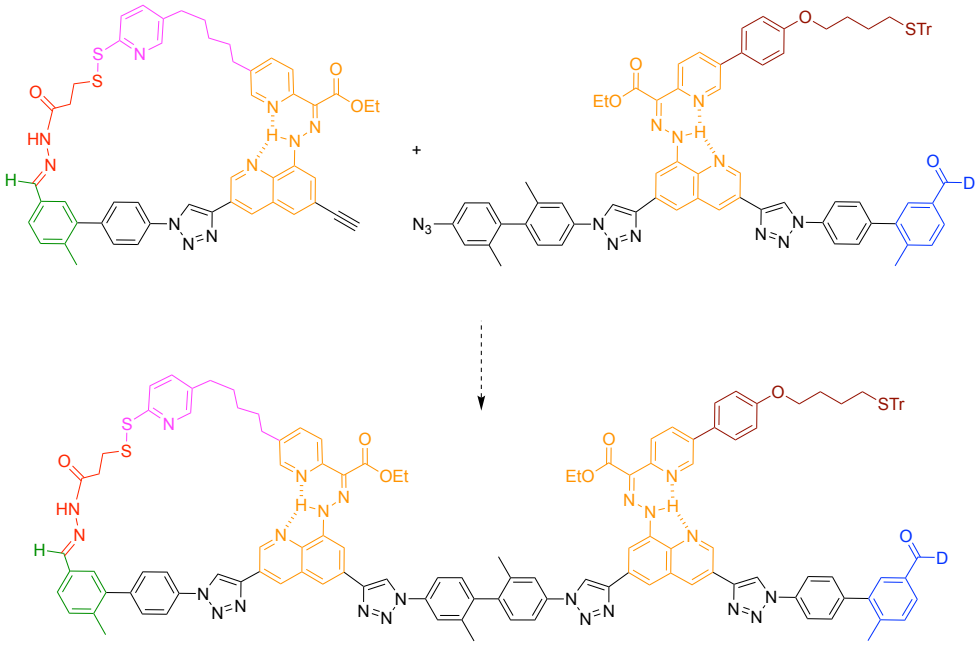


Figure 20. Intended final CuAAC for synthesising the molecular machine.

When an excess of reagents were used (5.4 eq. CuI, 5 eq. Na-ascorbate, 6.5 eq. *N*,*N*′-DMEDA, **Figure 21**), the CuAAC between model alkyne **74** and model azide **75** proceeded and the copper catalysts bound to the hydrazone switch could be removed by washing the crude with aqueous Na_4EDTA solution (1 M) and aqueous NH_3 solution (35%, *w*/*w*). In the ^1H NMR spectrum, the product had all characteristic signals of compound **76**

(labels **a** to **e** in **Figure 22**). The mass spectrum also showed the signals corresponding to **76** (**Figure 23**).

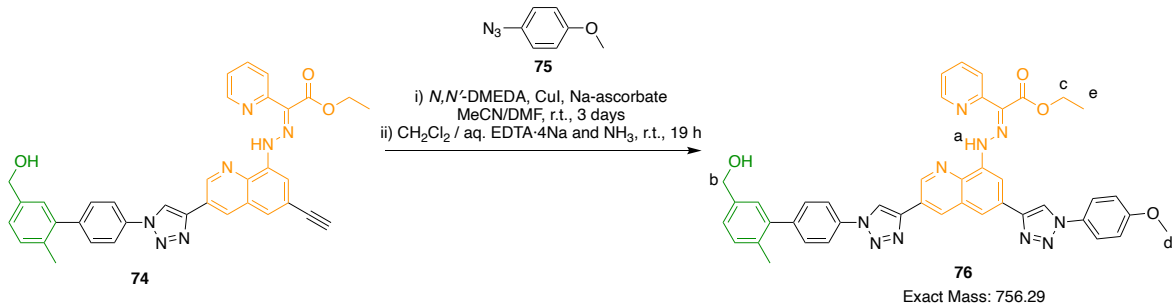


Figure 21. CuAAC in the presence of a hydrazone switch.

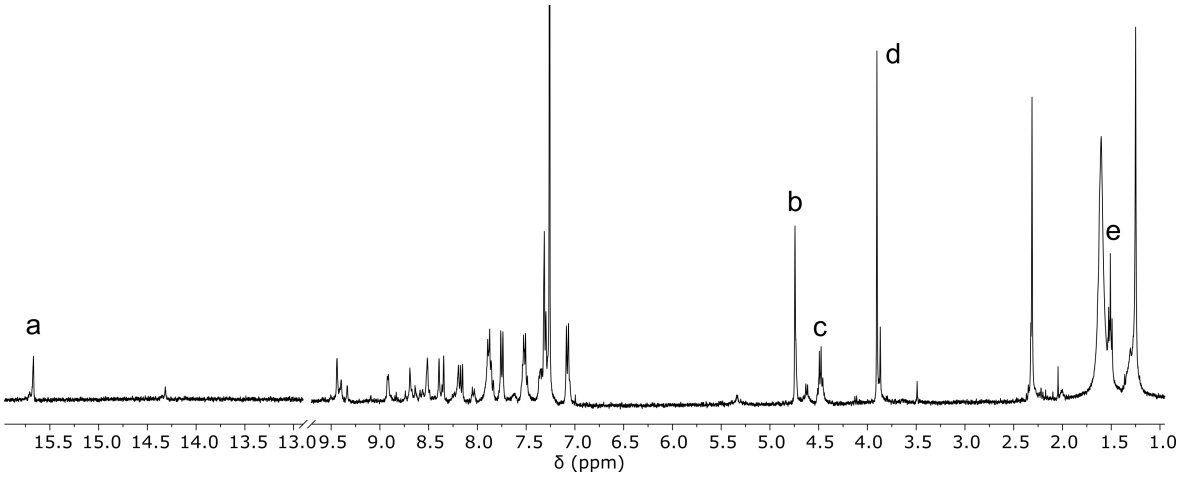


Figure 22. Partial ^{1}H NMR spectrum of the product after CuAAC and washing with aqueous Na_4EDTA solution (1 M) and aqueous NH_3 solution (35%, *w*/*w*) (600 MHz, $CDCl_3$, 295 K).

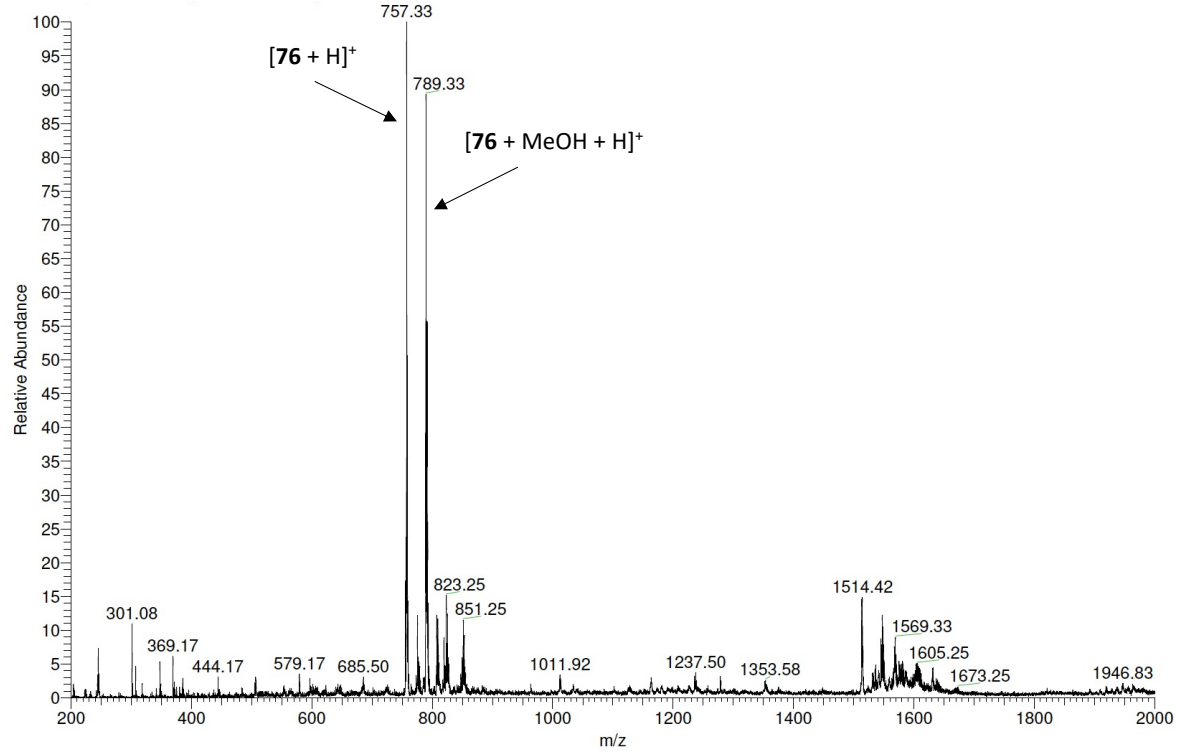
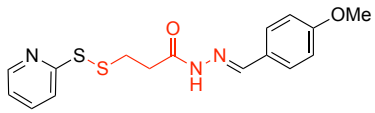


Figure 23. Mass spectrum of the product after CuAAC and washing with aqueous Na_4EDTA solution (1 M) and aqueous NH_3 solution (35%, *w*/*w*).

Encouraged by the successful initial CuAAC, the stability of the hydrazone and disulfide bonds of the cargo were examined under the reaction and washing conditions for the CuAAC. This is because the model hydrazone switch **74** lacks the reactive hydrazone and disulfide bonds of the cargo that are present in the actual left macrocycle (**Figure 20**), and thus the model study in **Figure 21** cannot tell if they are compatible with the CuAAC conditions. For this purpose, cargo compound **12**, which contains both the hydrazone and disulfide bond, was subjected to the reaction and washing conditions of CuAAC and its stability was evaluated (**Figure 24**). Initially, compound **12** was found to remain almost intact under the washing conditions employed in the model study in **Figure 21** (entry 1 of **Figure 24**). Subsequently, various reaction conditions of CuAAC were attempted on compound **12**, and the crude reaction mixture was submitted to the washing conditions (entry 2 to 5). Pleasingly, compound **12** was found to remain almost intact when a substoichiometric amount of the copper catalyst was employed without other additives (entry 5). Following these encouraging results, a CuAAC model study between compound **74** and **75** will again be attempted with these successful conditions to see if they facilitate the reaction.



12

Entry	Reaction conditions	Washing conditions	Results
1	-	CH_2Cl_2/aq. EDTA·4Na and aq. NH_3, r.t., 24 h	almost intact **12**
2	*N,N'*-DMEDA, CuI, Na-ascorbate dry DMF, r.t., 28 h	CH_2Cl_2/aq. EDTA·4Na and aq. NH_3, r.t., 24 h	decomposed
3	DIPEA, $[Cu(MeCN)_4]PF_6$ dry THF, 50 °C, 28 h	CH_2Cl_2/aq. EDTA·4Na and aq. NH_3, r.t., 24 h	decomposed
4	$[Cu(MeCN)_4]PF_6$ (5 eq) dry CH_2Cl_2/DMF, r.t., 23 h	CH_2Cl_2/aq. EDTA·4Na and aq. NH_3, r.t., 23 h	decomposed
5	$[Cu(MeCN)_4]PF_6$ (0.5 eq) dry CH_2Cl_2/DMF, r.t., 23 h	CH_2Cl_2/aq. EDTA·4Na and aq. NH_3, r.t., 23 h	almost intact **12**

Figure 24. Model study to test the stability of a hydrazone and disulfide bond in the conditions for CuAAC.

2.9. Conclusion and future work

To date, model studies of the cargo transfer between two arms have been conducted and the multi-step synthesis of this complex molecular machine is at an advanced stage. Three model studies of the cargo transfer showed that the property of the thiopyridone arm is strongly affected by the substituents: in the first (Section 2.3.) and the third (Section 2.6.) model studies, where unsubsituted and methyl-substituted thiopyridone arm was used, respectively, the cargo transfer to the alkyl thiol arm was observed. However, in the second model study (Section 2.5.), where an oxygen atom was connected to the thiopyridone arm, the cargo transfer was not observed. These results imply that parts of molecular machines that undergo formation and cleavage of bonds during operation need to be designed with great care to the steric and electronic environment around them to achieve intended outcomes. As for the synthesis, difficulty mainly arose from incompatibility of various functional groups and reaction conditions, and efforts have been made to circumvent this difficulty. For example, the hydrazone exchange to make the left hydrazone switch (**Figure 11a**) was initially attempted between the top part **14** and the bottom half with an acetal-protected aldehyde station, instead of the bottom half with a benzyl alcohol **47** as in the current procedure. However, this reaction was unsuccessful as the aldehyde station was deprotected under the acidic conditions and this station could also be involved in the hydrazone condensation. This problem was circumvented by using the bottom half with a benzyl alcohol **47**, and converting it to an aldehyde after the hydrazone exchange. The next step of the research will be to complete the synthesis of the full molecular machine. After the synthesis, the operation of the molecular machine will be attempted to transfer the cargo

molecule from one side of the machine to the other by the coordination of two hydrazone switches.

The success of this research will pave the way to molecular machines with multiple movable parts that are coordinated to achieve complex and sophisticated tasks. Such molecular machines will enable controlled transfer of the cargo in microscopic systems including biological environment. Another application will be an actuator made of the assembly of molecular machines. In such an application, the coordination of the dynamics of multiple movable parts will be essential. The hydrazone-type switch used in this research has also been applied to the stereodivergent synthesis of four possible stereoisomers starting from common reactants[25]. By applying molecular machines with multiple movable parts to the synthesis as in this study, it may be possible to realise a molecular 'assembly line', which synthesise complicated molecules step-by-step, just as automobiles are assembled on an assembly line in the macroscopic world.

The difficulty of synthesis of the transporter with two hydrazone switches suggests that it is not practical to incorporate too many types of functional groups in one molecule. As the variation of functional groups increases, finding reaction conditions that are compatible with all of them becomes increasingly difficult. One solution to this problem will be to use repeated modules to construct complicated structures and minimise the difference between modules. This is how biology synthesises a wide variety of proteins from a limited choice of building blocks, amino acids, and this strategy will be useful for designing future molecular machines as well.

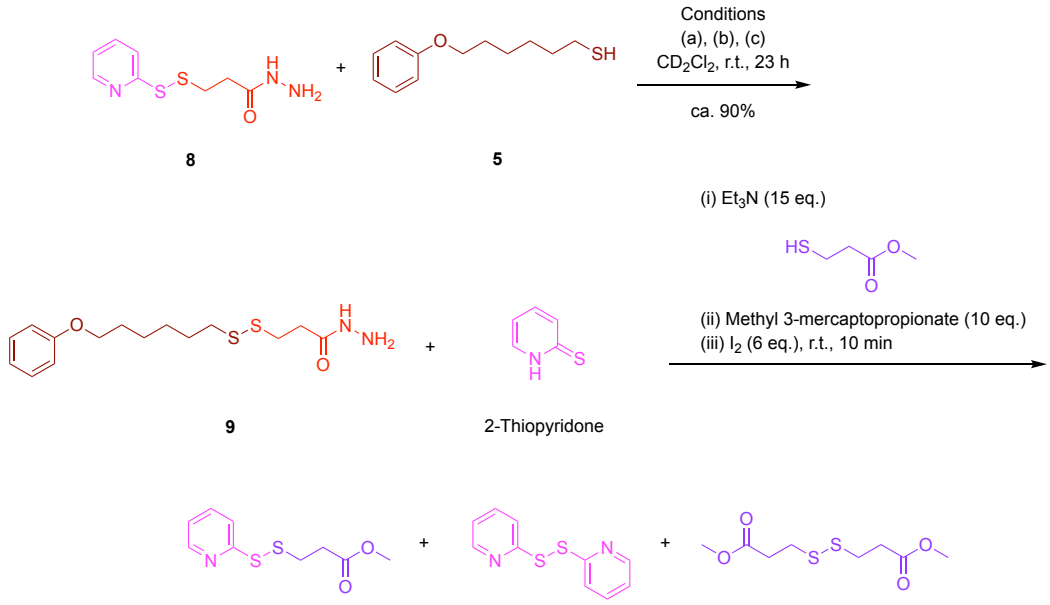
2.10. Experimental section

2.10.1. General information

Unless otherwise stated, all chemicals were purchased from commercial sources and were used without further purification. Dry CH_2Cl_2, DMF, Et_2O, THF and toluene were obtained by passing the solvent through an activated alumina column on a Phoenix solvent drying system (SDS) (JC Meyer Solvent Systems, CA, USA). Dry dioxane was purchased from Sigma-Aldrich. Solvents were degassed by argon bubbling and sonication. Flash column chromatography was performed using 60 Å silica gel from Sigma-Aldrich or Merck (40–63 μm particle size) as the stationary phase. Reactions were monitored by TLC using Macherey Nagel 0.2 mm pre coated TLC sheets with fluorescent indicator UV_{254}. The spots were observed under UV lamps or with prepared stains (cerium Molybdate and Ninhydrin). ^{1}H NMR and ^{13}C NMR were recorded on a Bruker Avance III instrument with an Oxford AS600 magnet equipped with a cryoprobe (5mm CPDCH ^{13}C-^{1}H/D). Chemical shifts are reported in parts per million (ppm) and referenced to residual solvent peaks. Coupling constants are reported in Hertz (Hz). ^{1}H NMR multiplicities are abbreviated as follows: s = singlet, d = doublet, t = triplet, q = quartet, p = pentet, m = multiplet, br = broad. ^{1}H and ^{13}C assignments were made using 2D-NMR methods (COSY, HSQC, HMBC). High-resolution mass spectrometry was performed by the Mass Spectrometry Service, Department of Chemistry, University of Manchester.

2.10.2. Model studies

Model study of the cargo transfer and cargo reloading

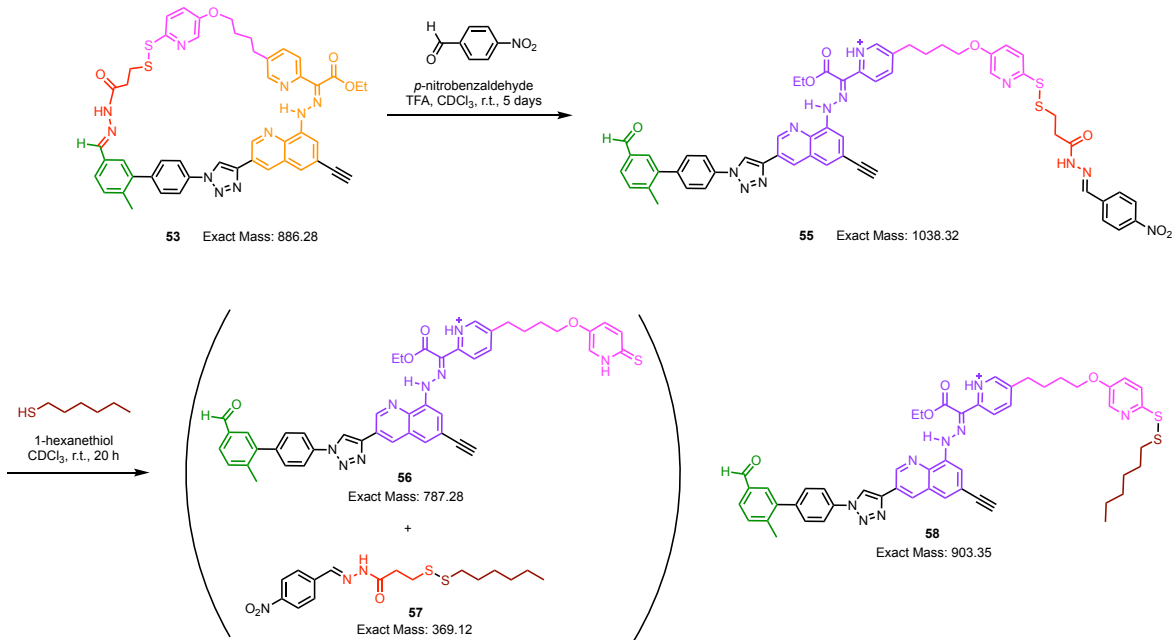


Conditions	**10**	Aldrithiol^{TM}-2	**11**
(a) AcOH (0.2 eq.)	1	0.25	7.5
(b) TFA (0.2 eq.)	1	0.25	7
(c) Et_3N (0.2 eq.)	1	0.25	8.5

Compound **9** was omitted in the table for clarity.

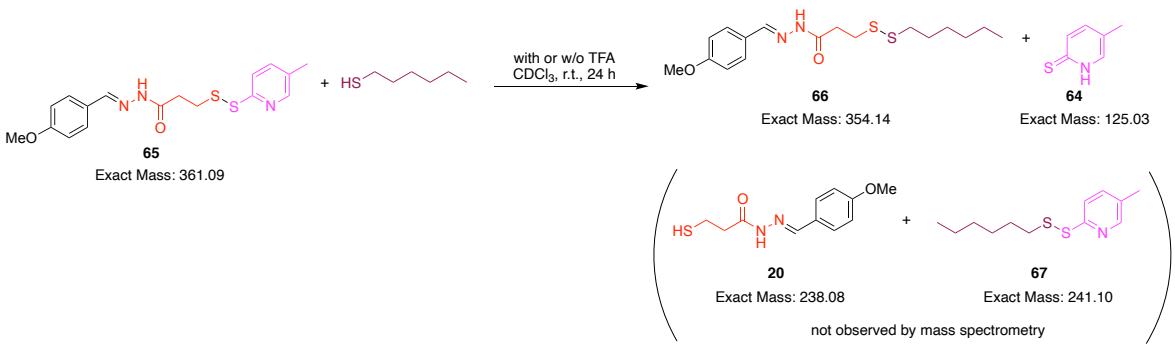
Three identical solutions of **8** (3.44 mg, 15 µmol, 1 eq.) and **5** (4.16 mg, 15 µmol, 1 eq.) in CD_2Cl_2 (0.4 mL) were prepared. To each of them, the solution of AcOH, TFA or Et_3N in CD_2Cl_2 (30 mM, 0.1 mL, 3 µmol, 0.2 eq.) was added. The reaction was conducted at r.t. for 23 h in NMR tubes. A solution of excess amount of Et_3N (2250 mM, 0.1 mL, 225 µmol, 15 eq.), a solution of methyl 3-mercaptopropionate (200 mM, 0.75 mL, 150 µmol, 10 eq.) and a solution of iodine (180 mM, 0.5 mL, 90 µmol, 6 eq.) in CD_2Cl_2 were added to each tube. The reaction was continued for 10 min and ^1H NMR spectra were recorded (**Figure 5**). The numbers in the table indicate the relative amount of substance.

Model study of ring opening, switching and cargo transfer with the left macrocycle



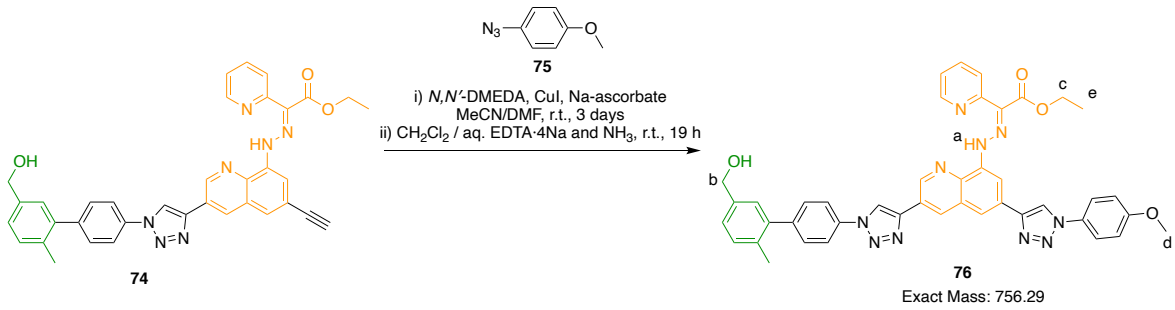
Compound **53** (0.07 mg, 0.077 µmol, 1 eq.) was dissolved in $CDCl_3$ (0.5 mL) and a solution of *p*-nitrobenzaldehyde in $CDCl_3$ (20 µL of solution with 0.14 mg (0.9 µmol, 12 eq.) of *p*-nitrobenzaldehyde) and a solution of TFA in $CDCl_3$ (5 µL of solution with 0.1 µL (1.32 µmol, 17 eq.) of TFA) were added. The reaction mixture was kept in an NMR tube at r.t. for 5 days and the progress was monitored by ^{1}H NMR spectroscopy (**Figure 13**) and mass spectrometry (**Figure 14**). Then, a solution of 1-hexanethiol in $CDCl_3$ (10 µL of solution with 0.053 mg (0.45 µmol, 5.8 eq.) of 1-hexanethiol) was added and the reaction was continued for 20 h more.

Model study of cargo transfer with a new design of the left arm



Compound **65** (0.16 mg, 0.45 µmol, 1 eq.) was dissolved in $CDCl_3$ (0.5 mL) and a solution of 1-hexanethiol in $CDCl_3$ (10 µL of solution with 0.053 mg (0.45 µmol, 1 eq.) of 1-hexanethiol) was added. Two sets of this reaction mixture were prepared in NMR tubes and to one of them was added a solution of TFA in $CDCl_3$ (10 µL of solution with 0.1 µL (1.35 µmol, 3 eq.) of TFA). The reaction mixtures were kept at r.t. for 24 h and submitted to mass spectrometry (**Figure 18**).

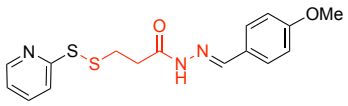
Model study of CuAAC in the presence of a hydrazone switch



Compound **74** (5.1 mg, 8 µmol, 1 eq.), **75** (1.9 mg, 13 µmol, 1.5 eq.), CuI (8.6 mg, 45 µmol, 5.4 eq.), Na-ascorbate (7.9 mg, 40 µmol, 5 eq.) and *N*,*N*′-DMEDA (5.6 µL, 52 µmol, 6.5 eq.) were combined in degassed MeCN:DMF (4:1, *v*/*v*, 0.5 mL) and the reaction mixture was stirred under argon atmosphere at r.t. for 3 days. The mixture was diluted with EtOAc (10 mL), washed with aqueous Na_4EDTA solution (1 M, 3 × 10 mL) and the organic layer was concentrated *in vacuo*. The crude was suspended in CH_2Cl_2 (10 mL) and aqueous Na_4EDTA solution (1 M, 5 mL) and aqueous NH_3 solution (35%, *w*/*w*, 5 mL) were added. The mixture was stirred vigorously at r.t. for 19 h. The reaction mixture was extracted with CH_2Cl_2 (2 × 10 mL) and the organic layer was dried with $MgSO_4$ and concentrated *in vacuo*. The crude product was purified by PTLC (SiO_2, CH_2Cl_2/EtOAc, 1:4, *v*/*v*) to give **76** (ca. 2 mg, 2.6 µmol) as a yellow solid. Full characterisation of the product is yet to be done due to

small amount of the compound. However, formation of compound **76** was confirmed by ^{1}H NMR spectroscopy (**Figure 22**) and mass spectrometry (**Figure 23**).

Model study to test the stability of a hydrazone and disulfide bond in the conditions for CuAAC



12

Entry	Reaction conditions	Washing conditions	Results
1	-	CH_2Cl_2/aq. EDTA·4Na and aq. NH_3, r.t., 24 h	almost intact **12**
2	*N,N'*-DMEDA, CuI, Na-ascorbate dry DMF, r.t., 28 h	CH_2Cl_2/aq. EDTA·4Na and aq. NH_3, r.t., 24 h	decomposed
3	DIPEA, $[Cu(MeCN)_4]PF_6$ dry THF, 50 °C, 28 h	CH_2Cl_2/aq. EDTA·4Na and aq. NH_3, r.t., 24 h	decomposed
4	$[Cu(MeCN)_4]PF_6$ (5 eq) dry CH_2Cl_2/DMF, r.t., 23 h	CH_2Cl_2/aq. EDTA·4Na and aq. NH_3, r.t., 23 h	decomposed
5	$[Cu(MeCN)_4]PF_6$ (0.5 eq) dry CH_2Cl_2/DMF, r.t., 23 h	CH_2Cl_2/aq. EDTA·4Na and aq. NH_3, r.t., 23 h	almost intact **12**

Compound **12** (ca. 2.8 mg, 8 µmol, 1 eq.) was subjected to the following reaction conditions and workup procedures:

Entry 2—CuI (7.7 mg, 40 µmol, 5 eq.), Na-ascorbate (7.8 mg, 39 µmol, 5 eq.), *N,N'*-DMEDA (5.6 µL, 52 µmol, 6.5 eq.) in dry degassed DMF (0.4 mL). The reaction mixture was stirred under nitrogen atmosphere at r.t. for 28 h. The reaction mixture was diluted with EtOAc (10 mL), washed with aqueous Na_4EDTA solution (1 M, 3 × 10 mL) and the organic layer was concentrated *in vacuo*.

Entry 3—DIPEA (11 µL, 64 µmol, 8 eq.), $[Cu(MeCN)_4]PF_6$ (14.5 mg, 40 µmol, 5 eq.) in dry degassed THF (0.4 mL). The reaction mixture was stirred under nitrogen atmosphere at 50 °C for 28 h. The reaction mixture was diluted with EtOAc (10 mL), washed with aqueous Na_4EDTA solution (1 M, 2 × 10 mL) and brine (10 mL) and the organic layer was concentrated *in vacuo*.

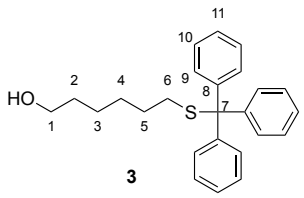
Entry 4—$[Cu(MeCN)_4]PF_6$ (15 mg, 40 μmol, 5 eq.) in dry degassed CH_2Cl_2:DMF (5:1, *v*/*v*, 3 mL). The reaction mixture was stirred under argon atmosphere at r.t. for 23 h. The reaction mixture was diluted with CH_2Cl_2 (5 mL), washed with aqueous Na_4EDTA solution (1 M, 3 × 10 mL). The organic layer was dried over $MgSO_4$ and concentrated *in vacuo*.

Entry 5—$[Cu(MeCN)_4]PF_6$ (1.5 mg, 4 μmol, 0.5 eq.) in dry degassed CH_2Cl_2:DMF (5:1, *v*/*v*, 3 mL). The reaction mixture was stirred under argon atmosphere at r.t. for 23 h. The reaction mixture was diluted with CH_2Cl_2 (5 mL), washed with aqueous Na_4EDTA solution (1 M, 3 × 10 mL). The organic layer was dried over $MgSO_4$ and concentrated *in vacuo*.

The crude (entry 2 to 5) or compound **12** (entry 1) was combined with CH_2Cl_2 (10 mL), aqueous Na_4EDTA solution (1 M, 5 mL) and aqueous NH_3 solution (35%, *w*/*w*, 5 mL) and the mixture was stirred vigorously at r.t. overnight. The reaction mixture was extracted with CH_2Cl_2 (2 × 10 mL) and the organic layer was dried with $MgSO_4$ and concentrated *in vacuo*. The crude after washing was analysed with ^1H NMR spectroscopy and TLC.

2.10.3. Synthesis

Synthesis of 3



3

6-Mercapto-1-hexanol (**2**, 500 μL, 3.65 mmol) was dissolved in dry CH_2Cl_2 (10 mL). To this solution, TrCl (1020 mg, 3.65 mmol, 1 eq.) and DMAP (45 mg, 0.37 mmol, 0.1 eq.) were added and the reaction mixture was stirred at r.t. for 1.5 h. The reaction mixture was partitioned between water (20 mL) and chloroform (20 mL). The water layer was extracted with chloroform (2 × 20 mL). The combined organic layer was washed with water (20 mL)

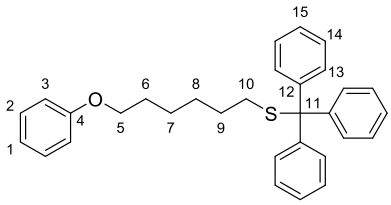
and brine (20 mL). The organic layer was dried with $MgSO_4$ and concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO_2, CH_2Cl_2/MeOH, 0% → 2%) to give **3** (900 mg, 2.39 mmol, 65%) as colorless oil.

^{1}H NMR (600 MHz, $(CD_3)_2SO$) δ 7.35 – 7.30 (m, 12H, HC_9 + HC_{10}), 7.24 (m, 3H, HC_{11}), 4.30 (t, J = 5.1 Hz, 1H, OH), 3.30 (td, J = 6.6, 5.1 Hz, 2H, HC_1), 2.07 (t, J = 7.3 Hz, 2H, HC_6), 1.33 – 1.25 (m, 4H, HC_2 + HC_5), 1.20 – 1.07 (m, 4H, HC_3 + HC_4).

^{13}C NMR (151 MHz, $(CD_3)_2SO$) δ 144.6 (C_8), 129.1 (C_9), 128.0 (C_{10}), 126.7 (C_{11}), 66.0 (C_7), 60.6 (C_1), 32.3 (C_2), 31.3 (C_6), 28.3 (C_3), 28.1 (C_5), 25.0 (C_4).

HRMS (HESI$^+$): m/z = 415.1476 $[M+K]^+$ (calcd. 415.1492 for $C_{25}H_{28}OSK$).

Synthesis of 4



4

di-2-Methoxyethyl azodicarboxylate (321 mg, 1.37 mmol, 1.6 eq.) in THF (10 mL) was added to a solution of PPh_3 (364 mg, 1.39 mmol, 1.6 eq.) in THF (10 mL) and the mixture was stirred for 30 min. A solution of **3** (320 mg, 0.85 mmol, 1 eq.) and phenol (115 mg, 1.23 mmol, 1.5 eq.) in THF (19 mL) was added and stirring was continued at r.t. for 17 h. The reaction mixture was partitioned between CH_2Cl_2 (40 mL) and water (40 mL). The aqueous phase was extracted with CH_2Cl_2 (2 × 40 mL). The combined organic phases were dried over Na_2SO_4, filtered and concentrated *in vacuo*. Flash column chromatography (SiO_2, Petroleum ether/EtOAc, 1% → 10%) gave the desired product **4** as colorless oil (260 mg, 0.57 mmol, 68%).

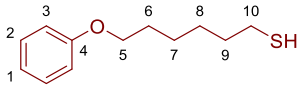
^{1}H NMR (600 MHz, $(CD_3)_2SO$) δ 7.35 – 7.30 (m, 12H, HC_{13} + HC_{14}), 7.28 – 7.25 (m, 2H, HC_2), 7.25 – 7.21 (m, 3H, HC_{15}), 6.90 (t, J = 7.2 Hz, 1H, HC_1), 6.88 (d, J = 7.2 Hz, 2H,

HC_3), 3.87 (t, J = 6.5 Hz, 2H, HC_5), 2.09 (t, J = 7.3 Hz, 2H, HC_{10}), 1.59 (m, 2H, HC_6), 1.31 (m, 2H, HC_9), 1.27 – 1.20 (m, 4H, HC_7 + HC_8).

^{13}C NMR (151 MHz, $(CD_3)_2SO$) δ 158.6 (C_4), 144.6 (C_{12}), 129.5 (C_2), 129.1 (C_{13}), 128.0 (C_{14}), 126.7 (C_{15}), 120.3 (C_1), 114.4 (C_3), 67.1 (C_5), 66.0 (C_{11}), 31.2 (C_{10}), 28.4 (C_6), 28.1 (C_8), 28.0 (C_9), 25.0 (C_7).

HRMS (HESI$^+$): m/z = 491.1793 $[M+K]^+$ (calcd. 491.1805 for $C_{31}H_{32}OSK$).

Synthesis of 5



5

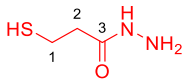
Compound **4** (180 mg, 0.40 mmol) was dissolved in dry CH_2Cl_2 (50 mL) and the solution was cooled to 0 °C. i-Pr_3SiH (0.17 mL, 0.83 mmol, 2.1 eq.) and TFA (2 mL) were added to the solution and the mixture was stirred at r.t. for 4 h. The reaction mixture was concentrated *in vacuo* and column chromatography (SiO_2, Petroleum ether/ CH_2Cl_2, 10% → 20%) gave the desired product **5** as pale yellow liquid (80 mg, 0.38 mmol, 95%).

^{1}H NMR (600 MHz, CD_2Cl_2) δ 7.26 (m, 2H, HC_2), 6.92 (tt, J = 7.3, 1.1 Hz, 1H, HC_1), 6.88 (m, 2H, HC_3), 3.95 (t, J = 6.5 Hz, 2H, HC_5), 2.53 (m, 2H, HC_{10}), 1.78 (p, J = 7.3 Hz, 2H, HC_6), 1.64 (p, J = 7.3 Hz, 2H, HC_9), 1.51 – 1.42 (m, 4H, HC_7 + HC_8), 1.38 (t, J = 7.8 Hz, 1H, SH).

^{13}C NMR (151 MHz, CD_2Cl_2) δ 159.7 (C_4), 129.9 (C_2), 120.9 (C_1), 114.9 (C_3), 68.2 (C_5), 34.6 (C_9), 29.7 (C_6), 28.7 (C_8), 26.1 (C_7), 25.0 (C_{10}).

HRMS (ASAP$^+$): m/z = 211.1143 $[M+H]^+$ (calcd. 211.1151 for $C_{12}H_{19}OS$).

Synthesis of 7 [24]



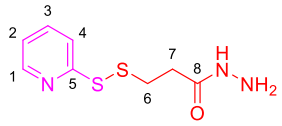
7

Compound **7** was synthesized by a previously reported procedure[24]. Methyl 3-mercaptopropionate (0.9 mL, 8.13 mmol, 1 eq.) and hydrazine hydrate (0.45 mL, 8.94 mmol, 1.1 eq.) were dissolved in methanol (3 mL). The solution was degassed with N_2 for 30 min. The mixture was then stirred at reflux for 7 h. The reaction mixture was concentrated *in vacuo* and column chromatography (SiO_2, Et_2O/MeOH, 10%) gave the desired product **7** as a colourless oil (0.28 g, 2.33 mmol, 29%).

^{1}H NMR (600 MHz, $(CD_3)_2SO$) δ 9.02 (s, 1H, NH), 4.19 (s, 2H, NH_2), 2.65 (dt, J = 8.0, 7.0 Hz, 2H, HC_1), 2.32 (t, J = 7.0 Hz, 2H, HC_2), 2.29 (t, J = 7.8 Hz, 1H, SH).

^{13}C NMR (151 MHz, $(CD_3)_2SO$) δ 169.7 (C_3), 37.5 (C_2), 20.0 (C_1).

Synthesis of 8



8

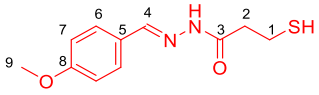
Aldrithiol™-2 (1540 mg, 6.99 mmol, 3 eq.) and Et_3N (65 µL, 0.47 mmol, 0.2 eq.) were dissolved in degassed CH_2Cl_2 (2 mL). A solution of **7** (280 mg, 2.33 mmol) in degassed CH_2Cl_2 (4.5 mL) was added dropwise. The reaction mixture was stirred at r.t. for 17 h and concentrated *in vacuo*. Column chromatography (SiO_2, Et_2O/MeOH, 1% → 20%) gave the desired product **8** as colourless oil (180 mg, 0.78 mmol, 34%).

^{1}H NMR (600 MHz, $(CD_3)_2SO$) δ 9.08 (s, 1H, NH), 8.46 (ddd, J = 4.8, 1.8, 1.0 Hz, 1H, HC_1), 7.84 (ddd, J = 8.1, 7.4, 1.8 Hz, 1H, HC_3), 7.76 (dt, J = 8.1, 1.0 Hz, 1H, HC_4), 7.25 (ddd, J = 7.4, 4.8, 1.0 Hz, 1H, HC_2), 4.20 (s, 2H, NH_2), 3.01 (t, J = 7.1 Hz, 2H, HC_6), 2.44 (t, J = 7.1 Hz, 2H, HC_7).

^{13}C NMR (151 MHz, $(CD_3)_2SO$) δ 169.2 (C_8), 159.1 (C_5), 149.6 (C_1), 137.9 (C_3), 121.2 (C_2), 119.2 (C_4), 33.9 (C_6), 32.8 (C_7).

HRMS (HESI$^+$): *m/z* = 252.0226 $[M+Na]^+$ (calcd. 252.0236 for $C_8H_{11}ON_3S_2Na$).

Synthesis of 20 [24]



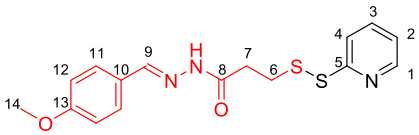
20

Compound **20** was synthesized by a previously reported procedure[24].

^{1}H NMR (600 MHz, $(CD_3)_2SO$) ~0.6:0.4 ratio of *cis*/*trans* amide rotamers; *cis*-**20** isomer. δ 11.24 (s, 1H, NH), 7.92 (s, 1H, HC_4), 7.60 (d, J = 8.3 Hz, 2H, HC_6), 7.00 (m, 2H, HC_7), 3.79 (s, 3H, HC_9), 2.93 (t, J = 7.0 Hz, 2H, HC_2), 2.73 (m, 2H, HC_1), 2.43 (t, J = 8.1 Hz, 1H, SH).

^{13}C NMR (151 MHz, $(CD_3)_2SO$) of *cis*-**20** isomer δ 172.3, 166.6, 160.8, 160.6, 146.0, 142.8, 128.6, 128.3, 126.8, 114.33, 114.31, 55.3, 38.3, 36.4, 19.7, 19.0.

Synthesis of 12



12

Aldrithiol^{TM}-2 (831 mg, 3.78 mmol, 3 eq.) was dissolved in CH_2Cl_2 (2.5 mL) and flushed with nitrogen. To the solution was added AcOH (14 µL, 0.25 mmol, 0.2 eq.) and the solution was stirred at r.t. for 30 min. To the mixture was added suspension of compound **20** (306 mg, 1.26 mmol) in CH_2Cl_2 (1.5 mL). The reaction mixture was stirred at r.t. for 21 h and concentrated *in vacuo*. The residue was purified with flash column chromatography (SiO_2, Petroleum ether/EtOAc, 5% → 90%, then Petroleum ether/EtOAc, 40% → 100%, and then

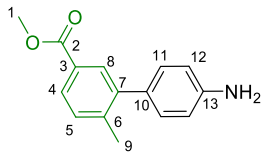
Petroleum ether/EtOAc, 50%) to give the product **12** as a yellow oil (240 mg, 0.69 mmol, 55%).

^{1}H NMR (600 MHz, CD_2Cl_2) ~0.8:0.2 ratio of *cis*/*trans* amide rotamers; *cis*-**12** isomer δ 8.73 (s, 1H, NH), 8.41 (d, J = 4.7 Hz, 1H, HC_1), 7.77 (d, J = 8.1 Hz, 1H, HC_4), 7.65 (s, 1H, HC_9), 7.63 (m, 1H, HC_3), 7.53 (m, 2H, HC_{11}), 7.06 (dd, J = 7.4, 4.7 Hz, 1H, HC_2), 6.91 (m, 2H, HC_{12}), 3.84 (s, 3H, HC_{14}), 3.16 (s, 4H, HC_6 + HC_7).

^{13}C NMR (151 MHz, CD_2Cl_2) of *cis*-**12** isomer δ 173.0 (C_8), 161.9 (C_{13}), 160.8 (C_5), 150.0 (C_1), 143.6 (C_9), 137.6 (C_3), 129.2 (C_{11}), 126.9 (C_{10}), 121.2 (C_2), 120.0 (C_4), 114.7 (C_{12}), 55.9 (C_{14}), 34.2 (C_6), 33.2 (C_7).

HRMS (HESI$^+$): m/z = 370.0640 $[M+Na]^+$ (calcd. 370.0654 for $C_{16}H_{17}O_2N_3NaS_2$).

Synthesis of 23



23

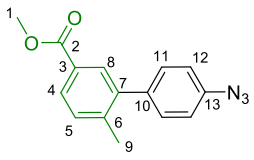
Compound **22** (3.60 g, 16.4 mmol, 1.0 eq.), **21** (4.50 g, 19.6 mmol, 1.2 eq.), $PdCl_2(PPh_3)_2$ (1.10 g, 1.6 mmol, 0.1 eq.) and K_2CO_3 (4.60 g, 32.8 mmol, 2 eq.) were mixed in a mixture of 30 mL of dioxane and 8 mL of water. The reaction mixture was stirred at 90 °C overnight. The mixture was filtered and concentrated *in vacuo*. Purification by flash column chromatography (SiO_2, CH_2Cl_2) afforded compound **23** (3.47 g, 14.4 mmol, 88%) as a dark yellow oil.

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.89 (d, *J* = 1.8 Hz, 1H, HC_8), 7.87 (dd, *J* = 7.9, 1.8 Hz, 1H, HC_4), 7.30 (d, *J* = 7.9 Hz, 1H, HC_5), 7.12 (m, 2H, HC_{11}), 6.74 (m, 2H, HC_{12}), 3.89 (s, 3H, HC_1), 3.72 (br, 2H, NH_2), 2.33 (s, 3H, HC_9).

^{13}C NMR (151 MHz, $CDCl_3$) δ 167.4 (C_2), 145.7 (C_{13}), 142.2 (C_7), 141.3 (C_6), 131.3 (C_{10}), 131.2 (C_8), 130.5 (C_5), 130.2 (C_{11}), 127.9 (C_3 + C_4), 114.9 (C_{12}), 52.1 (C_1), 21.0 (C_9).

HRMS (HESI$^+$): *m/z* = 264.0982 $[M+Na]^+$ (calcd. 264.0995 for $C_{15}H_{15}O_2NNa$).

Synthesis of 24



24

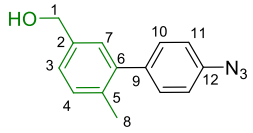
Compound **23** (5.5 g, 22.8 mmol, 1.0 eq.) was dissolved in a 1:1 mixture of CH_2Cl_2 and MeCN (40 mL). Then, at 0 °C *t*-BuONO (4.55 mL, 34.2 mmol, 1.5 eq.) was added followed by the addition of $TMSN_3$ (3.65 mL, 27.4 mmol, 1.2 eq.). The reaction mixture was stirred for 1 h at 0 °C and then another hour at r.t.. The reaction mixture was concentrated *in vacuo* and compound **24** (6.0 g, 22.4 mmol, 98%) was obtained as a yellow oil.

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.92 (dd, *J* = 7.9, 1.9 Hz, 1H, HC_4), 7.89 (d, *J* = 1.9 Hz, 1H, HC_8), 7.34 (d, *J* = 7.9 Hz, 1H, HC_5), 7.31 (m, 2H, HC_{11}), 7.10 (m, 2H, HC_{12}), 3.91 (s, 3H, HC_1), 2.31 (s, 3H, HC_9).

^{13}C NMR (151 MHz, $CDCl_3$) δ 167.2 (C_2), 141.2 (C_6 + C_7), 139.2 (C_{13}), 137.7 (C_{10}), 131.0 (C_8), 130.7 (C_5 + C_{11}), 128.6 (C_4), 128.1 (C_3), 119.0 (C_{12}), 52.2 (C_1), 20.9 (C_9).

Attempts to get HRMS led to decomposition of the compound.

Synthesis of 13



13

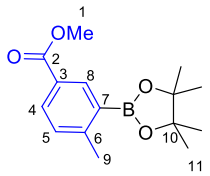
Compound **24** (2.00 g, 7.48 mmol, 1 eq.) was dissolved in dry CH_2Cl_2 (10 mL). Then, at −78 °C a solution of DIBAL-H in toluene (1 M, 22.5 mL, 3 eq.) was added dropwise. The reaction mixture was stirred at −78 °C for 1 h and then at r.t. for 2 h. To the mixture was added MeOH (10 mL) to quench the reaction. The mixture was washed with a saturated aqueous solution of $NaHCO_3$ and extracted twice with CH_2Cl_2. The combined organic layer was dried over $MgSO_4$ and concentrated *in vacuo*. The crude was purified by flash column chromatography (SiO_2, CH_2Cl_2/EtOAc, 5%) to afford compound **13** as a dark yellow oil (1.45 g, 6.06 mmol, 81%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.31 (m, 2H, HC_{10}), 7.27 (m, 2H, HC_3 + HC_4), 7.21 (s, 1H, HC_7), 7.08 (m, 2H, HC_{11}), 4.70 (s, 2H, HC_1), 2.26 (s, 3H, HC_8).

^{13}C NMR (151 MHz, $CDCl_3$) δ 141.2 (C_6), 138.9 (C_{12}), 138.6 (C_9), 138.5 (C_2), 135.0 (C_5), 130.9 (C_4), 130.7 (C_{10}), 128.7 (C_7), 126.3 (C_3), 118.9 (C_{11}), 65.3 (C_1), 20.4 (C_8).

Attempts to get HRMS led to decomposition of the compound.

Synthesis of 25 [27]



25

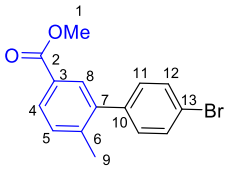
Compound **25** was synthesized by a previously reported procedure[27].

3-bromo-4-methylbenzoate (**21**, 10 g, 43.9 mmol, 1.0 eq.), bis(pinacolato)diboron (13.3 g, 52.4 mmol, 1.2 eq.), KOAc (13 g, 130.6 mmol, 3.0 eq.) and Pd(dppf)Cl_2 (950 mg, 1.3 mmol, 0.03 eq.) were added to a round bottom flask. The flask was flushed with N_2 three times. Then dry dioxane (150 mL) was added and the reaction mixture was stirred at 100 °C overnight. The reaction mixture was cooled to r.t. and diluted with EtOAc. The mixture was filtered and the filtrate was washed with water twice and brine once. The organic layer was dried with anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Flash column chromatography (SiO_2, Petroleum ether/EtOAc, 5% → 10%) gave the product **25** as a white solid (3.90 g, 14.5 mmol, 33%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 8.42 (d, J = 1.8 Hz, 1H, HC_8), 7.98 (dd, J = 8.1, 1.8 Hz, 1H, HC_4), 7.23 (d, J = 8.1 Hz, 1H, HC_5), 3.91 (s, 3H, HC_1), 2.59 (s, 3H, HC_9), 1.36 (s, 12H, HC_{11}).

^{13}C NMR (151 MHz, $CDCl_3$) δ 167.2, 150.4, 137.0, 131.8, 129.9, 129.6, 126.7, 83.7, 51.8, 24.9, 22.4.

Synthesis of 26



26

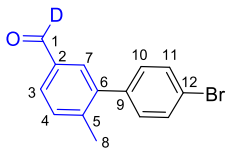
To a solution of **25** (5.52 g, 20.0 mmol, 1.0 eq.) and 4-bromo-1-iodobenzene (6.20 g, 22.0 mmol, 1.1 eq.) in dioxane (100 mL) was added a solution of Na_2CO_3 (4.87 g in 20.0 mL water, 46.0 mmol, 2.3 eq.) and the mixture was purged with nitrogen gas for 30 min. Then, $Pd(PPh_3)_4$ (462 mg, 0.400 mmol, 0.02 eq.) was added and the mixture was stirred at 80 °C for 20 h. The reaction mixture was allowed to cool to r.t. and concentrated *in vacuo*. The residue was dissolved in EtOAc (100 mL). The resulting solution was washed with brine and water, dried with Na_2SO_4, filtered and concentrated *in vacuo*. Flash column chromatography (SiO_2, Petroleum ether/EtOAc, 0% → 5%) gave the desired product **26** as pale yellow oil (4.29 g, 14.1 mmol, 71%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.93 (dd, J = 7.9, 1.8 Hz, 1H, HC_4), 7.87 (d, J = 1.8 Hz, 1H, HC_8), 7.56 (d, J = 8.1 Hz, 2H, HC_{12}), 7.34 (d, J = 7.9 Hz, 1H, HC_5), 7.19 (d, J = 8.1 Hz, 2H, HC_{11}), 3.90 (s, 3H, HC_1), 2.30 (s, 3H, HC_9).

^{13}C NMR (151 MHz, $CDCl_3$) δ 167.1 (C_2), 141.0 (C_6), 140.9 (C_7), 139.9 (C_{10}), 131.5 (C_{12}), 130.9 (C_{11}), 130.8 (C_8), 130.7 (C_5), 128.8 (C_4), 128.1 (C_3), 121.6 (C_{13}), 52.2 (C_1), 20.8 (C_9).

HRMS (ESI$^+$): m/z = 305.0157 $[M+H]^+$ (calcd. 305.0172 for $C_{15}H_{14}BrO_2$).

Synthesis of 27



27

A solution of **26** (2.00 g, 6.56 mmol) in dry CH_2Cl_2 (40 mL) was stirred at –78 °C for 30 min under nitrogen atmosphere. Diisobutylaluminum deuteride solution (0.7 M in toluene, 28 mL, 20 mmol, 3.0 eq.) was added in 5 minutes and then the reaction mixture was stirred for

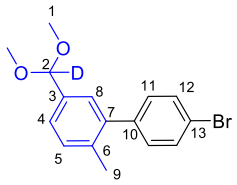
1 h. After warmed to r.t., the reaction was continued for another 2 h before being quenched with MeOH (5.0 mL) and 2 M aqueous HCl solution (5.0 mL). The mixture was extracted with CH_2Cl_2 (2 × 50 mL). The combined organic layer was washed with brine, dried with Na_2SO_4 and concentrated *in vacuo* to give the reduced product. It was then treated with Dess–Martin reagent (4.17 g, 9.84 mmol, 1.5 eq.) in CH_2Cl_2 (40 mL) at r.t.. After 2.5 h, saturated aqueous sodium thiosulfate solution (20 mL) was added. The mixture was extracted with CH_2Cl_2 and the organic phase was washed with saturated aqueous $NaHCO_3$, dried with Na_2SO_4 and concentrated *in vacuo*. Column chromatography (SiO_2, Petroleum ether/EtOAc, 15%) gave the desire compound **27** as colourless oil (1.47 g, 5.35 mmol, 81%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.79 (dd, J = 7.8, 1.8 Hz, 1H, HC_3), 7.71 (d, J = 1.8 Hz, 1H, HC_7), 7.57 (d, J = 8.1 Hz, 2H, HC_{11}), 7.44 (d, J = 7.8 Hz, 1H, HC_4), 7.20 (d, J = 8.1 Hz, 2H, HC_{10}), 2.33 (s, 3H, HC_8).

^{13}C NMR (151 MHz, $CDCl_3$) δ 191.7 (t, J = 26.4 Hz, C_1), 143.0 (C_5), 141.6 (C_6), 139.5 (C_9), 134.6 (t, J = 3.3 Hz, C_2), 131.7 (C_{11}), 131.4 (C_4), 131.1 (C_7), 130.9 (C_{10}), 128.8 (C_3), 121.9 (C_{12}), 21.1 (C_8).

HRMS (APCI$^+$): m/z = 276.0125 $[M+H]^+$ (calcd. 276.0129 for $C_{14}H_{11}DBrO$).

Synthesis of 28



28

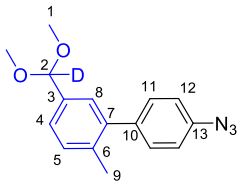
A solution of **27** (1.47 g, 5.35 mmol), *p*-TsOH·H_2O (51.0 mg, 0.268 mmol, 0.05 eq.) and trimethyl orthoformate (2.9 mL, 27 mmol, 5.0 eq.) in MeOH (107 mL) was stirred at r.t. for 16 h. The reaction mixture was partitioned between CH_2Cl_2 (100 mL) and saturated aqueous $NaHCO_3$ (100 mL). The aqueous phase was extracted with CH_2Cl_2 (2 × 50 mL) and the combined organic phase was dried with Na_2SO_4, filtered and concentrated *in vacuo*. Column chromatography (SiO_2, Petroleum ether/EtOAc, 10%) gave the desired product **28** as pale yellow oil (1.40 g, 4.36 mmol, 82%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.53 (d, *J* = 8.4 Hz, 2H, HC_{12}), 7.34 (dd, *J* = 7.8, 1.6 Hz, 1H, HC_4), 7.28 (d, *J* = 1.6 Hz, 1H, HC_8), 7.27 (d, *J* = 7.8 Hz, 1H, HC_5), 7.20 (d, *J* = 8.4 Hz, 2H, HC_{11}), 3.34 (s, 6H, HC_1), 2.25 (s, 3H, HC_9).

^{13}C NMR (151 MHz, $CDCl_3$) δ 140.8 (C_{10}), 140.6 (C_7), 135.9 (C_3), 135.6 (C_6), 131.3 (C_{12}), 131.0 (C_{11}), 130.5 (C_5), 128.0 (C_8), 126.0 (C_4), 121.2 (C_{13}), 102.7 (t, *J* = 24.2 Hz, C_2), 52.9 (C_1), 20.4 (C_9).

HRMS (APCI$^+$): *m/z* = 290.0280 $[M–OCH_3]^+$ (calcd. 290.0285 for $C_{15}H_{13}DBrO$).

Synthesis of 19



19

28 (321 mg, 1.00 mmol), NaN_3 (130 mg, 2.00 mmol, 2.0 eq.), CuI (19.0 mg, 0.100 mmol, 0.1 eq.), sodium ascorbate (9.9 mg, 0.050 mmol, 0.05 eq.), and EtOH/water (5 mL, 7:3, *v*/*v*) were mixed together. The mixture was bubbled with nitrogen gas for 10 min. Then, *N*,*N*'-dimethylethylenediamine (21 μL, 0.20 mmol, 0.2 eq.) was added to the sealed flask. The

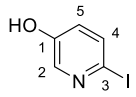
reaction mixture was heated to 80 °C for 2.5 h, after which it was allowed to cool to r.t. and diluted by water (20 mL). The mixture was extracted with CH_2Cl_2 (2 × 20 mL). The combined organic phase was dried with Na_2SO_4, filtered and concentrated *in vacuo*. Flash column chromatography (SiO_2, Petroleum ether/EtOAc, 15%) gave the product **19** as pale yellow oil (267 mg, 0.939 mmol, 95%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.34 (dd, *J* = 7.8, 1.6 Hz, 1H, HC_4), 7.33 (d, *J* = 8.4 Hz, 2H, HC_{11}), 7.29 (d, *J* = 1.6 Hz, 1H, HC_8), 7.27 (d, *J* = 7.8 Hz, 1H, HC_5), 7.07 (d, *J* = 8.4 Hz, 2H, HC_{12}), 3.35 (s, 6H, HC_1), 2.27 (s, 3H, HC_9).

^{13}C NMR (151 MHz, $CDCl_3$) δ 140.9 (C_7), 138.8 (C_{13}), 138.7 (C_{10}), 135.8 (C_3), 135.7 (C_6), 130.8 (C_{11}), 130.5 (C_5), 128.2 (C_8), 125.8 (C_4), 118.9 (C_{12}), 102.8 (t, *J* = 24.2 Hz, C_2), 52.9 (C_1), 20.4 (C_9).

HRMS (ESI$^+$): *m/z* = 253.1191 $[M–OCH_3]^+$ (calcd. 253.1194 for $C_{15}H_{13}DN_3O$).

Synthesis of 30



30

2-Bromo-5-hydroxypyridine (**29**, 5.00 g, 28.7 mmol), NaI (8.59 g, 57.0 mmol, 2 eq.) and CuI (276 mg, 1.44 mmol, 0.05 eq.) were mixed and flushed with argon three times. After the addition of dry degassed dioxane (28 mL) and *trans*-*N,N'*-Dimethylcyclohexane-1,2-diamine (0.45 mL, 2.9 mmol, 0.1 eq.), the reaction mixture was refluxed under argon atmosphere for 21 h. The reaction mixture was concentrated under vacuum and the resulting solid was mixed with water (180 mL). The mixture was extracted with chloroform : *i*-PrOH (3:1, *v*/*v*, 4 × 30 mL). The organic layer was dried with $MgSO_4$ and concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO_2, Petroleum

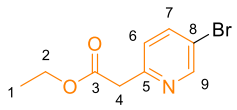
ether/EtOAc, 20% → 30%) to give the mixture of **30** (4.87 g, 22.1 mmol, 77%) and 2-bromo-5-hydroxypyridine (0.79 g, 4.50 mmol, 16%) as yellow solid.

^{1}H NMR (600 MHz, $(CD_3)_2SO$) δ 10.20 (s, 1H, OH), 7.95 (d, J = 3.1 Hz, 1H, HC_2), 7.58 (d, J = 8.5 Hz, 1H, HC_4), 6.95 (dd, J = 3.1, 8.5 Hz, 1H, HC_5).

^{13}C NMR (151 MHz, $(CD_3)_2SO$) δ 154.2 (C_1), 139.7 (C_2), 134.7 (C_4), 125.6 (C_5), 104.3 (C_3).

HRMS (HESI$^+$): m/z = 221.9405 $[M+H]^+$ (calcd. 221.9410 for C_5H_5ONI).

Synthesis of 31 [24]



31

Compound **31** was synthesized by a previously reported procedure[24]. 5-Bromo-2-iodopyridine (5.00 g, 17.6 mmol) was dissolved in dry dioxane (50 mL) and the solution was degassed. Cs_2CO_3 (17.2 g, 52.8 mmol, 3.0 eq.), picolinic acid (430 mg, 3.49 mmol, 0.2 eq.), CuI (300 mg, 1.58 mmol, 0.1 eq.) and diethylmalonate (3.22 mL, 21.2 mmol, 1.2 eq.) were added to the solution under nitrogen atmosphere. The reaction mixture was heated to 70 °C for 48 h. The reaction mixture was then filtered and concentrated *in vacuo*. The residual oil was dissolved in EtOAc (50 mL) and washed with water (2 × 100 mL) and brine. The organic layer was dried with $MgSO_4$, filtered and concentrated *in vacuo*. The crude mixture (7.80 g) was used in the next step without further purification.

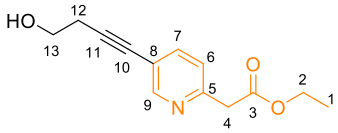
To the solution of the crude (7.80 g) in DMSO (25 mL) was added LiCl (1.10 g, 25.9 mmol, 1.05 eq.). After being flushed with nitrogen gas three times, the reaction mixture was stirred at 160 °C for 3 h. The reaction mixture was allowed to cool down to r.t. and partitioned between Et_2O (120 mL) and water (400 mL). The aqueous layer was extracted with Et_2O (2 × 90 mL) and the combined organic layer was washed with brine (90 mL) and water (90

mL). The organic layer was dried with $MgSO_4$ and concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO_2, Petroleum ether/EtOAc, 5% → 10%) to give **31** (2.40 g, 9.83 mmol, 56% in two steps) as yellow oil.

^1H NMR (600 MHz, $CDCl_3$) δ 8.61 (d, J = 2.4 Hz, 1H, HC_9), 7.78 (dd, J = 8.3, 2.4 Hz, 1H, HC_7), 7.21 (d, J = 8.3 Hz, 1H, HC_6), 4.18 (q, J = 7.1 Hz, 2H, HC_2), 3.80 (s, 2H, HC_4), 1.26 (t, J = 7.1 Hz, 3H, HC_1).

^{13}C NMR (151 MHz, $CDCl_3$) δ 170.3, 153.2, 150.7, 139.3, 125.4, 119.4, 61.4, 43.4, 14.3.

Synthesis of 32



32

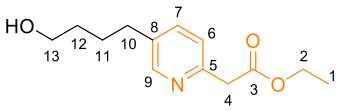
CuI (561 mg, 2.95 mmol, 0.36 eq.), $Pd(PPh_3)_4$ (1.70 g, 1.47 mmol, 0.18 eq.) and 3-butyn-1-ol (0.87 mL, 11.5 mmol, 1.4 eq.) were combined in a degassed mixture of dry THF:Et_3N (1:1, *v*/*v*, 16 mL) under argon atmosphere. The mixture was stirred at r.t. for 15 min and then the solution of **31** (2.02 g, 8.29 mmol) in a degassed 1:1 mixture of dry THF:Et_3N (1:1, *v*/*v*, 16 mL) was added. The mixture was stirred at 70 °C for 14 h and then diluted with EtOAc (100 mL). The organic layer was washed with saturated aqueous NH_4Cl solution (150 mL), aqueous Na_4EDTA solution (0.5 M, 150 mL), 5% aqueous LiCl solution (150 mL) and brine (150 mL). The water layer was combined and extracted with EtOAc (80 mL). The organic layer was dried with $MgSO_4$ and concentrated *in vacuo.* The crude product was purified by flash column chromatography (SiO_2, CH_2Cl_2/EtOAc, 40% → 70%) to give **32** (1.86 g, 7.97 mmol, 96%) as brown solid.

^{1}H NMR (600 MHz, $CDCl_3$) δ 8.58 (d, J = 2.2 Hz, 1H, HC_9), 7.66 (dd, J = 8.0, 2.2 Hz, 1H, HC_7), 7.24 (d, J = 8.0 Hz, 1H, HC_6), 4.18 (q, J = 7.1 Hz, 2H, HC_2), 3.84 (t, J = 6.3 Hz, 2H, HC_{13}), 3.82 (s, 2H, HC_4), 2.71 (t, J = 6.3 Hz, 2H, HC_{12}), 1.25 (t, J = 7.1 Hz, 3H, HC_1).

^{13}C NMR (151 MHz, $CDCl_3$) δ 170.5 (C_3), 153.4 (C_5), 152.1 (C_9), 139.4 (C_7), 123.4 (C_6), 119.1 (C_8), 90.3 (C_{11}), 79.1 (C_{10}), 61.3 (C_2), 61.1 (C_{13}), 43.9 (C_4), 24.0 (C_{12}), 14.3 (C_1).

HRMS (HESI$^+$): m/z = 256.0932 $[M+Na]^+$ (calcd. 256.0944 for $C_{13}H_{15}O_3NNa$).

Synthesis of 33



33

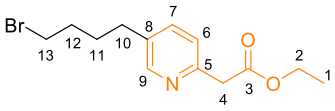
To the solution of **32** (1.85 g, 7.93 mmol) in EtOH (43 mL), 10% Pd on C (441 mg, 0.4 mmol of Pd, 0.05 eq.) was added. The reaction mixture was stirred at r.t. under hydrogen atmosphere for 4 h. The catalyst was removed by filtration through a pad of Celite, and the filtrate was concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO_2, CH_2Cl_2/EtOAc, 60% → 80%) to give **33** (1.30 g, 5.48 mmol, 69%) as pale brown oil.

^{1}H NMR (600 MHz, $CDCl_3$) δ 8.38 (d, J = 2.2 Hz, 1H, HC_9), 7.48 (dd, J = 7.9, 2.2 Hz, 1H, HC_7), 7.21 (d, J = 7.9 Hz, 1H, HC_6), 4.18 (q, J = 7.1 Hz, 2H, HC_2), 3.80 (s, 2H, HC_4), 3.67 (q, J = 6.1 Hz, 2H, HC_{13}), 2.63 (t, J = 7.7 Hz, 2H, HC_{10}), 1.70 (m, 2H, HC_{11}), 1.61 (m, 2H, HC_{12}), 1.26 (t, J = 7.1 Hz, 3H, HC_1).

^{13}C NMR (151 MHz, $CDCl_3$) δ 171.0 (C_3), 152.1 (C_5), 149.6 (C_9), 136.7 (C_7), 136.0 (C_8), 123.6 (C_6), 62.7 (C_{13}), 61.1 (C_2), 43.7 (C_4), 32.5 (C_{10}), 32.3 (C_{12}), 27.4 (C_{11}), 14.3 (C_1).

HRMS (HESI$^+$): m/z = 238.1431 $[M+H]^+$ (calcd. 238.1438 for $C_{13}H_{20}O_3N$).

Synthesis of 34



34

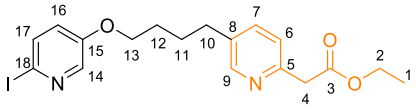
To the solution of **33** (1.19 g, 5.01 mmol) in dry CH_2Cl_2 (9 mL), CBr_4 (2.40 g, 7.24 mmol, 1.44 eq.) was added. The solution was flushed with argon and cooled to 0 °C. PPh_3 (1.90 g, 7.24 mmol, 1.44 eq.) in dry CH_2Cl_2 (9 mL) was added dropwise to the solution and the resulting mixture was stirred for 4 h at r.t.. The reaction mixture was concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO_2, Petroleum ether/EtOAc, 30% → 40%) to give **34** (0.89 g, 2.96 mmol, 59%) as pale yellow oil.

^{1}H NMR (600 MHz, $CDCl_3$) δ 8.38 (d, J = 2.3 Hz, 1H, HC_9), 7.48 (dd, J = 7.9, 2.3 Hz, 1H, HC_7), 7.23 (d, J = 7.9 Hz, 1H, HC_6), 4.18 (q, J = 7.2 Hz, 2H, HC_2), 3.81 (s, 2H, HC_4), 3.42 (t, J = 6.6 Hz, 2H, HC_{13}), 2.63 (t, J = 7.7 Hz, 2H, HC_{10}), 1.90 (m, 2H, HC_{12}), 1.78 (m, 2H, HC_{11}), 1.26 (t, J = 7.2 Hz, 3H, HC_1).

^{13}C NMR (151 MHz, $CDCl_3$) δ 171.0 (C_3), 152.2 (C_5), 149.6 (C_9), 136.7 (C_7), 135.5 (C_8), 123.7 (C_6), 61.2 (C_2), 43.6 (C_4), 33.5 (C_{13}), 32.2 (C_{12}), 31.9 (C_{10}), 29.6 (C_{11}), 14.3 (C_1).

HRMS (HESI$^+$): m/z = 300.0581 $[M+H]^+$ (calcd. 300.0594 for $C_{13}H_{19}O_2NBr$).

Synthesis of 35



35

30 (178 mg, 0.80 mmol, 1.1 eq.) was dissolved in MeCN (3 mL). K_2CO_3 (187 mg, 1.35 mmol, 1.9 eq.) was added and the mixture was stirred at 60 °C for 30 min. **34** (209 mg, 0.70 mmol, 1 eq.) in MeCN (3 mL) was added and the temperature was raised to 80 °C. The reaction mixture was stirred for 16 h and then concentrated. The resulting solid was mixed with water (20 mL) and extracted with EtOAc (3 × 20 mL). The combined organic layer was

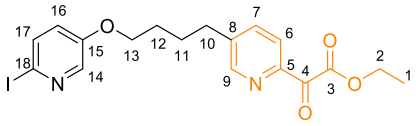
dried with $MgSO_4$ and concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO_2, Petroleum ether/EtOAc, 40% → 50%) to give **35** (250 mg, 0.57 mmol, 81%) as colorless oil.

^{1}H NMR (600 MHz, $CDCl_3$) δ 8.40 (d, J = 2.3 Hz, 1H, HC_9), 8.07 (d, J = 3.2 Hz, 1H, HC_{14}), 7.56 (d, J = 8.6 Hz, 1H, HC_{17}), 7.49 (dd, J = 7.9, 2.3 Hz, 1H, HC_7), 7.23 (d, J = 7.9 Hz, 1H, HC_6), 6.88 (dd, J = 8.6, 3.2 Hz, 1H, HC_{16}), 4.18 (q, J = 7.2 Hz, 2H, HC_2), 3.98 (t, J = 5.8 Hz, 2H, HC_{13}), 3.81 (s, 2H, HC_4), 2.67 (t, J = 7.2 Hz, 2H, HC_{10}), 1.81 (m, 4H, HC_{11} + HC_{12}), 1.26 (t, J = 7.2 Hz, 3H, HC_1).

^{13}C NMR (151 MHz, $CDCl_3$) δ 171.0 (C_3), 155.6 (C_{15}), 152.3 (C_5), 149.6 (C_9), 138.8 (C_{14}), 136.7 (C_7), 135.6 (C_8), 134.9 (C_{17}), 124.3 (C_{16}), 123.7 (C_6), 105.9 (C_{18}), 68.3 (C_{13}), 61.2 (C_2), 43.7 (C_4), 32.4 (C_{10}), 28.7 (C_{12}), 27.5 (C_{11}), 14.3 (C_1).

HRMS (HESI$^+$): m/z = 441.0652 $[M+H]^+$ (calcd. 441.0670 for $C_{18}H_{22}O_3N_2I$).

Synthesis of 36



36

35 (267 mg, 0.61 mmol) was dissolved in dry dioxane (8 mL). SeO_2 (90 mg, 0.81 mmol, 1.3 eq.) was added and the mixture was stirred at 90 °C for 2 h under argon atmosphere. After cooling to r.t., the reaction mixture was concentrated *in vacuo*. The residue was dissolved in CH_2Cl_2, filtered with Celite and concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO_2, Petroleum ether/EtOAc, 30%) to give **36** (223 mg, 0.49 mmol, 80%) as pale orange oil.

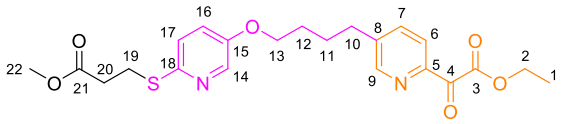
^{1}H NMR (600 MHz, $CDCl_3$) δ 8.59 (d, J = 2.1 Hz, 1H, HC_9), 8.07 (d, J = 3.1 Hz, 1H, HC_{14}), 8.05 (d, J = 8.0 Hz, 1H, HC_6), 7.72 (dd, J = 8.0, 2.1 Hz, 1H, HC_7), 7.57 (d, J = 8.6 Hz, 1H,

HC_{17}), 6.88 (dd, J = 8.6, 3.1 Hz, 1H, HC_{16}), 4.48 (q, J = 7.1 Hz, 2H, HC_2), 3.99 (m, 2H, HC_{13}), 2.79 (m, 2H, HC_{10}), 1.85 (m, 4H, HC_{11} + HC_{12}), 1.42 (t, J = 7.1 Hz, 3H, HC_1).

^{13}C NMR (151 MHz, $CDCl_3$) δ 187.6 (C_4), 165.6 (C_3), 155.5 (C_{15}), 150.2 (C_9), 148.6 (C_5), 143.1 (C_8), 138.7 (C_{14}), 137.0 (C_7), 134.9 (C_{17}), 124.3 (C_{16}), 123.5 (C_6), 106.1 (C_{18}), 68.1 (C_{13}), 62.3 (C_2), 33.1 (C_{10}), 28.7 (C_{12}), 27.3 (C_{11}), 14.3 (C_1).

HRMS (HESI$^+$): m/z = 455.0447 $[M+H]^+$ (calcd. 455.0462 for $C_{18}H_{20}O_4N_2I$).

Synthesis of 14



14

Under nitrogen atmosphere, DIPEA (170 µL, 0.97 mmol, 2.5 eq.) was dissolved in dry degassed dioxane (1 mL). The solution was flushed with nitrogen gas and cooled to 0 °C. To the solution was added methyl 3-mercaptopropionate (52 µL, 0.47 mmol, 1.2 eq.). The reaction mixture was stirred at r.t. for 80 min. In another flask, **36** (175 mg, 0.39 mmol) was dissolved in dry degassed dioxane (4 mL). To the solution was added $Pd_2(dba)_3$ (34 mg, 0.037 mmol, 0.1 eq.), Xantphos (27 mg, 0.047 mmol, 0.12 eq.). The solution was flushed with argon three times and stirred at r.t. for 30 min. Two solutions were mixed together, flushed with argon, and stirred at 110 °C for 25 h. The reaction mixture was cooled down to r.t. and dried *in vacuo*. To the residue was added water (20 mL) and brine (10 mL). The mixture was extracted with EtOAc (3 × 20 mL). The combined organic layer was dried with $MgSO_4$ and concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO_2, Petroleum ether/EtOAc, 30% → 40%) to give **14** (136 mg, 0.30 mmol, 78%) as pale orange oil.

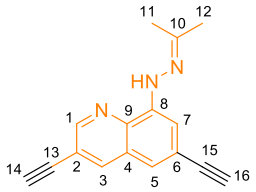
^1H NMR (600 MHz, CD_2Cl_2) δ 8.58 (d, J = 2.1 Hz, 1H, HC_9), 8.14 (d, J = 2.6 Hz, 1H, HC_{14}), 8.02 (d, J = 8.0 Hz, 1H, HC_6), 7.75 (dd, J = 8.0, 2.1 Hz, 1H, HC_7), 7.10 (m, 2H, HC_{16} +

HC_{17}), 4.43 (q, J = 7.1 Hz, 2H, HC_2), 4.00 (m, 2H, HC_{13}), 3.66 (s, 3H, HC_{22}), 3.34 (t, J = 7.2 Hz, 2H, HC_{19}), 2.80 (m, 2H, HC_{10}), 2.71 (t, J = 7.2 Hz, 2H, HC_{20}), 1.84 (m, 4H, HC_{11} + HC_{12}), 1.38 (t, J = 7.1 Hz, 3H, HC_1).

^{13}C NMR (151 MHz, CD_2Cl_2) δ 188.3 (C_4), 172.9 (C_{21}), 166.2 (C_3), 153.5 (C_{15}), 150.6 (C_9), 149.0 (C_{18}), 148.8 (C_5), 144.1 (C_8), 137.5 (C_7 + C_{14}), 123.6 (C_6 + C_{16} + C_{17}), 68.7 (C_{13}), 62.6 (C_2), 52.1 (C_{22}), 35.1 (C_{20}), 33.4 (C_{10}), 29.2 (C_{12}), 27.7 (C_{11}), 26.3 (C_{19}), 14.5 (C_1).

HRMS (ASAP$^+$): m/z = 447.1576 $[M+H]^+$ (calcd. 447.1584 for $C_{22}H_{27}O_6N_2S$).

Synthesis of 15 [25]



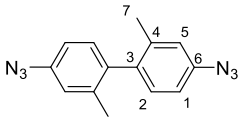
15

Compound **15** was synthesized by a previously reported procedure[25].

^1H NMR (600 MHz, $CDCl_3$) δ 8.98 (s, 1H, NH), 8.74 (d, J = 2.0 Hz, 1H, HC_1), 8.17 (d, J = 2.0 Hz, 1H, HC_3), 7.62 (d, J = 1.7 Hz, 1H, HC_7), 7.33 (d, J = 1.7 Hz, 1H, HC_5), 3.30 (s, 1H, HC_{14}), 3.16 (s, 1H, HC_{16}), 2.16 (s, 3H, HC_{12}), 2.08 (s, 3H, HC_{11}).

^{13}C NMR (151 MHz, $CDCl_3$) δ 149.7 (C_1), 147.2 (C_{10}), 141.2 (C_8), 139.0 (C_3), 135.4 (C_9), 126.9 (C_6), 122.3 (C_4), 119.9 (C_5), 117.1 (C_2), 111.3 (C_7), 84.0 (C_{15}), 80.8 (C_{13}), 80.6 (C_{14}), 77.7 (C_{16}), 25.4 (C_{12}), 16.2 (C_{11}).

Synthesis of 16



16

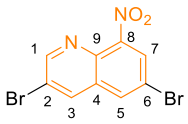
2,2'-dimethyl-[1,1'-biphenyl]-4,4'-diamine (**37**, 50 mg, 0.24 mmol) was dissolved in MeCN (1.0 mL) in a 10 mL flask and cooled to –20 °C. To this mixture, *t*-butyl nitrite (85 µL, 0.72 mmol, 3.0 eq.) was added followed by TMS-N_3 (77 µL, 0.58 mmol, 2.4 eq.). The reaction mixture was then allowed to warm to r.t. slowly, and stirred for another 3 h. The reaction mixture was concentrated *in vacuo* and purified by flash column chromatography (SiO_2, Petroleum ether/ CH_2Cl_2, 50%) to give the desired product **16** as colourless oil (44 mg, 0.17 mmol, 70%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.06 (d, J = 8.0 Hz, 2H, HC_2), 6.93 (d, J = 2.4 Hz, 2H, HC_5), 6.90 (dd, J = 8.0, 2.4 Hz, 2H, HC_1), 2.04 (s, 6H, HC_7).

^{13}C NMR (151 MHz, $CDCl_3$) δ 139.1 (C_6), 138.1 (C_4), 137.6 (C_3), 130.9 (C_2), 120.4 (C_5), 116.5 (C_1), 20.1 (C_7).

Attempts to get HRMS led to decomposition of the compound.

Synthesis of 39



39

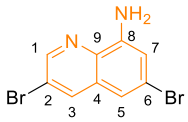
8-Nitroquinoline (**38**, 10 g, 57.4 mmol) and NBS (30.0 g, 169 mmol, 2.9 eq.) were dissolved in AcOH (200 mL) and the solution was stirred at 125 °C for three days. The reaction mixture was allowed to cool down to r.t. and neutralized with saturated aqueous Na_2CO_3 solution and 20% (*w*/*w*) aqueous NaOH solution. The mixture was extracted with chloroform twice. The organic layer was dried with $MgSO_4$ and concentrated *in vacuo.* The crude product was purified by flash column chromatography (SiO_2, Petroleum ether/EtOAc, 20%). The fractions with impurity were further purified by recrystallization (EtOH). Overall, **39** (11.4 g, 34.4 mmol, 60%) was obtained as off-white solid.

^{1}H NMR (600 MHz, $(CD_3)_2SO$) δ 9.14 (d, *J* = 2.2 Hz, 1H, HC_1), 8.90 (d, *J* = 2.2 Hz, 1H, HC_3), 8.65 (d, *J* = 2.1 Hz, 1H, HC_7), 8.57 (d, *J* = 2.1 Hz, 1H, HC_5).

^{13}C NMR (151 MHz, $(CD_3)_2SO$) δ 153.8 (C_1), 148.2 (C_9), 137.3 (C_3), 135.4 (C_4), 133.2 (C_5), 130.5 (C_8), 126.6 (C_7), 119.7 (C_2), 119.0 (C_6).

HRMS (ESI^-): *m/z* = 328.8562 $[M-H]^-$ (calcd. 328.8567 for $C_9H_3O_2N_2Br_2$).

Synthesis of 40



40

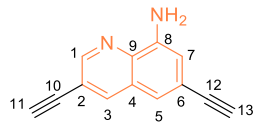
39 (5.37 g, 16.2 mmol) was dissolved in EtOH : water (90:10, *v*/*v*, 230 mL). Iron (3.25 g, 58.2 mmol, 3.6 eq.) and catalytic amount of aqueous HCl (37%, *w*/*w*, 0.8 mL) were added and the reaction mixture was heated to 90 °C for 6 h. After cooling it down, the mixture was concentrated under reduced pressure. To the resulting solid, aqueous $NaHCO_3$ (200 mL) was added and the aqueous layer was extracted with CH_2Cl_2 (3 × 50 mL). After the drying with $MgSO_4$, the organic layer was filtered through celite and concentrated *in vacuo* to give **40** (4.12 g, 13.6 mmol, 84%) as black solid.

^{1}H NMR (600 MHz, $(CD_3)_2SO$) δ 8.75 (d, *J* = 2.2 Hz, 1H, HC_1), 8.49 (d, *J* = 2.2 Hz, 1H, HC_3), 7.23 (d, *J* = 2.0 Hz, 1H, HC_5), 6.97 (d, *J* = 2.0 Hz, 1H, HC_7), 6.35 (s, 2H, NH_2).

^{13}C NMR (151 MHz, $(CD_3)_2SO$) δ 147.6 (C_1), 147.3 (C_9), 136.4 (C_3), 134.2 (C_8), 130.8 (C_4), 122.3 (C_6), 118.1 (C_2), 114.0 (C_5), 111.2 (C_7).

HRMS ($APCI^+$): *m/z* = 300.8964 $[M+H]^+$ (calcd. 300.8971 for $C_9H_7N_2Br_2$).

Synthesis of 17



17

CuI (568 mg, 2.99 mmol, 0.4 eq.), $Pd(PPh_3)_2Cl_2$ (1.05 g, 1.50 mmol, 0.2 eq.) and TMS-acetylene (3.10 mL, 22.5 mmol, 3.0 eq.) were combined in a degassed mixture of dry DMF:Et_3N (1:1, *v*/*v*, 20 mL) under argon atmosphere. The mixture was stirred at r.t. for 5 min and then the solution of **40** (2.26 g, 7.48 mmol) in a degassed 1:1 mixture of dry DMF:Et_3N (1:1, *v*/*v*, 5 mL) was added. The mixture was stirred at 60 °C for 24 h and then diluted with EtOAc (100 mL). The organic layer was washed with saturated aqueous NH_4Cl solution (50 mL), aqueous Na_4EDTA solution (2 M, 20 mL), water (50 mL) and brine (50 mL). The organic layer was dried with $MgSO_4$ and concentrated *in vacuo* to give a brown oil as the crude product.

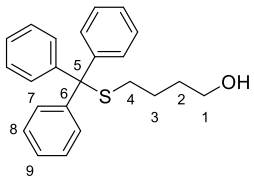
The crude product was dissolved in THF (50 mL). The solution was treated with TBAF (1 M in THF, 16.4 mL, 2.2 eq.) and the reaction mixture was stirred at r.t. for 10 min. To the reaction mixture, saturated aqueous $NaHCO_3$ solution (20 mL) and water (20 mL) were added. The mixture was extracted with EtOAc (3 × 30 mL) and the combined organic layer was dried with $MgSO_4$ and concentrated *in vacuo.* The crude product was purified by flash column chromatography (SiO_2, petroleum ether / EtOAc, 10%) to give **17** (828 mg, 4.31 mmol, 58% over two steps) as yellow solid.

^1H NMR (600 MHz, $CDCl_3$) δ 8.76 (d, J = 2.2 Hz, 1H, HC_1), 8.13 (d, J = 2.2 Hz, 1H, HC_3), 7.28 (d, J = 1.47 Hz, 1H, HC_5), 6.97 (d, J = 1.47 Hz, 1H, HC_7), 5.00 (br, 2H, NH_2), 3.27 (s, 1H, HC_{11}), 3.12 (s, 1H, HC_{13}).

^{13}C NMR (151 MHz, $CDCl_3$) δ 150.2 (C_1), 144.0 (C_8), 139.2 (C_3), 137.0 (C_9), 127.3 (C_4), 121.9 (C_6), 120.2 (C_5), 117.3 (C_2), 113.3 (C_7), 83.8 (C_{12}), 81.0 (C_{10}), 80.8 (C_{11}), 77.8 (C_{13}).

HRMS (ESI^+): *m/z* = 193.0757 $[M+H]^+$ (calcd. 193.0767 for $C_{13}H_9N_2$).

Synthesis of 45



45

To a solution of γ-thiobutyrolactone (**44**, 1.00 g, 9.80 mmol) in EtOH (20 mL) was added sodium borohydride (1.86 g, 49.0 mmol, 5.0 eq.) in one portion, and the mixture was refluxed for 20 min. After being cooled down to r.t., the mixture was concentrated to give a slurry. It was then dissolved in water and acidified to pH = 2 with 2 M aqueous HCl solution. The mixture was extracted with CH_2Cl_2 (3 × 50 mL) and the combined organic phase was washed with brine, dried with $MgSO_4$ and concentrated *in vacuo* to give the crude product as a yellow oil (0.99 g, 9.34 mmol). The crude product (4-mercapto-1-butanol) was used in the next step without further purification.

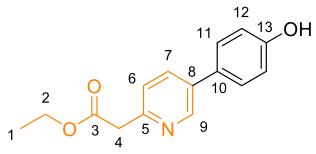
A solution of 4-mercapto-1-butanol (0.99 g, 9.34 mmol), trityl chloride (2.61 g, 9.34 mmol, 1.0 eq.) and DMAP (114 mg, 0.934 mmol, 0.10 eq.) in CH_2Cl_2 (50 mL) was stirred at r.t. for 2 h. The reaction mixture was partitioned between water (50 mL) and CH_2Cl_2 (50 mL). The aqueous phase was extracted with CH_2Cl_2 (2 × 20 mL) and the combined organic layer was washed with brine (20 mL), dried with $MgSO_4$ and concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO_2, CH_2Cl_2 / MeOH, 0% → 2%) to give **45** (1.93 g, 5.54 mmol, 57%) as a colorless oil.

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.42 (d, J = 7.7 Hz, 6H, HC_7), 7.28 (m, 6H, HC_8), 7.21 (m, 3H, HC_9), 3.50 (t, J = 6.2 Hz, 2H, HC_1), 2.18 (t, J = 7.3 Hz, 2H, HC_4), 1.51 (m, 2H, HC_2), 1.45 (m, 2H, HC_3), 1.16 (br, 1H, OH).

^{13}C NMR (151 MHz, $CDCl_3$) δ 145.1 (C_6), 129.7 (C_7), 128.0 (C_8), 126.7 (C_9), 66.6 (C_5), 62.5 (C_1), 32.1 (C_2), 31.8 (C_4), 25.0 (C_3).

HRMS (ESI$^+$): *m/z* = 371.1435 $[M+Na]^+$ (calcd. 371.1440 for $C_{23}H_{24}OSNa$).

Synthesis of 46



46

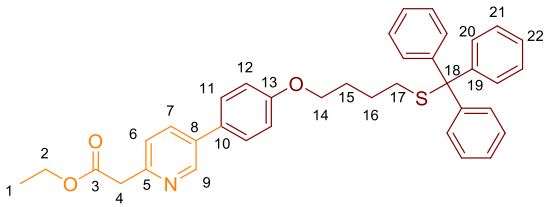
4-hydroxyphenylboronic acid (304 mg, 2.20 mmol, 1.1 eq.), pyridine **31** (486 mg, 2.00 mmol), Na_2CO_3 (636 mg, 6.00 mmol, 3.0 eq.) and $Pd(AmPhos)_2Cl_2$ (28 mg, 0.040 mmol, 0.02 eq.) were combined in a 100 mL Schlenk tube and flushed with nitrogen gas three times. Then, EtOH/water (5:1, *v*/*v*, 20 mL, degassed) was added to the tube and the mixture was heated to 60 °C for 2 h. The mixture was cooled to r.t., diluted with water (200 mL) and extracted with EtOAc (2 × 200 mL). The combined organic layers were dried with $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (SiO_2, Petroleum ether/EtOAc, 50%) gave the desired product **46** as pale yellow solid (481 mg, 1.87 mmol, 94%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 8.71 (d, *J* = 2.4 Hz, 1H, HC_9), 7.81 (dd, *J* = 8.1, 2.4 Hz, 1H, HC_7), 7.65 (br, 1H, OH), 7.40 (d, *J* = 8.4 Hz, 2H, HC_{11}), 7.36 (d, *J* = 8.1 Hz, 1H, HC_6), 6.91 (d, *J* = 8.4 Hz, 2H, HC_{12}), 4.19 (q, *J* = 7.1 Hz, 2H, HC_2), 3.91 (s, 2H, HC_4), 1.26 (t, *J* = 7.1 Hz, 3H, HC_1).

^{13}C NMR (151 MHz, $CDCl_3$) δ 171.1 (C_3), 157.0 (C_{13}), 152.1 (C_5), 147.1 (C_9), 135.4 (C_8), 135.1 (C_7), 129.2 (C_{10}), 128.4 (C_{11}), 124.3 (C_6), 116.4 (C_{12}), 61.5 (C_2), 43.2 (C_4), 14.3 (C_1).

HRMS (ESI$^+$): *m/z* = 258.1116 $[M+H]^+$ (calcd. 258.1125 for $C_{15}H_{16}O_3N$).

Synthesis of 18



18

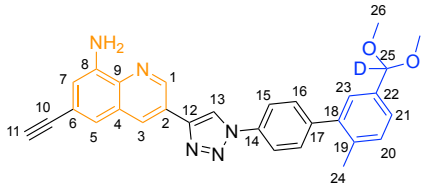
Di-2-methoxyethylazodicarboxylate (161 mg, 0.690 mmol, 1.2 eq.) in THF (10 mL) was added to a solution of PPh_3 (181 mg, 0.690 mmol, 1.2 eq.) in THF (10 mL) and the mixture was stirred at r.t. for 30 min. A solution of **45** (200 mg, 0.575 mmol, 1.0 eq.) and **46** (162 mg, 0.632 mmol, 1.1 eq.) in THF (20 mL) was added and stirring was continued for 16 h. The reaction mixture was partitioned between CH_2Cl_2 (40 mL) and water (40 mL). The water layer was extracted with CH_2Cl_2 (2 × 20 mL). The combined organic layer was dried over $MgSO_4$ and concentrated *in vacuo*. Flash column chromatography (SiO_2, CH_2Cl_2/EtOAc, 0% → 10%) gave the desired product **18** as a yellow oil which solidified into a yellow crystal (275 mg, 0.468 mmol, 81%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 8.74 (d, J = 2.2 Hz, 1H, HC_9), 7.80 (dd, J = 8.1, 2.6 Hz, 1H, HC_7), 7.47 (d, J = 8.8 Hz, 2H, HC_{11}), 7.42 (d, J = 7.4 Hz, 6H, HC_{20}), 7.34 (d, J = 8.1 Hz, 1H, HC_6), 7.27 (m, 6H, HC_{21}), 7.21 (m, 3H, HC_{22}), 6.91 (d, J = 8.8 Hz, 2H, HC_{12}), 4.21 (q, J = 7.1 Hz, 2H, HC_2), 3.87 (s, 2H, HC_4), 3.85 (t, J = 6.2 Hz, 2H, HC_{14}), 2.24 (t, J = 7.3 Hz, 2H, HC_{17}), 1.76 (m, 2H, HC_{15}), 1.59 (m, 2H, HC_{16}), 1.28 (t, J = 7.1 Hz, 3H, HC_1).

^{13}C NMR (151 MHz, $CDCl_3$) δ 171.0 (C_3), 159.2 (C_{13}), 152.6 (C_5), 147.7 (C_9), 145.1 (C_{19}), 134.9 (C_8), 134.7 (C_7), 130.0 (C_{10}), 129.7 (C_{20}), 128.2 (C_{11}), 128.0 (C_{21}), 126.7 (C_{22}), 123.8 (C_6), 115.2 (C_{12}), 67.5 (C_{14}), 66.7 (C_{18}), 61.3 (C_2), 43.7 (C_4), 31.7 (C_{17}), 28.6 (C_{15}), 25.4 (C_{16}), 14.4 (C_1).

HRMS (ESI$^+$): *m/z* = 588.2559 $[M+H]^+$ (calcd. 588.2567 for $C_{38}H_{38}NO_3S$).

Synthesis of 54



54

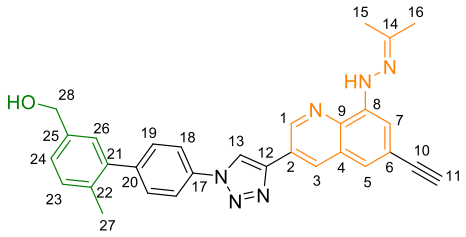
TBTA (65.8 mg, 0.124 mmol, 0.10 eq.) and $[Cu(MeCN)_4]PF_6$ (23.1 mg, 0.062 mmol, 0.05 eq.) were mixed in degassed dry THF (1.0 mL) and stirred under argon for 30 min. To this mixture, a solution of amine **17** (262 mg, 1.36 mmol, 1.1 eq.) and azide **19** (352 mg, 1.24 mmol, 1.0 eq.) in degassed THF (15.0 mL) was added. The reaction mixture was stirred at r.t. for 16 h and concentrated *in vacuo*. The residue was purified by flash column chromatography (SiO_2, Petroleum ether/EtOAc, 30% → 50%) to give compound **54** as a yellow solid (340 mg, 0.714 mmol, 58%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 9.20 (d, *J* = 1.8 Hz, 1H, HC_1), 8.49 (d, *J* = 2.2 Hz, 1H, HC_3), 8.39 (s, 1H, HC_{13}), 7.84 (d, *J* = 8.4 Hz, 2H, HC_{15}), 7.50 (d, *J* = 8.4 Hz, 2H, HC_{16}), 7.38 (m, 1H, HC_{21}), 7.36 (m, 1H, HC_{23}), 7.34 (m, 1H, HC_5), 7.30 (d, *J* = 7.8 Hz, 1H, HC_{20}), 6.92 (d, *J* = 1.5 Hz, 1H, HC_7), 5.05 (br, 2H, NH_2), 3.36 (s, 6H, HC_{26}), 3.12 (s, 1H, HC_{11}), 2.29 (s, 3H, HC_{24}).

^{13}C NMR (151 MHz, $CDCl_3$) δ 146.1 (C_1), 145.5 (C_{12}), 144.0 (C_8), 142.8 (C_{17}), 140.2 (C_{18}), 137.8 (C_9), 136.0 (C_{22}), 135.61 (C_{19}), 135.59 (C_{14}), 131.8 (C_3), 130.7 (C_{16}), 130.6 (C_{20}), 128.1 (C_4), 128.0 (C_{23}), 126.3 (C_{21}), 124.4 (C_2), 121.6 (C_6), 120.7 (C_5), 120.3 (C_{15}), 118.4 (C_{13}), 112.6 (C_7), 102.7 (t, *J* = 24.2 Hz, C_{25}), 84.0 (C_{10}), 77.6 (C_{11}), 52.9 (C_{26}), 20.4 (C_{24}).

HRMS (ESI$^+$): *m/z* = 515.1694 $[M+K]^+$ (calcd. 515.1703 for $C_{29}H_{24}N_5O_2DK$).

Synthesis of 47



47

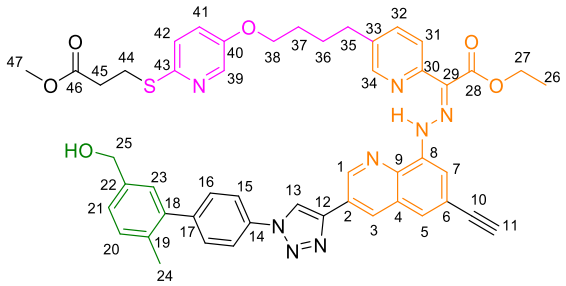
TBTA (32 mg, 0.061 mmol, 0.13 eq.) and $[Cu(MeCN)_4]PF_6$ (11.3 mg, 0.030 mmol, 0.06 eq.) were mixed in degassed THF (6 mL) and stirred under argon for 1 h. To this mixture, a solution of quinoline **15** (129 mg, 0.52 mmol, 1.1 eq.) and azide **13** (113 mg, 0.47 mmol, 1.0 eq.) in degassed THF (6 mL) was added. The reaction mixture was stirred at r.t. under argon atmosphere for 23 h and concentrated *in vacuo*. The residue was purified by flash column chromatography (SiO_2, CH_2Cl_2/Et_2O, 0% → 25%) to give compound **47** as a brown foam (66 mg, 0.14 mmol, 29%).

^{1}H NMR (600 MHz, CD_2Cl_2) δ 9.27 (d, J = 2.2 Hz, 1H, HC_1), 9.07 (br, 1H, NH), 8.58 (d, J = 2.2 Hz, 1H, HC_3), 8.48 (s, 1H, HC_{13}), 7.89 (m, 2H, HC_{18}), 7.56 (m, 2H, HC_{19}), 7.55 (d, J = 1.8 Hz, 1H, HC_7), 7.45 (d, J = 1.8 Hz, 1H, HC_5), 7.31 (m, 2H, HC_{23} + HC_{24}), 7.29 (s, 1H, HC_{26}), 4.70 (s, 2H, HC_{28}), 3.22 (s, 1H, HC_{11}), 2.32 (s, 3H, HC_{27}), 2.13 (s, 3H, HC_{16}), 2.08 (s, 3H, HC_{15}).

^{13}C NMR (151 MHz, CD_2Cl_2) δ 147.8 (C_{14}), 146.4 (C_1), 145.9 (C_{12}), 143.4 (C_{20}), 141.9 (C_8), 141.0 (C_{21}), 139.6 (C_{25}), 136.7 (C_9), 136.3 (C_{17}), 135.2 (C_{22}), 132.2 (C_3), 131.3 (C_{23}), 131.2 (C_{19}), 128.8 (C_{26}), 128.4 (C_4), 127.0 (C_{24}), 125.2 (C_2), 122.6 (C_6), 120.9 (C_5 + C_{18}), 119.2 (C_{13}), 110.8 (C_7), 84.7 (C_{10}), 77.9 (C_{11}), 65.3 (C_{28}), 25.6 (C_{16}), 20.5 (C_{27}), 16.4 (C_{15}).

HRMS (HESI$^+$): m/z = 487.2226 $[M+H]^+$ (calcd. 487.2241 for $C_{30}H_{27}N_6O$).

Synthesis of 48



***EZ*-48**

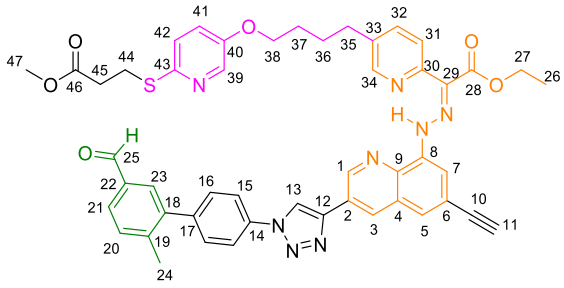
In the mixture of THF (1.4 mL) and EtOH : water (5:1, *v*/*v*, 1 mL), compound **14** (70 mg, 0.16 mmol, 1.1 eq.), compound **47** (70 mg, 0.14 mmol, 1 eq.) and TFA (33 µL, 0.43 mmol, 3.1 eq.) were dissolved. The reaction mixture was heated to 40 °C and stirred for 17 h. The reaction mixture was concentrated *in vacuo* and mixed with saturated aqueous $NaHCO_3$ solution (10 mL). The mixture was extracted with chloroform : *i*-PrOH (3:1, *v*/*v*, 4 × 10 mL). The organic layer was dried over $MgSO_4$ and concentrated *in vacuo*. The residue was purified by flash column chromatography (SiO_2, CH_2Cl_2/EtOAc, 10% → 50%, with 1% Et_3N) to give compound **48** as a brown paste (92 mg, 0.11 mmol, 73%).

^{1}H NMR (600 MHz, CD_2Cl_2) ~1:1 ratio of *E*/*Z*; *E*-**48** isomer δ 15.58 (s, 1H, NH), 9.49 (d, *J* = 2.1 Hz, 1H, HC_1), 8.79 (d, *J* = 2.1 Hz, 1H, HC_{34}), 8.63 (d, *J* = 2.1 Hz, 1H, HC_3), 8.54 (s, 1H, HC_{13}), 8.16 (m, 1H, HC_{39}), 8.10 (d, *J* = 8.4 Hz, 1H, HC_{31}), 7.94 (d, *J* = 1.7 Hz, 1H, HC_7), 7.90 (m, 2H, HC_{15}), 7.74 (dd, *J* = 8.4, 2.3 Hz, 1H, HC_{32}), 7.68 (m, 1H, HC_5), 7.56 (m, 2H, HC_{16}), 7.31 (m, 2H, HC_{20} + HC_{21}), 7.29 (s, 1H, HC_{23}), 7.12 (m, 2H, HC_{41} + HC_{42}), 4.70 (s, 2H, HC_{25}), 4.43 (q, *J* = 7.1 Hz, 2H, HC_{27}), 4.04 (m, 2H, HC_{38}), 3.64 (s, 3H, HC_{47}), 3.34 (m, 2H, HC_{44}), 3.29 (s, 1H, HC_{11}), 2.83 (t, *J* = 7.1 Hz, 2H, HC_{35}), 2.71 (m, 2H, HC_{45}), 2.31 (s, 3H, HC_{24}), 1.89 (m, 4H, HC_{36} + HC_{37}), 1.47 (t, *J* = 7.1 Hz, 3H, HC_{26}).

^{13}C NMR (151 MHz, CD_2Cl_2) of *E*-**48** isomer δ 172.9 (C_{46}), 166.1 (C_{28}), 153.6 (C_{40}), 150.4 (C_{30}), 149.0 (C_{43}), 147.9 (C_1), 147.8 (C_{34}), 145.8 (C_{12}), 145.6 (C_{29}), 143.4 (C_{17}), 141.0 (C_{18}), 140.7 (C_8), 139.6 (C_{22}), 137.8 (C_{33}), 137.6 (C_9 + C_{39}), 137.3 (C_{32}), 136.2 (C_{14}), 135.1 (C_{19}), 132.0 (C_3), 131.2 (C_{16} + C_{20}), 128.8 (C_{23}), 128.5 (C_4), 127.0 (C_{21}), 125.5 (C_2), 124.9 (C_5 + C_{31}), 123.6 (C_{41} + C_{42}), 122.3 (C_6), 120.9 (C_{15}), 119.3 (C_{13}), 113.3 (C_7), 84.1 (C_{10}), 78.7 (C_{11}), 68.9 (C_{38}), 65.3 (C_{25}), 61.8 (C_{27}), 52.1 (C_{47}), 35.1 (C_{45}), 33.2 (C_{35}), 29.3 (C_{37}), 27.9 (C_{36}), 26.3 (C_{44}), 20.5 (C_{24}), 14.8 (C_{26}).

HRMS (HESI$^+$): *m/z* = 875.3307 $[M+H]^+$ (calcd. 875.3334 for $C_{49}H_{47}N_8O_6S$).

Synthesis of 49



***EZ*-49**

Compound **48** (87 mg, 0.099 mmol) was dissolved in dry CH_2Cl_2 (1.5 mL) and $NaHCO_3$ (18 mg, 0.21 mmol, 2.2 eq.) was added. The reaction mixture was flushed with argon and cooled to 0 °C. To the mixture was added Dess-Martin periodinane (46 mg, 0.11 mmol, 1.1 eq.) and the mixture was stirred at r.t. for 2 h. Another portion of Dess-Martin periodinane (12 mg, 0.028 mmol, 0.3 eq.) was added and the reaction was continued for 1 h. The reaction mixture was quenched with saturated aqueous $Na_2S_2O_3$ solution (4 mL) and saturated aqueous $NaHCO_3$ solution (4 mL) and stirred vigorously for 1 h. The mixture was extracted with chloroform : *i*-PrOH (3:1, *v*/*v*, 4 × 5 mL). The organic layer was dried over $MgSO_4$ and

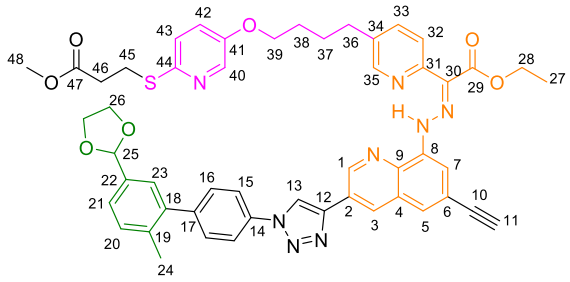
concentrated *in vacuo*. The residue was purified by PTLC (SiO_2, CH_2Cl_2/EtOAc, 12%) to give compound **49** as a brown paste (79 mg, 0.09 mmol, 91%).

^{1}H NMR (600 MHz, CD_2Cl_2) ~0.8:0.2 ratio of *E*/*Z*; *E*-**49** isomer δ 15.59 (s, 1H, NH), 10.03 (s, 1H, HC_{25}), 9.50 (d, J = 2.1 Hz, 1H, HC_1), 8.79 (d, J = 2.1 Hz, 1H, HC_{34}), 8.64 (d, J = 2.1 Hz, 1H, HC_3), 8.56 (s, 1H, HC_{13}), 8.16 (d, J = 2.7 Hz, 1H, HC_{39}), 8.10 (d, J = 8.4 Hz, 1H, HC_{31}), 7.95 (m, 3H, HC_7 + HC_{15}), 7.83 (m, 1H, HC_{21}), 7.80 (m, 1H, HC_{23}), 7.75 (m, 1H, HC_{32}), 7.69 (d, J = 1.6 Hz, 1H, HC_5), 7.59 (m, 2H, HC_{16}), 7.52 (d, J = 7.9 Hz, 1H, HC_{20}), 7.12 (m, 2H, HC_{41} + HC_{42}), 4.44 (q, J = 7.1 Hz, 2H, HC_{27}), 4.05 (m, 2H, HC_{38}), 3.64 (s, 3H, HC_{47}), 3.32 (t, J = 7.2 Hz, 2H, HC_{44}), 3.29 (s, 1H, HC_{11}), 2.83 (t, J = 7.1 Hz, 2H, HC_{35}), 2.70 (t, J = 7.2 Hz, 2H, HC_{45}), 2.41 (s, 3H, HC_{24}), 1.90 (m, 4H, HC_{36} + HC_{37}), 1.47 (t, J = 7.1 Hz, 3H, HC_{26}).

^{13}C NMR (151 MHz, CD_2Cl_2) of *E*-**49** isomer δ 192.2 (C_{25}), 172.9 (C_{46}), 166.1 (C_{28}), 153.6 (C_{40}), 150.4 (C_{30}), 149.0 (C_{43}), 147.9 (C_1), 147.8 (C_{34}), 145.9 (C_{12}), 143.6 (C_{19}), 142.1 (C_{17}), 141.8 (C_{18}), 140.7 (C_8), 137.9 (C_{33}), 137.7 (C_9), 137.6 (C_{29} + C_{39}), 137.3 (C_{32}), 136.7 (C_{14}), 135.3 (C_{22}), 132.0 (C_3), 131.9 (C_{20}), 131.3 (C_{16} + C_{23}), 129.5 (C_{21}), 128.6 (C_4), 125.4 (C_2), 124.9 (C_{31}), 124.8 (C_5), 123.6 (C_{41} + C_{42}), 122.4 (C_6), 121.1 (C_{15}), 119.3 (C_{13}), 113.4 (C_7), 84.1 (C_{10}), 78.6 (C_{11}), 68.9 (C_{38}), 61.8 (C_{27}), 52.1 (C_{47}), 35.1 (C_{45}), 33.2 (C_{35}), 29.3 (C_{37}), 27.9 (C_{36}), 26.3 (C_{44}), 21.3 (C_{24}), 14.8 (C_{26}).

HRMS (HESI$^+$): *m/z* = 873.3147 $[M+H]^+$ (calcd. 873.3177 for $C_{49}H_{45}N_8O_6S$).

Synthesis of 50



***EZ*-50**

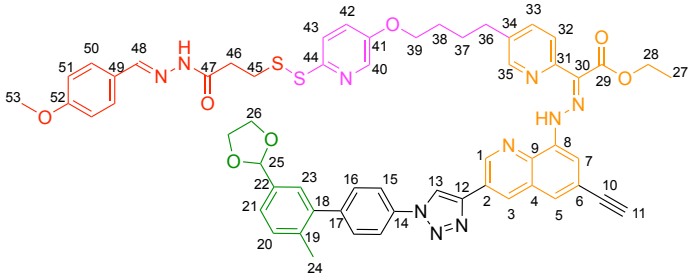
Compound **49** (67 mg, 0.077 mmol) was dissolved in dry THF (1.5 mL). To the mixture was added ethylene glycol (65 µL, 1.16 mmol, 15 eq.), triethyl orthoformate (51 µL, 0.31 mmol, 4 eq.) and *p*-TsOH·H_2O (37 mg, 0.19 mmol, 2.5 eq.) and the mixture was stirred under argon atmosphere at r.t. for 20 h. The reaction mixture was mixed with saturated aqueous $NaHCO_3$ solution (20 mL) extracted with chloroform : *i*-PrOH (3:1, *v*/*v*, 3 × 10 mL). The organic layer was washed with water (15 mL), dried over $MgSO_4$ and concentrated *in vacuo*. The residue was purified by flash column chromatography (SiO_2 pre-neutralized with Et_3N, hexane/CH_2Cl_2, 70% → 100% with 1% of Et_3N) to give compound **50** as a brown solid (55 mg, 0.060 mmol, 78%).

^{1}H NMR (600 MHz, CD_2Cl_2) ~0.4:0.6 ratio of *E*/*Z*; *E*-**50** isomer δ 15.58 (s, 1H, NH), 9.50 (d, J = 2.2 Hz, 1H, HC_1), 8.79 (d, J = 2.4 Hz, 1H, HC_{35}), 8.63 (d, J = 2.2 Hz, 1H, HC_3), 8.54 (s, 1H, HC_{13}), 8.16 (m, 1H, HC_{40}), 8.10 (d, J = 8.3 Hz, 1H, HC_{32}), 7.94 (d, J = 1.6 Hz, 1H, HC_7), 7.92 (m, 2H, HC_{15}), 7.74 (dd, J = 8.3, 2.4 Hz, 1H, HC_{33}), 7.70 (m, 1H, HC_5), 7.57 (m, 2H, HC_{16}), 7.41 (d, J = 8.7 Hz, 1H, HC_{21}), 7.39 (d, J = 2.0 Hz, 1H, HC_{23}), 7.34 (d, J = 7.8 Hz, 1H, HC_{20}), 7.11 (m, 2H, HC_{42} + HC_{43}), 5.79 (s, 1H, HC_{25}), 4.44 (q, J = 7.1 Hz, 2H, HC_{28}), 4.12 (m, 2H, HC_{26}), 4.02 (m, 4H, HC_{26} + HC_{39}), 3.63 (s, 3H, HC_{48}), 3.33 (t, J = 7.2 Hz, 2H, HC_{45}), 3.29 (s, 1H, HC_{11}), 2.83 (t, J = 7.0 Hz, 2H, HC_{36}), 2.71 (t, J = 7.2 Hz, 2H, HC_{46}), 2.33 (s, 3H, HC_{24}), 1.89 (m, 4H, HC_{37} + HC_{38}), 1.47 (t, J = 7.1 Hz, 3H, HC_{27}).

^{13}C NMR (151 MHz, CD_2Cl_2) of *E*-**50** isomer δ 172.9 (C_{47}), 166.1 (C_{29}), 153.6 (C_{41}), 150.4 (C_{31}), 149.0 (C_{44}), 147.8 (C_1 + C_{35}), 145.8 (C_{12}), 143.2 (C_{17}), 140.9 (C_{18}), 140.7 (C_8), 137.9 (C_{34}), 137.7 (C_9 + C_{40}), 137.3 (C_{33}), 137.1 (C_{19}), 136.6 (C_{22}), 136.3 (C_{14}), 132.0 (C_3), 131.3 (C_{16}), 131.1 (C_{20}), 128.6 (C_4), 128.4 (C_{23}), 126.6 (C_{21}), 125.5 (C_2), 124.9 (C_{32}), 124.8 (C_5), 123.6 (C_{42} + C_{43}), 122.3 (C_6), 120.9 (C_{15}), 119.4 (C_{13}), 113.3 (C_7), 104.1 (C_{25}), 84.1 (C_{10}), 78.6 (C_{11}), 68.9 (C_{39}), 65.9 (C_{26}), 61.8 (C_{28}), 52.1 (C_{48}), 35.1 (C_{46}), 33.2 (C_{36}), 29.3 (C_{38}), 27.9 (C_{37}), 26.3 (C_{45}), 20.6 (C_{24}), 14.8 (C_{27}). The signal of C_{30} could not be found.

HRMS (HESI$^+$): *m/z* = 917.3439 $[M+H]^+$ (calcd. 917.3404 for $C_{51}H_{49}N_8O_7S$).

Synthesis of 52



EZ-**52**

Compound **50** (21 mg, 0.023 mmol) was dissolved in dry degassed THF (1.5 mL) and flushed with argon. The solution was cooled to −78 °C and *t*-BuOK solution in THF (1 M, 95 µL, 0.095 mmol of *t*-BuOK, 4 eq.) was added dropwise. The reaction mixture was stirred at −78 °C for 2 h and warmed to r.t.. To the mixture was added saturated aqueous NH_4Cl solution (3 mL), saturated aqueous $NaHCO_3$ solution (3 mL), and TCEP·HCl (20 mg). The pH of the solution was ca. 7. The mixture was extracted with chloroform : *i*-PrOH (3:1, *v*/*v*, 4 × 5 mL). The organic layer was dried over $MgSO_4$ and concentrated *in vacuo* to obtain the crude as orange solid.

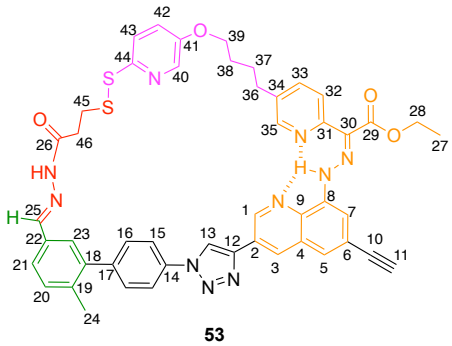
The crude and compound **12** (9.8 mg, 0.028 mmol, 1.2 eq.) were dissolved in degassed CH_2Cl_2 (4 mL), flushed with argon and stirred at r.t. for 17 h. The reaction mixture was purified with PTLC (SiO_2, CH_2Cl_2/EtOAc, 70% with 0.5% of Et_3N) and then with size exclusion column chromatography (Bio-Beads S-X1, CH_2Cl_2 eluent) to give compound **52** as a brown solid (12 mg, 0.011 mmol, 49%).

^{1}H NMR (600 MHz, CD_2Cl_2) ~0.9:0.1 ratio of *E*/*Z*; *E*-**52** isomer δ 15.59 (s, 1H, NH of the hydrazone switch), 9.50 (d, *J* = 2.2 Hz, 1H, HC_1), 8.78 (d, *J* = 2.3 Hz, 1H, HC_{35}), 8.63 (d, *J* = 2.2 Hz, 1H, HC_3), 8.60 (s, 1H, NH of the amide), 8.55 (s, 1H, HC_{13}), 8.14 (d, *J* = 2.9 Hz, 1H, HC_{40}), 8.10 (d, *J* = 8.4 Hz, 1H, HC_{32}), 7.94 (d, *J* = 1.6 Hz, 1H, HC_7), 7.91 (m, 2H, HC_{15}), 7.73 (m, 1H, HC_{33}), 7.69 (d, *J* = 1.6 Hz, 1H, HC_5), 7.63 (d, *J* = 8.8 Hz, 1H, HC_{43}), 7.60 (s, 1H, HC_{48}), 7.56 (m, 2H, HC_{16}), 7.49 (m, 2H, HC_{50}), 7.41 (dd, *J* = 7.8, 1.9 Hz, 1H, HC_{21}), 7.39 (d, *J* = 1.9 Hz, 1H, HC_{23}), 7.34 (d, *J* = 7.8 Hz, 1H, HC_{20}), 7.18 (dd, *J* = 8.8, 2.9 Hz, 1H, HC_{42}), 6.88 (m, 2H, HC_{51}), 5.79 (s, 1H, HC_{25}), 4.44 (q, *J* = 7.2 Hz, 2H, HC_{28}), 4.13 (m, 2H, HC_{26}), 4.02 (m, 2H, HC_{26}), 3.95 (m, 2H, HC_{39}), 3.79 (s, 3H, HC_{53}), 3.29 (s, 1H, HC_{11}), 3.12 (s, 4H, HC_{45} + HC_{46}), 2.80 (m, 2H, HC_{36}), 2.33 (s, 3H, HC_{24}), 1.85 (m, 4H, HC_{37} + HC_{38}), 1.47 (t, *J* = 7.2 Hz, 3H, HC_{27}).

^{13}C NMR (151 MHz, CD_2Cl_2) of *E*-**52** isomer δ 172.9 (C_{47}), 166.1 (C_{29}), 161.9 (C_{52}), 154.4 (C_{41}), 151.0 (C_{44}), 150.4 (C_{31}), 147.9 (C_1), 147.8 (C_{35}), 145.8 (C_{12}), 143.3 (C_{48}), 143.2 (C_{17}), 140.9 (C_{18}), 140.7 (C_8), 137.9 (C_9 + C_{40}), 137.7 (C_{34}), 137.3 (C_{33}), 137.1 (C_{19}), 136.6 (C_{22}), 136.3 (C_{14}), 132.0 (C_3), 131.3 (C_{16}), 131.1 (C_{20}), 129.1 (C_{50}), 128.6 (C_4), 128.4 (C_{23}), 126.9 (C_{49}), 126.6 (C_{21}), 125.5 (C_2), 124.9 (C_{32}), 124.8 (C_5), 123.6 (C_{42}), 122.4 (C_6), 121.8 (C_{43}), 120.9 (C_{15}), 119.4 (C_{13}), 114.7 (C_{51}), 113.3 (C_7), 104.1 (C_{25}), 84.1 (C_{10}), 78.6 (C_{11}), 68.8 (C_{39}), 66.0 (C_{26}), 61.8 (C_{28}), 55.9 (C_{53}), 34.4 (C_{45}), 33.1 (C_{36} + C_{46}), 29.2 (C_{38}), 27.8 (C_{37}), 20.7 (C_{24}), 14.8 (C_{27}). The signal of C_{30} was obscured.

HRMS (HESI$^+$): *m/z* = 1089.3505 $[M+Na]^+$ (calcd. 1089.3511 for $C_{58}H_{54}N_{10}O_7S_2Na$).

Synthesis of 53



53

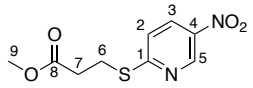
Compound **52** (13 mg, 12.2 µmol, 1 eq.) was dissolved in chloroform (12 mL) and TFA (65 µL, 0.85 mmol, 70 eq.) was added. The reaction mixture was stirred at r.t. for 3 h and quenched with Et_3N (250 µL, 1.8 mmol, 150 eq.). The organic layer was washed with water (5 mL), saturated aqueous $NaHCO_3$ solution (2 × 5 mL), water (5 mL) and concentrated *in vacuo*. The crude product was purified by size exclusion column chromatography (Bio-Beads S-X1, CH_2Cl_2 eluent) to give compound **53** as a yellow solid (6 mg, 6.8 µmol, 56%).

^{1}H NMR (600 MHz, $(CD_3)_2SO$) δ 15.33 (s, 1H, NH of the hydrazone switch), 11.48 (s, 1H, NH of the amide), 9.66 (s, 1H, HC_{13}), 9.41 (d, J = 2.1 Hz, 1H, HC_1), 9.02 (d, J = 2.1 Hz, 1H, HC_3), 8.75 (s, 1H, HC_{35}), 8.36 (d, J = 2.9 Hz, 1H, HC_{40}), 8.05 (s, 1H, HC_{25}), 8.04–8.00 (m, 3H, HC_{15} + HC_{32}), 7.97 (m, 1H, HC_{33}), 7.95 (d, J = 1.7 Hz, 1H, HC_7), 7.85 (s, 1H, HC_{23}), 7.74–7.69 (m, 4H, HC_5 + HC_{16} + HC_{43}), 7.54 (dd, J = 8.8, 3.0 Hz, 1H, HC_{42}), 7.45–7.40 (m, 2H, HC_{20} + HC_{21}), 4.44 (s, 1H, HC_{11}), 4.39 (q, J = 7.1 Hz, 2H, HC_{28}), 4.19 (m, 2H, HC_{39}), 3.12 (m, 2H, HC_{46}), 3.08 (m, 2H, HC_{45}), 2.82 (m, 2H, HC_{36}), 2.36 (s, 3H, HC_{24}), 1.89 (br, 4H, HC_{37} + HC_{38}), 1.39 (t, J = 7.1 Hz, 3H, HC_{27}).

^{13}C NMR (151 MHz, $(CD_3)_2SO$) δ 172.1, 164.8, 153.9, 149.1, 148.9, 148.2, 146.7, 144.2, 142.6, 141.5, 140.5, 139.3, 138.1, 137.8, 137.7, 137.0, 136.1, 135.6, 132.2, 131.1, 130.6, 130.0, 129.3, 128.0, 125.2, 124.6, 124.5, 124.4, 123.5, 121.3, 121.2, 120.7, 111.5, 83.5, 82.0, 67.8, 61.0, 32.5, 31.7, 30.5, 28.1, 26.8, 20.3, 14.3.

Attempts to obtain HRMS led to decomposition of the compound.

Synthesis of 59



59

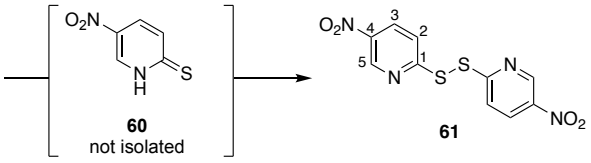
2-Bromo-5-nitropyridine (1.00 g, 4.93 mmol, 1 eq.), $Pd_2(dba)_3$ (460 mg, 0.50 mmol, 0.1 eq.) and Xantphos (338 mg, 0.58 mmol, 0.12 eq.) were dissolved in dry degassed dioxane (50 mL) and stirred at r.t. under nitrogen atmosphere for 1 h. In another flask, DIPEA (2.2 mL, 12.6 mmol, 2.6 eq.) and methyl 3-mercaptopropionate (0.7 mL, 6.35 mmol, 1.3 eq.) were dissolved in dry degassed dioxane (15 mL). The solution was flushed with nitrogen gas and stirred at r.t. for 45 min. The two solutions were mixed together, flushed with argon gas and stirred at 100 °C for 65 h. The reaction mixture was cooled to r.t. and brine (100 mL) was added. The mixture was extracted with CH_2Cl_2 (3 × 60 mL). The organic layer was combined, dried with $MgSO_4$ and concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO_2, Petroleum ether/EtOAc, 2% → 10%) to give **59** (1.07 g, 4.42 mmol, 90%) as a pale yellow oil.

^{1}H NMR (600 MHz, $CDCl_3$) δ 9.24 (d, J = 2.5 Hz, 1H, HC_5), 8.22 (dd, J = 8.9, 2.7 Hz, 1H, HC_3), 7.28 (d, J = 8.9 Hz, 1H, HC_2), 3.72 (s, 3H, HC_9), 3.51 (t, J = 7.0 Hz, 2H, HC_6), 2.82 (t, J = 7.0 Hz, 2H, HC_7).

^{13}C NMR (151 MHz, $CDCl_3$) δ 172.3, 167.0, 145.2, 130.6, 121.8, 52.1, 34.0, 25.5.

HRMS (ASAP$^+$): m/z = 243.0427 $[M+H]^+$ (calcd. 243.0434 for $C_9H_{11}N_2O_4S$).

Synthesis of 61 *via* 60



Compound **59** (1.07 g, 4.42 mmol) was dissolved in dry THF (30 mL) and flushed with argon. The solution was cooled to −78 °C and *t*-BuOK solution in THF (1 M, 8.5 mL, 8.5 mmol of *t*-BuOK, 1.9 eq.) was added dropwise. The reaction mixture was stirred at −78 °C for 1 h and warmed to r.t.. To the mixture was added saturated aqueous NH_4Cl solution (50 mL), brine (50 mL), and chloroform (50 mL). TFA was added until the pH of the water layer becomes 1. The mixture was extracted with chloroform (2 × 50 mL). The organic layer was dried over $MgSO_4$ and concentrated *in vacuo* to obtain the crude **60** (1.06 g) as orange solid.

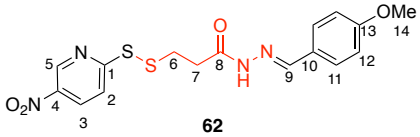
The crude was dissolved in THF (10 mL) and Et_3N (0.7 mL, 5 mmol, 1.1 eq.) was added. Solution of I_2 in THF was added dropwise until the yellow colour of the solution disappeared. The resulting mixture was directly dry loaded to silica gel and purified by flash column chromatography (SiO_2, Petroleum ether/EtOAc, 10% → 20%) to give **61** (0.60 g, 1.93 mmol, 87%) as off-white solid.

^{1}H NMR (600 MHz, $CDCl_3$) δ 9.29 (d, *J* = 2.6 Hz, 1H, HC_5), 8.40 (dd, *J* = 8.9, 2.6 Hz, 1H, HC_3), 7.73 (d, *J* = 8.9 Hz, 1H, HC_2).

^{13}C NMR (151 MHz, $CDCl_3$) δ 165.1, 145.5, 142.8, 132.3, 120.0.

HRMS (ASAP$^+$): *m/z* = 310.9895 $[M+H]^+$ (calcd. 310.9903 for $C_{10}H_7N_4O_4S_2$).

Synthesis of 62



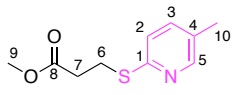
Compound **61** (600 mg, 1.93 mmol, 2.9 eq.) was dissolved in dry degassed THF (20 mL) and compound **20** (160 mg, 0.67 mmol, 1 eq.) was added to the solution portion-wise. The mixture was stirred at r.t. under argon atmosphere for 1 h and concentrated *in vacuo*. The residue was purified by flash column chromatography (SiO_2, Petroleum ether/EtOAc, 20% → 60%) to give **62** (250 mg, 0.64 mmol, 95%) as a yellow solid.

^{1}H NMR (600 MHz, $CDCl_3$) δ 9.21 (d, J = 2.6 Hz, 1H, HC_5), 8.81 (s, 1H, NH), 8.34 (dd, J = 8.9, 2.6 Hz, 1H, HC_3), 7.95 (d, J = 8.8 Hz, 1H, HC_2), 7.63 (s, 1H, HC_9), 7.48 (m, 2H, HC_{11}), 6.90 (m, 2H, HC_{12}), 3.86 (s, 3H, HC_{14}), 3.23 (m, 2H, HC_6), 3.17 (m, 2H, HC_7).

^{13}C NMR (151 MHz, $CDCl_3$) δ 172.5, 169.1, 161.7, 145.2, 143.8, 142.1, 131.8, 128.8, 125.9, 119.4, 114.5, 55.6, 33.8, 32.4.

HRMS (ASAP$^+$): m/z = 415.0502 $[M+Na]^+$ (calcd. 415.0505 for $C_{16}H_{16}N_4O_4S_2Na$).

Synthesis of 63



63

2-Bromo-5-methylpyridine (360 mg, 2.08 mmol, 1 eq.), $Pd_2(dba)_3$ (183 mg, 0.20 mmol, 0.1 eq.) and Xantphos (135 mg, 0.23 mmol, 0.12 eq.) were dissolved in dry degassed dioxane (24 mL) and stirred at r.t. under argon atmosphere for 1.5 h. In another flask, DIPEA (0.88 mL, 5.04 mmol, 2.5 eq.) and methyl 3-mercaptopropionate (0.28 mL, 2.54 mmol, 1.2 eq.) were dissolved in dry degassed dioxane (5 mL). The solution was flushed with argon gas and stirred at r.t. for 2 h. The two solutions were mixed together, flushed with argon gas and stirred at 110 °C for 19 h. The reaction mixture was cooled to r.t. and concentrated *in vacuo*. To the residue was added brine (100 mL) and the mixture was extracted with CH_2Cl_2 (3 × 60 mL). The organic layer was combined, dried with $MgSO_4$ and concentrated *in vacuo*. The

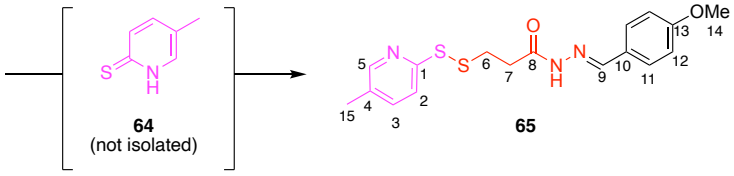
crude product was purified by flash column chromatography (SiO_2, Petroleum ether/CH_2Cl_2, 80% → 100%) to give **63** (440 mg, 2.08 mmol, quant.) as a brown oil.

^{1}H NMR (600 MHz, $CDCl_3$) δ 8.26 (d, J = 2.3 Hz, 1H, HC_5), 7.30 (dd, J = 8.2, 2.3 Hz, 1H, HC_3), 7.07 (d, J = 8.1 Hz, 1H, HC_2), 3.70 (s, 3H, HC_9), 3.40 (t, J = 7.1 Hz, 2H, HC_6), 2.77 (t, J = 7.1 Hz, 2H, HC_7), 2.26 (s, 3H, HC_{10}).

^{13}C NMR (151 MHz, $CDCl_3$) δ 172.8, 154.7, 149.9, 137.1, 129.2, 122.2, 51.9, 34.7, 25.2, 18.0.

HRMS (HESI$^+$): m/z = 234.0554 $[M+Na]^+$ (calcd. 234.0559 for $C_{10}H_{13}NO_2SNa$).

Synthesis of 65 *via* 64



Compound **63** (11 mg, 52 µmol, 2.5 eq.) was dissolved in dry degassed THF (0.8 mL) and flushed with argon. The solution was cooled to −78 °C and *t*-BuOK solution in THF (1 M, 90 µL, 90 µmol of *t*-BuOK, 4.5 eq.) was added dropwise. The reaction mixture was stirred at −78 °C for 1 h and warmed to r.t.. To the mixture was added saturated aqueous NH_4Cl solution (5 mL), TCEP·HCl (20 mg) and degassed chloroform (5 mL). The pH of the water layer was ca. 4. The mixture was extracted with degassed chloroform (2 × 5 mL). The organic layer was dried over $MgSO_4$ and concentrated *in vacuo* to obtain the crude **64** (10 mg) as yellow solid.

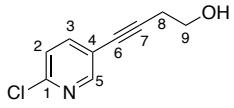
Compound **62** (8.0 mg, 20 µmol, 1 eq.) was dissolved in dry degassed THF (2 mL) and this solution was added to the crude. The reaction mixture was stirred at r.t. for 4 h under nitrogen atmosphere, and then concentrated *in vacuo*. The residue was purified by PTLC (SiO_2, hexane/EtOAc, 50%) twice to give **65** (6.0 mg, 17 µmol, 83%) as a pale brown oil.

^{1}H NMR (600 MHz, $CDCl_3$) ~0.8:0.2 ratio of *cis*/*trans* amide rotamers; *cis*-**65** isomer δ 8.66 (s, 1H, NH), 8.26 (d, *J* = 2.6 Hz, 1H, HC_5), 7.64 (d, *J* = 8.2 Hz, 1H, HC_2), 7.62 (s, 1H, HC_9), 7.52 (m, 2H, HC_{11}), 7.42 (dd, *J* = 8.3, 2.6 Hz, 1H, HC_3), 6.91 (m, 2H, HC_{12}), 3.85 (s, 3H, HC_{14}), 3.22 – 3.13 (m, 4H, HC_6 + HC_7), 2.22 (s, 3H, HC_{15}).

^{13}C NMR (151 MHz, $CDCl_3$) of *cis*-**65** isomer δ 173.0, 161.5, 157.0, 150.0, 143.3, 138.0, 130.5, 128.9, 126.2, 119.7, 114.4, 55.6, 33.6, 32.7, 18.0.

HRMS (HESI$^+$): *m/z* = 384.0805 $[M+Na]^+$ (calcd. 384.0811 for $C_{17}H_{19}N_3O_2S_2Na$).

Synthesis of 68



68

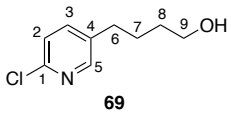
$Pd(PPh_3)_4$ (2.43 g, 2.1 mmol, 0.18 eq.), CuI (0.8 g, 4.2 mmol, 0.36 eq.) and 3-butyn-1-ol (0.9 mL, 11.9 mmol, 1.02 eq.) were mixed in dry degassed THF/Et_3N (1:1, *v*/*v*, 24 mL) and the mixture was stirred at r.t. under nitrogen atmosphere. After 40 min, solution of 5-bromo-2-chloropyridine (2.27 g, 11.7 mmol, 1 eq.) in dry degassed THF/Et_3N (1:1, *v*/*v*, 24 mL) was added and the reaction mixture was stirred at 70 °C under argon atmosphere for 26 h. The mixture was cooled to r.t. and diluted with EtOAc (150 mL). The organic layer was washed with saturated aqueous NH_4Cl solution (2 × 50 mL), aqueous Na_4EDTA solution (1 M, 50 mL) and brine (50 mL). The water layer was combined and extracted with EtOAc (2 × 20 mL). The organic layer was dried with $MgSO_4$ and concentrated *in vacuo.* The crude product was purified by flash column chromatography (SiO_2, CH_2Cl_2/EtOAc, 15% → 25%) to give **68** (1.58 g, 8.70 mmol, 74%) as brown solid.

^{1}H NMR (600 MHz, $CDCl_3$) δ 8.42 (d, *J* = 2.4 Hz, 1H, HC_5), 7.64 (dd, *J* = 8.3, 2.4 Hz, 1H, HC_3), 7.27 (d, *J* = 8.3 Hz, 1H, HC_2), 3.84 (q, *J* = 6.2 Hz, 2H, HC_9), 2.71 (t, *J* = 6.3 Hz, 2H, HC_8), 1.77 (t, *J* = 6.2 Hz, 1H, OH).

^{13}C NMR (151 MHz, $CDCl_3$) δ 152.4, 150.3, 141.3, 124.0, 119.6, 91.6, 78.0, 61.1, 24.0.

HRMS (ASAP$^+$): *m/z* = 182.0362 $[M+H]^+$ (calcd. 182.0367 for C_9H_9NOCl).

Synthesis of 69



69

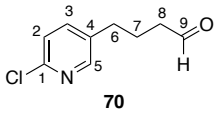
To the solution of **68** (500 mg, 2.75 mmol) in EtOH (25 mL), 20% (*w*/*w*) $Pd(OH)_2$ on C (96 mg, 0.14 mmol of $Pd(OH)_2$, 0.05 eq.) was added. The reaction mixture was stirred at r.t. under hydrogen atmosphere for 25 h. The catalyst was removed by filtration through a pad of Celite, and the filtrate was concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO_2, CH_2Cl_2/Et_2O, 50%) to give **69** (270 mg, 1.45 mmol, 53%) as a colourless oil.

^{1}H NMR (600 MHz, $CDCl_3$) δ 8.21 (d, *J* = 2.5 Hz, 1H, HC_5), 7.47 (dd, *J* = 8.1, 2.5 Hz, 1H, HC_3), 7.24 (d, *J* = 8.1 Hz, 1H, HC_2), 3.68 (m, 2H, HC_9), 2.63 (t, *J* = 7.7 Hz, 2H, HC_6), 1.70 (m, 2H, HC_7), 1.60 (m, 2H, HC_8), 1.34 (br, 1H, OH).

^{13}C NMR (151 MHz, $CDCl_3$) δ 149.7, 149.1, 138.9, 136.6, 124.0, 62.6, 32.2, 32.1, 27.4.

HRMS (HESI$^+$): *m/z* = 186.0680 $[M+H]^+$ (calcd. 186.0680 for $C_9H_{13}NOCl$).

Synthesis of 70



70

Compound **69** (260 mg, 1.40 mmol, 1 eq.) and $NaHCO_3$ (237 mg, 2.80 mmol, 2 eq.) were mixed with dry CH_2Cl_2 (18 mL) and flushed with nitrogen. The mixture was cooled to 0 °C and DMP reagent (669 mg, 1.58 mmol, 1.1 eq.) was added. The reaction mixture was stirred at r.t. for 4 h and quenched with saturated aqueous $Na_2S_2O_3$ solution (15 mL) and saturated

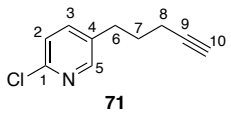
aqueous $NaHCO_3$ solution (15 mL). The mixture was stirred further for 15 min and extracted with CH_2Cl_2 (3 × 20 mL). The organic layer was dried with $MgSO_4$ and concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO_2, CH_2Cl_2/Et_2O, 0% → 2%) to give **70** (250 mg, 1.36 mmol, 97%) as a pale yellow oil.

^1H NMR (600 MHz, $CDCl_3$) δ 9.78 (t, *J* = 1.3 Hz, 1H, HC_9), 8.21 (d, *J* = 2.5 Hz, 1H, HC_5), 7.49 (dd, *J* = 8.2, 2.5 Hz, 1H, HC_3), 7.26 (d, *J* = 8.2 Hz, 1H, HC_2), 2.64 (d, *J* = 7.4 Hz, 2H, HC_6), 2.50 (td, *J* = 7.2, 1.3 Hz, 2H, HC_8), 1.94 (m, 2H, HC_7).

^{13}C NMR (151 MHz, $CDCl_3$) δ 201.5, 149.7, 149.5, 138.9, 135.6, 124.2, 43.0, 31.5, 23.3.

HRMS ($HESI^+$): *m/z* = 184.0520 $[M+H]^+$ (calcd. 184.0524 for $C_9H_{11}NOCl$).

Synthesis of 71



71

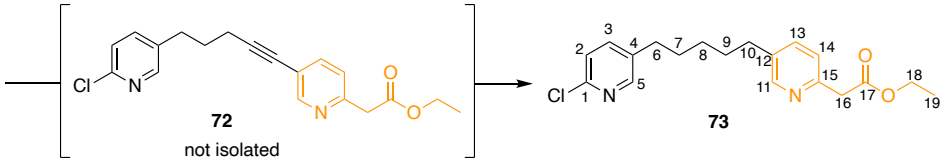
Compound **70** (240 mg, 1.31 mmol, 1 eq.) and K_2CO_3 (371 mg, 2.68 mmol, 2.05 eq.) were mixed with MeOH (20 mL) and flushed with nitrogen. Ohira-Bestmann reagent in MeCN (10%, *w*/*w*, 3.5 mL, 1.72 mmol, 1.3 eq. of the reagent) was added and the reaction mixture was stirred at r.t. for 24 h. The mixture was diluted with Et_2O (50 mL) and washed with saturated aqueous $NaHCO_3$ solution (30 mL) and water (50 mL). The organic layer was dried with $MgSO_4$ and concentrated *in vacuo* to give **71** (210 mg, 1.17 mmol, 89%) as a pale brown oil.

^1H NMR (600 MHz, $CDCl_3$) δ 8.24 (d, *J* = 2.5 Hz, 1H, HC_5), 7.49 (dd, *J* = 8.1, 2.5 Hz, 1H, HC_3), 7.25 (d, *J* = 8.1 Hz, 1H, HC_2), 2.74 (t, *J* = 7.7 Hz, 2H, HC_6), 2.22 (td, *J* = 6.9, 2.6 Hz, 2H, HC_8), 2.02 (t, *J* = 2.6 Hz, 1H, HC_{10}), 1.83 (m, 2H, HC_7).

^{13}C NMR (151 MHz, $CDCl_3$) δ 149.8, 149.4, 139.0, 135.8, 124.1, 83.4, 69.5, 30.9, 29.6, 17.8.

HRMS (APCI$^+$): *m/z* = 180.0569 $[M+H]^+$ (calcd. 180.0575 for $C_{10}H_{11}NCl$).

Synthesis of 73 *via* 72



Compound **71** (200 mg, 1.11 mmol, 1 eq.), CuI (78 mg, 0.41 mmol, 0.37 eq.) and $Pd(PPh_3)_4$ (245 mg, 0.21 mmol, 0.19 eq.) were combined in a degassed mixture of dry THF:Et_3N (1:1, *v*/*v*, 4.5 mL) under argon atmosphere. The mixture was stirred at r.t. for 1 h and then compound **31** (299 mg, 1.23 mmol, 1.1 eq.) was added. The mixture was stirred at 70 °C for 24 h and then concentrated *in vacuo*. The residue was combined with EtOAc (15 mL) and washed with saturated aqueous NH_4Cl solution (10 mL), aqueous Na_4EDTA solution (1 M, 10 mL) and brine (10 mL). The organic layer was dried with $MgSO_4$ and concentrated *in vacuo*. The crude product was partially purified by flash column chromatography (SiO_2, CH_2Cl_2/EtOAc, 5% → 10%) to give **72** with impurity (280 mg, assume 0.82 mmol) as a brown oil. The compound was used in the next step without further purification.

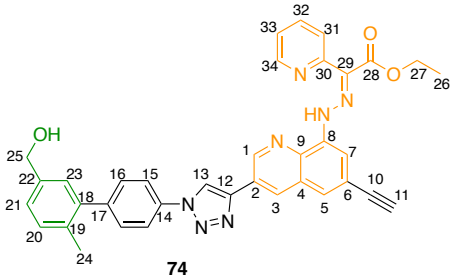
To the solution of **72** with impurity (99 mg, assume 0.29 mmol) in EtOH (3 mL), 20% (*w*/*w*) $Pd(OH)_2$ on C (41 mg, 0.058 mmol of $Pd(OH)_2$, 0.2 eq.) was added. The reaction mixture was stirred vigorously at r.t. under hydrogen atmosphere for 8 h. The catalyst was removed by filtration through a pad of Celite, and the filtrate was concentrated *in vacuo*. The crude product was purified by PTLC (SiO_2, CH_2Cl_2/Et_2O, 30%) to give **73** (52 mg, 0.15 mmol, 38% in two steps) as a colourless oil.

^{1}H NMR (600 MHz, $CDCl_3$) δ 8.36 (d, *J* = 2.3 Hz, 1H, HC_{11}), 8.19 (d, *J* = 2.5 Hz, 1H, HC_5), 7.46–7.42 (m, 2H, HC_3 + HC_{13}), 7.25–7.18 (m, 2H, HC_2 + HC_{14}), 4.18 (q, *J* = 7.2 Hz, 2H, HC_{18}), 3.81 (s, 2H, HC_{16}), 2.58 (m, 4H, HC_6 + HC_{10}), 1.64 (m, 4H, HC_7 + HC_9), 1.37 (m, 2H, HC_8), 1.26 (t, *J* = 7.2 Hz, 3H, HC_{19}).

^{13}C NMR (151 MHz, $CDCl_3$) δ 171.1, 152.1, 149.7, 149.6, 149.1, 138.9, 136.7, 136.6, 136.0, 124.0, 123.6, 61.2, 43.7, 32.7, 32.2, 31.0, 28.7, 14.3.

HRMS (HESI$^+$): *m/z* = 369.1339 $[M+Na]^+$ (calcd. 369.1340 for $C_{19}H_{23}N_2O_2ClNa$).

Synthesis of 74



74

Compound **47** (89 mg, 0.18 mmol, 1 eq.) and ethyl 2-oxo-2-(pyridin-2-yl)acetate (32 mg, 0.18 mmol, 1 eq.) were combined in a vial and dissolved in CH_2Cl_2 (1.8 mL), water (1.8 mL) and EtOH (3.6 mL). To the solution was added TFA (50 µL) and the reaction mixture was stirred at 40 °C for 18 h. The mixture was then diluted with CH_2Cl_2 (20 mL) and washed with saturated aqueous $NaHCO_3$ solution (2 × 10 mL), water (10 mL) and brine (10 mL). The organic layer was dried with $MgSO_4$ and concentrated *in vacuo.* The crude product was purified by size exclusion column chromatography (Bio-Beads S-X1, CH_2Cl_2 eluent) to give compound **74** as a brown solid (82 mg, 0.14 mmol, 75%).

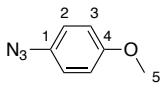
^{1}H NMR (600 MHz, CD_2Cl_2) ~0.8:0.2 ratio of *E*/*Z*; *E*-**74** isomer δ 15.57 (s, 1H, NH), 9.47 (d, *J* = 2.1 Hz, 1H, HC_1), 8.92 (m, 1H, HC_{34}), 8.64 (d, *J* = 2.1 Hz, 1H, HC_3), 8.52 (s, 1H, HC_{13}), 8.16 (d, *J* = 8.4 Hz, 1H, HC_{31}), 7.95 (d, *J* = 1.6 Hz, 1H, HC_7), 7.92–7.88 (m, 3H, HC_{15} + HC_{32}), 7.70 (d, *J* = 1.6 Hz, 1H, HC_5), 7.56 (m, 2H, HC_{16}), 7.40 (ddt, *J* = 7.0, 4.0, 1.1 Hz, 1H, HC_{33}), 7.31 (m, 2H, HC_{20} + HC_{21}), 7.29 (s, 1H, HC_{23}), 4.70 (s, 2H, HC_{25}), 4.44 (q, *J* = 7.1 Hz, 2H, HC_{27}), 3.29 (s, 1H, HC_{11}), 2.32 (s, 3H, HC_{24}), 1.47 (t, *J* = 7.1 Hz, 3H, HC_{26}).

^{13}C NMR (151 MHz, CD_2Cl_2) of *E*-**74** isomer δ 165.8, 152.4, 147.8, 147.7, 145.6, 143.2, 140.8, 140.4, 139.4, 137.7, 137.3, 136.0, 135.0, 131.8, 131.1 (3H), 129.9, 128.7, 128.4,

126.8, 125.3, 125.1, 124.8, 123.3, 122.1, 120.7 (2H), 119.1, 113.3, 83.9, 78.4, 65.1, 61.6, 20.3, 14.6.

HRMS (HESI$^+$): *m/z* = 630.2211 $[M+Na]^+$ (calcd. 630.2224 for $C_{36}H_{29}N_7O_3Na$).

Synthesis of 75



75

4-Methoxyaniline (100 mg, 0.81 mmol, 1 eq.) was dissolved in MeCN (10 mL) and the solution was cooled to −10 °C. To the solution was added *t*-BuONO (200 µL, 1.66 mmol, 2.05 eq.) and the mixture was stirred for 10 min. $TMSN_3$ (220 µL, 1.66 mmol, 2.05 eq.) was added and the reaction mixture was stirred at r.t. for 65 h. Solvent was removed *in vacuo* and the crude was purified by flash column chromatography (SiO_2, CH_2Cl_2) to give **75** (100 mg, 0.67 mmol, 83%) as a yellow oil.

^{1}H NMR (600 MHz, CD_2Cl_2) δ 6.96 (m, 2H, HC_2), 6.90 (m, 2H, HC_3), 3.78 (s, 3H, HC_5).

^{13}C NMR (151 MHz, CD_2Cl_2) δ 157.5, 132.6, 120.3, 115.4, 55.9.

Attempts to obtain HRMS led to decomposition of the compound.

2.11. References

1 Von Ballmoos, C., Wiedenmann, A. & Dimroth, P. Essentials for ATP synthesis by F_1F_0 ATP synthases. *Annu. Rev. Biochem.* **78**, 649–672 (2009).

2 Yonath, A. Hibernating bears, antibiotics, and the evolving ribosome (Nobel Lecture). *Angew. Chem. Int. Ed.* **49**, 4340–4354 (2010).

3 Vale, R. D. & Milligan, R. A. The way things move: looking under the hood of molecular motor proteins. *Science* **288**, 88–95 (2000).

4 Schliwa, M. & Woehlke, G. Molecular motors. *Nature* **422**, 759–765 (2003).

5 Sowa, Y. & Berry, R. M. Bacterial flagellar motor. *Q. Rev. Biophys.* **41**, 103–132 (2008).

6 Maier, T., Leibundgut, M. & Ban, N. The crystal structure of a mammalian fatty acid synthase. *Science* **321**, 1315–1322 (2008).

7 Brignole, E. J., Smith, S. & Asturias, F. J. Conformational flexibility of metazoan fatty acid synthase enables catalysis. *Nat. Struct. Mol. Biol.* **16**, 190–197 (2009).

8 Chan, D. I. & Vogel, H. J. Current understanding of fatty acid biosynthesis and the acyl carrier protein. *Biochem. J.* **430**, 1–19 (2010).

9 Feynman, R. P. There's plenty of room at the bottom. *Eng. Sci.* **23**, 22–36 (1960).

10 Drexler, K. E. Molecular engineering: an approach to the development of general capabilities for molecular manipulation. *Proc. Natl. Acad. Sci. USA* **78**, 5275–5278 (1981).

11 Drexler, K. E. *Nanosystems: Molecular Machinery, Manufacturing, and Computation*. (Wiley, 1992).

12 Smalley, R. E. Of chemistry, love and nanobots. *Sci. Am.* **285**, 76–77 (2001).

13 Whitesides, G. M. The once and future nanomachine. *Sci. Am.* **285**, 78–83 (2001).

14 Jones, R. A. L. *Soft Machines: Nanotechnology and Life*. (Oxford Univ. Press, 2004).

15 Muraoka, T., Kinbara, K. & Aida, T. Mechanical twisting of a guest by a photoresponsive host. *Nature* **440**, 512–515 (2006).

16 Berná, J. *et al.* Macroscopic transport by synthetic molecular machines. *Nat. Mater.* **4**, 704–710 (2005).

17 Liu, Y. *et al.* Linear Artificial molecular muscles. *J. Am. Chem. Soc.* **127**, 9745–9759 (2005).

18 Eelkema, R. *et al.* Nanomotor rotates microscale objects. *Nature* **440**, 163 (2006).

19 Li, Q. *et al.* Macroscopic contraction of a gel induced by the integrated motion of light-driven molecular motors. *Nat. Nanotechnol.* **10**, 161–165 (2015).

20 Ding, B. & Seeman, N. C. Operation of a DNA robot arm inserted into a 2D DNA crystalline substrate. *Science* **314**, 1583–1585 (2006).

21 Gu, H., Chao, J., Xiao, S.-J. & Seeman, N. C. A proximity-based programmable DNA nanoscale assembly line. *Nature* **465**, 202–205 (2010).

22 Landge, S. M. & Aprahamian, I. A pH activated configurational rotary switch: controlling the *E*/*Z* isomerization in hydrazones. *J. Am. Chem. Soc.* **131**, 18269–18271 (2009).

23 Su, X. & Aprahamian, I. Switching around two axles: controlling the configuration and conformation of a hydrazone-based switch. *Org. Lett.* **13**, 30–33 (2010).

24 Kassem, S., Lee, A. T., Leigh, D. A., Markevicius, A. & Solà, J. Pick-up, transport and release of a molecular cargo using a small-molecule robotic arm. *Nat. Chem.* **8**, 138–143 (2016).

25 Kassem, S. *et al.* Stereodivergent synthesis with a programmable molecular machine. *Nature* **549**, 374–378 (2017).

26 Trankle, W. G. & Kopach, M. E. Green chemical synthesis of 2-benzenesulfonyl-pyridine and related derivatives. *Org. Process Res. Dev.* **11**, 913–917 (2007).
27 Li, H., Wang, H., Liu, Y. & Liu, Z. A benzoboroxole-functionalized monolithic column for the selective enrichment and separation of cis-diol containing biomolecules. *Chem. Commun.* **48**, 4115–4117 (2012).

Chapter 3.

A Catalysis-Driven Artificial Molecular Pump

Blank Page

Declaration

The research presented in this chapter has been published as a research article in a peer reviewed journal and the paper has been adapted only for making the format consistent within the thesis. The reference is as follows:

Amano, S., Fielden, S. D. P. & Leigh, D. A. A catalysis-driven artificial molecular pump. *Nature* **594**, 529–534 (2021).

The author of articles published by Springer Nature do not usually need to seek permission for re-use of their material as long as the journal is credited with initial publication. Ownership of copyright in in original research articles remains with the Author, and provided that, when reproducing the contribution or extracts from it or from the Supplementary Information, the Author acknowledges first and reference publication in the Journal, the Author retains the following non-exclusive rights:
To reproduce the contribution in whole or in part in any printed volume (book or thesis) of which they are the author(s).
(https://www.nature.com/nature-portfolio/reprints-and-permissions/permissions-requests)

Acknowledgements

The author's contribution to this work includes the design of the system, model studies of each step, synthesis of the machine molecule and autonomous operation. Dr. Stephen D. P. Fielden assisted with the design of the system and experiments. Prof. David A. Leigh directed the research. All authors contributed to analysing the results and assisting with the writing and proofreading of the manuscript and Supporting Information. All the authors of the paper are gratefully thanked for their contribution. Authorship order is alphabetical.

Synopsis

All biomolecular pumps are autonomous catalysts; they run continuously, drawing energy for the pumping process through their decomposition of a chemical fuel (generally ATP or GTP), and stop operating once the fuel is consumed. In contrast, all of the artificial chemically-driven molecular pumps reported to date require repetitious stepwise operations, either repeated sequential additions of chemical reagents or the varying of an electric potential during each cycle of pump operation. Here we report on an autonomous chemically-fueled molecular pump that, following initial addition of the fuel, continuously pumps crown ether macrocycles from bulk solution onto a molecular axle without the need for further external intervention. The result of the pump's action is dissipatively captured in the form of thermodynamically unstable [*n*]rotaxanes (characterized with *n* up to 4). The pump operation is experimentally demonstrated by progressively pumping up to three macrocycles onto the axle under conditions where stopper formation and removal occur continuously.

3.1. Introduction

All biological pumps are autonomous catalysts; they maintain the out-of-equilibrium conditions of the cell by harnessing the energy released from their catalytic decomposition of a chemical fuel[1–3]. A number of artificial molecular pumps have been reported to date[4], but they are all either fuelled by light[5–10] or require repetitive sequential additions of reagents or varying of an electric potential during each cycle to operate[11–16]. Here we describe an autonomous chemically fuelled information ratchet[17–20] that in the presence of fuel continuously pumps crown ether macrocycles from bulk solution onto a molecular axle without the need for further intervention. The mechanism uses the position of a crown ether on an axle both to promote barrier attachment behind it upon threading and to suppress subsequent barrier removal until the ring has migrated to a catchment region. Tuning the dynamics of both processes[20,21] enables the molecular machine[22–25] to pump macrocycles continuously from their lowest energy state in bulk solution to a higher energy state on the axle. The ratchet action is experimentally demonstrated by the progressive pumping of up to three macrocycles onto the axle from bulk solution under conditions where barrier formation and removal occur continuously. The out-of-equilibrium [*n*]rotaxanes (characterized with *n* up to 4) are maintained for as long as unreacted fuel is present, after which the rings slowly de-thread. The use of catalysis to drive artificial molecular pumps opens up new opportunities, insights and research directions at the interface of catalysis and molecular machinery.

3.2. Results and discussion

The structure and mode of operation of the catalysis-driven molecular pump, **1**, is shown in **Figure 1**. One end of the axle is permanently blocked by a bulky triarylmethine group

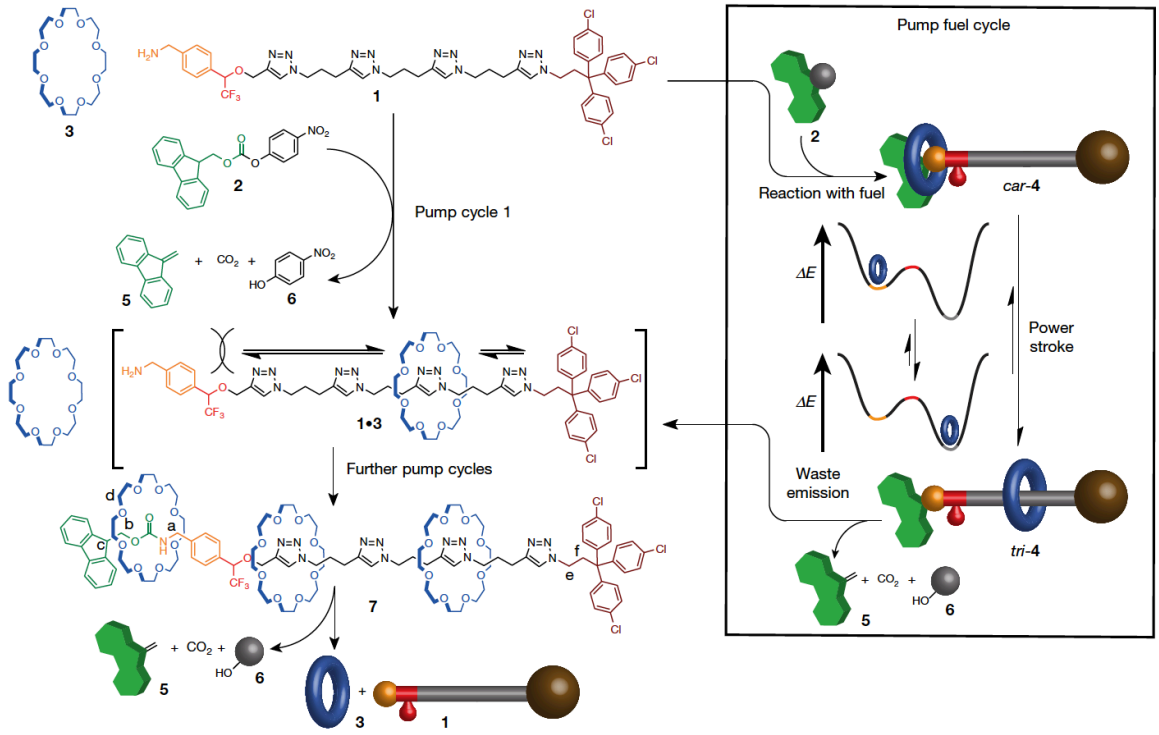


Figure 1. Structure and operation of a catalysis-driven artificial molecular pump. Continuous pumping of 24-crown-8 rings (**3**) onto a molecular axle (**1**) in the presence of fuel (**2**), causing the rings to be dissipatively captured in the form of [*n*]rotaxanes such as **7**. Typical operating conditions are **1** (1 equiv.), **3** (10 equiv.), *i*-Pr_2NH (15 equiv.), toluene (0.2 M with respect to **1**), **2** (either 10+ equiv. added at time zero, or added continuously via syringe pump at a rate of 0.7 equiv. per hour), at room temperature. The inset shows the pump fuel cycle. Kinetically gated addition of the fuel to the pump causes active template threading of the ring onto the axle, initially forming rotaxane co-conformer *car*-**4** (with the macrocycle positioned over the carbamate axle site). The rotaxane relaxes to the lower energy state, *tri*-**4**, through a power stroke (although in the formation of [4]rotaxane **7** from **1**•(**3**)$_2$ the equivalent step is not exergonic) of the ring past the trifluoromethyl 'speed bump' (red) onto the catchment region (grey part of the axle in the cartoon). Kinetically gated decomposition of the blocking group (green) to waste liberates the open end of the axle for further pumping while threaded rings remain kinetically trapped in the catchment region, completing the pumping cycle. equiv., equivalents.

(brown); the other end of the axle (the *N*-terminus) is open for threading when in the form of a benzyl amine group (orange). The axle includes a chain of triazole heterocycles linked by short, propyl (that is, –$(CH_2)_3$–), spacers. Triazoles interact only weakly with crown ethers[26,27], so the proportion of crown ethers transferred from bulk solution through the open end of the axle and onto the chain is vanishingly small at equilibrium. The pumping mechanism is based on a recently discovered[28] metal-free[29,30] active template rotaxane-forming reaction[31], in which addition of a primary amine to an electrophile is accelerated

through the cavity of a crown ether. Pump **1** reacts with Fmoc–$OC_6H_4NO_2$ (the fuel[32], **2**; where Fmoc = 9-fluorenylmethyloxycarbonyl) and 24-crown-8 (a crown ether, **3**; blue) causing active template threading of the ring, initially forming *car*-**4** (**Figure 1**; the italicized prefix indicates the position of the ring on the axle). The newly formed Fmoc–carbamate group (green) prevents dethreading of the ring back over the axle *N*-terminus. Adjacent to the benzyl amine group is a trifluoromethyl (CF_3) group (red), which is of sufficient size to substantially slow the passage of 24-crown-8 (a 'speed bump'[9]). The triazole groups on the axle should form more favourable interactions with the crown ether[26,27] than the sterically crowded benzylic carbamate in *car*-**4**, and so provide a 'catchment region' for rings threaded onto the axle. Accordingly, over time the crown ether in *car*-**4** slowly passes over the trifluoromethyl group to the catchment region, forming co-conformer *tri*-**4** ('co-conformers' are structures that differ in the relative positions of their components[33]). Rotaxane *tri*-**4** is of lower energy than *car*-**4**, but still of higher energy than the unthreaded components **1** and **3**. The presence of a base (diisopropylamine (*i*-Pr_2NH) in **Figure 1**) is necessary to complete the pumping cycle. The base deprotonates the Fmoc–carbamate group, triggering its fragmentation to dibenzofulvene (**5**), CO_2 and 4-nitrophenol ($NO_2C_6H_4OH$; **6**), liberating the amine *N*-terminus of the molecular machine for use in further pumping cycles. In this way, rings are captured from bulk solution, constrained within a localized region of the axle, and maintained in an out-of-equilibrium state by the continuous, autonomous, consumption of the chemical fuel.

The pump and fuel structure, in particularly the stereoelectronic characteristics of the CF_3 speedbump, the triazole groups of the catchment region, and the 4-nitrophenol leaving group of the fuel, were selected through a series of model and optimization studies (Section 3.4.5.). The pump operating conditions (see **Figure 1**) were chosen so that the Fmoc-removal reaction and the passage of the rings from the catchment region of the axle back over the

CF_3 group (resulting in dethreading when the pump *N*-terminus is not blocked) are both much slower than active template reaction of the axle benzylic amine group with unreacted Fmoc–$OC_6H_4NO_2$ fuel. This means that the *N*-terminus of **1•3** undergoes further active template reactions before dethreading occurs, pumping a second (and subsequently third) ring onto the axle. The molecular machine continues to pump rings onto the axle for as long as there is sufficient space on the chain and unreacted fuel present.

Pump **1** was synthesized in 13 steps (Section 3.4.2.) and operated successfully in three distinct ways: (i) stepwise, adding the fuel and base in separate steps, which facilitated isolation of the rotaxanes; (ii) autonomously, with all of the fuel added to the rings, pump and base at time zero; (iii) continuous fuel addition, which minimizes background decomposition of the fuel by the base. The pumping mechanism was confirmed by operating **1** in a stepwise manner. A single pumping cycle, other than removal of the Fmoc barrier, was performed by combining **1**, **2** and **3** in toluene under basic conditions (**Figure 1**), affording [2]rotaxane **4** in 61% yield after 10 min (the amount of rotaxane did not increase using longer reaction times), with concurrent formation of the Fmoc-derivatized unthreaded axle in 10% yield. This shows that >80% of the fuel molecules (**2**) that react with **1** result in threading. The formation of [2]rotaxane under conditions where the release of the blocking group (waste-forming step) is negligible shows that during the fuel addition step of the pumping mechanism, the pumping is kinetically gated (that is, threading accelerates Fmoc-carbamate formation of the axle).

The interconversion of *car*-**4** and *tri*-**4** is slow on the ^{1}H NMR (proton nuclear magnetic resonance spectroscopy) timescale at 295 K in $C_6D_5CD_3$ (**Figure 2c**), allowing the ratio of the [2]rotaxane co-conformers to be determined. Various axle protons are deshielded by the

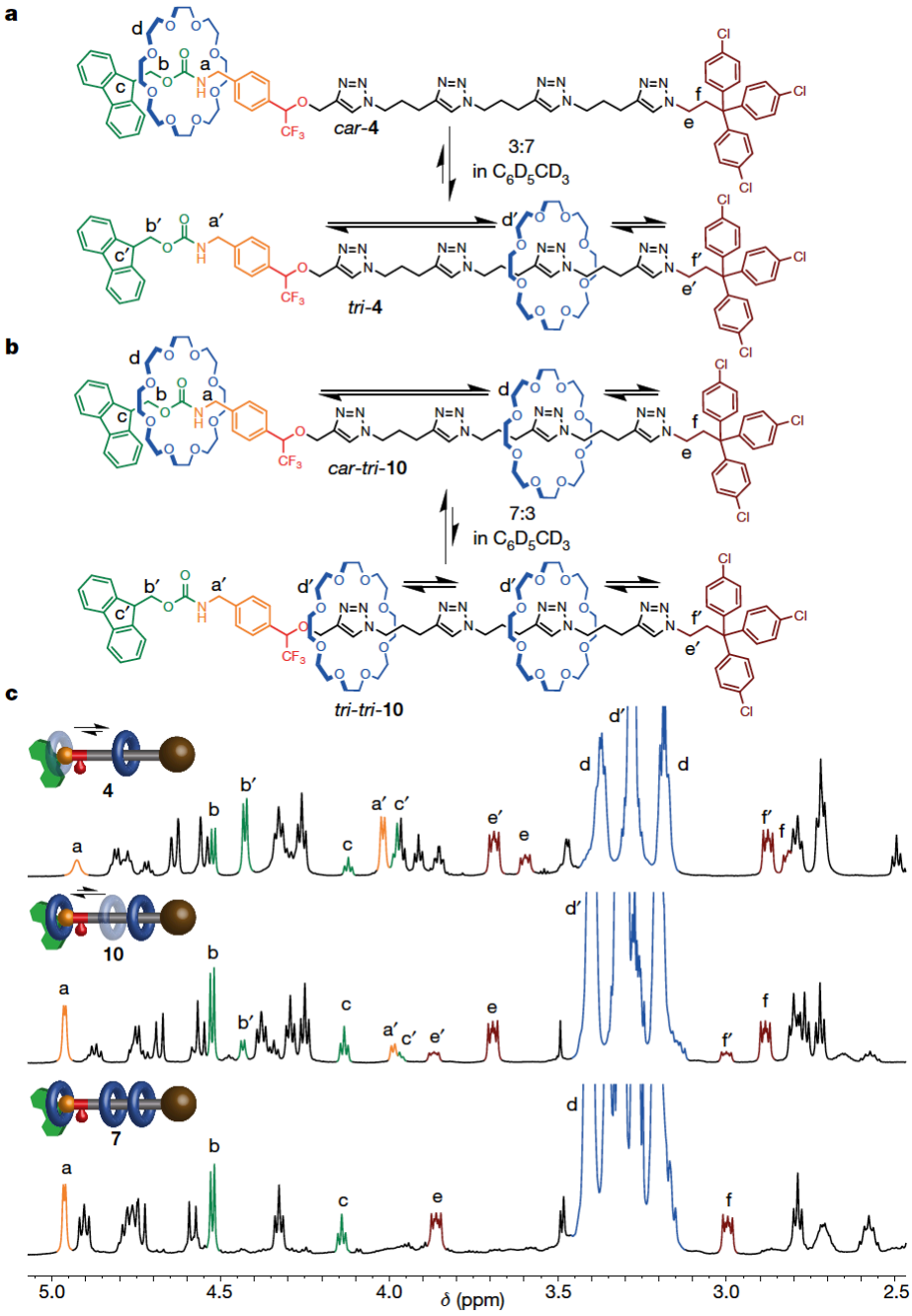


Figure 2. Macrocycle distribution in [*n*]rotaxane co-conformers. **a**, Macrocycle positions in [2]rotaxane co-conformers *car*-**4** and *tri*-**4** (295 K, $C_6D_5CD_3$). **b**, Macrocycle positions in [3]rotaxane co-conformers *car-tri*-**10** and *tri-tri*-**10** (295 K, $C_6D_5CD_3$). **c**, Partial ^{1}H NMR (600 MHz, 295 K, $C_6D_5CD_3$) spectra of **4**, **10** and [4]rotaxane **7**. Lettering relates protons in the chemical structures to the corresponding signals in the ^{1}H NMR spectra. Primes are used to distinguish the protons in *tri*-**4** and *tri-tri*-**10** from those in co-conformers *car*-**4** and *car-tri*-**10**.

presence of the macrocycle: in particular, H_a, H_b and H_c of *car*-**4** (labelling shown in **Figure 2a**), and $H_{e'}$ and $H_{f'}$ of *tri*-**4**. Signal integration gives a *car*-**4**:*tri*-**4** ratio (that is, local

equilibrium distribution of a single crown ether between *car*- and *tri*- sites) of 3:7 at 295 K in $C_6D_5CD_3$.

To complete the pump cycle, the Fmoc group needs to be removed from the *N*-terminus of **4** to regenerate the primary amine for another threading event. The position of the macrocycle on the axle affects the rate of Fmoc removal. When treated with *i*-Pr_2NH in toluene, the half-life of a model rotaxane, **8**, which has 24-crown-8 constrained on a short axle and forced into close proximity to the Fmoc group, was found to be ~3× longer than that of the corresponding non-interlocked thread **9** (**Figure 3a**). During the operation of pump **1**, the ring's steric inhibition of the attack of *i*-Pr_2NH means that the Fmoc group is removed more rapidly from the *N*-terminus in co-conformers where all the threaded macrocycles are in the catchment region, that is, *tri*-**4** and *tri*-*tri*-**10** react faster than *car*-**4** and *car*-*tri*-**10**, suppressing dethreading. This means that the waste removal step (the Fmoc–carbamate-to-waste-products reaction) is also kinetically gated in the pumping mechanism of **1** (that is, deblocking of the *N*-terminus occurs more rapidly when all of the threaded macrocycles are located in the catchment region).

Fmoc removal from **4** with the strong base 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) was complete within 30 s, allowing the half-life of pseudorotaxane **1•3** to be measured as 230 min in $C_6D_5CD_3$ (**Figure 3b**); the rate-determining step is passage of the ring over the CF_3 group from the catchment region to the *N*-terminus. As ring-dethreading from **1•3** is much slower than the 10 min required to complete an active template threading-and-stoppering reaction, in the presence of excess fuel **1•3** collects additional rings through a continuous pumping process to give first [3]rotaxane **10** and then [4]rotaxane **7**. Further iterations of the pumping cycle did not yield any [5]rotaxane, suggesting that the axle is already full in the [4]rotaxane.

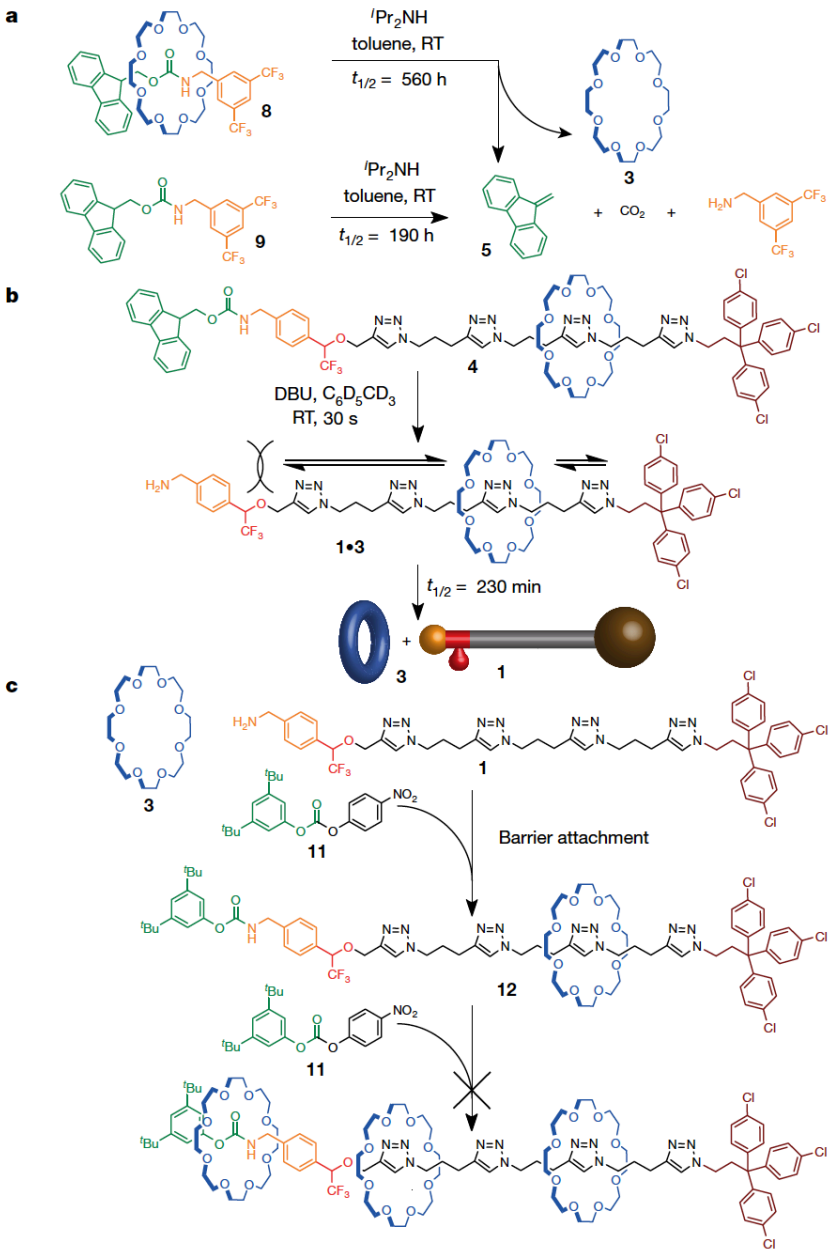


Figure 3. Fmoc removal, pseudorotaxane dethreading, and irreversible rotaxane formation experiments. **a**, Reagents and conditions for Fmoc removal from model [2]rotaxane **8** and model non-interlocked thread **9**: **8** or **9** (1 equiv.), *i*-Pr$_2$NH (10 equiv.), toluene (0.14 M), at room temperature (RT). **b**, Reagents and conditions for Fmoc removal and subsequent dethreading: rotaxane **4** (1 equiv.), DBU (50 equiv.), $C_6D_5CD_3$ (21 mM), room temperature, 30 s, then diluted to 4.2 mM with $C_6D_5CD_3$ and monitored by ^{1}H NMR. **c**, Reagents and conditions for formation of [2]rotaxane **12** using carbonate **11**: pump **1** (1 equiv.), 24-crown-8 **3** (10 equiv.), *i*-Pr$_2$NH (15 equiv.), toluene (0.2 M), reagent **11** (added continuously via syringe pump as a solution in toluene at a rate of 0.4 equiv. per hour), at room temperature, for 16 h. DBU, 1,8-diazabicyclo[5.4.0]undec-7-ene.

For the rotaxane with two rings pumped onto the axle (**10**; **Figure 2b**), two co-conformers are observed by ^{1}H NMR: *car-tri*-**10**, with one macrocycle residing over the

carbamate group and the other in the catchment region, and *tri-tri*-**10** with both macrocycles in the catchment region. The ratio of *car-tri*-**10** to *tri-tri*-**10** is 7:3 at 295 K in $C_6D_5CD_3$. In contrast, only one co-conformer is observed for the [4]rotaxane, *car-tri-tri*-**7** isomer (**Figure 2c**), which is consistent with the results of the pumping studies done in this work that suggest that the axle of the [4]rotaxane is only long enough to accommodate three macrocycles. Extending the catchment region to include more triazole rings (or other groups that weakly stabilize the crown ether) is complicated by the need for the pump to be soluble in toluene, which is the optimal solvent[28–30] for the active template reaction.

Considered together, these experimental results provide insight into the role and effect of the pump's power stroke. The ^{1}H NMR spectra (**Figure 2c**) show that although the threaded ring is at lower energy in the catchment region than in the carbamate region in the [2]rotaxane (*car*-**4**:*tri*-**4** 3:7; 295 K, $C_6D_5CD_3$), the [3]rotaxane is actually more thermodynamically stable when the additional ring is over the carbamate group rather than in the catchment region (*car-tri*-**10**:*tri-tri*-**10** 7:3; 295 K, $C_6D_5CD_3$). Steric considerations mean that [4]rotaxane can only be produced from active template threading of the *tri-tri*- co-conformer of **1**•(**3**)$_2$ (that is, both rings in the catchment region of the open-ended pump). Therefore, pumping of the [3]rotaxane (which is a 7:3 mixture of *car-tri*-**10**:*tri-tri*-**10**) to the [4]rotaxane (**7**) can be viewed as proceeding as follows: from (i) *car-tri*-**10** (the lowest-energy co-conformer of the [3]rotaxane) to (ii) *tri-tri*-**10** (a state that undergoes fast deblocking of the pump terminus) to (iii) **1**•(**3**)$_2$ (open-end pump available for active template threading, while the speed bump inhibits dethreading of already captured rings) to (iv) *car-tri-tri*-**7** (the [4]rotaxane); a sequence that actually involves an energetically uphill co-conformational change (*car-tri*-**10** to *tri-tri*-**10**)—a necessary, but endergonic, analogue of the power stroke that occurs to form the [2]- and [3]rotaxane.

Control experiments confirmed that there is no net displacement of 24-crown-8 rings from bulk solution onto **1** in the absence of fuel. At equilibrium, the amount of **1•3** present in toluene is lower than the limits detectable by spectroscopy (Section 3.4.3.6.). Furthermore, when an excess of an alternative carbonate, **11**, was used instead of fuel **2**, resulting in a *N*-terminus carbamate that is not decomposed by base, [2]rotaxane **12** was the only threaded product isolated (**Figure 3c**). This demonstrates that only one ring is pumped onto the axle by the consumption of each fuel molecule.

The pump can also operate if excess fuel and *i*-Pr_2NH are present from the start, allowing multiple fuel cycles to occur autonomously to generate [*n*]rotaxanes until all of the fuel is consumed. However, in the presence of *i*-Pr_2NH fuel **2** decomposes at a similar rate to the Fmoc–carbamate axle, so it is difficult to solubilize sufficient fuel at time zero to generate more than very small amounts of [4]rotaxane. Accordingly, pump operation is more efficient with fuel continuously drip-fed to the pump rather than delivered in a single pulse. Treatment of pump **1** with 10 equivalents of 24-crown-8 (**3**), 15 equivalents of *i*-Pr_2NH and 0.7 equivalents fuel (**2**) per hour for 16 h generated [4]rotaxane **7**, along with [2]rotaxane **4** and [3]rotaxane **10** (**Figure 4b**). The threaded products are formed dissipatively[34–37] under these operating conditions: Fmoc–carbamate decomposition continued for a further 19 days after ceasing the addition of fuel **2**, causing **3** to be slowly released back into solution resulting in virtually all of the original amount of unthreaded axle and rings, **1** and **3**, being recovered (Section 3.4.3.10.).

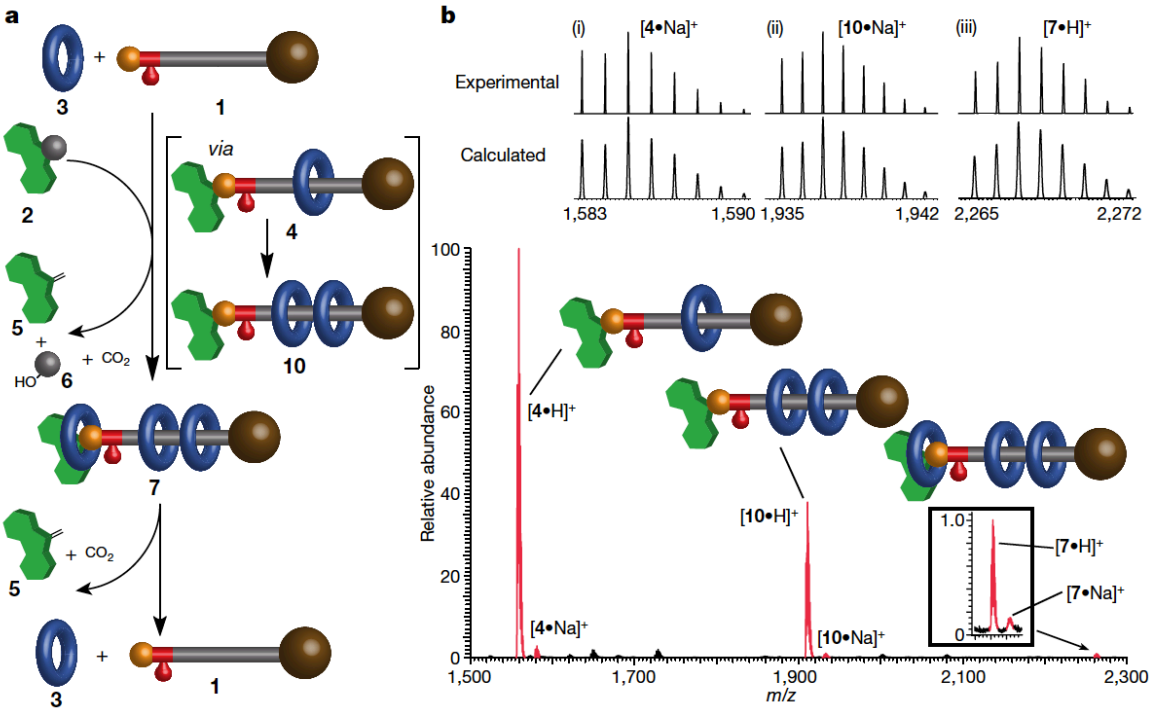


Figure 4. Out-of-equilibrium state produced by the operation of pump **1**. **a**, Reagents and conditions: pump **1** (1 equiv.), 24-crown-8 **3** (10 equiv.), *i*-Pr_2NH (15 equiv.), toluene (0.2 M), fuel **2** (added continuously via syringe pump as a solution in toluene at a rate of 0.7 equiv. per hour), at room temperature, for 16 h. After stopping the addition of fuel **2**, half of the reaction mixture was left at room temperature for 19 days, regenerating the unthreaded components, **1** and **3**. **b**, Electrospray ionization mass spectrum of the crude reaction mixture worked up just after stopping the addition of fuel **2**. Signal intensity does not correspond to the amount of product formed. The signal corresponding to [4]rotaxane **7** is shown at ×30 intensity. Isotopic patterns shown for [**4**•Na]$^+$, [**10**•Na]$^+$ and [**7**•H]$^+$. The cartoons indicate the number of rings in each rotaxane.

3.3. Conclusion

Synthetic chemically fuelled molecular pump **1** drives macrocycles from bulk solution into an organized, high-energy, threaded assembly, and dissipatively maintains the out-of-equilibrium state for as long as unreacted chemical fuel remains. The pump uses a doubly kinetically gated catalysis-driven information ratchet mechanism that is wholly artificial and minimalist in design[38], yet fundamentally analogous to that of biomolecular pumps. It is kinetic asymmetry in the act of threading under catalysis that drives the pumping away from equilibrium[20]. The mechanism of pumping features a step that, for the first two rings, has the characteristics of a power stroke—'a viscoelastic, free-energy releasing, large-amplitude conformational change'[39,40]—as the initially captured, high-energy crown ethers close to the

carbamate group (for example, *car*-**4**) relax to the catchment region (for example, *tri*-**4**). However, the analogous co-conformational change required to form the [4]rotaxane is an energetically uphill process (an 'anti-power stroke'). There is ongoing debate concerning the role of power strokes in catalysis-driven biomolecular machines[39,41]. In this catalysis-driven small-molecule pump the change of position of the ring(s) on the axle from the carbamate to the catchment region serves to retard the rate of release of the captured substrate. This co-conformational change is necessary for the operating mechanism irrespective of whether it is exergonic (for example, during pumping of the first two rings) or not (as is required to pump on the third ring). The mechanism of **1** is general (not restricted to particular substrates, functional groups or catalysts), and we anticipate that catalysis will play an increasingly important part in the design of artificial molecular machines[20,42–48] and, thereby[21], aid in the understanding of their biomolecular counterparts.

3.4. Experimental section

3.4.1. General information

Unless stated otherwise, reagents were obtained from commercial sources and used without purification. Dry THF (HPLC grade, Fischer Scientific), Et_2O (>99%, Fischer Scientific), CH_2Cl_2 (HPLC grade, Fischer Scientific), DMF (HPLC grade, Fischer Scientific) and toluene (>99%, Fischer Scientific) were obtained by passing the solvent through an activated alumina column on a Phoenix SDS (solvent drying system; JC Meyer Solvent Systems, CA, USA). ^{1}H NMR spectra were recorded on a Bruker Avance III instrument with an Oxford AS600 magnet equipped with a cryoprobe [5mm CPDCH 13C-1H/D] (600 MHz) at 298 K. Chemical shifts are reported in parts per million (ppm) from high to low frequency using the residual solvent peak as the internal reference ($CDCl_3$ = 7.26 ppm and $C_6D_5CD_3$ = 2.09 ppm). All ^{1}H resonances are reported to the nearest 0.01 ppm. The

multiplicity of ^{1}H signals are indicated as: s = singlet; d = doublet; t = triplet; q = quartet; p = pentet; m = multiplet; br = broad; or combinations of thereof. Coupling constants (*J*) are quoted in Hz and reported to the nearest 0.1 Hz. Where appropriate, averages of the signals from peaks displaying multiplicity were used to calculate the value of the coupling constant. ^{13}C NMR spectra were recorded on the same spectrometer at 298 K with the central resonance of the solvent peak as the internal reference ($CDCl_3$ = 77.16 ppm and $C_6D_5CD_3$ = 20.40 ppm). All ^{13}C resonances are reported to the nearest 0.01 ppm. DEPT, COSY, HSQC and HMBC experiments were used to aid structural determination and spectral assignment. Fully characterized compounds were chromatographically homogeneous. Flash column chromatography was carried out using Silica 60 Å (particle size 40 – 63 μm, Sigma Aldrich, UK) as the stationary phase. Preparative TLC was performed using either PLC 20 × 20 cm, 60 F254 preparatory plates (Merck) or Silica Gel GF 20 × 20 cm, U254 preparatory plates (Analtech) of various thicknesses (250 – 2000 μm). Analytical TLC was performed on precoated silica gel plates (0.25 mm thick, 60 F254, Merck, Germany) and visualized using both short and long waved ultraviolet light in combination with standard laboratory stains (acidic potassium permanganate, acidic ammonium molybdate and ninhydrin). Low resolution ESI mass spectrometry was performed with a Thermo Scientific LCQ Fleet Ion Trap Mass Spectrometer or an Agilent Technologies 1200 LC system with an Advion Expression CMS L single quadrupole MS detector. High-resolution mass spectrometry was carried out at the Mass Spectrometry Service, Department of Chemistry, University of Manchester. Pure samples were prepared for analysis by mass spectrometry by dissolution of a small quantity (*ca.* 1 mg) in the minimum amount of a suitable solvent (typically CH_2Cl_2 or acetone), followed by dilution with CH_3OH or CH_3CN (to a final concentration of approximately 1 mg/mL) and filtration. Samples of pump operations for analysis by mass spectrometry were prepared by diluting the sample for ^{1}H NMR measurement of the crude

(in $CDCl_3$) with CH_3CN (to give a final volume of 1 mL and a final concentration of approximately 1 mg/mL). The samples were then filtered and analyzed immediately. Compound **8** was synthesised according to a literature procedure[29]. Infrared spectroscopy was performed with Brucker Compact FT-IR Spectrometer: ALPHA II. The neat sample was directly applied to the instrument to measure IR spectra. Pump operations were performed in disposable glass vials. Operation reaction mixtures were homogenous. Heating or sonication were not required to dissolve reagents.

3.4.2. Synthesis

3.4.2.1. Synthesis of **S8**

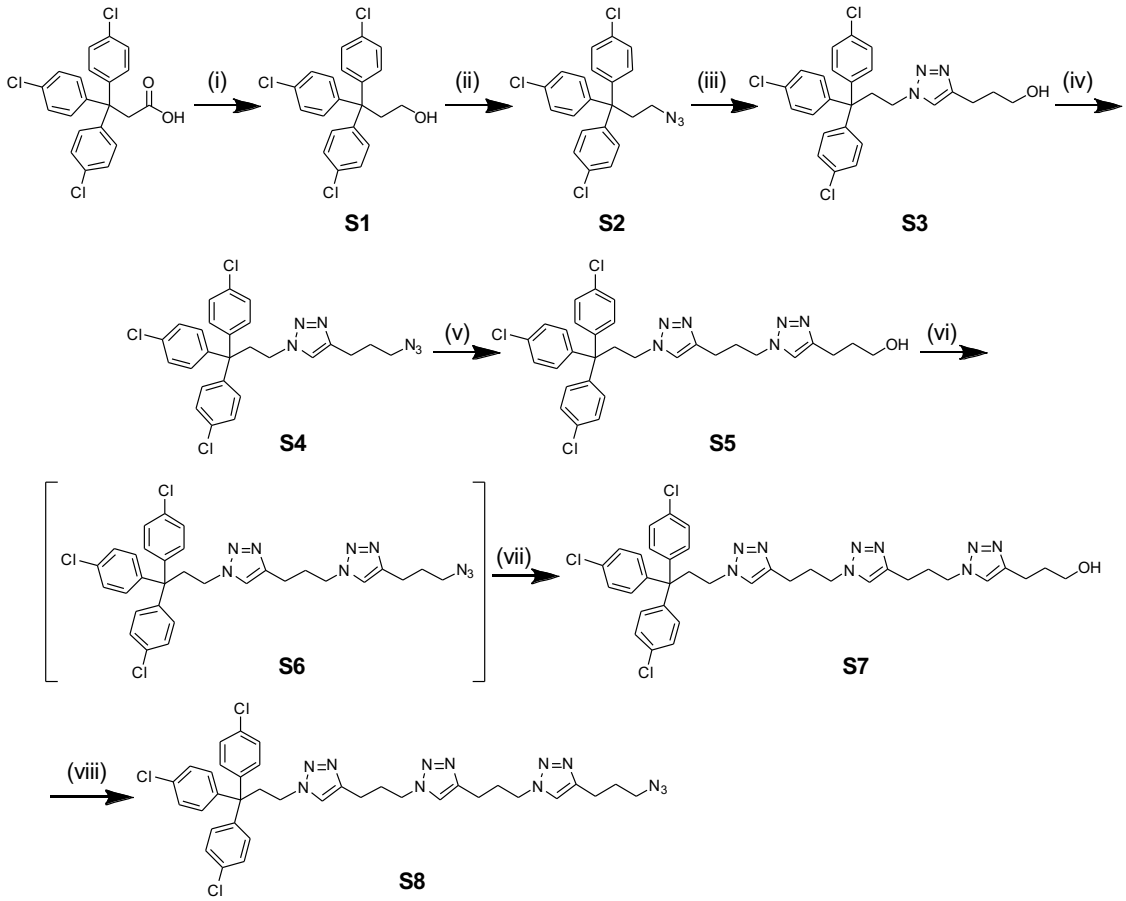
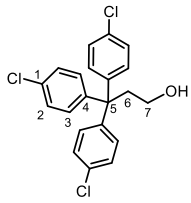


Figure S1. Synthesis of **S8.** Reagents and conditions: (i) BH_3·THF, THF, 0 °C to 60 °C, 16 h, quant.. (ii) ADMP, DBU, CH_2Cl_2, r.t. to reflux, 16 h, 91%. (iii) 4-Pentyn-1-ol, *i*-Pr_2NEt, $[Cu(MeCN)_4]PF_6$, THF/*t*-BuOH, 50 °C, 19 h, 87%. (iv) ADMP, DBU, CH_2Cl_2, 0 °C to r.t., 1 h, 84%. (v) 4-Pentyn-1-ol, *i*-Pr_2NEt, $[Cu(MeCN)_4]PF_6$, THF/*t*-BuOH, 50 °C, 17 h, 90%. (vi) ADMP, DBU, THF, 0 °C to r.t., 2 h, then (vii) 4-Pentyn-1-ol, *i*-Pr_2NEt, $[Cu(MeCN)_4]PF_6$, THF/*t*-BuOH, r.t., 16 h, 29% in two steps. (viii) ADMP, DBU, THF, 0 °C to r.t., 3 h, 80%.

Synthesis of **S1**



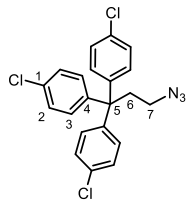
3,3,3-Tris(4-chlorophenyl)propionic acid (6.03 g, 14.8 mmol, 1 eq.) was dissolved in dry THF (25 mL). The flask was flushed with argon and cooled to 0 °C. To the solution was added BH_3•THF (1 M in THF, 55 mL, 55 mmol, 3.7 eq.) dropwise over 5 min. The mixture was stirred at 60 °C for 16 h. Saturated aqueous $NaHCO_3$ (10 mL) and water (5 mL) were added dropwise and the THF was removed under reduced pressure. To the residue was added saturated aqueous $NaHCO_3$ (100 mL) and brine (40 mL). The resulting suspension was extracted with EtOAc (3 × 100 mL). The combined organic layers were then washed with brine (50 mL), dried over $MgSO_4$ and concentrated under reduced pressure to give **S1** (5.8 g, 14.8 mmol, quant.) as a colourless solid.

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.25 (m, 6H, H_2), 7.15 (m, 6H, H_3), 3.43 (m, 2H, H_7), 2.84 (m, 2H, H_6), 1.17 (t, J = 5.0 Hz, 1H, OH).

^{13}C NMR (151 MHz, $CDCl_3$) δ 144.73, 132.61, 130.18, 128.57, 60.31, 54.24, 42.71.

HRMS (ESI^-): m/z = 425.0042 $[M+Cl]^-$ (calcd. 425.0039 for $C_{21}H_{17}OCl_4$).

Synthesis of **S2**



To a stirring solution of **S1** (2.1 g, 5.4 mmol, 1.0 eq.) and ADMP (3.1 g, 10.8 mmol, 2.0 eq.) in dry CH_2Cl_2 (50 mL) under argon at r.t. was added DBU (2.4 mL, 16.2 mmol, 3.0 eq.) dropwise. The resulting orange solution was heated to reflux for 16 h. After concentration under reduced pressure the crude mixture was purified by flash column chromatography (SiO_2, petroleum ether/10% → 20% CH_2Cl_2) to afford **S2** as a colourless foam (2.0 g, 4.9 mmol, 91%).

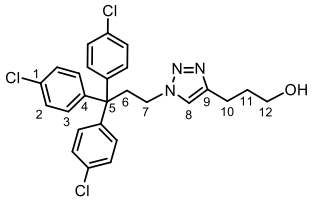
^{1}H NMR (600 MHz, $CDCl_3$) δ 7.27 (m, 6H, H_2), 7.13 (m, 6H, H_3), 3.02 (m, 2H, H_7), 2.79 (m, 2H, H_6).

^{13}C NMR (151 MHz, $CDCl_3$) δ 144.04, 132.88, 130.06, 128.75, 54.37, 48.54, 39.08.

This compound was not observed by HRMS (ESI, APCI, ASAP ionisation attempted).

IR (neat, υ/cm^{-1}): 2932, 2093, 1489, 1095, 1011, 811.

Synthesis of **S3**



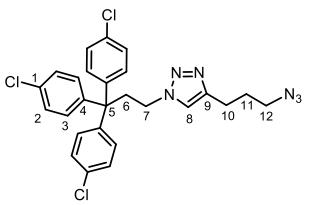
S2 (745 mg, 1.79 mmol, 1 eq.), 4-pentyn-1-ol (202 µL, 2.16 mmol, 1.2 eq.), *i*-Pr_2NEt (126 µL, 0.72 mmol, 0.4 eq.) and $[Cu(MeCN)_4]PF_6$ (134 mg, 0.36 mmol, 0.2 eq.) were dissolved in degassed THF/*t*-BuOH = 95/5 (20 mL) and the reaction mixture was stirred under an argon atmosphere at 50 °C for 19 h. The resulting mixture was concentrated under reduced pressure and then purified by flash column chromatography (SiO_2, CH_2Cl_2/30% → 70% EtOAc, then CH_2Cl_2/5% MeOH) to give **S3** as a colourless solid (780 mg, 1.56 mmol, 87%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.29 (m, 6H, H_2), 7.21 (m, 6H, H_3), 7.13 (s, 1H, H_8), 4.00 (m, 2H, H_7), 3.73 (br, 2H, H_{12}), 3.15 (m, 2H, H_6), 2.81 (t, J = 7.2 Hz, 2H, H_{10}), 1.92 (m, 2H, H_{11}).

^{13}C NMR (151 MHz, $CDCl_3$) δ 147.66, 143.73, 133.06, 130.08, 128.91, 121.25, 62.10, 54.58, 47.47, 40.53, 31.97, 22.27.

HRMS (ASAP$^+$): m/z = 500.1069 $[M+H]^+$ (calcd. 500.1058 for $C_{26}H_{25}N_3OCl_3$).

Synthesis of **S4**



S3 (770 mg, 1.54 mmol, 1 eq.) was suspended in dry CH_2Cl_2 (15 mL). The flask was flushed with argon and cooled to 0 °C. ADMP (795 mg, 2.79 mmol, 1.8 eq.) and DBU (700 µL, 4.69 mmol, 3 eq.)

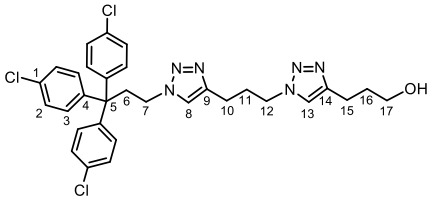
were added and the reaction mixture was stirred at r.t. for 1 h. Saturated aqueous NH_4Cl (15 mL) was added and the mixture was extracted with CH_2Cl_2 (3 × 20 mL). The combined organic layers were dried over $MgSO_4$ and concentrated under reduced pressure. The residue was purified by flash column chromatography (SiO_2, CH_2Cl_2/0% → 10% EtOAc) to give **S4** as a colourless foam (680 mg, 1.29 mmol, 84%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.30 (m, 6H, H_2), 7.22 (m, 6H, H_3), 7.11 (s, 1H, H_8), 4.01 (m, 2H, H_7), 3.34 (t, J = 6.7 Hz, 2H, H_{12}), 3.16 (m, 2H, H_6), 2.78 (t, J = 7.5 Hz, 2H, H_{10}), 1.98 (m, 2H, H_{11}).

^{13}C NMR (151 MHz, $CDCl_3$) δ 146.78, 143.73, 133.08, 130.08, 128.93, 121.23, 54.58, 50.74, 47.45, 40.57, 28.63, 22.71.

HRMS (ASAP$^+$): m/z = 525.1138 $[M+H]^+$ (calcd. 525.1123 for $C_{26}H_{24}N_6Cl_3$).

Synthesis of **S5**



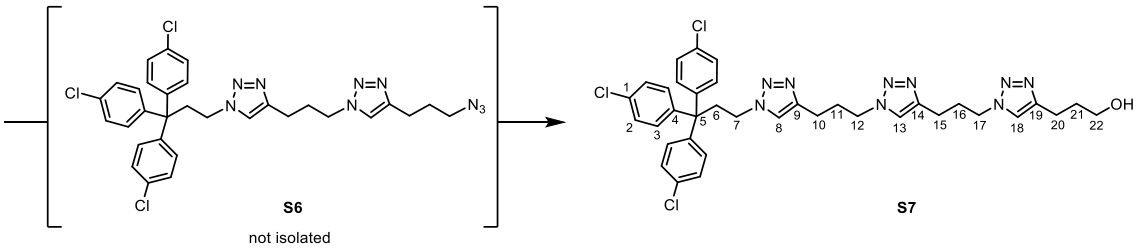
S4 (306 mg, 0.58 mmol, 1 eq.), 4-pentyn-1-ol (64 µL, 0.69 mmol, 1.2 eq.), *i*-Pr_2NEt (40 µL, 0.23 mmol, 0.4 eq.) and $[Cu(MeCN)_4]PF_6$ (43 mg, 0.12 mmol, 0.2 eq.) were dissolved in degassed THF/*t*-BuOH = 95/5 (7 mL) and the reaction mixture was stirred under an argon atmosphere at 50 °C for 17 h. The resulting mixture was concentrated under reduced pressure and then purified by flash column chromatography (SiO_2, CH_2Cl_2/3% → 5% MeOH) to give **S5** as a colourless foam (320 mg, 0.52 mmol, 90%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.41 (s, 1H, H_{13}), 7.33 (m, 6H, H_2), 7.25 (m, 6H, H_3), 7.20 (s, 1H, H_8), 4.40 (t, J = 6.8 Hz, 2H, H_{12}), 4.03 (m, 2H, H_7), 3.74 (q, J = 5.6 Hz, 2H, H_{17}), 3.18 (m, 2H, H_6), 2.85 (t, J = 7.2 Hz, 2H, H_{15}), 2.70 (t, J = 7.0 Hz, 2H, H_{10}), 2.30 (p, J = 6.9 Hz, 2H, H_{11}), 2.18 (br, 1H), 1.96 (m, 2H, H_{16}).

^{13}C NMR (151 MHz, $CDCl_3$) δ 147.73, 146.16, 143.71, 133.09, 130.08, 128.94, 121.75, 121.31, 62.20, 54.58, 49.11, 47.50, 40.51, 31.98, 29.84, 22.37, 22.23.

HRMS (ASAP$^+$): *m/z* = 609.1717 $[M+H]^+$ (calcd. 609.1698 for $C_{31}H_{32}N_6OCl_3$).

Synthesis of **S7** via **S6**



S5 (304 mg, 0.51 mmol, 1 eq.) was dissolved in dry THF (5 mL). The flask was flushed with argon and cooled to 0 °C. ADMP (236 mg, 0.83 mmol, 1.6 eq.) and DBU (220 µL, 1.47 mmol, 2.9 eq.) were added and the reaction mixture was stirred at r.t. for 2 h. Saturated aqueous NH_4Cl (20 mL) was added and the mixture was extracted with CH_2Cl_2 (3 × 20 mL). The combined organic layers were dried over $MgSO_4$ and concentrated under reduced pressure. The residue was partially purified by flash column chromatography (SiO_2, CH_2Cl_2/50% → 70% EtOAc then CH_2Cl_2/5% MeOH) to give a mixture of **S6** and 1,3-dimethyl-2-imidazolidinone in a 1:1.3 ratio (determined by ^{1}H NMR, **Figure S24**), a by-product originating from ADMP, as a colourless gum (260 mg in total, approx. 0.33 mmol of **S6**). The crude mixture was used without further purification. 1,3-dimethyl-2-imidazolidinone was readily removed in the next step.

To a stirring solution of crude **S6** (1.0 g in total, 1.3 mmol of **S6**, 1.0 eq.), 4-pentyn-1-ol (450 µL, 4.8 mmol, 3.7 eq.) and *i*-Pr_2NEt (110 µL, 0.64 mmol, 0.5 eq.) in degassed THF/*t*-BuOH = 99/1 (10 mL) was added $[Cu(MeCN)_4]PF_6$ (119 mg, 0.32 mmol, 0.25 eq.). The reaction mixture was stirred under an argon atmosphere at r.t. for 16 h and then diluted with EtOAc (50 mL). The combined organic layers were washed with saturated aqueous EDTA (3 × 20 mL), aqueous HCl (1 M, 20 mL) and brine (20 mL). Concentration under reduced pressure

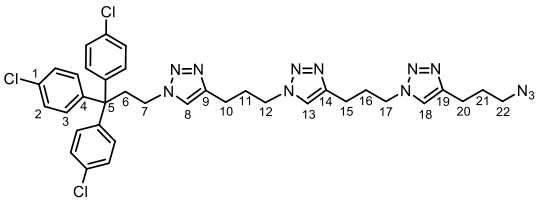
followed by flash column chromatography (SiO_2, CH_2Cl_2/5% → 10% MeOH) afforded **S7** as a colourless gum (409 mg, 0.57 mmol, 29% after two steps).

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.41 (s, 1H, H_{13}), 7.38 (s, 1H, H_{18}), 7.29 (m, 6H, H_2), 7.22 (m, 6H, H_3), 7.20 (s, 1H, H_8), 4.38 (t, J = 6.8 Hz, 4H, H_{12} + H_{17}), 4.01 (m, 2H, H_7), 3.71 (t, J = 6.1 Hz, 2H, H_{22}), 3.16 (m, 2H, H_6), 2.82 (t, J = 7.2 Hz, 2H, H_{20}), 2.68 (m, 4H, H_{10} + H_{15}), 2.28 (m, 4H, H_{11} + H_{16}), 1.93 (m, 2H, H_{21}).

^{13}C NMR (151 MHz, $CDCl_3$) δ 147.65, 146.11, 146.08, 143.71, 133.06, 130.08, 128.91, 121.84, 121.74, 121.34, 62.11, 54.57, 49.13, 49.10, 47.49, 40.49, 32.01, 29.83, 29.80, 22.33, 22.30, 22.20.

HRMS (ESI$^+$): *m/z* = 740.2153 $[M+Na]^+$ (calcd. 740.2157 for $C_{36}H_{38}N_9OCl_3Na$).

Synthesis of **S8**



S7 (373 mg, 0.52 mmol, 1 eq.) and DBU (250 µL, 1.68 mmol, 3.2 eq.) were dissolved in dry THF (4 mL). The flask was flushed with argon and cooled to 0 °C. ADMP (319 mg, 1.11 mmol, 2.1 eq.) was added and the reaction mixture was stirred at r.t. for 3 h. Saturated aqueous NH_4Cl (20 mL) was added and the mixture was extracted with CH_2Cl_2 (3 × 20 mL). The combined organic layers were washed with saturated aqueous NH_4Cl (20 mL), water (20 mL) and brine (20 mL), dried over $MgSO_4$ and concentrated under reduced pressure. The residue was purified by flash column chromatography (SiO_2, CH_2Cl_2/2% → 5% MeOH) to give **S8** as a colourless foam (310 mg, 0.42 mmol, 80%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.42 (s, 1H, H_{13}), 7.40 (s, 1H, H_{18}), 7.30 (m, 6H, H_2), 7.22 (m, 6H, H_3), 7.20 (s, 1H, H_8), 4.39 (m, 4H, H_{12} + H_{17}), 4.02 (m, 2H, H_7), 3.35 (t, J = 6.7 Hz,

2H, H_{22}), 3.16 (m, 2H, H_6), 2.80 (t, J = 7.5 Hz, 2H, H_{20}), 2.69 (m, 4H, H_{10} + H_{15}), 2.29 (m, 4H, H_{11} + H_{16}), 1.99 (m, 2H, H_{21}).

^{13}C NMR (151 MHz, $CDCl_3$) δ 146.80, 146.12, 146.09, 143.72, 133.08, 130.09, 128.93, 121.84, 121.74, 121.37, 54.58, 50.81, 49.14, 49.12, 47.50, 40.51, 29.89, 29.85, 28.67, 22.81, 22.32, 22.22.

HRMS (ESI$^+$): m/z = 765.2223 $[M+Na]^+$ (calcd. 765.2222 for $C_{36}H_{37}N_{12}Cl_3Na$).

3.4.2.2. Synthesis of **1**

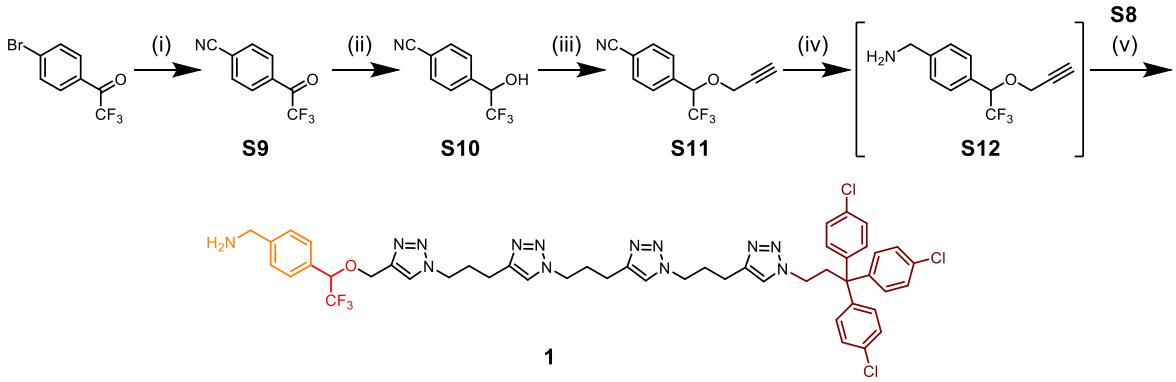
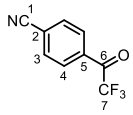


Figure S2. Synthesis of **1.** Reagents and conditions: (i) $Pd(PPh_3)_4$, $Zn(CN)_2$, DMF, 80 °C, 28 h, 66%. (ii) $NaBH_4$, THF/MeOH, 0 °C to r.t., 1.5 h, quant.. (iii) Propargyl bromide in toluene, Cs_2CO_3, DMF, 0 °C to r.t., 23 h, 96%. (iv) $LiAlH_4$, Et_2O, r.t., 7 h, then (v) *i*-Pr_2NEt, $[Cu(MeCN)_4]PF_6$, THF/*t*-BuOH, 50 °C, 22 h, 53% over two steps.

Synthesis of **S9**



4′-Bromo-2,2,2-trifluoroacetophenone (1.51 g, 5.98 mmol, 1 eq.), $Pd(PPh_3)_4$ (275 mg, 0.24 mmol, 0.04 eq.) and $Zn(CN)_2$ (430 mg, 3.66 mmol, 0.6 eq.) were dissolved in dry degassed DMF (8 mL). The reaction mixture was stirred at 80 °C for 28 h under argon atmosphere. To the mixture was added aqueous NH_3 (2 M, 30 mL). The mixture was extracted with Et_2O (3 × 20 mL). The combined organic layers were washed with aqueous NH_3 (2 M, 20 mL), water (20 mL) and brine (20 mL), dried over $MgSO_4$ and concentrated under reduced pressure. The residue was purified by flash column

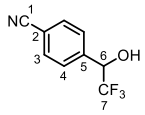
chromatography (SiO_2, petroleum ether/70% → 100% CH_2Cl_2) to give **S9** as an off-white solid (780 mg, 3.92 mmol, 66%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 8.18 (d, *J* = 7.8 Hz, 2H, H_4), 7.87 (d, *J* = 7.8 Hz, 2H, H_3).

^{13}C NMR (151 MHz, $CDCl_3$) δ 179.62 (q, *J* = 36.4 Hz), 132.98, 130.56 (q, *J* = 2.3 Hz), 118.91, 117.35, 116.39 (q, *J* = 290.4 Hz).

HRMS (ASAP$^-$): *m/z* = 199.0245 $[M]^-$ (calcd. 199.0250 for $C_9H_4NOF_3$).

Synthesis of **S10**



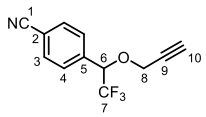
In a dried flask, **S9** (1.12 g, 5.62 mmol, 1 eq.) was dissolved in dry THF (6 mL) and MeOH (6 mL). The flask was flushed with argon and cooled to 0 °C. $NaBH_4$ (323 mg, 8.54 mmol, 1.5 eq.) was added portionwise over 5 mins and the reaction mixture was then stirred at r.t. for 1.5 h. The mixture was concentrated under reduced pressure and diluted with EtOAc (60 mL). The resulting suspension was washed with brine (2 × 20 mL), water (20 mL) and brine (20 mL). The combined aqueous layers were extracted a further time with EtOAc (20 mL). The combined organic layers were dried over $MgSO_4$ and concentrated under reduced pressure to give **S10** (1.13 g, 5.62 mmol, quant.) as a colourless solid.

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.72 (m, 2H, H_3), 7.63 (m, 2H, H_4), 5.12 (q, *J* = 6.5 Hz, 1H, H_6), 2.77 (br, 1H, OH).

^{13}C NMR (151 MHz, $CDCl_3$) δ 138.73, 132.48, 128.35, 123.89 (q, *J* = 282.3 Hz), 118.36, 113.67, 72.17 (q, *J* = 32.3 Hz).

HRMS (ASAP$^-$): *m/z* = 236.0089 $[M+Cl]^-$ (calcd. 236.0095 for $C_9H_6NOClF_3$).

Synthesis of **S11**



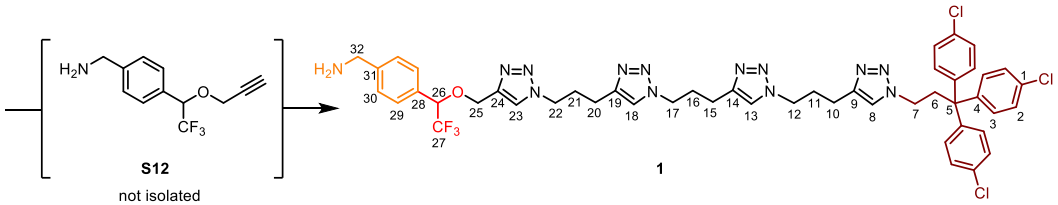
In a dried flask, **S10** (1.13 g, 5.62 mmol, 1 eq.) was dissolved in dry DMF (23 mL) and propargyl bromide (80% *w*/*w* solution in toluene, 1.44 mL, 2.3 eq.) was added. The flask was flushed with argon and cooled to 0 °C. Whilst the solution was vigorously stirred, Cs_2CO_3 (2.42 g, 7.44 mmol, 1.3 eq.) was added. The resulting suspension was stirred at r.t. for 23 h. Water (100 mL) was added and the mixture was extracted with Et_2O (3 × 50 mL). The combined organic layers were washed with water (2 × 40 mL), dried over $MgSO_4$ and concentrated under reduced pressure to give **S11** (1.29 g, 5.39 mmol, 96%) as a pale brown solid.

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.73 (m, 2H, H_3), 7.59 (d, J = 8.0 Hz, 2H, H_4), 5.02 (q, J = 6.4 Hz, 1H, H_6), 4.40 (dd, J = 16.1, 2.4 Hz, 1H, H_8), 4.15 (dd, J = 16.1, 2.4 Hz, 1H, H_8), 2.51 (t, J = 2.4 Hz, 1H, H_{10}).

^{13}C NMR (151 MHz, $CDCl_3$) δ 137.24, 132.59, 129.25, 123.41 (q, J = 282.1 Hz), 118.29, 113.99, 77.20, 77.11, 76.88, 57.79.

HRMS (APCI$^-$): m/z = 238.0478 $[M–H]^-$ (calcd. 238.0485 for $C_{12}H_7NOF_3$).

Synthesis of **1** via **S12**



In a dried flask, $LiAlH_4$ (1 M in THF, 6 mL, 6 mmol, 2.1 eq.) was concentrated under reduced pressure. **S11** (680 mg, 2.84 mmol, 1 eq.) as a solution in dry Et_2O (20 mL) was added dropwise. The resulting slurry was stirred at r.t. under argon atmosphere for 7 h. The reaction was quenched with water (200 µL), 15% (*w*/*w*) aqueous NaOH (200 µL) and water (600 µL) and then stirred at r.t. for 1.5 h. The mixture was filtered through a pad of Celite® and concentrated under reduced pressure to give partially purified **S12** (582 mg, assume 2.40

mmol) as a pale yellow oil. The crude (^{1}H NMR in **Figure S35**) was used in the next step without further purification.

Crude **S12** (35 mg, assume 0.14 mmol, 1 eq.), **S8** (123 mg, 0.17 mmol, 1.2 eq.), *i*-Pr_2NEt (19.4 μL, 0.11 mmol, 0.8 eq.) and $[Cu(MeCN)_4]PF_6$ (21.5 mg, 0.058 mmol, 0.4 eq.) were dissolved in degassed THF/*t*-BuOH = 95/5 (1.5 mL) and the reaction mixture was stirred under nitrogen atmosphere at 50 °C for 22 h. The reaction mixture was concentrated under reduced pressure and the residue was purified by PTLC (SiO_2, CH_2Cl_2/2% MeOH/1% Et_3N). The product was then dissolved in CH_2Cl_2 and washed with saturated aqueous $NaHCO_3$. The organic layer was dried over $MgSO_4$ and concentrated under reduced pressure to give **1** as an off-white solid (90 mg, 0.09 mmol, 53% over two steps).

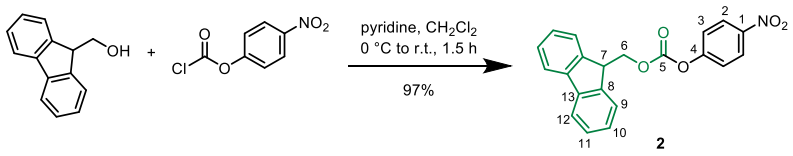
^{1}H NMR (600 MHz, $CDCl_3$) δ 7.64 (s, 1H, H_{23}), 7.45 (s, 1H, $H_{13/18}$), 7.43 (m, 3H, $H_{13/18}$ + H_{29}), 7.36 (d, *J* = 7.9 Hz, 2H, H_{30}), 7.30 (m, 6H, H_2), 7.22 (m, 6H, H_3), 7.20 (s, 1H, H_8), 4.78 (q, *J* = 6.6 Hz, 1H, H_{26}), 4.74 (d, *J* = 12.5 Hz, 1H, H_{25}), 4.63 (d, *J* = 12.5 Hz, 1H, H_{25}), 4.42 (t, *J* = 7.0 Hz, 2H, H_{22}), 4.39 (m, 4H, H_{12} + H_{17}), 4.01 (m, 2H, H_7), 3.90 (s, 2H, H_{32}), 3.16 (m, 2H, H_6), 2.74 – 2.63 (m, 6H, H_{10} + H_{15} + H_{20}), 2.35 – 2.22 (m, 6H, H_{11} + H_{16} + H_{21}).

^{13}C NMR (151 MHz, $CDCl_3$) δ 146.09, 146.01, 144.08, 143.70, 133.06, 130.89, 130.07, 128.92, 128.78, 127.55, 123.41, 121.86, 121.83, 121.74, 78.42, 63.32, 54.56, 49.37, 49.19, 49.12, 47.49, 40.48, 29.85, 29.81, 22.31, 22.25, 22.21.

HRMS (ESI^+): *m/z* = 986.3230 $[M+H]^+$ (calcd. 986.3273 for $C_{48}H_{50}N_{13}OCl_3F_3$).

3.4.2.3. Synthesis of fuel **2**, model compound **9** and reagent **11**

Synthesis of fuel **2**



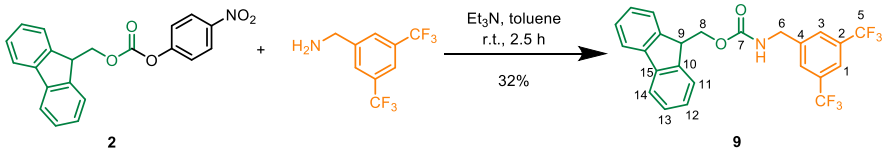
9-Fluorenemethanol (504 mg, 2.57 mmol, 1 eq.) was dissolved in dry CH_2Cl_2 (6 mL). The flask was flushed with argon and cooled to 0 °C. To the solution was added dry pyridine (0.42 mL, 5.2 mmol, 2 eq.) and a solution of 4-nitrophenyl chloroformate (614 mg, 3.05 mmol, 1.2 eq.) in dry CH_2Cl_2 (2 mL) in a dropwise manner. The mixture was stirred at r.t. for 1.5 h. Saturated aqueous NH_4Cl (30 mL) was added to the mixture and it was extracted with CH_2Cl_2 (3 × 30 mL). The combined organic layers were washed with brine (30 mL) and water (30 mL), dried over $MgSO_4$ and concentrated under reduced pressure. The residue was purified by flash column chromatography (SiO_2, petroleum ether/5% → 10% EtOAc) to give **2** as a colourless solid (900 mg, 2.49 mmol, 97%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 8.27 (d, J = 9.1 Hz, 2H, H_3), 7.80 (d, J = 7.6 Hz, 2H, H_{12}), 7.64 (d, J = 7.5 Hz, 2H, H_9), 7.45 (t, J = 7.5 Hz, 2H, H_{11}), 7.38 – 7.32 (m, 5H, H_2 + H_{10}), 4.60 (d, J = 7.1 Hz, 2H, H_6), 4.34 (t, J = 7.1 Hz, 1H, H_7).

^{13}C NMR (151 MHz, $CDCl_3$) δ 155.62, 152.59, 145.59, 142.98, 141.55, 128.29, 127.44, 125.47, 125.16, 121.96, 120.39, 71.05, 46.78.

HRMS This compound was not observed by HRMS (ESI, APCI, ASAP ionisation attempted).

Synthesis of **9**



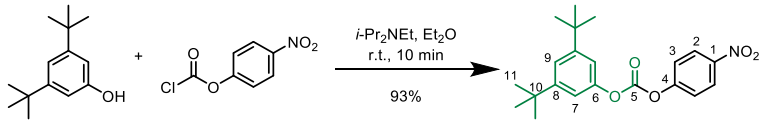
3,5-Bis(trifluoromethyl)benzylamine (11.6 mg, 0.048 mmol, 1 eq.) and Et_3N (14 µL, 0.10 mmol, 2.1 eq.) were dissolved in toluene (200 µL) and **2** (23.4 mg, 0.064 mmol, 1.3 eq.) was added. The reaction mixture was stirred at r.t. for 2.5 h. CF_3COOH (20 µL) and saturated aqueous NH_4Cl (5 mL) were added to the mixture and it was extracted with CH_2Cl_2 (3 × 5 mL). The combined extracts were concentrated under reduced pressure. The residue was purified by PTLC (SiO_2, hexane/80% CH_2Cl_2) to give **9** as a colourless solid (7.1 mg, 0.015 mmol, 32%).

^{1}H NMR (600 MHz, $C_6D_5CD_3$) δ 7.61 (s, 1H, H_1), 7.55 (d, J = 7.5 Hz, 2H, H_{14}), 7.33 (d, J = 7.4 Hz, 2H, H_{11}), 7.22 (m, 4H, H_3 + H_{13}), 7.17 (t, J = 7.4 Hz, 2H, H_{12}), 4.37 (d, J = 5.7 Hz, 2H, H_8), 3.82 (m, 2H, H_9 + NH), 3.57 (d, J = 6.4 Hz, 2H, H_6).

^{13}C NMR (151 MHz, $C_6D_5CD_3$) δ 156.11, 144.36, 142.24, 141.92, 131.74 (q, J = 33.2 Hz), 127.75, 127.20, 124.90, 122.96, 121.01, 120.22, 66.03, 47.68, 43.76.

HRMS (ESI$^+$): m/z = 488.1044 $[M+Na]^+$ (calcd. 488.1056 for $C_{24}H_{17}NO_2F_6Na$).

Synthesis of **11**



11

To a stirring solution of 3,5-di-*tert*-butylphenol (206 mg, 1.0 mmol, 1.0 eq.) and 4-nitrophenyl chloroformate (212 mg, 1.05 mmol, 1.05 eq.) in Et_2O (10 mL) at r.t. was added *i*-Pr_2NEt (178 µL, 1.0 mmol, 1.0 eq.) dropwise. The resulting mixture was stirred for 10 min during which time a white precipitate formed. The suspension was diluted with Et_2O (10 mL) and washed with water (20 mL) and brine (20 mL). The organic extract was dried with $MgSO_4$ and concentrated under reduced pressure to afford reagent **11** as a colourless solid (346 mg, 0.93 mmol, 93%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 8.32 (d, J = 9.5 Hz, 2H, H_3), 7.50 (d, J = 9.5 Hz, 2H, H_2), 7.35 (t, J = 1.7 Hz, 1H, H_9), 7.09 (d, J = 1.7 Hz, 2H, H_7), 1.33 (s, 18H, H_{11}).

^{13}C NMR (151 MHz, $CDCl_3$) δ 155.57, 152.95, 151.44, 150.62, 145.68, 125.52, 121.96, 120.81, 115.01, 35.23, 31.48.

HRMS (ESI^+): m/z = 372.1802 $[M+H]^+$ (calcd. 372.1805 for $C_{21}H_{26}NO_5$).

3.4.2.4. Stepwise synthesis of [2]rotaxane **4**, [3]rotaxane **10**, [4]rotaxane **7** and [2]rotaxane **12**

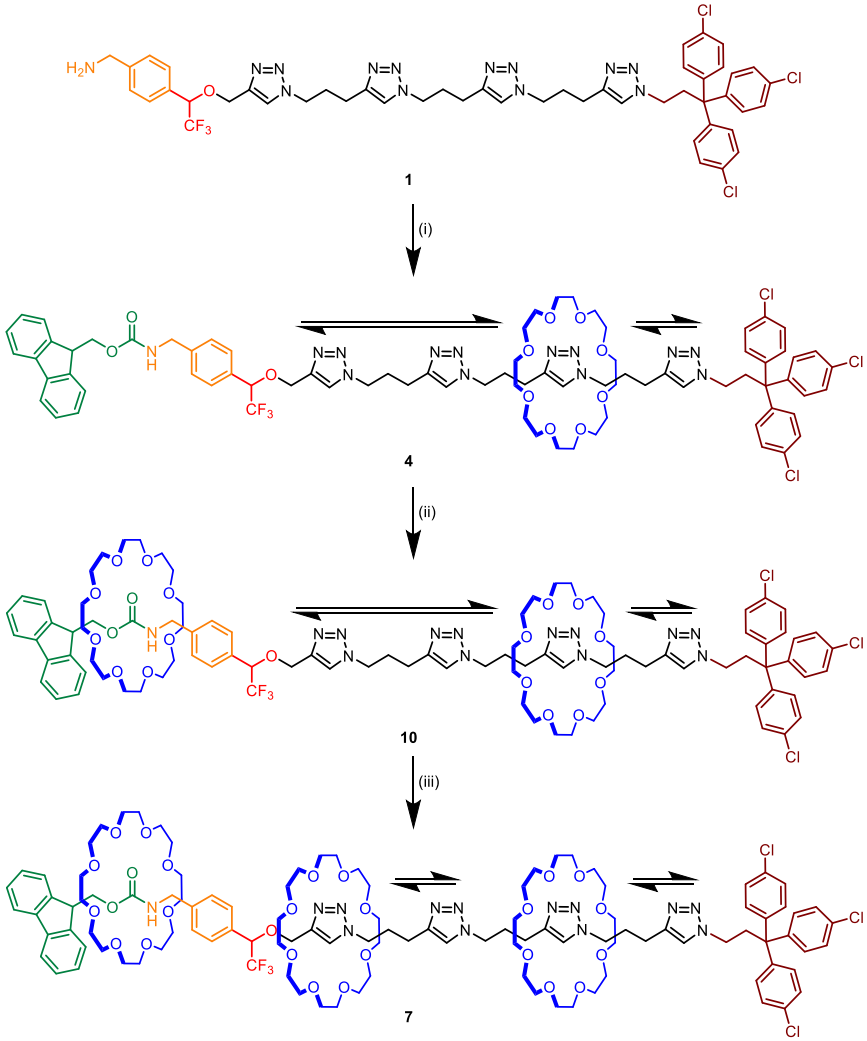
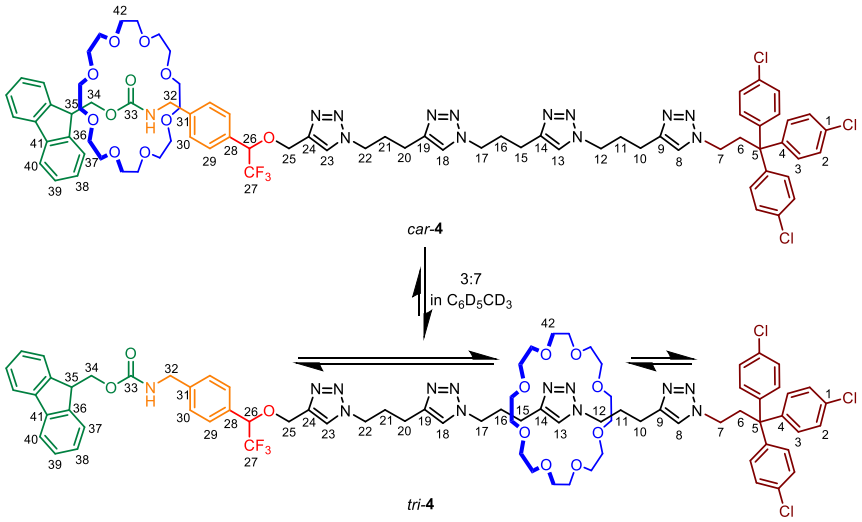


Figure S3. Stepwise synthesis of [2]rotaxane **4**, [3]rotaxane **10** and [4]rotaxane **7**. Reagents and conditions: (i) **3**, **2**, Et_3N, toluene, r.t., 1 h, 54%. (ii) **3**, DBU, toluene, r.t., 1 min then **2**, Et_3N, r.t., 30 min, 23%. (iii) **3**, *i*-Pr_2NH, **2** (added continuously), toluene, r.t., 23 h, 6%.

Synthesis of **4**



2 (13 mg, 34 µmol, 3 eq.) was added to a solution of **1** (11.2 mg, 11 µmol, 1 eq.), **3** (40 mg, 110 µmol, 10 eq.) and Et_3N (7.9 µL, 57 µmol, 5 eq.) in toluene (50 µL). The resulting mixture was stirred at r.t. for 1 h. To the reaction were added CF_3COOH (10 µL) and saturated aqueous NH_4Cl (5 mL) and the suspension was extracted with CH_2Cl_2 (3 × 5 mL). The combined organic extracts were concentrated under reduced pressure and the residue was purified by SEC (Bio-Beads S-X1, CH_2Cl_2) followed by PTLC (SiO_2, CH_2Cl_2/5% MeOH) to give **4** as an off-white solid (9.7 mg, 6 µmol, 54%).

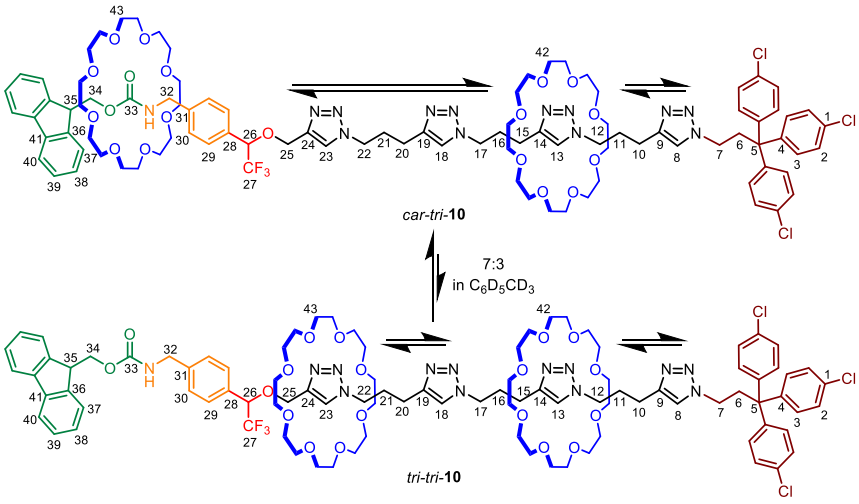
^{1}H NMR (600 MHz, $C_6D_5CD_3$) δ 8.12 (d, J = 7.6 Hz, 2H, H_{30}^{car}), 7.70 (d, J = 7.5 Hz, 2H, H_{37}^{car}), 7.60 – 7.50 (m, 5H, H_{13}^{tri} + H_{18}^{tri} + H_{23}^{tri} + H_{40}^{car} + H_{40}^{tri}), 7.45 (d, J = 7.6 Hz, 4H, H_{29}^{car} + H_{37}^{tri}), 7.33 (d, J = 7.8 Hz, 2H, H_{29}^{tri}), 7.23 (m, 2H, H_{39}^{car} + H_{39}^{tri}), 7.17 (m, 2H, H_{38}^{car} + H_{38}^{tri}), 7.05 (m, 7H, H_2^{car} + H_2^{tri} + H_8^{tri}), 7.00 (br, 2H, H_{30}^{tri}), 6.98 (s, 1H, $H_{13}^{car}{}_{/18}^{car}{}_{/23}^{car}$), 6.91 (s, 1H, NH^{car}), 6.86 (m, 7H, H_3^{tri} + $H_{13}^{car}{}_{/18}^{car}{}_{/23}^{car}$), 6.83 (m, 6H, H_3^{car}), 6.75 (s, 1H, $H_{13}^{car}{}_{/18}^{car}{}_{/23}^{car}$), 6.47 (s, 1H, H_8^{car}), 4.93 (br, 2H, H_{32}^{car}), 4.81 (m, 1H, H_{26}^{tri}), 4.77 (m, 1H, NH^{tri}), 4.72 (m, 1H, H_{26}^{car}), 4.64 (d, J = 12.1 Hz, 1H, H_{25}^{car} + H_{25}^{tri}), 4.55 (d, J = 12.1 Hz, 1H, H_{25}^{car} + H_{25}^{tri}), 4.52 (d, J = 6.7 Hz, 2H, H_{34}^{car}), 4.43 (d, J = 6.3 Hz, 2H, H_{34}^{tri}), 4.38 – 4.23 (m, 6H, H_{12}^{tri} + H_{17}^{tri} + H_{22}^{tri}), 4.12 (t, J = 6.6 Hz, 1H, H_{35}^{car}), 4.02 (d, J

= 6.4 Hz, 2H, H_{32}^{tri}), 3.97 (m, 3H, $H_{12}^{car}/_{17}^{car}/_{22}^{car}$ + H_{35}^{tri}), 3.91 (t, *J* = 6.6 Hz, 2H, $H_{12}^{car}/_{17}^{car}/_{22}^{car}$), 3.85 (td, *J* = 6.8, 2.6 Hz, 2H, $H_{12}^{car}/_{17}^{car}/_{22}^{car}$), 3.69 (m, 2H, H_{7}^{tri}), 3.60 (m, 2H, H_{7}^{car}), 3.44 – 3.07 (m, 32H, H_{42}^{car} + H_{42}^{tri}), 2.88 (m, 2H, H_{6}^{tri}), 2.82 (m, 2H, H_{6}^{car}), 2.81 – 2.68 (m, 6H, H_{10}^{tri} + H_{15}^{tri} + H_{20}^{tri}), 2.50 (t, *J* = 7.1 Hz, 2H, $H_{10}^{car}/_{15}^{car}/_{20}^{car}$), 2.44 (t, *J* = 6.9 Hz, 2H, $H_{10}^{car}/_{15}^{car}/_{20}^{car}$), 2.38 (t, *J* = 7.1 Hz, 2H, $H_{10}^{car}/_{15}^{car}/_{20}^{car}$), 2.33 – 2.18 (m, 6H, H_{11}^{tri} + H_{16}^{tri} + H_{21}^{tri}), 2.04 (m, 2H, $H_{11}^{car}/_{16}^{car}/_{21}^{car}$), 1.98 (m, 2H, $H_{11}^{car}/_{16}^{car}/_{21}^{car}$), 1.90 (m, 2H, $H_{11}^{car}/_{16}^{car}/_{21}^{car}$).

^{13}C NMR (151 MHz, $C_6D_5CD_3$) δ 156.44, 147.11, 146.88, 146.39, 145.92, 145.83, 145.72, 145.34, 144.65, 144.25, 144.11, 143.18, 141.89, 141.02, 132.99, 132.93, 132.33, 132.05, 130.36, 130.29, 128.92, 128.82, 127.70, 127.59, 127.54, 127.28, 127.20, 126.05, 124.16, 123.01, 122.49, 122.44, 122.22, 121.62, 121.54, 120.20, 119.92, 78.40, 71.07, 70.74, 66.16, 65.63, 63.69, 63.40, 54.58, 49.32, 49.10, 48.97, 48.79, 48.69, 48.61, 48.23, 47.91, 47.04, 45.42, 44.60, 40.53, 40.45, 32.43, 30.31, 30.04, 29.99, 29.92, 29.87, 29.59, 23.18, 22.73, 22.41, 22.35, 22.31.

HRMS (ESI$^+$): *m/z* = 1582.5839 $[M+Na]^+$ (calcd. 1582.5871 for $C_{79}H_{91}N_{13}O_{11}Cl_3F_3Na$).

Synthesis of **10**



DBU (2.5 μL, 17 μmol, 1.3 eq.) was added to a solution of **4** (20 mg, 13 μmol, 1 eq.) and **3** (45 mg, 130 μmol, 10 eq.) in toluene (70 μL). The mixture was vigorously shaken at r.t. for 1 min. **2** (16 mg, 44 μmol, 3.4 eq.) and Et_3N (3.6 μL, 26 μmol, 2 eq.) were then added and the mixture was stirred at r.t. for a further 30 min. CF_3COOH (7 μL) and saturated aqueous NH_4Cl (5 mL) were added to the mixture and it was extracted with CH_2Cl_2 (3 × 5 mL). The combined organic layers were concentrated under reduced pressure. The residue was purified by PTLC (SiO_2, CH_2Cl_2/5% MeOH), followed by SEC (Bio-Beads S-X1, CH_2Cl_2) and then PTLC (SiO_2, CH_2Cl_2/5% MeOH) to give **10** as a colourless film (6 mg, 3 μmol, 23%).

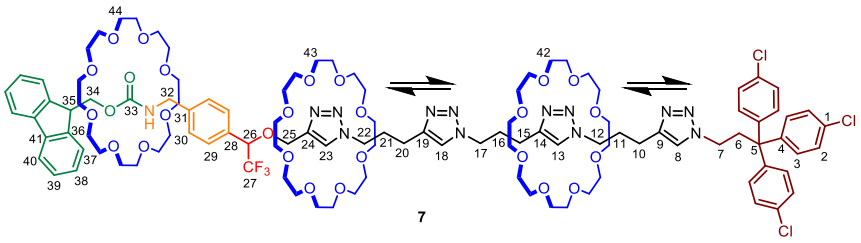
^{1}H NMR (600 MHz, $C_6D_5CD_3$) δ 8.28 (s, 1H, $H_{23}^{tri\text{-}tri}$), 8.16 (d, J = 7.7 Hz, 2H, $H_{30}^{car\text{-}tri}$), 7.98 (s, 1H, $H_{13}^{tri\text{-}tri}$/$_{18}^{tri\text{-}tri}$), 7.94 (s, 1H, $H_{13}^{tri\text{-}tri}$/$_{18}^{tri\text{-}tri}$), 7.85 (s, 1H, $H_8^{tri\text{-}tri}$), 7.71 (m, 3H, $H_{23}^{car\text{-}tri}$ + $H_{37}^{car\text{-}tri}$), 7.62 (s, 1H, $H_{13}^{car\text{-}tri}$/$_{18}^{car\text{-}tri}$), 7.58 (d, J = 7.6 Hz, 2H, $H_{40}^{tri\text{-}tri}$), 7.56 (d, J = 7.5 Hz, 2H, $H_{40}^{car\text{-}tri}$), 7.48 (m, 3H, $H_{13}^{car\text{-}tri}$/$_{18}^{car\text{-}tri}$ + $H_{29}^{car\text{-}tri}$), 7.45 (d, J = 7.5 Hz, 2H, $H_{37}^{tri\text{-}tri}$), 7.39 (d, J = 7.8 Hz, 2H, $H_{29}^{tri\text{-}tri}$), 7.23 (t, J = 7.5 Hz, 2H, $H_{39}^{car\text{-}tri}$ + $H_{39}^{tri\text{-}tri}$), 7.17 (t, J = 7.5 Hz, 2H, $H_{38}^{car\text{-}tri}$ + $H_{38}^{tri\text{-}tri}$), 7.09 – 7.04 (m, 7H, $H_2^{car\text{-}tri}$ + $H_2^{tri\text{-}tri}$ + $H_8^{car\text{-}tri}$), 6.94 (m, 8H, $H_3^{tri\text{-}tri}$ + $H_{30}^{tri\text{-}tri}$), 6.92 (br, 1H, $NH^{car\text{-}tri}$), 6.89 – 6.83 (m, 6H, $H_3^{car\text{-}tri}$), 4.97 (d, J = 4.4 Hz, 2H, $H_{32}^{car\text{-}tri}$), 4.88 (m, 3H, $H_{17}^{tri\text{-}tri}$/$_{22}^{tri\text{-}tri}$ + $H_{26}^{tri\text{-}tri}$), 4.75 (m, 3H, $H_{12}^{tri\text{-}tri}$ + $H_{26}^{car\text{-}tri}$), 4.68 (d, J = 12.3 Hz, 1H, $H_{25}^{car\text{-}tri}$ + $H_{25}^{tri\text{-}tri}$), 4.57 (m, 1H, $H_{25}^{car\text{-}tri}$ + $H_{25}^{tri\text{-}tri}$), 4.53 (m, 3H, $NH^{tri\text{-}tri}$ + $H_{34}^{car\text{-}tri}$), 4.44 (d, J = 6.3 Hz, 2H, $H_{34}^{tri\text{-}tri}$), 4.39 (t, J = 7.7 Hz, 2H, $H_{17}^{car\text{-}tri}$/$_{22}^{car\text{-}tri}$), 4.34 (t, J = 7.2 Hz, 2H, $H_{17}^{tri\text{-}tri}$/$_{22}^{tri\text{-}tri}$), 4.30 (m, 2H, $H_{12}^{car\text{-}tri}$), 4.25 (t, J = 7.2 Hz, 2H, $H_{17}^{car\text{-}tri}$/$_{22}^{car\text{-}tri}$), 4.14 (t, J = 6.7 Hz, 1H, $H_{35}^{car\text{-}tri}$), 4.00 (d, J = 6.4 Hz, 2H, $H_{32}^{tri\text{-}tri}$), 3.97 (m, 1H, $H_{35}^{tri\text{-}tri}$), 3.86 (m, 2H, $H_7^{tri\text{-}tri}$), 3.69 (m, 2H, $H_7^{car\text{-}tri}$), 3.47 – 3.14 (m, 66H, $H_{10}^{tri\text{-}tri}$ + $H_{15}^{tri\text{-}tri}$/$_{20}^{tri\text{-}tri}$ + $H_{42}^{car\text{-}tri}$ + $H_{42}^{tri\text{-}tri}$ + $H_{43}^{car\text{-}tri}$ + $H_{43}^{tri\text{-}tri}$), 3.00 (m, 2H, $H_6^{tri\text{-}tri}$), 2.88 (m, 2H, $H_6^{car\text{-}tri}$), 2.81 (t, J = 7.4 Hz, 2H, $H_{15}^{car\text{-}tri}$/$_{20}^{car\text{-}tri}$), 2.77 (t, J = 7.5 Hz, 4H, $H_{10}^{car\text{-}tri}$ + $H_{15}^{tri\text{-}$

$^{tri}{}_{/20}{}^{tri\text{-}tri}$), 2.72 (t, J = 7.3 Hz, 2H, $H_{15}{}^{car\text{-}tri}{}_{/20}{}^{car\text{-}tri}$), 2.66 (m, 2H, $H_{16}{}^{tri\text{-}tri}{}_{/21}{}^{tri\text{-}tri}$), 2.58 (m, 2H, $H_{11}{}^{tri\text{-}tri}$), 2.36 – 2.18 (m, 8H, $H_{11}{}^{car\text{-}tri}$ + $H_{16}{}^{car\text{-}tri}$ + $H_{21}{}^{car\text{-}tri}$ + $H_{16}{}^{tri\text{-}tri}{}_{/21}{}^{tri\text{-}tri}$).

^{13}C NMR (151 MHz, $C_6D_5CD_3$) δ 156.62, 147.13, 146.97, 146.36, 145.34, 144.53, 144.25, 143.48, 142.14, 141.87, 132.90, 132.79, 132.52, 130.53, 130.36, 129.23, 128.83, 127.56, 127.28, 127.19, 126.06, 125.45, 123.95, 122.44, 122.19, 120.21, 119.92, 71.08, 71.01, 70.74, 65.59, 63.38, 54.58, 53.27, 50.14, 50.03, 49.75, 49.41, 48.99, 48.90, 48.24, 47.02, 45.46, 40.50, 30.29, 29.98, 29.85, 29.60, 22.85, 22.72, 22.67.

HRMS (ESI$^+$): m/z = 1934.7914 $[M+Na]^+$ (calcd. 1934.7968 for $C_{95}H_{123}N_{13}O_{19}Cl_3F_3Na$).

Synthesis of **7**



7

To a solution of **10** (39 mg, 20 µmol, 1 eq.) and **3** (65 mg, 180 µmol, 9 eq.) in toluene (88 µL) was added *i*-Pr_2NH (43 µL, 306 µmol, 15 eq.). A solution of **2** (83 mg, 230 µmol, 11.5 eq.) in toluene (500 µL total volume) was continuously added via syringe pump at r.t. for 23 h. CF_3COOH (63 µL) and saturated aqueous NH_4Cl (10 mL) were added to the mixture and it was extracted with CH_2Cl_2 (3 × 10 mL). The combined organic layers were concentrated under reduced pressure. The residue was purified by PTLC (SiO_2, CH_2Cl_2/5% MeOH), followed by SEC (Bio-Beads S-X1, CH_2Cl_2) and then PTLC (SiO_2, CH_2Cl_2/5% MeOH) to give **7** as a colourless film (2.5 mg, 1.1 µmol, 6%).

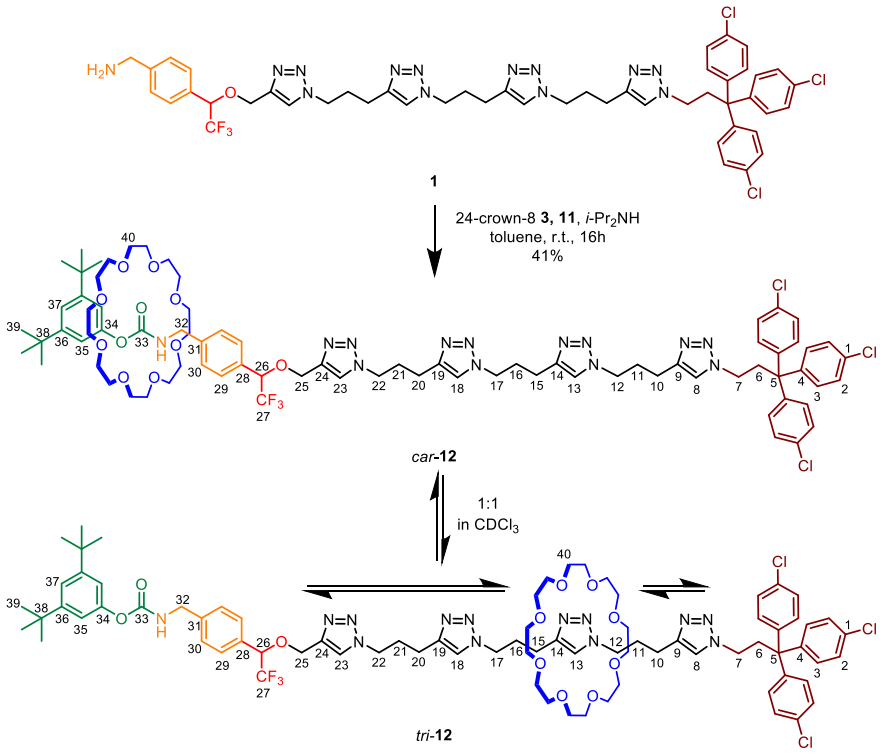
^{1}H NMR (600 MHz, $C_6D_5CD_3$) δ 8.38 (s, 1H, H_{23}), 8.16 (d, J = 7.8 Hz, 2H, H_{30}), 8.01 (s, 1H, $H_{8/13/18}$), 7.94 (s, 1H, $H_{8/13/18}$), 7.86 (s, 1H, $H_{8/13/18}$), 7.71 (d, J = 7.4 Hz, 2H, H_{37}), 7.56 (d, J = 7.5 Hz, 2H, H_{40}), 7.51 (d, J = 7.9 Hz, 2H, H_{29}), 7.23 (t, J = 7.4 Hz, 2H, H_{39}), 7.17 (m, 2H, H_{38}), 7.06 (m, 6H, H_2), 6.94 (m, 6H, H_3), 6.91 (br, 1H, NH), 4.98 (d, J = 4.4 Hz, 2H,

H_{32}), 4.92 (m, 2H, $H_{12/17/22}$), 4.78 (m, 3H, $H_{12/17/22}$ + H_{26}), 4.74 (m, 1H, H_{25}), 4.58 (d, J = 11.7 Hz, 1H, H_{25}), 4.53 (d, J = 6.9 Hz, 2H, H_{34}), 4.32 (t, J = 7.2 Hz, 2H, $H_{12/17/22}$), 4.15 (t, J = 6.9 Hz, 1H, H_{35}), 3.86 (m, 2H, H_7), 3.46 – 3.13 (m, 98H, H_{42} + H_{43} + H_{44} + $H_{10/15/20}$ + $H_{10/15/20}$), 2.99 (m, 2H, H_6), 2.79 (t, J = 7.3 Hz, 2H, $H_{10/15/20}$), 2.72 (m, 2H, $H_{11/16/21}$), 2.58 (p, J = 8.0 Hz, 2H, $H_{11/16/21}$), 2.32 (p, J = 7.5 Hz, 2H, $H_{11/16/21}$).

^{13}C NMR (151 MHz, $C_6D_5CD_3$) δ 156.58, 148.79, 148.50, 145.84, 145.38, 144.55, 142.77, 142.04, 141.88, 132.97, 132.81, 132.50, 130.55, 130.36, 130.27, 128.85, 127.74, 127.53, 127.20, 126.10, 125.05, 123.13, 122.86, 122.76, 119.92, 78.86, 71.14, 71.08, 71.03, 70.77, 65.61, 63.48, 54.68, 50.11, 49.75, 49.07, 48.28, 47.01, 45.51, 40.47, 30.32, 29.82, 29.13, 23.39, 23.15, 23.04.

HRMS (ESI$^+$): m/z = 2265.0212 $[M+H]^+$ (calcd. 2265.0246 for $C_{111}H_{156}N_{13}O_{27}Cl_3F_3$).

Synthesis of [2]rotaxane **12**



To a solution of **1** (6.0 mg, 6.1 μmol, 1 eq.) and **3** (21.5 mg, 61 μmol, 10 eq.) in toluene (27 μL) was added *i*-Pr_2NH (13 μL, 92 μmol, 15 eq.). A solution of **11** (13 mg, 34 μmol, 6 eq.) in toluene (129 μL total volume) was continuously added via syringe pump at r.t. for 16 h. CF_3COOH (16 μL) and saturated aqueous NH_4Cl (5 mL) were added to the mixture and it was extracted with CH_2Cl_2 (3 × 5 mL). The combined organic extracts were concentrated under reduced pressure and the residue was purified by PTLC (SiO_2, CH_2Cl_2/5% MeOH) followed by SEC (Bio-Beads S-X1, CH_2Cl_2) to give **12** as a colourless film (3.9 mg, 2.5 μmol, 41%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.97 (m, 3H, H_{30}^{car} + H_{23}^{tri}), 7.80 (s, 1H, H_{18}^{tri}/$_{13}^{tri}$), 7.71 (s, 1H, H_{18}^{tri}/$_{13}^{tri}$), 7.70 (s, 1H, H_{23}^{car}), 7.48 (d, *J* = 7.9 Hz, 2H, H_{29}^{tri}), 7.47 – 7.39 (m, 7H, H_{29}^{car} + H_{30}^{tri} + H_{18}^{car} + H_{13}^{car} + H_8^{tri}), 7.29 (m, 6H, H_2^{car} + H_2^{tri}), 7.26 (s, 1H, H_{37}^{tri}), 7.22 (d, *J* = 8.5 Hz, 6H, H_3^{car} + H_3^{tri}), 7.20 (s, 2H, H_8^{car} + H_{37}^{car}), 6.99 (m, 2H, H_{35}^{car}), 6.97 (m, 2H, H_{35}^{tri}), 5.45 (t, *J* = 6.1 Hz, 1H, NH^{tri}), 4.85 (q, *J* = 6.6 Hz, 1H, H_{26}^{tri}), 4.80 – 4.72 (m, 5H, H_{32}^{car} + H_{26}^{car} + H_{25}^{car} + H_{25}^{tri}), 4.61 (m, 4H, H_{25}^{car} + H_{25}^{tri} + H_{22}^{tri}), 4.51 – 4.43 (m, 8H, H_{32}^{tri} + H_{12}^{tri} + H_{17}^{tri} + H_{22}^{car}), 4.39 (dt, *J* = 9.4, 6.8 Hz, 4H, H_{12}^{car} + H_{17}^{car}), 4.01 (m, 2H, H_7^{car} + H_7^{tri}), 3.63 – 3.31 (m, 32H, H_{40}^{car} + H_{40}^{tri}), 3.16 (m, 2H, H_6^{car}), 3.13 (m, 2H, H_6^{tri}), 2.84 (t, *J* = 7.6 Hz, 2H, H_{20}^{tri}), 2.80 – 2.66 (m, 6H, H_{20}^{car} + H_{15}^{car} + H_{15}^{tri} + H_{10}^{car} + H_{10}^{tri}), 2.42 (m, 2H, H_{21}^{tri}), 2.37 – 2.24 (m, 6H, H_{21}^{car} + H_{16}^{car} + H_{16}^{tri} + H_{11}^{car} + H_{11}^{tri}), 1.31 (s, 18H, H_{39}^{tri}), 1.30 (s, 18H, H_{39}^{car}).

The signal of NH^{car} couldn't be identified.

^{13}C NMR (151 MHz, $CDCl_3$) δ 155.16, 152.24, 151.62, 147.17, 146.60, 146.10, 143.80, 143.72, 143.11, 139.93, 133.08, 133.03, 130.11, 130.09, 129.05, 128.93, 128.89, 128.10, 124.43, 122.66, 122.51, 122.23, 121.84, 121.80, 121.74, 119.64, 116.53, 115.88, 70.96, 70.67, 63.40, 63.21, 54.58, 54.56, 49.72, 49.43, 49.32, 49.21, 49.13, 47.50, 47.36, 45.07,

40.56, 40.51, 35.09, 35.01, 31.59, 31.53, 29.61, 29.51, 29.42, 29.04, 22.84, 22.64, 22.41, 22.34, 22.23.

HRMS (ESI$^+$): *m/z* = 1592.6610 $[M+Na]^+$ (calcd. 1592.6653 for $C_{79}H_{101}N_{13}O_{11}Cl_3F_3Na$).

3.4.3. Model study and operation

3.4.3.1. Rotaxane formation study

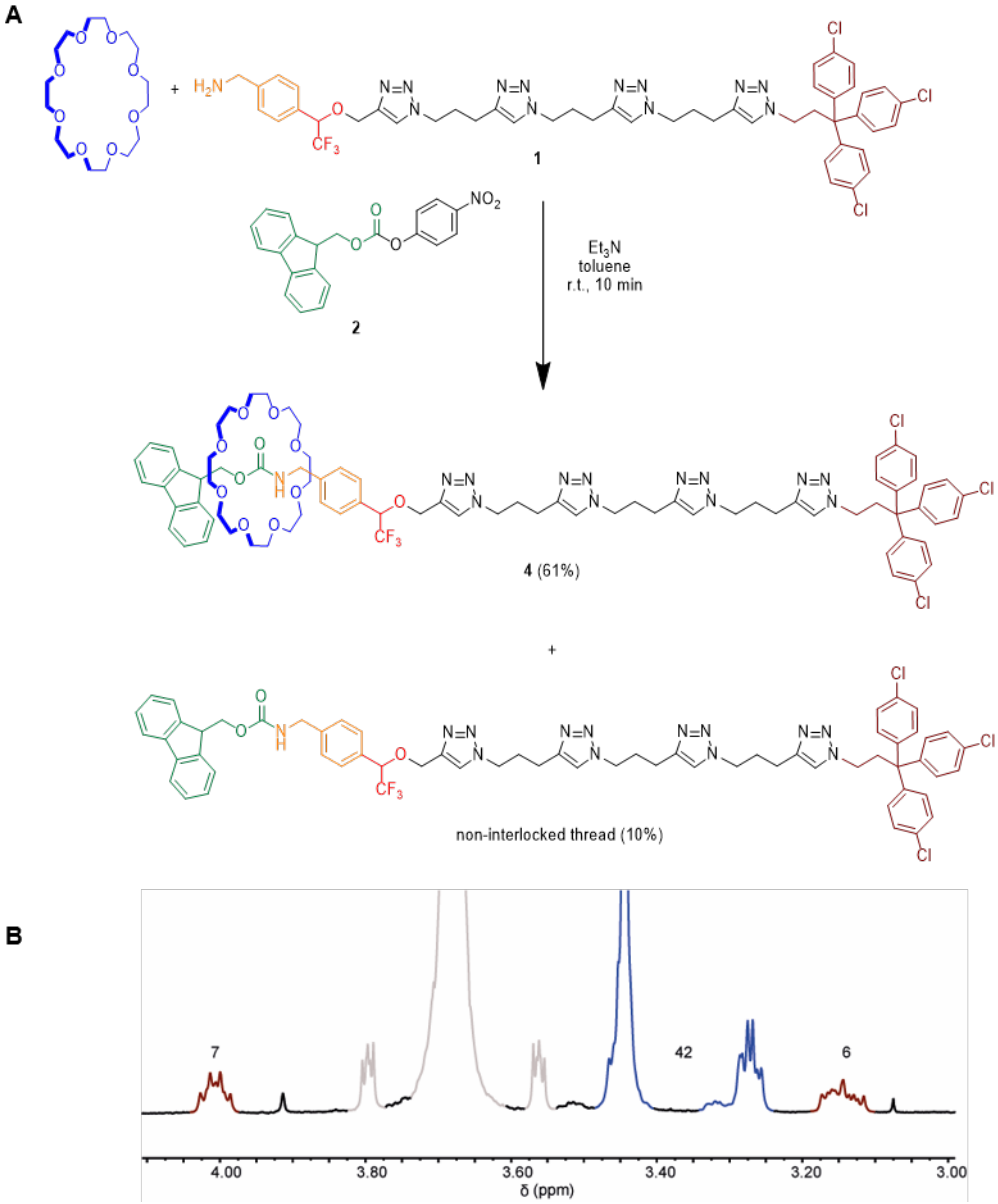


Figure S4. Rotaxane formation study. (**A**) Reagents and conditions: Pump **1**, fuel **2**, 24-crown-8 **3**, Et_3N, toluene, r.t., 10 min. (**B**) Partial ^1H NMR ($CDCl_3$, 600 MHz, 298 K) of the reaction mixture after 10 minutes.

To a solution of pump **1** (11.2 mg, 11 μmol, 1 eq.) in dry toluene (50 μL) was added 24-crown-8 **3** (40 mg, 114 μmol, 10 eq.), Et_3N (7.9 μL, 57 μmol, 5 eq.) and fuel **2** (13.0 mg, 36 μmol, 3.3 eq.). The resultant mixture was stirred at r.t.. An aliquot (*ca.* 2 μL) of the mixture was taken after 10 minutes and immediately diluted with $CDCl_3$ (0.5 mL) and analyzed by ^1H NMR. The extent of rotaxane **4** (61%) and non-interlocked thread (10%) formation was calculated by comparison of ^1H NMR (600 MHz, 298 K) peak integrals for protons H_6, H_{35} and H_{42}. The amount of rotaxane did not increase using longer reaction times. In the timescale of 10 minutes, the Fmoc removal process is so slow that it is negligible. The particular base used (Et_3N or *i*-Pr_2NH) does not affect the result.

3.4.3.2. Macrocycle displacement study for [2]rotaxane **4**, [3]rotaxane **10** and [4]rotaxane **7**

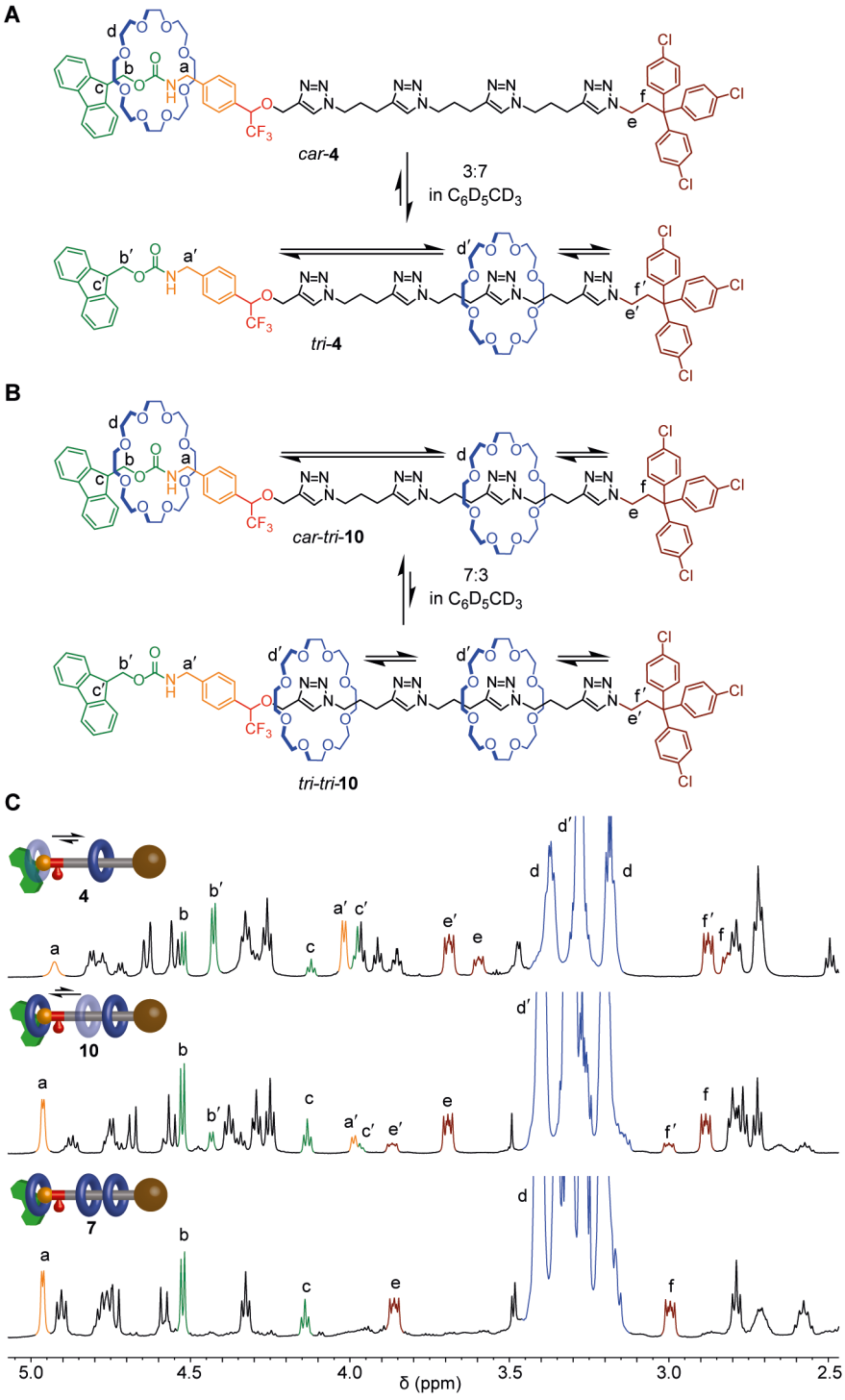


Figure S5. Macrocycle displacement study of **4**, **10** and **7**. Displacement equilibrium of (**A**) [2]rotaxane **4** and (**B**) [3]rotaxane **10** in $C_6D_5CD_3$. [4]rotaxane **7** takes only one co-conformation. (**C**) Partial ^{1}H NMR (600 MHz, 295 K, $C_6D_5CD_3$) spectra of **4**, **10** and **7**.

4, **10** or **7** (*ca*. 2 mg each) were left to equilibrate in $C_6D_5CD_3$ (0.5 mL) overnight at r.t. before recording ^{1}H NMR (600 MHz, 298 K) spectra. The displacement equilibrium was

calculated by comparison of the peak integrals for $H_{b/b'}$ and $H_{e/e'}$, which are clearly resolved from other peaks for both co-conformers.

3.4.3.3. Barrier removal study

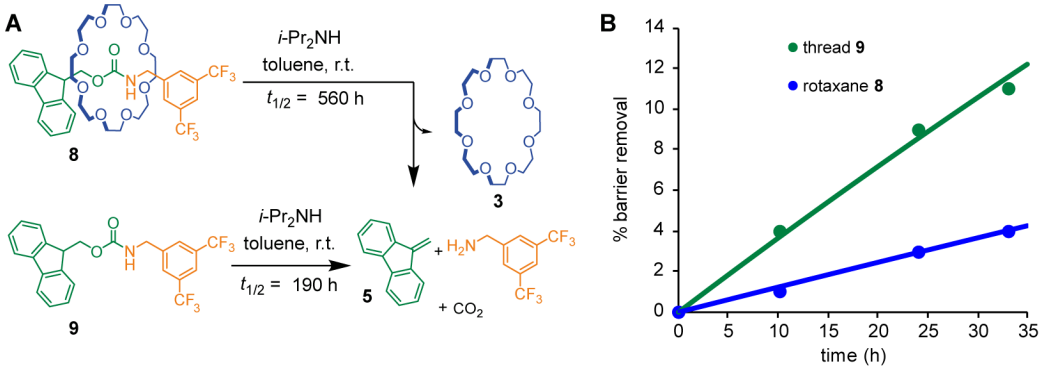


Figure S6. Barrier removal study. (**A**) Reagents and conditions for barrier removal of model [2]rotaxane **8** and model non-interlocked thread **9**: **8** or **9**, *i*-Pr_2NH, toluene, r.t.. (**B**) Extent of barrier removal over time. Solid lines represent data fitting according to a kinetic model of a pseudo-first order reaction.

To a solution of **8** (5.7 mg, 7 µmol, 1 eq.) or **9** (3.3 mg, 7 µmol, 1 eq.) in toluene (50 µL) was added *i*-Pr_2NH (9.8 µL, 70 µmol, 10 eq.). The reaction mixture was stirred at r.t.. An aliquot (*ca.* 2 µL) of the mixture was taken at 10 h, 24 h and 33 h and immediately diluted with $CDCl_3$ (0.5 mL) for ^1H NMR analysis. The extent of barrier removal was calculated by comparison of ^1H NMR (600 MHz, 298 K) peak integrals for the protons of H_8 of rotaxane **8** and thread **9** with that of the alkene peak of dibenzofulvene decomposition product **5**[49].

3.4.3.4. Fuel decomposition study

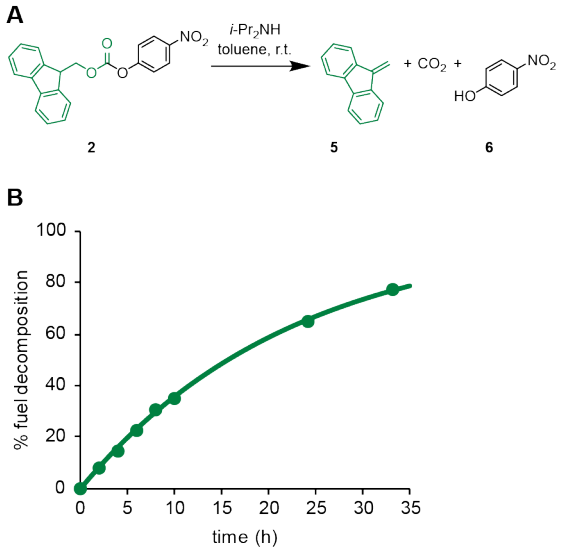


Figure S7. Fuel decomposition study. (**A**) Reagents and conditions: **2**, *i*-Pr_2NH, toluene, r.t.. (**B**) Extent of decomposition over time. The solid line represents data fitting according to a kinetic model of a pseudo-first order reaction.

To a solution of **2** (2.5 mg, 7 µmol) in dry toluene (50 µL) was added *i*-Pr_2NH (9.8 µL, 70 µmol, 10 eq.) The resultant mixture was stirred at r.t.. An aliquot (*ca.* 2 µL) of the mixture was taken at 2 h, 4 h, 6 h, 8 h, 10 h, 24 h and 33 h and immediately diluted with $CDCl_3$ (0.5 mL) for ^1H NMR analysis. The extent of fuel decomposition was calculated by comparison of ^1H NMR (600 MHz, 298 K) peak integrals for the H_6 protons of **2** with that of the alkene peak of dibenzofulvene decomposition product **5**.

3.4.3.5. Dethreading study

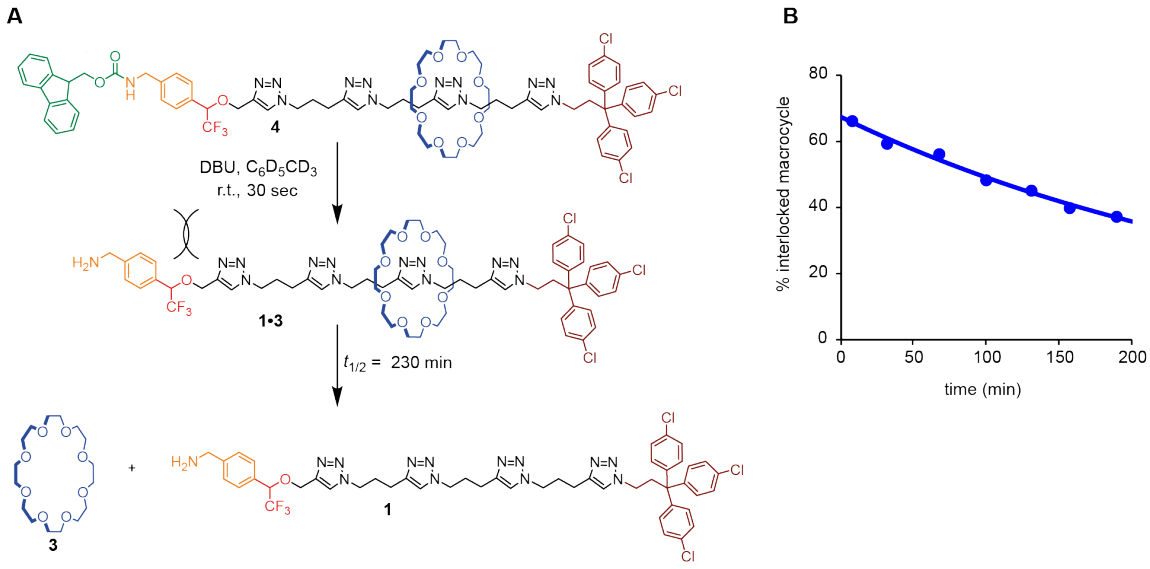


Figure S8. Dethreading study. (**A**) Reagents and conditions: rotaxane **4**, DBU, $C_6D_5CD_3$, r.t., 30 sec. Then dilution with $C_6D_5CD_3$ for ^{1}H NMR (600 MHz, 298 K) analysis. (**B**) Amount of pseudo-rotaxane **1·3** over time. The solid line represents data fitting according to a kinetic model of a first order reaction.

To a solution of **4** (3.3 mg, 2.1 μmol) in $C_6D_5CD_3$ (100 μL) was added DBU (16 μL, 105 μmol, 50 eq.) and the reaction mixture was vigorously mixed for 30 sec at r.t.. The mixture was then diluted with $C_6D_5CD_3$ (400 μL) for ^{1}H NMR analysis at the indicated time intervals. The extent of dethreading of **1·3** was calculated by comparison of ^{1}H NMR (600 MHz, 298 K) peak integrals for the protons of non-interlocked **3** with that of the alkene peak of dibenzofulvene decomposition product **5**.

3.4.3.6. Equilibrium threading experiment

This study was performed to determine whether threading of **3** onto **1** was thermodynamically favorable and thus occurred at equilibrium without the need for fuel **2**. No formation of pseudo-rotaxane **1·3** was observed by ^{1}H NMR ($C_6D_5CD_3$) analysis of a mixture of **1** and **3** that had equilibrated over 3 days at r.t. under analogous conditions for the synthesis of [2]rotaxane **4**. Macrocycle threading thus only occurs by active templated

threading, i.e. as a consequence of the stabilization of the transition state for the reaction of **1** and **2** within the cavity of **3**.

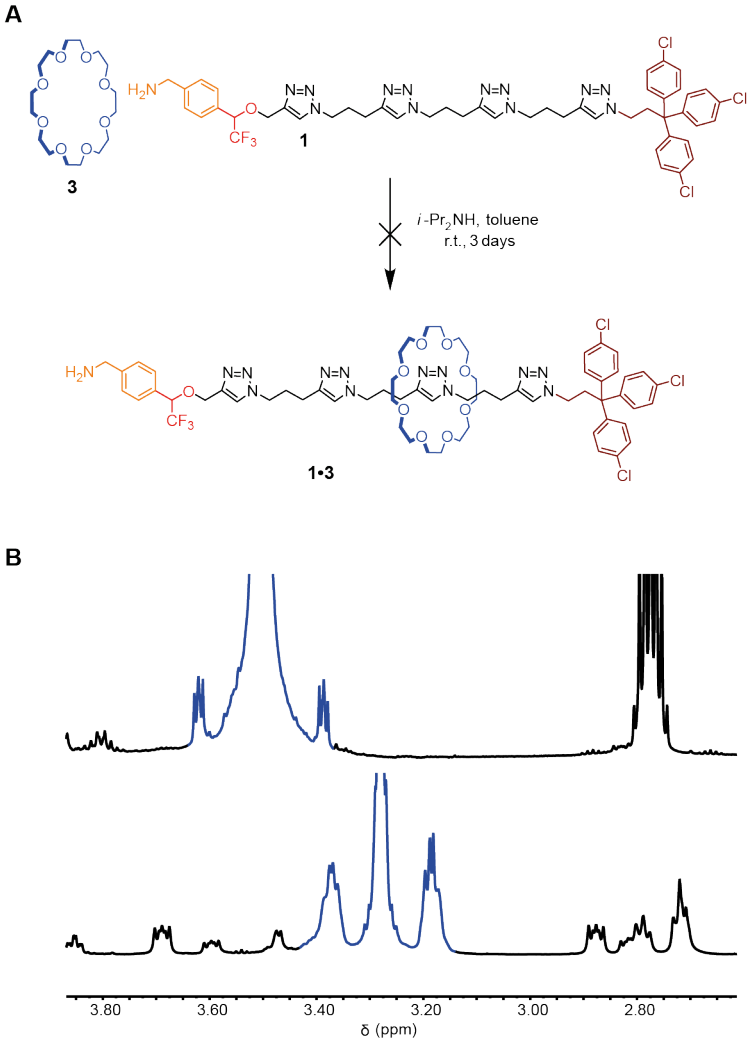


Figure S9. Equilibrium threading experiment. (**A**) Reagents and conditions: Pump **1**, 24-crown-8 **3**, *i*-Pr_2NH, toluene, r.t., 3 days. (**B**) Partial ^1H NMR spectrum ($C_6D_5CD_3$, 600 MHz, 298 K) of macrocycle (shown in blue) contained in the reaction mixture (above) and threaded in **4** (below). No threaded macrocycle is observed in the equilibrium threading experiment.

A solution of **1** (4.9 mg, 5.0 µmol, 1 eq.), **3** (20.5 mg, 58 µmol, 12 eq.) and *i*-Pr_2NH (10.5 µL, 75 µmol, 15 eq.) in toluene (22 µL) was kept at r.t. for 3 days. An aliquot (*ca.* 2 µL) of the mixture was diluted with $C_6D_5CD_3$ (500 µL) and immediately analysed by ^1H NMR (600 MHz, 298 K). Signals corresponding to interlocked macrocycle residing in the oligotriazole region at 3.28 ppm (such as with *tri*-**4**) were not observed.

3.4.3.7. Operation with reagent **11**

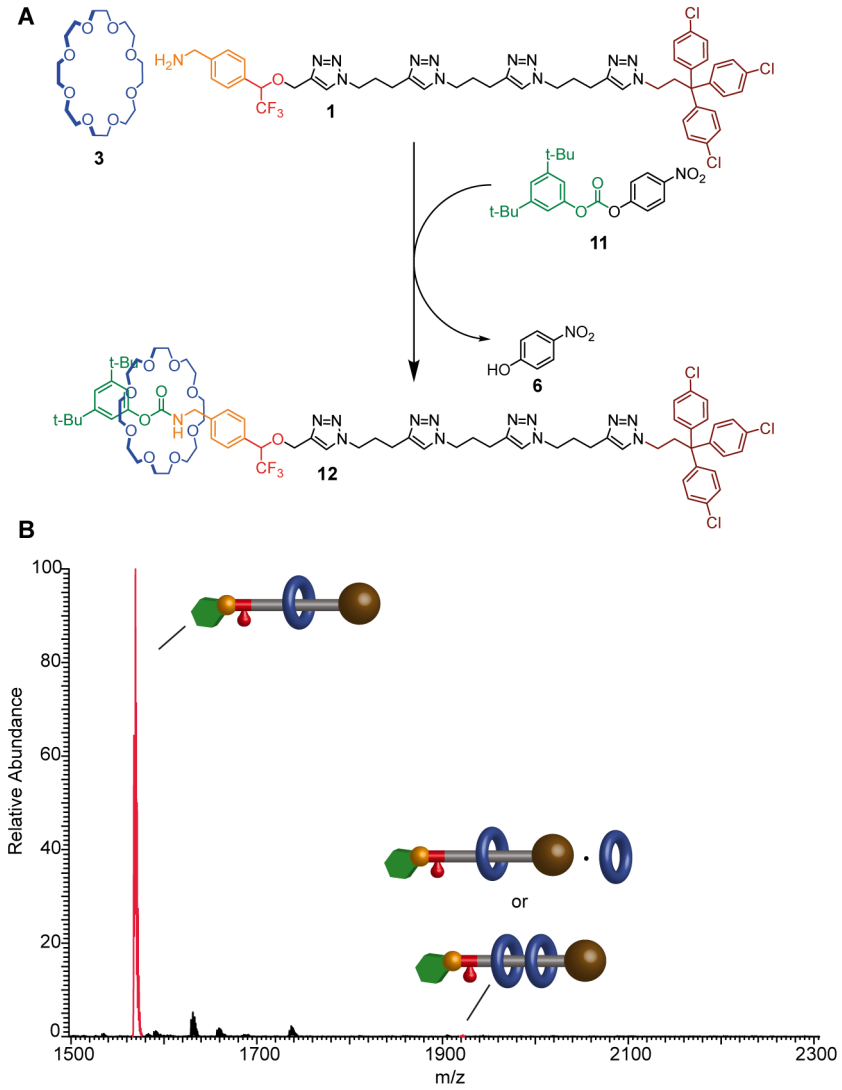


Figure S10. Operation with reagent **11**. (**A**) Reagents and conditions: pump **1**, 24-crown-8 **3**, *i*-Pr_2NH, toluene, reagent **11** (added continuously via syringe pump as a solution in toluene), r.t., 16 h. (**B**) Electrospray ionisation mass spectrum of the crude reaction mixture.

This model experiment was performed in order to demonstrate that multiple fuel cycles need to occur for more than one macrocycle to be threaded. Pump **1** can only react with **11** once, as the resultant carbamate does not decompose under the reaction conditions. Thus, [2]rotaxane **12** should be the only interlocked product of this reaction.

The procedure of operation and workup is the same as described in the synthesis of **12**. The crude residue was analysed by ESI-MS prior to purification by PTLC. The major signals

observed corresponded to [**12**+H]$^+$ and [**12**+Na]$^+$ along with a very weak signal (<1% relative abundance versus [**12**+H]$^+$) with an *m/z* ratio consistent with [{**12**•**3**}+H]$^+$. It could not be determined whether this species was a threaded adduct, i.e. a [3]rotaxane, as it could not be isolated.

3.4.3.8. Different methods of operation

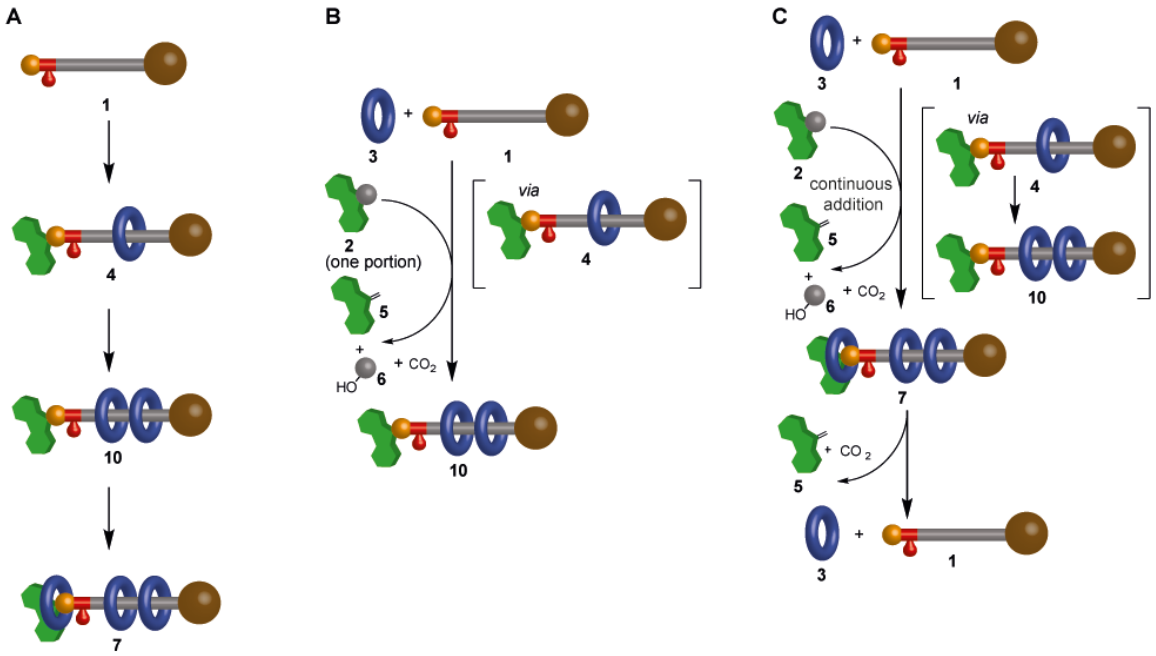


Figure S11. Methods of operation. (**A**) Stepwise. (**B**) Autonomous. (**C**) Continuous fuel addition.

A. Stepwise Operation: Carried out according to experimental procedures in Section 3.4.2.4. Products analyzed by ^{1}H NMR/^{13}C NMR and HRMS (ESI$^+$).

B. Autonomous operation (Section 3.4.3.9.): performed by adding fuel in one portion. Analyzed by ESI$^+$.

C. Operation with continuous fuel addition (Section 3.4.3.10.): performed by adding fuel by syringe pump. Analyzed by ESI$^+$.

3.4.3.9. Autonomous operation

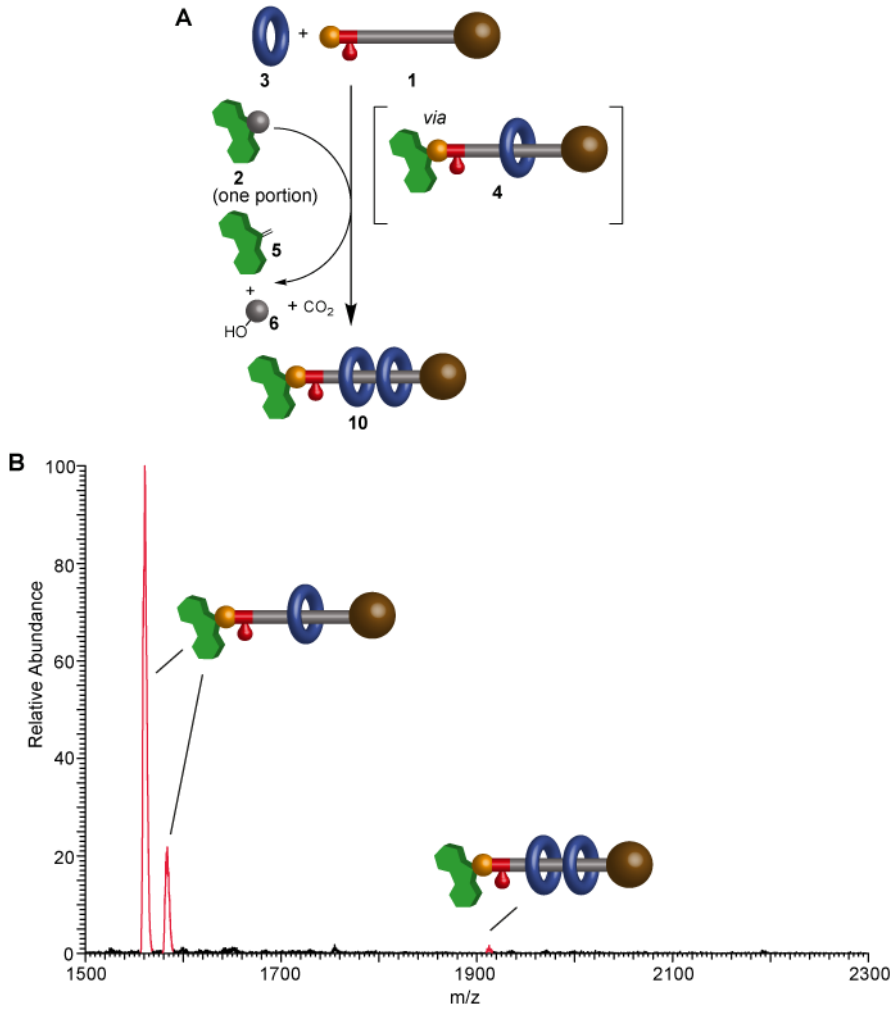


Figure S12. Autonomous operation with chemical fuel **2**. (**A**) Reagents and conditions: pump **1**, 24-crown-8 **3**, *i*-Pr_2NH, toluene, fuel **2** (added as one portion at the beginning of the reaction), r.t., 17 h. (**B**) Electrospray ionisation mass spectrum of the crude reaction mixture.

To a solution of **1** (5.6 mg, 5.7 µmol, 1 eq.) and **3** (21 mg, 57 µmol, 10 eq.) in dry toluene (25 µL) was added *i*-Pr_2NH (12 µL, 88 µmol, 15 eq.). A solution of **2** (24.8 mg, 69 µmol, 12 eq.) in toluene (134 µL) was added as one portion and the reaction was continued at r.t. for 17 h after which time analysis by ESI-MS indicated the presence of [2]rotaxane **4** and [3]rotaxane **10**.

3.4.3.10. Operation with continuous fuel addition

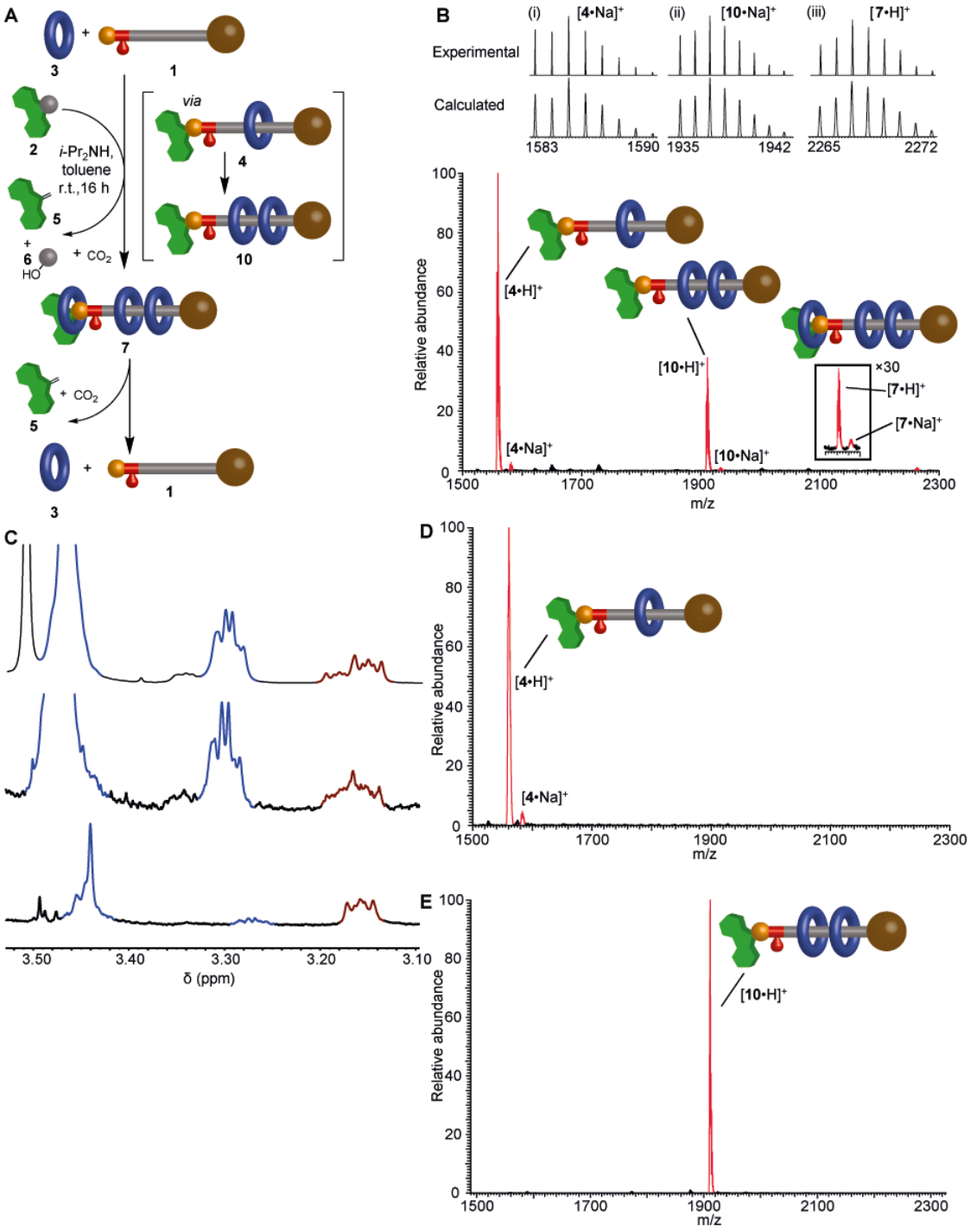


Figure S13. Operation with continuous addition of chemical fuel **2**. (**A**) Reagents and conditions: pump **1**, 24-crown-8 **3**, *i*-Pr_2NH, toluene, fuel **2** (added continuously via syringe pump as a solution in toluene), r.t., 16 h. After stopping the addition of fuel **2**, half of the reaction mixture was left at r.t. for 19 days. (**B**) Electrospray ionisation mass spectrum of the crude reaction mixture immediately after stopping the addition of **2**. (**C**) Stack plot of partial ^{1}H NMR spectra ($CDCl_3$, 600MHz, 298K) of purified rotaxane **4** (top) reaction mixture just after stopping the addition of fuel **2** (middle) and after 19 days (bottom). Blue and brown peaks correspond to the interlocked macrocycle and proton H_f, respectively. (**D**) Electrospray ionisation mass spectrum of the mixture of **4** and **3**. (**E**) Electrospray ionisation mass spectrum of the mixture of **10** and **3**.

To a solution of **1** (5.8 mg, 5.9 μmol, 1 eq.) and **3** (21 mg, 60 μmol, 10 eq.) in dry toluene (27 μL) was added *i*-Pr_2NH (12 μL, 88 μmol, 15 eq.). A solution of **2** (23 mg, 63 μmol, 11 eq.) in toluene (133 μL total volume) was continuously added via syringe pump at r.t. for 16 h. An aliquot (*ca.* 2 μL) of the mixture was diluted with $CDCl_3$ (500 μL) and immediately analysed by ^1H NMR (600 MHz, 298 K) (**Figure S13**, **C**, middle). Comparison of this spectrum with the ^1H NMR (600 MHz, 298 K) of purified **4** in $CDCl_3$ (**Figure S13**, **C**, top) shows peaks at 3.30 and 3.46 ppm were present, corresponding to the interlocked macrocycle.

The reaction was then partitioned into two further equal aliquots (each approximately 100 μL).

To one aliquot was added CF_3COOH (10 μL) and saturated aqueous NH_4Cl (5 mL). The mixture was extracted with CH_2Cl_2 (3 × 5 mL). The combined organic extracts were concentrated under reduced pressure. The residue was analysed by ESI-MS (**Figure S13**, **B**), which indicated the presence of [2]rotaxane **4**, [3]rotaxane **10** and [4]rotaxane **7**.

In order to verify that the signals at *m/z* = 1912–1919 and *m/z* = 2265–2272 corresponded to the adducts [**10**+H]$^+$ and [**7**+H]$^+$ rather than non-interlocked adducts, such as [{**4**·**3**}+H]$^+$ or [{**10**·**3**}+H]$^+$, control experiments were performed whereby a mixture of pre-formed **4** or **10** (1 eq.) and **3** (16 eq.) was analysed by ESI-MS. No signals were observed at *m/z* = 1912–1919 or *m/z* = 2265–2272 respectively (**Figure S13**, **D** and **E**), therefore proving that [3]rotaxane **10** and [4]rotaxane **7** are indeed formed during autonomous operation.

The other aliquot was left to react at r.t. for an additional 19 days. Additional toluene was occasionally added to maintain a constant reaction volume. After this time, the aliquot was diluted with $CDCl_3$ (500 μL) and analysed by ^1H NMR (600 MHz, 298 K) (**Figure S13**, **C**, bottom).

Immediately after addition of fuel **2** was stopped, analysis by ^{1}H NMR (**Figure S13**, **C**, middle) showed that 1.17 eq. of 24-crown-8 **3** (based on the initial amount of pump **1**) had been pumped into either rotaxanes **4**, **10** or **7**. Analysis of the aliquot that had reacted for a further 19 days by ^{1}H NMR showed the interlocked macrocycle peaks at 3.30 and 3.46 ppm had reduced in intensity. Measuring the integral of these peaks revealed that only 0.13 eq. of 24-crown-8 **3** remained pumped in **4**, **10** or **7**, i.e. approximately 90% of pumped **3** had been released back into solution. This demonstrates that the threaded products dissipate back to the original unthreaded components under the basic reaction conditions.

3.4.4. NMR spectra

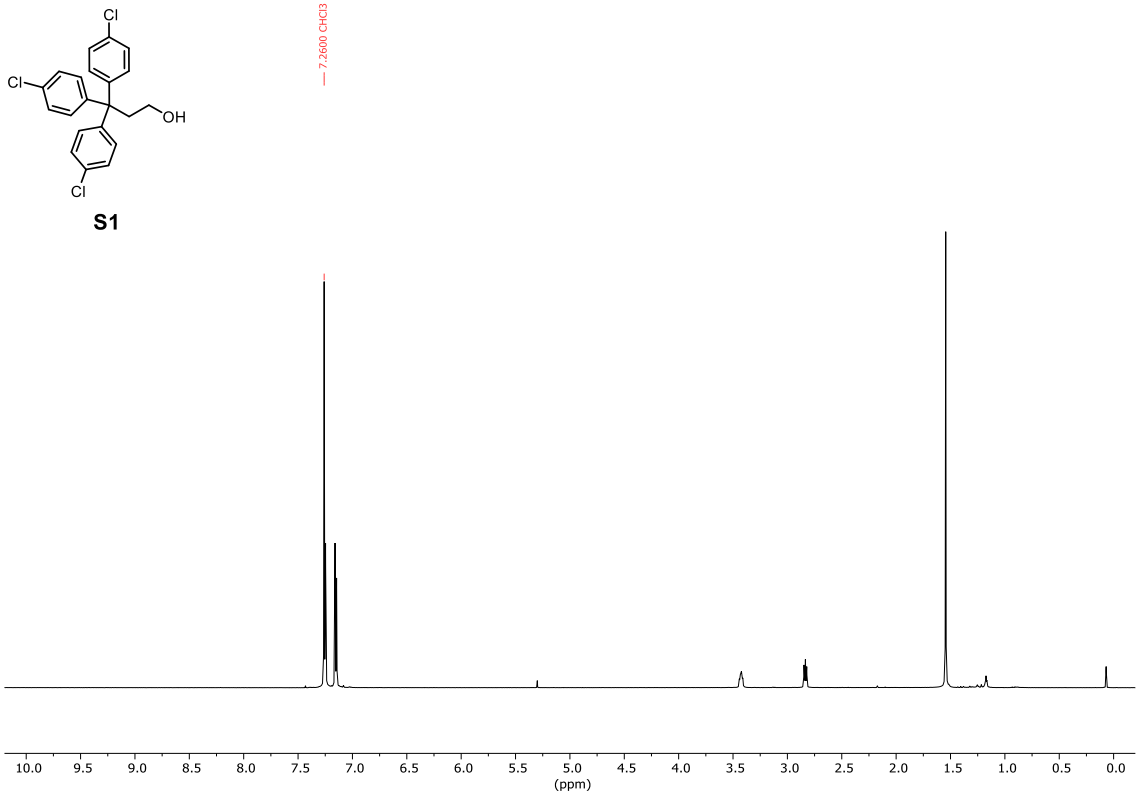


Figure S14. ^{1}H NMR ($CDCl_3$, 600 MHz, 298 K) of **S1**.

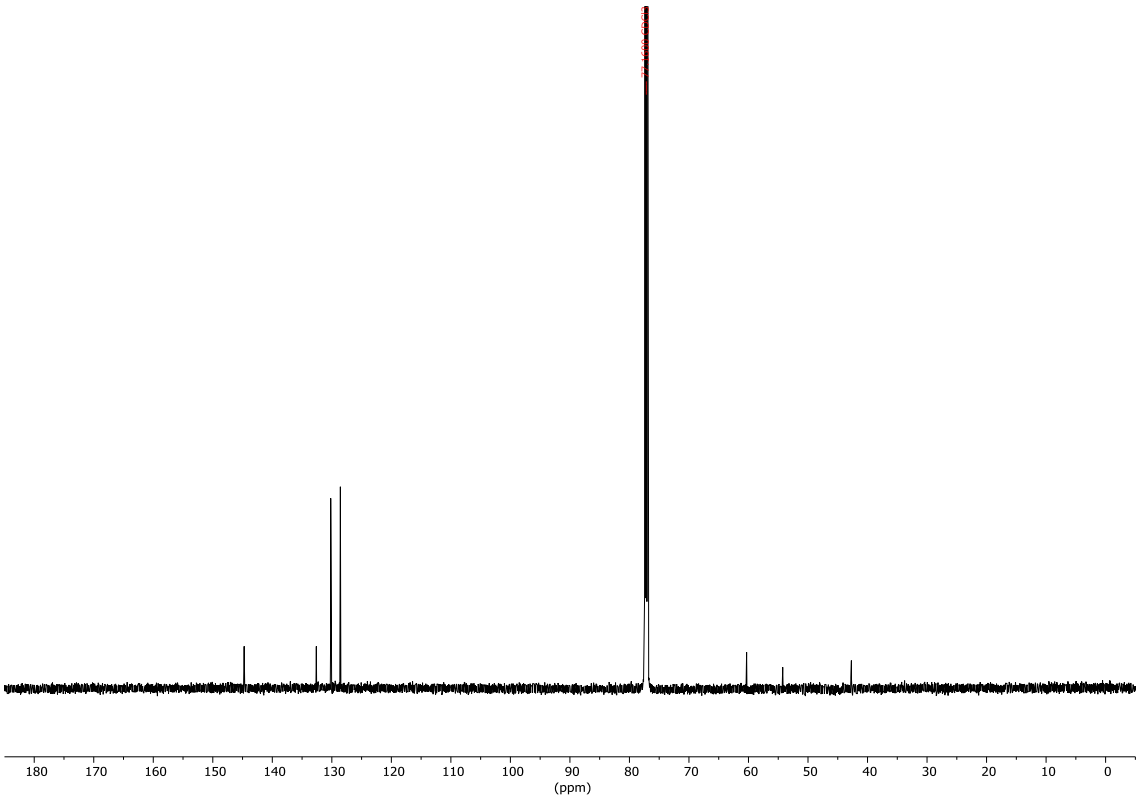


Figure S15. ^{13}C NMR ($CDCl_3$, 151 MHz, 298 K) of **S1**.

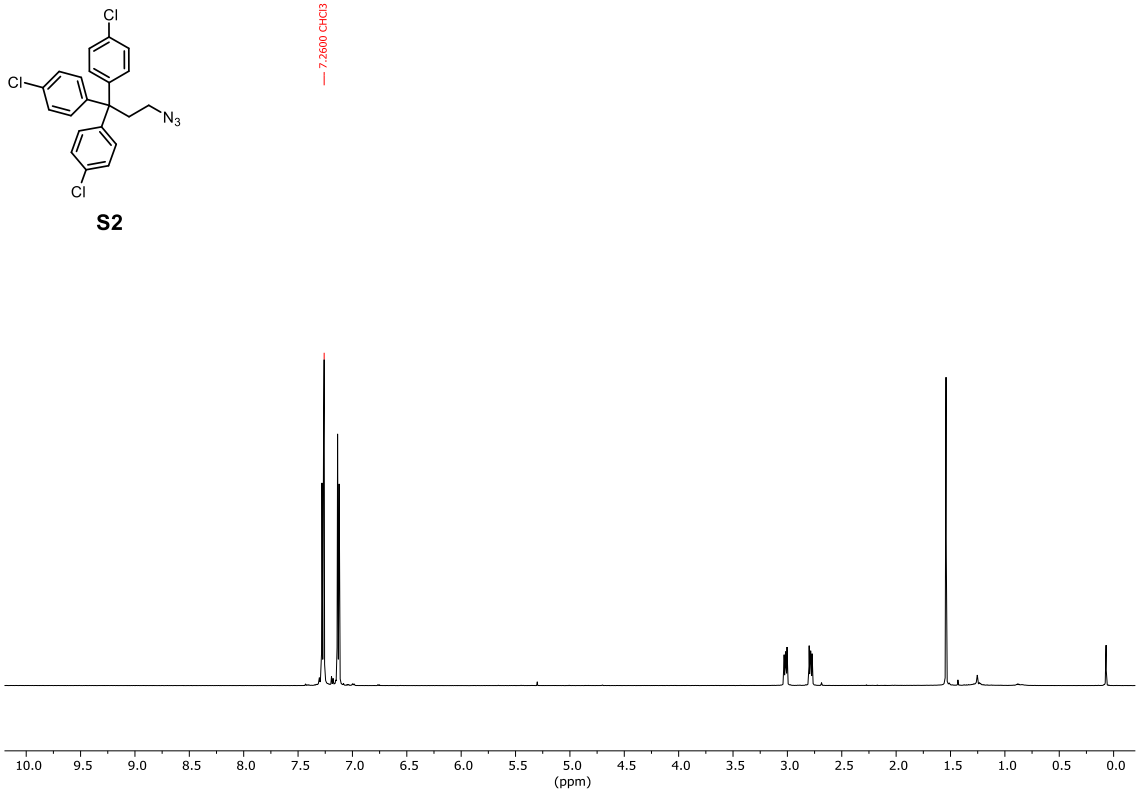


Figure S16. ^{1}H NMR ($CDCl_3$, 600 MHz, 298 K) of **S2**.

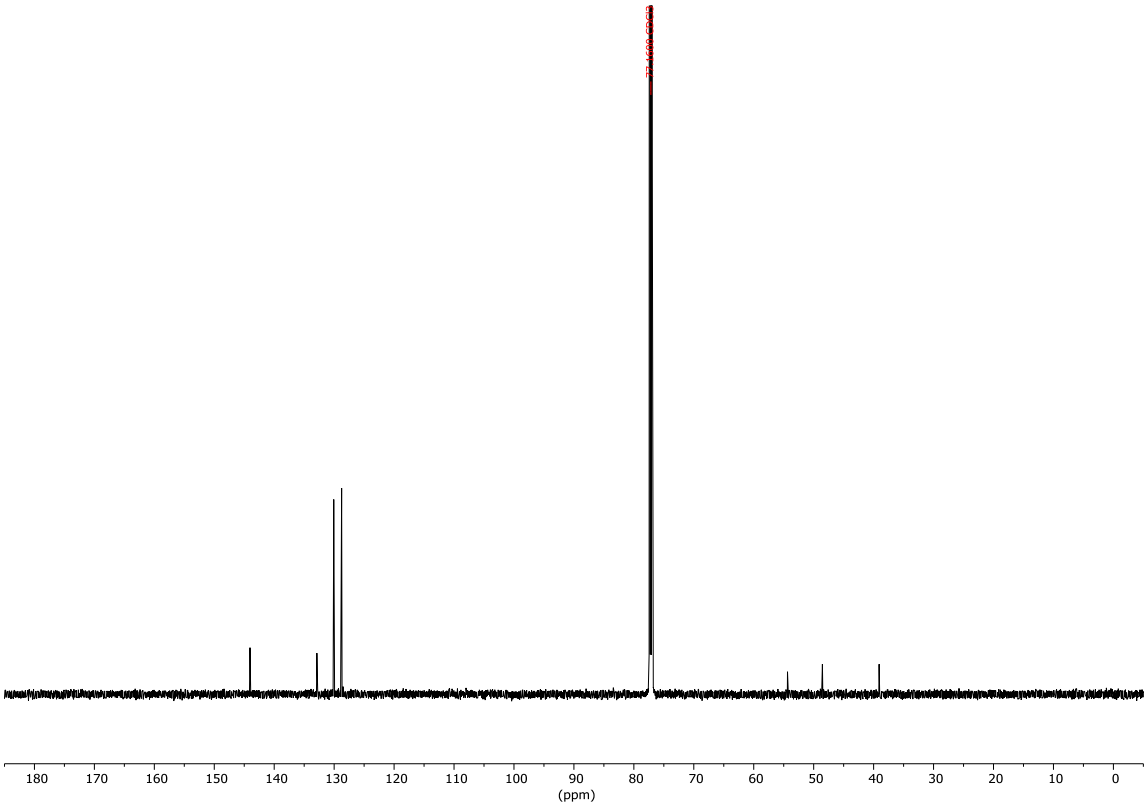


Figure S17. ^{13}C NMR ($CDCl_3$, 151 MHz, 298 K) of **S2**.

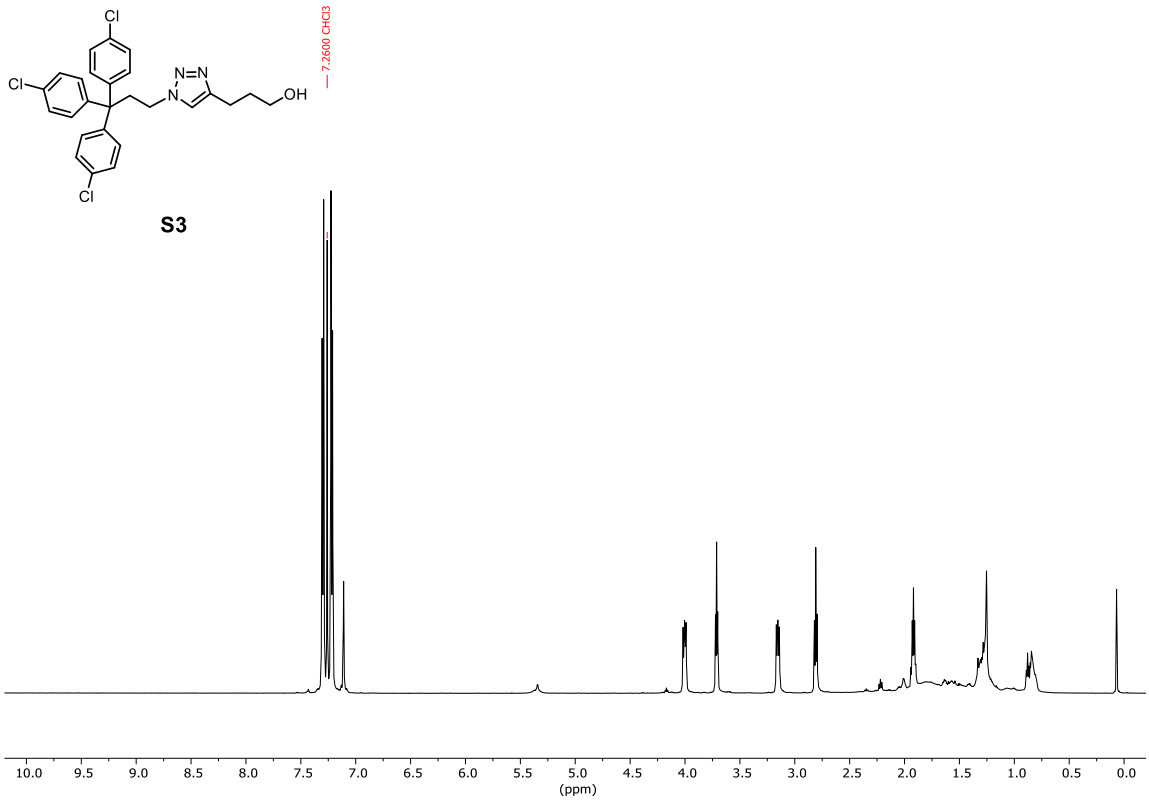


Figure S18. ^{1}H NMR ($CDCl_3$, 600 MHz, 298 K) of **S3**.

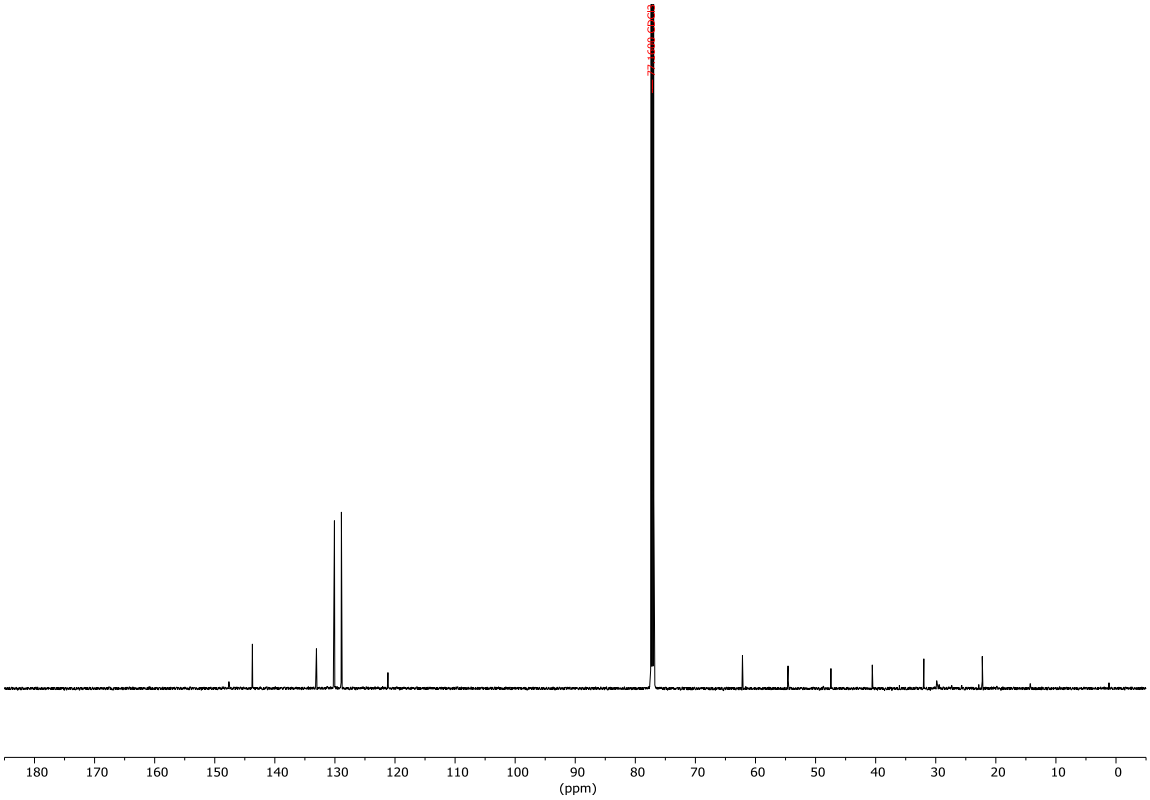


Figure S19. ^{13}C NMR ($CDCl_3$, 151 MHz, 298 K) of **S3**.

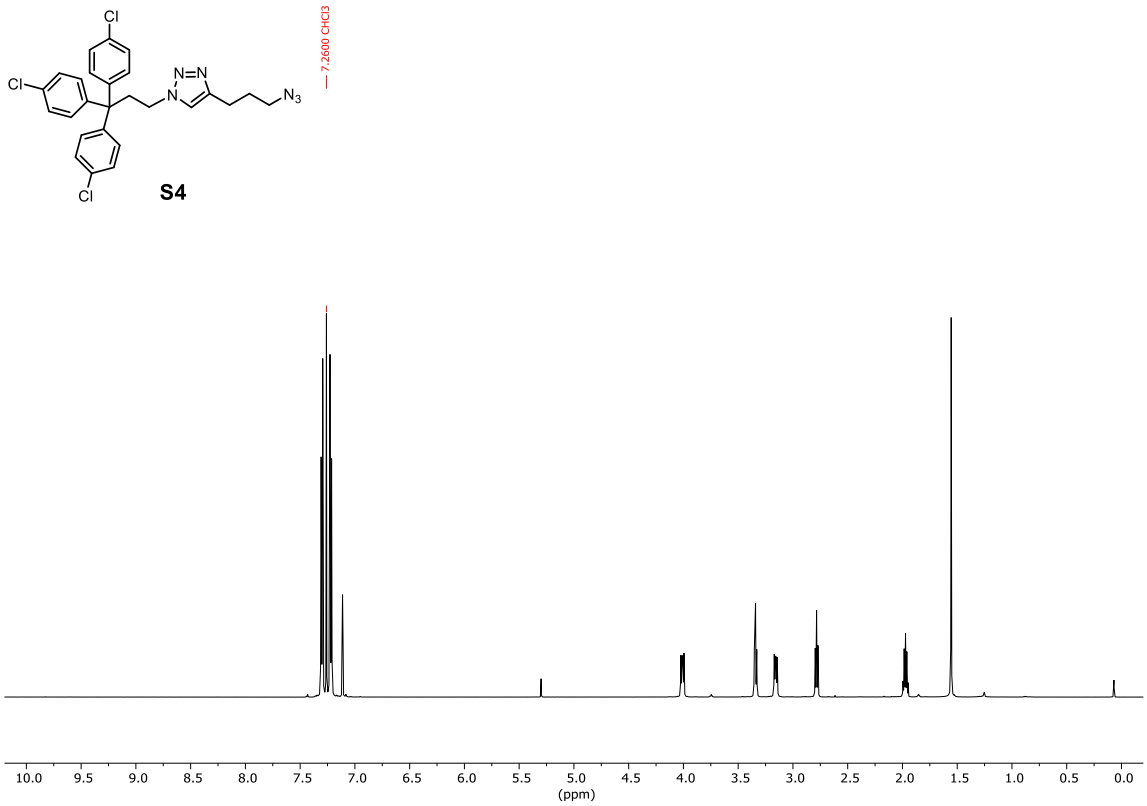


Figure S20. ^{1}H NMR ($CDCl_3$, 600 MHz, 298 K) of **S4**.

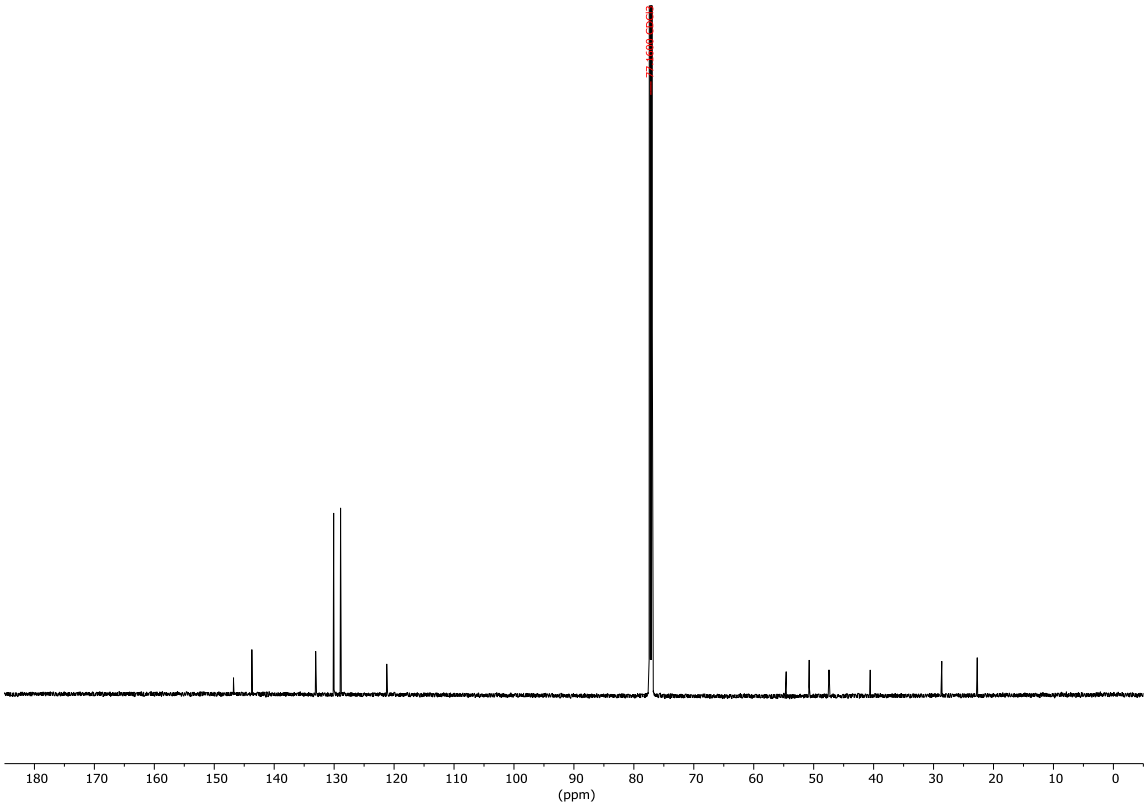


Figure S21. ^{13}C NMR ($CDCl_3$, 151 MHz, 298 K) of **S4**.

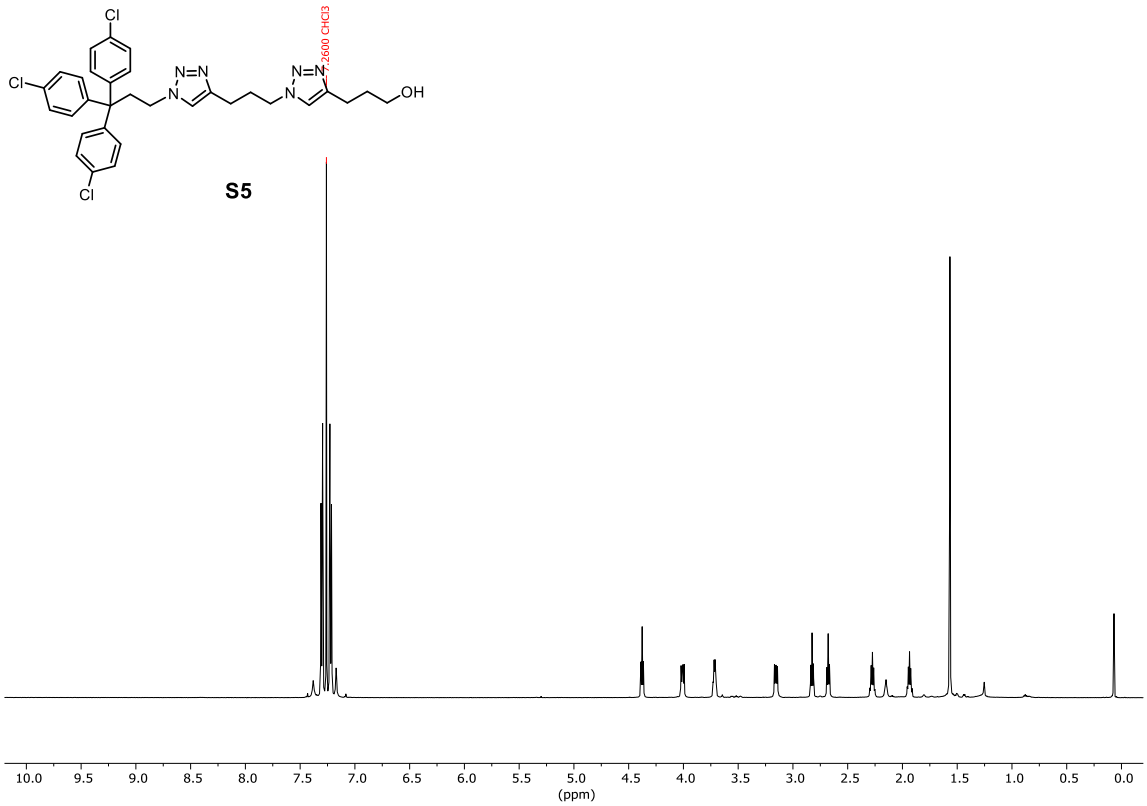


Figure S22. ^{1}H NMR ($CDCl_3$, 600 MHz, 298 K) of **S5**.

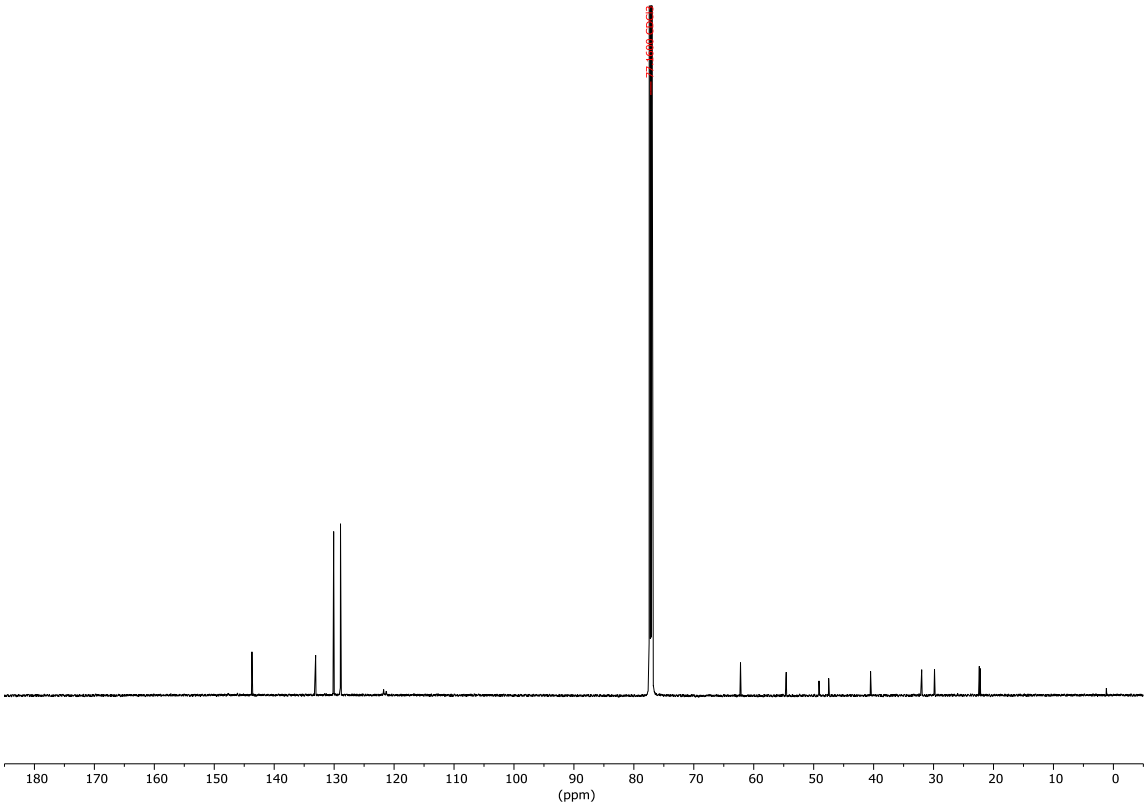


Figure S23. ^{13}C NMR ($CDCl_3$, 151 MHz, 298 K) of **S5**.

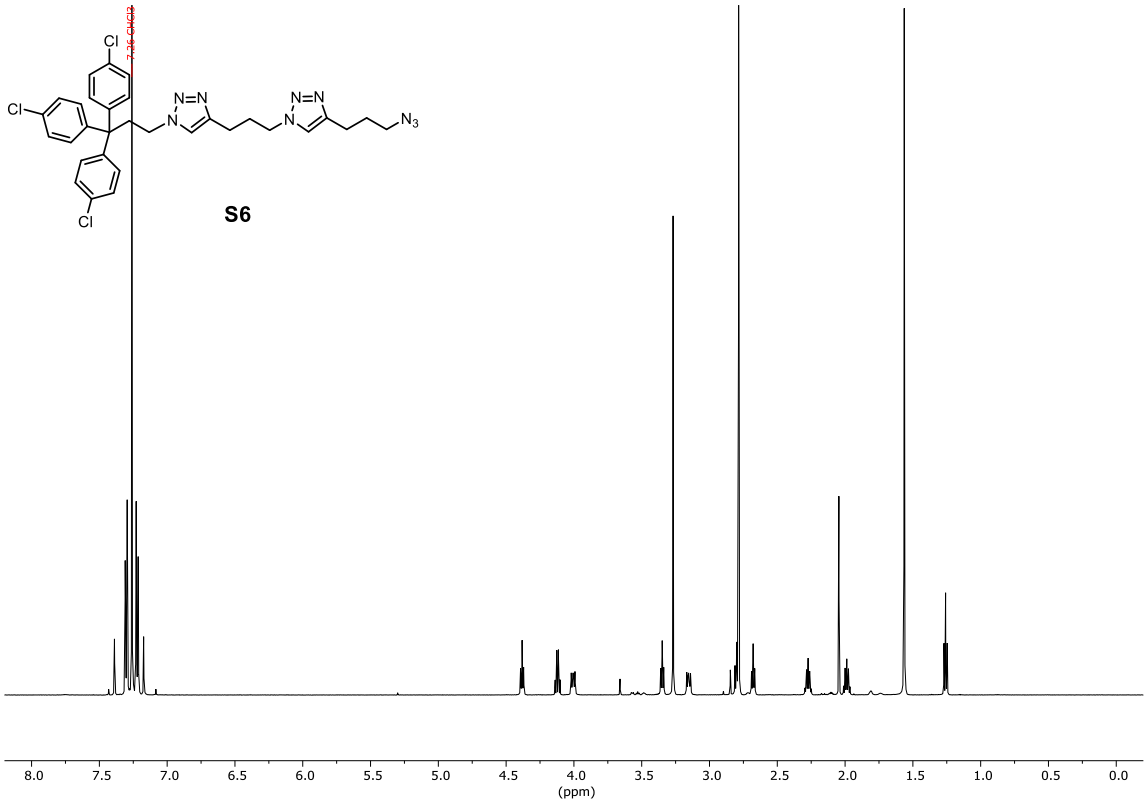


Figure S24. ^{1}H NMR ($CDCl_3$, 600 MHz, 298 K) of crude **S6** (with 1,3-dimethyl-2-imidazolidinone).

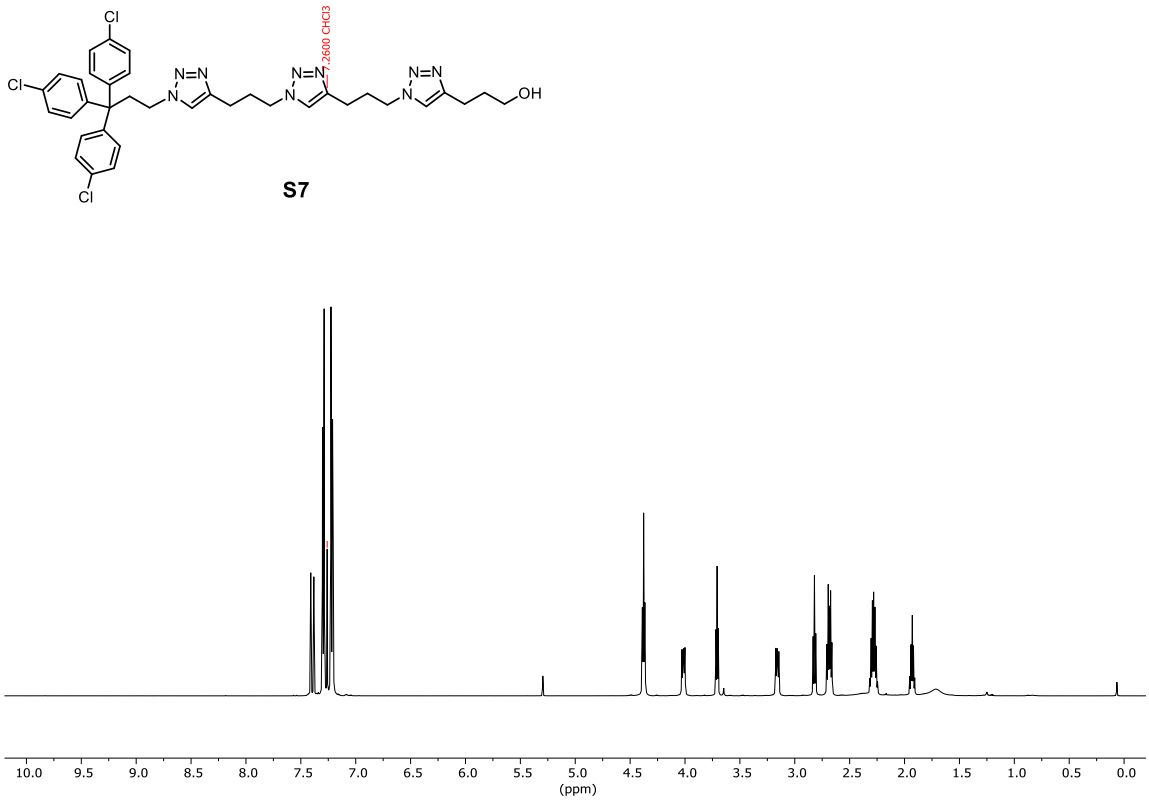


Figure S25. ^{1}H NMR ($CDCl_3$, 600 MHz, 298 K) of **S7**.

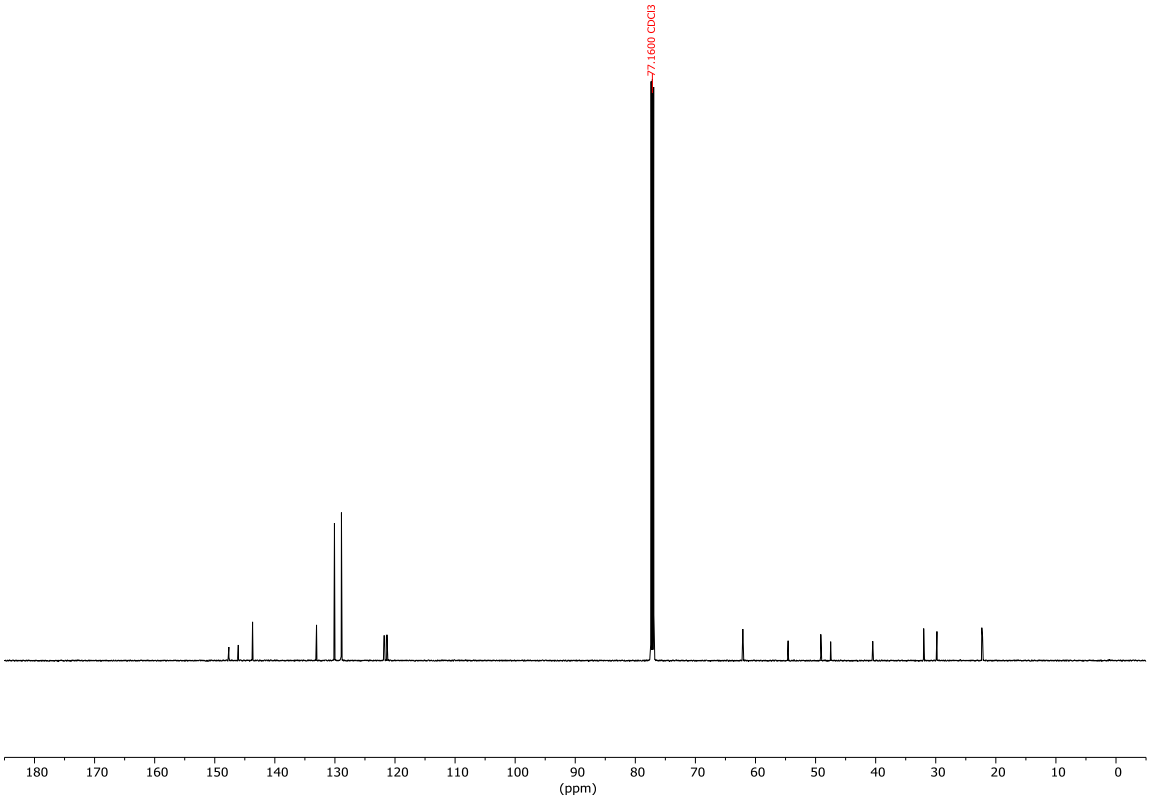


Figure S26. ^{13}C NMR ($CDCl_3$, 151 MHz, 298 K) of **S7**.

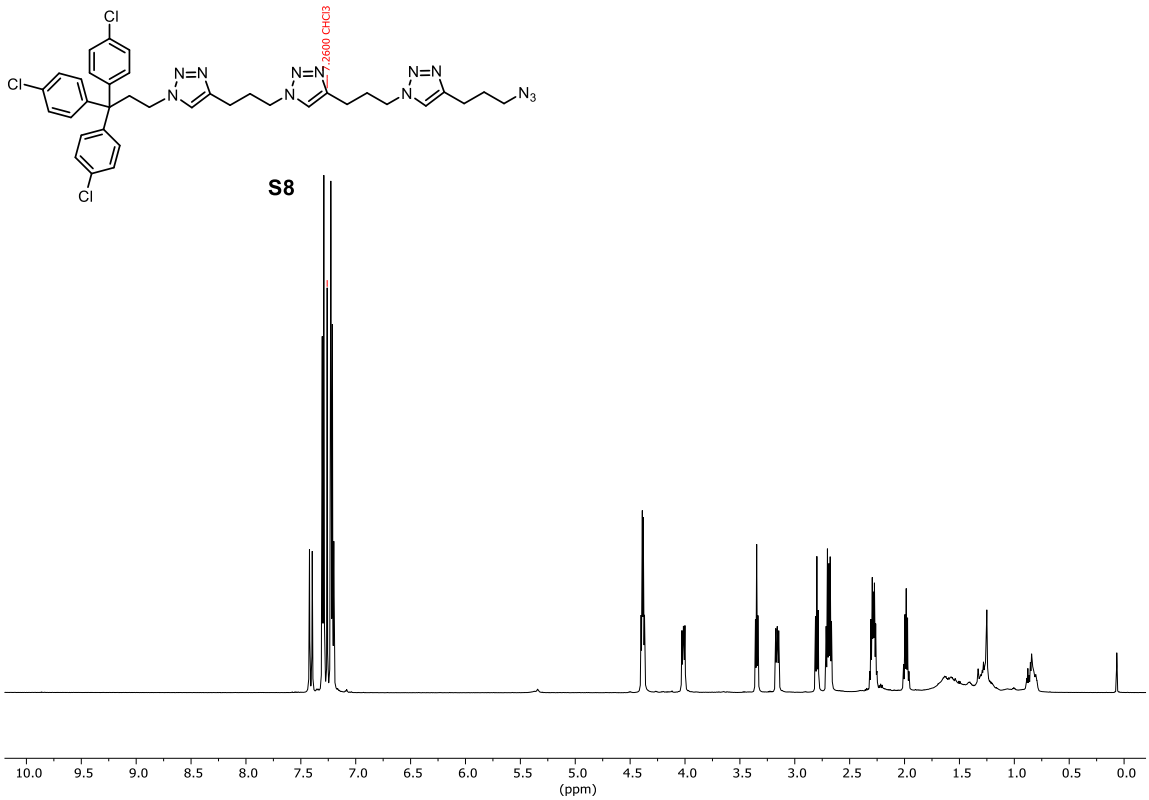


Figure S27. ^{1}H NMR ($CDCl_3$, 600 MHz, 298 K) of **S8**.

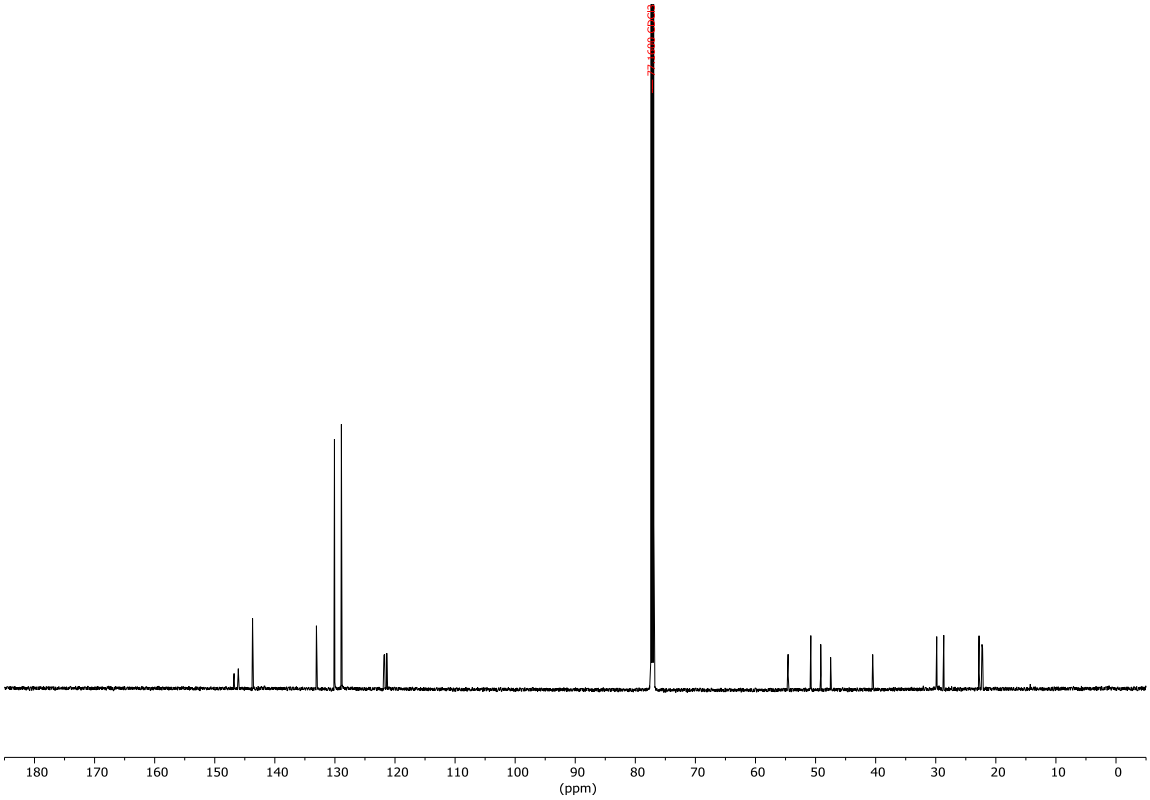
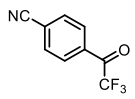


Figure S28. ^{13}C NMR ($CDCl_3$, 151 MHz, 298 K) of **S8**.



S9

Figure S29. ^{1}H NMR ($CDCl_3$, 600 MHz, 298 K) of **S9**.

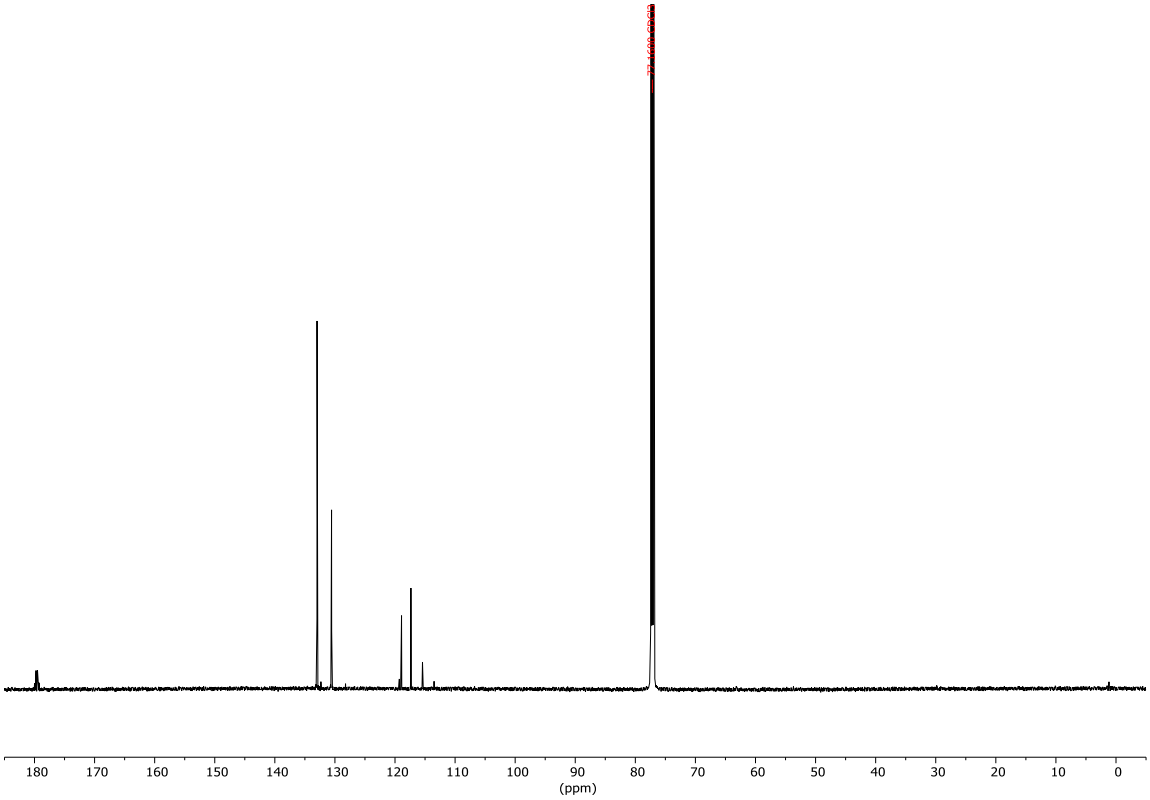


Figure S30. ^{13}C NMR ($CDCl_3$, 151 MHz, 298 K) of **S9**.

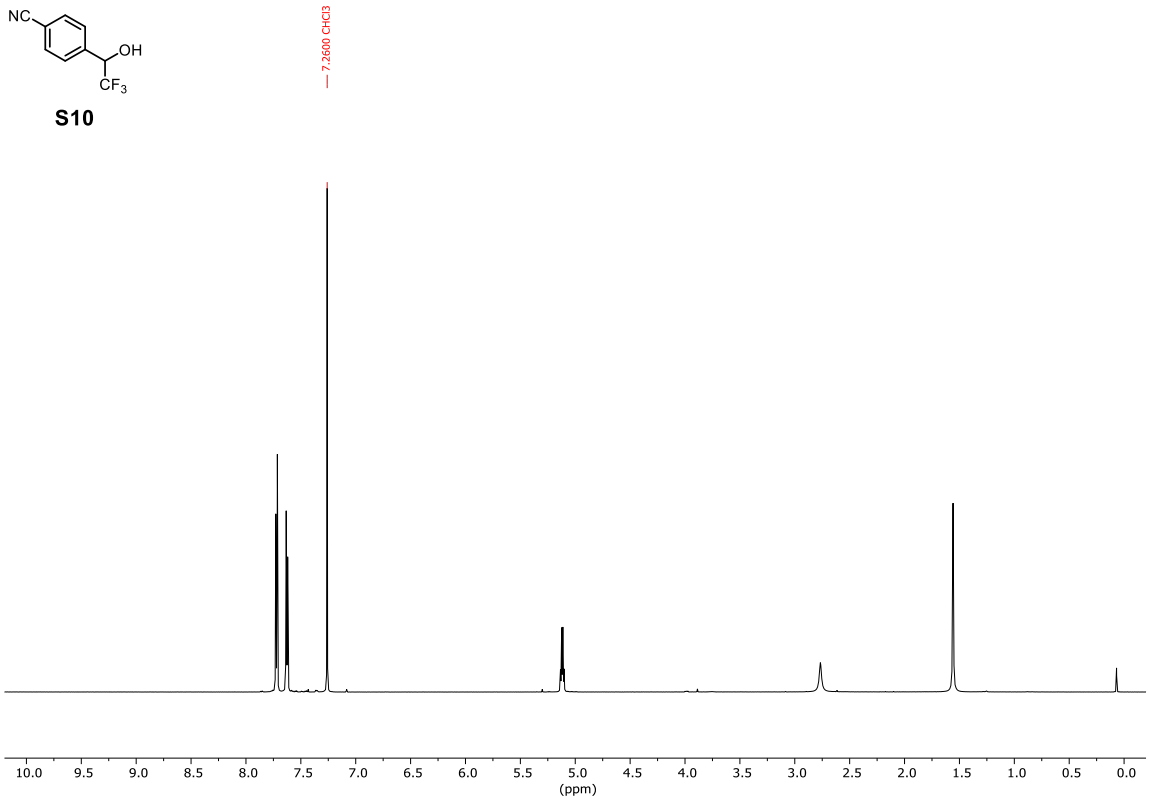


Figure S31. ^{1}H NMR ($CDCl_3$, 600 MHz, 298 K) of **S10**.

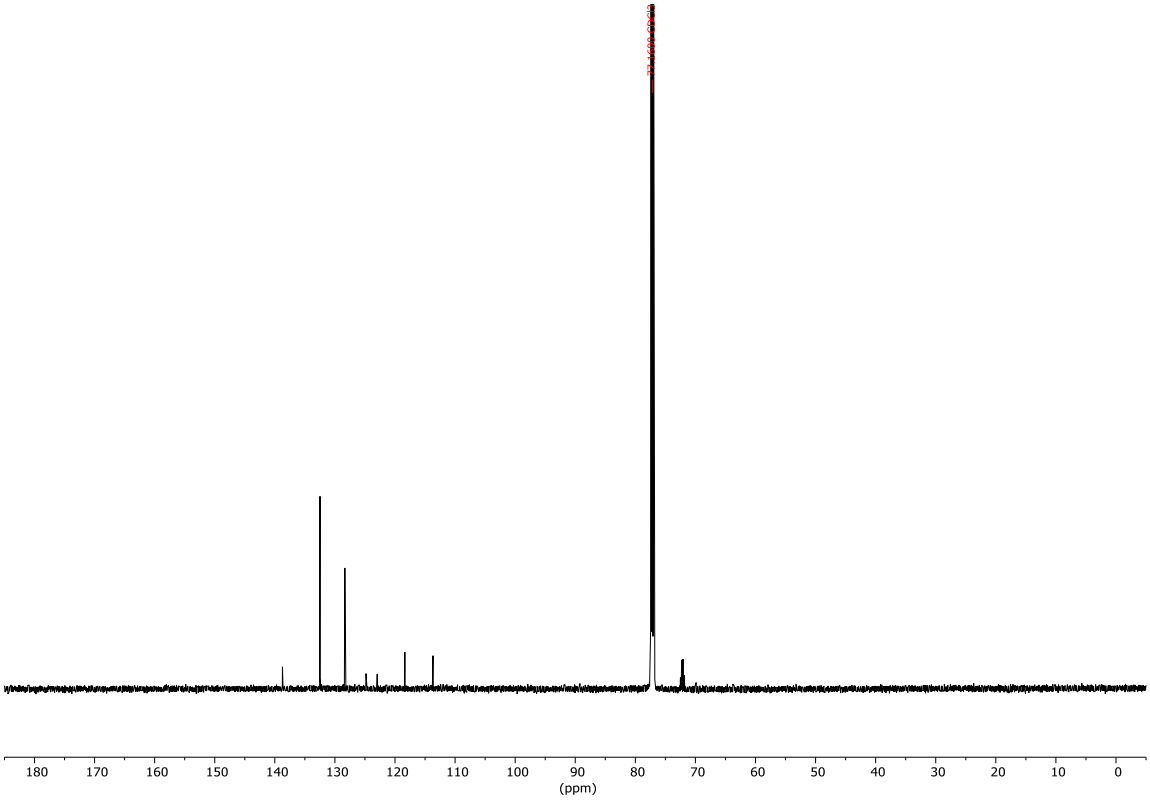
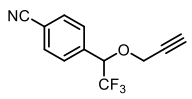


Figure S32. ^{13}C NMR ($CDCl_3$, 151 MHz, 298 K) of **S10**.



S11

Figure S33. ^{1}H NMR ($CDCl_3$, 600 MHz, 298 K) of **S11**.

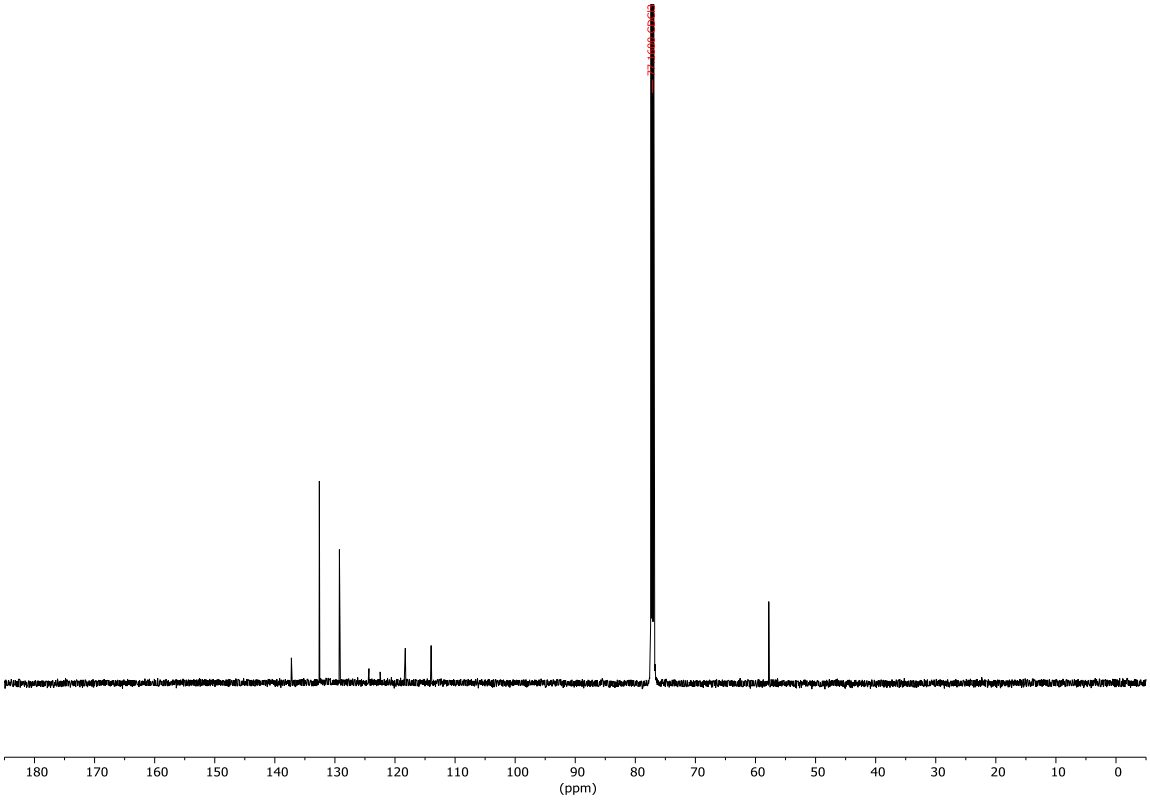


Figure S34. ^{13}C NMR ($CDCl_3$, 151 MHz, 298 K) of **S11**.

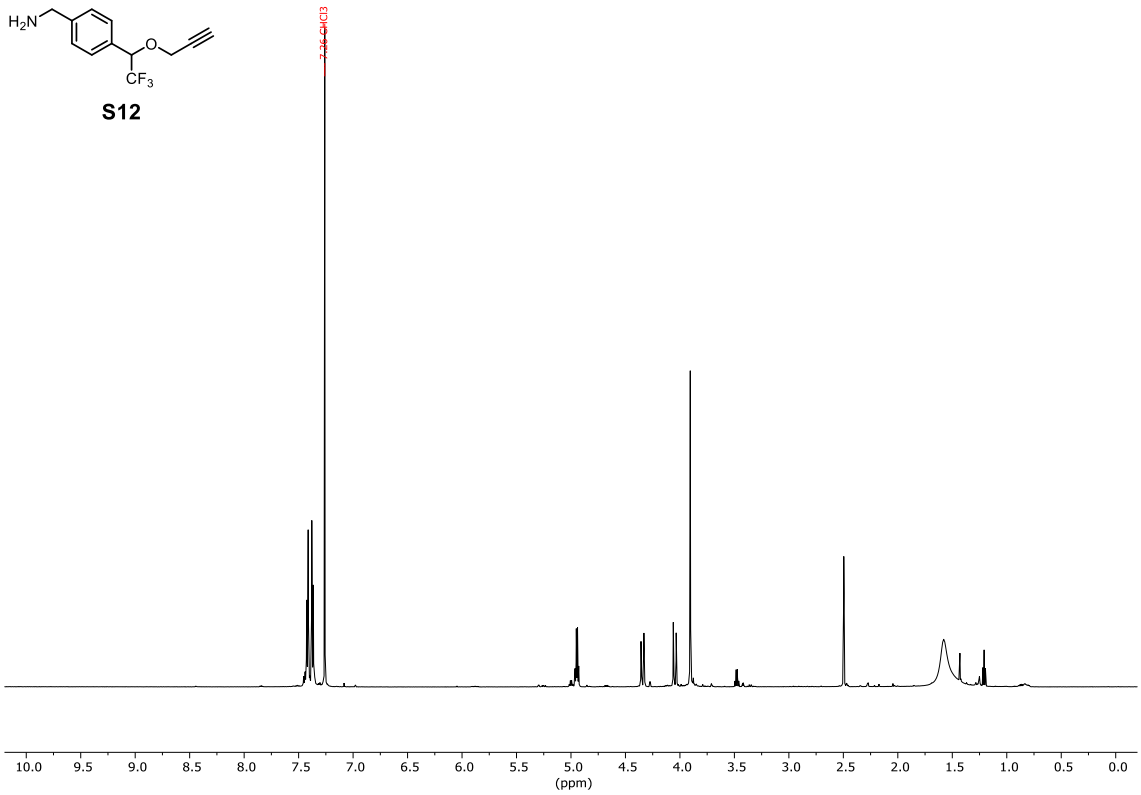
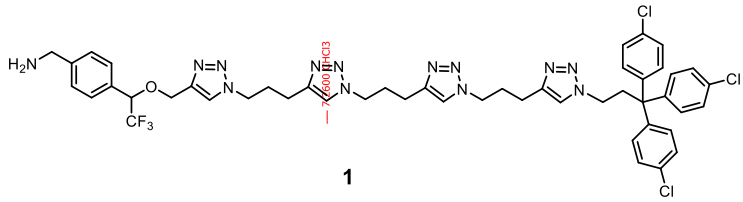


Figure S35. ^{1}H NMR ($CDCl_3$, 600 MHz, 298 K) of crude **S12**.



1

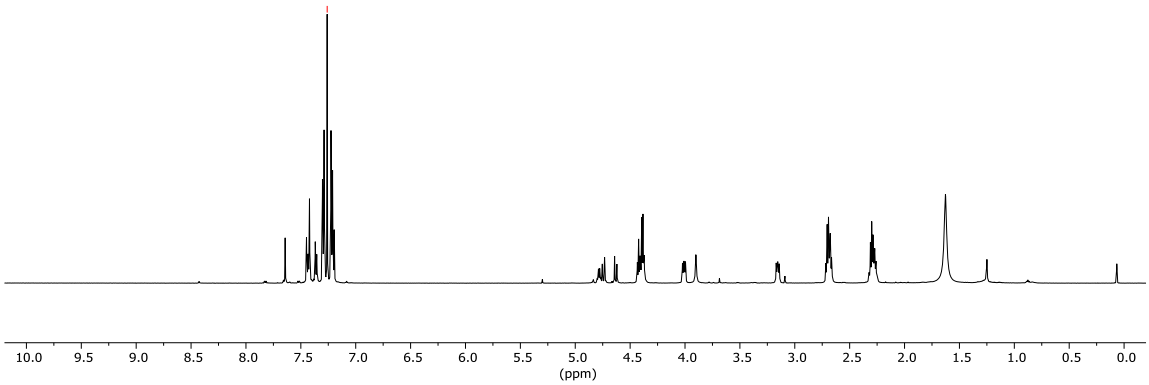


Figure S36. ^{1}H NMR ($CDCl_3$, 600 MHz, 298 K) of **1**.

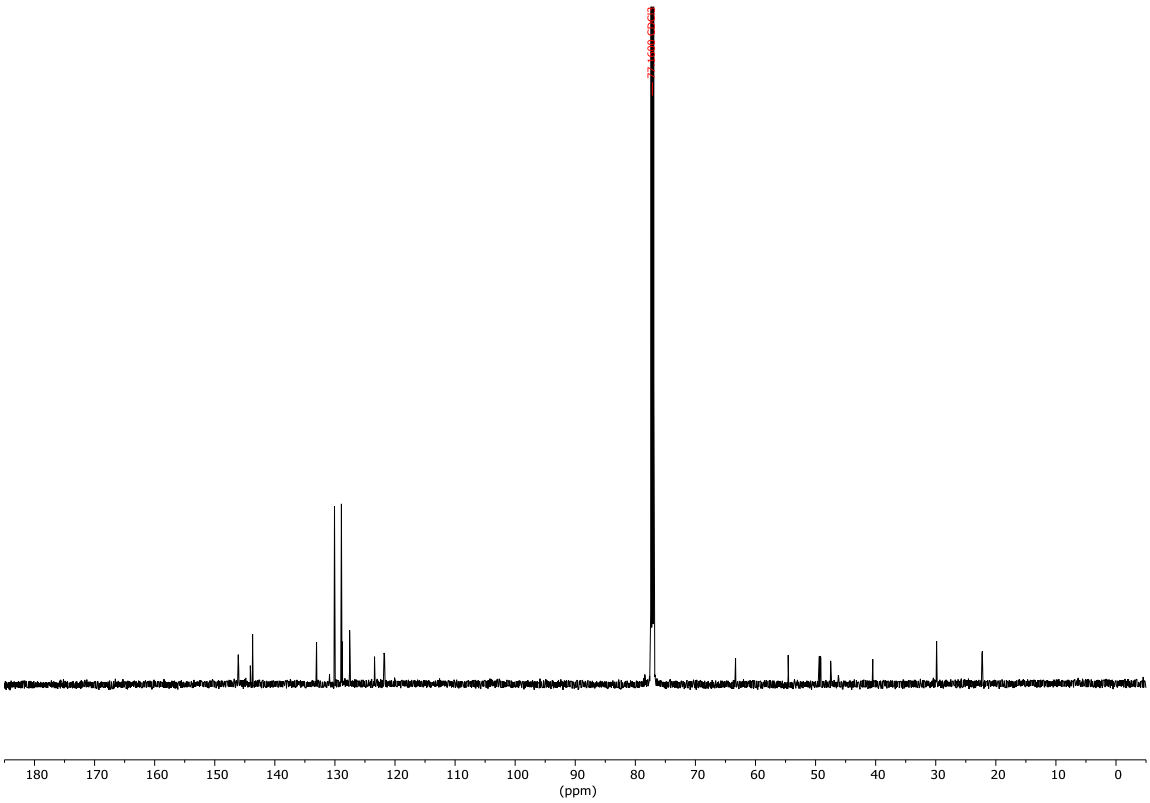


Figure S37. ^{13}C NMR ($CDCl_3$, 151 MHz, 298 K) of **1**.

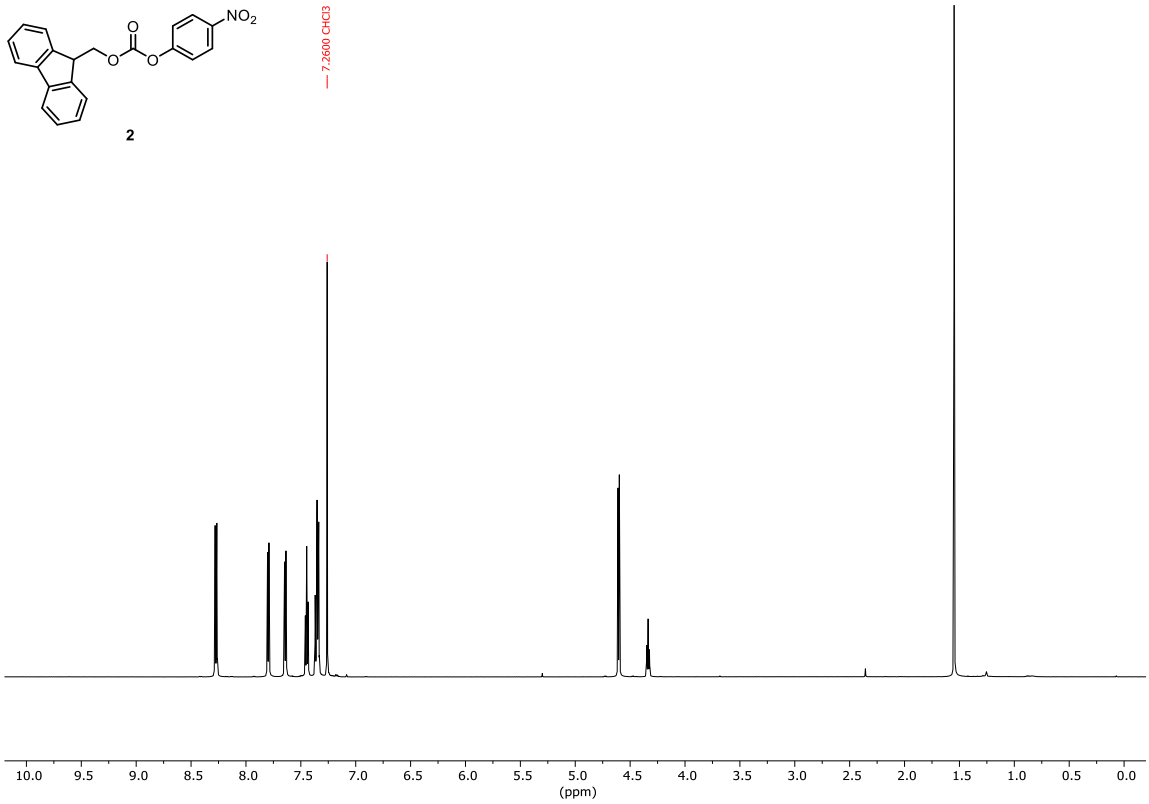


Figure S38. ^{1}H NMR ($CDCl_3$, 600 MHz, 298 K) of **2**.

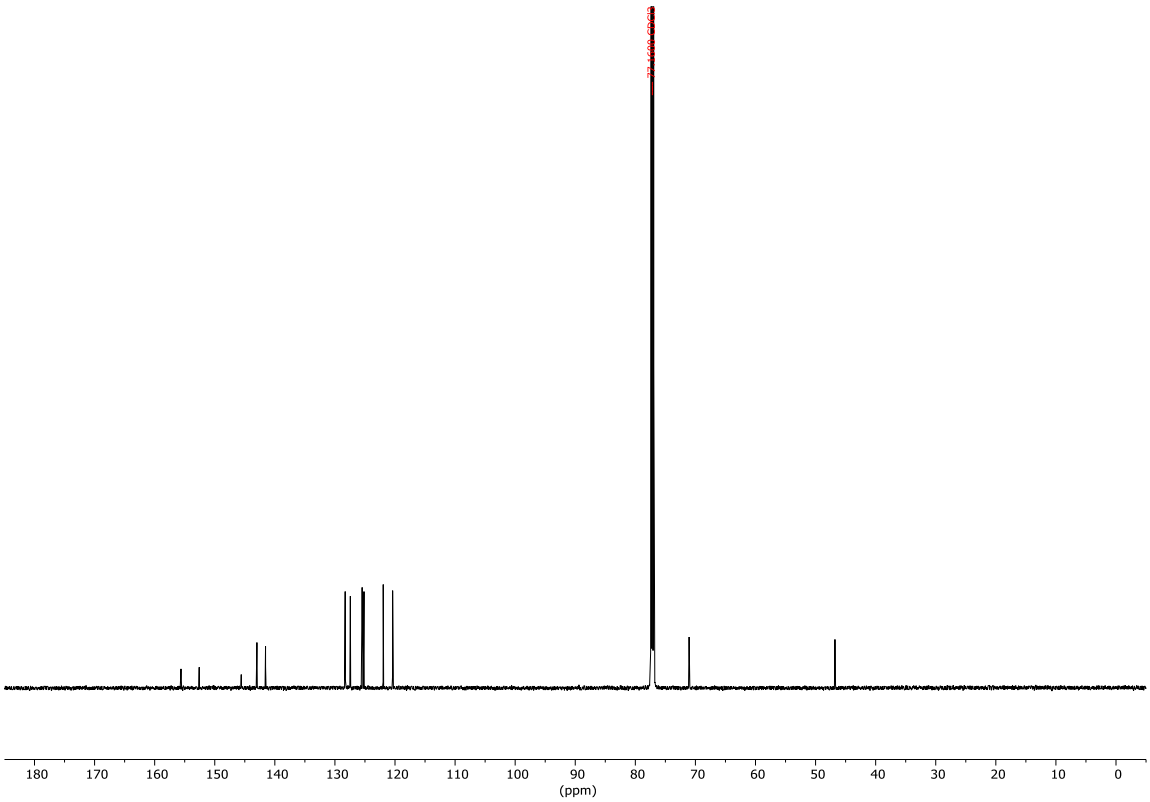


Figure S39. ^{13}C NMR ($CDCl_3$, 151 MHz, 298 K) of **2**.

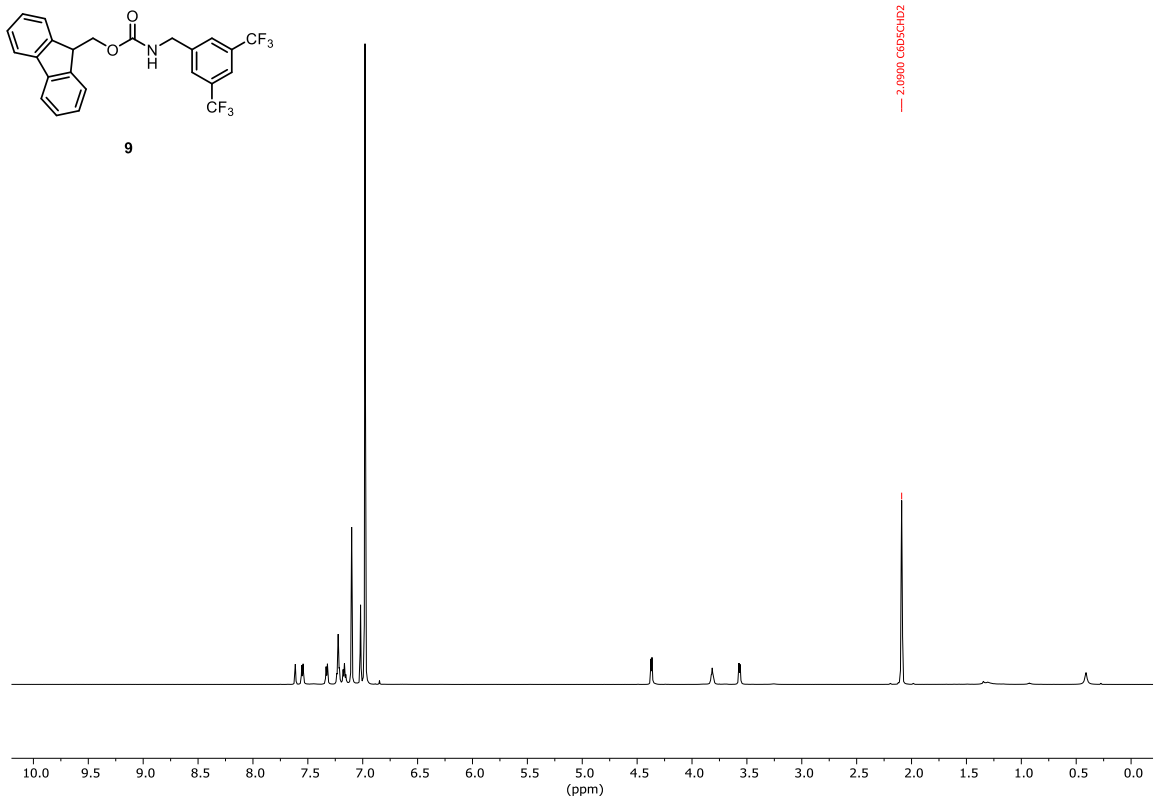


Figure S40. ^{1}H NMR ($C_6D_5CD_3$, 600 MHz, 298 K) of **9**.

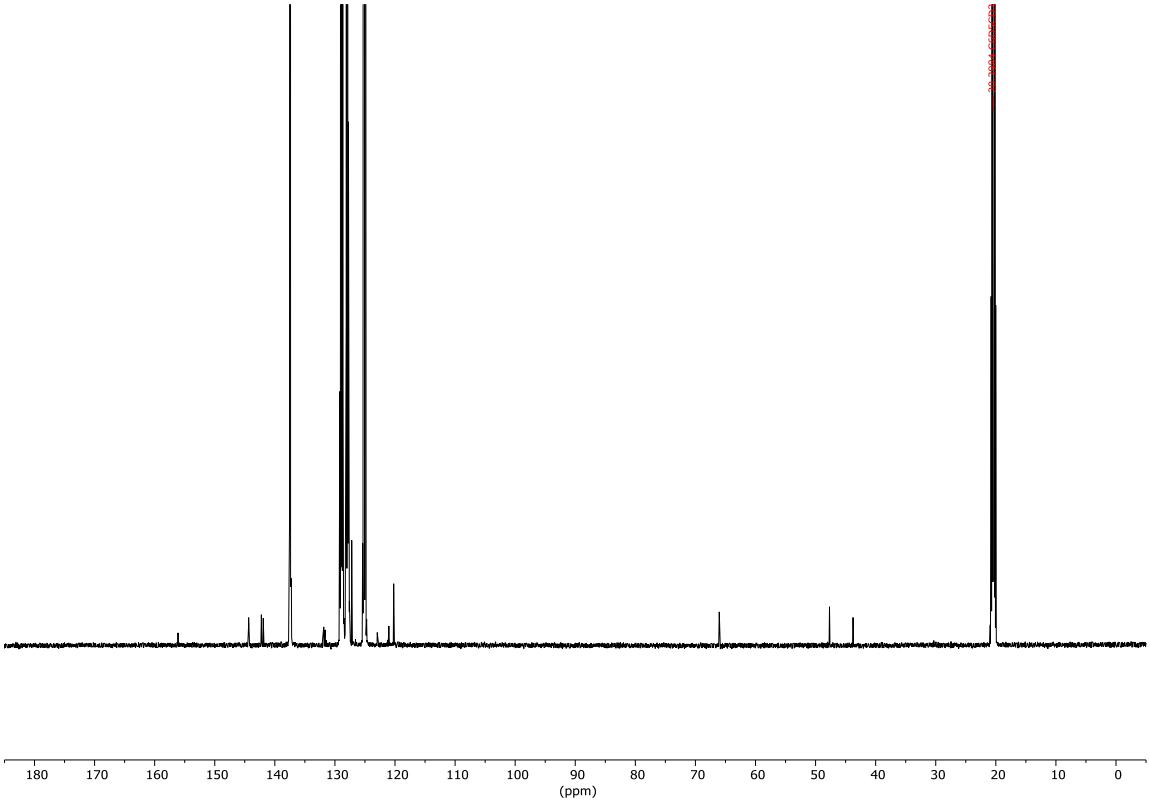


Figure S41. ^{13}C NMR ($C_6D_5CD_3$, 151 MHz, 298 K) of **9**.

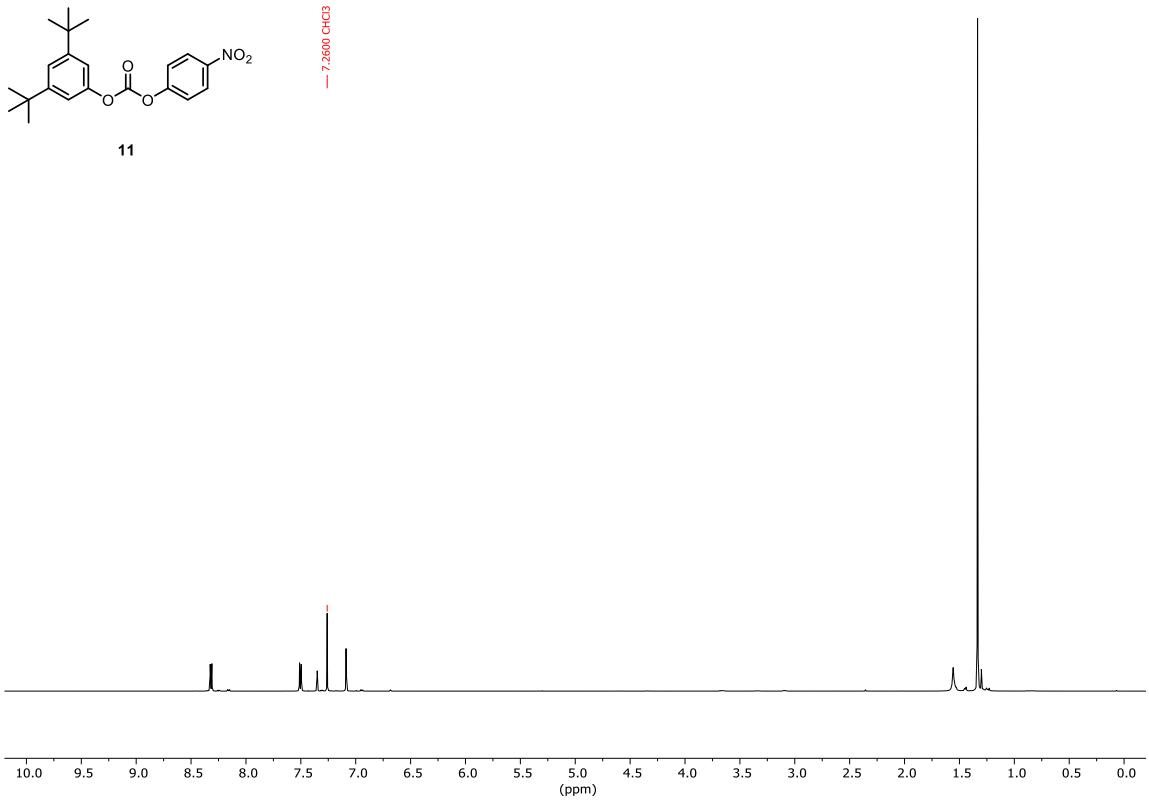


Figure S42. ^{1}H NMR ($CDCl_3$, 600 MHz, 298 K) of **11**.

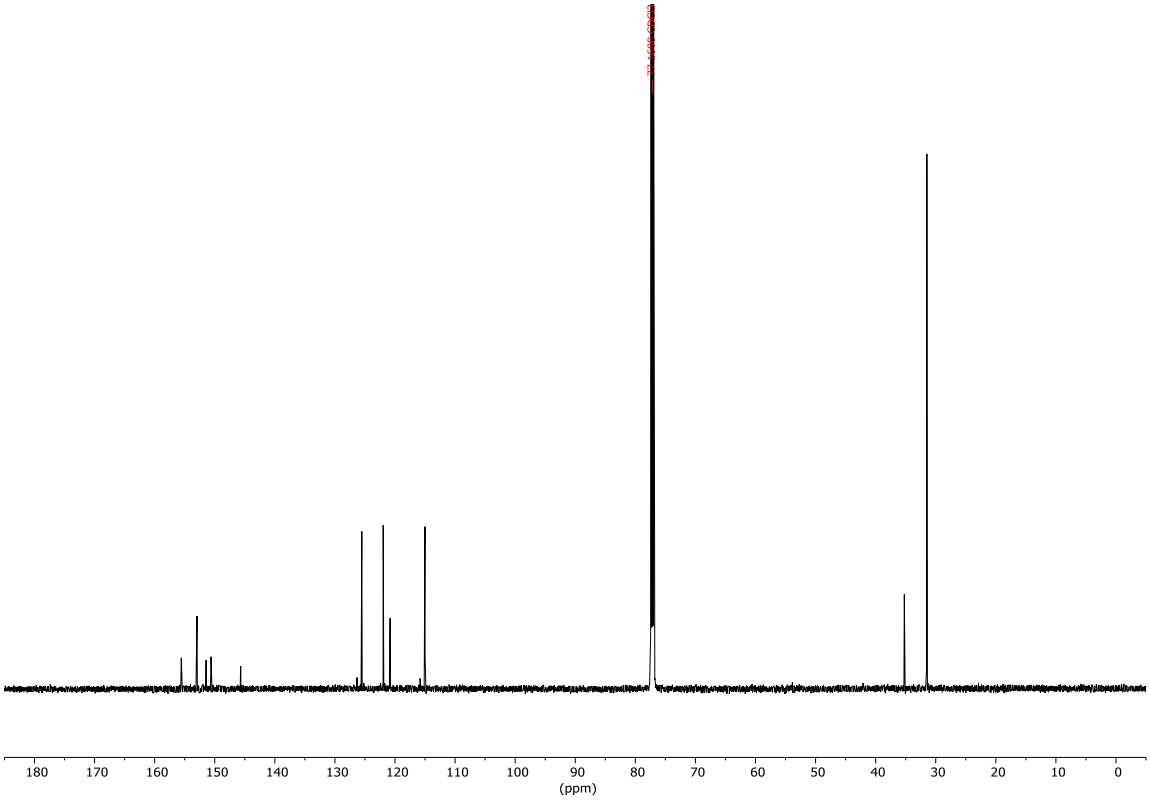


Figure S43. ^{13}C NMR ($CDCl_3$, 151 MHz, 298 K) of **11**.

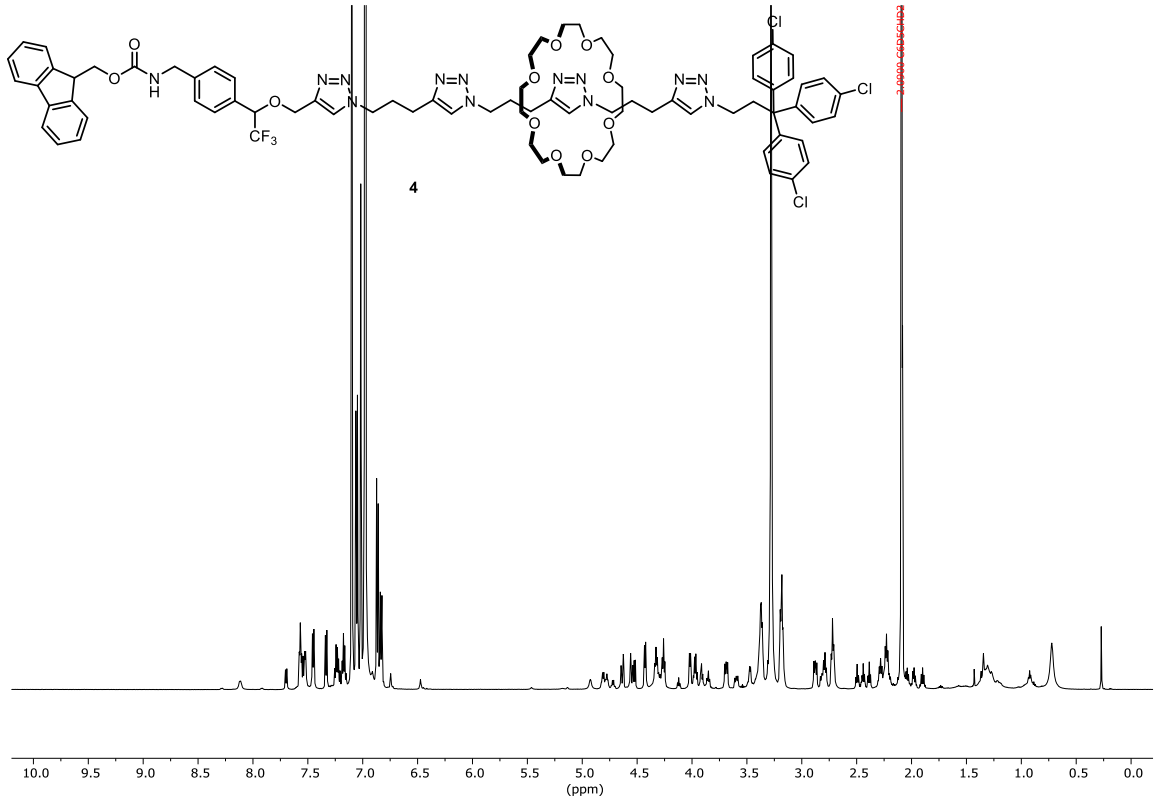


Figure S44. ^{1}H NMR ($C_6D_5CD_3$, 600 MHz, 298 K) of **4**.

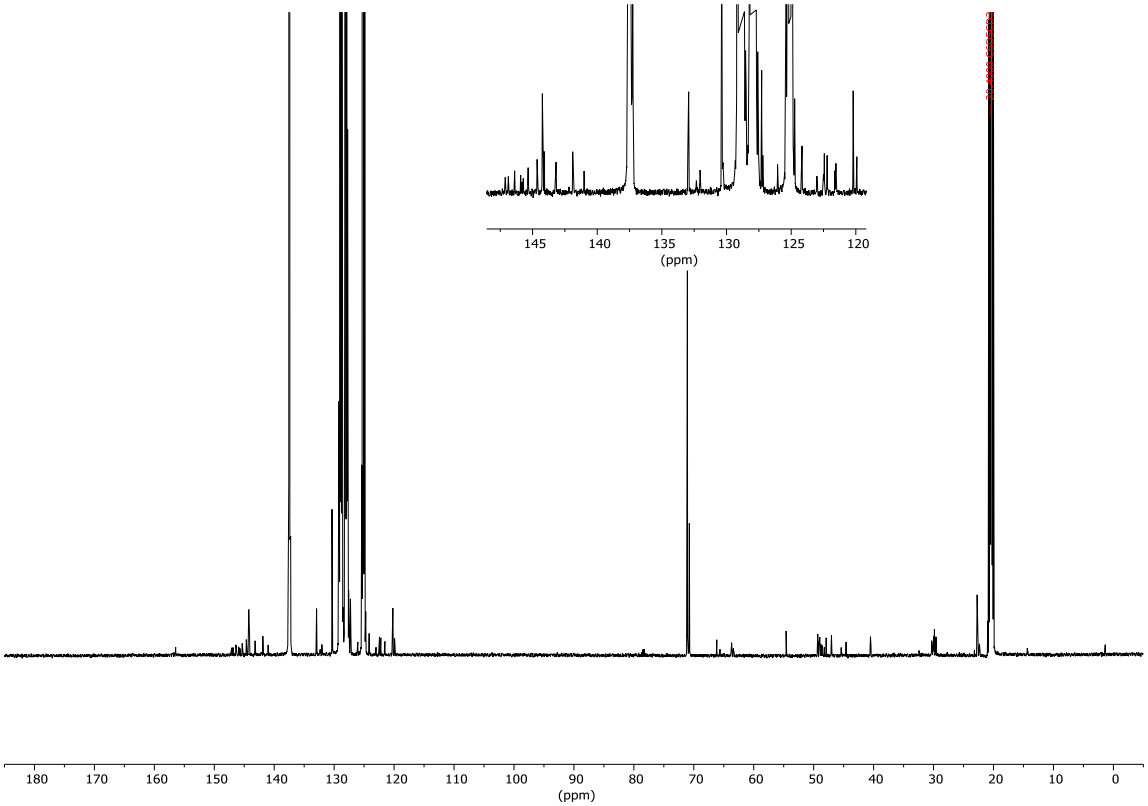


Figure S45. ^{13}C NMR ($C_6D_5CD_3$, 151 MHz, 298 K) of **4**.

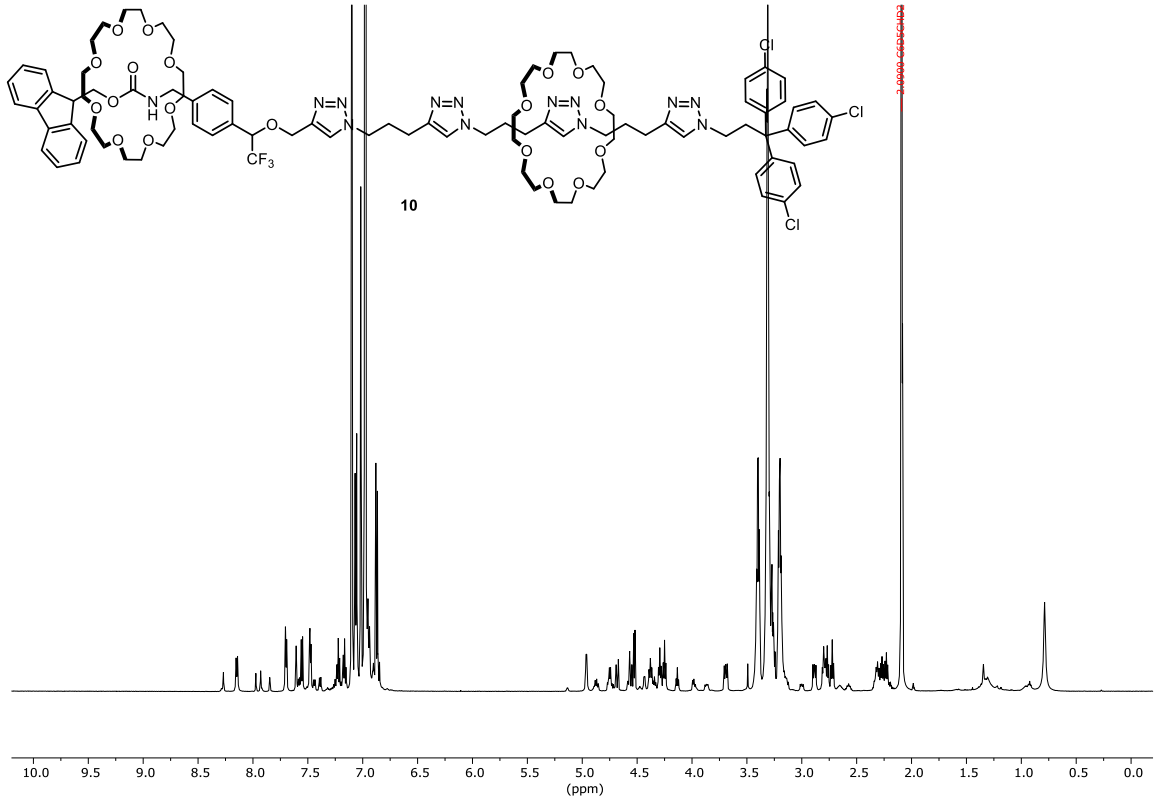


Figure S46. ^{1}H NMR ($C_6D_5CD_3$, 600 MHz, 298 K) of **10**.

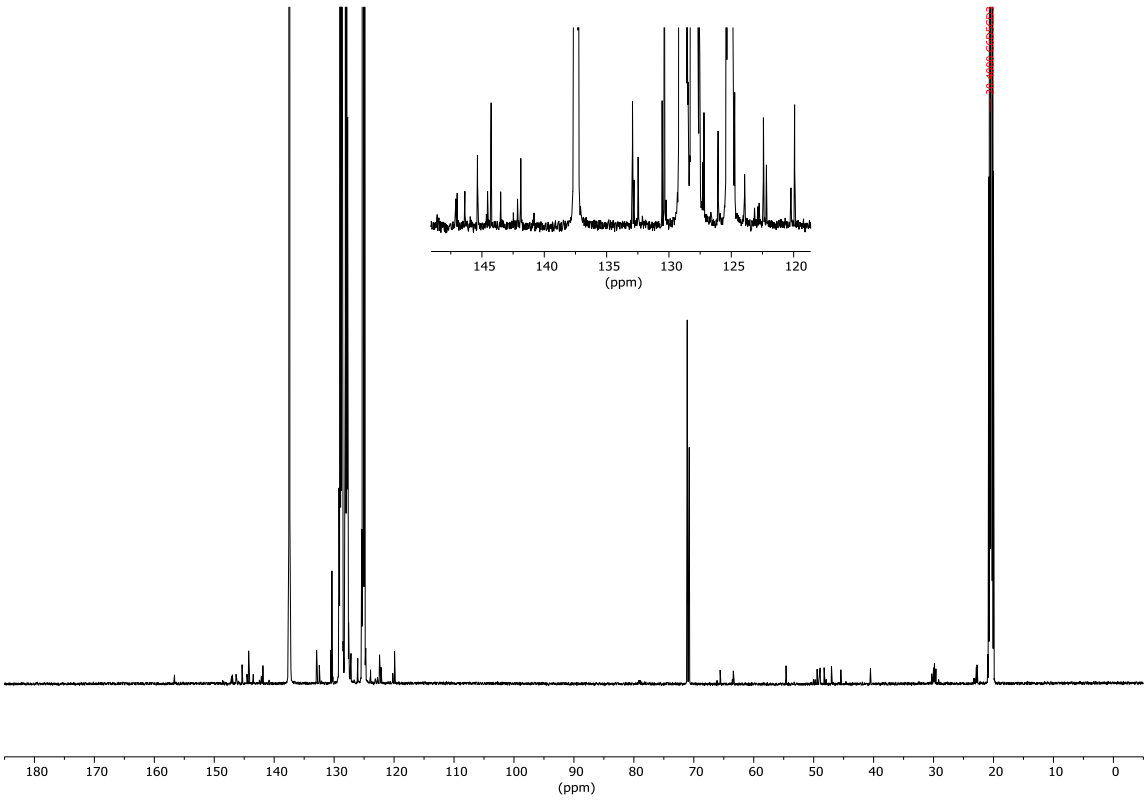


Figure S47. ^{13}C NMR ($C_6D_5CD_3$, 151 MHz, 298 K) of **10**.

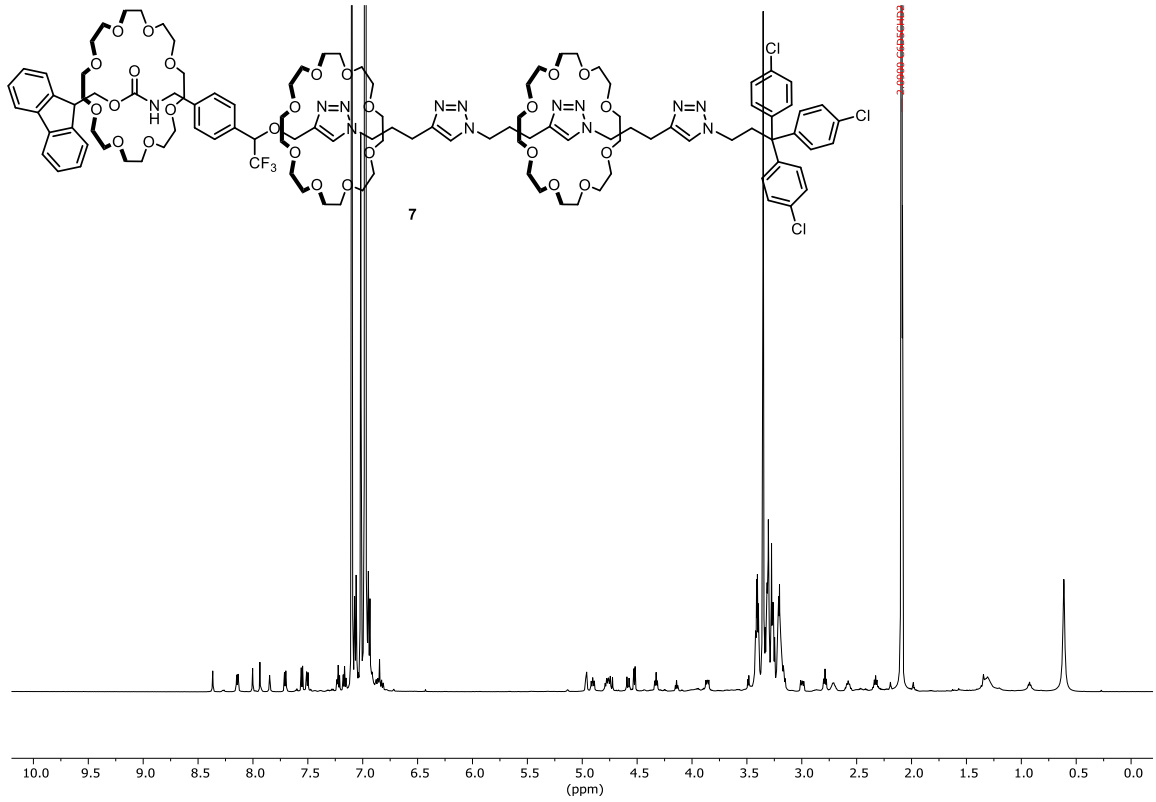


Figure S48. ^{1}H NMR ($C_6D_5CD_3$, 600 MHz, 298 K) of **7**.

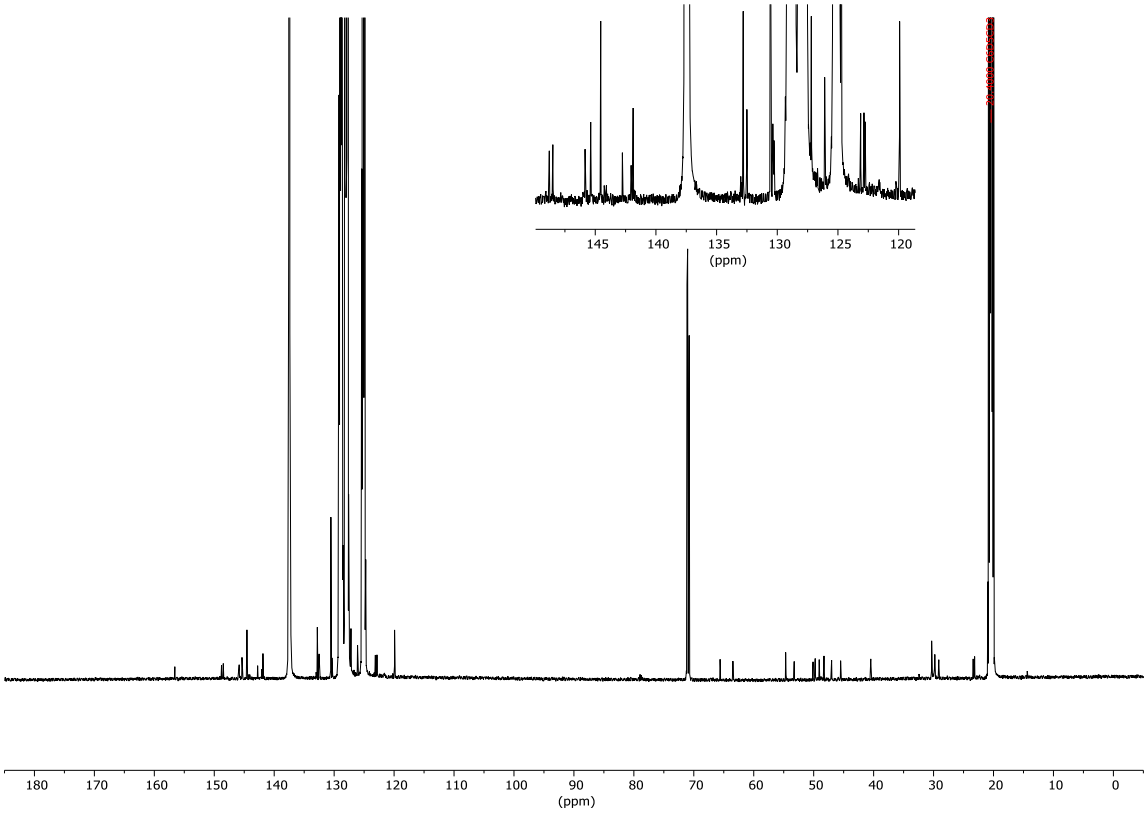


Figure S49. ^{13}C NMR ($C_6D_5CD_3$, 151 MHz, 298 K) of **7**.

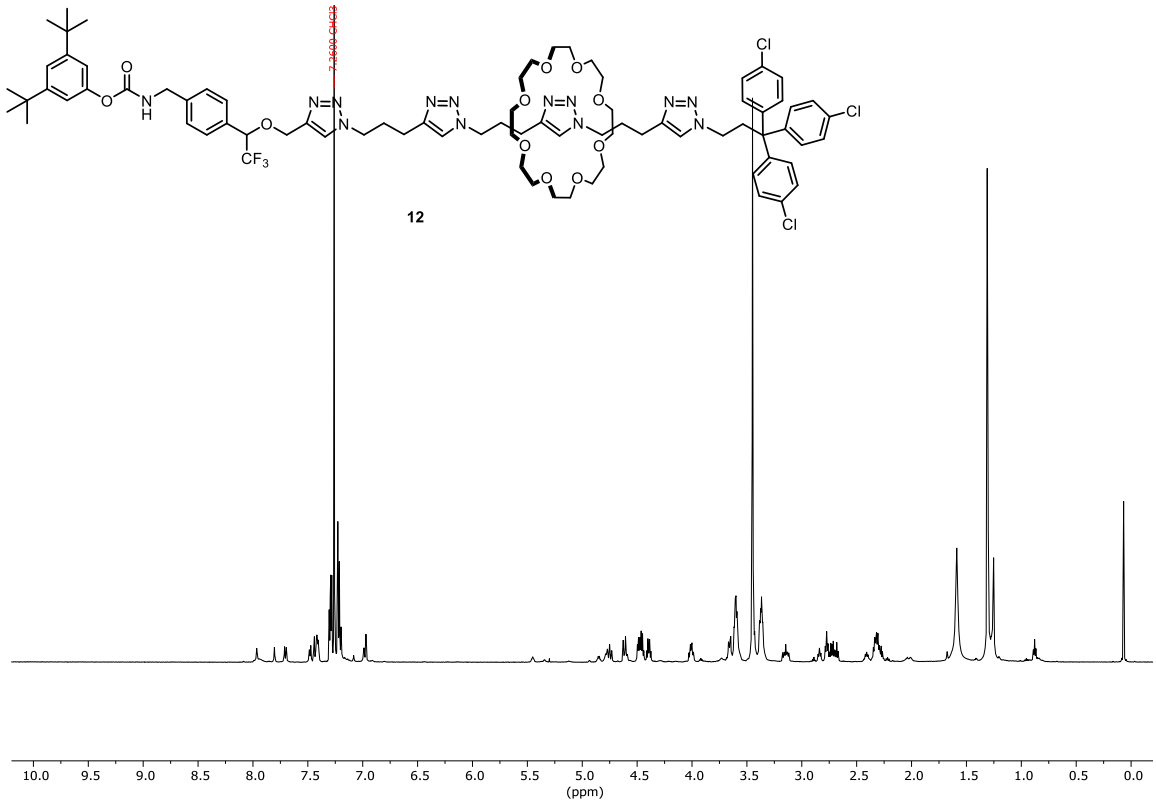


Figure S50. ^{1}H NMR ($CDCl_3$, 600 MHz, 298 K) of **12**.

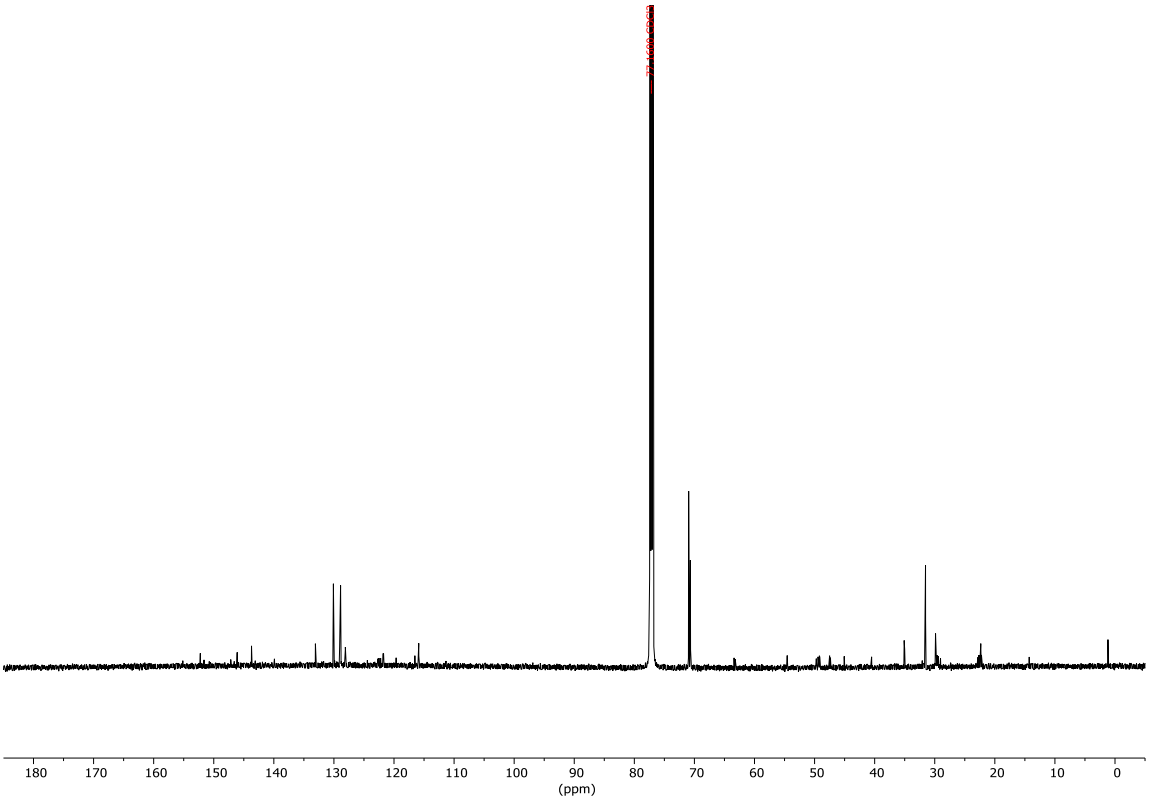


Figure S51. ^{13}C NMR ($CDCl_3$, 151 MHz, 298 K) of **12**.

3.4.5. Optimisation of pumping conditions

It was recently shown[29] that Fmoc-Cl is a suitable electrophile for metal-free active template rotaxane synthesis with crown ether **3** (**Figure S52**).

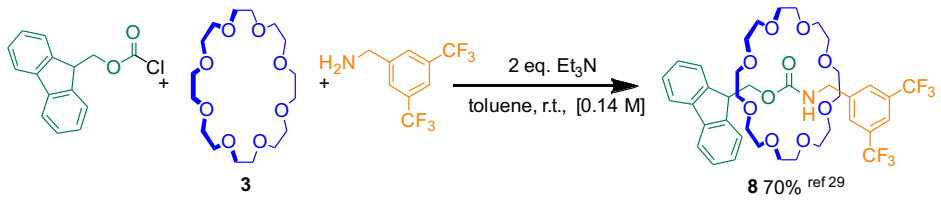


Figure S52. Metal-free active template rotaxane synthesis using Fmoc-Cl as the electrophile.

This active template reaction required an additional base to be present to sequester the HCl by-product and so prevent deactivation of the amine nucleophile by protonation. We reasoned that adding more of this base would cause removal of the Fmoc group, to regenerate starting amine and crown ether **3**, under the same conditions as rotaxane formation. If both processes occur simultaneously, the rotaxane is continuously formed and decomposed (**Figure S53**).

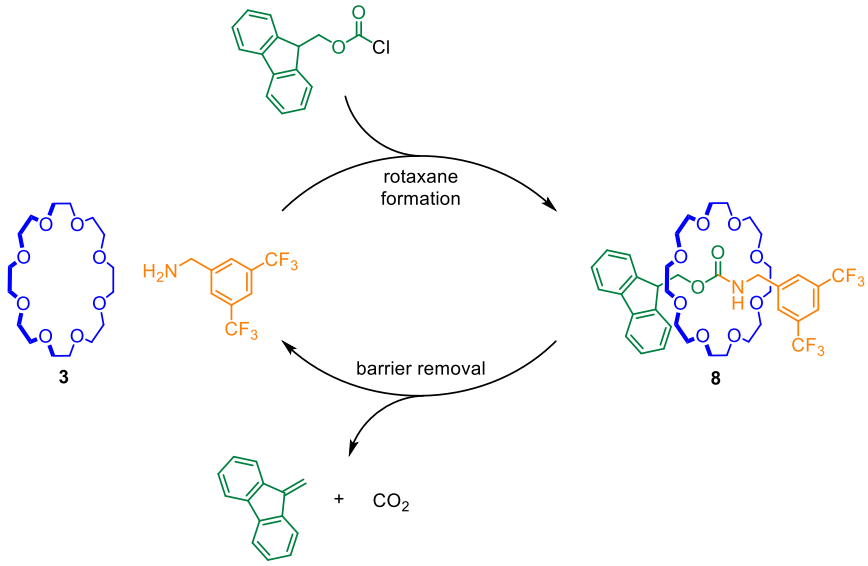


Figure S53. Simultaneous Fmoc addition and removal

We proposed that a pump could be realized by adding a kinetic barrier ('speed bump') and catchment region to the amine nucleophile. These would inhibit dethreading of intermediate pseudo-rotaxanes that are present between pumping cycles. This would allow autonomous pumping of multiple crown ethers onto the same thread if displacement of the crown ether over the speed bump is slower than the active template reaction.

All studies were performed using previously optimized reaction conditions for metal-free active template synthesis unless otherwise stated: toluene as a solvent, room temperature, high concentration (0.1 – 0.2 M).[28, 29] For multiple pumping cycles to occur the reaction should remain homogenous throughout pump operation. For this reason, we switched to using Fmoc-$OC_6H_4NO_2$ (4-nitrophenol leaving group) **2** as the fuel, because the build-up of HCl salts that occurs when using Fmoc-Cl results in partial gelation of the reaction medium. We used unsubstituted 24-crown-8 **3** because it is the optimum size for giving the highest rotaxane yields,[28, 29] it has a simple ^{1}H NMR spectrum and is freely soluble in toluene.

We first screened commercially available benzylic amines as nucleophiles to determine how quickly the crown ether passed over potential speed barriers. This was determined by analyzing the amount of free crown ether released back into solution over time. Benzylic amines were used because they form rotaxanes with easily interpretable ^{1}H NMR spectra. Benzyl amines containing a trisubstituted aryl group, e.g. **S13**, gave stable rotaxanes (**Figure S54**, **A**). We found that pseudo-rotaxane **S14•3**, derived from 4-*tert*-butylbenzylamine, had a suitable half-life of ca. 90 h, which is slower than rotaxane formation (complete in 10 min).

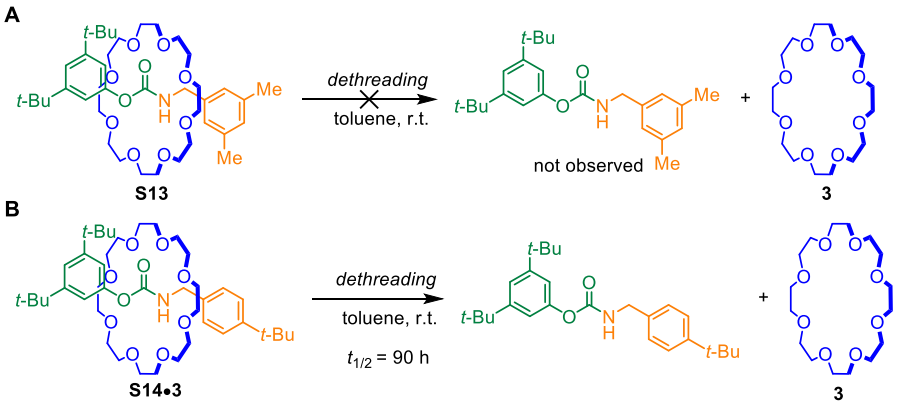


Figure S54. Determining optimal speed bump size. (**A**) Trisubstituted amine stoppers do not permit dethreading. (**B**) Psuedo-rotaxane **S14•3**, derived from 4-*tert*-butylbenzylamine, slowly dethreads.

The first model pump **S15** was therefore designed with a geminal-dimethyl (-$(CH_3)_2$) speed bump (**Figure S55**). Using an ether, rather than alkyl, linkage simplified synthesis. This design also included a catchment region containing a single amide unit to act as a secondary binding station for the crown ether. An amide was chosen because it has a similar H-bond donor ability to a carbamate group[50] and so should promote displacement over the speed bump. The corresponding [2]rotaxane **S16** was accessed in ca. 50% yield. This existed exclusively as the *car*-**S16** co-conformer (i.e. crown ether binding to the carbamate station) under the reaction conditions. Displacement to give *ami*-**S16** (i.e. crown ether binding to the amide station) was only observed when *car*-**S16** was heated to 40 °C in $CDCl_3$: after 24 h the co-conformer distribution was 78:22 *car*-**S16**:*ami*-**S16**. Slow displacement was attributed to the differences in speed bump shape between **S15/S16** and *tert*-butylbenzylamine.

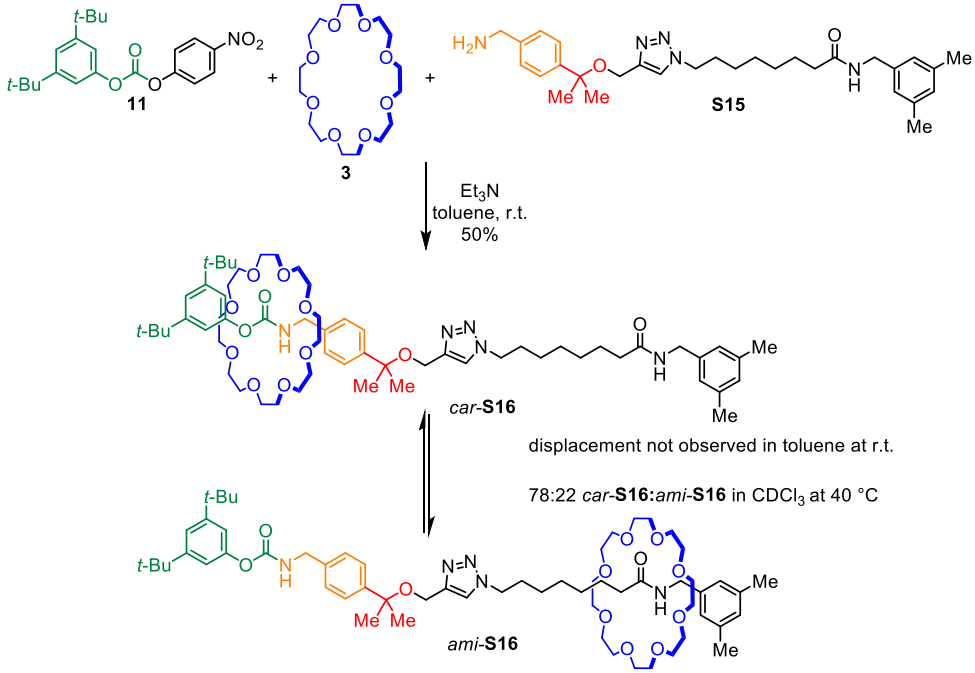
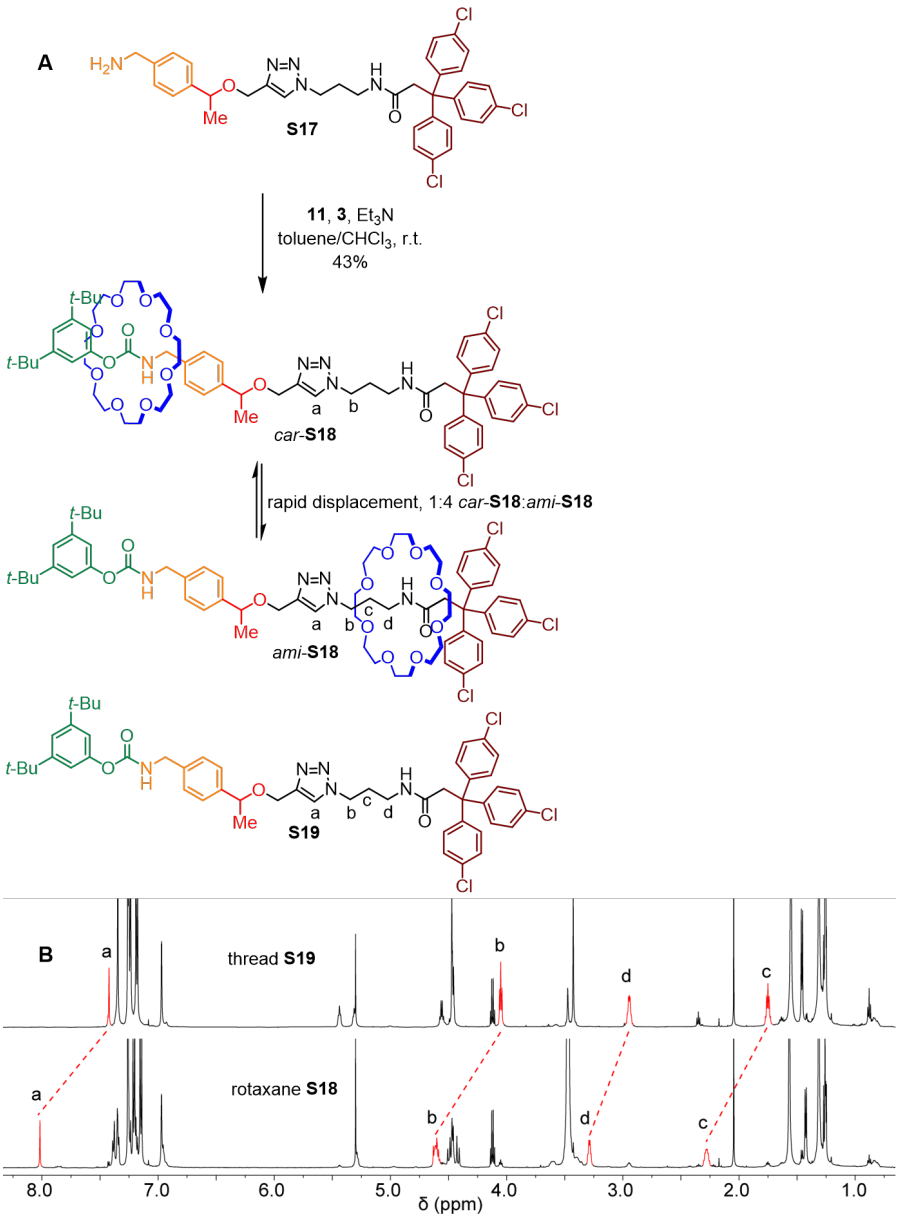


Figure S55. First model pump **S15**.

The second model pump **S17** was designed with a mono-methyl (-CH_3) speed bump to increase displacement rate (**Figure S56**). As **S15** was sparingly soluble in toluene, the 3,5-dimethylphenyl stopper was also replaced with a *tris*-(4-chlorophenyl)methine group. Previous studies found that incorporating this moiety into molecules increases solubility in toluene.[28] However, **S17** was still poorly soluble in toluene, presumably due to the amide unit. The corresponding rotaxane **S18** was accessed in 43% yield when using a mixture of toluene/$CHCl_3$ as reaction solvent. Analysis of the reaction crude showed both co-conformers *car*-**S18** and *ami*-**S18** were now present in a 1:4 ratio. Displacement was found to occur at a comparable rate to a 2D ROESY NMR experiment (**Figure S56**, **C**): cross peaks between the two co-conformers were observed in the spectrum, indicative of chemical exchange occurring over <10 min. As displacement was now too rapid, a further redesign was necessary. Significant downfield shifts (**Figure S56**, **B**) were observed in the ^{1}H NMR spectrum of *ami*-**S18** (versus free thread **S19**) for H_a (δ = +0.7 ppm) and H_b (δ = +0.6 ppm), protons proximal to the triazole linker. This is indicative of positioning of the macrocycle

close to the triazole groups.[26, 27] As **S17** was still poorly soluble in toluene, we decided to remove the amide unit in future designs and instead link multiple triazole rings to form the secondary binding station in the catchment region.



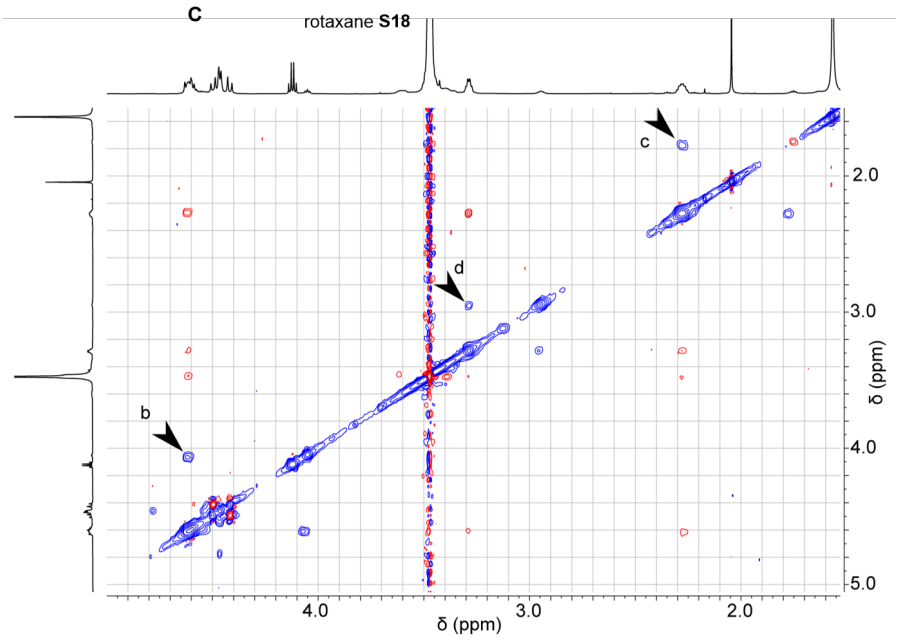


Figure S56. Second model pump **S17** and results of NMR study. (**A**) Second model pump **S17**. (**B**) Partial ^{1}H NMR ($CDCl_3$, 600 MHz, 298 K) spectra of free thread **S19** and rotaxane **S18**, revealing the presence of **3** close to the triazole unit. (**C**) ^{1}H-^{1}H ROESY NMR ($CDCl_3$, 298 K) study revealing rapid displacement of macrocycle in [2]rotaxane **S18**. Cross peaks between two co-conformers are shown by wedge-shaped signs.

We designed a third model pump, **S20**, that contained intermediate sized trifluoromethyl (-CF_3) speed bump and a catchment region consisting of two triazole units linked by a propyl (-$(CH_2)_3$-), linker (**Figure S57**). This speed bump proved to have appropriate stereoelectronics for the pump design: the half-life of pseudo-rotaxane **S21•3**, formed from the benzylic amine fragment of the pump (**Figure S57**, **A**), was 24 min. The corresponding rotaxane **S22** (this time using chemical fuel **2** as the electrophile) was formed in 51% yield, which exists as a 2:3 ratio of *car*-**S22**:*tri*-**S22** in d_8-toluene. This demonstrated that linked triazole rings are sufficient for promoting displacement of crown ethers over the speed bump. The final pump design, **1**, was a simple redesign of **S20**, where two additional triazole rings (the limit imposed by solubility) were added to the catchment region to allow more crown ethers to be pumped.

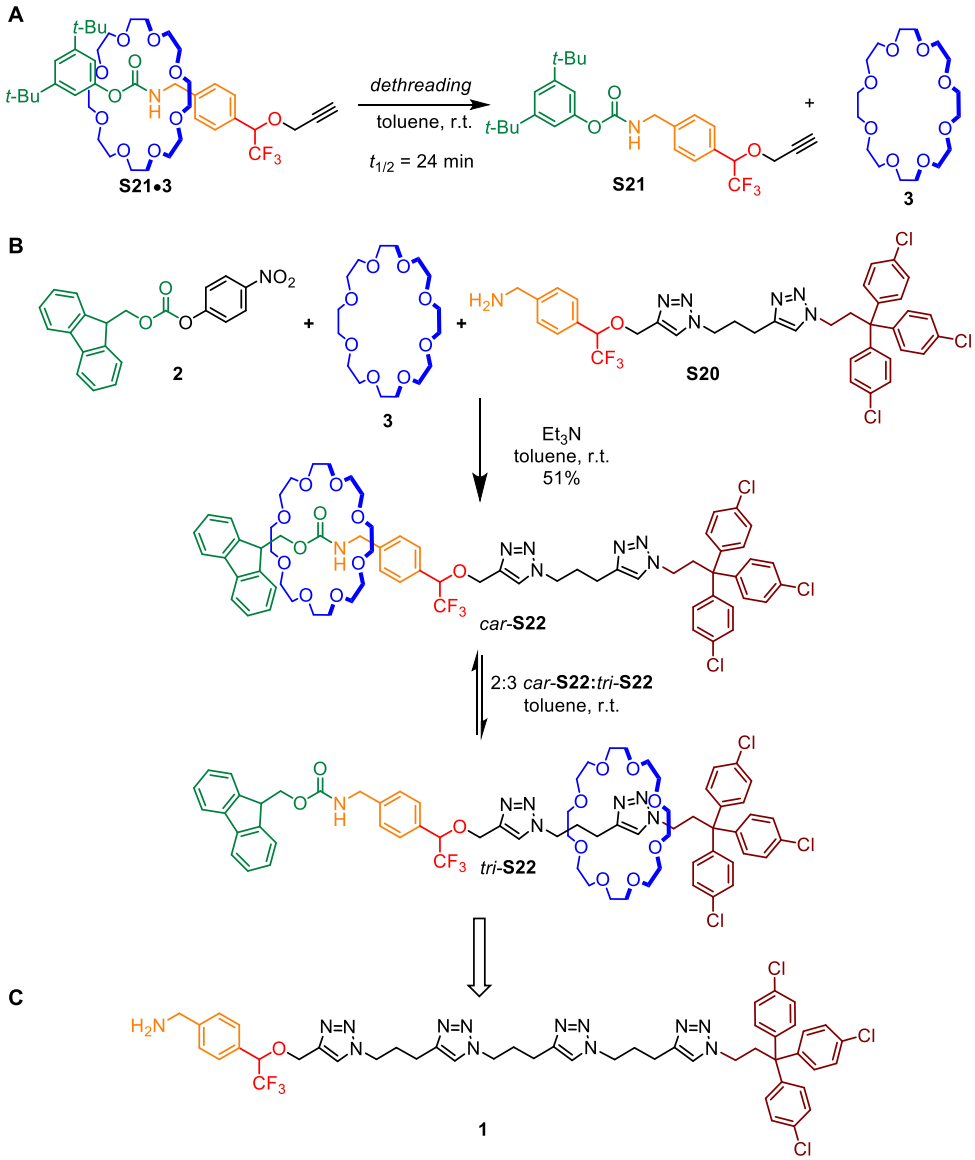


Figure S57. Third model pump **S20** and the final design. (**A**) Dethreading test of optimized speed bump design. (**B**) Third model pump **S20**. (**C**) Final pump **1**.

3.5. References

1 Skou, J. C. The identification of the sodium–potassium pump (Nobel Lecture). *Angew. Chem. Int. Ed.* **37**, 2320–2328 (1998).

2 Lodish, H., Berk, A., Zipursky, S. L., Matsudaira, P., Baltimore, D. & Darnell, J. 'Transport across Cell Membranes' in *Molecular Cell Biology* (W. H. Freeman, New York, 2000), vol. 4, chap. 15.

3 Du, D., Wang-Kan, X., Neuberger, A., van Veen, H. W., Pos, K. M., Piddock, L. J. V. & Luisi, B. F. Multidrug efflux pumps: structure, function and regulation. *Nat. Rev. Microbiol.* **16**, 523–539 (2018).

4 Qiu, Y., Feng, Y., Guo, Q.-H., Astumian, R. D. & Stoddart, J. F. Pumps through the ages. *Chem* **6**, 1952–1977 (2020).

5 Steinberg-Yfrach, G., Liddell, P. A., Hung, S.-C., Moore, A. L., Gust, D. & Moore, T. A. Conversion of light energy to proton potential in liposomes by artificial photosynthetic reaction centres. *Nature* **385**, 239–241 (1997).

6 Bennett, I. M., Farfano, H. M. V., Bogani, F., Primak, A., Liddell, P. A., Otero, L., Sereno, L., Silber, J. J., Moore, A. L., Moore, T. A. & Gust, D. Active transport of Ca^{2+} by an artificial photosynthetic membrane. *Nature* **420**, 398–401 (2002).

7 Bhosale, S., Sisson, A. L., Talukdar, P., Fürstenberg, A., Banerji, N., Vauthey, E., Bollot, G., Mareda, J., Röger, C., Würthner, F., Sakai, N. & Matile, S. Photoproduction of proton gradients with π-stacked fluorophore scaffolds in lipid bilayers. *Science* **313**, 84–86 (2006).

8 Serreli, V., Lee, C.-F., Kay, E. R. & Leigh, D. A. A molecular information ratchet. *Nature* **445**, 523–527 (2007).

9 Baroncini, M., Silvi, S., Venturi, M. & Credi, A. Photoactivated directionally controlled transit of a non-symmetric molecular axle through a macrocycle. *Angew. Chem. Int. Ed.* **51**, 4223–4226 (2012).

10 Ragazzon, G., Baroncini, M., Silvi, S., Venturi, M. & Credi, A. Light-powered autonomous and directional molecular motion of a dissipative self-assembling system. *Nat. Nanotechnol.* **10**, 70–75 (2015).

11 Cheng, C., McGonigal, P. R., Schneebeli, S. T., Li, H., Vermeulen, N. A., Ke, C. & Stoddart, J. F. An artificial molecular pump. *Nat. Nanotechnol.* **10**, 547–553 (2015).

12 Pezzato, C., Nguyen, M. T., Cheng, C., Kim, D. J., Otley, M. T. & Stoddart, J. F. An efficient artificial molecular pump. *Tetrahedron* **73**, 4849–4857 (2017).

13 Erbas-Cakmak, S., Fielden, S. D. P., Karaca, U., Leigh, D. A., McTernan, C. T., Tetlow, D. J. & Wilson, M. R. Rotary and linear molecular motors driven by pulses of a chemical fuel. *Science* **358**, 340–343 (2017).

14 Pezzato, C., Nguyen, M. T., Kim, D. J., Anamimoghadam, O., Mosca, L. & Stoddart, J. F. Controlling dual molecular pumps electrochemically. *Angew. Chem. Int. Ed.* **57**, 9325–9329 (2018).

15 Qiu, Y., Zhang, L., Pezzato, C., Feng, Y., Li, W., Nguyen, M. T., Cheng, C., Shen, D., Guo, Q.-H., Shi, Y., Cai, K., Alsubaie, F. M., Astumian, R. D. & Stoddart, J. F. A molecular dual pump. *J. Am. Chem. Soc.* **141**, 17472–17476 (2019).

16 Qiu, Y., Song, B., Pezzato, C., Shen, D., Liu, W., Zhang, L., Feng, Y., Guo, Q.-H., Cai, K., Li, W., Chen, H., Nguyen, M. T., Shi, Y., Cheng, C., Astumian, R. D., Li, X. & Stoddart, J. F. A precise polyrotaxane synthesizer. *Science* **368**, 1247–1253 (2020).

17 Astumian, R. D. & Bier, M. Mechanochemical coupling of the motion of molecular motors to ATP hydrolysis. *Biophys. J.* **70**, 637–653 (1996).

18 Kay, E. R., Leigh, D. A. & Zerbetto, F. Synthetic molecular motors and mechanical machines. *Angew. Chem. Int. Ed.* **46**, 72–191 (2007).

19 Astumian, R. D. Design principles for Brownian molecular machines: how to swim in molasses and walk in a hurricane. *Phys. Chem. Chem. Phys.* **9**, 5067–5083 (2007).

20 Astumian, R. D. Kinetic asymmetry allows macromolecular catalysts to drive an information ratchet. *Nat. Commun.* **10**, 3837 (2019).

21 Astumian, R. D. How molecular motors work – insights from the molecular machinist's toolbox: the Nobel prize in Chemistry 2016. *Chem. Sci.* **8**, 840–845 (2017).

22 Pezzato, C., Cheng, C., Stoddart, J. F. & Astumian, R. D. Mastering the non-equilibrium assembly and operation of molecular machines. *Chem. Soc. Rev.* **46**, 5491–5507 (2017).

23 Kassem, S., van Leeuwen, T., Lubbe, A. S., Wilson, M. R., Feringa, B. L. & Leigh, D. A. Artificial molecular motors. *Chem. Soc. Rev.* **46**, 2592–2621 (2017).

24 Wang, Q., Chen, D. & Tian, H. Artificial molecular machines that can perform work. *Sci. China Chem*. **61**, 1261–1273 (2018).

25 Findlay, J. A. & Crowley, J. D. Functional nanomachines: recent advances in synthetic molecular machinery. *Tetrahedron Lett.* **59**, 334–346 (2018).

26 Coutrot, F. A focus on triazolium as a multipurpose molecular station for pH-sensitive interlocked crown-ether-based molecular machines. *ChemistryOpen* **4**, 556–576 (2015).

27 Zheng, H., Zhou, W., Lv, J., Yin, X., Li, Y., Liu, H. & Li, Y. A dual-response [2]rotaxane based on a 1,2,3-triazole ring as a novel recognition station. *Chem. Eur. J.* **15**, 13253–13262 (2009).

28 Fielden, S. D. P., Leigh, D. A., McTernan, C. T., Pérez-Saavedra, B. & Vitorica-Yrezabal, I. J. Spontaneous assembly of rotaxanes from a primary amine, crown ether and electrophile. *J. Am. Chem. Soc.* **140**, 6049–6052 (2018).

29 Tian, C., Fielden, S. D. P., Whitehead, G. F. S., Vitorica-Yrezabal, I. J. & Leigh, D. A. Weak functional group interactions revealed through metal-free active template rotaxane synthesis. *Nat. Commun.* **11**, 744 (2020).

30 Tian, C., Fielden, S. D. P., Pérez-Saavedra, B., Vitorica-Yrezabal, I. J. & Leigh, D. A. Single-step enantioselective synthesis of mechanically planar chiral [2]rotaxanes using a chiral leaving group strategy. *J. Am. Chem. Soc*. **142**, 9803–9808 (2020).

31 Denis, M. & Goldup, S. M. The active template approach to interlocked molecules: Principles, progress and applications. *Nat. Rev. Chem.* **1**, 0061 (2017).

32 Wilson, M. R., Solá, J., Carlone, A., Goldup, S. M., Lebrasseur, N. & Leigh, D. A. An autonomous chemically fuelled small-molecule motor. *Nature* **534**, 235–240 (2016).

33 Fyfe, M. C. T., Glink, P. T., Menzer, S., Stoddart, J. F., White, A. J. P. & Williams, D. J. Anion assisted self-assembly. *Angew. Chem., Int. Ed. Engl.* **36**, 2068–2070 (1997).

34 della Sala, F., Neri, S., Maiti, S., Chen, J. L.-Y. & Prins, L. J. Transient self-assembly of molecular nanostructures driven by chemical fuels. *Curr. Opin. Biotech.* **46**, 27–33 (2017).

35 van Rossum, S. A. P., Tena-Solsona, M., van Esch, J. H., Eelkema, R. & Boekhoven, J. Dissipative out-of-equilibrium assembly of man-made supramolecular materials. *Chem. Soc. Rev.* **46**, 5519–5535 (2017).

36 Ragazzon, G. & Prins, L. J. Energy consumption in chemical fuel-driven self-assembly. *Nat. Nanotechnol.* **13**, 882–889 (2018).

37 Rieß, B., Grötsch, R. & Boekhoven, J. The design of dissipative molecular assemblies driven by chemical reaction cycles. *Chem* **6**, 552–578 (2020).

38 Borsley, S., Leigh, D. A. & Roberts, B. M. W. A doubly kinetically-gated information ratchet autonomously driven by carbodiimide hydration. *J. Am. Chem. Soc*. **143**, 4414–4420 (2021).

39 Astumian, R. D. Irrelevance of the power stroke for the directionality, stopping force, and optimal efficiency of chemically driven molecular machines. *Biophys. J*. **108**, 291–303 (2015).

40 Howard, J. Protein power strokes. *Curr. Biol.* **16**, R517–R519 (2006).

41 Hwang, W. & Karplus, M. Structural basis for power stroke vs. Brownian ratchet mechanisms of motor proteins. *Proc. Natl. Acad. Sci. USA* **116**, 19777–19785 (2019).

42 van Dijk, L., Tilby, M. J., Szpera, R., Smith, O. A., Bunce, H. A. P. & Fletcher, S. P. Molecular machines for catalysis. *Nat. Rev. Chem.* **2**, 0117 (2018).

43 Zhang, L., Marcos, V. & Leigh, D. A. Molecular machines with bio-inspired mechanisms. *Proc. Natl. Acad. Sci. USA* **115**, 9397–9404 (2018).

44 Biagini, C., Fielden, S. D. P., Leigh, D. A., Schaufelberger, F., Di Stefano, S. & Thomas, D. Dissipative catalysis with a molecular machine. *Angew. Chem. Int. Ed.* **58**, 9876–9880 (2019).

45 Astumian, R. D., Pezzato, C., Feng, Y., Qiu, Y., McGonigal, P. R., Cheng, C. & Stoddart, J. F. Non-equilibrium kinetics and trajectory thermodynamics of synthetic molecular pumps. *Mater. Chem.* **4**, 1304–1314 (2020).

46 Heard, A. W. & Goldup, S. M. Simplicity in the design, operation and applications of mechanically interlocked molecular machines. *ACS Cent. Sci*. **6**, 117–128 (2020).

47 Aprahamian, I. The future of molecular machines. *ACS Cent. Sci.* **6**, 347–358 (2020).

48 Moulin, E., Faour, L., Carmona-Vargas, C. C. & Giuseppone, N. From molecular machines to stimuli-responsive materials. *Adv. Mater.* **32**, e1906036 (2020).

49 Lee, J., Yu, E., Park, J., Ryu, I., Shin, D. & Gong, Y. Solid-phase synthesis of 7-arylbenzo[*b*][1,4]oxazin-3(4*H*)-one derivatives on a BOMBA resin utilizing the Smiles rearrangement. *Bull. Korean Chem. Soc.* **30**, 1325–1330 (2009).
50 Hunter, C. A. Quantifying intermolecular interactions: guidelines for the molecular recognition toolbox. *Angew. Chem. Int. Ed.* **43**, 5310–5324 (2004).

Chapter 4.

Analysis of a Synthetic Molecular Motor with Information Thermodynamics

Blank Page

Declaration

The research presented in this chapter has been uploaded to the *ChemRxiv* and submitted for publication in a peer reviewed journal with the following authorship and title:

Amano, S., Esposito, M., Kreidt, E., Leigh, D. A., Penocchio, E. & Roberts, B. M. W. Insights from an information thermodynamics analysis of a synthetic molecular motor. *ChemRxiv* doi: 10.33774/chemrxiv-2021-60k1r **(**2021).

This work will be submitted by two co-workers (Benjamin M. W. Roberts and Emanuele Penocchio) in support of an application for a degree of PhD from the University of Manchester and the University of Luxembourg, respectively.

Acknowledgements

The author initially conceived this project. Further contributions from the author to this work include the preparation of the proposal of collaboration to Prof. Massimiliano Esposito's group, preliminary theoretical analysis and the design of numerical simulation. Emanuele Penocchio did the majority of theoretical analysis and numerical simulation. Benjamin M. W. Roberts assisted with the preparation of the proposal of collaboration and the design of

numerical simulation. Dr. Elisabeth Kreidt assisted with the preparation of the proposal of collaboration and prepared the figures in the manuscript. Prof. David A. Leigh and Prof. Massimiliano Esposito directed the research. All authors contributed to analysing the results and writing and proofreading of the manuscript and Supporting Information. All the authors of the paper are gratefully thanked for their contribution. Authorship order is alphabetical.

Synopsis

Information is a physical quantity, the realisation of which transformed the physics of measurement and communication in the latter half of the 20th Century. However, the relationship and flow between information, energy and mechanics in chemical systems and mechanisms remains largely unexplored. Here we analyze a minimalist experimental example of an autonomous artificial chemically-driven molecular motor - a molecular information ratchet - in terms of information thermodynamics, a framework that quantitatively relates information to other thermodynamic parameters. This treatment reveals how directional motion is generated by free energy transfer from the chemical to the mechanical processes involving the motor. We find that the free energy transfer consists of two distinct contributions that can be considered as “energy flow” and “information flow”. We identify the efficiency with which the chemical fuel powers the free energy transfer and show that this is a useful quantity with which to compare and evaluate mechanisms of, and guide designs for, molecular machines. The study provides a thermodynamic level of understanding of molecular motors that is general, complements previous analyses based on kinetics, and has practical implications for designing and improving synthetic molecular machines, regardless of the particular type of machine or chemical structure. In particular, the study confirms that, in line with kinetic analysis, power strokes do not affect the directionality of chemically-driven molecular machines. However, we also find that under

some conditions power strokes can modulate the molecular motor current (how fast the components rotate), efficiency with respect to how free energy is dissipated, and the number of fuel molecules consumed per cycle. This may help explain the role of such conformational changes in biomolecular machine mechanisms and illustrates the interplay between energy and information in chemical systems.

4.1. Introduction

Understanding how and why a machine works in the way it does is crucial for optimizing designs and inventing new ones. For macroscopic machines such an understanding can be deduced from engineering principles and Newtonian physics. In contrast, at the molecular level there is no simple explanation for why individual components of biomachines move in a particular way. Is the movement of a specific amino acid in ATP synthase a key requisite for the mechanism, or does it occur incidentally as part of an evolutionary pathway that was overall successful? Biomolecular machines operate autonomously, apparently through Brownian ratchet mechanisms.[1–5] They use energy and information to rectify the directionality of random thermal movements of their components so that work can be performed. This type of machine includes biological pumps and motors, in which the energy is generally provided in the form of a chemical potential gradient.[6,7] Chemists are learning how to design synthetic analogues of such machines,[8–13] with examples including a minimalist autonomous chemically-fueled molecular rotary motor[14] (**Figure 1**).

Kinetic models have proved useful[15–17] in describing the behavior of molecular machines, but such analysis is inherently unable to give an account of free energy transfer from the fuel to the machine. The consideration of thermodynamics is essential to understand free energy transduction, how this drives directional motion and generates the capacity to perform work[18], and, hence, how to optimize motor's design. Up to now attempts[12,19–24] to design molecular motors have been led by chemical intuition, with little opportunity to reliably judge the effectiveness of a machine's design or performance unless and until it has been realized experimentally.

Recently, stochastic thermodynamics has emerged as a method for studying systems that operate at energies of the order of thermal fluctuations.[25–28] The theory is valid even when systems are driven far from equilibrium[25,29,30] and has been used to study

nonequilibrium nanoscale systems.[31–36] A major achievement in this field was the connection made between stochastic thermodynamics and information theory,[37] which gave birth to "information thermodynamics".[38,39] Information thermodynamics relates information to other thermodynamic quantities, such as free energy and entropy, and has proved particularly successful in resolving apparent thermodynamic paradoxes, such as Maxwell's demon.[40]

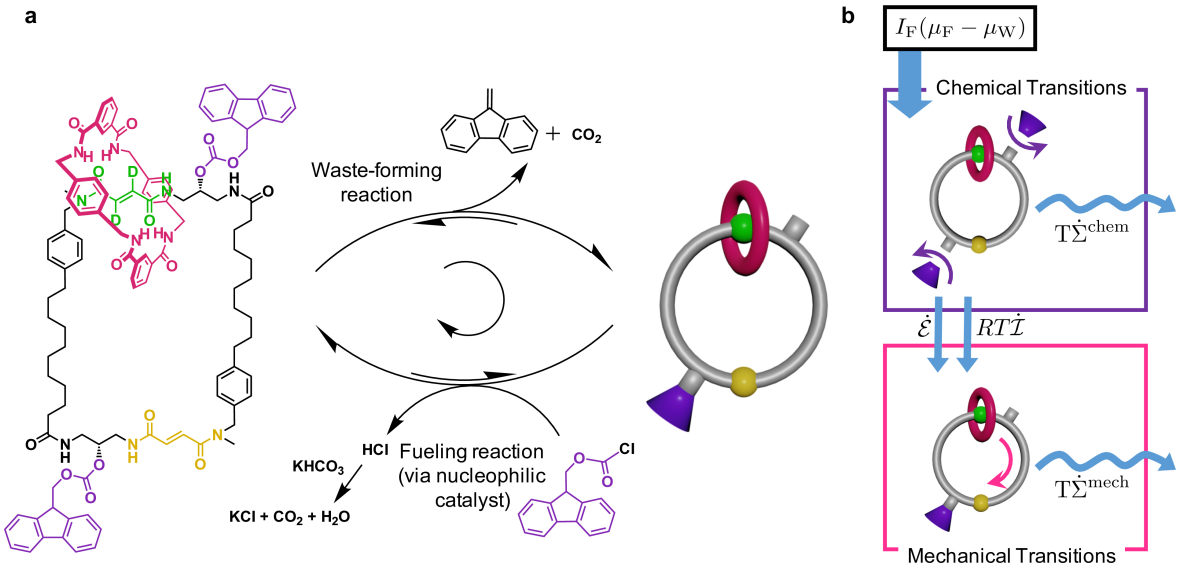


Figure 1. A Rosetta Stone for chemical (reactions and co-conformational dynamics) and information thermodynamics descriptions of a molecular motor; two distinct but complementary accounts of the processes involved in a minimalist, autonomous, chemically-fueled, molecular rotary motor. **a**, Chemical structure of the rotary motor and the chemical reactions involved in its operation. The motor comprises a benzylic amide macrocycle (magenta) and a track with two fumaramide binding sites (yellow, non-deuterated; green, deuterated for analytical purposes). The macrocycle randomly shuttles between the two fumaramide sites when its path is not blocked by Fmoc groups (purple). The fueling reaction consumes the fuel (Fmoc–Cl) and attaches an Fmoc group to the track, while the waste-forming reaction removes the Fmoc group (allowing passage of the macrocycle) and generates waste species (dibenzofulvene and CO_2). The fueling reaction is catalyzed by a pyridine-based nucleophilic catalyst (e.g., 4-dimethylaminopyridine or the bulky catalyst shown in **Figure 4a**). Note that both reactions are considered reversible, even when the backward reactions (i.e. regeneration of fuel via barrier removal and waste products reacting to give barrier) are extremely rare events.[41] **b**, Information thermodynamics description of free energy transduction in the rotary motor. In the chemical transitions, free energy is supplied to the motor ($I_F(\mu_F - \mu_W)$), part of which is dissipated ($T\dot{\Sigma}^{chem}$). I_F is the rate at which the motor reacts with the fuel, $(\mu_F - \mu_W)$ the chemical potential difference between fuel and waste species, T the temperature, $\dot{\Sigma}^{chem}$ the entropy production rate of the chemical transitions. The rest of the free energy is supplied to the mechanical transitions either as energy ($\dot{\mathcal{E}}$) or information ($RT\dot{I}$) flow, where R is the gas constant. Mechanical transitions dissipate this free energy ($T\dot{\Sigma}^{mech}$), generating directional motion of the macrocycle. When $\dot{\mathcal{E}} + RT\dot{I} = 0$, mechanical transitions are at equilibrium and no net mechanical displacement of the macrocycle can arise.

Here, we develop a quantitative understanding of the processes that drive an autonomous chemically-fueled molecular motor[14] (**Figure 1**) using an analysis that

incorporates concepts from information thermodynamics within the framework of nonequilibrium thermodynamics of open chemical reaction networks.[42,43] Contrary to common models in stochastic thermodynamics[25,30], our framework neglects fluctuations and describes macroscopic ensembles of chemical species characterized by experimentally measurable concentrations. The approach is consistent with kinetic models[16,17,44] but goes further by introducing a quantitative thermodynamic understanding of how autonomous molecular motors work. Two features, "information flow" and "energy flow", contribute to the transfer of free energy from the fuel to the machine that is the origin of current (i.e. the net rate of displacement of the macrocycle directionally along the track, see **Box I**) in the motor. The effect of changing chemical gating,[15] power strokes[45–47] ("a viscoelastic, free energy releasing, large-amplitude conformational change"[45,47]), and overall rates on current and efficiency (**Box I**) are examined through simulations, revealing design principles for molecular motors. Particular insight is gained in terms of the role of power strokes in tuning a motor's performance while remaining consistent with core aspects of kinetic models,[16,17,44–48] informing the current debate concerning the role of power strokes in biomolecular machines.[44,46,47]

As a result, the mechanism of operation of the rotary motor can be understood in several different ways: through chemical design,[14] reaction kinetics,[15,49] molecular dynamics,[50] and now, nonequilibrium information thermodynamics. Accordingly, this minimalist molecular motor can act as a "Rosetta Stone" for relating these disparate frameworks[16,17,35,44,49–51], aiding the translation of concepts and relationships between energy, information, kinetics and molecular structure.

Fueling reaction	The reaction that consumes the fuel (Fmoc−Cl) and attaches an Fmoc group to the motor's track, generating HCl as waste product. Forward ($k^{\mathrm{p/d}}_{+\mathrm{F}}$) and backward ($k^{\mathrm{p/d}}_{-\mathrm{F}}$) rate constants may depend on the macrocycle being proximal (p) or distal (d) to the reaction site.
Waste-forming reaction	The reaction that removes the Fmoc group from the motor's track, generating dibenzofulvene and CO_2 as waste products. Forward ($k^{\mathrm{p/d}}_{-\mathrm{W}}$) and backward ($k^{\mathrm{p/d}}_{+\mathrm{W}}$) rate constants may depend on the macrocycle being proximal (p) or distal (d) to the reaction site.
Chemical gating	The bias of distal over proximal rate constants in fueling and waste-forming reactions, quantified by $k^{\mathrm{d}}_{+\mathrm{F}}/k^{\mathrm{p}}_{+\mathrm{F}}$ (fueling gating) and $k^{\mathrm{p}}_{-\mathrm{W}}/k^{\mathrm{d}}_{-\mathrm{W}}$ (waste-forming gating).
Overall reaction rate	The sum of the rate constants for a particular process: *e.g.*, Fmoc addition through fueling reactions ($k^{\mathrm{d}}_{+\mathrm{F}} + k^{\mathrm{p}}_{+\mathrm{F}}$), Fmoc removal through waste-forming reactions ($k^{\mathrm{p}}_{-\mathrm{W}} + k^{\mathrm{d}}_{-\mathrm{W}}$), mechanical shuttling ($k_{+\Delta} + k_{-\Delta}$ or $k'_{+\Delta} + k'_{-\Delta}$).
Current	The net rate of directional displacement of the macrocycle along the track, quantified by equation (4).
Efficiency	The free energy available to the motor to dissipate as mechanical motion compared to the free energy available from the fuel-to-waste reaction, quantified by equation (5). This definition of efficiency does not take into account the background fuel decomposition.
Power Stroke	The viscoelastic, free energy releasing, mechanical shuttling of the macrocycle along the track. The free energy released is quantified by the change in standard chemical potential due to a net mechanical displacement: $\mu^{\circ}_{1\mathrm{D}_{\mathrm{H}}} - \mu^{\circ}_{1\mathrm{D}_{\mathrm{D}}}$ and $\mu^{\circ}_{1\mathrm{H}_{\mathrm{D}}} - \mu^{\circ}_{1\mathrm{H}_{\mathrm{H}}} \neq 0$.
Kinetic Asymmetry	The kinetic preference for one direction over the other in a chemomechanical cycle, embodied by the ratcheting constant: $K_{\mathrm{r}} = \frac{k_{+\Delta}k'_{+\Delta}(k^{\mathrm{p}}_{-\mathrm{F}}[\mathrm{HCl}]+k^{\mathrm{p}}_{-\mathrm{W}})^2([\mathrm{Fmoc-Cl}]k^{\mathrm{d}}_{+\mathrm{F}}+[\mathrm{CO_2}][\mathrm{DBF}]k^{\mathrm{d}}_{+\mathrm{W}})^2}{k_{-\Delta}k'_{-\Delta}([\mathrm{Fmoc-Cl}]k^{\mathrm{p}}_{+\mathrm{F}}+[\mathrm{CO_2}][\mathrm{DBF}]k^{\mathrm{p}}_{+\mathrm{W}})^2(k^{\mathrm{d}}_{-\mathrm{F}}[\mathrm{HCl}]+k^{\mathrm{d}}_{-\mathrm{W}})^2}$, where DBF stands for dibenzofulvene. When $K_{\mathrm{r}} > 1$, the motor preferentially cycles in the forward direction as apparent from equation (4).

Box I. Definitions as applied to the minimalist rotary motor.

4.2. Results and discussion

4.2.1. A bipartite chemical reaction network for the minimalist rotary motor

The rotary motor in **Figure 1** comprises a cyclic track with two degenerate binding sites (distinguishable by one being labelled with deuterium) for a macrocycle.[14] Fluorenylmethoxycarbonyl (Fmoc) groups, that sterically prevent passage of the macrocycle, can be attached to hydroxy residues on the track (**Figure 2**). When only one barrier is in place, macrocycle shuttling enables exchange between two co-conformers ("co-conformers" are structures that differ in the relative positions of the components[52]) that have the macrocycle either proximal or distal to the free hydroxy group. The fueling reaction kinetically discriminates between the two co-conformers, favoring the reaction of the distal co-conformer over the proximal co-conformer. Under basic conditions, the waste-forming reaction removes barriers without any chemical gating. The chemical gating afforded by the biased fueling reaction, and the free energy supplied by the fuel-to-waste conversion, result in directional movement of the macrocycle around the track. The rotary motor can be represented by a chemomechanical network of reactions (**Figure 2**) in which mechanical

and chemical transitions are coupled, as in common models for biological molecular motors[44] and minimal Brownian motors with external dichotomous noise[53]. Since no transitions enable simultaneous change in the mechanical and chemical state, this network is said to be bipartite.[39,54]

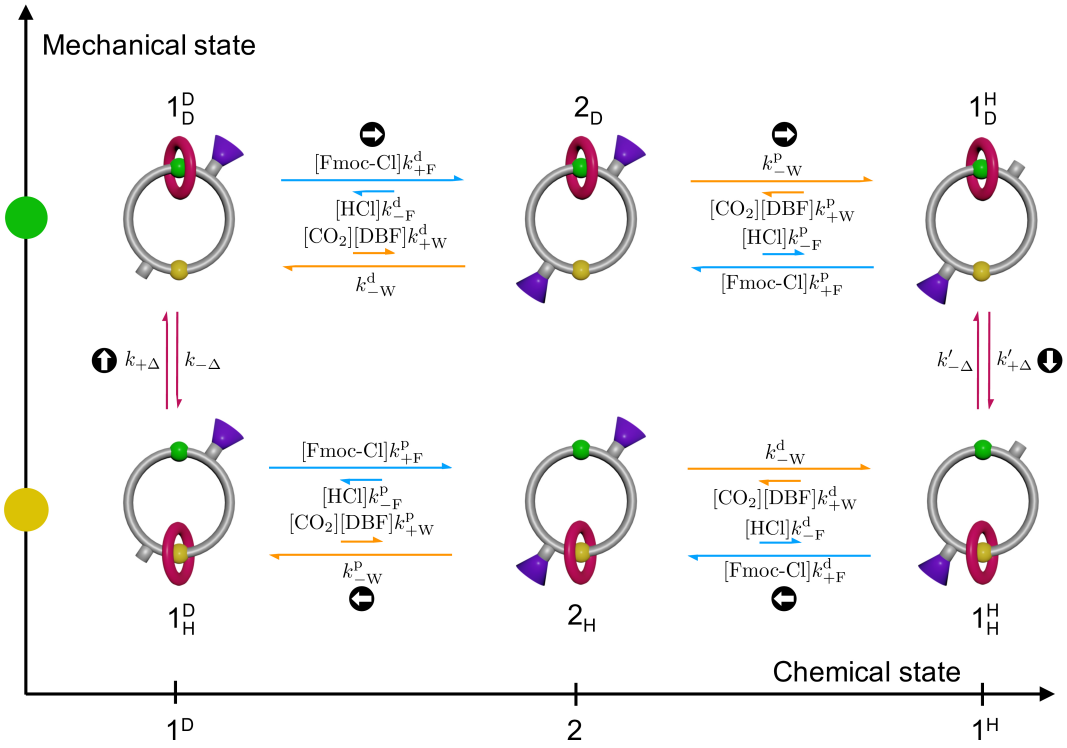


Figure 2. Rotary motor as an open and bipartite chemical reaction network. The state of the motor can be represented as a combination of two mechanical states (subscript D or H, expressing whether the macrocycle is binding to the deuterated or non-deuterated site on the track) and three chemical states (1^{D}, 2 or 1^{H}). The number in the chemical state shows the number of the Fmoc groups attached to the track; the superscript H or D denotes the binding site close to the attached Fmoc group. Note that chemical states without any Fmoc group on the track are neglected. The mechanical transitions involve displacement of the macrocycle, the rate constants for which are denoted $k_{+\Delta}$, $k_{-\Delta}$, $k'_{+\Delta}$ and $k'_{-\Delta}$. Subscript Δ signifies that they are the rate constants of mechanical transitions, which are only coupled to the thermal reservoir. Rate constants with and without the prime (′) are for the interconversion between 1^H_D and 1^H_H, and the interconversion between 1^D_D and 1^D_H, respectively. Subscript's sign shows the direction of the transition (+ for the clockwise and – for the counterclockwise direction). The rate constants of the fueling reaction and its reverse are denoted as $k^{p/d}_{+F}$ and $k^{p/d}_{-F}$, respectively. Superscript shows whether the macrocycle is proximal (p) or distal (d) to the reacting hydroxy group. The rate constants of the waste-forming reaction and its reverse are denoted as $k^{p/d}_{-W}$ and $k^{p/d}_{+W}$, respectively. In the experimental rotary motor, a clockwise current across the network is generated because $k^p_{+F} < k^d_{+F}$, $k^p_{-W} = k^d_{-W}$ and free energy is supplied to the motor by the fuel-to-waste conversion. The chemical gating in the fueling reaction arises due to the steric crowding in transition states of the proximal co-conformers (1^H_D and 1^D_H) compared to the distal co-conformers (1^D_D and 1^H_H). Circled arrows represent the most probable (i.e., most frequent) clockwise path. DBF stands for dibenzofulvene.

4.2.2. Information thermodynamic analysis

As detailed in Section 4.5., the rotary motor is modelled as the isothermal open bipartite chemical reaction network[39,42,43] shown in **Figure 2**. The concentrations of the six motor species $2_H, 2_D, 1_D^H, 1_D^D, 1_H^H, 1_H^D$ (see **Figure 2** caption) evolve according to the rate constants of each reaction following mass-action kinetics. The system is open because the concentrations of fuel (Fmoc–Cl) and waste (HCl — liberated as part of the fueling reaction, and subsequently neutralized by $KHCO_3$ that is present to produce KCl, CO_2 and H_2O — and dibenzofulvene and CO_2 produced during the waste-forming reaction) species are kept constant through addition from, or removal to, an external source. An important quantity for our analysis is the chemical potential gradient between the fuel and waste species ($\mu_{Fmoc-Cl} - \mu_{HCl} - \mu_{dibenzofulvene} - \mu_{CO_2}$) which we denote $\mu_F - \mu_W$. For any thermodynamically consistent set of parameters (see Section 4.4.), the system will evolve towards a stationary state in which the concentrations of all the motor species (i.e., all co-conformations and chemical states of the motor) are constant in time, as are the thermodynamic properties of the system. The entropy production rate (the entropy changes in the system and in the reservoirs per unit time) denoted $\dot{\Sigma}$ measures how far from equilibrium the motor operates and has to be non-negative by virtue of the Second Law of Thermodynamics.[55] When multiplied by temperature (T) it corresponds to the amount of free energy that is instantaneously dissipated by the motor ($T\dot{\Sigma}$). In a stationary state, the only kind of state that will be considered in this paper, all the free energy from the conversion of the fuel-to-waste is dissipated by the motor: $T\dot{\Sigma} = I_F(\mu_F - \mu_W) \geq 0$, where I_F is the rate at which the fuel is consumed by the motor (we neglect the fuel-to-waste background reaction).[43] This indicates that a non-null chemical potential gradient between fuel and waste species (we consider the case where $\mu_F > \mu_W$) is necessary to drive the system out of equilibrium, and produce directed motion. However, this expression gives no information as

to how free energy is consumed nor the amount of dissipation specifically devoted to sustain directional motion of the components.

To obtain a more in-depth understanding, we use information thermodynamics to split the free energy dissipation rate into two separately non-negative contributions (**Figure 1b**), one due to the chemical transitions ($T\dot{\Sigma}^{chem}$) and the other due to the mechanical transitions ($T\dot{\Sigma}^{mech}$). This is possible since the Fmoc-motor chemical reaction network is bipartite:[39]

$$T\dot{\Sigma} = \overbrace{I_{\mathrm{F}}(\mu_{\mathrm{F}} - \mu_{\mathrm{W}}) - \dot{\mathcal{E}} - RT\dot{\mathcal{I}}}^{T\dot{\Sigma}^{\mathrm{chem}} \geq 0} + \overbrace{\dot{\mathcal{E}} + RT\dot{\mathcal{I}}}^{T\dot{\Sigma}^{\mathrm{mech}} \geq 0} \geq 0\,. \tag{1}$$

This dissection of terms in equation (1) underlines that the free energy supplied by the fuel is only partially dissipated via the chemical processes as $T\dot{\Sigma}^{chem}$ while the remaining part is transferred to the mechanical processes and dissipated as $T\dot{\Sigma}^{mech}$. This transfer of free energy is composed of a standard free energy part, denoted "energy flow":

$$\dot{\mathcal{E}} = J(\mu^{\circ}_{1^{\mathrm{D}}_{\mathrm{H}}} - \mu^{\circ}_{1^{\mathrm{D}}_{\mathrm{D}}} + \mu^{\circ}_{1^{\mathrm{H}}_{\mathrm{D}}} - \mu^{\circ}_{1^{\mathrm{H}}_{\mathrm{H}}})\,, \tag{2}$$

and a mutual information part,[37] denoted "information flow":

$$RT\dot{\mathcal{I}} = JRT \log \frac{[1^{\mathrm{D}}_{\mathrm{H}}][1^{\mathrm{H}}_{\mathrm{D}}]}{[1^{\mathrm{D}}_{\mathrm{D}}][1^{\mathrm{H}}_{\mathrm{H}}]}\,, \tag{3}$$

where R is the gas constant and J is the stationary clockwise (as viewed in **Figure 2**) current at which the motor operates (see **Box I**). This current can be expressed as

$$J = k_{+\Delta}[1^{\mathrm{D}}_{\mathrm{H}}] - k_{-\Delta}[1^{\mathrm{D}}_{\mathrm{D}}] = k'_{+\Delta}[1^{\mathrm{H}}_{\mathrm{D}}] - k'_{-\Delta}[1^{\mathrm{H}}_{\mathrm{H}}] = \Gamma(K_{\mathrm{r}} - 1)\,. \tag{4}$$

where both Γ and K_r are positive quantities (for derivation, see Section 4.5.5.2.), the latter denoting the ratcheting constant (recently applied in the context of dissipative self-assembly[17,49]) which quantifies the kinetic asymmetry of the motor[15,16,56] (see **Box I**).

Under the experimental conditions in which the motor was originally operated,[14] the rotary motor is driven purely by information flow as the macrocycle binds with equal affinity

to fumaramide stations adjacent to a hydroxy group and an Fmoc group (within the detection limits of ^{1}H nuclear magnetic resonance (NMR) measurements), so there is no energy flow. However, when the standard chemical potentials of the distal and proximal co-conformers differ, energy flow arises according to equation (2). An example of this could arise if there was, say, a stabilizing interaction between the macrocycle and the Fmoc group, which would lead to an increase in energy of the macrocycle upon removal of the Fmoc-barrier close to it. As a consequence, the standard chemical potential would decrease, leading to the release of heat, each time net mechanical displacement occurs in the forward direction. This fits the definition of a "power stroke"[45–47] (see **Box I**). Therefore, according to equation (2), the energy flow accounts for the part of the fueling free energy that contributes to destabilizing the macrocycle during chemical transitions and which is subsequently dissipated in a power stroke.

Mutual information quantifies the correlation between the two parts of a bipartite system[39,54], here the chemical and the mechanical states. For instance, when $[1_D^H]$ (the concentration of the species 1_D^H) and $[1_H^D]$ are larger than $[1_H^H]$ and $[1_D^D]$, respectively, a correlation is present between the mechanical and chemical states: when the motor's chemical state is 1^{H}, its mechanical state is more likely to be D than H. Likewise, when the motor's chemical state is 1^{D}, its mechanical state is more likely to be H than D. A concentration distribution with this kind of correlation has smaller (Shannon-like[37]) entropy than one without ($[1_D^H]$ and $[1_H^D]$ equal to $[1_H^H]$ and $[1_D^D]$, respectively). Therefore, correlation between the mechanical and chemical states (mutual information) generates an entropic driving force for a directional current (from $[1_D^H]$ to $[1_H^H]$ and from $[1_H^D]$ to $[1_D^D]$). From the thermodynamic viewpoint, mutual information constitutes the entropic contribution of the free energy that comes from the fuel. As mutual information is constant in the stationary state, changes in mutual information due to the chemical processes and mechanical processes

are balanced[39]. If, as in the above situation, $\dot{I} > 0$, the chemical transitions are producing mutual information that is consumed by the mechanical ones. Therefore, according to equation (3), the information flow accounts for the part of fueling free energy that contributes to increasing the system's mutual information during chemical transitions and which is subsequently erased by mechanical shuttling.

Regimes where the free energy supplied by the chemical to the mechanical processes is exclusively due to the information flow ($RT\dot{I}$), and thus lack any energy flow ($\dot{\mathcal{E}}$), are denoted as pure Maxwell demon regimes.[39]

This analysis demonstrates that the free energy supplied to the mechanical processes by the energy and information flows, $RT\dot{I} + \dot{\mathcal{E}}$, is the origin of net directional motion of the macrocycle around the track. Indeed, in the absence of such flow ($RT\dot{I} + \dot{\mathcal{E}} = 0$) the mechanical transitions are at thermodynamic equilibrium ($\dot{\Sigma}^{mech} = 0$), meaning a zero directional shuttling current ($J = 0$; see equation (4)). From an information thermodynamics perspective, the molecular motor operates by using chemical processes to transduce the free energy supplied by the fuel into the free energy supplied to the mechanical processes. The efficiency of this transduction is the ratio of the latter to the former and is bounded between zero and one due to the non-negativity of $\dot{\Sigma}^{chem}$:

$$0 \leq \eta = \frac{RT\dot{\mathcal{I}} + \dot{\mathcal{E}}}{I_{\mathrm{F}}(\mu_{\mathrm{F}} - \mu_{\mathrm{W}})} = 1 - \frac{T\dot{\Sigma}^{\mathrm{chem}}}{I_{\mathrm{F}}(\mu_{\mathrm{F}} - \mu_{\mathrm{W}})} \leq 1 \tag{5}$$

In this setup, the energy and information flows are entirely dissipated by the shuttling of the macrocycle as $T\dot{\Sigma}^{mech}$. If the mechanical steps of the motor were to work against a force (for example if a load were attached to the macrocycle), $T\dot{\Sigma}^{mech}$ would incorporate a negative work term in addition to $RT\dot{I} + \dot{\mathcal{E}}$, which could serve to define the efficiency of the energy and information flows being converted into output work, instead of just being dissipated. The energy and information flows would thus constitute the maximum work

output that can be delivered by the motor. A traditional thermodynamic analysis of such a motor[25,26,57] would exclusively focus on the efficiency with which the input free energy supplied by the fuel-to-waste chemical potential gradient is converted into output work, thus over-estimating the maximum work output as the overall free energy input $I_F(\mu_F - \mu_W)$. The present approach refines this analysis by showing how the input-to-output transduction is mediated by the free energy transfer within the motor, whose efficiency η (equation (5)) limits the maximum work output potentially deliverable by the motor. It also formally defines a thermodynamic efficiency that can be applied to motors while they perform no appreciable output work, as is the case for most of the synthetic molecular motors made to date, and can serve to compare the efficiencies of their operation.

The framework we have outlined can also be used to re-derive previous results obtained using kinetic arguments as a consequence of the Second Law of Thermodynamics in bipartite systems. In Section 4.5.5.2., we show that the condition $T\dot{\Sigma}^{mech} = 0$ implies $K_r = 1$, while the condition $T\dot{\Sigma}^{mech} > 0$ implies $K_r \neq 1$, with forward movement when $K_r > 1$. This shows how the nonequilibrium thermodynamic framework, which focuses on energetic aspects quantified by the dissipation $T\dot{\Sigma}^{mech}$, is consistent with previous analysis[15–17,49] focusing on kinetic aspects quantified by K_r, which determines the sign of the current J according to equation (4). This reiterates the effectiveness of this information thermodynamics-based approach and, again, demonstrates the usefulness of this minimalist molecular motor as a Rosetta Stone for the translation of meaning and understanding between different frameworks for describing phenomena.

Previously,[17] the ratcheting constant has also been related to the ability of a dissipative self-assembly system to store free energy, but this connection is only valid in an operating regime where chemical transitions are significantly faster than mechanical ones. The information thermodynamics framework offers a general understanding of dissipative

chemical systems and establishes limits to the maximum work deliverable by them that are valid in any operating regime, thanks to concepts such as the efficiency η introduced in equation (5).

4.2.3. Design principles for molecular motors

To demonstrate the use of our framework as a design tool, we explored the effects of altering the design features of the rotary motor on its current and efficiency (see Section 4.5.6. for details). Basing the simulations on experimentally derived parameters[14], we established that under the experimental conditions employed, the rotary motor is driven purely by information flow as only the macrocycle distribution, rather than binding site affinity (within the detection limits of ^{1}H nuclear magnetic resonance (NMR) measurements), is altered during operation. Under experimental conditions ([motor] = 10 mM, [Fmoc–Cl] = 30 mM, which is sustained by constant addition, [bulky catalyst] = 50 mM, [Et_3N] = 15 mM, [$KHCO_3$] = 200 mM, CH_2Cl_2, room temperature),[14] the current was estimated to be $2.1 \cdot 10^{-8}$ mol dm^{-3} s^{-1}, requiring an average of seven fuel molecules per cycle per motor, though only 10^{-6}% of the free energy provided by the fuel is used to sustain the current (efficiency, $\eta = 10^{-8}$, see equation (5)).

4.2.3.1. Kinetic modifications

Varying model parameters allows consideration of the effects of potential structural and chemical changes on the rotary motor (Section 4.5.6.1.). Greater chemical gating for either the fueling (**Figure 3a**, **b**) or waste-forming reaction (Section 4.5.6.2.) increases the current and efficiency by increasing information flow. The former has been achieved by increasing the steric bulk of the barrier-formation catalyst[14,58,59] (**Figure 4a**) and the latter by catalysis of barrier removal by a proximal macrocycle (**Figure 4b**).[24] Gating of both fueling and

waste-forming reactions (Section 4.5.6.3.) was recently demonstrated in a [2]rotaxane information ratchet.[24] Inverting chemical gating is predicted to reverse the direction of the motor and could be achieved if the macrocycle activates, rather than hinders, proximal barrier formation (**Figure 4c**). In the absence of a kinetic preference, or when gating from the fueling and waste-forming reactions cancel out, the motor stops working, as this precludes information flow and hence mechanical dissipation ($\dot{\Sigma}^{mech} = 0$). These results are consistent with kinetic models where kinetic asymmetry predicts the direction of the current.[15–17]

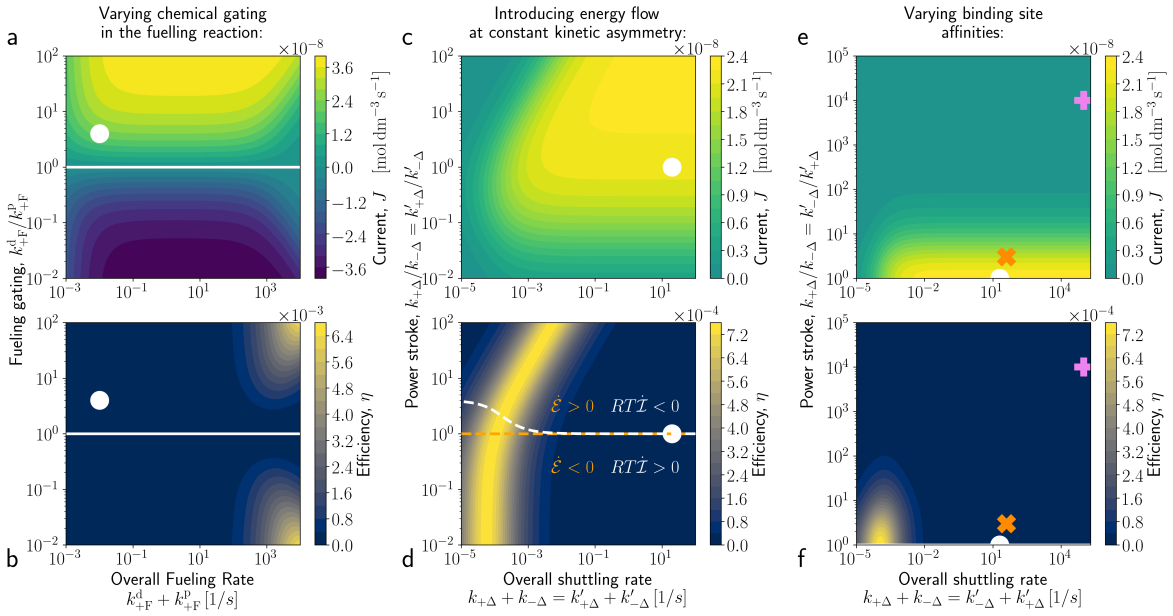


Figure 3. Numerical simulations of different molecular motor design modifications. Simulated current (J, top) and efficiency (η, bottom). Solid white lines indicate the regions of null current (and efficiency). White circles indicate the approximate parameters of the experimental rotary motor, giving a current of $2.1 \cdot 10^{-8}$ mol dm^{-3} s^{-1} and efficiency of 10^{-6}%. Directional motion in the experimental rotary motor arises because distal fueling reactions are ~4 × faster than proximal fueling reactions (chemical gating, $k^d_{+F}/k^p_{+F} \approx 4$), due to steric hindrance of the proximal transition state, while there was no chemical gating of waste-forming reactions ($k^p_{-W}/k^d_{-W} \approx 1$). **a,b**, Overall rate of the fueling reactions (x-axis) and fueling gating (y-axis) are varied. **c,d**, Overall shuttling rate (x-axis) and the ratio of forward and backward shuttling rates (indicative of power stroke magnitude, see Section 4.4.) are varied. Power strokes are introduced in a way that induces energy flow ($\dot{\mathcal{E}}$) while keeping K_r constant. Below the dashed white lines, information flow ($RT\dot{I}$) is positive and thus contributes to mechanical dissipation $\dot{\Sigma}^{mech}$, while above these lines, information flow is negative, requiring the motor to be driven by positive energy flow ($\dot{\mathcal{E}} > 0$). Negative energy flow is found below the orange dashed line. **e,f**, Overall shuttling rate (x-axis) and ratio of shuttling rate between stations (y-axis) are varied. When this ratio diverges from one (indicating non-degenerate binding sites, see Section 4.4.), both the current and efficiency drop. The pink + symbol indicates the position expected by changing one

fumaramide group to a succinamide.[60] The orange × symbol indicates a position with a 3:1 difference in binding site affinity for the macrocycle.

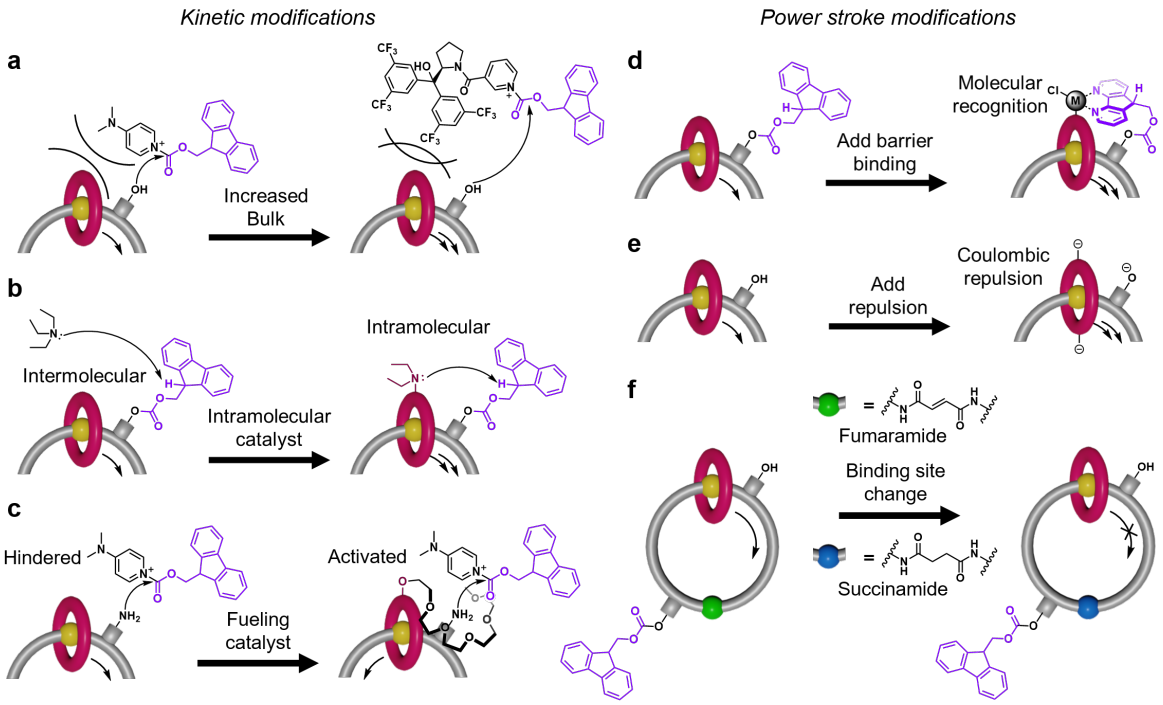


Figure 4. Potential ways of achieving different molecular motor design modifications. **a**, Increasing the steric bulk of the catalyst could be used to increase the fueling chemical gating by slowing the undesired proximal reaction. **b**, Addition of a catalyst for the waste-forming reaction to the macrocycle could increase waste-formation chemical gating by increasing the rate of the proximal waste-forming reaction. **c**, Addition of a catalyst to the macrocycle that accelerates rather than hinders proximal barrier formation could negate chemical gating arising from steric hindrance, leading to an inversion of the directionality of the motor. **d**, Allowing the formation of a complex that binds the barrier and the macrocycle could allow the introduction of a power stroke by stabilizing the co-conformer in which the macrocycle is adjacent to the barrier. **e**, Conversely, a power stroke might be introduced by destabilizing the co-conformer with the macrocycle adjacent to the site with no barrier, for example by introduction of coulombic repulsion. **f**, Swapping one fumaramide binding site for a succinamide should, in principle, be sufficient to stall the rotary motor under the previous experimental operating conditions.

4.2.3.2. Power stroke modifications

The relevance of power strokes in molecular machinery is contentious as power strokes are often experimentally observed in biological molecular motors,[44–46] but according to analysis based on kinetic asymmetry, the magnitude of the free energy released by such conformational changes does not affect the properties of chemically driven molecular machines such as directionality and stopping force, and cannot improve the efficiency of a

motor to work against an external force.[47] To reconcile differing viewpoints as to the significance of power strokes in molecular machines, we attempted to use our framework to understand the ways in which a power stroke can affect a molecular motor while staying kinetically and thermodynamically consistent. Power strokes can be used to induce energy flow (equation (2)) and, in principle, could be added to the rotary motor by introducing attractive interactions between the barrier and the macrocycle (**Figure 4d**), stabilizing distal co-conformers 1_D^D and 1_H^H, or by adding repulsive interactions between the free barrier site and the macrocycle (**Figure 4e**), destabilizing proximal co-conformers 1_H^D and 1_D^H. Our simulations show that power strokes can change the magnitude of the current and efficiency of internal energy transduction (as defined in equations (4) and (5), respectively) despite kinetic asymmetry remaining unaltered (see Section 4.4. and 4.5.6.4.). This is because, while not altering K_r, power strokes can still increase the value of Γ in equation (4), reflecting their ability to favor forward cycles by inducing energy flow out of equlibrium (**Figure 3c**, **d**). However, power strokes cannot drive directional motion in the absence of kinetic asymmetry ($RT\dot{I} = -\dot{\mathcal{E}}$ when $K_r = 1$), nor can they invert directionality while the kinetic asymmetry remains constant. In these simulations, altering power strokes and kinetic asymmetry together (Section 4.5.6.5.) gave the greatest simulated efficiency suggesting that modifying both aspects may be important for optimizing the design of synthetic molecular motors. We note that improvements that occur through power strokes arise from the induced energy flow rather than from any special role for the energetically downhill nature of the power stroke in determining the motor's behaviour. As our analysis shows, a prerequisite for the motor's operation is the availability of free energy to dissipate through mechanical motion. Therefore, any design feature that enhances free energy transduction from the chemical to the mechanical transitions could equally well foster improvements in performance.

If power strokes cancel out over the motor cycle, then no net energy flow is introduced: free energy gained from one mechanical transition is lost in the other (see Section 4.4. and 4.5.6.6.). This could be realized in molecular form by using non-degenerate binding sites, for example, changing one fumaramide site to a more weakly-binding succinamide unit (**Figure 3e**, **f** and **Figure 4f**) with a typical difference in binding energy of 23 kJ mol^{-1} (pink + symbol, **Figure 3e**, **f**, equilibrium distribution >99:1)[60] under experimental conditions similar to those used for rotary motor operation. Kinetic asymmetry remains unaltered in this scenario and, correspondingly, the direction of the motor could not be inverted in the simulation.[15–17] However, our analysis suggests that such a change would be sufficient to effectively stall the motor, if operated under the original experimental conditions, despite the unchanged kinetic asymmetry. The simulations predict that, with power strokes cancelling out, any change from degenerate binding sites lowers the current and efficiency, though a smaller difference would leave the motor functional, albeit less effective. A 3:1 bias[58] (orange × symbol, **Figure 3e**, **f**) is predicted to reduce the current by ~20%, rendering a design with non-degenerate binding sites plausible but less effective than a motor with binding sites of equal affinities.

In all of the cases considered, the highest efficiencies are predicted for when the rates of all the forward processes are approximately equal, leading to a cycle with no single rate-limiting step. Rate-limiting mechanical steps promote futile cycles, in which fuel is consumed without taking a forward step, as the unfavorable fueling reaction is kinetically favored over shuttling, decreasing both current and efficiency. Rate-limiting chemical reactions decrease the information flow in the steady-state since fast shuttling hinders the generation of a concentration bias compared to mechanical equilibrium, resulting in lower thermodynamic efficiency without reducing current or substantially changing fuel consumption per cycle. The strong dependence of efficiency on shuttling rate indicates that,

like macroscopic engines, the efficiency of a motor will be dependent on the load it is working against.[35] To use molecular machines most efficiently, they must either be tailored to the job they perform — such as using diesel engines for heavy loads — or they must use the equivalent of gears for macroscopic engines, to ensure they are working under optimal conditions.

4.3. Conclusions

Information thermodynamics-based analysis of a minimalist autonomous chemically-driven molecular motor shows how information and energy flow, the two components of free energy transfer from chemical to mechanical transitions, enable the generation of directional motion from free energy supplied by a chemical fuel. The experimental rotary motor is a pure "chemical Maxwell's demon", as information flow is the sole driving force. However, energy flow could potentially be introduced using power strokes, one of several design variations explored using our model. The predicted effect of energy flow is in line with observations made in biological motors[44] and contributes to the ongoing debate regarding the role of power strokes in molecular motors[46–48]. Information thermodynamics confirms that, in line with kinetic analysis, power strokes do not affect some key properties of chemically-driven molecular machines such as directionality. However, the magnitude of power strokes is able to affect the magnitude of the current (how fast the motor components rotate), the efficiency in terms of how free energy is dissipated, and the number of fuel molecules consumed per cycle. But these results should not be misinterpreted as supporting special significance of power strokes compared to other processes in the chemomechanical cycle. The information thermodynamics framework used in this paper should be generally applicable to other types of synthetic molecular machines, such as non-autonomous[10,12] and light-driven[8,9,11] motors, providing a quantitative basis through which to compare molecular

machine designs. Additionally, it could, in principle, be extended to other types of (supra)molecular systems (such as dissipative self-assembly[17]) powered by chemical engines[61]. We have uncovered for the first time significant roles for 'energy flow' and 'information flow' in the mechanism of the transduction of free energy from chemical reactions by molecular machinery, although the exact nature of the connection of energy and information flow to energy and information ratcheting[1,62] remains to be clarified. The minimalist autonomous chemically-driven molecular motor acts as a Rosetta Stone for relating energy, information, kinetics and molecular structure by aiding the translation of concepts and relationships between the 'languages' (i.e. frameworks) of chemical kinetics, thermodynamics and chemical reactions.

4.4. Methods

Wegscheider's conditions: To ensure that the system reaches thermodynamic equilibrium (detailed balance), when there is no chemical potential gradient between fuel and waste species ($\mu_F - \mu_W = 0$), the rate constants must satisfy Wegscheider's conditions.[42] These are equivalent to the constraints on the rate constants imposed in previous kinetic analyses[4,16,48] and dictate that the product of the forward rate constants along each independent cyclic pathway of reactions in the network, with neither net consumption nor net production of fuel or waste species, must equal the product of the corresponding backward rate constants:

$$k^{\mathrm{p}}_{-\mathrm{F}} k_{+\Delta} k^{\mathrm{d}}_{+\mathrm{F}} k^{\mathrm{p}}_{-\mathrm{F}} k'_{+\Delta} k^{\mathrm{d}}_{+\mathrm{F}} = k^{\mathrm{p}}_{+\mathrm{F}} k_{-\Delta} k^{\mathrm{d}}_{-\mathrm{F}} k^{\mathrm{p}}_{+\mathrm{F}} k'_{-\Delta} k^{\mathrm{d}}_{-\mathrm{F}} \tag{M1a}$$

$$k^{\mathrm{p}}_{-\mathrm{W}} k_{+\Delta} k^{\mathrm{d}}_{+\mathrm{W}} k^{\mathrm{p}}_{-\mathrm{W}} k'_{+\Delta} k^{\mathrm{d}}_{+\mathrm{W}} = k^{\mathrm{p}}_{+\mathrm{W}} k_{-\Delta} k^{\mathrm{d}}_{-\mathrm{W}} k^{\mathrm{p}}_{+\mathrm{W}} k'_{-\Delta} k^{\mathrm{d}}_{-\mathrm{W}} , \tag{M1b}$$

see Section 4.5.3. for derivation. These conditions were always imposed in numerical simulations to guarantee thermodynamic consistency.[41]

Local Detailed Balance: In **Figure 3**, variations in the power stroke magnitude have been related to changes in the ratio of shutting rate constants by virtue of the so-called "principle of local detailed balance" (see below), which relates the log-ratio of forward and backward rate constants of a single chemical reaction to the difference in standard chemical potentials between its reagents and products[42] (see Section 4.5.3., equation S.33). For instance, it implies the relation

$$RT \log\left(\frac{k_{+\Delta}k'_{+\Delta}}{k_{-\Delta}k'_{-\Delta}}\right) = \mu^{\circ}_{1^{\mathrm{D}}_{\mathrm{H}}} - \mu^{\circ}_{1^{\mathrm{D}}_{\mathrm{D}}} + \mu^{\circ}_{1^{\mathrm{H}}_{\mathrm{D}}} - \mu^{\circ}_{1^{\mathrm{H}}_{\mathrm{H}}}\,, \tag{M2}$$

which was employed in numerical simulations.

In addition, Wegscheider's conditions (equations (M1a) and (M1b)) imply that a variation in the power stroke magnitude must always be compensated by a variation in the fueling and waste-forming rate constants, since the following constraint (equation (M3)) must always hold for thermodynamic consistency:

$$\frac{k_{+\Delta}}{k_{-\Delta}} \cdot \frac{k'_{+\Delta}}{k'_{-\Delta}} = \frac{k^{\mathrm{p}}_{+\mathrm{F}}k^{\mathrm{d}}_{-\mathrm{F}}}{k^{\mathrm{p}}_{-\mathrm{F}}k^{\mathrm{d}}_{+\mathrm{F}}} \cdot \frac{k^{\mathrm{p}}_{+\mathrm{W}}k^{\mathrm{d}}_{-\mathrm{W}}}{k^{\mathrm{p}}_{-\mathrm{W}}k^{\mathrm{d}}_{+\mathrm{W}}}\,. \tag{M3}$$

When energy-flow-inducing power strokes were introduced in **Figure 3c** and **d**, the constraint in equation (M3) was imposed by changing rate constants k^{p}_{+W} and k^{p}_{+F} according to variations in shuttling rate constants. By doing so, kinetic asymmetry (K_r) is not altered during the simulation, but the magnitude of the current in equation (4) can still change by virtue of alterations in the value of the positive factor Γ (for its mathematical expression, see Section 4.5.5.2.). Instead, in Section 4.5.6.5. (**Figure S5**) rate constants k^{p}_{+W} and k^{p}_{-F} were changed to vary energy flow and K_r together. Experimentally, this could correspond to introducing an interaction between the macrocycle and the Fmoc group that affects (**Figure S5**) or not (**Figure 3c** and **d**) the transition state of the proximal fueling reaction, without affecting the transition state of the proximal waste-forming reaction. Note that, when binding

affinities are modified as in **Figure 3e** and **f**, the left-hand side of equation (M3) stays constant and the constraint is automatically satisfied.

We end by noting that the terminology "local detailed balance" comes from statistical physics[63], where it has become the central concept to formulate thermodynamically consistent dynamics[64,65]. Its chemical counterpart[42] (see Section 4.5.3., equation S.33) is fully equivalent to usual conditions imposed on the rate constants to ensure that microscopic reversibility holds[47] (see for instance equation 5 in reference 16). We note that use of the term "local detailed balance" in this context is considered contentious by some.[16] For a more detailed discussion see Section 4.5.3..

4.5. Details of theoretical analysis

4.5.1. The model

As stated in Section 4.2., we treat the autonomous chemically-fueled molecular motor[14] as an open bipartite chemical reaction network. We consider an isothermal, isobaric, and well-stirred ideal dilute solution containing species undergoing the 10 reactions represented in **Figure 2**, which we collect in the vector $\boldsymbol{\rho} = \{1F^{(p)}, 1W^{(p)}, 2\Delta, 3F^{(d)}, 3W^{(d)}, 4F^{(p)}, 4W^{(p)}, 5\Delta', 6F^{(d)}, 6W^{(d)}\}$:

$$
\begin{aligned}
&1\mathrm{F}^{(\mathrm{p})}: \quad 2_{\mathrm{H}} + \mathrm{HCl} \underset{k^{\mathrm{p}}_{+\mathrm{F}}}{\overset{k^{\mathrm{p}}_{-\mathrm{F}}}{\rightleftharpoons}} 1^{\mathrm{D}}_{\mathrm{H}} + \mathrm{Fmoc{-}Cl} \qquad
&&4\mathrm{F}^{(\mathrm{p})}: \quad 2_{\mathrm{D}} + \mathrm{HCl} \underset{k^{\mathrm{p}}_{+\mathrm{F}}}{\overset{k^{\mathrm{p}}_{-\mathrm{F}}}{\rightleftharpoons}} 1^{\mathrm{H}}_{\mathrm{D}} + \mathrm{Fmoc{-}Cl} \\
&1\mathrm{W}^{(\mathrm{p})}: \quad 2_{\mathrm{H}} \underset{k^{\mathrm{p}}_{+\mathrm{W}}}{\overset{k^{\mathrm{p}}_{-\mathrm{W}}}{\rightleftharpoons}} 1^{\mathrm{D}}_{\mathrm{H}} + \mathrm{DBF} + \mathrm{CO_2} \qquad
&&4\mathrm{W}^{(\mathrm{p})}: \quad 2_{\mathrm{D}} \underset{k^{\mathrm{p}}_{+\mathrm{W}}}{\overset{k^{\mathrm{p}}_{-\mathrm{W}}}{\rightleftharpoons}} 1^{\mathrm{H}}_{\mathrm{D}} + \mathrm{DBF} + \mathrm{CO_2} \\
&2\Delta: \quad 1^{\mathrm{D}}_{\mathrm{H}} \underset{k_{-\Delta}}{\overset{k_{+\Delta}}{\rightleftharpoons}} 1^{\mathrm{D}}_{\mathrm{D}} \qquad
&&5\Delta': \quad 1^{\mathrm{H}}_{\mathrm{D}} \underset{k'_{-\Delta}}{\overset{k'_{+\Delta}}{\rightleftharpoons}} 1^{\mathrm{H}}_{\mathrm{H}} \\
&3\mathrm{F}^{(\mathrm{d})}: \quad 1^{\mathrm{D}}_{\mathrm{D}} + \mathrm{Fmoc{-}Cl} \underset{k^{\mathrm{d}}_{-\mathrm{F}}}{\overset{k^{\mathrm{d}}_{+\mathrm{F}}}{\rightleftharpoons}} 2_{\mathrm{D}} + \mathrm{HCl} \qquad
&&6\mathrm{F}^{(\mathrm{d})}: \quad 1^{\mathrm{H}}_{\mathrm{H}} + \mathrm{Fmoc{-}Cl} \underset{k^{\mathrm{d}}_{-\mathrm{F}}}{\overset{k^{\mathrm{d}}_{+\mathrm{F}}}{\rightleftharpoons}} 2_{\mathrm{H}} + \mathrm{HCl} \\
&3\mathrm{W}^{(\mathrm{d})}: \quad 1^{\mathrm{D}}_{\mathrm{D}} + \mathrm{DBF} + \mathrm{CO_2} \underset{k^{\mathrm{d}}_{-\mathrm{W}}}{\overset{k^{\mathrm{d}}_{+\mathrm{W}}}{\rightleftharpoons}} 2_{\mathrm{D}} \qquad
&&6\mathrm{W}^{(\mathrm{d})}: \quad 1^{\mathrm{H}}_{\mathrm{H}} + \mathrm{DBF} + \mathrm{CO_2} \underset{k^{\mathrm{d}}_{-\mathrm{W}}}{\overset{k^{\mathrm{d}}_{+\mathrm{W}}}{\rightleftharpoons}} 2_{\mathrm{H}}
\end{aligned}
$$

The abbreviation DBF stands for dibenzofulvene. The symbols Δ and Δ' label thermal shuttlings (i.e., mechanical processes) which change the mechanical state of the system,

while all the other reactions constitute the chemical processes changing the chemical state of the system. Note that the latter are not elementary reactions, as detailed later in this section. The system is said to be open because during the experiment fuel (F = {Fmoc−Cl}) and waste (W = {HCl, DBF, CO_2}) species undergo other (not included in $\boldsymbol{\rho}$) processes: Fmoc−Cl is constantly added to the solution to compensate for its depletion, CO_2 exits the system as a gas, and HCl is readily buffered by the presence of $KHCO_3$ in the solution. With the exception of DBF, whose accumulation does not affect the rest of the system as it is unreactive under the experimental conditions, the concentrations of fuel and waste species are kept approximately constant by external processes. This is modeled with the concept of *chemostats*: large chemical reservoirs in contact with the system which fix the chemical potentials (and thus the concentrations in the ideal dilute solution setting) of F and W species in the system. These species are therefore said to be chemostatted.

In the experimental case, the following symmetries between rate constants hold:

$$k_{+\Delta} = k_{-\Delta} = k'_{+\Delta} = k'_{-\Delta} \coloneqq k_{\Delta}\,, \tag{S.1a}$$

$$k^{\mathrm{P}}_{+\mathrm{W}} = k^{\mathrm{d}}_{+\mathrm{W}} \coloneqq k_{+\mathrm{W}} \quad \text{and} \quad k^{\mathrm{P}}_{-\mathrm{W}} = k^{\mathrm{d}}_{-\mathrm{W}} \coloneqq k_{-\mathrm{W}}\,. \tag{S.1b}$$

Coarse graining

Consider the fueling reaction pathway, catalyzed by a pyridine-based nucleophilic catalyst (denoted Cat):

$$\mathrm{Fmoc{-}Cl} + \mathrm{Cat} \underset{k_{-1\mathrm{F}}}{\overset{k_{+1\mathrm{F}}}{\rightleftharpoons}} \mathrm{Fmoc{-}Cat \cdot Cl} \tag{S.2a}$$

$$\mathrm{Fmoc{-}Cat \cdot Cl} + \underbrace{\mathrm{HO{-}R}}_{e.g.:\,1^{\mathrm{D}}_{\mathrm{D}}} \underset{k_{-2\mathrm{F}}}{\overset{k_{+2\mathrm{F}}}{\rightleftharpoons}} \mathrm{Fmoc{-}Cat{-}O{-}R} + \mathrm{HCl} \tag{S.2b}$$

$$\mathrm{Fmoc{-}Cat{-}O{-}R} \underset{k_{-3\mathrm{F}}}{\overset{k_{+3\mathrm{F}}}{\rightleftharpoons}} \underbrace{\mathrm{Fmoc{-}O{-}R}}_{e.g.:\,2_{\mathrm{D}}} + \mathrm{Cat} \tag{S.2c}$$

which has the net effect of adding one Fmoc group to the motor. In principle, we could explicitly include the above steps in our analysis and this would not change the main outcomes of the work, thus uselessly complicating the mathematical treatment. For the sake

of simplifying our model without losing its relevant thermodynamic features, we then consider a coarse-grained version of the fueling process based on the assumption that k_{+2F} and k_{+3F} are much larger than k_{-2F} and k_{-3F}. In addition, given that in the original experiment Fmoc−Cl is continuously added to the system, the catalyst Cat, whose total amount is denoted by L_{Cat} can be assumed to be saturated by the Fmoc−Cl ($[Fmoc - Cat \cdot Cl] \approx L_{Cat}$). This is an additional level of control on the fueling process, since the availability of Fmoc−Cl to react with the motor species is regulated by the amount of catalyst in the system. As a consequence of these assumptions, scheme (S.2) can be coarse-grained via a standard steady-state approximation of the intermediates to get[66]:

$$\mathrm{HO{-}R + Fmoc{-}Cl} \underset{k_{-F}}{\overset{k_{+F}}{\rightleftharpoons}} \mathrm{Fmoc{-}O{-}R + HCl} \tag{S.3}$$

where we defined the rate constants

$$k_{+F} = \Omega \cdot k_{+1F} k_{+2F} k_{+3F} \quad , \qquad k_{-F} = \Omega \cdot k_{-1F} k_{-2F} k_{-3F} \, , \tag{S.4}$$

with the kinetic factor

$$\Omega = \frac{L_{\mathrm{Cat}}}{[\mathrm{Fmoc{-}Cl}] k_{+1F} k_{+3F} + [\mathrm{HCl}][\mathrm{Fmoc{-}Cl}] k_{+1F} k_{-2F} + [\mathrm{HCl}][\mathrm{Fmoc{-}O{-}R}] k_{-2F} k_{-3F}} \approx \frac{L_{\mathrm{Cat}}}{[\mathrm{Fmoc{-}Cl}] k_{+1F} k_{+3F}} \, . \tag{S.5}$$

The final approximation highlights that, in the experimental conditions, the fueling rate is strongly determined by $[Fmoc - Cat \cdot Cl] \approx L_{Cat}$, regardless the actual value of [Fmoc−Cl]. Indeed, we could also have considered the species Fmoc−Cat·Cl as the actual chemostatted fuel without altering the results of our analysis. However, the way we proceeded here is more general and equation (S.3) holds also for situations where the catalysts is not saturated[67].

Similar considerations apply to the waste-forming reaction pathway catalyzed by triethylamine (NEt_3), which has the net effect of removing one Fmoc group from the track. Let us consider the following reaction pathway:

$$\underbrace{\text{Fmoc}-\text{O}-\text{R}}_{e.g.:\ 2_D} + \text{NEt}_3 \underset{k_{-1W}}{\overset{k_{+1W}}{\rightleftharpoons}} \text{HNEt}_3^+ \cdot \text{Fmoc}-\text{O}-\text{R}^- \tag{S.6a}$$

$$\text{HNEt}_3^+ \cdot \text{Fmoc}-\text{O}-\text{R}^- \underset{k_{-2W}}{\overset{k_{+2W}}{\rightleftharpoons}} \text{HNEt}_3^+ \cdot \text{R}-\text{OCOO}^- + \text{DBF} \tag{S.6b}$$

$$\text{HNEt}_3^+ \cdot \text{R}-\text{OCOO}^- \underset{k_{-3W}}{\overset{k_{+3W}}{\rightleftharpoons}} \text{HNEt}_3^+ \cdot \text{R}-\text{O}^- + \text{CO}_2 \tag{S.6c}$$

$$\text{HNEt}_3^+ \cdot \text{R}-\text{O}^- \underset{k_{-4W}}{\overset{k_{+4W}}{\rightleftharpoons}} \underbrace{\text{R}-\text{OH}}_{e.g.:\ 1_D^D} + \text{NEt}_3 \tag{S.6d}$$

By considering the forward processes to be faster than the backward ones, here the coarse-graining of the intermediates through a stationary state approximation leads to

$$\text{Fmoc}-\text{O}-\text{R} \underset{k_{+W}}{\overset{k_{-W}}{\rightleftharpoons}} \text{R}-\text{OH} + \text{DBF} + \text{CO}_2 \tag{S.7}$$

where

$$k_{-W} = \Psi \cdot k_{+1W} k_{+2W} k_{+3W} k_{+4W} \quad , \qquad k_{+W} = \Psi \cdot k_{-1W} k_{-2W} k_{-3W} k_{-4W} \, , \tag{S.8}$$

with kinetic factor

$$\Psi = \frac{L_{\text{NEt}_3}}{k_{+2W}k_{+3W}k_{+4W} + k_{-1W}(k_{+3W}k_{+4W} + [\text{DBF}]k_{-2W}([\text{CO}_2]k_{-3W} + k_{+4W}))} \approx \frac{L_{\text{NEt}_3}}{k_{+3W}k_{+4W}(k_{-1W} + k_{+2W})} \, . \tag{S.9}$$

The symbol L_{NEt_3} denotes the total amount of NEt_3 added to the system to catalyze the waste-forming process. The final approximation is based on the assumption that k_{-2W} and k_{-3W} are small compared to the other rate constants.

Expressions (S.3) and (S.7) are commonly adopted in the literature to model this kind of fueling and waste-forming processes[15,16,49]. Here, we have shown under which (reasonable) assumptions they can be considered as the coarse-grained versions of elementary reaction pathways, whose fast dynamics can be safely neglected. We stress that the results in the following do not rely on the chosen elementary model and can be extended also to more complicated schemes. Crucially, it has been shown that the coarse-graining procedure described above is thermodynamically consistent, namely it exactly retains the stationary state entropy production of the system, which is the key observable in our analysis[67,68].

4.5.2. Dynamics

The evolution in time of the concentrations of the motor species $\boldsymbol{X} = \{2_H, 2_D, 1_H^H, 1_D^H, 1_H^D, 1_D^D\}$ is ruled by the following rate equations. Note that, from now on, we will abbreviate Fmoc−Cl as F, as it is the only fuel species.

$$\underbrace{\mathrm{d}_t \begin{pmatrix} [2_{\mathrm{H}}] \\ [2_{\mathrm{D}}] \\ [1_{\mathrm{H}}^{\mathrm{D}}] \\ [1_{\mathrm{D}}^{\mathrm{D}}] \\ [1_{\mathrm{H}}^{\mathrm{H}}] \\ [1_{\mathrm{D}}^{\mathrm{H}}] \end{pmatrix}}_{[X]} = \underbrace{\begin{array}{c} \begin{matrix} 1\mathrm{F}^{(\mathrm{p})} & 1\mathrm{W}^{(\mathrm{p})} & 2\Delta & 3\mathrm{F}^{(\mathrm{d})} & 3\mathrm{W}^{(\mathrm{d})} & 4\mathrm{F}^{(\mathrm{p})} & 4\mathrm{W}^{(\mathrm{p})} & 5\Delta' & 6\mathrm{F}^{(\mathrm{d})} & 6\mathrm{W}^{(\mathrm{d})} \end{matrix} \\ \begin{pmatrix} -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & -1 & -1 & 0 & 0 & 0 \\ 1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & -1 & 0 & 0 \end{pmatrix} \end{array}}_{\mathbb{S}^{\mathrm{X}}} \cdot \underbrace{\begin{pmatrix} k_{-\mathrm{F}}^{\mathrm{p}}[2_{\mathrm{H}}][\mathrm{HCl}] - k_{+\mathrm{F}}^{\mathrm{p}}[\mathrm{F}][1_{\mathrm{H}}^{\mathrm{D}}] \\ k_{-\mathrm{W}}^{\mathrm{p}}[2_{\mathrm{H}}] - k_{+\mathrm{W}}^{\mathrm{p}}[\mathrm{CO_2}][\mathrm{DBF}][1_{\mathrm{H}}^{\mathrm{D}}] \\ k_{+\Delta}[1_{\mathrm{H}}^{\mathrm{D}}] - k_{-\Delta}[1_{\mathrm{D}}^{\mathrm{D}}] \\ k_{+\mathrm{F}}^{\mathrm{d}}[\mathrm{F}][1_{\mathrm{D}}^{\mathrm{D}}] - k_{-\mathrm{F}}^{\mathrm{d}}[\mathrm{HCl}][2_{\mathrm{D}}] \\ k_{+\mathrm{W}}^{\mathrm{d}}[\mathrm{CO_2}][\mathrm{DBF}][1_{\mathrm{D}}^{\mathrm{D}}] - k_{-\mathrm{W}}^{\mathrm{d}}[2_{\mathrm{D}}] \\ k_{-\mathrm{F}}^{\mathrm{p}}[\mathrm{HCl}][2_{\mathrm{D}}] - k_{+\mathrm{F}}^{\mathrm{p}}[\mathrm{F}][1_{\mathrm{H}}^{\mathrm{D}}] \\ k_{-\mathrm{W}}^{\mathrm{p}}[2_{\mathrm{D}}] - k_{+\mathrm{W}}^{\mathrm{p}}[\mathrm{CO_2}][\mathrm{DBF}][1_{\mathrm{H}}^{\mathrm{D}}] \\ k'_{+\Delta}[1_{\mathrm{D}}^{\mathrm{H}}] - k'_{-\Delta}[1_{\mathrm{H}}^{\mathrm{H}}] \\ k_{+\mathrm{F}}^{\mathrm{d}}[\mathrm{F}][1_{\mathrm{H}}^{\mathrm{H}}] - k_{-\mathrm{F}}^{\mathrm{d}}[\mathrm{HCl}][2_{\mathrm{H}}] \\ k_{+\mathrm{W}}^{\mathrm{d}}[\mathrm{CO_2}][\mathrm{DBF}][1_{\mathrm{H}}^{\mathrm{H}}] - k_{-\mathrm{W}}^{\mathrm{d}}[2_{\mathrm{H}}] \end{pmatrix}}_{J = J_+ - J_-} . \tag{S.10}$$

The balance equations for concentrations of chemostatted species $\boldsymbol{Y} = \{F, HCl, DBF, CO_2\}$ read

$$\mathbf{0} = \mathrm{d}_t \underbrace{\begin{pmatrix} [\mathrm{F}] \\ [\mathrm{HCl}] \\ [\mathrm{DBF}] \\ [\mathrm{CO_2}] \end{pmatrix}}_{[Y]} = \underbrace{\begin{array}{c} \begin{matrix} 1\mathrm{F}^{(\mathrm{p})} & 1\mathrm{W}^{(\mathrm{p})} & 2\Delta & 3\mathrm{F}^{(\mathrm{d})} & 3\mathrm{W}^{(\mathrm{d})} & 4\mathrm{F}^{(\mathrm{p})} & 4\mathrm{W}^{(\mathrm{p})} & 5\Delta' & 6\mathrm{F}^{(\mathrm{d})} & 6\mathrm{W}^{(\mathrm{d})} \end{matrix} \\ \begin{pmatrix} 1 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & -1 \end{pmatrix} \end{array}}_{\mathbb{S}^{\mathrm{Y}}} \cdot \underbrace{\begin{pmatrix} k_{-\mathrm{F}}^{\mathrm{p}}[\mathrm{HCl}][2_{\mathrm{H}}] - k_{+\mathrm{F}}^{\mathrm{p}}[\mathrm{F}][1_{\mathrm{H}}^{\mathrm{D}}] \\ k_{-\mathrm{W}}^{\mathrm{p}}[2_{\mathrm{H}}] - k_{+\mathrm{W}}^{\mathrm{p}}[\mathrm{CO_2}][\mathrm{DBF}][1_{\mathrm{H}}^{\mathrm{D}}] \\ k_{+\Delta}[1_{\mathrm{H}}^{\mathrm{D}}] - k_{-\Delta}[1_{\mathrm{D}}^{\mathrm{D}}] \\ k_{+\mathrm{F}}^{\mathrm{d}}[\mathrm{F}][1_{\mathrm{D}}^{\mathrm{D}}] - k_{-\mathrm{F}}^{\mathrm{d}}[\mathrm{HCl}][2_{\mathrm{D}}] \\ k_{+\mathrm{W}}^{\mathrm{d}}[\mathrm{CO_2}][\mathrm{DBF}][1_{\mathrm{D}}^{\mathrm{D}}] - k_{-\mathrm{W}}^{\mathrm{d}}[2_{\mathrm{D}}] \\ k_{-\mathrm{F}}^{\mathrm{p}}[\mathrm{HCl}][2_{\mathrm{D}}] - k_{+\mathrm{F}}^{\mathrm{p}}[\mathrm{F}][1_{\mathrm{H}}^{\mathrm{D}}] \\ k_{-\mathrm{W}}^{\mathrm{p}}[2_{\mathrm{D}}] - k_{+\mathrm{W}}^{\mathrm{p}}[\mathrm{CO_2}][\mathrm{DBF}][1_{\mathrm{H}}^{\mathrm{D}}] \\ k'_{+\Delta}[1_{\mathrm{D}}^{\mathrm{H}}] - k'_{-\Delta}[1_{\mathrm{H}}^{\mathrm{H}}] \\ k_{+\mathrm{F}}^{\mathrm{d}}[\mathrm{F}][1_{\mathrm{H}}^{\mathrm{H}}] - k_{-\mathrm{F}}^{\mathrm{d}}[\mathrm{HCl}][2_{\mathrm{H}}] \\ k_{+\mathrm{W}}^{\mathrm{d}}[\mathrm{CO_2}][\mathrm{DBF}][1_{\mathrm{H}}^{\mathrm{H}}] - k_{-\mathrm{W}}^{\mathrm{d}}[2_{\mathrm{H}}] \end{pmatrix}}_{J = J_+ - J_-} + \underbrace{\begin{pmatrix} I_{\mathrm{F}} \\ I_{\mathrm{HCl}} \\ I_{\mathrm{DBF}} \\ I_{\mathrm{CO_2}} \end{pmatrix}}_{\boldsymbol{I}}, \tag{S.11}$$

with $\boldsymbol{I}$ collecting currents of the chemostatting processes which keep the corresponding concentrations constant. For instance, from equation (S.11) we obtain the following expression for the fuelling current:

$$I_F = J_{3F^{(d)}} + J_{6F^{(d)}} - J_{1F^{(p)}} - J_{4F^{(p)}} \,. \tag{S.12}$$

To conclude this part, we introduce an alternative representation of the dynamics which will prove useful in some of the following derivations. The reader who is not interested in detailed derivations may safely skip the following formula (as well as all the equations marked with an asterisk).

$$\mathrm{d}_t \underbrace{\begin{pmatrix} [2_H] \\ [2_D] \\ [1_H^D] \\ [1_D^D] \\ [1_H^H] \\ [1_D^H] \end{pmatrix}}_{[X]} = \underbrace{\begin{pmatrix} -(\tilde{k}_-^p + \tilde{k}_-^d) & 0 & \tilde{k}_+^p & 0 & \tilde{k}_+^d & 0 \\ 0 & -(\tilde{k}_-^d + \tilde{k}_-^p) & 0 & \tilde{k}_+^d & 0 & \tilde{k}_+^p \\ \tilde{k}_-^p & 0 & -(\tilde{k}_+^p + \tilde{k}_{+\Delta}) & \tilde{k}_{-\Delta} & 0 & 0 \\ 0 & \tilde{k}_-^d & \tilde{k}_{+\Delta} & -(\tilde{k}_+^d + \tilde{k}_{-\Delta}) & 0 & 0 \\ \tilde{k}_-^d & 0 & 0 & 0 & -(\tilde{k}_+^d + \tilde{k}'_{-\Delta}) & \tilde{k}'_{+\Delta} \\ 0 & \tilde{k}_-^p & 0 & 0 & \tilde{k}'_{-\Delta} & -(\tilde{k}_+^p + \tilde{k}'_{+\Delta}) \end{pmatrix}}_{\mathbb{W}} \cdot \underbrace{\begin{pmatrix} [2_H] \\ [2_D] \\ [1_H^D] \\ [1_D^D] \\ [1_H^H] \\ [1_D^H] \end{pmatrix}}_{[X]} . \tag{S.13*}$$

The above equations are equivalent to equations (S.10), and we will make use of the $\mathbb{W}$ matrix in the following. Note that we defined new pseudo first-order rate constants marked with a tilde: the ones referring to mechanical transitions are identical to the ones introduced before (e.g., $\tilde{k}_{+\Delta} = k_{+\Delta}$), while for chemical transitions fuelling and waste-forming rate constants are lumped together (e.g., $\tilde{k}_+^d = k_{+F}^d[F] + k_{+W}^d[DBF][CO_2]$ and $\tilde{k}_-^p = k_{-F}^p[HCl] + k_{-W}^p$). Despite symbols redundancy, the use of the latter rate constants will be convenient in Section 4.5.5.2..

4.5.2.1. Topological properties

The stoichiometric matrix $\mathbb{S} \equiv (\mathbb{S}^X, \mathbb{S}^Y)^T$ in Equations (S.10) and (S.11) encodes the topological properties of the chemical reaction network (CRN). The first class of them can be accessed by determining the cokernel of $\mathbb{S}$, which is spanned by the following vectors

$$\boldsymbol{\ell}_{\mathrm{M}} = \begin{pmatrix} {}^{2_{\mathrm{H}}} & {}^{2_{\mathrm{D}}} & {}^{1^{\mathrm{D}}_{\mathrm{H}}} & {}^{1^{\mathrm{D}}_{\mathrm{D}}} & {}^{1^{\mathrm{H}}_{\mathrm{H}}} & {}^{1^{\mathrm{H}}_{\mathrm{D}}} & \mathrm{F} & \mathrm{HCl} & \mathrm{DBF} & \mathrm{CO_2} \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{S.14}$$

$$\boldsymbol{\ell}_{\mathrm{CO_2}} = \begin{pmatrix} {}^{2_{\mathrm{H}}} & {}^{2_{\mathrm{D}}} & {}^{1^{\mathrm{D}}_{\mathrm{H}}} & {}^{1^{\mathrm{D}}_{\mathrm{D}}} & {}^{1^{\mathrm{H}}_{\mathrm{H}}} & {}^{1^{\mathrm{H}}_{\mathrm{D}}} & \mathrm{F} & \mathrm{HCl} & \mathrm{DBF} & \mathrm{CO_2} \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}, \tag{S.15}$$

$$\boldsymbol{\ell}_{\mathrm{DBF}} = \begin{pmatrix} {}^{2_{\mathrm{H}}} & {}^{2_{\mathrm{D}}} & {}^{1^{\mathrm{D}}_{\mathrm{H}}} & {}^{1^{\mathrm{D}}_{\mathrm{D}}} & {}^{1^{\mathrm{H}}_{\mathrm{H}}} & {}^{1^{\mathrm{H}}_{\mathrm{D}}} & \mathrm{F} & \mathrm{HCl} & \mathrm{DBF} & \mathrm{CO_2} \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}, \tag{S.16}$$

$$\boldsymbol{\ell}_{\mathrm{Cl}} = \begin{pmatrix} {}^{2_{\mathrm{H}}} & {}^{2_{\mathrm{D}}} & {}^{1^{\mathrm{D}}_{\mathrm{H}}} & {}^{1^{\mathrm{D}}_{\mathrm{D}}} & {}^{1^{\mathrm{H}}_{\mathrm{H}}} & {}^{1^{\mathrm{H}}_{\mathrm{D}}} & \mathrm{F} & \mathrm{HCl} & \mathrm{DBF} & \mathrm{CO_2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}. \tag{S.17}$$

The first of these vectors identifies a conserved quantity

$$L_{\mathrm{M}} = \boldsymbol{\ell}_{\mathrm{M}} \cdot \begin{pmatrix} [X] \\ [Y] \end{pmatrix} = [2_{\mathrm{H}}] + [2_{\mathrm{D}}] + [1^{\mathrm{D}}_{\mathrm{H}}] + [1^{\mathrm{D}}_{\mathrm{D}}] + [1^{\mathrm{H}}_{\mathrm{H}}] + [1^{\mathrm{H}}_{\mathrm{D}}]$$
$$\mathrm{d}_t L_{\mathrm{M}} = 0\,, \tag{S.18}$$

which is proved through the rate equations (S.10) and (S.11). It expresses the fact that the total concentration of the motor species is determined by the initial conditions and cannot be altered by the dynamics. The other three vectors identify what we call *broken* conserved quantities

$$L_{\mathrm{CO_2}} = \boldsymbol{\ell}_{\mathrm{CO_2}} \cdot \begin{pmatrix} [X] \\ [Y] \end{pmatrix} = [2_{\mathrm{H}}] + [2_{\mathrm{D}}] + [\mathrm{F}] + [\mathrm{CO_2}]\,, \tag{S.19}$$

$$L_{\mathrm{DBF}} = \boldsymbol{\ell}_{\mathrm{DBF}} \cdot \begin{pmatrix} [X] \\ [Y] \end{pmatrix} = [2_{\mathrm{H}}] + [2_{\mathrm{D}}] + [\mathrm{F}] + [\mathrm{DBF}]\,, \tag{S.20}$$

$$L_{\mathrm{Cl}} = \boldsymbol{\ell}_{\mathrm{Cl}} \cdot \begin{pmatrix} [X] \\ [Y] \end{pmatrix} = [\mathrm{F}] + [\mathrm{HCl}]\,. \tag{S.21}$$

They express the total concentrations of moieties (specific groups of atoms) whose number in the system is not determined by the initial conditions, as they are exchanged with the exterior through the chemostatting. Indeed, by using again the rate equations, it can be shown that

$$\mathrm{d}_t L_{\mathrm{CO_2}} = I_{\mathrm{F}} + I_{\mathrm{CO_2}}\,, \tag{S.22}$$
$$\mathrm{d}_t L_{\mathrm{DBF}} = I_{\mathrm{F}} + I_{\mathrm{DBF}}\,, \tag{S.23}$$
$$\mathrm{d}_t L_{\mathrm{Cl}} = I_{\mathrm{F}} + I_{\mathrm{HCl}}\,. \tag{S.24}$$

Namely, quantities L_{CO_2}, L_{DBF} and L_{Cl} change only due to the exchange of fuel and waste species with the chemostats. If the CRN was closed, they would be conserved by the dynamics like L_M.

The other useful piece of information contained in the stoichiometric matrix $\mathbb{S}$ is given by cycles. A cycle is a vector in the kernel of $\mathbb{S}^X$, which corresponds to a series of reactions which upon completion leaves the concentrations of each motor species unchanged. They play an important role at the steady state, where reaction currents can only flow across cycles.

$$\boldsymbol{c}_{\mathrm{F}}^{\mathrm{T}} = \begin{pmatrix} {}^{1\mathrm{F}^{(\mathrm{p})}} & {}^{1\mathrm{W}^{(\mathrm{p})}} & {}^{2\Delta} & {}^{3\mathrm{F}^{(\mathrm{d})}} & {}^{3\mathrm{W}^{(\mathrm{d})}} & {}^{4\mathrm{F}^{(\mathrm{p})}} & {}^{4\mathrm{W}^{(\mathrm{p})}} & {}^{5\Delta'} & {}^{6\mathrm{F}^{(\mathrm{d})}} & {}^{6\mathrm{W}^{(\mathrm{d})}} \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \end{pmatrix}, \tag{S.25a}$$

$$\boldsymbol{c}_{\mathrm{W}}^{\mathrm{T}} = \begin{pmatrix} {}^{1\mathrm{F}^{(\mathrm{p})}} & {}^{1\mathrm{W}^{(\mathrm{p})}} & {}^{2\Delta} & {}^{3\mathrm{F}^{(\mathrm{d})}} & {}^{3\mathrm{W}^{(\mathrm{d})}} & {}^{4\mathrm{F}^{(\mathrm{p})}} & {}^{4\mathrm{W}^{(\mathrm{p})}} & {}^{5\Delta'} & {}^{6\mathrm{F}^{(\mathrm{d})}} & {}^{6\mathrm{W}^{(\mathrm{d})}} \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}, \tag{S.25b}$$

$$\boldsymbol{c}_{\mathrm{a}}^{\mathrm{T}} = \begin{pmatrix} {}^{1\mathrm{F}^{(\mathrm{p})}} & {}^{1\mathrm{W}^{(\mathrm{p})}} & {}^{2\Delta} & {}^{3\mathrm{F}^{(\mathrm{d})}} & {}^{3\mathrm{W}^{(\mathrm{d})}} & {}^{4\mathrm{F}^{(\mathrm{p})}} & {}^{4\mathrm{W}^{(\mathrm{p})}} & {}^{5\Delta'} & {}^{6\mathrm{F}^{(\mathrm{d})}} & {}^{6\mathrm{W}^{(\mathrm{d})}} \\ 1 & -1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \end{pmatrix}, \tag{S.25c}$$

$$\boldsymbol{c}_{\mathrm{b}}^{\mathrm{T}} = \begin{pmatrix} {}^{1\mathrm{F}^{(\mathrm{p})}} & {}^{1\mathrm{W}^{(\mathrm{p})}} & {}^{2\Delta} & {}^{3\mathrm{F}^{(\mathrm{d})}} & {}^{3\mathrm{W}^{(\mathrm{d})}} & {}^{4\mathrm{F}^{(\mathrm{p})}} & {}^{4\mathrm{W}^{(\mathrm{p})}} & {}^{5\Delta'} & {}^{6\mathrm{F}^{(\mathrm{d})}} & {}^{6\mathrm{W}^{(\mathrm{d})}} \\ -1 & 1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{S.25d}$$

$$\boldsymbol{c}_{\epsilon}^{\mathrm{T}} = \begin{pmatrix} {}^{1\mathrm{F}^{(\mathrm{p})}} & {}^{1\mathrm{W}^{(\mathrm{p})}} & {}^{2\Delta} & {}^{3\mathrm{F}^{(\mathrm{d})}} & {}^{3\mathrm{W}^{(\mathrm{d})}} & {}^{4\mathrm{F}^{(\mathrm{p})}} & {}^{4\mathrm{W}^{(\mathrm{p})}} & {}^{5\Delta'} & {}^{6\mathrm{F}^{(\mathrm{d})}} & {}^{6\mathrm{W}^{(\mathrm{d})}} \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}. \tag{S.25e}$$

In the above vectors, each entry represents the number of times the corresponding reaction has to be performed in order to complete the corresponding cycle. Minus signs indicate that the corresponding reaction is performed backwards in the cycle, considering the clockwise cycling in **Figure 2** as the forward direction (the same convention is adopted for the reactions at the beginning of Section 4.5.1., where left-to-right is considered as forward direction). Cycles $\boldsymbol{c}_F$, $\boldsymbol{c}_W$, $\boldsymbol{c}_a$ and $\boldsymbol{c}_b$ would be present also if the system was closed (fuel and waste species not chemostatted and free to vary) and we call them *internal* cycles (mathematically, they are also right-null vectors of the whole $\mathbb{S}$ matrix[42]). They set thermodynamic constraints on the rate constants that will be introduced later (Wegscheider's conditions, equations (S.36)). The cycle $\boldsymbol{c}_\epsilon$ involves a net exchange of chemicals with chemostats, and we call it *emergent* cycle because it arises only when the system is open. Note that the set of vectors (S.25) is a complete basis for the cycles in the system, but it is not unique. Another useful basis that we will exploit in Section 4.5.3. is the following, obtained from linear combinations of vectors (S.25):

$$\boldsymbol{c}_{\epsilon}^{\mathrm{T}} = \begin{pmatrix} {}^{1\mathrm{F}^{(\mathrm{p})}} & {}^{1\mathrm{W}^{(\mathrm{p})}} & {}^{2\Delta} & {}^{3\mathrm{F}^{(\mathrm{d})}} & {}^{3\mathrm{W}^{(\mathrm{d})}} & {}^{4\mathrm{F}^{(\mathrm{p})}} & {}^{4\mathrm{W}^{(\mathrm{p})}} & {}^{5\Delta'} & {}^{6\mathrm{F}^{(\mathrm{d})}} & {}^{6\mathrm{W}^{(\mathrm{d})}} \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}, \tag{S.26a}$$

$$\boldsymbol{c}_{1\mathrm{p}}^{\mathrm{T}} = \frac{1}{2}(\boldsymbol{c}_{\epsilon}^{\mathrm{T}} - \boldsymbol{c}_{\mathrm{F}}^{\mathrm{T}} - \boldsymbol{c}_{\mathrm{a}}^{\mathrm{T}}) = \begin{pmatrix} {}^{1\mathrm{F}^{(\mathrm{p})}} & {}^{1\mathrm{W}^{(\mathrm{p})}} & {}^{2\Delta} & {}^{3\mathrm{F}^{(\mathrm{d})}} & {}^{3\mathrm{W}^{(\mathrm{d})}} & {}^{4\mathrm{F}^{(\mathrm{p})}} & {}^{4\mathrm{W}^{(\mathrm{p})}} & {}^{5\Delta'} & {}^{6\mathrm{F}^{(\mathrm{d})}} & {}^{6\mathrm{W}^{(\mathrm{d})}} \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{S.26b}$$

$$\boldsymbol{c}_{3\mathrm{d}}^{\mathrm{T}} = \boldsymbol{c}_{1\mathrm{p}}^{\mathrm{T}} - \boldsymbol{c}_{\mathrm{b}}^{\mathrm{T}} = \begin{pmatrix} {}^{1\mathrm{F}^{(\mathrm{p})}} & {}^{1\mathrm{W}^{(\mathrm{p})}} & {}^{2\Delta} & {}^{3\mathrm{F}^{(\mathrm{d})}} & {}^{3\mathrm{W}^{(\mathrm{d})}} & {}^{4\mathrm{F}^{(\mathrm{p})}} & {}^{4\mathrm{W}^{(\mathrm{p})}} & {}^{5\Delta'} & {}^{6\mathrm{F}^{(\mathrm{d})}} & {}^{6\mathrm{W}^{(\mathrm{d})}} \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{S.26c}$$

$$\boldsymbol{c}_{4\mathrm{p}}^{\mathrm{T}} = \boldsymbol{c}_{\mathrm{a}}^{\mathrm{T}} + \boldsymbol{c}_{1\mathrm{p}}^{\mathrm{T}} = \begin{pmatrix} {}^{1\mathrm{F}^{(\mathrm{p})}} & {}^{1\mathrm{W}^{(\mathrm{p})}} & {}^{2\Delta} & {}^{3\mathrm{F}^{(\mathrm{d})}} & {}^{3\mathrm{W}^{(\mathrm{d})}} & {}^{4\mathrm{F}^{(\mathrm{p})}} & {}^{4\mathrm{W}^{(\mathrm{p})}} & {}^{5\Delta'} & {}^{6\mathrm{F}^{(\mathrm{d})}} & {}^{6\mathrm{W}^{(\mathrm{d})}} \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \end{pmatrix}, \tag{S.26d}$$

$$\boldsymbol{c}_{6\mathrm{d}}^{\mathrm{T}} = \boldsymbol{c}_{\epsilon}^{\mathrm{T}} - \boldsymbol{c}_{\mathrm{W}}^{\mathrm{T}} - \boldsymbol{c}_{3\mathrm{d}}^{\mathrm{T}} = \begin{pmatrix} {}^{1\mathrm{F}^{(\mathrm{p})}} & {}^{1\mathrm{W}^{(\mathrm{p})}} & {}^{2\Delta} & {}^{3\mathrm{F}^{(\mathrm{d})}} & {}^{3\mathrm{W}^{(\mathrm{d})}} & {}^{4\mathrm{F}^{(\mathrm{p})}} & {}^{4\mathrm{W}^{(\mathrm{p})}} & {}^{5\Delta'} & {}^{6\mathrm{F}^{(\mathrm{d})}} & {}^{6\mathrm{W}^{(\mathrm{d})}} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}. \tag{S.26e}$$

For instance, the vector $\boldsymbol{c}_{1p}$ describes the sequence of reactions ${}^{2}\mathrm{H} \xrightarrow{{}^{1\mathrm{W}^{(\mathrm{p})}}} {}^{1}_{\mathrm{H}}\mathrm{D} \xrightarrow{{}^{1\mathrm{F}^{(\mathrm{p})}}} {}^{2}\mathrm{H}$. Cycles $\boldsymbol{c}_{1p}$, $\boldsymbol{c}_{3d}$, $\boldsymbol{c}_{4p}$, and $\boldsymbol{c}_{6d}$ are called *futile* cycles, since they consume fuel molecules without contributing to a forward step of the motor.

4.5.2.2. Stationary state

Given a set of values for the kinetic parameters and the total concentration of motor species L_M (defined in equation (S.18)), for any choice of the chemostatted concentrations in $\boldsymbol{Y}$ the system will reach a stationary state distribution (mathematically, this is because the network is complex balanced and confined[69]) such that $d_t[\boldsymbol{X}]_{SS} = 0$ in equation (S.10). It can be analytically computed using graph-theoretical techniques (see Appendix A of reference 67 for an introduction):

$$[X_\alpha]_{\mathrm{ss}} = \frac{L_{\mathrm{M}}}{\mathcal{N}} \sum_{t \in \mathcal{T}_\alpha} \prod_{\lambda \in t} \tilde{k}_\lambda \,. \tag{S.27*}$$

In the above expression, $\mathcal{N} = \sum_\alpha \sum_{t \in \mathcal{T}_\alpha} \prod_{\lambda \in t} \tilde{k}_\lambda$ is a normalizing denominator and $\mathcal{T}_\alpha$ is the set of spanning trees rooted in vertex α of the chemical reaction network representing the motor. The index λ runs over all forward and backward reactions such that $\tilde{k}_\lambda$ are the pseudo-first-order rate constants as defined in equation (S.13*). A rooted spanning tree is a spanning tree with its edges oriented such that all edges point towards the root.

For the experimental case where equations (S.1) hold, the above equations boil down to:

$$\begin{pmatrix}[2_H]_{ss}\\ [2_D]_{ss}\\ [1^D_H]_{ss}\\ [1^D_D]_{ss}\\ [1^H_H]_{ss}\\ [1^H_D]_{ss}\end{pmatrix} = \frac{L_M}{\mathcal{N}}\begin{pmatrix}(k^{p}_{+F}[F]+k_{+W}[CO_2][DBF])(k^{d}_{+F}[F]+k_{+W}[CO_2][DBF])+(k^{p}_{+F}[F]+k^{d}_{+F}[F]+2k_{+W}[CO_2][DBF])k_{\Delta}\\ (k^{p}_{+F}[F]+k_{+W}[CO_2][DBF])(k^{d}_{+F}[F]+k_{+W}[CO_2][DBF])+(k^{p}_{+F}[F]+k^{d}_{+F}[F]+2k_{+W}[CO_2][DBF])k_{\Delta}\\ (k^{p}_{-F}[HCl]+k_{-W})(k^{d}_{+F}[F]+k_{+W}[CO_2][DBF])+(k^{p}_{-F}[HCl]+k^{d}_{-F}[HCl]+2k_{-W})k_{\Delta}\\ (k^{d}_{-F}[HCl]+k_{-W})(k^{p}_{+F}[F]+k_{+W}[CO_2][DBF])+(k^{p}_{-F}[HCl]+k^{d}_{-F}[HCl]+2k_{-W})k_{\Delta}\\ (k^{d}_{-F}[HCl]+k_{-W})(k^{p}_{+F}[F]+k_{+W}[CO_2][DBF])+(k^{p}_{-F}[HCl]+k^{d}_{-F}[HCl]+2k_{-W})k_{\Delta}\\ (k^{p}_{-F}[HCl]+k_{-W})(k^{d}_{+F}[F]+k_{+W}[CO_2][DBF])+(k^{p}_{-F}[HCl]+k^{d}_{-F}[HCl]+2k_{-W})k_{\Delta}\end{pmatrix}. \tag{S.28}$$

It is easy to see that if the steric hindrance caused by the macrocycle did not affect the rates of the fueling reaction (i.e., $k^{p}_{+F} = k^{d}_{+F}$ and $k^{p}_{-F} = k^{d}_{-F}$), the stationary state would be symmetric ($[1^H_H]_{SS} = [1^H_D]_{SS} = [1^D_H]_{SS} = [1^D_D]_{SS}$).

From equation (S.10), it follows that the stationary state current vector $\boldsymbol{J}_{SS}$ must be a right null vector of the matrix $\mathbb{S}^X$. As anticipated in the previous section (Section 4.5.2.1.), this implies that the stationary state currents can be decomposed on a basis of cycles:

$$\boldsymbol{J}_{ss} = \sum_i \psi_i \boldsymbol{c}_i\,, \tag{S.29}$$

where the index i runs over the elements of a complete basis of cycles and terms ψ_i are the currents flowing across each cycle. Considering vectors (S.25) and (S.26) as basis sets:

$$\begin{pmatrix}J^{ss}_{1F(p)}\\ J^{ss}_{1W(p)}\\ J^{ss}_{2\Delta}\\ J^{ss}_{3F(d)}\\ J^{ss}_{3W(d)}\\ J^{ss}_{4F(p)}\\ J^{ss}_{4W(p)}\\ J^{ss}_{5\Delta'}\\ J^{ss}_{6F(d)}\\ J^{ss}_{6W(d)}\end{pmatrix} = \begin{pmatrix}\psi_{c_F}+\psi_{c_a}-\psi_{c_b}\\ \psi_{c_W}-\psi_{c_a}+\psi_{c_b}+\psi_{c_\epsilon}\\ \psi_{c_F}+\psi_{c_W}+\psi_{c_\epsilon}\\ \psi_{c_F}-\psi_{c_b}+\psi_{c_\epsilon}\\ \psi_{c_W}+\psi_{c_b}\\ \psi_{c_F}-\psi_{c_a}\\ \psi_{c_W}+\psi_{c_a}+\psi_{c_\epsilon}\\ \psi_{c_F}+\psi_{c_W}+\psi_{c_\epsilon}\\ \psi_{c_F}+\psi_{c_\epsilon}\\ \psi_{c_W}\end{pmatrix} = \begin{pmatrix}-\psi_{c_{1p}}\\ \psi_{c_{1p}}+\psi'_{c_\epsilon}\\ \psi'_{c_\epsilon}\\ \psi_{c_{3d}}+\psi'_{c_\epsilon}\\ -\psi_{c_{3d}}\\ -\psi_{c_{4p}}\\ \psi_{c_{4p}}+\psi'_{c_\epsilon}\\ \psi'_{c_\epsilon}\\ \psi_{c_{6d}}+\psi'_{c_\epsilon}\\ -\psi_{c_{6d}}\end{pmatrix}. \tag{S.30}$$

Note that, depending on the chosen basis set, the same cycle may be associated with a different cycle current (*i.e.*, $\psi_{c_\epsilon} \neq \psi'_{c_\epsilon}$) in equation (S.30). We will come back to this point

at the end of the next section (Section 4.5.3.), when the set of fundamental cycles and corresponding currents will be identified. Here, we conclude by noticing that equation (S.30) implies the following relations between steady state currents:

$$J^{ss} := J^{ss}_{1F(p)} + J^{ss}_{1W(p)} = J^{ss}_{2\Delta} = J^{ss}_{3F(d)} + J^{ss}_{3W(d)} = J^{ss}_{4F(p)} + J^{ss}_{4W(p)} = J^{ss}_{5\Delta'} = J^{ss}_{6F(d)} + J^{ss}_{6W(d)} = \psi_{c_\epsilon} + \psi_{c_W} + \psi_{c_F} = \psi'_{c_\epsilon}\,, \quad \text{(S.31a)}$$

$$I^{ss}_F = J^{ss}_{3F(d)} + J^{ss}_{6F(d)} - J^{ss}_{1F(p)} - J^{ss}_{4F(p)} = 2\psi_{c_\epsilon} = 2\psi'_{c_\epsilon} + \psi_{c_{1p}} + \psi_{c_{3d}} + \psi_{c_{4p}} + \psi_{c_{6d}}\,, \quad \text{(S.31b)}$$

where J^{SS} corresponds to the rate of net displacement of the macrocycle along the track in the stationary state. Note that J^{SS} and I_F^{SS} are respectively expressed as J and I_F in Section 4.2., where we unambiguously refer to the stationary state.

4.5.3. Thermodynamics

Each species is thermodynamically characterized by chemical potentials of the form

$$\mu_X = \mu_X^\circ + RT\ln[X]\,, \quad \mu_Y = \mu_Y^\circ + RT\ln[Y]\,, \quad \text{(S.32)}$$

where μ°_X and μ°_Y are standard-state chemical potentials. We will denote the sum of the chemical potentials of waste species as $\mu_W = \mu_{HCl} + \mu_{DBF} + \mu_{CO_2}$. Note that in the experimental Fmoc rotary motor (see equations (S.1)) we have $\mu^\circ_{1_H^D} = \mu^\circ_{1_D^D} = \mu^\circ_{1_D^H} = \mu^\circ_{1_H^H} \equiv \mu^\circ_1$, since the benzylic amide macrocycle only interacts with the two identical fumaramide sites. The equality $\mu^\circ_{2_H} = \mu^\circ_{2_D} \equiv \mu^\circ_2$ also holds in the experimental case.

Dynamics and thermodynamics can be connected via the so-called "principle of local detailed balance"[42] which relates the ratio of forward and backward rate constants of a single elementary reaction to the difference of standard-state chemical potentials between its reagents and products, and thus to the free-energy exchanged with the reservoirs (thermal bath and chemostats) each time the reaction happens:

$$RT\ln\frac{k_{+\rho}}{k_{-\rho}} = -\sum_X \mu_X^\circ \mathbb{S}_\rho^X - \sum_Y \mu_Y^\circ \mathbb{S}_\rho^Y\,, \quad \text{(S.33)}$$

where indexes X and Y run over motor's and chemostatted species, respectively, while index ρ runs over all the 10 reactions introduced in Section 4.5.1.. For closed systems, this relation is essentially the law of mass action which has been known since the 19th century in chemistry. Its open systems formulation was later extensively used and was derived to ensure that microscopic reversibility holds at equilibrium[57,70,71]. In the broader context of statistical mechanics, similar relations were explicitly formulated in reference 72 and to the best of our knowledge the terminology "local detailed balance" first appeared in reference 63 to denote a constructive principle for formulating physically sound models describing nonequilibrium phenomena. It is nowadays the central concept underlying stochastic thermodynamics, the theory that formulates the general principles for building nonequilibrium thermodynamics for a variety of dynamics,[64,65] not always related to chemical processes. Under certain conditions, local detailed balance also holds for nonelementary reactions as shown in references 67 and 68. This is indeed the case for the coarse-grained rate constants of the fueling and waste-forming reaction (see Section 4.5.1.), as commonly assumed[15,16].

We note that the local detailed balance as stated in (S.33), is different from how local detailed balance is discussed in references 16 and 73. Indeed, only in contexts where there is only one possible transition (*e.g.* one chemical reaction reaction) connecting a pair of states (*e.g.* two motor species), local detailed balance reduces to a constraint on the forward and backward transition rates connecting the two. But in the cases where there are multiple transitions between a couple of states (*e.g.* two reactions involving the same motor species but different chemostats), local detailed balance only holds at the level of each single transition[63,64,74], as done in equation (S.33), and not at the level of the net (*i.e.* summed) transition rates. This important point is also discussed in comment 75.

At equilibrium, the thermodynamic forces driving each reaction (also called *affinities* in the literature, a term not adopted here to avoid confusion with binding affinities) vanish

$$A_\rho^{\mathrm{eq}} = -\sum_X \mu_X^{\mathrm{eq}} \mathbb{S}_\rho^X - \sum_Y \mu_Y^{\mathrm{eq}} \mathbb{S}_\rho^Y = 0\,, \tag{S.34}$$

as well as all reaction currents

$$J_\rho^{\mathrm{eq}} = J_{+\rho}^{\mathrm{eq}} - J_{-\rho}^{\mathrm{eq}} = 0\,. \tag{S.35}$$

Equations (S.33) imply the Wegscheider's conditions for the internal cycles: the product of the forward rate constants along each internal cycle must be equal to that of the backward rate constants

$$\text{from } \boldsymbol{c}_{\mathrm{F}}: \quad k_{-\mathrm{F}}^{\mathrm{p}} k_{+\Delta} k_{+\mathrm{F}}^{\mathrm{d}} k_{-\mathrm{F}}^{\mathrm{p}} k'_{+\Delta} k_{+\mathrm{F}}^{\mathrm{d}} = k_{+\mathrm{F}}^{\mathrm{p}} k_{-\Delta} k_{-\mathrm{F}}^{\mathrm{d}} k_{+\mathrm{F}}^{\mathrm{p}} k'_{-\Delta} k_{-\mathrm{F}}^{\mathrm{d}}\,, \tag{S.36a}$$

$$\text{from } \boldsymbol{c}_{\mathrm{W}}: \quad k_{-\mathrm{W}}^{\mathrm{p}} k_{+\Delta} k_{+\mathrm{W}}^{\mathrm{d}} k_{-\mathrm{W}}^{\mathrm{p}} k'_{+\Delta} k_{+\mathrm{W}}^{\mathrm{d}} = k_{+\mathrm{W}}^{\mathrm{p}} k_{-\Delta} k_{-\mathrm{W}}^{\mathrm{d}} k_{+\mathrm{W}}^{\mathrm{p}} k'_{-\Delta} k_{-\mathrm{W}}^{\mathrm{d}}\,, \tag{S.36b}$$

$$\text{from } \boldsymbol{c}_{\mathrm{a}}: \quad k_{-\mathrm{F}}^{\mathrm{p}} k_{+\mathrm{W}}^{\mathrm{p}} k_{+\mathrm{F}}^{\mathrm{p}} k_{-\mathrm{W}}^{\mathrm{p}} = k_{+\mathrm{F}}^{\mathrm{p}} k_{-\mathrm{W}}^{\mathrm{p}} k_{-\mathrm{F}}^{\mathrm{p}} k_{+\mathrm{W}}^{\mathrm{p}}\,, \tag{S.36c}$$

$$\text{from } \boldsymbol{c}_{\mathrm{b}}: \quad k_{+\mathrm{F}}^{\mathrm{p}} k_{-\mathrm{W}}^{\mathrm{p}} k_{-\mathrm{F}}^{\mathrm{d}} k_{+\mathrm{W}}^{\mathrm{d}} = k_{-\mathrm{F}}^{\mathrm{p}} k_{+\mathrm{W}}^{\mathrm{p}} k_{+\mathrm{F}}^{\mathrm{d}} k_{-\mathrm{W}}^{\mathrm{d}}\,. \tag{S.36d}$$

They set thermodynamic constraints that the kinetic parameters must satisfy, thus reducing freedom in their choice. Note that, due to symmetries in the system (*i.e.*, rate constants of reactions $1\mathrm{F}^{(\mathrm{p})}$ and $4\mathrm{F}^{(\mathrm{p})}$ are considered identical), condition (S.36c) is trivially satisfied and condition (S.36d) is implied by (S.36a) and (S.36b), which are therefore the only ones reported in Section 4.4..

The dissipation of the process is captured by the entropy production rate times the temperature T:

$$T\dot{\Sigma} = RT \sum_\rho J_\rho \ln \frac{J_{+\rho}}{J_{-\rho}} \geq 0\,, \tag{S.37}$$

which also vanishes at equilibrium. Equation (S.37) shows that each of the 10 reactions in the vector $\boldsymbol{\rho}$ (see Section 4.5.1.) brings a positive contribution to the total dissipation. We will use this property in Section 4.5.4. to identify the specific contributions of chemical and mechanical processes.

By using the rate equations and the local detailed balance (S.33), the dissipation in equation (S.37) can be rewritten as

$$T\dot{\Sigma} = -\mathrm{d}_t \mathcal{G} + \dot{\mathcal{W}}_{\mathrm{fuel}}\,, \tag{S.38}$$

where

$$\mathcal{G} = [1^{\mathrm{D}}_{\mathrm{H}}]\mu_{1^{\mathrm{D}}_{\mathrm{H}}} + [1^{\mathrm{D}}_{\mathrm{D}}]\mu_{1^{\mathrm{D}}_{\mathrm{D}}} + [1^{\mathrm{H}}_{\mathrm{H}}]\mu_{1^{\mathrm{H}}_{\mathrm{H}}} + [1^{\mathrm{H}}_{\mathrm{D}}]\mu_{1^{\mathrm{H}}_{\mathrm{D}}} + [2_{\mathrm{H}}]\left(\mu_{2_{\mathrm{H}}} - \mu_{\mathrm{DBF}} - \mu_{\mathrm{CO}_2}\right) + [2_{\mathrm{D}}]\left(\mu_{2_{\mathrm{D}}} - \mu_{\mathrm{DBF}} - \mu_{\mathrm{CO}_2}\right) + \\ + [\mathrm{F}]\,(\mu_{\mathrm{F}} - \mu_{\mathrm{W}}) - RT\,(L_{\mathrm{M}} + [\mathrm{F}] + [\mathrm{HCl}] + [\mathrm{DBF}] + [\mathrm{CO}_2]) \tag{S.39}$$

is the *semigrand* Gibbs potential of the system[42], and

$$\dot{\mathcal{W}}_{\mathrm{fuel}} = I_{\mathrm{F}}(\mu_{\mathrm{F}} - \mu_{\mathrm{W}}) \tag{S.40}$$

is the fueling chemical work per unit of time (*i.e.*, the fueling power). If $\mu_F = \mu_W$, equation (S.38) shows that $\mathcal{G}$ is a monotonically decreasing function in time, given that $T\dot{\Sigma} \geq 0$.

Steady state entropy production and J^{SS}/I_F^{SS}

As pointed out in Section 4.2.2., at stationary state (see Section 4.5.2.2.) the dissipation in equation (S.38) boils down to:

$$T\dot{\Sigma}^{\mathrm{ss}} = \dot{\mathcal{W}}^{\mathrm{ss}}_{\mathrm{fuel}} = I^{\mathrm{ss}}_{\mathrm{F}}(\mu_{\mathrm{F}} - \mu_{\mathrm{W}})\,. \tag{S.41}$$

It is instructive to look at the equivalent expressions of equation (S.41) which are obtained by decomposing the stationary I_F^{SS} current on both cycles basis (S.25) and (S.26):

$$T\dot{\Sigma}^{\mathrm{ss}} = 2\psi_{c_\epsilon}(\mu_{\mathrm{F}} - \mu_{\mathrm{W}}) = (\psi_{c_{1\mathrm{p}}} + \psi_{c_{3\mathrm{d}}} + \psi_{c_{4\mathrm{p}}} + \psi_{c_{6\mathrm{d}}} + 2\psi'_{c_\epsilon})(\mu_{\mathrm{F}} - \mu_{\mathrm{W}})\,. \tag{S.42}$$

The first equality (decomposition according to the basis of vectors (S.25)) identifies the cycle current ψ_{c_ϵ} as the fundamental current in the stationary regime: it accounts for total dissipation once multiplied by its conjugated fundamental force $2(\mu_F - \mu_W)$, also denoted *cycle's affinity*; this is a consequence of the presence of a unique emergent cycle. The second one (decomposition according to the basis of vectors (S.26)) highlights how the consumption of fuel can be thought of as happening due to 5 different processes: the directional cycling of the macrocycle (ψ'_{c_ϵ}, which one would like to enhance), and four futile cycles consuming fuel without inducing directional movement of the macrocycle ($\psi_{c_{1p}}$, $\psi_{c_{3d}}$, $\psi_{c_{4p}}$, and $\psi_{c_{6d}}$ which one would like to suppress). From the latter perspective, we use relations (S.31) to

introduce the following *coefficient of performance* (J^{SS}/I_F^{SS}) quantifying how many fuel molecules are needed on average to complete a directional cycle:

$$\frac{J^{\mathrm{ss}}}{I_{\mathrm{F}}^{\mathrm{ss}}} = \frac{\psi'_{c_\epsilon}}{\psi_{c_{1\mathrm{p}}} + \psi_{c_{3\mathrm{d}}} + \psi_{c_{4\mathrm{p}}} + \psi_{c_{6\mathrm{d}}} + 2\psi'_{c_\epsilon}} \,. \tag{S.43}$$

In the best possible scenario, namely when futile currents vanish and the fuel consumption is tightly coupled with macrocycle directional movement, we have an average of 2 fuel molecules consumed per cycle, leading to an optimal J^{SS}/I_F^{SS} of 0.5. Equation (S.43) allows to quantify how much the motor deviates from the optimal condition and will be employed in the simulations of Section 4.5.6. to further compare different designs.

4.5.4. Information thermodynamics perspective

If we stick to the description above, the motor appears as an autonomous system exploiting the chemical gradient $\mu_F - \mu_W$ to sustain currents. We can evaluate its performance through equation (S.43), but questions such as how effectively the free energy harvested from the fuel is transduced by the motor to produce directional motion remain open. The detailed mechanism is better understood when the motor is thought of as a bipartite system, where we distinguish between "mechanical" states $\boldsymbol{mech} = \{H, D\}$ referring to the position of the macrocycle, and "chemical" ones $\boldsymbol{chem} = \{2, 1^H, 1^D\}$ indicating the state of the track (how many stoppers attached, if one, where).

With this structure in mind, the entropy production in equation (S.37) can be split as:

$$
\begin{aligned}
T\dot{\Sigma} &= RT\sum_{\rho} J_{\rho}\ln\frac{J_{+\rho}}{J_{-\rho}} = \underbrace{T\,R\sum_{\rho} J_{\rho}\ln\left(\frac{k_{+\rho}}{k_{-\rho}}Y_{\sigma}^{-\mathbb{S}^{Y\sigma}_{\rho}}\right)}_{\dot{S}_{\mathrm{r}}} + \underbrace{T\,R\sum_{\rho} J_{\rho}\ln X_{\sigma}^{-\mathbb{S}^{X\sigma}_{\rho}}}_{\mathrm{d}_t S} = \\
&= \underbrace{RTJ_{2\Delta}\ln\frac{k_{+\Delta}}{k_{-\Delta}} + RTJ_{5\Delta'}\ln\frac{k'_{+\Delta}}{k'_{-\Delta}}}_{T\dot{S}_{\mathrm{r}}^{\mathrm{mech}}} + \underbrace{RTJ_{2\Delta}\ln\frac{[1^{\mathrm{D}}_{\mathrm{H}}]}{[1^{\mathrm{D}}_{\mathrm{D}}]} + RTJ_{5\Delta'}\ln\frac{[1^{\mathrm{H}}_{\mathrm{D}}]}{[1^{\mathrm{H}}_{\mathrm{H}}]}}_{T\mathrm{đ}_t S^{\mathrm{mech}}} + \\
&+ \underbrace{RT(J_{1\mathrm{F}^{(\mathrm{p})}} + J_{4\mathrm{F}^{(\mathrm{p})}})\ln\frac{k^{\mathrm{p}}_{-\mathrm{F}}[\mathrm{HCl}]}{k^{\mathrm{p}}_{+\mathrm{F}}[\mathrm{F}]} + RT(J_{3\mathrm{F}^{(\mathrm{d})}} + J_{6\mathrm{F}^{(\mathrm{d})}})\ln\frac{k^{\mathrm{d}}_{+\mathrm{F}}[\mathrm{F}]}{k^{\mathrm{d}}_{-\mathrm{F}}[\mathrm{HCl}]}}_{T\dot{S}_{\mathrm{r}}^{\mathrm{chem,F}}} + \\
&+ \underbrace{RT(J_{1\mathrm{W}^{(\mathrm{p})}} + J_{4\mathrm{W}^{(\mathrm{p})}})\ln\frac{k^{\mathrm{p}}_{-\mathrm{W}}}{k^{\mathrm{p}}_{+\mathrm{W}}[\mathrm{DBF}][\mathrm{CO_2}]} + RT(J_{3\mathrm{W}^{(\mathrm{d})}} + J_{6\mathrm{W}^{(\mathrm{d})}})\ln\frac{k^{\mathrm{d}}_{+\mathrm{W}}[\mathrm{DBF}][\mathrm{CO_2}]}{k^{\mathrm{d}}_{-\mathrm{W}}}}_{T\dot{S}_{\mathrm{r}}^{\mathrm{chem,W}}} + \\
&+ \underbrace{RT(J_{1\mathrm{F}^{(\mathrm{p})}} + J_{1\mathrm{W}^{(\mathrm{p})}})\ln\frac{[2_{\mathrm{H}}]}{[1^{\mathrm{D}}_{\mathrm{H}}]} + RT(J_{3\mathrm{F}^{(\mathrm{d})}} + J_{3\mathrm{W}^{(\mathrm{d})}})\ln\frac{[1^{\mathrm{D}}_{\mathrm{D}}]}{[2_{\mathrm{D}}]} + RT(J_{4\mathrm{F}^{(\mathrm{p})}} + J_{4\mathrm{W}^{(\mathrm{p})}})\ln\frac{[2_{\mathrm{D}}]}{[1^{\mathrm{H}}_{\mathrm{D}}]} + RT(J_{6\mathrm{F}^{(\mathrm{d})}} + J_{6\mathrm{W}^{(\mathrm{d})}})\ln\frac{[1^{\mathrm{H}}_{\mathrm{H}}]}{[2_{\mathrm{H}}]}}_{T\mathrm{đ}_t S^{\mathrm{chem}}} \geq 0\,,
\end{aligned}
\tag{S.44}
$$

where terms

$$
T\dot{S}_{\mathrm{r}}^{\mathrm{mech}} = J_{2\Delta}(\mu^{\circ}_{1^{\mathrm{D}}_{\mathrm{H}}} - \mu^{\circ}_{1^{\mathrm{D}}_{\mathrm{D}}}) + J_{5\Delta'}(\mu^{\circ}_{1^{\mathrm{H}}_{\mathrm{D}}} - \mu^{\circ}_{1^{\mathrm{H}}_{\mathrm{H}}}) \tag{S.45a}
$$

$$
\begin{aligned}
T\dot{S}_{\mathrm{r}}^{\mathrm{chem,F}} + T\dot{S}_{\mathrm{r}}^{\mathrm{chem,W}} = {} & \overbrace{J_{\mathrm{F}}(\mu_{\mathrm{F}} - \mu_{\mathrm{W}})}^{\dot{W}_{\mathrm{fuel}}} + (J_{1\mathrm{F}^{(\mathrm{p})}} + J_{1\mathrm{W}^{(\mathrm{p})}})(\mu^{\circ}_{2_{\mathrm{H}}} - \mu^{\circ}_{1^{\mathrm{D}}_{\mathrm{H}}} - \mu_{\mathrm{DBF}} - \mu_{\mathrm{CO_2}}) + \\
& + (J_{4\mathrm{F}^{(\mathrm{p})}} + J_{4\mathrm{W}^{(\mathrm{p})}})(\mu^{\circ}_{2_{\mathrm{D}}} - \mu^{\circ}_{1^{\mathrm{H}}_{\mathrm{D}}} - \mu_{\mathrm{DBF}} - \mu_{\mathrm{CO_2}}) + (J_{3\mathrm{F}^{(\mathrm{d})}} + J_{3\mathrm{W}^{(\mathrm{d})}})(\mu^{\circ}_{1^{\mathrm{D}}_{\mathrm{D}}} - \mu^{\circ}_{2_{\mathrm{D}}} + \mu_{\mathrm{DBF}} + \mu_{\mathrm{CO_2}}) + \\
& + (J_{6\mathrm{F}^{(\mathrm{d})}} + J_{6\mathrm{W}^{(\mathrm{d})}})(\mu^{\circ}_{1^{\mathrm{H}}_{\mathrm{H}}} - \mu^{\circ}_{2_{\mathrm{H}}} + \mu_{\mathrm{DBF}} + \mu_{\mathrm{CO_2}})
\end{aligned}
\tag{S.45b}
$$

keep track of the free energy exchanged with the reservoirs by each of the two sets of degrees of freedom. Note that fueling work by the chemostats is done on chemical degrees of freedom only. Terms $\mathrm{đ}_t S^{\mathrm{mech}}$ and $\mathrm{đ}_t S^{\mathrm{chem}}$ can be rewritten by introducing cumulative concentrations for the chemical subsystem: $[2] = [2_H] + [2_D], [1^H] = [1^H_D] + [1^H_H], [1^D] = [1^D_D] + [1^D_H]$; and for the mechanical one: $[H^{mech}] = [2_H] + [1^H_H] + [1^D_H], [D^{mech}] = [2_D] + [1^H_D] + [1^D_D]$. We thus have:

$$
\begin{aligned}
đ_t S^{\mathrm{mech}}/R &= \underbrace{(J_{2\Lambda} - J_{5\Lambda'}) \ln \frac{[\mathrm{H}^{\mathrm{mech}}]}{[\mathrm{D}^{\mathrm{mech}}]}}_{\mathrm{d}_t S^{\mathrm{mech}}/R} \underbrace{- J_{2\Lambda} \ln \frac{[\mathrm{H}^{\mathrm{mech}}][1_{\mathrm{D}}^{\mathrm{D}}]}{[\mathrm{D}^{\mathrm{mech}}][1_{\mathrm{H}}^{\mathrm{D}}]} - J_{5\Lambda'} \ln \frac{[\mathrm{D}^{\mathrm{mech}}][1_{\mathrm{H}}^{\mathrm{H}}]}{[\mathrm{H}^{\mathrm{mech}}][1_{\mathrm{D}}^{\mathrm{H}}]}}_{-\dot{I}^{\mathrm{mech}}} \\
đ_t S^{\mathrm{chem}}/R &= \underbrace{(J_{1\mathrm{F}^{(\mathrm{p})}} + J_{1\mathrm{W}^{(\mathrm{p})}} - J_{3\mathrm{F}^{(\mathrm{d})}} - J_{3\mathrm{W}^{(\mathrm{d})}}) \ln \frac{[2]}{[1^{\mathrm{D}}]} + (J_{4\mathrm{F}^{(\mathrm{p})}} + J_{4\mathrm{W}^{(\mathrm{p})}} - J_{6\mathrm{F}^{(\mathrm{d})}} - J_{6\mathrm{W}^{(\mathrm{d})}}) \ln \frac{[2]}{[1^{\mathrm{H}}]}}_{\mathrm{d}_t S^{\mathrm{chem}}/R} + \\
&\underbrace{-(J_{1\mathrm{F}^{(\mathrm{p})}} + J_{1\mathrm{W}^{(\mathrm{p})}}) \ln \frac{[2][1_{\mathrm{H}}^{\mathrm{D}}]}{[1^{\mathrm{D}}][2_{\mathrm{H}}]} - (J_{3\mathrm{F}^{(\mathrm{d})}} + J_{3\mathrm{W}^{(\mathrm{d})}}) \ln \frac{[1^{\mathrm{D}}][2_{\mathrm{D}}]}{[2][1_{\mathrm{D}}^{\mathrm{D}}]} - (J_{4\mathrm{F}^{(\mathrm{p})}} + J_{4\mathrm{W}^{(\mathrm{p})}}) \ln \frac{[2][1_{\mathrm{D}}^{\mathrm{H}}]}{[1^{\mathrm{H}}][2_{\mathrm{D}}]} - (J_{6\mathrm{F}^{(\mathrm{d})}} + J_{6\mathrm{W}^{(\mathrm{d})}}) \ln \frac{[1^{\mathrm{H}}][2_{\mathrm{H}}]}{[2][1_{\mathrm{H}}^{\mathrm{H}}]}}_{-\dot{I}^{\mathrm{chem}}} .
\end{aligned}
\tag{S.46}
$$

Now the interpretation of $d_t S^{mech}$ and $d_t S^{chem}$ is clear: they are the time derivatives of the subsystem's Shannon-like entropies[39,42]. If we could measure the two subsystems separately, the entropy production rates we would assign to them would be:

$$
\sigma^{\mathrm{mech}} := \dot{S}_{\mathrm{r}}^{\mathrm{mech}} + \mathrm{d}_t S^{\mathrm{mech}} \tag{S.47}
$$

$$
\sigma^{\mathrm{chem}} := \dot{S}_{\mathrm{r}}^{\mathrm{chem,F}} + \dot{S}_{\mathrm{r}}^{\mathrm{chem,W}} + \mathrm{d}_t S^{\mathrm{chem}} . \tag{S.48}
$$

Concerning the other terms, we now justify the names $\dot{I}^{mech}$ and $\dot{I}^{chem}$ by showing that their sum gives the time variation of the mutual information between the two subsystems. The reader who is not interested in technical details can directly go to equation (S.51).

If we focus on the stochastic behavior of a single motor unit, the joint probability to find it in a certain mechanical and chemical state evolves according to the following master equation:

$$
\mathrm{d}_t \boldsymbol{p}(\mathrm{chem}, \mathrm{mech}) = \sum_{\mathrm{mech}', \mathrm{chem}'} \left[\mathbb{W}_{(\mathrm{chem},\mathrm{mech}),(\mathrm{chem}',\mathrm{mech}')} p(\mathrm{chem}', \mathrm{mech}') \right] . \tag{S.49*}
$$

For instance, it is analogous to the one employed in reference 76 to study another example of a molecular motor. Note how the structure of equation (S.49*) is analogous to the equation (S.13*) evolving the concentrations of the various motor species. Indeed, when the number N of single motor units in the system is large, the probabilities $\boldsymbol{p}(mech, chem)$ for each chemical species become sharply peaked on their average values. Thus, we can perform the substitution $\boldsymbol{p}(mech, chem) = \frac{\boldsymbol{N}(mech,chem)}{N} = \frac{[\boldsymbol{X}]}{L_{\boldsymbol{M}}}$, which shows that equation (S.13*) is the macroscopic limit of equation (S.49*). This is crucial, because it implies that for such a case

taking the macroscopic limit of stochastic quantities just means to switch from the probability of a certain state to its concentrations. In this case, the stochastic mutual information between mechanical and chemical subsystems reads:

$$\mathcal{I}_{\mathrm{M}} = \sum_{\mathrm{mech,chem}} p(\mathrm{mech,\,chem}) \ln \frac{p(\mathrm{mech,\,chem})}{p(\mathrm{mech})p(\mathrm{chem})} \geq 0\,, \tag{S.50*}$$

which is a measure of the correlation between the two sets of degrees of freedom. In the above equation, $p(mech) = \sum_{chem} p(mech, chem)$ and $p(chem) = \sum_{mech} p(mech, chem)$. Thus, the mutual information in terms of macroscopic concentrations can be defined as:

$$\begin{aligned} \mathcal{I} := L_{\mathrm{M}}\mathcal{I}_{\mathrm{M}} = \Bigg(& [2_{\mathrm{H}}] \ln \frac{[2_{\mathrm{H}}]}{[\mathrm{H}^{\mathrm{mech}}][2]} + [2_{\mathrm{D}}] \ln \frac{[2_{\mathrm{D}}]}{[\mathrm{D}^{\mathrm{mech}}][2]} + [1^{\mathrm{D}}_{\mathrm{H}}] \ln \frac{[1^{\mathrm{D}}_{\mathrm{H}}]}{[\mathrm{H}^{\mathrm{mech}}][1^{\mathrm{D}}]} + \\ & + [1^{\mathrm{D}}_{\mathrm{D}}] \ln \frac{[1^{\mathrm{D}}_{\mathrm{D}}]}{[\mathrm{D}^{\mathrm{mech}}][1^{\mathrm{D}}]} + [1^{\mathrm{H}}_{\mathrm{D}}] \ln \frac{[1^{\mathrm{H}}_{\mathrm{D}}]}{[\mathrm{D}^{\mathrm{mech}}][1^{\mathrm{H}}]} + [1^{\mathrm{H}}_{\mathrm{H}}] \ln \frac{[1^{\mathrm{H}}_{\mathrm{H}}]}{[\mathrm{H}^{\mathrm{mech}}][1^{\mathrm{H}}]} \Bigg) + L_{\mathrm{M}} \ln(L_{\mathrm{M}})\,, \end{aligned} \tag{S.51}$$

and by straightforward computation and equation (S.18) we get

$$\mathrm{d}_t \mathcal{I} = \dot{\mathcal{I}}^{\mathrm{mech}} + \dot{\mathcal{I}}^{\mathrm{chem}}\,. \tag{S.52}$$

To summarize, we showed that the entropy production of the full system can be written as a sum of two individually positive terms:

$$T\dot{\Sigma} = T\dot{\Sigma}^{\mathrm{mech}} + T\dot{\Sigma}^{\mathrm{chem}} = \underbrace{T\sigma^{\mathrm{mech}} - RT\dot{\mathcal{I}}^{\mathrm{mech}}}_{\geq 0} + \underbrace{T\sigma^{\mathrm{chem}} - RT\dot{\mathcal{I}}^{\mathrm{chem}}}_{\geq 0}\,. \tag{S.53}$$

At stationary state, we have that $d_t I = 0 = \dot{I}_{SS}^{mech} + \dot{I}_{SS}^{chem}$ (equation (S.51)), and also $d_t S^{mech}$ and $d_t S^{chem}$ vanish, as they are state functions. By specializing equation (S.53) to the stationary state (where relations (S.31) are employed to simplify the final expression) and defining $\dot{I}_{SS}^{mech} = -\dot{I}_{SS}^{chem} = -\dot{I}$ and $T\dot{S}_{r,SS}^{mech} = \dot{\mathcal{E}}$, we finally get the equation (1) in Section 4.2.2. (where the ss symbol has been dropped since there we unambiguously refer to the stationary state):

$$T\dot{\Sigma}^{\mathrm{ss}} = T\dot{\Sigma}^{\mathrm{chem}}_{\mathrm{ss}} + T\dot{\Sigma}^{\mathrm{mech}}_{\mathrm{ss}} = \underbrace{I^{\mathrm{ss}}_{\mathrm{F}}(\mu_{\mathrm{F}} - \mu_{\mathrm{W}}) - \dot{\mathcal{E}} - RT\dot{\mathcal{I}}}_{\geq 0} + \underbrace{\dot{\mathcal{E}} + RT\dot{\mathcal{I}}}_{\geq 0} \geq 0\,. \tag{S.54}$$

4.5.5. Thermodynamic constraints on the stationary state dynamics and connection with kinetic asymmetry: the rotary motor as a Rosetta Stone for kinetic and thermodynamic analysis

In this section, we examine in detail how equation (S.54) implies previous results obtained in the literature, namely the kinetic asymmetry rule[15,16]. In general, we can distinguish three possible stationary states for this system. The equilibrium (i) is always reached when the chemical potential gradient acting on the system is null ($\mu_F = \mu_W$), no matter the values of the rate constants (provided they fulfill Wegscheider's conditions in equations (S.36)). At equilibrium each reaction current vanishes ($J_\rho^{eq} = J_{+\rho}^{eq} - J_{-\rho}^{eq} = 0$) and no average net displacement of the macrocycle with respect to the track can be observed. Whenever the chemical potentials of fuel and waste species are kept different (*e.g.*, $\mu_F > \mu_W$), a nonequilibrium stationary state with net currents is reached in the long time limit (see Section 4.5.2.2.). This situation can be "symmetric" (ii) – always when the position of the macrocycle does not affect the rates of the fueling and waste-forming reactions – or "asymmetric" (iii) – the general case for the experiment we are considering. The latter is the interesting one from the point of view of directionality, because it corresponds to an average directional motion of the macrocycle along the track characterized by a non-null stationary state current J^{SS} (equation (S.31)).

4.5.5.1. Condition for directional J^{SS}

By considering the decomposition of the dissipation in equation (S.54):

$$T\dot{\Sigma}^{ss} = \overbrace{I_F^{ss}(\mu_F - \mu_W) - \dot{\mathcal{E}} - RT\dot{\mathcal{I}}}^{T\dot{\Sigma}_{ss}^{chem}} + \overbrace{\underbrace{J^{ss}RT\ln\frac{k_{+\Delta}k'_{+\Delta}}{k_{-\Delta}k'_{-\Delta}}}_{\dot{\mathcal{E}}} + \underbrace{J^{ss}RT\ln\frac{[1_H^D]_{ss}[1_D^H]_{ss}}{[1_D^D]_{ss}[1_H^H]_{ss}}}_{RT\dot{\mathcal{I}}}}^{T\dot{\Sigma}_{ss}^{mech}}, \quad \text{(S.55)}$$

the following inequalities hold by virtue of the second law of thermodynamics for bipartite systems:

$$T\dot{\Sigma}_{\mathrm{ss}}^{\mathrm{chem}} = \mathrm{I_F}^{\mathrm{ss}}(\mu_{\mathrm{F}} - \mu_{\mathrm{W}}) - T\dot{\Sigma}_{\mathrm{ss}}^{\mathrm{mech}} \geq 0 \tag{S.56a}$$

$$T\dot{\Sigma}_{\mathrm{ss}}^{\mathrm{mech}} = \dot{\mathcal{E}} + RT\dot{\mathcal{I}} = J^{\mathrm{ss}}RT\left(\ln\frac{k_{+\Delta}k'_{+\Delta}}{k_{-\Delta}k'_{-\Delta}} + \ln\frac{[1_{\mathrm{H}}^{\mathrm{D}}]_{\mathrm{ss}}[1_{\mathrm{D}}^{\mathrm{H}}]_{\mathrm{ss}}}{[1_{\mathrm{D}}^{\mathrm{D}}]_{\mathrm{ss}}[1_{\mathrm{H}}^{\mathrm{H}}]_{\mathrm{ss}}}\right) \geq 0\,. \tag{S.56b}$$

As highlighted in Section 4.2.2., net directional displacement of the macrocycle in the stationary state can be observed only in the presence of mechanical dissipation ($T\dot{\Sigma}_{SS}^{mech} > 0$), as directed motion is a nonequilibrium behaviour. This implies that in order to have a non null J^{SS}, a positive fueling work is required ($I_F^{SS}(\mu_F - \mu_W) > 0$ in equation (S.56a), otherwise the stationary state would be an equilibrium one). This is not sufficient, since the sum of energy and information flow in equation (S.56b) must be positive too, namely a net free energy transfer from chemical to mechanical processes must be present.

The condition for a net directional displacement of the macrocycle in the stationary state can then be written as:

$$\ln\frac{k_{+\Delta}k'_{+\Delta}}{k_{-\Delta}k'_{-\Delta}} + \ln\frac{[1_{\mathrm{H}}^{\mathrm{D}}]_{\mathrm{ss}}[1_{\mathrm{D}}^{\mathrm{H}}]_{\mathrm{ss}}}{[1_{\mathrm{D}}^{\mathrm{D}}]_{\mathrm{ss}}[1_{\mathrm{H}}^{\mathrm{H}}]_{\mathrm{ss}}} = \ln\frac{J^{\mathrm{ss}}_{+2\Delta}J^{\mathrm{ss}}_{+5\Delta'}}{J^{\mathrm{ss}}_{-2\Delta}J^{\mathrm{ss}}_{-5\Delta'}} \neq 0\,. \tag{S.57}$$

Thanks to equation (S.31), it is easy to realize that whenever the quantity in equation (S.57) is positive, J^{SS} is positive too, and vice-versa, thus guaranteeing equation (S.56b) to always hold. We will now show how the condition (S.57) (obtained with nonequilibrium thermodynamic arguments) is equivalent to previous results obtained in the literature (based on kinetic arguments).

4.5.5.2. Connection to kinetic asymmetry and K_r

According to references 17 and 49, the condition for having directional currents in nonequilibrium chemical systems ($J^{SS} \neq 0$) can be expressed in terms of the ratcheting constant, which for the motor under study reads

$$K_{\mathrm{r}}=\frac{k_{+\Delta}k'_{+\Delta}(k^{\mathrm{p}}_{-\mathrm{F}}[\mathrm{HCl}]+k^{\mathrm{p}}_{-\mathrm{W}})^2([\mathrm{F}]k^{\mathrm{d}}_{+\mathrm{F}}+[\mathrm{CO_2}][\mathrm{DBF}]k^{\mathrm{d}}_{+\mathrm{W}})^2}{k_{-\Delta}k'_{-\Delta}([\mathrm{F}]k^{\mathrm{p}}_{+\mathrm{F}}+[\mathrm{CO_2}][\mathrm{DBF}]k^{\mathrm{p}}_{+\mathrm{W}})^2(k^{\mathrm{d}}_{-\mathrm{F}}[\mathrm{HCl}]+k^{\mathrm{d}}_{-\mathrm{W}})^2}\neq 1\,, \tag{S.58}$$

with $K_r > 1$ implying a positive forward current, and vice-versa. By using Wegscheider's conditions (S.36), K_r can be rewritten as[15–17,56]:

$$K_{\mathrm{r}}=\underbrace{\frac{k_{+\Delta}k'_{+\Delta}(k^{\mathrm{p}}_{-\mathrm{F}}[\mathrm{HCl}])^2([\mathrm{F}]k^{\mathrm{d}}_{+\mathrm{F}})^2}{k_{-\Delta}k'_{-\Delta}([\mathrm{F}]k^{\mathrm{p}}_{+\mathrm{F}})^2(k^{\mathrm{d}}_{-\mathrm{F}}[\mathrm{HCl}])^2}}_{=1,\text{ see equation (S.36a)}}\frac{(1+\frac{k^{\mathrm{p}}_{-\mathrm{W}}}{k^{\mathrm{p}}_{-\mathrm{F}}[\mathrm{HCl}]})^2(1+\frac{[\mathrm{CO_2}][\mathrm{DBF}]k^{\mathrm{d}}_{+\mathrm{W}}}{[\mathrm{F}]k^{\mathrm{d}}_{+\mathrm{F}}})^2}{(1+\frac{[\mathrm{CO_2}][\mathrm{DBF}]k^{\mathrm{p}}_{+\mathrm{W}}}{[\mathrm{F}]k^{\mathrm{p}}_{+\mathrm{F}}})^2(1+\frac{k^{\mathrm{d}}_{-\mathrm{W}}}{k^{\mathrm{d}}_{-\mathrm{F}}[\mathrm{HCl}]})^2}=\frac{(1+\frac{k^{\mathrm{p}}_{-\mathrm{W}}}{k^{\mathrm{p}}_{-\mathrm{F}}[\mathrm{HCl}]})^2(1+\frac{[\mathrm{CO_2}][\mathrm{DBF}]k^{\mathrm{d}}_{+\mathrm{W}}}{[\mathrm{F}]k^{\mathrm{d}}_{+\mathrm{F}}})^2}{(1+\frac{[\mathrm{CO_2}][\mathrm{DBF}]k^{\mathrm{p}}_{+\mathrm{W}}}{[\mathrm{F}]k^{\mathrm{p}}_{+\mathrm{F}}})^2(1+\frac{k^{\mathrm{d}}_{-\mathrm{W}}}{k^{\mathrm{d}}_{-\mathrm{F}}[\mathrm{HCl}]})^2}\,. \tag{S.59}$$

In the following, we show how the condition $K_r \neq 1$ for a directional current $J^{SS} \neq 0$ is implied by equations (S.57).

Equilibrium condition

The expression of the ratcheting constant in equation (S.59) can be further rearranged thanks to equations (S.32) and (S.33) to get (see also equation (5) in reference 15):

$$K_{\mathrm{r}}=\frac{(1+\frac{k^{\mathrm{p}}_{-\mathrm{W}}}{k^{\mathrm{p}}_{-\mathrm{F}}[\mathrm{HCl}]})^2(1+\frac{k^{\mathrm{d}}_{-\mathrm{W}}}{k^{\mathrm{d}}_{-\mathrm{F}}[\mathrm{HCl}]}\mathrm{e}^{-\frac{\mu_{\mathrm{F}}-\mu_{\mathrm{W}}}{RT}})^2}{(1+\frac{k^{\mathrm{p}}_{-\mathrm{W}}}{k^{\mathrm{p}}_{-\mathrm{F}}[\mathrm{HCl}]}\mathrm{e}^{-\frac{\mu_{\mathrm{F}}-\mu_{\mathrm{W}}}{RT}})^2(1+\frac{k^{\mathrm{d}}_{-\mathrm{W}}}{k^{\mathrm{d}}_{-\mathrm{F}}[\mathrm{HCl}]})^2}\,, \tag{S.60}$$

which nicely shows that $K_r = 1$ (and then $J^{SS} = 0$) whenever $\mu_F = \mu_W$, in analogy with equation (S.56a). Regardless of the values of the rate constants, if there is no thermodynamic force ($\mu_F - \mu_W = 0$), the motor will relax to an equilibrium stationary state with no directional currents, and the quantity in equation (S.57) will be zero.

Proof of the implication $\dot{\Sigma}^{mech} > 0 \rightarrow K_r \neq 1$

We now prove that the condition on K_r in equation (S.58) directly follows from equation (S.57). We do so by relying on some results and techniques from graph theory (see appendixes of reference 67 for an introduction). The reader who is not familiar with this kind

of techniques can directly go to the next paragraph, where a simpler proof valid for the experimental case (where equations (S.1) hold) is presented.

First of all, we define K_r in terms of the pseudo rate constants introduced in equation (S.13*):

$$K_{\mathrm{r}} = \frac{\tilde{k}_{+\Delta}\tilde{k}'_{+\Delta}(\tilde{k}^{\mathrm{p}}_{-}\tilde{k}^{\mathrm{d}}_{+})^2}{\tilde{k}_{-\Delta}\tilde{k}'_{-\Delta}(\tilde{k}^{\mathrm{p}}_{+}\tilde{k}^{\mathrm{d}}_{-})^2} = \frac{\mathcal{P}_{\mathrm{cw}}}{\mathcal{P}_{\mathrm{ccw}}}, \tag{S.61}$$

where $\mathcal{P}_{CW}$ and $\mathcal{P}_{CCW}$ are the product of all forward clockwise and counterclockwise pseudo reaction constants, respectively. By plugging the analytical expression (S.27) for the steady state concentrations into the condition (S.57), the latter can be rewritten as

$$\left(\tilde{k}_{+\Delta}\sum_{t\in\mathcal{T}^{\mathrm{D}}_{\mathrm{I}\,\mathrm{H}}}\prod_{\rho\in t}\tilde{k}_{\rho}\right)\cdot\left(\tilde{k}'_{+\Delta}\sum_{t\in\mathcal{T}^{\mathrm{H}}_{\mathrm{I}\,\mathrm{D}}}\prod_{\rho\in t}\tilde{k}_{\rho}\right)-\left(\tilde{k}_{-\Delta}\sum_{t\in\mathcal{T}^{\mathrm{D}}_{\mathrm{I}\,\mathrm{D}}}\prod_{\rho\in t}\tilde{k}_{\rho}\right)\cdot\left(\tilde{k}'_{-\Delta}\sum_{t\in\mathcal{T}^{\mathrm{H}}_{\mathrm{I}\,\mathrm{H}}}\prod_{\rho\in t}\tilde{k}_{\rho}\right)\neq 0\,. \tag{S.62*}$$

We can manipulate expression (S.62*) to get

$$\begin{aligned}
&\left(\tilde{k}_{+\Delta}\sum_{t\in\mathcal{T}^{\mathrm{D}}_{\mathrm{I}\,\mathrm{H}}\wedge\tilde{k}_{-\Delta}\in t}\prod_{\rho\in t}\tilde{k}_{\rho}+\mathcal{P}_{\mathrm{cw}}\right)\cdot\left(\tilde{k}'_{+\Delta}\sum_{t\in\mathcal{T}^{\mathrm{H}}_{\mathrm{D}}\wedge\tilde{k}'_{-\Delta}\in t}\prod_{\rho\in t}\tilde{k}_{\rho}+\mathcal{P}_{\mathrm{cw}}\right)\\
&-\left(\tilde{k}_{-\Delta}\sum_{t\in\mathcal{T}^{\mathrm{D}}_{\mathrm{I}\,\mathrm{D}}\wedge\tilde{k}_{+\Delta}\in t}\prod_{\rho\in t}\tilde{k}_{\rho}+\mathcal{P}_{\mathrm{ccw}}\right)\cdot\left(\tilde{k}'_{-\Delta}\sum_{t\in\mathcal{T}^{\mathrm{H}}_{\mathrm{I}\,\mathrm{H}}\wedge\tilde{k}'_{+\Delta}\in t}\prod_{\rho\in t}\tilde{k}_{\rho}+\mathcal{P}_{\mathrm{ccw}}\right)=\\
&0+(\mathcal{P}^2_{\mathrm{cw}}-\mathcal{P}^2_{\mathrm{ccw}})+\mathcal{P}_{\mathrm{cw}}\underbrace{\left[\tilde{k}_{+\Delta}\sum_{t\in\mathcal{T}^{\mathrm{D}}_{\mathrm{I}\,\mathrm{H}}\wedge\tilde{k}_{-\Delta}\in t}\prod_{\rho\in t}\tilde{k}_{\rho}+\tilde{k}'_{+\Delta}\sum_{t\in\mathcal{T}^{\mathrm{H}}_{\mathrm{I}\,\mathrm{D}}\wedge\tilde{k}'_{-\Delta}\in t}\prod_{\rho\in t}\tilde{k}_{\rho}\right]}_{\Phi}\\
&-\mathcal{P}_{\mathrm{ccw}}\underbrace{\left[\tilde{k}_{-\Delta}\sum_{t\in\mathcal{T}^{\mathrm{D}}_{\mathrm{I}\,\mathrm{D}}\wedge\tilde{k}_{+\Delta}\in t}\prod_{\rho\in t}\tilde{k}_{\rho}+\tilde{k}'_{-\Delta}\sum_{t\in\mathcal{T}^{\mathrm{H}}_{\mathrm{I}\,\mathrm{H}}\wedge\tilde{k}'_{+\Delta}\in t}\prod_{\rho\in t}\tilde{k}_{\rho}\right]}_{\Phi}=(\mathcal{P}_{\mathrm{cw}}-\mathcal{P}_{\mathrm{ccw}})\cdot(\mathcal{P}_{\mathrm{cw}}+\mathcal{P}_{\mathrm{ccw}}+\Phi)\neq 0\,.
\end{aligned} \tag{S.63*}$$

Since $\mathcal{P}_{CW}$, $\mathcal{P}_{CCW}$, and Φ are positive quantities, the above condition is satisfied if and only if $\mathcal{P}_{CW}-\mathcal{P}_{CCW}\neq 0$ or, alternatively, $\mathcal{P}_{CW}/\mathcal{P}_{CCW}\neq 1$, thus fully proving the equivalence between condition (S.58) from references 15, 16, 17 and 49 and condition (S.57) derived from the second law (S.56) for bipartite systems[39].

Special proof for the experimental case

Here, we repeat the above proof for the special case of the experimental rotary motor, where the symmetries in equations (S.1) hold and the expressions for stationary state concentrations is given by equations (S.28). By plugging the latter in equation (S.57), the condition for directional current boils down to:

$$\frac{(k^{\mathrm{p}}_{-\mathrm{F}}[\mathrm{HCl}] + k_{-\mathrm{W}})(k^{\mathrm{d}}_{+\mathrm{F}}[\mathrm{F}] + k_{+\mathrm{W}}[\mathrm{CO_2}][\mathrm{DBF}]) + (k^{\mathrm{p}}_{-\mathrm{F}}[\mathrm{HCl}] + k^{\mathrm{d}}_{-\mathrm{F}}[\mathrm{HCl}] + 2k_{-\mathrm{W}})k_{\Delta}}{(k^{\mathrm{d}}_{-\mathrm{F}}[\mathrm{HCl}] + k_{-\mathrm{W}})(k^{\mathrm{p}}_{+\mathrm{F}}[\mathrm{F}] + k_{+\mathrm{W}}[\mathrm{CO_2}][\mathrm{DBF}]) + (k^{\mathrm{p}}_{-\mathrm{F}}[\mathrm{HCl}] + k^{\mathrm{d}}_{-\mathrm{F}}[\mathrm{HCl}] + 2k_{-\mathrm{W}})k_{\Delta}} \neq 1\,, \tag{S.64}$$

which is equivalent to

$$(k^{\mathrm{p}}_{-\mathrm{F}}[\mathrm{HCl}] + k_{-\mathrm{W}})(k^{\mathrm{d}}_{+\mathrm{F}}[\mathrm{F}] + k_{+\mathrm{W}}[\mathrm{CO_2}][\mathrm{DBF}]) \neq (k^{\mathrm{d}}_{-\mathrm{F}}[\mathrm{HCl}] + k_{-\mathrm{W}})(k^{\mathrm{p}}_{+\mathrm{F}}[\mathrm{F}] + k_{+\mathrm{W}}[\mathrm{CO_2}][\mathrm{DBF}])\,, \tag{S.65}$$

and so

$$\frac{(k^{\mathrm{p}}_{-\mathrm{F}}[\mathrm{HCl}] + k_{-\mathrm{W}})(k^{\mathrm{d}}_{+\mathrm{F}}[\mathrm{F}] + k_{+\mathrm{W}}[\mathrm{CO_2}][\mathrm{DBF}])}{(k^{\mathrm{d}}_{-\mathrm{F}}[\mathrm{HCl}] + k_{-\mathrm{W}})(k^{\mathrm{p}}_{+\mathrm{F}}[\mathrm{F}] + k_{+\mathrm{W}}[\mathrm{CO_2}][\mathrm{DBF}])} \neq 1\,, \tag{S.66}$$

which is exactly the square root of K_r in equation (S.58) specialized for the experimental case.

On the magnitude of the stationary state current J^{SS}

We conclude this section by deriving the expression of the stationary state current displayed in equation (4) of Section 4.2.2., which proves particularly useful to rationalize and discuss numerical simulations in Section 4.5.6..

First of all, we rewrite the stationary state current defined in equation (S.31a) and equation (4) of Section 4.2.2. in terms of the pseudo rate constants introduced in equation (S.13*):

$$J^{\mathrm{ss}} = \tilde{k}_{+\Delta}[1^{\mathrm{D}}_{\mathrm{H}}]_{\mathrm{ss}} - \tilde{k}_{-\Delta}[1^{\mathrm{D}}_{\mathrm{D}}]_{\mathrm{ss}} = \tilde{k}'_{+\Delta}[1^{\mathrm{H}}_{\mathrm{D}}]_{\mathrm{ss}} - \tilde{k}'_{-\Delta}[1^{\mathrm{H}}_{\mathrm{H}}]_{\mathrm{ss}}\,. \tag{S.67}$$

This rewriting is trivial, since $\tilde{k}_{+\Delta} = k_{+\Delta}, \tilde{k}_{-\Delta} = k_{-\Delta}, \tilde{k'}_{+\Delta} = k'_{+\Delta}$ and $\tilde{k'}_{-\Delta} = k'_{-\Delta}$ (see equation (S.13*)), but by plugging the analytical expression (S.27) for stationary state concentrations into equation (S.67) we get

$$J^{ss} = \frac{L_M}{\mathcal{N}} \left(\tilde{k}_{+\Delta} \sum_{t \in \mathcal{T}_{iD}^{D}} \prod_{\lambda \in t} \tilde{k}_\lambda - \tilde{k}_{-\Delta} \sum_{t \in \mathcal{T}_{iD}^{D}} \prod_{\lambda \in t} \tilde{k}_\lambda \right) =$$

$$= \frac{L_M}{\mathcal{N}} \left(\tilde{k}_{+\Delta} \sum_{t \in \mathcal{T}_{iD}^{D} \wedge \tilde{k}_{-\Delta} \in t} \prod_{\rho \in t} \tilde{k}_\rho + \mathcal{P}_{cw} - \tilde{k}_{-\Delta} \sum_{t \in \mathcal{T}_{iD}^{D} \wedge \tilde{k}_{+\Delta} \in t} \prod_{\rho \in t} \tilde{k}_\rho - \mathcal{P}_{ccw} \right) = \frac{L_M}{\mathcal{N}} (\mathcal{P}_{cw} - \mathcal{P}_{ccw}) =$$

$$= \frac{L_M \mathcal{P}_{ccw}}{\mathcal{N}} (K_r - 1) = \Gamma (K_r - 1), \qquad \text{(S.68*)}$$

where we introduced the positive quantity Γ appearing in equation (4) of Section 4.2.2.. All the other symbols appearing in the above equation have been defined in equations (S.27) and (S.61). Equation (S.68) shows that the sign of the current J^{SS} is determined by K_r being smaller or greater than 1, with no current for $K_r = 1$, while its magnitude depends on the actual value of both K_r and Γ. Crucially, the value of Γ can be varied independently from the value of K_r, thus allowing for design modifications which can improve or stall a motor's performance (but not altering the direction of cycling) without varying its kinetic asymmetry. This kind of effects are explored in Section 4.2.3. and in the next section with the help of numerical simulations.

4.5.6. Numerical simulations

In order to better understand the implications of the model for the design of molecular motors, we performed a series of numerical simulations solving numerically dynamical equations (S.10) while varying key rate constants and plotting the predicted effect on the stationary state current (J in equation (4) in Section 4.2.2., J^{SS} in equations (S.31a) and (S.68)), the thermodynamic efficiency (η, equation (5) in Section 4.2.2.), and the quantity J/I_F (stationary state current divided by the rate at which fuel is consumed at the stationary state), which equates to the average number of cycles achieved per molecule of fuel (see equation (S.43)). As stated in Section 4.5.3., this last measure is a coefficient of performance which is distinct from efficiency because it is independent of both the energy content of the

fuel, instead looking at the effect of each fuel molecule, and of whether the energy is dissipated mechanically or chemically, only reflecting how well fuel use correlates with unidirectional movement. In all the simulations, thermodynamic consistency was always guaranteed by enforcing Wegscheider's conditions (S.36).

It should be noted that the current is unbounded and can be positive or negative, indicating forward or backward movement respectively, thermodynamic efficiency is bounded between 0 and 1, and J/I_F is bounded between 0.5 and −0.5 (the sign indicating direction). The value has a maximum of 0.5 because at minimum 2 molecules of fuel are required for a complete cycle as the rotary motor is a two-stroke design, with two barriers to move past per cycle.

The parameters in **Table I** were adopted as reference in the numerical simulations of the model introduced in Section 4.5.1.. The chosen values are intended to be realistic and give a correct qualitative outcome. A precise quantitative analysis would require more experimental data and is out of the scope of the present work. Apart where explicitly said or shown in the graphs, reference parameters are kept constant in the numerical simulations.

$k^{\mathrm{d}}_{+\mathrm{F}}$	$8 \cdot 10^{-3}$ mol^{-1} dm^{3}s^{-1}		
$k^{\mathrm{d}}_{-\mathrm{F}}$	$1 \cdot 10^{-27}$ mol^{-1} dm^{3}s^{-1}	L_{M}	$1 \cdot 10^{-2}$ mol dm^{-3}
$k^{\mathrm{p}}_{+\mathrm{F}}$	$2 \cdot 10^{-3}$ mol^{-1} dm^{3}s^{-1}	[Fmoc−Cl]	$3 \cdot 10^{-2}$ mol dm^{-3}
$k^{\mathrm{p}}_{-\mathrm{F}}$	$2.5 \cdot 10^{-28}$ mol^{-1} dm^{3}s^{-1}	[DBF] · [CO_2]	$1 \cdot 10^{-5}$ mol^2 dm^{-6}
$k_{-\mathrm{W}}$	$8 \cdot 10^{-6}$ s^{-1}	[HCl]	$1 \cdot 10^{-18}$ mol dm^{-3}
$k_{+\mathrm{W}}$	$1 \cdot 10^{-26}$ mol^{-2}dm^{6}s^{-1}		
k_{Δ}	$1 \cdot 10^{1}$ s^{-1}		

Table I. Reference parameters employed in the numerical simulations. Rate constants $k^d_{+F}, k^p_{+F}, k_{-W}$ and k_Δ are estimated based on data collected in reference 14, where equations (S.1) hold, and related rotaxanes[77]. Rate constants k^d_{-F} and k_{+W} are calculated by using equation (S.33) and assuming, based on known enthalpy-of-formation data, a difference of the order of 10^2 kJ mol^{-1} in the standard chemical potentials between fuel and waste species ($\mu^\circ_F - \mu^\circ_W$). The rate constant k^p_{-F} is obtained by using equation (S.36a) to guarantee thermodynamic consistency. The total concentration of motor's species L_M is the same as in reference 14. The parameters involving chemostatted concentrations ([Fmoc–Cl], [DBF] · [CO_2] and [HCl]) are estimated based on the

original experiment[14] and other known data.[78] The order of magnitude of parameters [DBF] · [CO_2] and [HCl] does not significantly affect the results of simulations, since the microscopic reverses of fueling and waste-forming reactions (rate constants $k_{-F}^{p/d}$ and k_{+W}) are highly unfavorable.

4.5.6.1. Chemical gating of the fueling reaction

In this simulation, the sum of the rates of the fueling reactions and the chemical gating ratio of the fueling reaction were varied (**Figure 3a, b** in Section 4.2.3., **Figure S1** below). This is the only source of chemical gating as no chemical gating is introduced via the waste forming reactions in this simulation, as was the case in the experimental rotary motor (white dot). With no chemical gating (white line) there can be no current in the motor and therefore thermodynamic efficiency and the number of cycles per unit of fuel are also always zero. For intermediate rates of the fueling reaction, the current increases as the chemical gating increases. This is because an increased kinetic bias means the motor spends more time moving in the forward direction. A plateau is reached at higher values of chemical gating, indicating that other values (for instance, shuttling rate or waste formation rate) may become limiting to the current under these conditions. In this case, the waste-forming reaction is likely to be limiting, as it can be seen that the corresponding plateau in the 'cycles per fuel molecule' graph is at a value of approximately ±0.25. This is the theoretical limit (with perfect fueling gating) for a two stroke motor with no chemical gating of the waste-forming reaction (minimum two fuel molecules per cycle because there are two barriers, but half the fuel molecules will lead to no step because of a backwards waste-forming reaction).

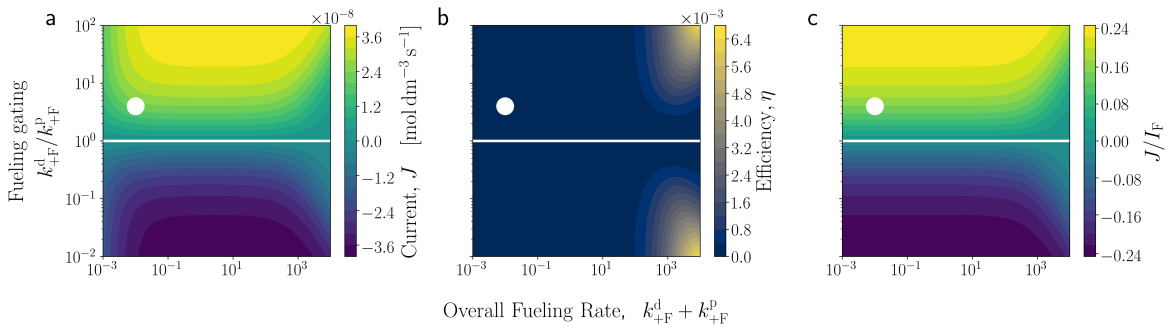


Figure S1. Graphs depicting the variation in current, efficiency, and J/I_F as overall fueling rate (x-axis) and fueling gating (y-axis) are changed.

Very slow fueling rates (far left of the graphs) slow the motor to a near-stall, as the rate of reaction limits how fast the rotation can be. The graph of cycles per unit fuel shows that the motor is not actually stalled, just extremely slow as the motor still achieves the same number of cycles for each unit of fuel under these conditions. Less intuitively, very fast fueling rates (far right of the graphs) can actually stall the motor. This is because the fueling reaction is so fast that the mechanical distribution is effectively trapped as there is not sufficient time for mechanical equilibration before both co-conformers react. This promotes futile cycles, so fuel is still used despite less directional motion being achieved and therefore fewer cycles are performed for each fuel molecule used. Directional motion is still possible, but a greater chemical gating ratio is needed to achieve it. Under these conditions with a relatively very fast fueling reaction, a motor can be effectively stalled even if the chemomechanical cycle has kinetic asymmetry arising from the chemical gating.

The optimal thermodynamic efficiency balances the competing factors of a fast fueling rate, allowing the mechanical state to be pushed further from equilibrium, and fueling rates that are slow enough to limit futile cycles. The symmetric hotspots (thermodynamic efficiency is independent of direction) are therefore on the right of the graph, towards faster fueling rates where this balance is struck best. The current and J/I_F graphs are anti-symmetric around the horizontal axis with a chemical gating ratio of one (*i.e.*, with the chemical gating ratio inverted, the same current and J/I_F are achieved but in the opposite direction).

4.5.6.2. Chemical gating of the waste-forming reaction

In this simulation, the overall rate and chemical gating ratio of the waste-forming reaction were varied, while the chemical gating ratio of the fueling reaction was retained at a value of 4 (based on experimental measurements[14], **Figure S2**, shown by the white dot in the figures). Therefore, even in the absence of waste-formation chemical gating, there is a driving force for the directional mechanical movement in the rotary motor from a positive information flow. Primarily this simulation shows the same as the fueling gating equivalent, while also showing that the waste-forming gating can cancel out the free energy transfer to the mechanical processes from the fueling gating and stall the motor. This gives the chemomechanical cycle no net kinetic asymmetry which is a condition synonymous with zero mechanical dissipation due to the absence of net free-energy flow between chemical and mechanical transitions. Where this direct negation occurs (white line), the current, efficiency and number of cycles per unit fuel are all inherently zero.

An interesting difference is that fast waste-forming reaction do not appear to cause a loss of current, efficiency, or cycles per fuel molecule. This is because the states preceding a waste-forming reaction (2_H or 2_D) are not able to undergo mechanical exchange and are therefore inherently always at mechanical equilibrium (precluding the mechanism by which fast fueling leads to a lower current, efficiency, and cycles per fuel). However, as the current model neglects the possibility for the rotary motor to lose both barriers, which is more likely to occur with faster waste-forming reactions, deviations from this behavior are expected. If this detail were included, it would be expected that faster waste-forming reactions would also lead to a loss of current, efficiency, and cycles per fuel.

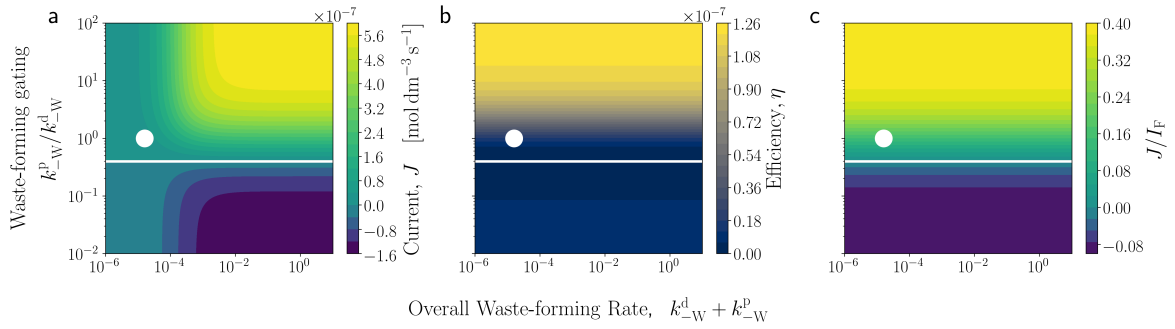


Figure S2. Graphs depicting the variation in current, efficiency, and J/I_F as overall waste-forming rate (x-axis) and waste-forming gating (y-axis) are changed.

4.5.6.3. Further exploration of chemical gating

Better currents and efficiencies are theoretically obtainable with double-gated machines, in which both the fueling and waste-forming processes are chemically gated[24]. A simulation in which the gating ratios of both reactions could be independently varied (**Figure S3**) supports this idea, showing that the highest currents and efficiencies are predicted for motors in which both chemical processes are gated (top right and bottom left). Furthermore, no matter the magnitude of gating for one process, the motor would be stalled if the other process is biased in the opposite direction (white line, bottom right to top left of the graphs), leading to zero total kinetic asymmetry and precluding mechanical dissipation. Although the overall rates of the fueling and waste-forming reactions are different within the model, it appears that the contributions of fueling and waste-forming gating are symmetric and of equal importance for current, efficiency, and J/I_F. Additionally, this was the only simulation to approach the theoretical maximum limit of 0.5 cycles per fuel molecule, demonstrating that this requires large chemical gating values for both chemical processes. Under the conditions in which this limit is reached (top right and bottom left), it follows that further improvements to current and efficiency would have to come from other parts of the chemomechanical cycle. The limiting factor in this simulation is most probably the overall

rates of the fueling and waste-forming reactions, though could plausibly be shuttling rate under other conditions.

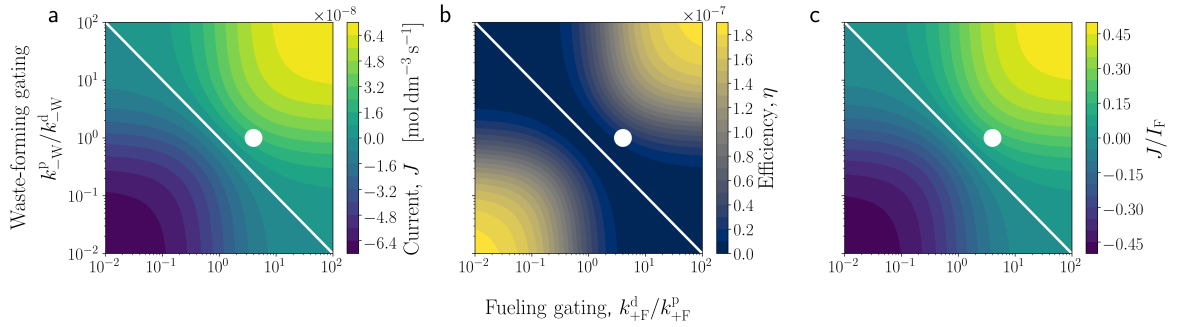


Figure S3. Graphs depicting the variation in current, efficiency, and J/I_F as chemical gating of the fueling reactions (x-axis) and the waste-forming reactions (y-axis) are changed. Shuttling rates and overall rates of the fueling and waste-forming reactions are kept constant at experimental values throughout.

4.5.6.4. Introducing energy flow with power strokes without varying kinetic asymmetry

In this simulation, the shuttling rate and ratio of clockwise and anticlockwise shuttling rates were varied (**Figure 3 c,d** in Section 4.2.3., **Figure S4** below), introducing energy flow via two power strokes. The simulation treats both halves of the rotary motor cycle (**Figure 2**) as identical ($k_{+\Delta} = k'_{+\Delta}, k_{-\Delta} = k'_{-\Delta}$). As explained in Section 4.4., in this simulation Wegscheider's consistency conditions (equations (S.36)) were imposed by varying rate constants k^{p}_{+W} and k^{p}_{+F} according to variations in shuttling rate constants, such that the kinetic asymmetry (K_r) of the chemomechanical cycle is not changed by the introduction of power strokes (effectively keeping the absolute transition state energies of the chemical transitions constant), which can nevertheless vary the positive factor Γ in equation (S.68). This allows us to explore changes caused only by power strokes to identify if, and how, power strokes have any effect on the behavior of the molecular motor.

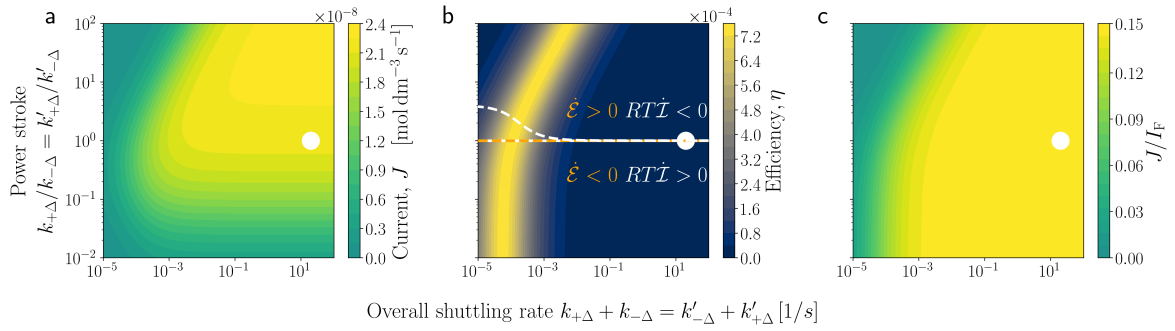


Figure S4. Graphs depicting the variation in current, efficiency, and J/I_F as overall shuttling rate (x-axis) and power stroke magnitude (y-axis) are changed. The power strokes are added in such a way as to keep the transition state energies and hence the kinetic asymmetry constant by compensating for changes to the mechanical transitions with the rates of the chemical transitions (see Section 4.4.).

For relatively slow shuttling rates (far left of the graphs) the motor is effectively stalled, with no significant current, thermodynamic efficiency, or cycles per unit fuel. As well as slow shuttling leading directly to low current, this situation is synonymous with the fast fueling reaction condition described above, in which the mechanical equilibration is slow enough that the mechanical distribution is trapped out by the relatively fast fueling reaction. Fast shuttling rates (far right of the graphs) do not appear detrimental to either the current or the number of cycles per fuel molecule, but thermodynamic efficiency with respect to how free energy is dissipated is strongly adversely affected. This is because fast shuttling hinders the generation of a concentration bias with respect to the equilibrium distribution in the mechanical states by more rapidly enabling equilibration. The position of optimal thermodynamic efficiency is determined by a trade-off between shuttling slow enough that the mechanical states are significantly far from their equilibrium distribution in the steady state, but fast enough not to too strongly promote futile cycles. It is particularly noticeable from the graphs that the optimal region of thermodynamic efficiency contours around the region in which J/I_F, the number of cycles per fuel molecule, moves from a non-productive plateau (left of the graphs) to a productive plateau (right of the graphs). This is because slow shuttling decreases both efficiency and J/I_F, while the latter is independent of the path by

which the energy from the fuel is dissipated. As a result, it is unaffected as the degree of energy dissipated through the mechanical processes decreases because of faster shuttling.

A key difference can be seen between graphs in the region describing negative energy flow from backwards power strokes (bottom, where $\frac{k_{+\Delta}}{k_{-\Delta}} = \frac{k'_{+\Delta}}{k'_{-\Delta}} < 1$, so clockwise shuttling requires gain of chemical potential). In this region, the thermodynamic efficiency and the number of cycles per unit fuel are independent of the magnitude of the power stroke and dependent only on the overall shuttling rate. Conversely, the current is strongly dependent on the power stroke magnitude. This arises because the current is directly reduced by the slow forward mechanical transitions associated with backward power strokes which limit the rate of the forwards chemomechanical cycles. However, efficiency and J/I_F are not affected as the decrease in the current (J) is offset by the decrease in the fuel consumption rate (I_F).

In the region with a positive energy flow arising from forwards power strokes, (top, where $\frac{k_{+\Delta}}{k_{-\Delta}} = \frac{k'_{+\Delta}}{k'_{-\Delta}} > 1$, so clockwise shuttling lowers standard chemical potential) all three quantities are dependent on both the overall shuttling rate and the magnitude of the power stroke. If the power stroke is too large or overall shuttling rate is too small (top and left of the graphs), futile cycles are favored, decreasing the current and thermodynamic efficiency and requiring more fuel per cycle on average. This is because the proximal species (1_D^H or 1_H^D) are sufficiently destabilized so that the unwanted proximal fueling reaction is faster than mechanical shuttling. It is especially unfavorable if a large power stroke and slow overall shuttling rate are combined.

The maximum current is seen in a region with positive energy flow from forward power strokes but faster mechanical transitions (top right of the graphs). This is because, at steady state, the energy flow is increased by adding power strokes faster than information flow is decreased by biasing the mechanical equilibrium. This is a specific property of

nonequilibrium regimes, where the standard chemical potentials and mutual information contributions to the total free-energy flow, which must inherently cancel at equilibrium, may differ in a nonequilibrium steady state. In mathematical terms we have, for the sum of energy and information flows:

$$\dot{\mathcal{E}} + RT\dot{I} = J(\mu^{\circ}_{1^{\mathrm{D}}_{\mathrm{H}}} - \mu^{\circ}_{1^{\mathrm{D}}_{\mathrm{D}}} + \mu^{\circ}_{1^{\mathrm{H}}_{\mathrm{D}}} - \mu^{\circ}_{1^{\mathrm{H}}_{\mathrm{H}}} + RT \log \frac{[1^{\mathrm{D}}_{\mathrm{H}}][1^{\mathrm{H}}_{\mathrm{D}}]}{[1^{\mathrm{D}}_{\mathrm{D}}][1^{\mathrm{H}}_{\mathrm{H}}]}) \geq 0\,. \tag{S.69}$$

At mechanical equilibrium, both the current J and the quantity in parentheses are individually zero, since the concentrations are distributed accordingly to the exponential of the difference in standard chemical potentials divided by RT. As soon as the system is driven out of equilibrium, the difference in standard chemical potentials stays constant (and positive in the region we are considering). In the presence of kinetic gating in the forward direction, the magnitude of the logarithm (negative, in the region we are considering) decreases. This is because, while the equilibrium favors 1^D_D over 1^H_D and 1^H_H over 1^D_H, the fueling kinetic gating is such to flatten this correlation, thus reducing mutual information. As a result, the mutual information contribution decreases towards more negative values slower than the standard chemical potential contribution increases towards positive values, so that the sum in the parentheses is positive. In the presence of fast shuttling, the concentration distribution stays very close to the equilibrium one and the quantity in parentheses is very small, thus making the efficiency very low. In other words, when the shuttling is fast, the mechanical transitions stay very close to equilibrium. This does not contradict the fact that we observe a high current in the stationary state, because ratio $\frac{k_{+\Delta}[1^D_H]_{SS}}{k_{-\Delta}[1^D_D]_{SS}} = \frac{k'_{+\Delta}[1^H_D]_{SS}}{k'_{-\Delta}[1^H_H]_{SS}}$ being close to 1 does not imply that the difference giving the current $J = k_{+\Delta}[1^D_H]_{SS} - k_{-\Delta}[1^D_D]_{SS} = k'_{+\Delta}[1^H_D]_{SS} - k'_{-\Delta}[1^H_H]_{SS}$ is small. Indeed, this is a clear example showing why the entropy production, not the current or J/I_F, is a measure of how far from equilibrium this system operates. Despite the increased current, the average number of cycles per fuel molecule

remains unchanged between the high-current area (top right of the graphs) and the region with no or negative energy flow (bottom of the graphs), showing that only the current is affected. In the same area, the efficiency plot shows that a very low percentage of the fueling free-energy is made available to the mechanical transitions, indicating that this regime, while optimizing current (J) and J/I_F, may not be optimal to perform work (e.g., transporting cargoes).

4.5.6.5. Introducing energy flow with a power stroke together with varying kinetic asymmetry

In contrast to the previous case, if Wegscheider's consistency conditions are imposed by varying rate constants k^p_{+W} and k^p_{-F} (and not k^p_{+W} and k^p_{+F}) according to variations in shuttling rate constants, the introduction of a power stroke will change the kinetic asymmetry of the chemomechanical cycle (**Figure S5**). Unsurprisingly, the graphs produced for this scenario appear to display a combination of the features seen when power stroke magnitude (**Figure S4**) and kinetic asymmetry (**Figure S1** and **S2**) are altered independently. In the region with a positive power stroke and the associated larger kinetic asymmetry, the current and J/I_F increase. As with the fueling gating variation (**Figure S1**) J/I_F, the number of cycles per fuel molecule, approaches the theoretical cap of ±0.25 for a motor with an unbiased waste-forming reaction. This is because the heights of the waste-forming transition states do not change (and hence remain identical to each other within the model) so any change in the power stroke is compensated by a corresponding change in the rates of the waste-forming reactions. The maximum efficiency in the region with positive power strokes is the highest seen in any simulation, appearing in the region with slower absolute shuttling for same reasons as previously discussed (see Section 4.5.6.1., 4.5.6.2. and 4.5.6.4.),

indicating that this combined approach, introducing power strokes with concomitant increase in kinetic asymmetry, may be an efficient design for molecular motors.

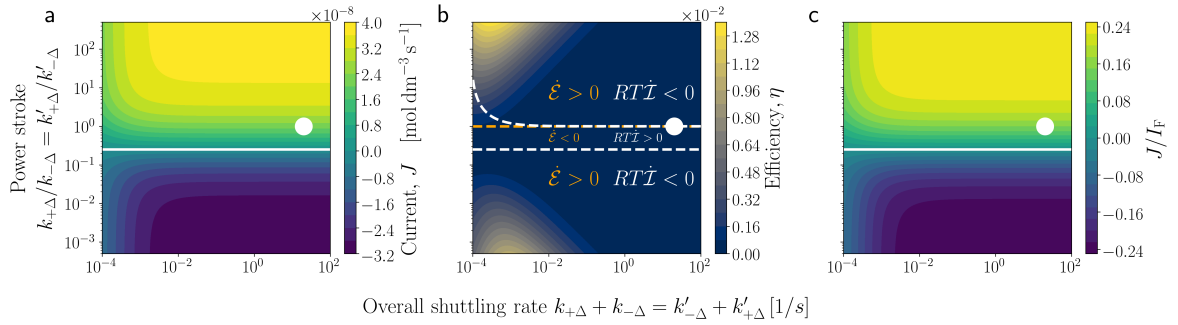


Figure S5. Graphs depicting the variation in current, efficiency, and J/I_F as overall shuttling rate (x-axis) and power stroke magnitude (y-axis) are changed. The total power stroke magnitude is not compensated by changes in fuel addition and waste removal transition rates (k^p_{+F} and k^p_{-W}), therefore allowing the kinetic asymmetry to change. Instead, the transition rates of reverse processes (waste addition and fuel removal, k^p_{+W} and k^p_{-F}), are altered to retain Wegscheider's conditions and thermodynamic validity (see Section 4.4.).

As with the simulations varying kinetic asymmetry by altering chemical gating of the fueling or waste-forming reactions, the direction of the rotary motor can be inverted with sufficiently negative power strokes. However, these graphs are not symmetrical and a lower maximum current and efficiency are seen in the negative current region as the chemical gating inherent to the system (*i.e.*, with no power strokes) favors a positive (clockwise) current.

4.5.6.6. Causing power strokes to cancel out

In this simulation (**Figure S6**), shuttling rate and power stroke magnitude are, again, altered. This time, however, the direction of one power stroke is always opposite to the other while they retain the same magnitude, cancelling out each other's contribution. Therefore, no energy flow is introduced across the full chemomechanical cycle and kinetic asymmetry remains inherently unaltered. This is most easily realized by switching one binding site for a more weakly or more strongly coordinating station. In the figures, the orange cross marks

the predicted position of a motor equivalent to the experimental motor but with one station that binds the macrocycle 2.7 kJ mol^{-1} more strongly than the other, equating to a 3:1 bias in occupancy. This motor is predicted to operate at about 80% of the current of the original design as it appears futile cycles would be promoted and ≈ 20% fewer cycles would be achieved, despite the same overall rate of fuel use. If the difference in binding strength is increased to 23 kJ mol^{-1}, the difference measured for the binding of the benzylic amide macrocycle used in the rotary motor to a fumaramide site compared to a succinamide site[60], the motor is effectively stalled, achieving no significant current, efficiency, or number of cycles per fuel molecule.

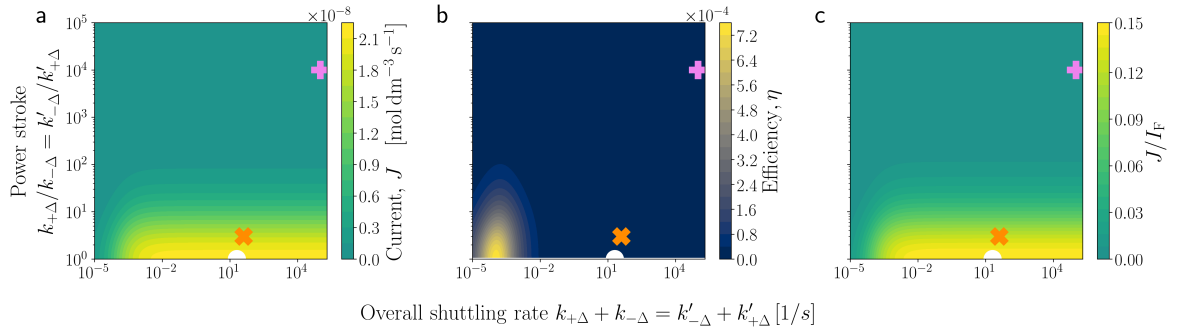


Figure S6. Graphs depicting the variation in current, efficiency, and J/I_F as overall shuttling rate (x-axis) and power stroke magnitude (y-axis) are changed. The power strokes are introduced so that they cancel out (*i.e.*, energy release by the first mechanical transition is always mirrored by an identical energy gain in the second) and lead to no net energy flow within the chemomechanical cycle. The pink plus and orange cross indicate the properties expected from changing one fumaramide group to a succinamide. The pink plus indicates the values obtained from single-molecule experiments[60], while the orange cross shows the values found in a related rotaxane-based system[58].

As before (see Section 4.5.6.1., 4.5.6.2., 4.5.6.4. and 4.5.6.5.), slow shuttling (left of the graphs) can be seen to stall the motor as futile cycles are favored, and once again, the area of maximum efficiency comes where this effect best balances with the degree to which the mechanical states are kept away from equilibrium in the steady state. Fast shuttling does not affect the current or the number of cycles per fuel molecule.

All graphs indicate that any change away from degenerate stations (moving up the y-axis) will always be detrimental for the motor in terms of current, efficiency, and the number of cycles per fuel molecule. This arises from a decreased capacity to support information flow in the motor as one station becomes more favored, which in turn decreases the ability of the motor to sustain a current. The full details and origins of this effect are not fully explored in this current work and warrant further experimental and theoretical investigation.

4.5.6.7. Conclusion

While the above by no-means represents an exhaustive analysis of all possible design feature alterations, even for the specific small-molecule motor analyzed in this paper, several important considerations have been identified. Furthermore, as future experimental systems become available to study, we envisage that similar analyses will help establish more general design considerations for molecular motors. These may become invaluable when putting molecular motors to work and may help clarify the driving forces behind the functioning of biological molecular motors.

4.6. References

1 Chatterjee, M. N., Kay, E. R. & Leigh, D. A. Beyond switches: ratcheting a particle energetically uphill with a compartmentalized molecular machine. *J. Am. Chem. Soc.* **128**, 4058–4073 (2006).

2 Coskun, A., Banaszak, M., Astumian, R. D., Stoddart, J. F. & Grzybowski, B. A. Great expectations: can artificial molecular machines deliver on their promise? *Chem. Soc. Rev.* **41**, 19–30 (2012).

3 Erbas-Cakmak, S., Leigh, D. A., McTernan, C. T. & Nussbaumer, A. L. Artificial molecular machines. *Chem. Rev.* **115**, 10081–10206 (2015).

4 Qiu, Y., Feng, Y., Guo, Q.-H., Astumian, R. D. & Stoddart, J. F. Pumps through the ages. *Chem* **6**, 1952–1977 (2020).

5 Aprahamian, I. The future of molecular machines. *ACS Cent. Sci.* **6**, 347–358 (2020).

6 Howard, J. *Mechanics of Motor Proteins and the Cytoskeleton* (Oxford University Press, Oxford, 2001).

7 Astumian, R. D., Mukherjee, S. & Warshel, A. The physics and physical chemistry of molecular machines. *ChemPhysChem* **17**, 1719–1741 (2016).

8 Koumura, N., Zijlstra, R. W. J., van Delden, R. A., Harada, N. & Feringa, B. L. Light driven monodirectional molecular rotor. *Nature* **401**, 152–155 (1999).

9 Serreli, V., Lee, C.-F., Kay, E. R. & Leigh, D. A. A molecular information ratchet. *Nature* **445**, 523–527 (2007).

10 Cheng, C. *et al.* An artificial molecular pump. *Nat. Nanotechnol.* **10**, 547–553 (2015).

11 Ragazzon, G., Baroncini, M., Silvi, S., Venturi, M. & Credi, A. Light-powered autonomous and directional molecular motion of a dissipative self-assembling system. *Nat. Nanotechnol.* **10**, 70–75 (2015).

12 Erbas-Cakmak, S. *et al.* Rotary and linear molecular motors driven by pulses of a chemical fuel. *Science* **358**, 340–343 (2017).

13 Amano, S., Fielden, S. D. P. & Leigh, D. A. A catalysis-driven artificial molecular pump. *Nature* **594**, 529–534 (2021).

14 Wilson, M. R. *et al.* An autonomous chemically fuelled small-molecule motor. *Nature* **534**, 235–240 (2016).

15 Astumian, R. D. How molecular motors work — insights from the molecular machinist's toolbox: the nobel prize in chemistry 2016. *Chem. Sci.* **8**, 840–845 (2017).

16 Astumian, R. D. Kinetic asymmetry allows macromolecular catalysts to drive an information ratchet. *Nat. Commun.* **10**, 3837 (2019).

17 Ragazzon, G. & Prins, L. J. Energy consumption in chemical fuel-driven self-assembly. *Nat. Nanotechnol.* **13**, 882–889 (2018).

18 Esposito, M. Open questions on nonequilibrium thermodynamics of chemical reaction networks. *Commun. Chem.* **3**, 107 (2020).

19 Kelly, T. R., De Silva, H. & Silva, R. A. Unidirectional rotary motion in a molecular system. *Nature* **401**, 150–152 (1999).

20 Mock, W. L. & Ochwat, K. J. Theory and example of a small-molecule motor. *J. Phys. Org. Chem.* **16**, 175–182 (2003).

21 Fletcher, S. P., Dumur, F., Pollard, M. M. & Feringa, B. L. A reversible, unidirectional molecular rotary motor driven by chemical energy. *Science* **310**, 80–82 (2005).

22 Collins, B. S. L., Kistemaker, J. C. M., Otten, E. & Feringa, B. L. A chemically powered unidirectional rotary molecular motor based on a palladium redox cycle. *Nat. Chem.* **8**, 860–866 (2016).

23 Zhang, Y. *et al.* A chemically driven rotary molecular motor based on reversible lactone formation with perfect unidirectionality. *Chem* **6**, 2420–2429 (2020).
24 Borsley, S., Leigh, D. A. & Roberts, B. M. W. A doubly kinetically-gated information ratchet autonomously driven by carbodiimide hydration. *J. Am. Chem. Soc.* **143**, 4414–4420 (2021).
25 Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* **75**, 126001 (2012).
26 Parrondo, J. M. R. & de Cisneros, B. J. Energetics of Brownian motors: a review. *Appl. Phys. A* **75**, 179–191 (2002).
27 Esposito, M., Lindenberg, K. & Van den Broeck, C. Universality of efficiency at maximum power. *Phys. Rev. Lett.* **102**, 130602 (2009).
28 Benenti, G., Casati, G., Saito, K. & Whitney, R. S. Fundamental aspects of steady-state conversion of heat to work at the nanoscale. *Phys. Rep.* **694**, 1–124 (2017).
29 Jarzynski, C. Equalities and inequalities: irreversibility and the second law of thermodynamics at the nanoscale. *Annu. Rev. Condens. Matter Phys.* **2**, 329–351 (2011).
30 Peliti, L. & Pigolotti, S. *Stochastic Thermodynamics: An Introduction* (Princeton University Press, Princeton, 2021).
31 Andrieux, D. & Gaspard, P. Fluctuation theorems and the nonequilibrium thermodynamics of molecular motors. *Phys. Rev. E* **74**, 011906 (2006).
32 Lipowsky, R. & Liepelt, S. Chemomechanical coupling of molecular motors: Thermodynamics, network representations, and balance conditions. *J. Stat. Phys.* **130**, 39–67 (2008).
33 Toyabe, S. *et al.* Nonequilibrium energetics of a single f_1-atpase molecule. *Phys. Rev. Lett.* **104**, 198103 (2010).
34 Ariga, T., Tomishige, M. & Mizuno, D. Nonequilibrium energetics of molecular motor kinesin. *Phys. Rev. Lett.* **121**, 218101 (2018).
35 Brown, A. I. & Sivak, D. A. Theory of nonequilibrium free energy transduction by molecular machines. *Chem. Rev.* **120**, 434–459 (2020).
36 Ciliberto, S. Experiments in stochastic thermodynamics: short history and perspectives. *Phys. Rev. X* **7**, 021051 (2017).
37 Cover, T. M. & Thomas, J. A. *Elements of Information Theory* (John Wiley & Sons, Inc., Hoboken, 2012).
38 Parrondo, J. M. R., Horowitz, J. M. & Sagawa, T. Thermodynamics of information. *Nat. Phys.* **11**, 131–139 (2015).
39 Horowitz, J. M. & Esposito, M. Thermodynamics with continuous information flow. *Phys. Rev. X* **4**, 031015 (2014).
40 Leff, H. S. & Rex, A. F. (eds.) *Maxwell's Demon: Entropy, Information, Computing* (Princeton University Press, Princeton, 1990).
41 Blackmond, D. G. "If pigs could fly" chemistry: a tutorial on the principle of microscopic reversibility. *Angew. Chem. Int. Ed.* **48**, 2648–2654 (2009).
42 Rao, R. & Esposito, M. Nonequilibrium thermodynamics of chemical reaction networks: wisdom from stochastic thermodynamics. *Phys. Rev. X* **6**, 041064 (2016).
43 Penocchio, E., Rao, R. & Esposito, M. Thermodynamic efficiency in dissipative chemistry. *Nat. Commun.* **10**, 3865 (2019).
44 Wagoner, J. A. & Dill, K. A. Mechanisms for achieving high speed and efficiency in biomolecular machines. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 5902–5907 (2019).
45 Howard, J. Protein power strokes. *Curr. Biol.* **16**, R517 (2006).

46 Hwang, W. & Karplus, M. Structural basis for power stroke vs. Brownian ratchet mechanisms of motor proteins. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 19777–19785 (2019).
47 Astumian, R. D. Irrelevance of the power stroke for the directionality, stopping force, and optimal efficiency of chemically driven molecular machines. *Biophys. J.* **108**, 291–303 (2015).
48 Pezzato, C., Cheng, C., Stoddart, J. F. & Astumian, R. D. Mastering the non-equilibrium assembly and operation of molecular machines. *Chem. Soc. Rev.* **46**, 5491–5507 (2017).
49 Das, K., Gabrielli, L. & Prins, L. J. Chemically-fueled self-assembly in biology and chemistry. *Angew. Chem. Int. Ed.* 10.1002/anie.202100274 (2021).
50 Albaugh, A. & Gingrich, T. R. Simulating a chemically-fueled molecular motor with nonequilibrium molecular dynamics. Preprint at http://arxiv.org/abs/2102.06298 (2021).
51 Horowitz, J. M., Sagawa, T. & Parrondo, J. M. R. Imitating chemical motors with optimal information motors. *Phys. Rev. Lett.* **111**, 010602 (2013).
52 Fyfe, M. C. T. *et al.* Anion-assisted self-assembly. *Angew. Chem. Int. Ed. Engl.* **36**, 2068–2070 (1997).
53 Astumian, R. D. Design principles for Brownian molecular machines: how to swim in molasses and walk in a hurricane. *Phys. Chem. Chem. Phys.* **9**, 5067–5083 (2007).
54 Hartich, D., Barato, A. C. & Seifert, U. Stochastic thermodynamics of bipartite systems: transfer entropy inequalities and a Maxwell's demon interpretation. *J. Stat. Mech.: Theory Exp.* **2014**, P02016 (2014).
55 Kondepudi, D. K. & Prigogine, I. *Modern Thermodynamics: From Heat Engines to Dissipative Structures* (John Wiley & Sons, Ltd., Chichester, 2015).
56 Astumian, R. & Bier, M. Mechanochemical coupling of the motion of molecular motors to ATP hydrolysis. *Biophys. J.* **70**, 637–653 (1996).
57 Hill, T. L. *Free energy transduction in biology* (Academic Press, New York, 1977).
58 Alvarez-Pérez, M., Goldup, S. M., Leigh, D. A. & Slawin, A. M. Z. A chemically-driven molecular information ratchet. *J. Am. Chem. Soc.* **130**, 1836–1838 (2008).
59 Carlone, A., Goldup, S. M., Lebrasseur, N., Leigh, D. A. & Wilson, A. A three-compartment chemically-driven molecular information ratchet. *J. Am. Chem. Soc.* **134**, 8321–8323 (2012).
60 Lussis, P. *et al.* A single synthetic small molecule that generates force against a load. *Nat. Nanotechnol.* **6**, 553–557 (2011).
61 Amano, S., Borsley, S., Leigh, D. A. & Sun, Z. Chemical engines: driving systems away from equilibrium through catalyst reaction cycles. *Nat. Nanotechnol.* **16**, 1057–1067 (2021).
62 Kay, E. R., Leigh, D. A. & Zerbetto, F. Synthetic molecular motors and mechanical machines. *Angew. Chem. Int. Ed.* **46**, 72–191 (2007).
63 Katz, S., Lebowitz, J. L. & Spohn, H. Phase transitions in stationary nonequilibrium states of model lattice systems. *Phys. Rev. B* **28**, 1655–1658 (1983).
64 Esposito, M. Stochastic thermodynamics under coarse graining. *Phys. Rev. E* **85**, 041125 (2012).
65 Maes, C. Local detailed balance. Preprint at https://arxiv.org/abs/2011.09200 (2021).
66 Laidler, K. J. *Chemical Kinetics* (HarperCollins Publishers, New York, 1987).
67 Wachtel, A., Rao, R. & Esposito, M. Thermodynamically consistent coarse graining of biocatalysts beyond Michaelis–Menten. *New J. Phys.* **20**, 042002 (2018).
68 Avanzini, F., Falasco, G. & Esposito, M. Thermodynamics of non-elementary chemical reaction networks. *New J. Phys.* **22**, 093040 (2020).

69 Feinberg, M. Complex balancing in general kinetic systems. *Arch. Ration. Mech. Anal.* **49**, 187–194 (1972).
70 Schnakenberg, J. Network theory of microscopic and macroscopic behavior of master equation systems. *Rev. Mod. Phys.* **48**, 571–585 (1976).
71 Mou, C. Y., Luo, J. & Nicolis, G. Stochastic thermodynamics of nonequilibrium steady states in chemical reaction systems. *J. Chem. Phys.* **84**, 7011–7017 (1986).
72 Lebowitz, J. L. & Bergmann, P. New approach to nonequilibrium process. *Phys. Rev.* **99**, 578–587 (1955).
73 Feng, Y. *et al.* Molecular pumps and motors. *J. Am. Chem. Soc.* **143**, 5569–5591 (2021).
74 Falasco, G. & Esposito, M. Local detailed balance across scales: from diffusions to jump processes and beyond. *Phys. Rev. E* **103**, 042114 (2021).
75 Gaspard, P. Comment on "validity of path thermodynamics in reactive systems". *Phys. Rev. E* **103**, 016101 (2021).
76 Geertsema, E. M., van der Molen, S. J., Martens, M. & Feringa, B. L. Optimizing rotary processes in synthetic molecular motors. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 16919–16924 (2009).
77 Leigh, D. A., Wong, J. K. Y., Dehez, F. & Zerbetto, F. Unidirectional rotation in a mechanically interlocked molecular rotor. *Nature* **424**, 174–179 (2003).
78 Shirono, K., Morimatsu, T. & Takemura, F. Gas solubilities (CO_2, O_2, Ar, N_2, H_2, and He) in liquid chlorinated methanes. *J. Chem. Eng. Data* **53**, 1867–1871 (2008).

Conclusion and Outlook

In this thesis, the author presented three projects that address fundamental challenges that need to be overcome to develop molecular machines with more sophisticated, life-like functions and to understand their operating mechanisms deeply. The first project, a unidirectional molecular transporter utilising multiple hydrazone-type switches (Chapter 2), attempted coordination of operation of two hydrazone-type switches to transport a cargo molecule, thus paving a way to synthetic molecular machines that achieve complex and sophisticated tasks by the coordination of multiple units. The second project, a catalysis-driven artificial molecular pump (Chapter 3), achieved a chemically driven autonomous molecular machine that operates *via* an information ratchet mechanism[1]. The fact that most of biological molecular machines operate *via* this mechanism implies that expanding the repertoire of this type of synthetic molecular machines will be crucial for future nanotechnology with life-like properties. The third project, analysis of a synthetic molecular motor with information thermodynamics (Chapter 4), offered thermodynamic level of understanding of an information ratchet[2], complementing previous analyses based on kinetics[3,4]. This analysis revealed how the free energy supplied from the chemical fuel is converted and utilised in molecular machines, thus enabling definition of efficiency of each process in molecular machines. Additionally, the analysis and numerical simulation based on it provided practical implications for designing and improving molecular machines. Overall, these three projects are effectively addressing the challenges in the field, leading to significant advancements.

The future development of the field of synthetic molecular machines will be built upon the synergy of creation and understanding as attempted in the above projects: creation of new synthetic systems and understanding of them (and their biological counterparts) from a theoretical perspective. Such cooperation will be especially important for developing

autonomous information ratchets, because they operate in out-of-equilibrium conditions and their design principles are different from molecular machines that operate by switching between multiple equilibrium states ('switch'-type molecular machines and energy ratchets). The works presented in Chapter 3 and 4 will be cornerstones for such synergetic development. Through this effort, it is expected that properties of autonomous molecular machines such as velocity, efficiency in free energy transduction and efficiency in usage of fuel molecules will be improved. Likewise, variety of chemical fuels and wavelength of light that power molecular machines, and the type of motility that molecular machines can achieve (*e.g.*, rotation around single/double bonds, translational motion, walking motion along the track, transmembrane transportation) will be expanded.

In addition to such efforts in fundamental aspects, another important direction of future development will be the application of synthetic molecular machines. There are already numerous examples of applying non-autonomous molecular machines to perform some tasks[5], and some examples of application of autonomous light-driven molecular machines, such as contraction of gel[6], breaking cell membranes[7] and triggering cell responses[8]. To date, there are no examples of the application of chemically driven synthetic autonomous molecular machines, although this is unsurprising given the time lag between the first example of a light-driven synthetic autonomous molecular machine[9] and the first example of a chemically driven synthetic autonomous molecular machine[10]. Considering diverse usage of chemically driven autonomous molecular machines in biology, it is expected that their synthetic counterparts will be utilised for various purposes such as synthesis, soft robotics and health care in future nanotechnology.

From a broader perspective, synthetic molecular machines can be regarded as an attempt to mimic sophisticated functions of biology. There are also other fields that aim at similar goals, such as the study of self-replicators[11], dissipative self-assembly[12] and systems with

feedback mechanisms[13]. These systems all operate in out-of-equilibrium conditions, as biology does, and have some elements in common. Indeed, a recent Perspective by the author and co-workers[14] (see Appendix, third article) have pointed out these common elements, such as catalytic cycles as a fueling reaction cycle, coupling between catalytic cycles and other dynamic processes (*e.g.*, mechanical processes, self-assembly processes), kinetic selection and synchronisation (which enables oscillation). Transferring design elements and theoretical understanding in one field to another will accelerate development of these fields, and help address challenges and understanding of various fundamental processes apparently intrinsic to life.

1 Amano, S., Fielden, S. D. P. & Leigh, D. A. A catalysis-driven artificial molecular pump. *Nature* **594**, 529–534 (2021).

2 Amano, S. *et al.* Insights from an information thermodynamics analysis of a synthetic molecular motor. *ChemRxiv*, doi:10.33774/chemrxiv-2021-60k1r (2021).

3 Astumian, R. D. Irrelevance of the power stroke for the directionality, stopping force, and optimal efficiency of chemically driven molecular machines. *Biophys. J.* **108**, 291–303 (2015).

4 Astumian, R. D. Kinetic asymmetry allows macromolecular catalysts to drive an information ratchet. *Nat. Commun.* **10**, 3837 (2019).

5 Erbas-Cakmak, S., Leigh, D. A., McTernan, C. T. & Nussbaumer, A. L. Artificial molecular machines. *Chem. Rev.* **115**, 10081–10206 (2015).

6 Li, Q. *et al.* Macroscopic contraction of a gel induced by the integrated motion of light-driven molecular motors. *Nat. Nanotechnol.* **10**, 161–165 (2015).

7 García-López, V. *et al.* Molecular machines open cell membranes. *Nature* **548**, 567 (2017).

8 Zheng, Y. *et al.* Optoregulated force application to cellular receptors using molecular motors. *Nat. Commun.* **12**, 3580 (2021).

9 Koumura, N., Zijlstra, R. W., van Delden, R. A., Harada, N. & Feringa, B. L. Light-driven monodirectional molecular rotor. *Nature* **401**, 152–155 (1999).

10 Wilson, M. R. *et al.* An autonomous chemically fuelled small-molecule motor. *Nature* **534**, 235–240 (2016).

11 Bissette, A. J. & Fletcher, S. P. Mechanisms of autocatalysis. *Angew. Chem. Int. Ed.* **52**, 12800–12826 (2013).

12 van Rossum, S. A. P., Tena-Solsona, M., van Esch, J. H., Eelkema, R. & Boekhoven, J. Dissipative out-of-equilibrium assembly of man-made supramolecular materials. *Chem. Soc. Rev.* **46**, 5519–5535 (2017).

13 Epstein, I. R. & Xu, B. Reaction-diffusion processes at the nano- and microscales. *Nat. Nanotechnol.* **11**, 312–319 (2016).

14 Amano, S., Borsley, S., Leigh, D. A. & Sun, Z. Chemical engines: driving systems away from equilibrium through catalyst reaction cycles. *Nat. Nanotechnol.* **16**, 1057–1067 (2021).

Appendix

Articles published in a peer-reviewed journal or submitted to a journal for publication:

Amano, S., Fielden, S. D. P. & Leigh, D. A. A catalysis-driven artificial molecular pump. *Nature* **594**, 529–534 (2021).

Amano, S., Esposito, M., Kreidt, E., Leigh, D. A., Penocchio, E. & Roberts, B. M. W. Insights from an information thermodynamics analysis of a synthetic molecular motor. *ChemRxiv* doi: 10.33774/chemrxiv-2021-60k1r **(**2021).
(Submitted to a peer-reviewed journal as well.)

Amano, S., Borsley, S., Leigh, D. A. & Sun, Z. Chemical engines: Driving systems away from equilibrium through catalyst reaction cycles. *Nat. Nanotechnol.* **16**, 1057–1067 (2021).

Article

A catalysis-driven artificial molecular pump

https://doi.org/10.1038/s41586-021-03575-3

Received: 23 February 2021

Accepted: 22 April 2021

Published online: 23 June 2021

Check for updates

Shuntaro Amano[1], Stephen D. P. Fielden[1] & David A. Leigh[1,2]✉

All biological pumps are autonomous catalysts; they maintain the out-of-equilibrium conditions of the cell by harnessing the energy released from their catalytic decomposition of a chemical fuel[1–3]. A number of artificial molecular pumps have been reported to date[4], but they are all either fuelled by light[5–10] or require repetitive sequential additions of reagents or varying of an electric potential during each cycle to operate[11–16]. Here we describe an autonomous chemically fuelled information ratchet[17–20] that in the presence of fuel continuously pumps crown ether macrocycles from bulk solution onto a molecular axle without the need for further intervention. The mechanism uses the position of a crown ether on an axle both to promote barrier attachment behind it upon threading and to suppress subsequent barrier removal until the ring has migrated to a catchment region. Tuning the dynamics of both processes[20,21] enables the molecular machine[22–25] to pump macrocycles continuously from their lowest energy state in bulk solution to a higher energy state on the axle. The ratchet action is experimentally demonstrated by the progressive pumping of up to three macrocycles onto the axle from bulk solution under conditions where barrier formation and removal occur continuously. The out-of-equilibrium [*n*]rotaxanes (characterized with *n* up to 4) are maintained for as long as unreacted fuel is present, after which the rings slowly de-thread. The use of catalysis to drive artificial molecular pumps opens up new opportunities, insights and research directions at the interface of catalysis and molecular machinery.

The structure and mode of operation of the catalysis-driven molecular pump, **1**, is shown in Fig. 1. One end of the axle is permanently blocked by a bulky triarylmethine group (brown); the other end of the axle (the N-terminus) is open for threading when in the form of a benzyl amine group (orange). The axle includes a chain of triazole heterocycles linked by short, propyl (that is, $-(CH_2)_3-$), spacers. Triazoles interact only weakly with crown ethers[26,27], so the proportion of crown ethers transferred from bulk solution through the open end of the axle and onto the chain is vanishingly small at equilibrium. The pumping mechanism is based on a recently discovered[28] metal-free[29,30] active template rotaxane-forming reaction[31], in which addition of a primary amine to an electrophile is accelerated through the cavity of a crown ether. Pump **1** reacts with Fmoc-$OC_6H_4NO_2$ (the fuel[32], **2**; where Fmoc = 9-fluorenylmethyloxycarbonyl) and 24-crown-8 (a crown ether, **3**; blue) causing active template threading of the ring, initially forming *car*-**4** (Fig. 1; the italicized prefix indicates the position of the ring on the axle). The newly formed Fmoc-carbamate group (green) prevents dethreading of the ring back over the axle N-terminus. Adjacent to the benzyl amine group is a trifluoromethyl (CF_3) group (red), which is of sufficient size to substantially slow the passage of 24-crown-8 (a 'speed bump'[9]). The triazole groups on the axle should form more favourable interactions with the crown ether[26,27] than the sterically crowded benzylic carbamate in *car*-**4**, and so provide a 'catchment region' for rings threaded onto the axle. Accordingly, over time the crown ether in *car*-**4** slowly passes over the trifluoromethyl group to the catchment region, forming co-conformer *tri*-**4** ('co-conformers' are structures that differ in the relative positions of their components[33]). Rotaxane *tri*-**4** is of lower energy than *car*-**4**, but still of higher energy than the unthreaded components **1** and **3**. The presence of a base (diisopropylamine (*i*-Pr_2NH) in Fig. 1) is necessary to complete the pumping cycle. The base deprotonates the Fmoc-carbamate group, triggering its fragmentation to dibenzofulvene (**5**), CO_2 and 4-nitrophenol ($NO_2C_6H_4OH$; **6**), liberating the amine N-terminus of the molecular machine for use in further pumping cycles. In this way, rings are captured from bulk solution, constrained within a localized region of the axle, and maintained in an out-of-equilibrium state by the continuous, autonomous, consumption of the chemical fuel.

The pump and fuel structure, in particularly the stereoelectronic characteristics of the CF_3 speed bump, the triazole groups of the catchment region, and the 4-nitrophenol leaving group of the fuel, were selected through a series of model and optimization studies (Supplementary Information). The pump operating conditions (see Fig. 1) were chosen so that the Fmoc-removal reaction and the passage of the rings from the catchment region of the axle back over the CF_3 group (resulting in dethreading when the pump N-terminus is not blocked) are both much slower than active template reaction of the axle benzylic amine group with unreacted Fmoc-$OC_6H_4NO_2$ fuel. This means that the N-terminus of **1•3** undergoes further active template reactions before dethreading occurs, pumping a second (and subsequently third) ring onto the axle. The molecular machine continues to pump rings onto the axle for as long as there is sufficient space on the chain and unreacted fuel present.

Pump **1** was synthesized in 13 steps (see Supplementary Information) and operated successfully in three distinct ways: (i) stepwise, adding

[1]Department of Chemistry, University of Manchester, Manchester, UK. [2]School of Chemistry and Molecular Engineering, East China Normal University, Shanghai, China.
✉e-mail: david.leigh@manchester.ac.uk

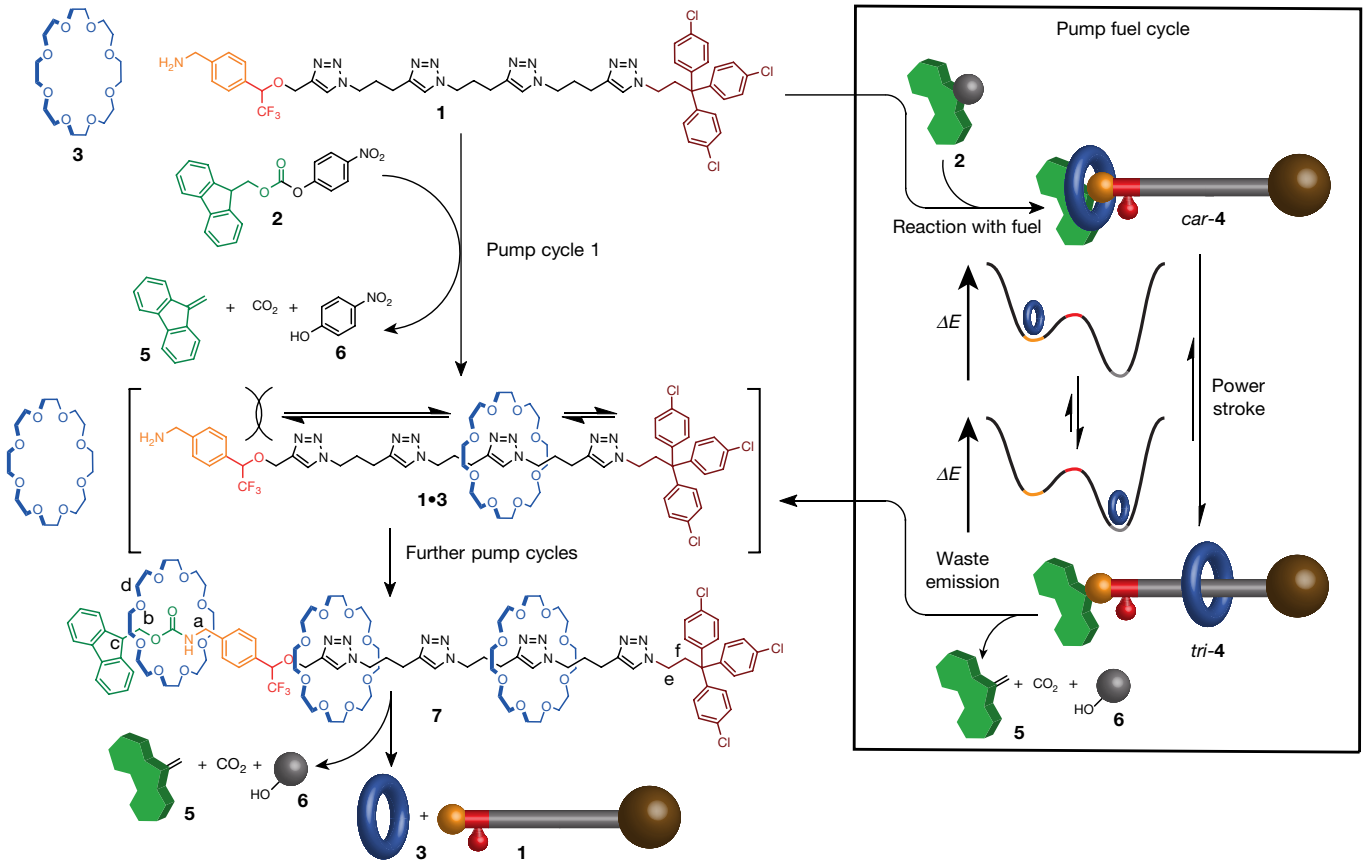


Fig. 1 | Structure and operation of a catalysis-driven artificial molecular pump. Continuous pumping of 24-crown-8 rings (**3**) onto a molecular axle (**1**) in the presence of fuel (**2**), causing the rings to be dissipatively captured in the form of [*n*]rotaxanes such as **7**. Typical operating conditions are **1** (1 equiv.), **3** (10 equiv.), *i*-Pr_2NH (15 equiv.), toluene (0.2 M with respect to **1**), **2** (either 10+ equiv. added at time zero, or added continuously via syringe pump at a rate of 0.7 equiv. per hour), at room temperature. The inset shows the pump fuel cycle. Kinetically gated addition of the fuel to the pump causes active template threading of the ring onto the axle, initially forming rotaxane co-conformer *car*-**4** (with the macrocycle positioned over the carbamate axle site). The rotaxane relaxes to the lower energy state, *tri*-**4**, through a power stroke (although in the formation of [4]rotaxane **7** from **1**•$(\mathbf{3})_2$ the equivalent step is not exergonic) of the ring past the trifluoromethyl 'speed bump' (red) onto the catchment region (grey part of the axle in the cartoon). Kinetically gated decomposition of the blocking group (green) to waste liberates the open end of the axle for further pumping while threaded rings remain kinetically trapped in the catchment region, completing the pumping cycle. equiv., equivalents.

the fuel and base in separate steps, which facilitated isolation of the rotaxanes; (ii) autonomously, with all of the fuel added to the rings, pump and base at time zero; (iii) continuous fuel addition, which minimizes background decomposition of the fuel by the base. The pumping mechanism was confirmed by operating **1** in a stepwise manner. A single pumping cycle, other than removal of the Fmoc barrier, was performed by combining **1**, **2** and **3** in toluene under basic conditions (Fig. 1), affording [2]rotaxane **4** in 61% yield after 10 min (the amount of rotaxane did not increase using longer reaction times), with concurrent formation of the Fmoc-derivatized unthreaded axle in 10% yield. This shows that >80% of the fuel molecules (**2**) that react with **1** result in threading. The formation of [2]rotaxane under conditions where the release of the blocking group (waste-forming step) is negligible shows that during the fuel addition step of the pumping mechanism, the pumping is kinetically gated (that is, threading accelerates Fmoc-carbamate formation of the axle).

The interconversion of *car*-**4** and *tri*-**4** is slow on the ^{1}H NMR (proton nuclear magnetic resonance spectroscopy) timescale at 295 K in $C_6D_5CD_3$ (Fig. 2c), allowing the ratio of the [2]rotaxane co-conformers to be determined. Various axle protons are deshielded by the presence of the macrocycle: in particular, H_a, H_b and H_c of *car*-**4** (labelling shown in Fig. 2a), and $H_{e'}$ and $H_{f'}$ of *tri*-**4**. Signal integration gives a *car*-**4**:*tri*-**4** ratio (that is, local equilibrium distribution of a single crown ether between *car*- and *tri*- sites) of 3:7 at 295 K in $C_6D_5CD_3$.

To complete the pump cycle, the Fmoc group needs to be removed from the N-terminus of **4** to regenerate the primary amine for another threading event. The position of the macrocycle on the axle affects the rate of Fmoc removal. When treated with *i*-Pr_2NH in toluene, the half-life of a model rotaxane, **8**, which has 24-crown-8 constrained on a short axle and forced into close proximity to the Fmoc group, was found to be around three times longer than that of the corresponding non-interlocked thread **9** (Fig. 3a). During the operation of pump **1**, the ring's steric inhibition of the attack of *i*-Pr_2NH means that the Fmoc group is removed more rapidly from the N-terminus in co-conformers where all the threaded macrocycles are in the catchment region, that is, *tri*-**4** and *tri-tri*-**10** react faster than *car*-**4** and *car-tri*-**10**, suppressing dethreading. This means that the waste removal step (the Fmoc-carbamate-to-waste-products reaction) is also kinetically gated in the pumping mechanism of **1** (that is, deblocking of the N-terminus occurs more rapidly when all of the threaded macrocycles are located in the catchment region).

Fmoc removal from **4** with the strong base 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) was complete within 30 s, allowing the half-life of pseudorotaxane **1•3** to be measured as 230 min in $C_6D_5CD_3$ (Fig. 3b); the rate-determining step is passage of the ring over the CF_3 group from the catchment region to the N-terminus. As ring-dethreading from **1•3** is much slower than the 10 min required to complete an active template threading-and-stoppering reaction, in the presence of excess fuel **1•3**

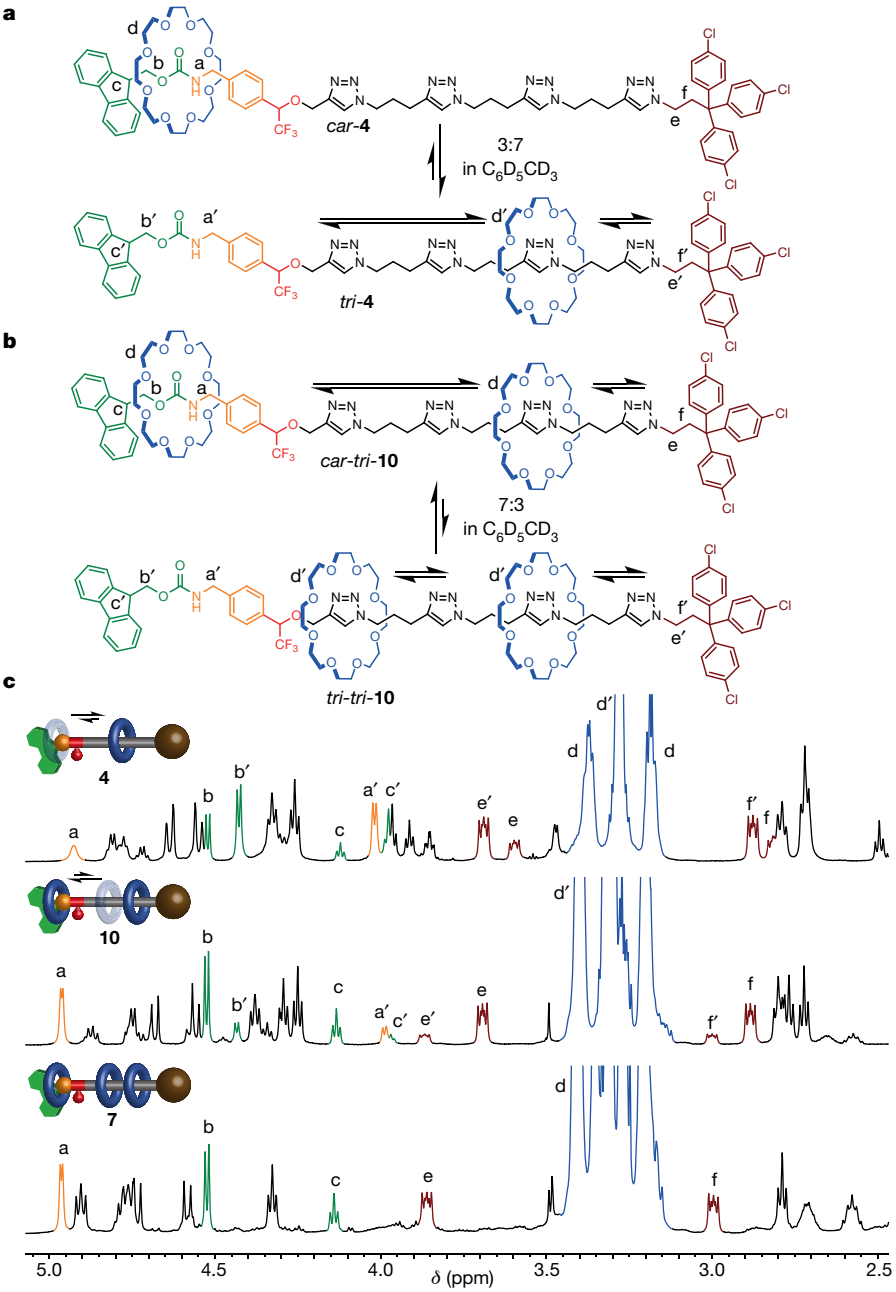


Fig. 2 | Macrocycle distribution in [*n*]rotaxane co-conformers. **a**, Macrocycle positions in [2]rotaxane co-conformers *car*-**4** and *tri*-**4** (295 K, $C_6D_5CD_3$). **b**, Macrocycle positions in [3]rotaxane co-conformers *car-tri*-**10** and *tri-tri*-**10** (295 K, $C_6D_5CD_3$). **c**, Partial ^{1}H NMR (600 MHz, 295 K, $C_6D_5CD_3$) spectra of **4**, **10** and [4]rotaxane **7**. Lettering relates protons in the chemical structures to the corresponding signals in the ^{1}H NMR spectra. Primes are used to distinguish the protons in *tri*-**4** and *tri-tri*-**10** from those in co-conformers *car*-**4** and *car-tri*-**10**.

collects additional rings through a continuous pumping process to give first [3]rotaxane **10** and then [4]rotaxane **7**. Further iterations of the pumping cycle did not yield any [5]rotaxane, suggesting that the axle is already full in the [4]rotaxane.

For the rotaxane with two rings pumped onto the axle (**10**; Fig. 2b), two co-conformers are observed by ^{1}H NMR: *car-tri*-**10**, with one macrocycle residing over the carbamate group and the other in the catchment region, and *tri-tri*-**10** with both macrocycles in the catchment region. The ratio of *car-tri*-**10** to *tri-tri*-**10** is 7:3 at 295 K in $C_6D_5CD_3$. In contrast, only one co-conformer is observed for the [4]rotaxane, *car-tri-tri*-**7** isomer (Fig. 2c), which is consistent with the results of the pumping studies done in this work that suggest that the axle of the [4]rotaxane is only long enough to accommodate three macrocycles. Extending the catchment region to include more triazole rings (or other groups that weakly stabilize the crown ether) is complicated by the need for the pump to be soluble in toluene, which is the optimal solvent[28–30] for the active template reaction.

Considered together, these experimental results provide insight into the role and effect of the pump's power stroke. The ^{1}H NMR spectra (Fig. 2c) show that although the threaded ring is at lower energy in the catchment region than in the carbamate region in the [2]rotaxane (*car*-**4**:*tri*-**4** 3:7; 295 K, $C_6D_5CD_3$), the [3]rotaxane is actually more thermodynamically stable when the additional ring is over the carbamate group rather than in the catchment region (*car-tri*-**10**:*tri-tri*-**10** 7:3; 295 K, $C_6D_5CD_3$). Steric considerations mean that [4]rotaxane can only be produced from active template threading of the *tri-tri*- co-conformer of **1**•(**3**)$_2$ (that is, both rings in the catchment region of the open-ended pump). Therefore, pumping of the [3]rotaxane (which is a 7:3 mixture of

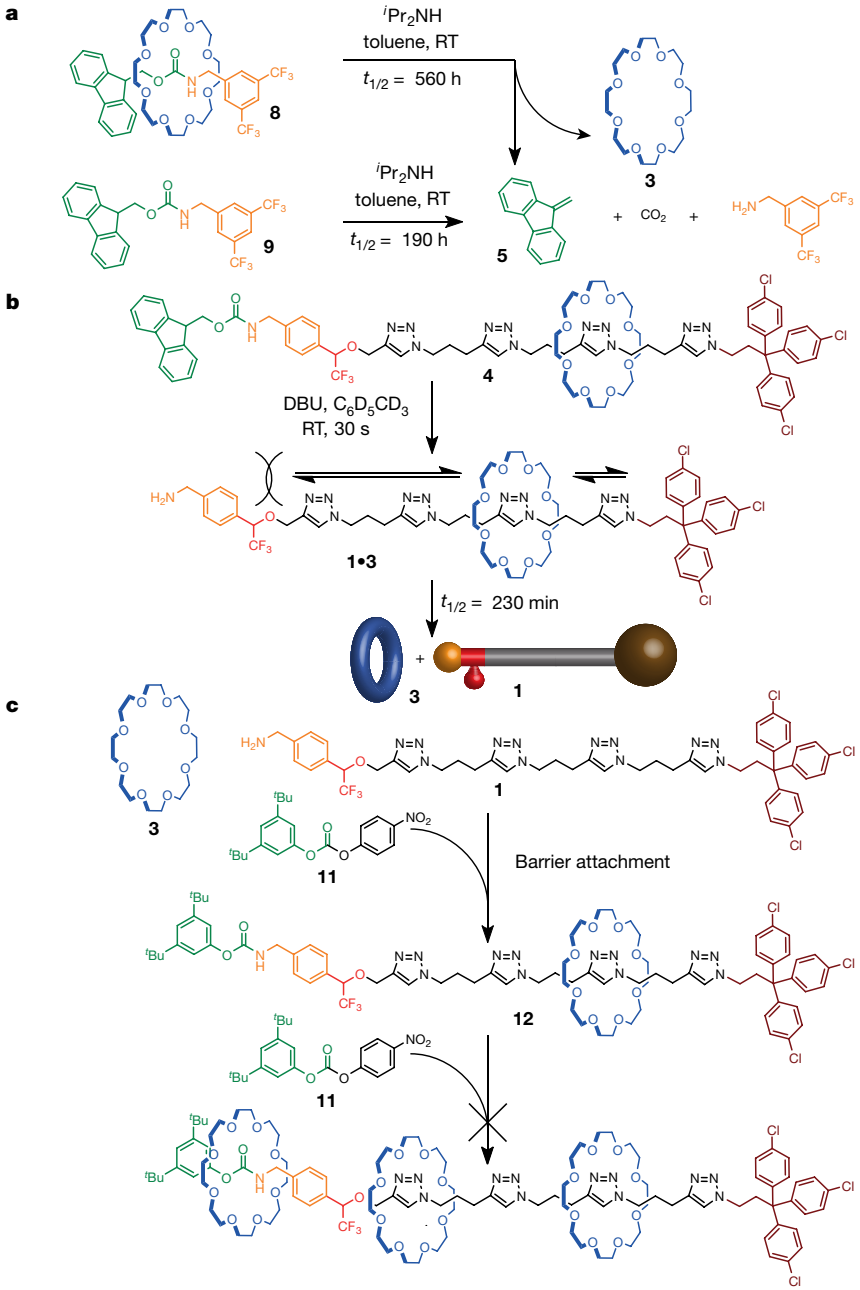


Fig. 3 | Fmoc removal, pseudorotaxane dethreading, and irreversible rotaxane formation experiments. a, Reagents and conditions for Fmoc removal from model [2]rotaxane **8** and model non-interlocked thread **9**: **8** or **9** (1 equiv.) *i*-Pr_2NH (10 equiv.), toluene (0.14 M), at room temperature (RT). **b**, Reagents and conditions for Fmoc removal and subsequent dethreading: rotaxane **4** (1 equiv.), DBU (50 equiv.), $C_6D_5CD_3$ (21 mM), room temperature, 30 s, then diluted to 4.2 mM with $C_6D_5CD_3$ and monitored by ^{1}H NMR. **c**, Reagents and conditions for formation of [2]rotaxane **12** using carbonate **11**: pump **1** (1 equiv.), 24-crown-8 **3** (10 equiv.), *i*-Pr_2NH (15 equiv.), toluene (0.2 M), reagent **11** (added continuously via syringe pump as a solution in toluene at a rate of 0.4 equiv. per hour), at room temperature, for 16 h. DBU, 1,8-diazabicyclo[5.4.0]undec-7-ene.

car-tri-**10**:*tri-tri*-**10**) to the [4]rotaxane (**7**) can be viewed as proceeding as follows: from (i) *car-tri*-**10** (the lowest-energy co-conformer of the [3]rotaxane) to (ii) *tri-tri*-**10** (a state that undergoes fast deblocking of the pump terminus) to (iii) **1**•(**3**)$_2$ (open-end pump available for active template threading, while the speed bump inhibits dethreading of already captured rings) to (iv) *car-tri-tri*-**7** (the [4]rotaxane); a sequence that actually involves an energetically uphill co-conformational change (*car-tri*-**10** to *tri-tri*-**10**)—a necessary, but endergonic, analogue of the power stroke that occurs to form the [2]- and [3]rotaxane.

Control experiments confirmed that there is no net displacement of 24-crown-8 rings from bulk solution onto **1** in the absence of fuel. At equilibrium, the amount of **1**•**3** present in toluene is lower than the limits detectable by spectroscopy (see Supplementary Information). Furthermore, when an excess of an alternative carbonate, **11**, was used instead of fuel **2**, resulting in a N-terminus carbamate that is not decomposed by base, [2]rotaxane **12** was the only threaded product isolated (Fig. 3c). This demonstrates that only one ring is pumped onto the axle by the consumption of each fuel molecule.

The pump can also operate if excess fuel and *i*-Pr_2NH are present from the start, allowing multiple fuel cycles to occur autonomously to generate [*n*]rotaxanes until all of the fuel is consumed. However, in the presence of *i*-Pr_2NH fuel **2** decomposes at a similar rate to the Fmoc-carbamate axle, so it is difficult to solubilize sufficient fuel at time zero to generate more than very small amounts of [4]rotaxane.

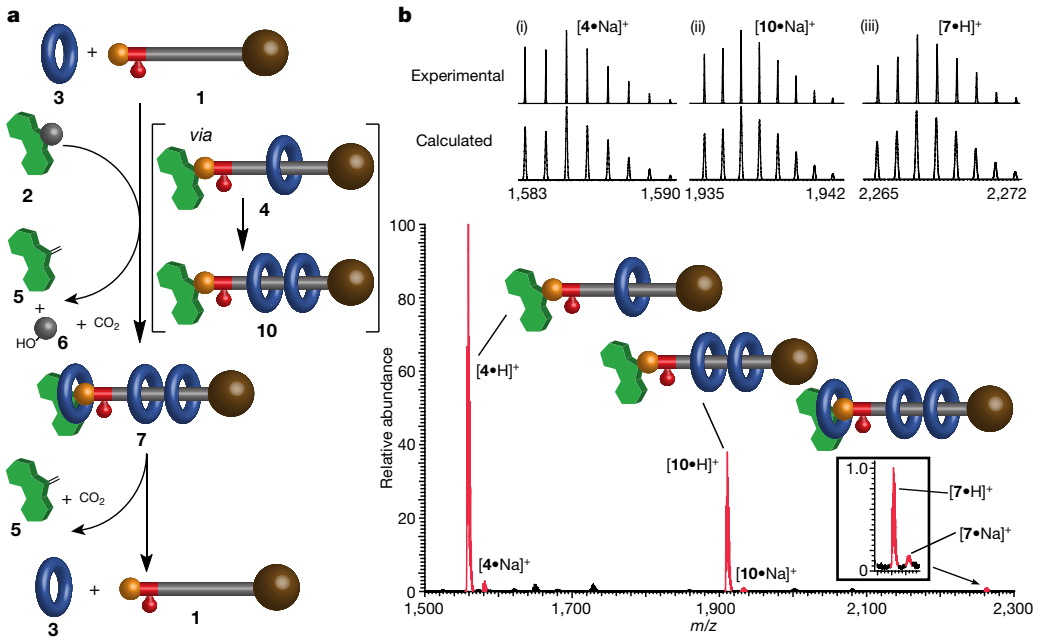


Fig. 4 | Out-of-equilibrium state produced by the operation of pump 1. **a**, Reagents and conditions: pump **1** (1 equiv.), 24-crown-8 **3** (10 equiv.), *i*-Pr_2NH (15 equiv.), toluene (0.2 M), fuel **2** (added continuously via syringe pump as a solution in toluene at a rate of 0.7 equiv. per hour), at room temperature, for 16 h. After stopping the addition of fuel **2**, half of the reaction mixture was left at room temperature for 19 days, regenerating the unthreaded components, **1** and **3**. **b**, Electrospray ionization mass spectrum of the crude reaction mixture worked up just after stopping the addition of fuel **2**. Signal intensity does not correspond to the amount of product formed. The signal corresponding to [4]rotaxane **7** is shown at ×30 intensity. Isotopic patterns shown for [**4**•Na]$^+$, [**10**•Na]$^+$ and [**7**•H]$^+$. The cartoons indicate the number of rings in each rotaxane.

Accordingly, pump operation is more efficient with fuel continuously drip-fed to the pump rather than delivered in a single pulse. Treatment of pump **1** with 10 equivalents of 24-crown-8 (**3**), 15 equivalents of *i*-Pr_2NH and 0.7 equivalents fuel (**2**) per hour for 16 h generated [4]rotaxane **7**, along with [2]rotaxane **4** and [3]rotaxane **10** (Fig. 4b). The threaded products are formed dissipatively[34–37] under these operating conditions: Fmoc-carbamate decomposition continued for a further 19 days after ceasing the addition of fuel **2**, causing **3** to be slowly released back into solution resulting in virtually all of the original amount of unthreaded axle and rings, **1** and **3**, being recovered (see Supplementary Information).

Synthetic chemically fuelled molecular pump **1** drives macrocycles from bulk solution into an organized, high-energy, threaded assembly, and dissipatively maintains the out-of-equilibrium state for as long as unreacted chemical fuel remains. The pump uses a doubly kinetically gated catalysis-driven information ratchet mechanism that is wholly artificial and minimalist in design[38], yet fundamentally analogous to that of biomolecular pumps. It is kinetic asymmetry in the act of threading under catalysis that drives the pumping away from equilibrium[20]. The mechanism of pumping features a step that, for the first two rings, has the characteristics of a power stroke—'a viscoelastic, free-energy-releasing, large-amplitude conformational change'[39,40]—as the initially captured, high-energy crown ethers close to the carbamate group (for example, *car*-**4**) relax to the catchment region (for example, *tri*-**4**). However, the analogous co-conformational change required to form the [4]rotaxane is an energetically uphill process (an 'anti-power stroke'). There is ongoing debate concerning the role of power strokes in catalysis-driven biomolecular machines[39,41]. In this catalysis-driven small-molecule pump the change of position of the ring(s) on the axle from the carbamate to the catchment region serves to retard the rate of release of the captured substrate. This co-conformational change is necessary for the operating mechanism irrespective of whether it is exergonic (for example, during pumping of the first two rings) or not (as is required to pump on the third ring). The mechanism of **1** is general (not restricted to particular substrates, functional groups or catalysts), and we anticipate that catalysis will play an increasingly important part in the design of artificial molecular machines[20,42–48] and, thereby[21], aid in the understanding of their biomolecular counterparts.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-03575-3.

1. Skou, J. C. The identification of the sodium–potassium pump (Nobel Lecture). *Angew. Chem. Int. Ed.* **37**, 2320–2328 (1998).
2. Lodish, H. et al. Transport across cell membranes. In *Molecular Cell Biology* Vol. 4, Ch. 15 (W. H. Freeman, 2000).
3. Du, D. et al. Multidrug efflux pumps: structure, function and regulation. *Nat. Rev. Microbiol.* **16**, 523–539 (2018).
4. Qiu, Y., Feng, Y., Guo, Q.-H., Astumian, R. D. & Stoddart, J. F. Pumps through the ages. *Chem* **6**, 1952–1977 (2020).
5. Steinberg-Yfrach, G. et al. Conversion of light energy to proton potential in liposomes by artificial photosynthetic reaction centres. *Nature* **385**, 239–241 (1997).
6. Bennett, I. M. et al. Active transport of Ca^{2+} by an artificial photosynthetic membrane. *Nature* **420**, 398–401 (2002).
7. Bhosale, S. et al. Photoproduction of proton gradients with π-stacked fluorophore scaffolds in lipid bilayers. *Science* **313**, 84–86 (2006).
8. Serreli, V., Lee, C.-F., Kay, E. R. & Leigh, D. A. A molecular information ratchet. *Nature* **445**, 523–527 (2007).
9. Baroncini, M., Silvi, S., Venturi, M. & Credi, A. Photoactivated directionally controlled transit of a non-symmetric molecular axle through a macrocycle. *Angew. Chem. Int. Ed.* **51**, 4223–4226 (2012).
10. Ragazzon, G., Baroncini, M., Silvi, S., Venturi, M. & Credi, A. Light-powered autonomous and directional molecular motion of a dissipative self-assembling system. *Nat. Nanotechnol.* **10**, 70–75 (2015).
11. Cheng, C. et al. An artificial molecular pump. *Nat. Nanotechnol.* **10**, 547–553 (2015).
12. Pezzato, C. et al. An efficient artificial molecular pump. *Tetrahedron* **73**, 4849–4857 (2017).
13. Erbas-Cakmak, S. et al. Rotary and linear molecular motors driven by pulses of a chemical fuel. *Science* **358**, 340–343 (2017).

14. Pezzato, C. et al. Controlling dual molecular pumps electrochemically. *Angew. Chem. Int. Ed.* **57**, 9325–9329 (2018).
15. Qiu, Y. et al. A molecular dual pump. *J. Am. Chem. Soc.* **141**, 17472–17476 (2019).
16. Qiu, Y. et al. A precise polyrotaxane synthesizer. *Science* **368**, 1247–1253 (2020).
17. Astumian, R. D. & Bier, M. Mechanochemical coupling of the motion of molecular motors to ATP hydrolysis. *Biophys. J.* **70**, 637–653 (1996).
18. Kay, E. R., Leigh, D. A. & Zerbetto, F. Synthetic molecular motors and mechanical machines. *Angew. Chem. Int. Ed.* **46**, 72–191 (2007).
19. Astumian, R. D. Design principles for Brownian molecular machines: how to swim in molasses and walk in a hurricane. *Phys. Chem. Chem. Phys.* **9**, 5067–5083 (2007).
20. Astumian, R. D. Kinetic asymmetry allows macromolecular catalysts to drive an information ratchet. *Nat. Commun.* **10**, 3837 (2019).
21. Astumian, R. D. How molecular motors work—insights from the molecular machinist's toolbox: the Nobel prize in Chemistry 2016. *Chem. Sci.* **8**, 840–845 (2017).
22. Pezzato, C., Cheng, C., Stoddart, J. F. & Astumian, R. D. Mastering the non-equilibrium assembly and operation of molecular machines. *Chem. Soc. Rev.* **46**, 5491–5507 (2017).
23. Kassem, S. et al. Artificial molecular motors. *Chem. Soc. Rev.* **46**, 2592–2621 (2017).
24. Wang, Q., Chen, D. & Tian, H. Artificial molecular machines that can perform work. *Sci. China Chem.* **61**, 1261–1273 (2018).
25. Findlay, J. A. & Crowley, J. D. Functional nanomachines: recent advances in synthetic molecular machinery. *Tetrahedron Lett.* **59**, 334–346 (2018).
26. Coutrot, F. A focus on triazolium as a multipurpose molecular station for pH-sensitive interlocked crown-ether-based molecular machines. *ChemistryOpen* **4**, 556–576 (2015).
27. Zheng, H. et al. A dual-response [2]rotaxane based on a 1,2,3-triazole ring as a novel recognition station. *Chem. Eur. J.* **15**, 13253–13262 (2009).
28. Fielden, S. D. P., Leigh, D. A., McTernan, C. T., Pérez-Saavedra, B. & Vitorica-Yrezabal, I. J. Spontaneous assembly of rotaxanes from a primary amine, crown ether and electrophile. *J. Am. Chem. Soc.* **140**, 6049–6052 (2018).
29. Tian, C., Fielden, S. D. P., Whitehead, G. F. S., Vitorica-Yrezabal, I. J. & Leigh, D. A. Weak functional group interactions revealed through metal-free active template rotaxane synthesis. *Nat. Commun.* **11**, 744 (2020).
30. Tian, C., Fielden, S. D. P., Pérez-Saavedra, B., Vitorica-Yrezabal, I. J. & Leigh, D. A. Single-step enantioselective synthesis of mechanically planar chiral [2]rotaxanes using a chiral leaving group strategy. *J. Am. Chem. Soc.* **142**, 9803–9808 (2020).
31. Denis, M. & Goldup, S. M. The active template approach to interlocked molecules: Principles, progress and applications. *Nat. Rev. Chem.* **1**, 0061 (2017).
32. Wilson, M. R. et al. An autonomous chemically fuelled small-molecule motor. *Nature* **534**, 235–240 (2016).
33. Fyfe, M. C. T. et al. Anion assisted self-assembly. *Angew. Chem. Int. Ed. Engl.* **36**, 2068–2070 (1997).
34. della Sala, F., Neri, S., Maiti, S., Chen, J. L.-Y. & Prins, L. J. Transient self-assembly of molecular nanostructures driven by chemical fuels. *Curr. Opin. Biotechnol.* **46**, 27–33 (2017).
35. van Rossum, S. A. P., Tena-Solsona, M., van Esch, J. H., Eelkema, R. & Boekhoven, J. Dissipative out-of-equilibrium assembly of man-made supramolecular materials. *Chem. Soc. Rev.* **46**, 5519–5535 (2017).
36. Ragazzon, G. & Prins, L. J. Energy consumption in chemical fuel-driven self-assembly. *Nat. Nanotechnol.* **13**, 882–889 (2018).
37. Rieß, B., Grötsch, R. & Boekhoven, J. The design of dissipative molecular assemblies driven by chemical reaction cycles. *Chem* **6**, 552–578 (2020).
38. Borsley, S., Leigh, D. A. & Roberts, B. M. W. A doubly kinetically-gated information ratchet autonomously driven by carbodiimide hydration. *J. Am. Chem. Soc.* **143**, 4414–4420 (2021).
39. Astumian, R. D. Irrelevance of the power stroke for the directionality, stopping force, and optimal efficiency of chemically driven molecular machines. *Biophys. J.* **108**, 291–303 (2015).
40. Howard, J. Protein power strokes. *Curr. Biol.* **16**, R517–R519 (2006).
41. Hwang, W. & Karplus, M. Structural basis for power stroke vs. Brownian ratchet mechanisms of motor proteins. *Proc. Natl Acad. Sci. USA* **116**, 19777–19785 (2019).
42. van Dijk, L., Tilby, M. J., Szpera, R., Smith, O. A., Bunce, H. A. P. & Fletcher, S. P. Molecular machines for catalysis. *Nat. Rev. Chem.* **2**, 0117 (2018).
43. Zhang, L., Marcos, V. & Leigh, D. A. Molecular machines with bio-inspired mechanisms. *Proc. Natl Acad. Sci. USA* **115**, 9397–9404 (2018).
44. Biagini, C. et al. Dissipative catalysis with a molecular machine. *Angew. Chem. Int. Ed.* **58**, 9876–9880 (2019).
45. Astumian, R. D. et al. Non-equilibrium kinetics and trajectory thermodynamics of synthetic molecular pumps. *Mater. Chem.* **4**, 1304–1314 (2020).
46. Heard, A. W. & Goldup, S. M. Simplicity in the design, operation and applications of mechanically interlocked molecular machines. *ACS Cent. Sci.* **6**, 117–128 (2020).
47. Aprahamian, I. The future of molecular machines. *ACS Cent. Sci.* **6**, 347–358 (2020).
48. Moulin, E., Faour, L., Carmona-Vargas, C. C. & Giuseppone, N. From molecular machines to stimuli-responsive materials. *Adv. Mater.* **32**, 1906036 (2020).

Methods

General methods for the operation of the pump with continuous fuel addition

To a solution of **1** (5.8 mg, 5.9 μmol, 1 equiv.) and **3** (21 mg, 60 μmol, 10 equiv.) in dry toluene (27 μl) was added *i*-Pr_2NH (12 μl, 88 μmol, 15 equiv.). A solution of **2** (23 mg, 63 μmol, 11 equiv.) in toluene (133 μl total volume) was continuously added via syringe pump at RT for 16 h. An aliquot (around 2 μl) of the mixture was diluted with $CDCl_3$ (500 μl) and immediately analysed by ^1H NMR (600 MHz, 298 K). To the reaction mixture was added CF_3COOH (10 μl) and saturated aqueous NH_4Cl (5 ml). The mixture was extracted with CH_2Cl_2 (3 × 5 ml). The combined organic extracts were concentrated under reduced pressure. The residue was analysed by ESI-MS (electrospray ionization-mass spectrometry), which indicated the presence of [2]rotaxane **4**, [3]rotaxane **10** and [4]rotaxane **7**.

Data availability

The data that support the findings of this study are available within the paper and its Supplementary Information, or are available from the Mendeley data repository (https://data.mendeley.com/) at https://doi.org/10.17632/r339vx45sz.1.

Acknowledgements We thank the Engineering and Physical Sciences Research Council (EPSRC; grant number EP/P027067/1) and the European Research Council (ERC; Advanced Grant number 786630) for funding. We also thank the University of Manchester's Department of Chemistry Services for mass spectrometry. D.A.L. is a Royal Society Research Professor.

Author contributions S.A. and S.D.P.F. carried out the synthesis and characterization studies. D.A.L. directed the research. All authors contributed to the analysis of the results and the writing of the manuscript.

Competing interests The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-03575-3.
Correspondence and requests for materials should be addressed to D.A.L.
Peer review information *Nature* thanks R. Dean Astumian and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.
Reprints and permissions information is available at http://www.nature.com/reprints.

Insights from an information thermodynamics analysis of a synthetic molecular motor

Shuntaro Amano,[1] Massimiliano Esposito,[2] Elisabeth Kreidt,[1]
David Leigh*,[1] Emanuele Penocchio*,[2] and Benjamin Roberts[1]

[1] *Department of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom*
[2] *Complex Systems and Statistical Mechanics, Department of Physics and Materials Science, University of Luxembourg, L-1511 Luxembourg City, G.D. Luxembourg*

*email: david.leigh@manchester.ac.uk, emanuele.penocchio@uni.lu

Information is a physical quantity, the realisation of which transformed the physics of measurement and communication. However, the relationship and flow between information, energy and mechanics in chemical systems and mechanisms remains largely unexplored. Here we analyze a minimalist experimental example of an autonomous artificial chemically-driven molecular motor - a molecular information ratchet - in terms of information thermodynamics, a framework that quantitatively relates information to other thermodynamic parameters. This treatment reveals how directional motion is generated by free energy transfer from the chemical to the mechanical processes involving the motor. We find that the free energy transfer consists of two distinct contributions that can be considered as "energy flow" and "information flow". We identify the efficiency with which the chemical fuel powers the free energy transfer and show that this is a useful quantity with which to compare and evaluate mechanisms of, and guide designs for, molecular machines. The study provides a thermodynamic level of understanding of molecular motors that is general, complements previous analyses based on kinetics, and has practical implications for designing and improving synthetic molecular machines, regardless of the particular type of machine or chemical structure. In particular, the study confirms that, in line with kinetic analysis, power strokes do not affect the directionality of chemically-driven molecular machines. However, we also find that under some conditions power strokes can modulate the molecular motor current (how fast the components rotate), efficiency with respect to how free energy is dissipated, and the number of fuel molecules consumed per cycle. This may help explain the role of such conformational changes in biomolecular machine mechanisms and illustrates the interplay between energy and information in chemical systems.

Introduction

Understanding how and why a machine works in the way it does is crucial for optimizing designs and inventing new ones. For macroscopic machines such an understanding can be deduced from engineering principles and Newtonian physics. In contrast, at the molecular level there is no simple explanation for why individual components of biomachines move in a particular way. Is the movement of a specific amino acid in ATP synthase a key requisite for the mechanism, or does it occur incidentally as part of an evolutionary pathway that was overall successful? Biomolecular machines operate autonomously, apparently through Brownian ratchet mechanisms.[1–5] They use energy and information to rectify the directionality of random thermal movements of their components so that work can be performed. This type of machine includes biological pumps and motors, in which the energy is generally provided in the form of a chemical potential gradient.[6,7] Chemists are learning how to design synthetic analogues of such machines,[8–13] with examples including a minimalist autonomous chemically-fueled molecular rotary motor[14] (Fig. 1).

Kinetic models have proved useful[15–17] in describing the behavior of molecular machines, but such analysis is inherently unable to give an account of free energy transfer from the fuel to the machine. The consideration of thermodynamics is essential to understand free energy transduction, how this drives directional motion and generates the capacity to perform work[18], and, hence, how to optimize motor's design. Up to now attempts[12,19–24] to design molecular motors have been led by chemical intuition, with little opportunity to reliably judge the effectiveness of a machine's design or performance unless and until it has been realized experimentally.

Recently, stochastic thermodynamics has emerged as a method for studying systems that operate at energies of the order of thermal fluctuations.[25–28] The theory is valid even when systems are driven far from equilibrium[25,29,30] and has been used to study nonequilibrium nanoscale systems.[31–36] A major achievement in this field was the connection made between stochastic thermodynamics and information theory,[37] which gave birth

to "information thermodynamics".[38,39] Information thermodynamics relates information to other thermodynamic quantities, such as free energy and entropy, and has proved particularly successful in resolving apparent thermodynamic paradoxes, such as Maxwell's demon.[40]

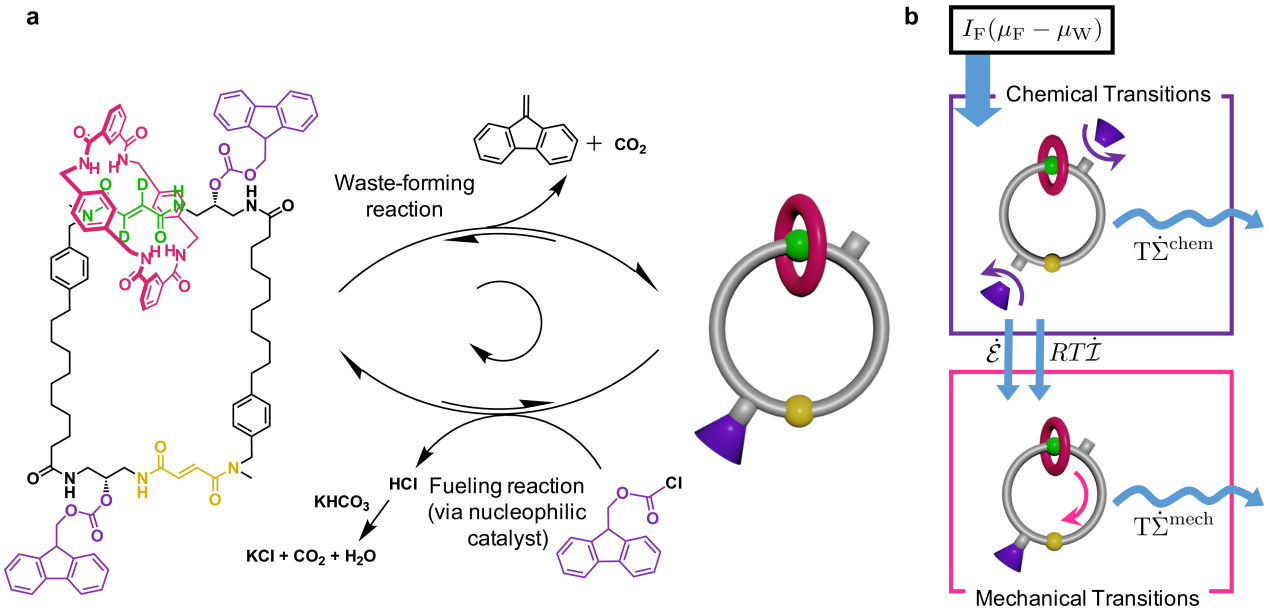


Figure 1. **A Rosetta Stone for chemical (reactions and co-conformational dynamics) and information thermodynamics descriptions of a molecular motor; two distinct but complementary accounts of the processes involved in a minimalist, autonomous, chemically-fueled, molecular rotary motor. a**, Chemical structure of the rotary motor and the chemical reactions involved in its operation. The motor comprises a benzylic amide macrocycle (magenta) and a track with two fumaramide binding sites (yellow, non-deuterated; green, deuterated for analytical purposes). The macrocycle randomly shuttles between the two fumaramide sites when its path is not blocked by Fmoc groups (purple). The fueling reaction consumes the fuel (Fmoc−Cl) and attaches an Fmoc group to the track, while the waste-forming reaction removes the Fmoc group (allowing passage of the macrocycle) and generates waste species (dibenzofulvene and CO_2). The fueling reaction is catalyzed by a pyridine-based nucleophilic catalyst (e.g., 4-dimethylaminopyridine or the bulky catalyst shown in Fig. 4a). Note that both reactions are considered reversible, even when the backward reactions (i.e. regeneration of fuel via barrier removal and waste products reacting to give barrier) are extremely rare events.[41] **b**, Information thermodynamics description of free energy transduction in the rotary motor. In the chemical transitions, free energy is supplied to the motor ($I_F(\mu_F - \mu_W)$), part of which is dissipated ($T\dot{\Sigma}^{chem}$). I_F is the rate at which the motor reacts with the fuel, $(\mu_F - \mu_W)$ the chemical potential difference between fuel and waste species, T the temperature, $\dot{\Sigma}^{chem}$ the entropy production rate of the chemical transitions. The rest of the free energy is supplied to the mechanical transitions either as energy ($\dot{\mathcal{E}}$) or information ($RT\dot{\mathcal{I}}$) flow, where R is the gas constant. Mechanical transitions dissipate this free energy ($T\dot{\Sigma}^{mech}$), generating directional motion of the macrocycle. When $\dot{\mathcal{E}} + RT\dot{\mathcal{I}} = 0$, mechanical transitions are at equilibrium and no net mechanical displacement of the macrocycle can arise.

Here, we develop a quantitative understanding of the processes that drive an autonomous chemically-fueled molecular motor[14] (Fig. 1) using an analysis that incorporates concepts from information thermodynamics within the framework of nonequilibrium thermodynamics of open chemical reaction networks.[42,43] Contrary to common models in stochastic thermodynamics[25,30], our framework neglects fluctuations and describes macroscopic ensembles of chemical species characterized by experimentally measurable concentrations. The approach is consistent with kinetic models[16,17,44] but goes further by introducing a quantitative thermodynamic understanding of how autonomous molecular motors work. Two features, "information flow" and "energy flow", contribute to the transfer of free energy from the fuel to the machine that is the origin of current (i.e. the net rate of displacement of the macrocycle directionally along the track, see Box I) in the motor. The effect of changing chemical gating,[15] power strokes[45–47] ("a viscoelastic, free energy releasing, large-amplitude conformational change"[45,47]), and overall rates on current and efficiency (Box I) are examined through simulations, revealing design principles for molecular motors. Particular insight is gained in terms of the role of power strokes in tuning a motor's performance while remaining consistent with core aspects of kinetic models,[16,17,44–48] informing the current debate concerning the role of power strokes in biomolecular machines.[44,46,47]

As a result, the mechanism of operation of the rotary motor can be understood in several different ways: through chemical design,[14] reaction kinetics,[15,49] molecular dynamics,[50] and now, nonequilibrium information

Box I. Definitions as applied to the minimalist rotary motor

Fueling reaction	The reaction that consumes the fuel (Fmoc−Cl) and attaches an Fmoc group to the motor's track, generating HCl as waste product. Forward ($k^{\mathrm{p/d}}_{+\mathrm{F}}$) and backward ($k^{\mathrm{p/d}}_{-\mathrm{F}}$) rate constants may depend on the macrocycle being proximal (p) or distal (d) to the reaction site.
Waste-forming reaction	The reaction that removes the Fmoc group from the motor's track, generating dibenzofulvene and CO_2 as waste products. Forward ($k^{\mathrm{p/d}}_{-\mathrm{W}}$) and backward ($k^{\mathrm{p/d}}_{+\mathrm{W}}$) rate constants may depend on the macrocycle being proximal (p) or distal (d) to the reaction site.
Chemical gating	The bias of distal over proximal rate constants in fueling and waste-forming reactions, quantified by $k^{\mathrm{d}}_{+\mathrm{F}}/k^{\mathrm{p}}_{+\mathrm{F}}$ (fueling gating) and $k^{\mathrm{p}}_{-\mathrm{W}}/k^{\mathrm{d}}_{-\mathrm{W}}$ (waste-forming gating).
Overall reaction rate	The sum of the rate constants for a particular process: *e.g.*, Fmoc addition through fueling reactions ($k^{\mathrm{d}}_{+\mathrm{F}} + k^{\mathrm{p}}_{+\mathrm{F}}$), Fmoc removal through waste-forming reactions ($k^{\mathrm{p}}_{-\mathrm{W}} + k^{\mathrm{d}}_{-\mathrm{W}}$), mechanical shuttling ($k_{+\Delta} + k_{-\Delta}$ or $k'_{+\Delta} + k'_{-\Delta}$).
Current	The net rate of directional displacement of the macrocycle along the track, quantified by equation (4).
Efficiency	The free energy available to the motor to dissipate as mechanical motion compared to the free energy available from the fuel-to-waste reaction, quantified by equation (5). This definition of efficiency does not take into account the background fuel decomposition.
Power Stroke	The viscoelastic, free energy releasing, mechanical shuttling of the macrocycle along the track. The free energy released is quantified by the change in standard chemical potential due to a net mechanical displacement: $\mu^{\circ}_{1^{\mathrm{D}}_{\mathrm{H}}} - \mu^{\circ}_{1^{\mathrm{D}}_{\mathrm{D}}}$ and $\mu^{\circ}_{1^{\mathrm{H}}_{\mathrm{D}}} - \mu^{\circ}_{1^{\mathrm{H}}_{\mathrm{H}}} \neq 0$.
Kinetic Asymmetry	The kinetic preference for one direction over the other in a chemomechanical cycle, embodied by the ratcheting constant: $K_{\mathrm{r}} = \frac{k_{+\Delta}k'_{+\Delta}(k^{\mathrm{p}}_{-\mathrm{F}}[\mathrm{HCl}]+k^{\mathrm{p}}_{-\mathrm{W}})^2([\mathrm{Fmoc-Cl}]k^{\mathrm{d}}_{+\mathrm{F}}+[\mathrm{CO_2}][\mathrm{DBF}]k^{\mathrm{d}}_{+\mathrm{W}})^2}{k_{-\Delta}k'_{-\Delta}([\mathrm{Fmoc-Cl}]k^{\mathrm{p}}_{+\mathrm{F}}+[\mathrm{CO_2}][\mathrm{DBF}]k^{\mathrm{p}}_{+\mathrm{W}})^2(k^{\mathrm{d}}_{-\mathrm{F}}[\mathrm{HCl}]+k^{\mathrm{d}}_{-\mathrm{W}})^2}$, where DBF stands for dibenzofulvene. When $K_{\mathrm{r}} > 1$, the motor preferentially cycles in the forward direction as apparent from equation (4).

thermodynamics. Accordingly, this minimalist molecular motor can act as a "Rosetta Stone" for relating these disparate frameworks[16,17,35,44,49–51], aiding the translation of concepts and relationships between energy, information, kinetics and molecular structure.

A bipartite chemical reaction network for the minimalist rotary motor

The rotary motor in Fig. 1 comprises a cyclic track with two degenerate binding sites (distinguishable by one being labelled with deuterium) for a macrocycle.[14] Fluorenylmethoxycarbonyl (Fmoc) groups, that sterically prevent passage of the macrocycle, can be attached to hydroxy residues on the track (Fig. 2). When only one barrier is in place, macrocycle shuttling enables exchange between two co-conformers ("co-conformers" are structures that differ in the relative positions of the components[52]) that have the macrocycle either proximal or distal to the free hydroxy group. The fueling reaction kinetically discriminates between the two co-conformers, favoring the reaction of the distal co-conformer over the proximal co-conformer. Under basic conditions, the waste-forming reaction removes barriers without any chemical gating. The chemical gating afforded by the biased fueling reaction, and the free energy supplied by the fuel-to-waste conversion, result in directional movement of the macrocycle around the track. The rotary motor can be represented by a chemomechanical network of reactions (Fig. 2) in which mechanical and chemical transitions are coupled, as in common models for biological molecular motors[44] and minimal Brownian motors with external dichotomous noise[53]. Since no transitions enable simultaneous change in the mechanical and chemical state, this network is said to be bipartite.[39,54]

Information thermodynamic analysis

As detailed in the Supplementary Information (SI), the rotary motor is modelled as the isothermal open bipartite chemical reaction network[39,42,43] shown in Fig. 2. The concentrations of the six motor species 2_{H},2_{D},$1^{\mathrm{H}}_{\mathrm{D}}$,$1^{\mathrm{D}}_{\mathrm{D}}$,$1^{\mathrm{H}}_{\mathrm{H}}$,$1^{\mathrm{D}}_{\mathrm{H}}$ (see Fig. 2 caption) evolve according to the rate constants of each reaction following mass-action kinetics. The system is open because the concentrations of fuel (Fmoc−Cl) and waste (HCl — liberated as part of the fueling reaction, and subsequently neutralized by $KHCO_3$ that is present to produce KCl, CO_2 and H_2O — and dibenzofulvene and CO_2 produced during the waste-forming reaction) species are kept constant through addition from, or removal to, an external source. An important quantity for our analysis is the chemical potential gradient between the fuel and waste species ($\mu_{\mathrm{Fmoc-Cl}} - \mu_{\mathrm{HCl}} - \mu_{\mathrm{dibenzofulvene}} - \mu_{\mathrm{CO_2}}$) which we denote $\mu_{\mathrm{F}} - \mu_{\mathrm{W}}$. For any thermodynamically consistent set of parameters (see Methods), the system will

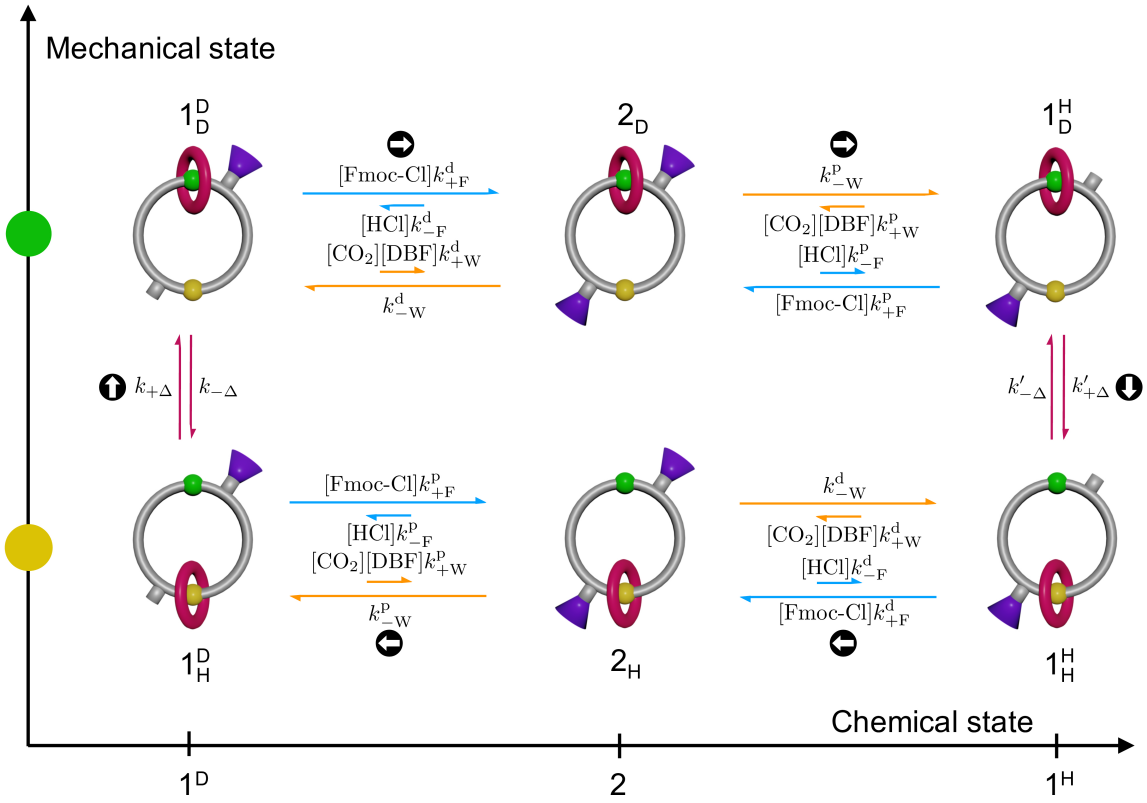


Figure 2. **Rotary motor as an open and bipartite chemical reaction network.** The state of the motor can be represented as a combination of two mechanical states (subscript D or H, expressing whether the macrocycle is binding to the deuterated or non-deuterated site on the track) and three chemical states (1^{D}, 2 or 1^{H}). The number in the chemical state shows the number of the Fmoc groups attached to the track; the superscript H or D denotes the binding site close to the attached Fmoc group. Note that chemical states without any Fmoc group on the track are neglected. The mechanical transitions involve displacement of the macrocycle, the rate constants for which are denoted $k_{+\Delta}$, $k_{-\Delta}$, $k'_{+\Delta}$ and $k'_{-\Delta}$. Subscript Δ signifies that they are the rate constants of mechanical transitions, which are only coupled to the thermal reservoir. Rate constants with and without the prime ($'$) are for the interconversion between $1^{\mathrm{H}}_{\mathrm{D}}$ and $1^{\mathrm{H}}_{\mathrm{H}}$, and the interconversion between $1^{\mathrm{D}}_{\mathrm{D}}$ and $1^{\mathrm{D}}_{\mathrm{H}}$, respectively. Subscript's sign shows the direction of the transition (+ for the clockwise and − for the counterclockwise direction). The rate constants of the fueling reaction and its reverse are denoted as $k^{\mathrm{p/d}}_{+\mathrm{F}}$ and $k^{\mathrm{p/d}}_{-\mathrm{F}}$, respectively. Superscript shows whether the macrocycle is proximal (p) or distal (d) to the reacting hydroxy group. The rate constants of the waste-forming reaction and its reverse are denoted as $k^{\mathrm{p/d}}_{-\mathrm{W}}$ and $k^{\mathrm{p/d}}_{+\mathrm{W}}$, respectively. In the experimental rotary motor, a clockwise current across the network is generated because $k^{\mathrm{p}}_{+\mathrm{F}} < k^{\mathrm{d}}_{+\mathrm{F}}$, $k^{\mathrm{p}}_{-\mathrm{W}} = k^{\mathrm{d}}_{-\mathrm{W}}$, and free energy is supplied to the motor by the fuel-to-waste conversion. The chemical gating in the fueling reaction arises due to the steric crowding in transition states of the proximal co-conformers ($1^{\mathrm{H}}_{\mathrm{D}}$ and $1^{\mathrm{D}}_{\mathrm{H}}$) compared to the distal co-conformers ($1^{\mathrm{D}}_{\mathrm{D}}$ and $1^{\mathrm{H}}_{\mathrm{H}}$). Circled arrows represent the most probable clockwise path. DBF stands for dibenzofulvene.

evolve towards a stationary state in which the concentrations of all the motor species (i.e., all co-conformations and chemical states of the motor) are constant in time, as are the thermodynamic properties of the system. The entropy production rate (the entropy changes in the system and in the reservoirs per unit time) denoted $\dot{\Sigma}$ measures how far from equilibrium the motor operates and has to be non-negative by virtue of the Second Law of Thermodynamics.[55] When multiplied by temperature (T) it corresponds to the amount of free energy that is instantaneously dissipated by the motor ($T\dot{\Sigma}$). In a stationary state, the only kind of state that will be considered in this paper, all the free energy from the conversion of the fuel-to-waste is dissipated by the motor: $T\dot{\Sigma} = I_{\mathrm{F}}(\mu_{\mathrm{F}} - \mu_{\mathrm{W}}) \geq 0$, where I_{F} is the rate at which the fuel is consumed by the motor (we neglect the fuel-to-waste background reaction).[43] This indicates that a non-null chemical potential gradient between fuel and waste species (we consider the case where $\mu_{\mathrm{F}} > \mu_{\mathrm{W}}$) is necessary to drive the system out of equilibrium, and produce directed motion. However, this expression gives no information as to how free energy is consumed nor the amount of dissipation specifically devoted to sustain directional motion of the components.

To obtain a more in-depth understanding, we use information thermodynamics to split the free energy dissipa-

tion rate into two separately non-negative contributions (Fig. 1b), one due to the chemical transitions ($T\dot{\Sigma}^{\mathrm{chem}}$) and the other due to the mechanical transitions ($T\dot{\Sigma}^{\mathrm{mech}}$). This is possible since the Fmoc-motor chemical reaction network is bipartite:[39]

$$T\dot{\Sigma} = \overbrace{I_{\mathrm{F}}(\mu_{\mathrm{F}} - \mu_{\mathrm{W}}) - \dot{\mathcal{E}} - RT\dot{\mathcal{I}}}^{T\dot{\Sigma}^{\mathrm{chem}} \geq 0} + \overbrace{\dot{\mathcal{E}} + RT\dot{\mathcal{I}}}^{T\dot{\Sigma}^{\mathrm{mech}} \geq 0} \geq 0\,. \tag{1}$$

This dissection of terms in equation (1) underlines that the free energy supplied by the fuel is only partially dissipated via the chemical processes as $T\dot{\Sigma}^{\mathrm{chem}}$ while the remaining part is transferred to the mechanical processes and dissipated as $T\dot{\Sigma}^{\mathrm{mech}}$. This transfer of free energy is composed of a standard free energy part, denoted "energy flow":

$$\dot{\mathcal{E}} = J(\mu^{\circ}_{1^{\mathrm{D}}_{\mathrm{H}}} - \mu^{\circ}_{1^{\mathrm{D}}_{\mathrm{D}}} + \mu^{\circ}_{1^{\mathrm{H}}_{\mathrm{D}}} - \mu^{\circ}_{1^{\mathrm{H}}_{\mathrm{H}}})\,, \tag{2}$$

and a mutual information part,[37] denoted "information flow":

$$RT\dot{\mathcal{I}} = JRT \log \frac{[1^{\mathrm{D}}_{\mathrm{H}}][1^{\mathrm{H}}_{\mathrm{D}}]}{[1^{\mathrm{D}}_{\mathrm{D}}][1^{\mathrm{H}}_{\mathrm{H}}]}\,, \tag{3}$$

where R is the gas constant and J is the stationary clockwise (as viewed in Fig. 2) current at which the motor operates (see Box I). This current can be expressed as

$$J = k_{+\Delta}[1^{\mathrm{D}}_{\mathrm{H}}] - k_{-\Delta}[1^{\mathrm{D}}_{\mathrm{D}}] = k'_{+\Delta}[1^{\mathrm{H}}_{\mathrm{D}}] - k'_{-\Delta}[1^{\mathrm{H}}_{\mathrm{H}}] = \Gamma(K_{\mathrm{r}} - 1)\,. \tag{4}$$

where both Γ and K_{r} are positive quantities (for derivation, see SI section V B4), the latter denoting the ratcheting constant (recently applied in the context of dissipative self-assembly[17,49]) which quantifies the kinetic asymmetry of the motor[15,16,56] (see Box I).

Under the experimental conditions in which the motor was originally operated,[14] the rotary motor is driven purely by information flow as the macrocycle binds with equal affinity to fumaramide stations adjacent to a hydroxy group and an Fmoc group (within the detection limits of ^{1}H nuclear magnetic resonance (NMR) measurements), so there is no energy flow. However, when the standard chemical potentials of the distal and proximal co-conformers differ, energy flow arises according to equation (2). An example of this could arise if there was, say, a stabilizing interaction between the macrocycle and the Fmoc group, which would lead to an increase in energy of the macrocycle upon removal of the Fmoc-barrier close to it. As a consequence, the standard chemical potential would decrease, leading to the release of heat, each time net mechanical displacement occurs in the forward direction. This fits the definition of a "power stroke"[45–47] (see Box I). Therefore, according to equation (2), the energy flow accounts for the part of the fueling free energy that contributes to destabilizing the macrocycle during chemical transitions and which is subsequently dissipated in a power stroke.

Mutual information quantifies the correlation between the two parts of a bipartite system[39,54], here the chemical and the mechanical states. For instance, when $[1^{\mathrm{H}}_{\mathrm{D}}]$ (the concentration of the species $1^{\mathrm{H}}_{\mathrm{D}}$) and $[1^{\mathrm{D}}_{\mathrm{H}}]$ are larger than $[1^{\mathrm{H}}_{\mathrm{H}}]$ and $[1^{\mathrm{D}}_{\mathrm{D}}]$, respectively, a correlation is present between the mechanical and chemical states: when the motor's chemical state is 1^{H}, its mechanical state is more likely to be D than H. Likewise, when the motor's chemical state is 1^{D}, its mechanical state is more likely to be H than D. A concentration distribution with this kind of correlation has smaller (Shannon-like[37]) entropy than one without ($[1^{\mathrm{H}}_{\mathrm{D}}]$ and $[1^{\mathrm{D}}_{\mathrm{H}}]$ equal to $[1^{\mathrm{H}}_{\mathrm{H}}]$ and $[1^{\mathrm{D}}_{\mathrm{D}}]$, respectively). Therefore, correlation between the mechanical and chemical states (mutual information) generates an entropic driving force for a directional current (from $[1^{\mathrm{H}}_{\mathrm{D}}]$ to $[1^{\mathrm{H}}_{\mathrm{H}}]$ and from $[1^{\mathrm{D}}_{\mathrm{H}}]$ to $[1^{\mathrm{D}}_{\mathrm{D}}]$). From the thermodynamic viewpoint, mutual information constitutes the entropic contribution of the free energy that comes from the fuel. As mutual information is constant in the stationary state, changes in mutual information due to the chemical processes and mechanical processes are balanced[39]. If, as in the above situation, $\dot{\mathcal{I}} > 0$, the chemical transitions are producing mutual information that is consumed by the mechanical ones. Therefore, according to equation (3), the information flow accounts for the part of fueling free energy that contributes to increasing the system's mutual information during chemical transitions and which is subsequently erased by mechanical shuttling.

Regimes where the free energy supplied by the chemical to the mechanical processes is exclusively due to the information flow ($RT\dot{\mathcal{I}}$), and thus lack any energy flow ($\dot{\mathcal{E}}$), are denoted as pure Maxwell demon regimes.[39]

This analysis demonstrates that the free energy supplied to the mechanical processes by the energy and information flows, $RT\dot{\mathcal{I}} + \dot{\mathcal{E}}$, is the origin of net directional motion of the macrocycle around the track. Indeed, in the absence of such flow ($RT\dot{\mathcal{I}} + \dot{\mathcal{E}} = 0$) the mechanical transitions are at thermodynamic equilibrium ($\dot{\Sigma}^{\mathrm{mech}} = 0$), meaning a zero directional shuttling current ($J = 0$; see equation (4)). From an information thermodynamics perspective, the molecular motor operates by using chemical processes to transduce the free

energy supplied by the fuel into the free energy supplied to the mechanical processes. The efficiency of this transduction is the ratio of the latter to the former and is bounded between zero and one due to the non-negativity of $\dot{\Sigma}^{\mathrm{chem}}$:

$$0 \leq \eta = \frac{RT\dot{\mathcal{I}} + \dot{\mathcal{E}}}{I_{\mathrm{F}}(\mu_{\mathrm{F}} - \mu_{\mathrm{W}})} = 1 - \frac{T\dot{\Sigma}^{\mathrm{chem}}}{I_{\mathrm{F}}(\mu_{\mathrm{F}} - \mu_{\mathrm{W}})} \leq 1 \tag{5}$$

In this setup, the energy and information flows are entirely dissipated by the shuttling of the macrocycle as $T\dot{\Sigma}^{\mathrm{mech}}$. If the mechanical steps of the motor were to work against a force (for example if a load were attached to the macrocycle), $T\dot{\Sigma}^{\mathrm{mech}}$ would incorporate a negative work term in addition to $RT\dot{\mathcal{I}} + \dot{\mathcal{E}}$, which could serve to define the efficiency of the energy and information flows being converted into output work, instead of just being dissipated. The energy and information flows would thus constitute the maximum work output that can be delivered by the motor. A traditional thermodynamic analysis of such a motor[25,26,57] would exclusively focus on the efficiency with which the input free energy supplied by the fuel-to-waste chemical potential gradient is converted into output work, thus over-estimating the maximum work output as the overall free energy input $I_F(\mu_{\mathrm{F}} - \mu_{\mathrm{W}})$. The present approach refines this analysis by showing how the input-to-output transduction is mediated by the free energy transfer within the motor, whose efficiency η (equation (5)) limits the maximum work output potentially deliverable by the motor. It also formally defines a thermodynamic efficiency that can be applied to motors while they perform no appreciable output work, as is the case for most of the synthetic molecular motors made to date, and can serve to compare the efficiencies of their operation.

The framework we have outlined can also be used to re-derive previous results obtained using kinetic arguments as a consequence of the Second Law of Thermodynamics in bipartite systems. In the SI (section V B), we show that the condition $T\dot{\Sigma}^{\mathrm{mech}}$=0 implies $K_{\mathrm{r}} = 1$, while the condition $T\dot{\Sigma}^{\mathrm{mech}} > 0$ implies $K_{\mathrm{r}} \neq 1$, with forward movement when $K_{\mathrm{r}} > 1$. This shows how the nonequilibrium thermodynamic framework, which focuses on energetic aspects quantified by the dissipation $T\dot{\Sigma}^{\mathrm{mech}}$, is consistent with previous analysis[15–17,49] focusing on kinetic aspects quantified by K_{r}, which determines the sign of the current J according to equation (4). This reiterates the effectiveness of this information-thermodynamics-based approach and, again, demonstrates the usefulness of this minimalist molecular motor as a Rosetta Stone for the translation of meaning and understanding between different frameworks for describing phenomena.

Previously,[17] the ratcheting constant has also been related to the ability of a dissipative self-assembly system to store free energy, but this connection is only valid in an operating regime where chemical transitions are significantly faster than mechanical ones. The information thermodynamics framework offers a general understanding of dissipative chemical systems and establishes limits to the maximum work deliverable by them that are valid in any operating regime, thanks to concepts such as the efficiency η introduced in equation (5).

Design principles for molecular motors

To demonstrate the use of our framework as a design tool, we explored the effects of altering the design features of the rotary motor on its current and efficiency (see SI section VI). Basing the simulations on experimentally derived parameters[14] (see SI section VI A), we established that under the experimental conditions employed, the rotary motor is driven purely by information flow as only the macrocycle distribution, rather than binding site affinity (within the detection limits of ^{1}H nuclear magnetic resonance (NMR) measurements), is altered during operation. Under experimental conditions ([motor] = 10 mM, [Fmoc−Cl] = 30 mM, which is sustained by constant addition, [bulky catalyst] = 50 mM, [Et_3N] = 15 mM, [$KHCO_3$] = 200 mM, CH_2Cl_2, room temperature),[14] the current was estimated to be $2.1 \cdot 10^{-8}$ mol dm^{-3} s^{-1}, requiring an average of seven fuel molecules per cycle per motor, though only 10^{-6}% of the free energy provided by the fuel is used to sustain the current (efficiency, $\eta = 10^{-8}$, see equation (5)).

Kinetic modifications

Varying model parameters allows consideration of the effects of potential structural and chemical changes on the rotary motor (SI section VI B1). Greater chemical gating for either the fueling (Fig. 3a, b) or waste-forming reaction (SI section VI B2) increases the current and efficiency by increasing information flow. The former has been achieved by increasing the steric bulk of the barrier-formation catalyst[14,58,59] (Fig. 4a) and the latter by catalysis of barrier removal by a proximal macrocycle (Fig. 4b).[24] Gating of both fueling and waste-forming reactions (SI section VI B3) was recently demonstrated in a [2]rotaxane information ratchet.[24] Inverting chemical gating is predicted to reverse the direction of the motor and could be achieved if the macrocycle activates, rather

than hinders, proximal barrier formation (Fig. 4c). In the absence of a kinetic preference, or when gating from the fueling and waste-forming reactions cancel out, the motor stops working, as this precludes information flow and hence mechanical dissipation ($\dot{\Sigma}^{\mathrm{mech}} = 0$). These results are consistent with kinetic models where kinetic asymmetry predicts the direction of the current.[15–17]

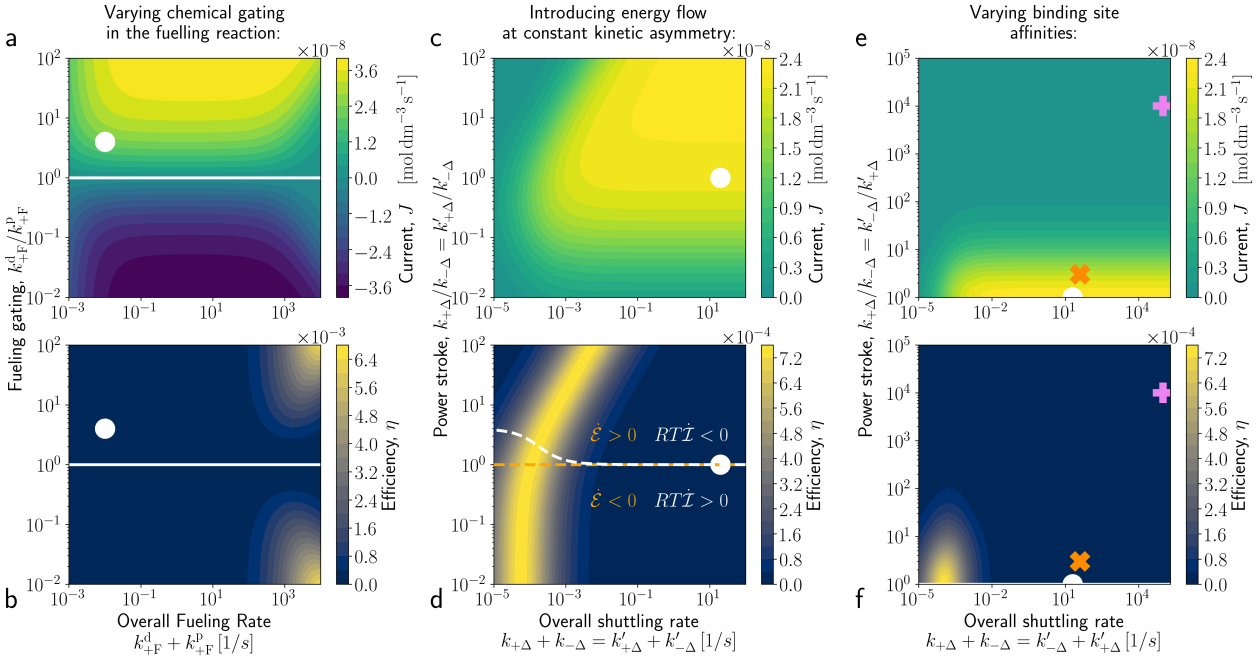


Figure 3. **Numerical simulations of different molecular motor design modifications.** Simulated current (J, top) and efficiency (η, bottom). Solid white lines indicate the regions of null current (and efficiency). White circles indicate the approximate parameters of the experimental rotary motor, giving a current of $2.1 \cdot 10^{-8}\,\mathrm{mol\,dm^{-3}\,s^{-1}}$ and efficiency of 10^{-6}%. Directional motion in the experimental rotary motor arises because distal fueling reactions are $\sim 4\times$ faster than proximal fueling reactions (chemical gating, $k^{\mathrm{d}}_{+\mathrm{F}}/k^{\mathrm{p}}_{+\mathrm{F}} \approx 4$), due to steric hindrance of the proximal transition state, while there was no chemical gating of waste-forming reactions ($k^{\mathrm{p}}_{-\mathrm{W}}/k^{\mathrm{d}}_{-\mathrm{W}} \approx 1$). **a**,**b**, Overall rate of the fueling reactions (x-axis) and fueling gating (y-axis) are varied. **c**,**d**, Overall shuttling rate (x-axis) and the ratio of forward and backward shuttling rates (indicative of power stroke magnitude, see Methods) are varied. Power strokes are introduced in a way that induces energy flow ($\dot{\mathcal{E}}$) while keeping K_{r} constant. Below the dashed white lines, information flow ($RT\dot{\mathcal{I}}$) is positive and thus contributes to mechanical dissipation $\dot{\Sigma}^{\mathrm{mech}}$, while above these lines, information flow is negative, requiring the motor to be driven by positive energy flow ($\dot{\mathcal{E}} > 0$). Negative energy flow is found below the orange dashed line. **e**,**f**, Overall shuttling rate (x-axis) and ratio of shuttling rate between stations (y-axis) are varied. When this ratio diverges from one (indicating non-degenerate binding sites, see Methods), both the current and efficiency drop. The pink + symbol indicates the position expected by changing one fumaramide group to a succinamide.[60] The orange × symbol indicates a position with a 3:1 difference in binding site affinity for the macrocycle.

Power stroke modifications

The relevance of power strokes in molecular machinery is contentious as power strokes are often experimentally observed in biological molecular motors,[44–46] but according to analysis based on kinetic asymmetry, the magnitude of the free energy released by such conformational changes does not affect the properties of chemically-driven molecular machines such as directionality and stopping force, and cannot improve the efficiency of a motor to work against an external force.[47] To reconcile differing viewpoints as to the significance of power strokes in molecular machines, we attempted to use our framework to understand the ways in which a power stroke can affect a molecular motor while staying kinetically and thermodynamically consistent. Power strokes can be used to induce energy flow (equation (2)) and, in principle, could be added to the rotary motor by introducing attractive interactions between the barrier and the macrocycle (Fig. 4d), stabilizing distal co-conformers $1^{\mathrm{D}}_{\mathrm{D}}$ and $1^{\mathrm{H}}_{\mathrm{H}}$, or by adding repulsive interactions between the free barrier site and the macrocycle (Fig. 4e), destabilizing proximal co-conformers $1^{\mathrm{D}}_{\mathrm{H}}$ and $1^{\mathrm{H}}_{\mathrm{D}}$. Our simulations show that power strokes can change the magnitude of the current and efficiency of internal energy transduction (as defined in equations (4) and (5),

respectively) despite kinetic asymmetry remaining unaltered (see Methods, SI section VI B4). This is because, while not altering K_r, power strokes can still increase the value of Γ in equation (4), reflecting their ability to favor forward cycles by inducing energy flow out of equlibrium (Fig. 3c, d). However, power strokes cannot drive directional motion in the absence of kinetic asymmetry ($RT\dot{\mathcal{I}} = -\dot{\mathcal{E}}$ when $K_r = 1$), nor can they invert directionality while the kinetic asymmetry remains constant. In these simulations, altering power strokes and kinetic asymmetry together (SI section VI B5) gave the greatest simulated efficiency suggesting that modifying both aspects may be important for optimizing the design of synthetic molecular motors. We note that improvements that occur through power strokes arise from the induced energy flow rather than from any special role for the energetically downhill nature of the power stroke in determining the motor's behaviour. As our analysis shows, a prerequisite for the motor's operation is the availability of free energy to dissipate through mechanical motion. Therefore, any design feature that enhances free energy transduction from the chemical to the mechanical transitions could equally well foster improvements in performance.

If power strokes cancel out over the motor cycle, then no net energy flow is introduced: free energy gained from one mechanical transition is lost in the other (see Methods, SI section VI B6). This could be realized in molecular form by using non-degenerate binding sites, for example, changing one fumaramide site to a more weakly-binding succinamide unit (Fig. 3e, f and 4f) with a typical difference in binding energy of $23\,\mathrm{kJ\,mol^{-1}}$ (pink + symbol, Fig. 3e, f, equilibrium distribution >99:1)[60] under experimental conditions similar to those used for rotary motor operation. Kinetic asymmetry remains unaltered in this scenario and, correspondingly, the direction of the motor could not be inverted in the simulation.[15–17] However, our analysis suggests that such a change would be sufficient to effectively stall the motor, if operated under the original experimental conditions, despite the unchanged kinetic asymmetry. The simulations predict that, with power strokes cancelling out, any change from degenerate binding sites lowers the current and efficiency, though a smaller difference would leave the motor functional, albeit less effective. A 3:1 bias[58] (orange × symbol, Fig. 3e, f,) is predicted to reduce the current by $\sim 20\%$, rendering a design with non-degenerate binding sites plausible but less effective than a motor with binding sites of equal affinities.

In all of the cases considered, the highest efficiencies are predicted for when the rates of all the forward processes are approximately equal, leading to a cycle with no single rate-limiting step. Rate-limiting mechanical steps promote futile cycles, in which fuel is consumed without taking a forward step, as the unfavorable fueling reaction is kinetically favored over shuttling, decreasing both current and efficiency. Rate-limiting chemical reactions decrease the information flow in the steady-state since fast shuttling hinders the generation of a concentration bias compared to mechanical equilibrium, resulting in lower thermodynamic efficiency without reducing current or substantially changing fuel consumption per cycle. The strong dependence of efficiency on shuttling rate indicates that, like macroscopic engines, the efficiency of a motor will be dependent on the load it is working against.[35] To use molecular machines most efficiently, they must either be tailored to the job they perform — such as using diesel engines for heavy loads — or they must use the equivalent of gears for macroscopic engines, to ensure they are working under optimal conditions.

Conclusions

Information thermodynamics-based analysis of a minimalist autonomous chemically-driven molecular motor shows how information and energy flow, the two components of free energy transfer from chemical to mechanical transitions, enable the generation of directional motion from free energy supplied by a chemical fuel. The experimental rotary motor is a pure "chemical Maxwell's demon", as information flow is the sole driving force. However, energy flow could potentially be introduced using power strokes, one of several design variations explored using our model. The predicted effect of energy flow is in line with observations made in biological motors[44] and contributes to the ongoing debate regarding the role of power strokes in molecular motors[46–48]. Information thermodynamics confirms that, in line with kinetic analysis, power strokes do not affect some key properties of chemically-driven molecular machines such as directionality. However, the magnitude of power strokes is able to affect the magnitude of the current (how fast the motor components rotate), the efficiency in terms of how free energy is dissipated, and the number of fuel molecules consumed per cycle. But these results should not be misinterpreted as supporting special significance of power strokes compared to other processes in the chemomechanical cycle. The information thermodynamics framework used in this paper should be generally applicable to other types of synthetic molecular machines, such as non-autonomous[10,12] and light-driven[8,9,11] motors, providing a quantitative basis through which to compare molecular machine designs. Additionally, it could, in principle, be extended to other types of (supra)molecular systems (such as dissipative self-assembly[17]) powered by chemical engines[61]. We have uncovered for the first time significant roles for 'energy flow' and 'information flow' in the mechanism of the transduction of free energy from chemical reactions by molecular machinery, although the exact nature of the connection of energy and information flow to energy and information

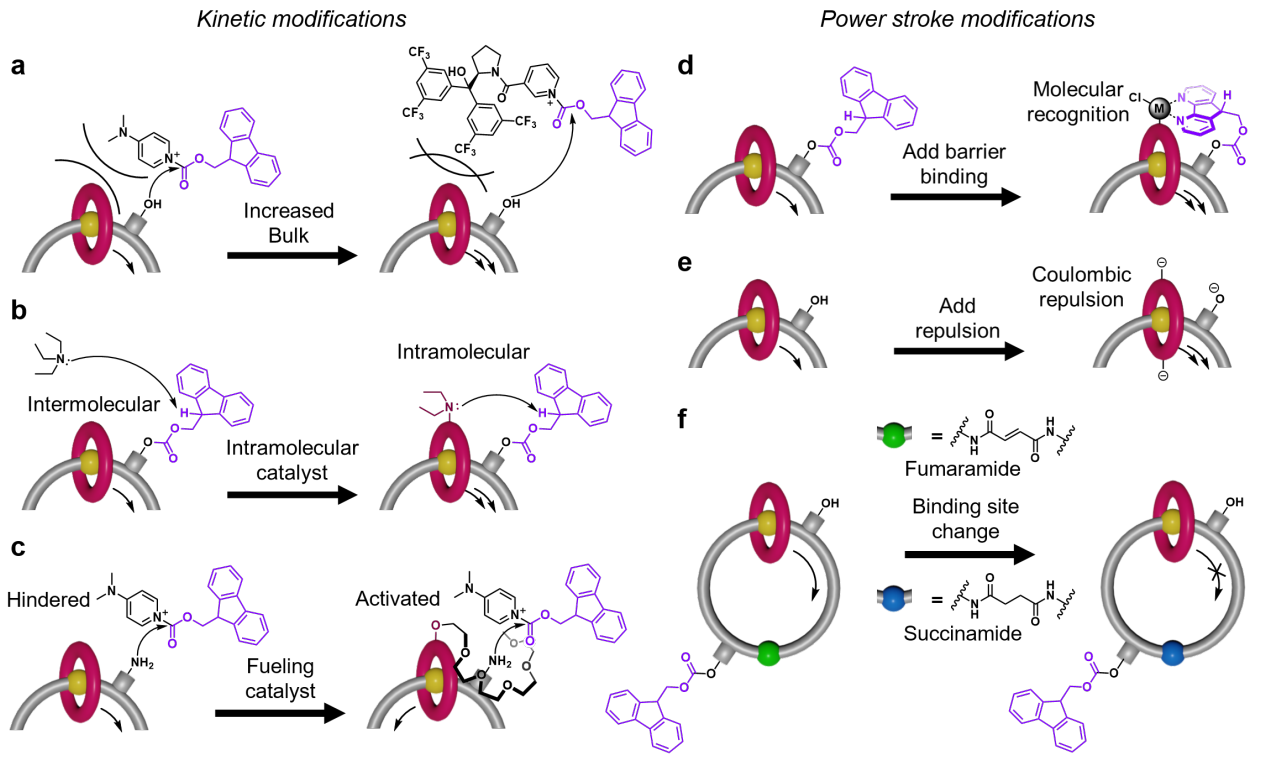


Figure 4. **Potential ways of achieving different molecular motor design modifications. a**, Increasing the steric bulk of the catalyst could be used to increase the fueling chemical gating by slowing the undesired proximal reaction. **b**, Addition of a catalyst for the waste-forming reaction to the macrocycle could increase waste-formation chemical gating by increasing the rate of the proximal waste-forming reaction. **c**, Addition of a catalyst to the macrocycle that accelerates rather than hinders proximal barrier formation could negate chemical gating arising from steric hindrance, leading to an inversion of the directionality of the motor. **d**, Allowing the formation of a complex that binds the barrier and the macrocycle could allow the introduction of a power stroke by stabilizing the co-conformer in which the macrocycle is adjacent to the barrier. **e**, Conversely, a power stroke might be introduced by destabilizing the co-conformer with the macrocycle adjacent to the site with no barrier, for example by introduction of coulombic repulsion. **f**, Swapping one fumaramide binding site for a succinamide should, in principle, be sufficient to stall the rotary motor under the previous experimental operating conditions.

ratcheting[1,62] remains to be clarified. The minimalist autonomous chemically-driven molecular motor acts as a Rosetta Stone for relating energy, information, kinetics and molecular structure by aiding the translation of concepts and relationships between the 'languages' (i.e. frameworks) of chemical kinetics, thermodynamics and chemical reactions.

Methods

Wegscheider's conditions: To ensure that the system reaches thermodynamic equilibrium (detailed balance), when there is no chemical potential gradient between fuel and waste species ($\mu_{\mathrm{F}} - \mu_{\mathrm{W}} = 0$), the rate constants must satisfy Wegscheider's conditions.[42] These are equivalent to the constraints on the rate constants imposed in previous kinetic analyses[4,16,48] and dictate that the product of the forward rate constants along each independent cyclic pathway of reactions in the network, with neither net consumption nor net production of fuel or waste species, must equal the product of the corresponding backward rate constants:

$$k^{\mathrm{p}}_{-\mathrm{F}} k_{+\Delta} k^{\mathrm{d}}_{+\mathrm{F}} k^{\mathrm{p}}_{-\mathrm{F}} k'_{+\Delta} k^{\mathrm{d}}_{+\mathrm{F}} = k^{\mathrm{p}}_{+\mathrm{F}} k_{-\Delta} k^{\mathrm{d}}_{-\mathrm{F}} k^{\mathrm{p}}_{+\mathrm{F}} k'_{-\Delta} k^{\mathrm{d}}_{-\mathrm{F}} \tag{M1a}$$

$$k^{\mathrm{p}}_{-\mathrm{W}} k_{+\Delta} k^{\mathrm{d}}_{+\mathrm{W}} k^{\mathrm{p}}_{-\mathrm{W}} k'_{+\Delta} k^{\mathrm{d}}_{+\mathrm{W}} = k^{\mathrm{p}}_{+\mathrm{W}} k_{-\Delta} k^{\mathrm{d}}_{-\mathrm{W}} k^{\mathrm{p}}_{+\mathrm{W}} k'_{-\Delta} k^{\mathrm{d}}_{-\mathrm{W}} \,, \tag{M1b}$$

see SI for derivation. These conditions were always imposed in numerical simulations to guarantee thermodynamic consistency.[41]

Local Detailed Balance: In Fig. 3, variations in the power stroke magnitude have been related to changes in the ratio of shutting rate constants by virtue of the local detailed balance principle.[42] Note that our use of

the term "local detailed balance" is the same as used in Refs 42 and in the SI of 43, but very different from how local detailed balance is discussed in Refs 16 and 63. "Local detailed balance", as used in this work, would refer to equation 5 in Ref. 16 (for a more detailed discussion see SI section III). Local detailed balance can be used to relate the log-ratio of rate constants to the differences in standard chemical potentials between reagents and products. For instance, the following relation (equation (M2)) holds

$$RT \log\left(\frac{k_{+\Delta}k'_{+\Delta}}{k_{-\Delta}k'_{-\Delta}}\right) = \mu^{\circ}_{1^{\mathrm{D}}_{\mathrm{H}}} - \mu^{\circ}_{1^{\mathrm{D}}_{\mathrm{D}}} + \mu^{\circ}_{1^{\mathrm{H}}_{\mathrm{D}}} - \mu^{\circ}_{1^{\mathrm{H}}_{\mathrm{H}}}\,, \tag{M2}$$

and was employed in numerical simulations. In addition, Wegscheider's conditions (equations (M1a) and (M1b)) imply that a variation in the power stroke magnitude must always be compensated by a variation in the fueling and waste-forming rate constants, since the following constraint (equation (M3)) must always hold for thermodynamic consistency:

$$\frac{k_{+\Delta}}{k_{-\Delta}} \cdot \frac{k'_{+\Delta}}{k'_{-\Delta}} = \frac{k^{\mathrm{p}}_{+\mathrm{F}}k^{\mathrm{d}}_{-\mathrm{F}}}{k^{\mathrm{p}}_{-\mathrm{F}}k^{\mathrm{d}}_{+\mathrm{F}}} \cdot \frac{k^{\mathrm{p}}_{+\mathrm{W}}k^{\mathrm{d}}_{-\mathrm{W}}}{k^{\mathrm{p}}_{-\mathrm{W}}k^{\mathrm{d}}_{+\mathrm{W}}}\,. \tag{M3}$$

When energy-flow-inducing power strokes were introduced in Fig. 3c and d, the constraint in equation (M3) was imposed by changing rate constants $k^{\mathrm{p}}_{+\mathrm{W}}$ and $k^{\mathrm{p}}_{+\mathrm{F}}$ according to variations in shuttling rate constants. By doing so, kinetic asymmetry (K_{r}) is not altered during the simulation, but the magnitude of the current in equation (4) can still change by virtue of alterations in the value of the positive factor Γ (for its mathematical expression, see SI section V B4). Instead, in Supplementary Fig. S.5 rate constants $k^{\mathrm{p}}_{+\mathrm{W}}$ and $k^{\mathrm{p}}_{-\mathrm{F}}$ were changed to vary energy flow and K_{r} together. Experimentally, this could correspond to introducing an interaction between the macrocycle and the Fmoc group that affects (Supplementary Fig. S.5) or not (Fig. 3c and d) the transition state of the proximal fueling reaction, without affecting the transition state of the proximal waste-forming reaction. Note that, when binding affinities are modified as in Fig. 3e and f, the left-hand side of equation (M3) stays constant and the constraint is automatically satisfied.

Acknowledgements

We thank the European Research Council (ERC Consolidator Grant n. 681456 to M.E. and funding E.P.; ERC Advanced Grant n. 786630 to D.A.L), the FQXi foundation, project "Information as a fuel in colloids and superconducting quantum circuits" (grant FQXi-IAF19-05 to M.E.), the Engineering and Physical Sciences Research Council (EPSRC; grant EP/P027067/1 to D.A.L.), the Deutsche Forschungsgemeinschaft (DFG) for a postdoctoral fellowship to E.K., and the University of Manchester and EPSRC for PhD studentships to S.A. and B.M.W.R.. D.A.L. is a Royal Society Research Professor.

Author contributions

SA, BR and EK proposed the collaboration. EP developed the theoretical model. SA, BR, EP and EK carried out the theoretical analysis and simulations. DL and ME directed the research. All authors contributed to the analysis of the results and the writing of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information is available for this paper. Correspondence and requests for materials should be addressed to DL and EP. Reprints and permissions information is available at www.nature.com/reprints.

Data availability

All data needed to reproduce numerical results are reported in the Supplementary Information.

Code availability

The code that generated the plots is available from EP upon request.

REFERENCES

[1] Chatterjee, M. N., Kay, E. R. & Leigh, D. A. Beyond switches: Ratcheting a particle energetically uphill with a compartmentalized molecular machine. *J. Am. Chem. Soc.* **128**, 4058–4073 (2006).

[2] Coskun, A., Banaszak, M., Astumian, R. D., Stoddart, J. F. & Grzybowski, B. A. Great expectations: can artificial molecular machines deliver on their promise? *Chem. Soc. Rev.* **41**, 19–30 (2012).

[3] Erbas-Cakmak, S., Leigh, D. A., McTernan, C. T. & Nussbaumer, A. L. Artificial molecular machines. *Chem. Rev.* **115**, 10081–10206 (2015).

[4] Qiu, Y., Feng, Y., Guo, Q.-H., Astumian, R. D. & Stoddart, J. F. Pumps through the ages. *Chem* **6**, 1952 – 1977 (2020).

[5] Aprahamian, I. The future of molecular machines. *ACS Cent. Sci.* **6**, 347–358 (2020).

[6] Howard, J. *Mechanics of Motor Proteins and the Cytoskeleton* (Oxford University Press, Oxford, 2001).

[7] Astumian, R. D., Mukherjee, S. & Warshel, A. The physics and physical chemistry of molecular machines. *ChemPhysChem* **17**, 1719–1741 (2016).

[8] Koumura, N., Zijlstra, R. W. J., van Delden, R. A., Harada, N. & Feringa, B. L. Light-driven monodirectional molecular rotor. *Nature* **401**, 152–155 (1999).

[9] Serreli, V., Lee, C.-F., Kay, E. R. & Leigh, D. A. A molecular information ratchet. *Nature* **445**, 523–527 (2007).

[10] Cheng, C. *et al.* An artificial molecular pump. *Nat. Nanotechnol.* **10**, 547–553 (2015).

[11] Ragazzon, G., Baroncini, M., Silvi, S., Venturi, M. & Credi, A. Light-powered autonomous and directional molecular motion of a dissipative self-assembling system. *Nat. Nanotechnol.* **10**, 70–75 (2015).

[12] Erbas-Cakmak, S. *et al.* Rotary and linear molecular motors driven by pulses of a chemical fuel. *Science* **358**, 340–343 (2017).

[13] Amano, S., Fielden, S. D. P. & Leigh, D. A. A catalysis-driven artificial molecular pump. *Nature* **594**, 529–534 (2021).

[14] Wilson, M. R. *et al.* An autonomous chemically fuelled small-molecule motor. *Nature* **534**, 235–240 (2016).

[15] Astumian, R. D. How molecular motors work — insights from the molecular machinist's toolbox: the nobel prize in chemistry 2016. *Chem. Sci.* **8**, 840–845 (2017).

[16] Astumian, R. D. Kinetic asymmetry allows macromolecular catalysts to drive an information ratchet. *Nat. Commun.* **10**, 3837 (2019).

[17] Ragazzon, G. & Prins, L. J. Energy consumption in chemical fuel-driven self-assembly. *Nat. Nanotechnol.* **13**, 882–889 (2018).

[18] Esposito, M. Open questions on nonequilibrium thermodynamics of chemical reaction networks. *Commun. Chem.* **3**, 107 (2020).

[19] Kelly, T. R., De Silva, H. & Silva, R. A. Unidirectional rotary motion in a molecular system. *Nature* **401**, 150–152 (1999).

[20] Mock, W. L. & Ochwat, K. J. Theory and example of a small-molecule motor. *J. Phys. Org. Chem.* **16**, 175–182 (2003).

[21] Fletcher, S. P., Dumur, F., Pollard, M. M. & Feringa, B. L. A reversible, unidirectional molecular rotary motor driven by chemical energy. *Science* **310**, 80–82 (2005).

[22] Collins, B. S. L., Kistemaker, J. C. M., Otten, E. & Feringa, B. L. A chemically powered unidirectional rotary molecular motor based on a palladium redox cycle. *Nat. Chem.* **8**, 860–866 (2016).

[23] Zhang, Y. *et al.* A chemically driven rotary molecular motor based on reversible lactone formation with perfect unidirectionality. *Chem* **6**, 2420–2429 (2020).

[24] Borsley, S., Leigh, D. A. & Roberts, B. M. W. A doubly kinetically-gated information ratchet autonomously driven by carbodiimide hydration. *J. Am. Chem. Soc.* **143**, 4414–4420 (2021).

[25] Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* **75**, 126001 (2012).

[26] Parrondo, J. M. R. & de Cisneros, B. J. Energetics of brownian motors: a review. *Appl. Phys. A* **75**, 179–191 (2002).

[27] Esposito, M., Lindenberg, K. & Van den Broeck, C. Universality of efficiency at maximum power. *Phys. Rev. Lett.* **102**, 130602 (2009).

[28] Benenti, G., Casati, G., Saito, K. & Whitney, R. S. Fundamental aspects of steady-state conversion of heat to work at the nanoscale. *Phys. Rep.* **694**, 1–124 (2017).

[29] Jarzynski, C. Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale. *Annu. Rev. Condens. Matter Phys.* **2**, 329–351 (2011).

[30] Peliti, L. & Pigolotti, S. *Stochastic Thermodynamics: An Introduction* (Princeton University Press, Princeton, 2021).

[31] Andrieux, D. & Gaspard, P. Fluctuation theorems and the nonequilibrium thermodynamics of molecular motors. *Phys. Rev. E* **74**, 011906 (2006).

[32] Lipowsky, R. & Liepelt, S. Chemomechanical coupling of molecular motors: Thermodynamics, network representations, and balance conditions. *J. Stat. Phys.* **130**, 39–67 (2008).

[33] Toyabe, S. *et al.* Nonequilibrium energetics of a single $\mathbf{f}_1$-atpase molecule. *Phys. Rev. Lett.* **104**, 198103 (2010).

[34] Ariga, T., Tomishige, M. & Mizuno, D. Nonequilibrium energetics of molecular motor kinesin. *Phys. Rev. Lett.* **121**, 218101 (2018).

[35] Brown, A. I. & Sivak, D. A. Theory of nonequilibrium free energy transduction by molecular machines. *Chem. Rev.* **120**, 434–459 (2020).

[36] Ciliberto, S. Experiments in stochastic thermodynamics: Short history and perspectives. *Phys. Rev. X* **7**, 021051 (2017).

[37] Cover, T. M. & Thomas, J. A. *Elements of Information Theory* (John Wiley & Sons, Inc., Hoboken, 2012).

[38] Parrondo, J. M. R., Horowitz, J. M. & Sagawa, T. Thermodynamics of information. *Nat. Phys.* **11**, 131–139 (2015).

[39] Horowitz, J. M. & Esposito, M. Thermodynamics with continuous information flow. *Phys. Rev. X* **4**, 031015 (2014).

[40] Leff, H. S. & Rex, A. F. (eds.) *Maxwell's Demon: Entropy, Information, Computing* (Princeton University Press, Princeton, 1990).

[41] Blackmond, D. G. "if pigs could fly" chemistry: A tutorial on the principle of microscopic reversibility. *Angew. Chem. Int. Ed.* **48**, 2648–2654 (2009).

[42] Rao, R. & Esposito, M. Nonequilibrium thermodynamics of chemical reaction networks: Wisdom from stochastic thermodynamics. *Phys. Rev. X* **6**, 041064 (2016).

[43] Penocchio, E., Rao, R. & Esposito, M. Thermodynamic efficiency in dissipative chemistry. *Nat. Commun.* **10**, 3865 (2019).

[44] Wagoner, J. A. & Dill, K. A. Mechanisms for achieving high speed and efficiency in biomolecular machines. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 5902–5907 (2019).

[45] Howard, J. Protein power strokes. *Curr. Biol.* **16**, R517 (2006).

[46] Hwang, W. & Karplus, M. Structural basis for power stroke vs. brownian ratchet mechanisms of motor proteins. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 19777–19785 (2019).

[47] Astumian, R. D. Irrelevance of the power stroke for the directionality, stopping force, and optimal efficiency of chemically driven molecular machines. *Biophys. J.* **108**, 291–303 (2015).

[48] Pezzato, C., Cheng, C., Stoddart, J. F. & Astumian, R. D. Mastering the non-equilibrium assembly and operation of molecular machines. *Chem. Soc. Rev.* **46**, 5491–5507 (2017).

[49] Das, K., Gabrielli, L. & Prins, L. J. Chemically-fueled self-assembly in biology and chemistry. *Angew. Chem. Int. Ed.* 10.1002/anie.202100274 (2021).

[50] Albaugh, A. & Gingrich, T. R. Simulating a chemically-fueled molecular motor with nonequilibrium molecular dynamics. Preprint at http://arxiv.org/abs/2102.06298 (2021).

[51] Horowitz, J. M., Sagawa, T. & Parrondo, J. M. R. Imitating chemical motors with optimal information motors. *Phys. Rev. Lett.* **111**, 010602 (2013).

[52] Fyfe, M. C. T. *et al.* Anion-assisted self-assembly. *Angew. Chem. Int. Ed. Engl.* **36**, 2068–2070 (1997).

[53] Astumian, R. D. Design principles for brownian molecular machines: how to swim in molasses and walk in a hurricane. *Phys. Chem. Chem. Phys.* **9**, 5067–5083 (2007).

[54] Hartich, D., Barato, A. C. & Seifert, U. Stochastic thermodynamics of bipartite systems: transfer entropy inequalities and a maxwell's demon interpretation. *J. Stat. Mech.: Theory Exp.* **2014**, P02016 (2014).

[55] Kondepudi, D. K. & Prigogine, I. *Modern Thermodynamics: From Heat Engines to Dissipative Structures* (John Wiley & Sons, Ltd., Chichester, 2015).

[56] Astumian, R. & Bier, M. Mechanochemical coupling of the motion of molecular motors to atp hydrolysis. *Biophysical Journal* **70**, 637–653 (1996).

[57] Hill, T. L. *Free energy transduction in biology* (Academic Press, New York, 1977).

[58] Alvarez-Pérez, M., Goldup, S. M., Leigh, D. A. & Slawin, A. M. Z. A chemically-driven molecular information ratchet. *J. Am. Chem. Soc.* **130**, 1836–1838 (2008).

[59] Carlone, A., Goldup, S. M., Lebrasseur, N., Leigh, D. A. & Wilson, A. A three-compartment chemically-driven molecular information ratchet. *J. Am. Chem. Soc.* **134**, 8321–8323 (2012).

[60] Lussis, P. *et al.* A single synthetic small molecule that generates force against a load. *Nat. Nanotechnol.* **6**, 553–557 (2011).

[61] Amano, S., Borsley, S., Leigh, D. A. & Sun, Z. Chemical engines: Driving systems away from equilibrium through catalyst reaction cycles. Submitted for publication (2021).

[62] Kay, E. R., Leigh, D. A. & Zerbetto, F. Synthetic molecular motors and mechanical machines. *Angew. Chem. Int. Ed.* **46**, 72–191 (2007).

[63] Feng, Y. *et al.* Molecular pumps and motors. *J. Am. Chem. Soc.* **143**, 5569–5591 (2021).

nature nanotechnology

PERSPECTIVE

https://doi.org/10.1038/s41565-021-00975-4

Chemical engines: driving systems away from equilibrium through catalyst reaction cycles

Shuntaro Amano[1], Stefan Borsley[1], David A. Leigh[1,2] ✉ and Zhanhu Sun[2]

Biological systems exhibit a range of complex functions at the micro- and nanoscales under non-equilibrium conditions (for example, transportation and motility, temporal control, information processing and so on). Chemists also employ out-of-equilibrium systems, for example in kinetic selection during catalysis, self-replication, dissipative self-assembly and synthetic molecular machinery, and in the form of chemical oscillators. Key to non-equilibrium behaviour are the mechanisms through which systems are able to extract energy from the chemical reactants ('fuel') that drive such processes. In this Perspective we relate different examples of such powering mechanisms using a common conceptual framework. We discuss how reaction cycles can be coupled to other dynamic processes through positive (acceleration) or negative (inhibition) catalysis to provide the thermodynamic impetus for diverse non-equilibrium behaviour, in effect acting as a 'chemical engine'. We explore the way in which the energy released from reaction cycles is harnessed through kinetic selection in a series of what have sometimes been considered somewhat disparate fields (systems chemistry, molecular machinery, dissipative assembly and chemical oscillators), highlight common mechanistic principles and the potential for the synchronization of chemical reaction cycles, and identify future challenges for the invention and application of non-equilibrium systems. Explicit recognition of the use of fuelling reactions to power structural change in catalysts may stimulate the investigation of known catalytic cycles as potential elements for chemical engines, a currently unexplored area of catalysis research.

Biological systems exhibit a broad range of complex functions, from transportation and motility[1–4] to temporal control[5,6] and information processing[7,8], under non-equilibrium conditions realized by dissipating the chemical potential of high-energy species (typically the hydrolysis of ATP or GTP[3–5,7]). Synthetic chemists are creating their own non-equilibrium systems, resulting in kinetic selection in catalysis, self-replicating systems[9,10], dissipative assembly[11–13], synthetic molecular machines[14–17], oscillating systems[18,19] and so on, a number of which are implicated in the emergence of complexity. The understanding of non-equilibrium phenomena originates largely from non-equilibrium thermodynamics[20,21], with some important common features already identified, including kinetic asymmetry[22–24], feedback mechanisms[25–28] and compartmentalization[25,27]. In this Perspective we consider catalysed reaction cycles as another ubiquitous feature of chemically driven non-equilibrium processes[13]. Reaction cycles harvest, and enable the utilization of, free energy supplied from a chemical fuel. When they are coupled to another dynamic process they form a chemical engine that powers function by transducing chemical energy into other forms (Fig. 1a; see Box 1 for the definition of some terms in the context used in this Perspective). Different non-equilibrium systems can be analysed in terms of factors influenced by and affecting such a chemical engine, such as coupling[29,30] and synchronization[31] of the reaction cycle (Fig. 1b and Box 1). Generally, different non-equilibrium systems (such as dissipative assembly, self-replicators, molecular motors and oscillators, Fig. 1b) have been analysed through different specific frameworks for different systems and processes. The holistic approach used in this Perspective is consistent with these previous analyses, but offers further insights by highlighting common elements, such as catalyst reaction cycles, coupling, kinetic selection and synchronization. This overview sometimes requires a degree of simplification (details of some processes are not always explicitly discussed), but such a concession has the advantage of allowing the broad mechanistic features of a system to be more easily related to others. It can also suggest a role for elements that were not originally considered in a mechanism (for example, Fig. 5a). We focus on processes driven by chemical reaction cycles, and omit light[32]- and diffusion[33,34]-driven systems as they have certain characteristics that differ from those of purely chemically driven systems. We conclude with a short discussion of common issues and challenges raised by this unifying treatment.

The chemical engine

Let us define a chemical engine as a network consisting of a reaction cycle of a catalyst ($\mathbf{Cat} \rightarrow [+\mathbf{F}] \rightarrow \mathbf{Cat}' \rightarrow [-\mathbf{W}] \rightarrow \mathbf{Cat}$) coupled to another dynamic process ($\mathbf{Cat_a} \rightleftharpoons \mathbf{Cat_b}$ and $\mathbf{Cat_a}' \rightleftharpoons \mathbf{Cat_b}'$) that typically changes or biases the mechanical or structural state of the catalysing species (Fig. 1a and Box 1). With such a module, the chemical energy harvested from the fuelling reaction is transduced to the coupled process, enabling work to be performed, such as directional motion, mechanical work, assembly, the storage of energy or the synchronization of chemical reaction cycles.

The chemical engine module comprises the following elements: chemical reaction cycles, fuel, catalysis, coupling and kinetic selection. Chemical reaction cycles are well established in the context of catalysis (Box 1). However, instead of viewing a catalysed reaction as the chemical transformation of a starting material to a product, it can be viewed as cycling of the catalytic species through two (or more) states with net directionality (Fig. 2a, left). From this viewpoint the starting material can be regarded as a fuel and the product of the catalysed reaction as waste. The conversion of fuel into waste supplies the free energy used to drive a coupled process involving the catalyst away from equilibrium. For example, biological molecular machines catalyse the hydrolysis of ATP and

[1]Department of Chemistry, University of Manchester, Manchester, UK. [2]School of Chemistry and Molecular Engineering, East China Normal University, Shanghai, China. ✉e-mail: david.leigh@manchester.ac.uk

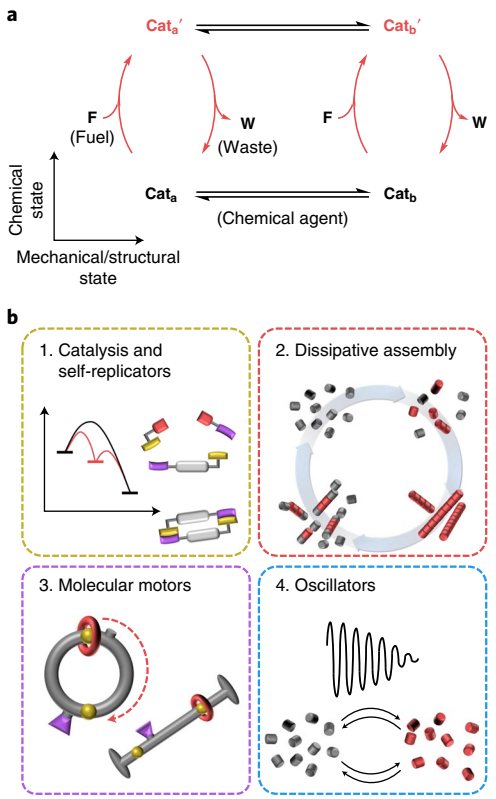


Fig. 1 | Driving systems with chemical engines. a, General scheme for a chemical engine based on a fuelled reaction cycle (**Cat** → [+**F**] → **Cat**′ → [−**W**] → **Cat**) coupled to a chemical agent (**Cat**) that undergoes a mechanical or structural dynamic change of state (**Cat$_a$** ⇌ **Cat$_b$**). Reaction cycles mediated by **Cat** provide a pathway for a fuel (**F**) to waste (**W**) reaction via a modified species (**Cat**′), where the prime denotes a state of the chemical agent bound to or reacted with the fuel. These reaction pathways can be coupled to dynamic exchange processes involving the chemical agent (**Cat$_a$** ⇌ **Cat$_b$** and **Cat$_a$**′ ⇌ **Cat$_b$**′), where subscript letters indicate the dynamic mechanical/structural state of the agent. We use the label **Cat** for the chemical agent as it is often a positive catalyst (that is, accelerates the fuel-to-waste reaction, Box 1). Kinetic parameters determine the behaviour of the chemical engine and how it drives non-equilibrium behaviour. Although the direct reverse of each pathway is accessible in line with the principle of microscopic reversibility, the reverse pathways (which are often extremely rare events), such as **Cat$_a$**′ → **Cat$_a$** converting waste to fuel, are omitted for clarity. **b**, Some examples of non-equilibrium systems driven by elements of a chemical engine.

in doing so extract the energy from this fuelling reaction to perform tasks[1,3,4]. This perspective of viewing the chemical reaction cycle to supply free energy to the system as a catalytic cycle has typically been associated with an information ratchet[14,35], which is one mechanism to drive molecular motors[23]. However, chemical reaction cycles within other systems can also be viewed as catalytic cycles. The recognition that such modules use fuelling reactions to power structural change in a catalyst suggests that already well understood catalytic cycles may prove new and effective elements for chemical engines.

Biological catalysts are generally highly efficient; almost no energy is wasted by way of a competing background fuel-to-waste reaction and the chemical-agent-mediated pathway is much lower in activation energy than the background rate (Fig. 2a, middle). However, reaction cycles that are coupled to follow higher-energy pathways (so-called 'negative catalysis'[36]) can also be used to harness the energy from a fuelling reaction (Fig. 2a, right). An example of this is a chemical fuel (9-fluorenylmethoxycarbonyl (Fmoc) nitrophenol ester) that autonomously powers a molecular pump, but is decomposed to waste products faster in the absence of the molecular machine[37]. An unusual type of chemical reaction cycle is found in self-replicators[9,38], whereby the product generated in an autocatalytic reaction cycle contains or constitutes a catalyst that promotes subsequent cycles. The positive feedback mechanism observed in autocatalysis also plays an important role in oscillating systems, examined later in this Perspective (Fig. 5).

The free energy released from the reactant-to-product conversion ($-\Delta G > 0$) of any type of chemical reaction can potentially be used to do work. What converts a chemical reaction to a fuelling reaction is a mechanism that allows that energy difference to be used to power another, coupled process. A mechanism for such free-energy transduction is kinetic selection (Box 1), which has already been utilized in some artificial molecular machines[37,39–42]. For example, the dynamic kinetic resolution of a mechanical point-chiral rotaxane shuttle[39] results in more rings trapped in the 'forward' (right-hand-side) compartment rather than the 'backward' (left-hand-side) one, constituting a kinetically robust non-equilibrium distribution (Fig. 2b). The ratcheted distribution is determined by the difference between the transition-state energies of two barrier attachment processes ($\Delta\Delta G^{\ddagger}$ in Fig. 2b), which would include the energy difference between the two co-conformers[43] if they differ[39]. This example couples (see Box 1 for the definition of 'coupling') macrocycle shuttling only to barrier formation (effectively the fuel-addition part of Fig. 1a), rather than to a full reaction cycle, although a full cycle was subsequently developed to generate an autonomous chemically fuelled rotary motor[41] (Fig. 4b). Nonetheless, this [2]rotaxane illustrates how two kinetically controlled reactions (barrier attachment to either mechanical state (co-conformation[43]) of the rotaxane) can be linked[40] to a dynamically reversible process (macrocycle shuttling) to bring about a mechanical bias (Fig. 2b, bottom).

Another striking combination of chemical reaction cycles with kinetic selection is the Soai reaction (Fig. 2c)[44–47]. Amplification of the major enantiomer occurs due to kinetic selection between two autocatalytic cycles catalysed by enantiomeric homochiral oligomers and the formation of catalytically inactive heterochiral oligomers that siphon off the minor enantiomer; effectively, the dominant enantiomer suppresses the formation of its mirror image.

In both of these examples, kinetic selection leads to a biased distribution in the products, which has a higher free energy than an unbiased distribution due to lower entropy. Thus, both the rotaxane ratchet (Fig. 2b) and Soai reactions can be recognized as being powered through similar elements of a chemical engine mechanism: kinetic selection enables part of the free energy supplied from the fuel-to-waste conversion ($-\Delta G$) to be converted to the free energy of the system in the form of a biased product distribution. Related selection and amplification phenomena are common in biology[5–8].

Kinetic selection also allows free-energy transduction from chemical reaction cycles to couple processes in other ways. In the dissipative assembly and information ratchet systems shown in Figs. 3a and 4b, the free energy released through fuel-to-waste conversion (*y* axis in Fig. 1a) can only be used to drive the coupled process (*x* axis in Fig. 1a) when the fuel adds preferentially to the catalyst in one mechanical/structural state (**Cat$_a$** in Fig. 1a) and/or waste products are released preferentially from a different mechanical/structural state (**Cat$_b$**′ in Fig. 1a)[22–24]. If catalysis of the fuel-to-waste reaction occurs through only one mechanical/structural state of the catalyst (for example, **Cat$_a$** → **Cat$_a$**′ in Fig. 1a) then the free energy released from the catalysed fuel-to-waste reaction cannot be used to power a mechanical/structural change involving the catalyst. The bias in the fuel-adding and/or waste-releasing reactions arises

Box 1 | Terminology used in this Perspective

Chemical reaction cycle
Process through which a particular chemical agent (**Cat**) reacts with reactants (fuel) and releases products (waste), regenerating the original chemical agent[13,80] (Fig. 2a). The chemical agent provides an alternative pathway to the background fuel-to-waste reaction. Often the chemical agent is a positive catalyst, that is, accelerates the fuel-to-waste reaction so that the agent-mediated pathway dominates the background reaction. However, this is not strictly necessary and the chemical agent can also be an inhibitor (negative catalyst) of the background reaction.

Autocatalytic cycle
Subcategory of chemical reaction cycle in which the released product is itself a catalyst for the cycle.

Chemical fuel and waste
Reactants (fuel) and products (waste) in the chemical reaction cycle, where the reactants supply free energy to drive the chemical agent states out of equilibrium and the products are emitted from the cycle during operation.

Kinetic selection
Where several reaction pathways are available, kinetic selection causes reaction preferentially via the pathway with the lowest activation energy.

Coupling
Mechanism that connects chemical reaction cycles with other processes, such as assembly or other structural or mechanical (conformational or co-conformational) processes. The mechanism transduces free energy harvested from chemical reaction cycles to the connected processes, enabling behaviour such as dissipative assembly or directional motion of the components of molecular motors. In the field of biological molecular machines, this is often referred to as chemomechanical or mechanochemical coupling[29,30].

Chemical engine
Module consisting of chemical reaction cycles coupled with (an) other dynamic process(es) (Fig. 1a). It transduces chemical free energy harvested from the chemical reaction cycles to other forms of energy, such as a biased (that is, non-equilibrium) distribution of species, increase of the system's free energy or performed mechanical or structural work.

Assembly process
Process in which multiple monomers aggregate. The aggregate species is not necessarily thermodynamically stable: when coupled with chemical reaction cycles, aggregation to form thermodynamically unfavourable species can take place, termed 'dissipative assembly'.

Mechanical process
Process that involves controlled, large-amplitude or directional displacement of one component of the system relative to another[14], such as the co-conformational change that occurs in shuttling of a macrocycle in a [2]rotaxane. 'Co-conformation' refers to the relative positions of non-covalently linked components with respect to each other[43].

Synchronization
Mechanism that causes the transition between states in the chemical reaction cycle to occur concurrently[31].

from the difference in reaction rates (activation energies) between the mechanical/structural states and determines the overall kinetic asymmetry of the process. Additionally, for fuelling to do work on a coupled process progressively (that is, more fuel can be consumed to add to, or in some cases sustain, previously done work) the fuel-to-waste reaction (*y* axis in Fig. 1a) and the coupled process (*x* axis in Fig. 1a) need to form a complete cycle (Fig. 1a). Exploring the free-energy transduction mechanism in transient (co-)conformational switching (Fig. 3b), energy ratchet[14,35] (Fig. 4a) and oscillating systems (Fig. 5) is an important future challenge for theory to aid the understanding and design of such systems.

The examples in Fig. 2 feature the coupling of an effectively irreversible reaction (resulting in barrier formation) with a dynamic exchange process (macrocycle shuttling), and kinetic selection between two reaction cycles that share common reactants. The chemical reaction cycles generate a non-equilibrium product distribution that once formed is kinetically stable. For the remainder of this Perspective we consider examples where complete reaction cycles are coupled with dynamic exchange processes, which is the complete module of what we term a chemical engine.

Dissipative assembly and (co-)conformational switching

Dissipative assembly refers to systems where a chemical reaction cycle is coupled to an assembly process, so that construction of the structure is first driven, then maintained, by a continuous input of energy[22–24]. In biology dissipative assembly is responsible for functions such as cell division[11,13]. Chemical reaction cycles for dissipative assembly typically involve two steps: (1) formation of activated monomers and (2) their deactivation. Activated monomers assemble to form aggregates, for example, fibres and gels. Through coupling of the reaction cycles and assembly process, free energy harvested by the former is supplied to the latter. The distribution of species is driven out of equilibrium by the chemical engine, enabling transient formation of the high-energy structure. While chemical reaction cycles for dissipative assembly have generally been regarded as a combination of the activation and deactivation of monomers, they can be perceived also as catalytic reaction cycles: they convert chemical fuel to waste through catalysis by the assembling species[13]. This, again, allows us to envisage such processes in terms of being powered by a chemical engine module.

A classic example is the methyl iodide-fuelled dissipative assembly of dicarboxylate salts/esters into fibres reported by the group of van Esch[48] (Fig. 3a). The chemical reaction cycle consists of the alkylation of carboxylates $\mathbf{1_a}$ or $\mathbf{1_b}$ with methyl iodide to form activated species $\mathbf{1_a}'$ or $\mathbf{1_b}'$, respectively, followed by hydrolysis of the esters. The diesters form fibres (leading to phase separation from solution), whereas the carboxylate salts do not, and so the reaction cycle is coupled to the assembly process. Note that kinetic asymmetry[23] in this process is also crucial; simply having reaction cycles connected to assembly processes is not sufficient for free-energy transduction. As the diesters are formed under conditions where they also hydrolyse, the fibres are transient, present only until the methyl iodide fuel is depleted through its consumption in the reaction cycle of carboxylate alkylation followed by hydrolysis. Activated aggregates can be obtained even when some species or processes in the chemical engine responsible for this assembly are absent. For example, even if the aggregate of deactivated species ($\mathbf{Cat_b}$ in Fig. 1a) and processes involving it are missing, monomers ($\mathbf{Cat_a}$) can still be activated to the state that allows assembly ($\mathbf{Cat_a}'$) by the fuel addition, resulting in the formation of aggregates ($\mathbf{Cat_b}'$). This distinction between self-assembly under dissipative conditions (where the activated

species undergo local thermodynamic exchange) versus driven assembly (where the states are cycled directionally through the chemical engine) has been noted[22,24]. Phenomenologically the two outcomes may appear similar, and, in practice, it can be difficult to establish which pathways within the chemical engine are operational within such systems.

Chemical reaction cycles can be coupled to mechanical molecular-level processes that result in dissipative promotion of a conformation or co-conformation[43] that performs a function. An example is fuel-activated transient catalysis by a rotaxane-based molecular shuttle[49] (Fig. 3b). Here the chemical reaction cycle consists of protonation of the amine of rotaxane $\mathbf{2}_a$ by the pulsed trichloroacetic acid fuel[50,51], followed by decarboxylation of trichloroacetate to generate carbon dioxide and chloroform, resulting in deprotonation of the ammonium group of the rotaxane. When the axle amine group is not protonated the crown ether preferentially binds the thiourea unit ($\mathbf{2}_a$), preventing the thiourea from participating in catalysis. However, the protonated form of the rotaxane favours the co-conformation with the macrocycle positioned over the dibenzylammonium station ($\mathbf{2}_b'$). Accordingly, protonation of the rotaxane by the trichloroacetic acid fuel displaces the macrocycle, revealing the thiourea group and enabling it to catalyse the hydrogenation of β-nitrostyrene. The rotaxane functions as a catalyst for this reaction until the fuel pulse is depleted. Note that in this example the trichloroacetic acid fuel is used to transiently change the nature of the environment, rather than the machine within a constant environment, so that the catalytic active state of the machine becomes thermodynamically favoured under the switched environmental conditions. This needs to be accomplished with pulses of fuel (each pulse can perform a single reaction cycle) rather than an autonomous or continuous fuel supply (which perform reaction cycles as long as fuel is present).

A common feature of dissipative assembly and dissipative (co-) conformational switching is that the process coupled to and fuelled by reaction cycles (assembly of the monomers into fibres/establishment of the catalytically active rotaxane co-conformation) is undone once the supply of fuel is depleted, because transient species are kinetically labile. However, if these species are kinetically stable, they will persist even after the fuel runs out. In biology there are numerous examples of exergonic reactions (for example, hydrolysis of ATP) being coupled to (in order to drive) endergonic processes in this way (for example, synthesis of the amino acid glutamine from glutamic acid and ammonia)[52]. The principles for dissipative

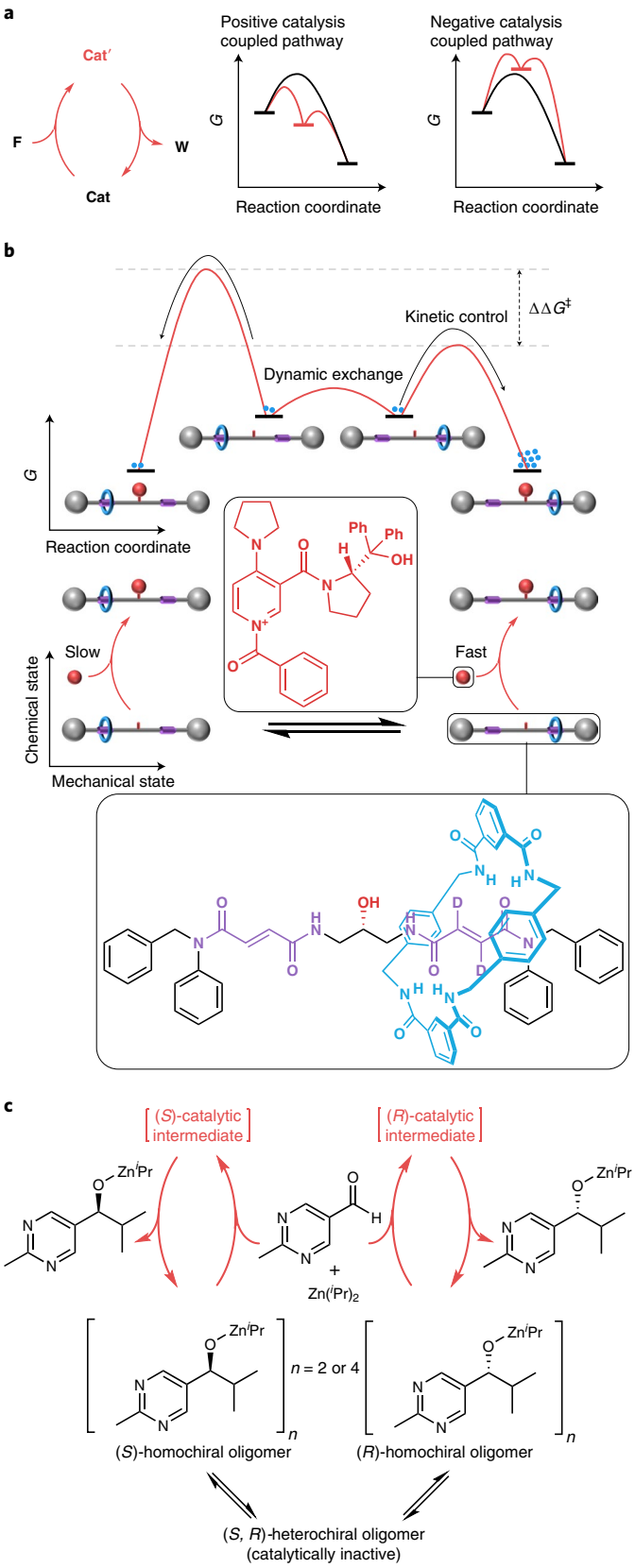


Fig. 2 | Chemical reaction cycles and kinetic selection. a, Extracting energy from chemical reaction cycles to drive non-equilibrium processes. The conversion of **F** to **W** via a pathway mediated by **Cat** enables the energy difference between **F** and **W** to be harnessed. The chemical-agent-mediated reaction (red pathway) may be either lower (positive catalysis, middle) or higher (negative catalysis, right) in energy than the background reaction (black pathway). **b**, Kinetic resolution of a rotaxane-based molecular shuttle operating via an information ratchet mechanism[39]. Top: a prochiral alcohol reactive site (red) in the middle of the rotaxane thread is a site for barrier (red ball) introduction, trapping the macrocycle (blue) on either side. Introduction of the barrier using a chiral electrophile causes preferential formation of the rotaxane with the macrocycle in one compartment (here shown as the right-hand compartment), due to the energy difference of the diastereomeric transition states $\Delta\Delta G^{\ddagger}$. Bottom: kinetic selection is achieved by coupling the barrier formation with the macrocycle shuttling process, a classic Curtin–Hammett principle[78] scenario. **c**, Self-replication within the Soai reaction[44]. In such systems the product is also the catalyst for the reaction cycle. Both enantiomeric cycles of the reaction catalysed by homochiral oligomers use the same achiral reagents. A stochastic excess of one enantiomer in the early stages of the reaction causes amplification of that species if the catalytic activity of the other enantiomer is suppressed by formation of heterochiral oligomers. After further cycles, this amplification ultimately leads to a substantially enantio-enriched product.

assembly and these mechanisms for synthesizing high-energy species are strikingly similar.

Reaction cycles can also be used to power actions that progressively build on the work done previously as additional molecules of fuel are consumed. These systems are ratchets[14], rather than simple switches[53], and are most familiar in the form of molecular motors, although the link to systems chemistry and driving exchanging distributions away from equilibrium was made[54] more than a decade ago. Theory shows that a condition for obtaining non-equilibrium distribution in dissipative assembly is equivalent to a condition for obtaining net directional motion in molecular ratchets, which is the 'ratcheting constant', K_r, being not equal to one[22,24]. This connection means that insights obtained from operating molecular ratchets, such as the utilization of catalytic cycles for supplying energy and various ways of introducing kinetic selection, should be applicable to the design of dissipative assembly systems, and vice versa. For example, increasing the degree of kinetic selection in dissipative assembly could push the distribution further away from equilibrium, with more of the high-energy species formed and sustained.

Molecular motors

Molecular motors are intrinsic to numerous processes in biology, including muscle contraction and intracellular transport[1–4]. Among other functions, they couple chemical reaction cycles with mechanical processes[29,30] to progressively generate directional motion. A key difference between chemically fuelled molecular motors and the dissipative rotaxane switch in Fig. 3b is that a molecular motor kinetically traps (ratchets) in work done during the fuel-promoted operation cycle[14]. This is manifested in a linear motor, such as kinesin, moving itself progressively further along a track with each step it takes, and, hypothetically, for every 360° directional rotation of a rotary motor, which returns the rotor to its starting position, being used to wind up more of an attached polymer chain.

The reaction cycle fuelled by the decomposition of trichloroacetic acid in the fuel-activated catalytic rotaxane (Fig. 3b) has also been used as an energy ratchet[55] (a type of Brownian ratchet mechanism[14,29,35,53]) to power a molecular motor (Fig. 4a). Trichloroacetic acid simultaneously protonates the track amine group and labilizes the hydrazone barrier, causing translocation of the crown ether from $\mathbf{3}_a$ to $\mathbf{3}_b'$ through the left-hand pathway via $\mathbf{3}_a'$. Decarboxylation of the fuel basifies the media, providing a thermodynamic driving force for the relocation of the crown ether to the methyl triazolium group. Under these conditions the hydrazone barrier is kinetically locked while the disulfide barrier is labilized, so the macrocycle returns to its original site via the right-hand pathway ($\mathbf{3}_b'$ to $\mathbf{3}_a$ via $\mathbf{3}_b$). One pulse of trichloroacetic acid fuel therefore results in 360° clockwise rotation of the crown ether around the larger macrocycle. Notably, the motion (or cycling around the chemical engine) is highly synchronized for this example; however, continuous autonomous operation is not realized due to the requirement to pulse the system with trichloroacetic acid with each cycle.

The coupling of the reaction cycle and macrocycle shuttling in the pulsed-fuel system shown in Fig. 4a has several noteworthy features and consequences. Although it does not operate autonomously, the pulsed-fuel requirement synchronizes the positions of the components. This feature may be adaptable to other types of system (dissipative assembly and so on) that employ reaction cycles to drive chemical engines. However, unlike the pulsed-fuel motor, biological molecular motors operate autonomously as long as fuel (typically ATP or GTP) is present, and each motor can be at a different stage in a reaction cycle from others in the ensemble. This is achieved through a different type of ratcheting mechanism, termed an information ratchet[14,23,35,39–42,56]. Unlike an energy ratchet[14,35,53], kinetic selection based on the chemomechanical state of the information ratchet drives the system away from equilibrium.

An example of an artificial chemically fuelled information ratchet is a catenane rotary motor that mediates the reaction cycle of the decomposition of Fmoc-Cl[41] (Fig. 4b). In this motor, the chemical reaction cycle consists of attachment of an Fmoc barrier to free hydroxyl groups, blocking translocation of the macrocycle, and base-induced decomposition of the Fmoc group, removing the barrier. The overall reaction is conversion of Fmoc-Cl to dibenzofulvene, carbon dioxide and hydrochloric acid (neutralized by the base that is present). The chemical reaction cycle is coupled with the mechanical translocation of the macrocycle: when an Fmoc group is removed from the track, the macrocycle can shuttle between the two fumaramide sites. The co-conformer with the macrocycle distal to the free hydroxyl group reacts with the Fmoc electrophile faster than the co-conformer with the macrocycle close to the hydroxyl group, because of less steric hindrance. This results in kinetic selection in terms of Fmoc attachment, similar to the kinetic resolution of the rotaxane shown in Fig. 2b. Although the Fmoc removal process is not kinetically selective in terms of macrocycle position, kinetic selection during Fmoc attachment is sufficient to supply free energy to the mechanical processes, enabling directional biasing of macrocycle shuttling as long as Fmoc-Cl fuel remains. Clearly, introducing kinetic selection in the barrier removal process could improve motor performance (double kinetic gating)[42].

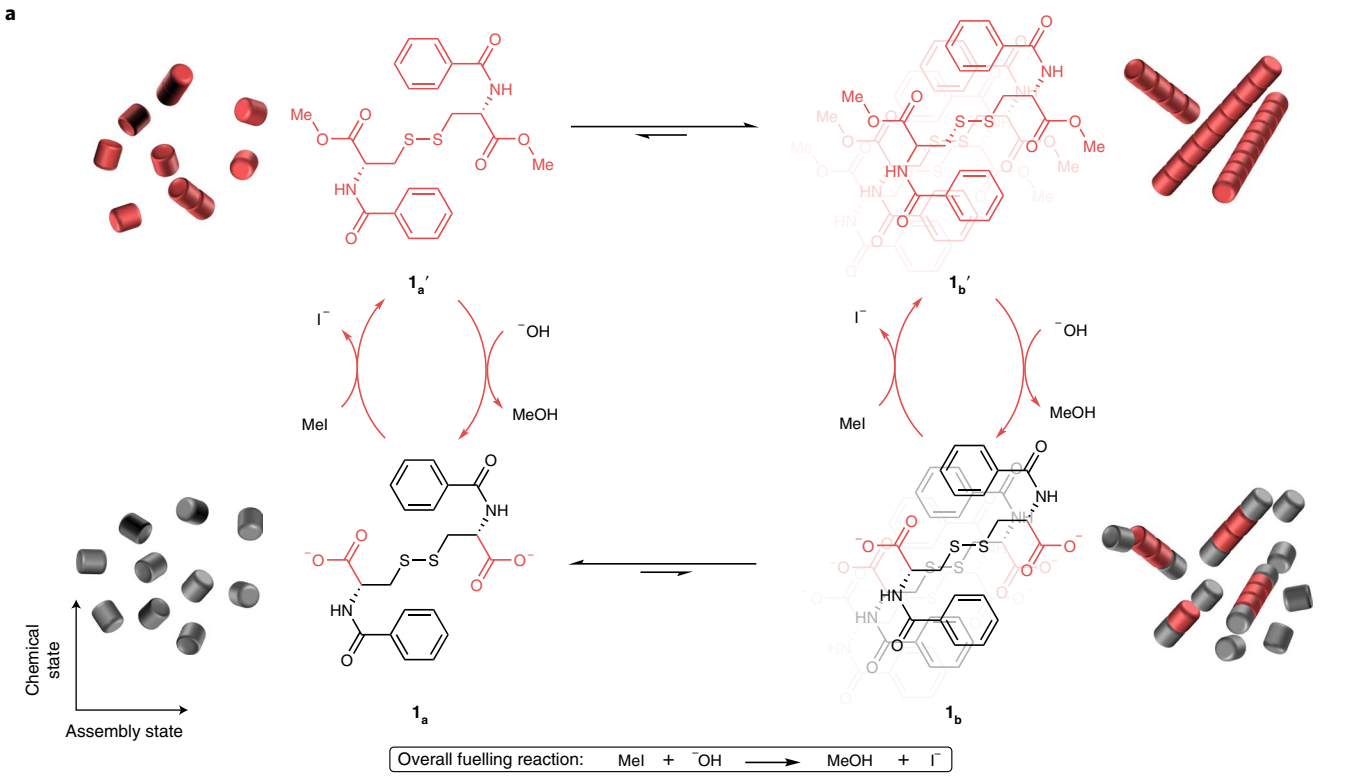
Both the energy ratchet and the information ratchet described here are driven by catalytic reaction cycles. In the information ratchet the catalytic cycle directly supplies free energy to drive the movement, while in the energy ratchet the catalytic cycle transiently changes the free-energy landscape within the kinetically defined area[14,35]. The catalytic cycle is a common element of the two systems, and recognizing such commonality aids the understanding of how a catalytic reaction cycle from one system can be used to drive another: for example, the decomposition of trichloroacetic acid in driving both transient catalysis (Fig. 3b) and an energy ratchet (Fig. 4a). Another example is the utilization of carbodiimide hydration, originally used to drive dissipative assembly[57–59], to drive an information ratchet[42]. The introduction of this efficient fuel from the field of dissipative assembly is spurring advances in molecular machine design that were not possible with earlier fuel systems.

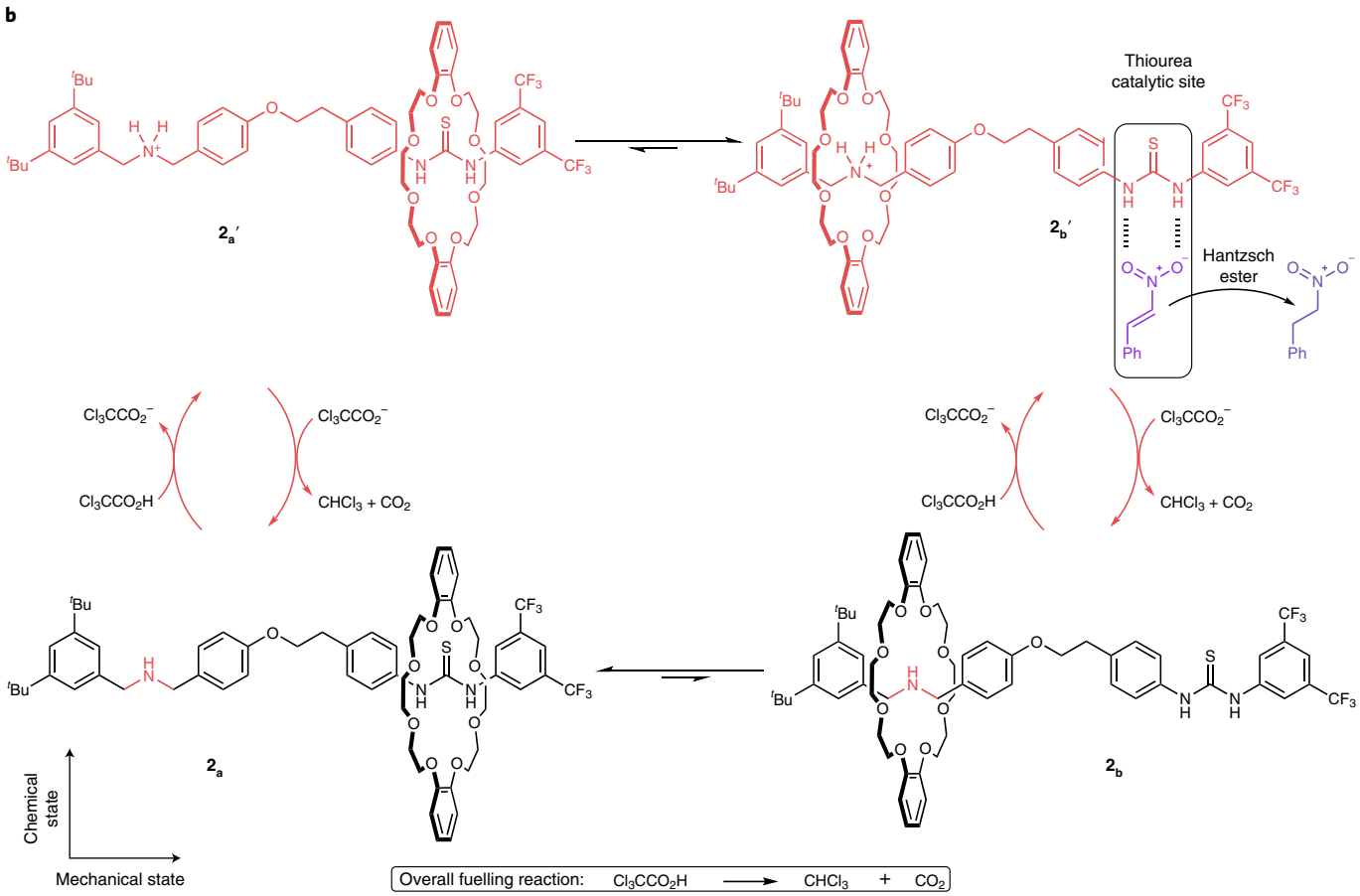
Oscillating systems

Oscillating systems in biology include processes such as circadian clocks and morphogenesis[5,6]. Since the discovery of the Belousov–

Fig. 3 | Dissipative systems fuelled by chemical reaction cycles. a, The hydrolysis of MeI under basic conditions to generate MeOH and I^- fuels the dissipative assembly of a dibenzoyl-L-cystine monomer[48]. Monomer $\mathbf{1}_a$ is methylated with MeI fuel to generate activated species $\mathbf{1}_a'$. Hydrolysis of $\mathbf{1}_a'$ regenerates $\mathbf{1}_a$. This chemical reaction cycle is coupled to an assembly process that favours the assembly of $\mathbf{1}_a'$ to form aggregate $\mathbf{1}_b'$ and the disassembly of $\mathbf{1}_b$ to form $\mathbf{1}_a$. This overall process sustains the concentration of high-energy aggregates until the fuel runs out. For clarity, the activated species is represented here as a dimethyl ester, though subsequent work under increased buffer concentration resulted in assembly of the monomethyl ester[79]. **b**, Base-catalysed decomposition of trichloroacetic acid to generate chloroform and carbon dioxide fuels the displacement of the macrocycle to reveal a catalytically active thiourea unit[49]. Rotaxane $\mathbf{2}_a$ is protonated with trichloroacetic acid fuel to generate $\mathbf{2}_a'$. Decarboxylation of trichloroacetate and subsequent deprotonation of $\mathbf{2}_a'$ regenerates $\mathbf{2}_a$. This chemical reaction cycle is coupled to the position of the macrocycle on the axle, which favours ring occupancy of the dibenzylammonium site in the protonated form, $\mathbf{2}_b'$, and the thiourea group in the amine form, $\mathbf{2}_a$. This overall process sustains the catalytically active species $\mathbf{2}_b'$ until the fuel pulse is depleted.

Zhabotinsky reaction[60], various synthetic oscillating systems have been reported[18,61–63]. The mechanism of oscillating systems is generally considered in terms of feedback processes, such as a combination of positive and delayed negative feedback[5]. However, we can relate oscillating systems to other chemically driven non-equilibrium systems by considering them also as synchronized chemical reaction





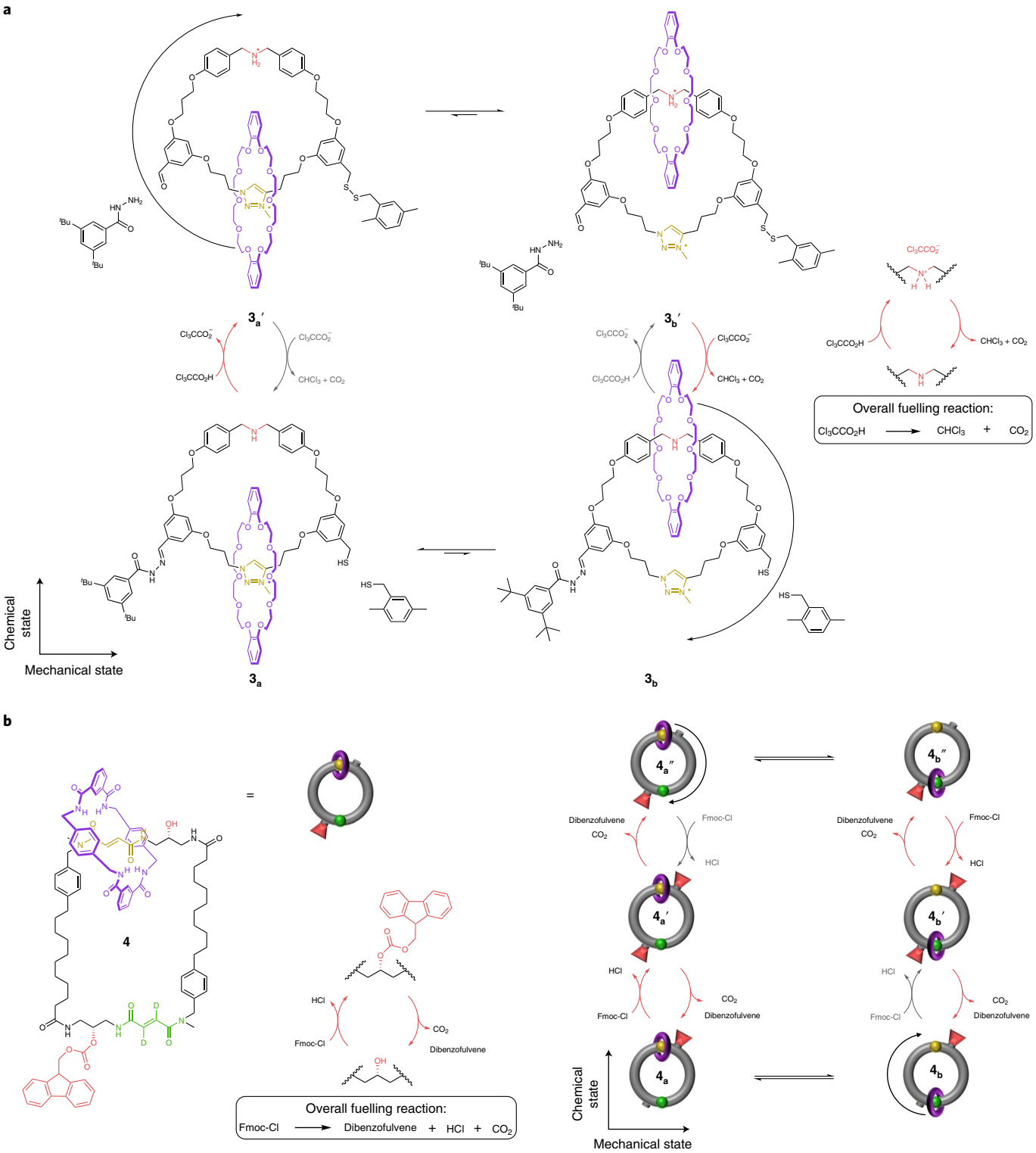


Fig. 4 | Molecular motors fuelled by chemical reaction cycles. a, A pulse of trichloroacetic acid fuel causes 360° clockwise rotation of the crown ether (purple) around the track through an energy ratchet mechanism[55]. The base-catalysed decomposition of trichloroacetic acid to generate chloroform and carbon dioxide fuels the rotation. [2]Catenane $\mathbf{3}_a$ is protonated with trichloroacetic acid fuel. The acidic conditions also labilize the hydrazone barrier and kinetically lock the disulfide barrier, generating $\mathbf{3}_a'$. Decarboxylation of trichloroacetate deprotonates $\mathbf{3}_a'$. The basic conditions also labilize the disulfide barrier and kinetically lock the hydrazone barrier, regenerating $\mathbf{3}_a$. This chemical reaction cycle is coupled to the mechanical movement of the macrocycle, which favours the occupancy of the dibenzylammonium site (red) when protonated ($\mathbf{3}_b'$), and the triazolium site (yellow) when deprotonated ($\mathbf{3}_a$). **b**, Fmoc-Cl fuels 360° directional rotation of the benzylic amide macrocycle (purple) around the track through an information ratchet mechanism[41]. The overall fuelling reaction is decomposition of Fmoc-Cl to generate dibenzofulvene, carbon dioxide and hydrochloric acid. The macrocycle can shuttle between the two fumaramide sites (yellow and green) along the unblocked pathway between $\mathbf{4}_a$ and $\mathbf{4}_b$ or between $\mathbf{4}_a''$ and $\mathbf{4}_b''$. Fmoc attachment occurs faster with $\mathbf{4}_a$ and $\mathbf{4}_b''$ than $\mathbf{4}_b$ and $\mathbf{4}_a''$, whereas the Fmoc removal reactions from $\mathbf{4}_a'$ and $\mathbf{4}_b'$ occur at the same rate. This kinetic selection leads to the directional rotation of the macrocycle through coupling of the chemical reaction cycles with the mechanical processes.

cycles[31] (Fig. 5) in which synchronization causes a dynamic transition between states that occurs concurrently for many molecules. This oscillation requires the input of free energy from the catalytic reaction cycles. The role of feedback mechanisms is to enable synchronization, leading to oscillations that are macroscopically observable. This is unlike the other non-equilibrium systems discussed in this Perspective, where transitions occur stochastically for each molecule (information ratchets and dissipative assembly) or are synchronized at each fuel addition (pulsed-fuel switches and energy ratchets).

Synthetic oscillators[60–65] are often characterized by complex chemical networks[64,65], but the supramolecular oscillator recently reported by Hermans' group[63] is comparatively simple. It oscillates the redox state, and consequently the assembly state, of a perylene diimide derivative (Fig. 5a). The system is reminiscent of a dissipative assembly network in which a chemical reaction cycle is coupled to an assembly process dependent on the charge state of the monomer (cf. Fig. 3a). However, for the system to oscillate the chemical reaction cycles and the assembly processes are synchronized as a result of feedback mechanisms. As neutral monomer $\mathbf{5}_a'$ assembles, the supramolecular stacks fragment to seed and accelerate further assembly, resulting in positive feedback. Reduction and disassembly of aggregate $\mathbf{5}_b'$ to regenerate $\mathbf{5}_a$ induce negative feedback in the assembly processes. The reduction and disassembly are size dependent, accelerating as the size of the aggregates decreases. The feedback processes in this system are a consequence of the assembly state of the monomer, with positive feedback occurring by a nucleation–elongation–fragmentation process, and negative feedback by the size-dependent reduction and disassembly. The oscillations drive the system through all four states of the chemical engine, unlike the simpler dissipative assembly, where the absence/inaccessibility of one of the states can result in phenomenologically the same dissipative behaviour (self-assembly under dissipative conditions[22,24]). In the originally proposed model[63], reduction was only considered in the assembled state and oxidation in the monomeric state. However, it seems likely that redox processes can take place in both states, given that both reductant and oxidant are always present. The chemical engine framework suggests that these processes, which were not explicit in the conventional model, may play a role in this oscillator, an observation that could be useful for the design of future systems.

The prototypical chemical engine diagram can also be used to describe the pH oscillator of Taylor's group[61] (Fig. 5b). The positive feedback by an autocatalytic generation of hydroxide (from $\mathbf{6}_a$ to $\mathbf{6}_a'$) and the delayed negative feedback by the hydrolysis of gluconolactone that consumes hydroxide (from $\mathbf{6}_a'$ to $\mathbf{6}_b'$) results in pH oscillation of the system. It is unclear to what extent the protonation of $\mathbf{6}_b'$ to generate $\mathbf{6}_b$ or the lactonization of $\mathbf{6}_b$ to regenerate $\mathbf{6}_a$ occurs under the reported reaction conditions. However, description with a chemical engine diagram suggests that incorporating these two processes may result in a completion of the oscillation cycle. Unlike Hermans' oscillator, Taylor's pH oscillator relies on chemical autocatalysis, rather than autocatalytic assembly processes, to generate feedback. Illustration of these systems with a chemical engine diagram indicates how feedback mechanisms could be built into various chemical-engine-powered phenomena to achieve synchronization. Indeed, Hermans' oscillator was realized by incorporating the elements of autocatalysis and feedback within an existing dissipative assembly system[66], illustrating the efficacy possible through such considerations.

There is much still to be understood about the effects of feedback loops in oscillating systems. However, as relatively simple oscillators can be described in terms of chemical engines, it may be that similar motifs are also components of more complex oscillating networks. There are generally considered to be four requirements for biochemical oscillations: negative feedback, time delay, sufficient nonlinearity of the reaction kinetics and proper balancing of the timescales of opposing reactions[5]. Positive feedback is one way to delay negative feedback, but not a crucial requirement (similar effects can be achieved through compartmentalization). Given the similarities between the mechanisms of oscillators (Fig. 5), dissipative assembly (Fig. 3) and molecular motors (Fig. 4), it seems likely that incorporating feedback mechanisms could be used to synchronize the operation of molecular motors. Examples of synchronized machines in biology include skeletal myosins, which synchronize in response to high loads to exert a stronger force[67].

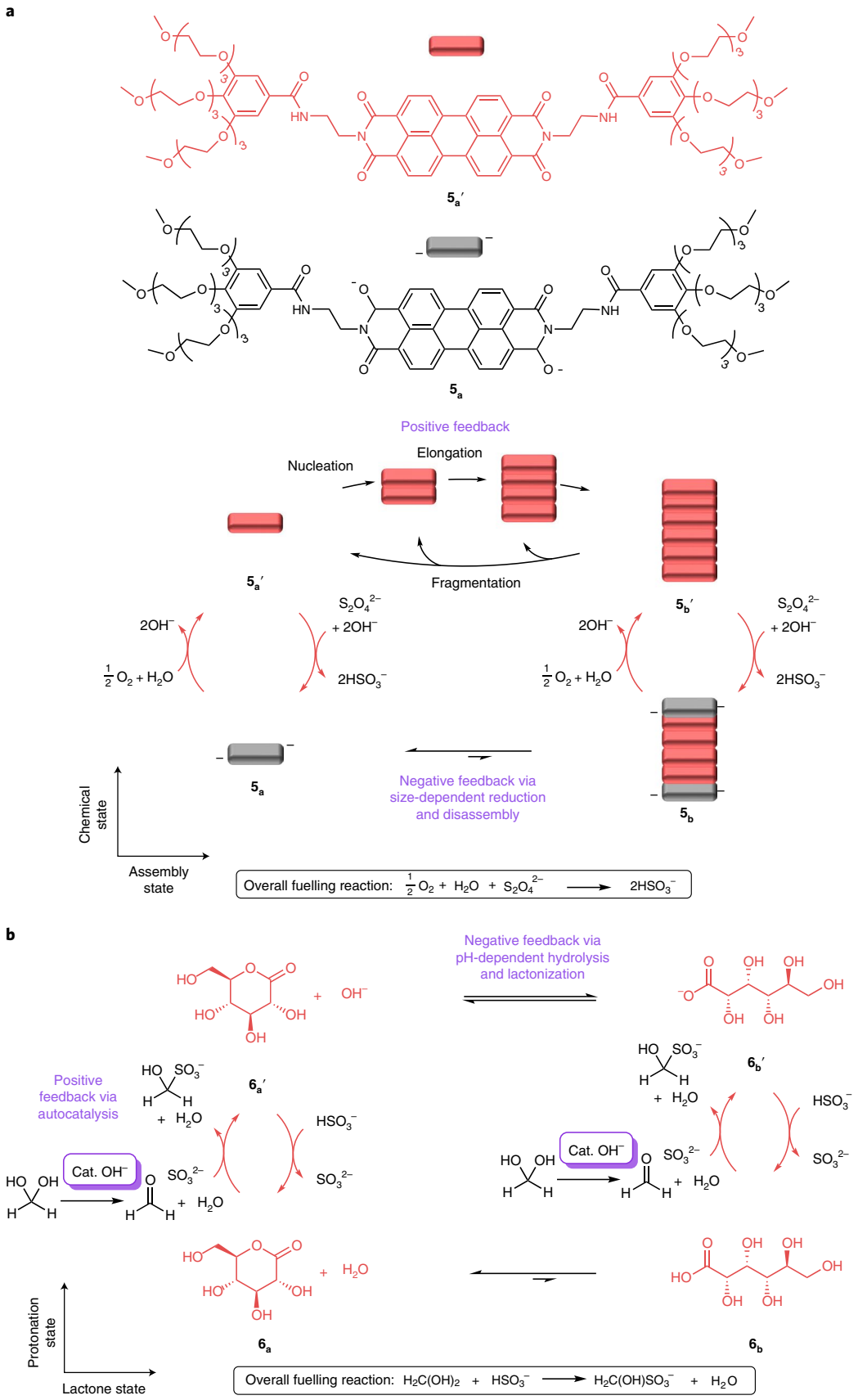
Conclusions

Chemical engines that couple chemical reaction cycles to other processes can be perceived as the archetypical module that provides the thermodynamic impetus to drive and maintain chemical systems away from equilibrium. Such modules harvest free energy from chemical fuels and power different types of non-equilibrium process. Conventional understanding of non-equilibrium systems such as those in Fig. 1b and others has been based on different specific frameworks for different systems and processes. The concept of a common chemical engine module can help identify common elements in apparently disparate mechanisms. It has already been demonstrated that some features of the chemical engine module (reaction cycles, catalysis, coupling, kinetic selection, synchronization) can be transferred from one type of system to another. For example, chemical reaction cycles originally used to drive dissipative assembly[12] have been used to power molecular ratchets and motors[42,68]. The converse may be possible, as well as the transfer of other elements, sometimes in modified form, between chemical engine systems. The different principles used to realize kinetic selection in various chemical engines (for example, energy difference of diastereomeric transition states (Fig. 2b), steric hindrance of components (Fig. 4b) and so on) may be transferable between different systems. Similarly, the mechanisms for synchronizing processes in chemical oscillators using positive and delayed negative feedback may be extendable to other non-equilibrium systems, such as synthetic molecular motors.

Perhaps the most important notion arising from the general structure for chemical engines is the recognition that such modules use fuelling reactions to power change (supramolecular assembly, co-conformational changes and so on) in a catalyst (while not necessarily requiring positive catalysis). This may stimulate the investigation of known catalytic cycles as elements of chemical engines (for molecular motors, dissipative assembly and other systems), an essentially unexplored direction for catalysis research. Most of the

Fig. 5 | Oscillating systems driven by chemical engines. a, Hermans' supramolecular oscillating system[63]. The coupling between chemical reaction cycles (vertical) and the assembly processes (horizontal) involving perylene diimide derivatives $\mathbf{5}_a$ and $\mathbf{5}_a'$ (feedback mechanisms indicated). Oxidation of $\mathbf{5}_a$ and $\mathbf{5}_b$ generates $\mathbf{5}_a'$ and $\mathbf{5}_b'$, respectively. The neutral species are reduced by $S_2O_4^{2-}$ to reform the anionic starting materials. The reduced species are mostly monomeric due to charge repulsion, while the oxidized species assemble to form a supramolecular polymer. Nucleation–elongation–fragmentation induces positive feedback, while size-dependent reduction and disassembly induce negative feedback for the assembly process. **b**, Taylor's pH oscillator[61]. Scheme showing the coupling between chemical reaction cycles that modulate the protonation state of the system (vertical) and the hydrolysis of gluconolactone (horizontal). The autocatalytic generation of hydroxide ($\mathbf{6}_a \rightarrow \mathbf{6}_a'$) generates positive feedback, increasing the pH. Gluconolactone is hydrolysed ($\mathbf{6}_a' \rightarrow \mathbf{6}_b'$) once a critical pH is reached, generating the gluconate anion and consuming hydroxide. The oscillation is probably completed by the protonation of $\mathbf{6}_b'$ to generate $\mathbf{6}_b$, and the lactonization of $\mathbf{6}_b$ to regenerate $\mathbf{6}_a$.

methods developed to date for driving systems out of equilibrium have combined two or more different reactions (for example, activation and deactivation of a monomer towards assembly). However, it is not easy to identify compatible reactions to develop a chemical reaction cycle from first principles, and the competing background decomposition of the fuel is often a problem[12]. By using



already well established catalytic reactions, problems of functional group compatibility within the reaction cycle may be largely avoided. Furthermore, traditional catalyst systems have usually been developed to bring about hard-to-perform transformations where there is effectively no background (uncatalysed) pathway to reduce the effectiveness of the fuel in powering the machine or assembly process.

Although we have focused on non-equilibrium systems driven by chemical reaction cycles, many of the same concepts also broadly apply to systems driven by diffusion[19], crystallization[69] and light[32,70]. The diverse examples discussed in this Perspective are all chemically driven non-equilibrium systems with features reminiscent of biology. The identification of the common features and components of chemical engines should facilitate communication and collaboration across fields[71–77], and help address challenges and understanding of various fundamental processes apparently intrinsic to life.

Received: 26 March 2021; Accepted: 3 August 2021;
Published online: 08 October 2021

References

1. Amos, L. A. Molecular motors: not quite like clockwork. *Cell Mol. Life Sci.* **65**, 509–515 (2008).
2. Sowa, Y. & Berry, R. M. Bacterial flagellar motor. *Q. Rev. Biophys.* **41**, 103–132 (2008).
3. Vale, R. D. & Milligan, R. A. The way things move: looking under the hood of molecular motor proteins. *Science* **288**, 88–95 (2000).
4. Schliwa, M. & Woehlke, G. Molecular motors. *Nature* **422**, 759–765 (2003).
5. Novak, B. & Tyson, J. J. Design principles of biochemical oscillators. *Nat. Rev. Mol. Cell Biol.* **9**, 981–991 (2008).
6. Gallego, M. & Virshup, D. M. Post-translational modifications regulate the ticking of the circadian clock. *Nat. Rev. Mol. Cell Biol.* **8**, 139–148 (2007).
7. Kholodenko, B. N. Cell-signalling dynamics in time and space. *Nat. Rev. Mol. Cell Biol.* **7**, 165–176 (2006).
8. Porter, S. L., Wadhams, G. H. & Armitage, J. P. Signal processing in complex chemotaxis pathways. *Nat. Rev. Microbiol.* **9**, 153–165 (2011).
9. Bissette, A. J. & Fletcher, S. P. Mechanisms of autocatalysis. *Angew. Chem. Int. Ed.* **52**, 12800–12826 (2013).
This review provides guidelines on the mechanisms of autocatalysis, particularly focusing on template mechanisms, asymmetric autocatalysis and prebiotically relevant processes.
10. Duim, H. & Otto, S. Towards open-ended evolution in self-replicating molecular systems. *Beilstein J. Org. Chem.* **13**, 1189–1203 (2017).
11. van Rossum, S. A. P., Tena-Solsona, M., van Esch, J. H., Eelkema, R. & Boekhoven, J. Dissipative out-of-equilibrium assembly of man-made supramolecular materials. *Chem. Soc. Rev.* **46**, 5519–5535 (2017).
12. Singh, N., Formon, G. J. M., De Piccoli, S. & Hermans, T. M. Devising synthetic reaction cycles for dissipative nonequilibrium self-assembly. *Adv. Mater.* **32**, 1906834 (2020).
13. Rieß, B., Grötsch, R. K. & Boekhoven, J. The design of dissipative molecular assemblies driven by chemical reaction cycles. *Chem.* **6**, 552–578 (2020).
This recent publication summarizes the design strategies for dissipative assembly, a subcategory of chemical engines, driven by chemical reaction cycles.
14. Kay, E. R., Leigh, D. A. & Zerbetto, F. Synthetic molecular motors and mechanical machines. *Angew. Chem. Int. Ed.* **46**, 72–191 (2007).
A major review that dispelled the 'motion = machine' view of artificial molecular machines, explained the difference between switches and motors, and introduced ratchet mechanisms from physics as design elements for synthetic molecular systems.
15. Erbas-Cakmak, S., Leigh, D. A., McTernan, C. T. & Nussbaumer, A. L. Artificial molecular machines. *Chem. Rev.* **115**, 10081–10206 (2015).
16. Lancia, F., Ryabchun, A. & Katsonis, N. Life-like motion driven by artificial molecular machines. *Nat. Rev. Chem.* **3**, 536–551 (2019).
17. Kathan, M. & Hecht, S. Photoswitchable molecules as key ingredients to drive systems away from the global thermodynamic minimum. *Chem. Soc. Rev.* **46**, 5536–5550 (2017).
18. van Roekel, H. W. H. et al. Programmable chemical reaction networks: emulating regulatory functions in living cells using a bottom-up approach. *Chem. Soc. Rev.* **44**, 7465–7483 (2015).
19. Epstein, I. R. & Xu, B. Reaction–diffusion processes at the nano- and microscales. *Nat. Nanotechnol.* **11**, 312–319 (2016).
20. Prigogine, I. *Introduction to Thermodynamics of Irreversible Processes* (Wiley, 1967).
21. de Groot, S. R. & Mazur, P. *Non-equilibrium Thermodynamics* (Dover, 1984).
22. Ragazzon, G. & Prins, L. J. Energy consumption in chemical fuel-driven self-assembly. *Nat. Nanotechnol.* **13**, 882–889 (2018).
This perspective provides a detailed analysis of the energy consumption in chemical fuel-driven self-assembly and defines a ratcheting constant to describe kinetic asymmetry.
23. Astumian, R. D. Kinetic asymmetry allows macromolecular catalysts to drive an information ratchet. *Nat. Commun.* **10**, 3837 (2019).
This seminal perspective elaborates the fundamental role of kinetic asymmetry in driving non-equilibrium systems through information ratchet mechanisms.
24. Das, K., Gabrielli, L. & Prins, L. J. Chemically-fueled self-assembly in biology and chemistry. *Angew. Chem. Int. Ed.* https://doi.org/10.1002/anie.202100274 (2021).
25. Grzybowski, B. A. & Huck, W. T. S. The nanotechnology of life-inspired systems. *Nat. Nanotechnol.* **11**, 584–591 (2016).
26. Lubbe, A. S., van Leeuwen, T., Wezenberg, S. J. & Feringa, B. L. Designing dynamic functional molecular systems. *Tetrahedron* **73**, 4837–4848 (2017).
27. Ashkenasy, G., Hermans, T. M., Otto, S. & Taylor, A. F. Systems chemistry. *Chem. Soc. Rev.* **46**, 2543–2554 (2017).
28. Merindol, R. & Walther, A. Materials learning from life: concepts for active, adaptive and autonomous molecular systems. *Chem. Soc. Rev.* **46**, 5588–5619 (2017).
29. Astumian, R. D. & Bier, M. Mechanochemical coupling of the motion of molecular motors to ATP hydrolysis. *Biophys. J.* **70**, 637–653 (1996).
30. Magnasco, M. O. Molecular combustion motors. *Phys. Rev. Lett.* **72**, 2656–2659 (1994).
31. Jolley, C. C., Ode, K. L. & Ueda, H. R. A design principle for a posttranslational biochemical oscillator. *Cell Rep.* **2**, 938–950 (2012).
32. van Leeuwen, T., Lubbe, A. S., Štacko, P., Wezenberg, S. J. & Feringa, B. L. Dynamic control of function by light-driven molecular motors. *Nat. Rev. Chem.* **1**, 0096 (2017).
33. Howe, E. N. W. & Gale, P. A. Fatty acid fueled transmembrane chloride transport. *J. Am. Chem. Soc.* **141**, 10654–10660 (2019).
34. Steinberg-Yfrach, G. et al. Light-driven production of ATP catalysed by F_0F_1-ATP synthase in an artificial photosynthetic membrane. *Nature* **392**, 479–482 (1998).
35. Astumian, R. D. & Derényi, I. Fluctuation driven transport and models of molecular motors and pumps. *Eur. Biophys. J.* **27**, 474–489 (1998).
36. Vögeli, B. & Erb, T. J. 'Negative' and 'positive catalysis': complementary principles that shape the catalytic landscape of enzymes. *Curr. Opin. Chem. Biol.* **47**, 94–100 (2018).
37. Amano, S., Fielden, S. D. P. & Leigh, D. A. A catalysis-driven artificial molecular pump. *Nature* **594**, 529–534 (2021).
38. Adamski, P. et al. From self-replication to replicator systems en route to de novo life. *Nat. Rev. Chem.* **4**, 386–403 (2020).
39. Alvarez-Perez, M., Goldup, S. M., Leigh, D. A. & Slawin, A. M. A chemically-driven molecular information ratchet. *J. Am. Chem. Soc.* **130**, 1836–1838 (2008).
40. Carlone, A., Goldup, S. M., Lebrasseur, N., Leigh, D. A. & Wilson, A. A three-compartment chemically-driven molecular information ratchet. *J. Am. Chem. Soc.* **134**, 8321–8323 (2012).
41. Wilson, M. R. et al. An autonomous chemically fuelled small-molecule motor. *Nature* **534**, 235–240 (2016).
The first synthetic autonomous chemically fuelled molecular motor, with directional rotation of the components driven by kinetic asymmetry through an information ratchet mechanism.
42. Borsley, S., Leigh, D. A. & Roberts, B. M. W. A doubly kinetically-gated information ratchet autonomously driven by carbodiimide hydration. *J. Am. Chem. Soc.* **143**, 4414–4420 (2021).
43. Fyfe, M. C. T. et al. Anion-assisted self-assembly. *Angew. Chem. Int. Ed.* **36**, 2068–2070 (1997).
44. Soai, K., Shibata, T., Morioka, H. & Choji, K. Asymmetric autocatalysis and amplification of enantiomeric excess of a chiral molecule. *Nature* **378**, 767–768 (1995).
45. Blackmond, D. G. Asymmetric autocatalysis and its implications for the origin of homochirality. *Proc. Natl Acad. Sci. USA* **101**, 5732–5736 (2004).
46. Blackmond, D. G. Autocatalytic models for the origin of biological homochirality. *Chem. Rev.* **120**, 4831–4847 (2020).
47. Athavale, S. V., Simon, A., Houk, K. N. & Denmark, S. E. Demystifying the asymmetry-amplifying, autocatalytic behaviour of the Soai reaction through structural, mechanistic and computational studies. *Nat. Chem.* **12**, 412–423 (2020).
48. Boekhoven, J. et al. Dissipative self-assembly of a molecular gelator by using a chemical fuel. *Angew. Chem. Int. Ed.* **49**, 4825–4828 (2010).
This seminal work established the field of dissipative assembly, using a chemical fuel to transiently assemble a monomer into a gel.

49. Biagini, C. et al. Dissipative catalysis with a molecular machine. *Angew. Chem. Int. Ed.* **58**, 9876–9880 (2019).
50. Brown, B. R. The mechanism of thermal decarboxylation. *Q. Rev. Chem. Soc.* **5**, 131–146 (1951).
51. Berrocal, J. A., Biagini, C., Mandolini, L. & Di Stefano, S. Coupling of the decarboxylation of 2-cyano-2-phenylpropanoic acid to large-amplitude motions: a convenient fuel for an acid–base-operated molecular switch. *Angew. Chem. Int. Ed.* **55**, 6997–7001 (2016).
52. Campbell, N. A. et al. *Biology: a Global Approach* (Pearson Education, 2020).
53. Chatterjee, M. N., Kay, E. R. & Leigh, D. A. Beyond switches: ratcheting a particle energetically uphill with a compartmentalized molecular machine. *J. Am. Chem. Soc.* **128**, 4058–4073 (2006).
54. von Delius, M., Geertsema, E. M., Leigh, D. A. & Tang, D.-T. D. Design, synthesis and operation of small molecules that walk along tracks. *J. Am. Chem. Soc.* **132**, 16134–16145 (2010).
55. Erbas-Cakmak, S. et al. Rotary and linear molecular motors driven by pulses of a chemical fuel. *Science* **358**, 340–343 (2017).
56. Serreli, V., Lee, C.-F., Kay, E. R. & Leigh, D. A. A molecular information ratchet. *Nature* **445**, 523–527 (2007).
The experimental realization of a non-adiabatic Maxwell 'pressure demon', the first example of a synthetic molecular information ratchet.
57. Tena-Solsona, M. et al. Non-equilibrium dissipative supramolecular materials with a tunable lifetime. *Nat. Commun.* **8**, 15895 (2017).
58. Kariyawasam, L. S. & Hartley, C. S. Dissipative assembly of aqueous carboxylic acid anhydrides fueled by carbodiimides. *J. Am. Chem. Soc.* **54139**, 11949–11955 (2017).
59. Bal, S., Das, K., Ahmed, S. & Das, D. Chemically fueled dissipative self-assembly that exploits cooperative catalysis. *Angew. Chem. Int. Ed.* **58**, 244–247 (2019).
60. Epstein, I. R. & Pojman, J. A. *An Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos* (Oxford Univ. Press, 1998).
61. Kovacs, K., McIlwaine, R. E., Scott, S. K. & Taylor, A. F. An organic-based pH oscillator. *J. Phys. Chem. A* **111**, 549–551 (2007).
62. Semenov, S. N. et al. Autocatalytic, bistable, oscillatory networks of biologically relevant organic reactions. *Nature* **537**, 656–660 (2016).
63. Leira-Iglesias, J., Tassoni, A., Adachi, T., Stich, M. & Hermans, T. M. Oscillations, travelling fronts and patterns in a supramolecular system. *Nat. Nanotechnol.* **13**, 1021–1027 (2018).
A synthetic supramolecular oscillator, incorporating feedback mechanisms into a dissipative assembly system to achieve oscillations in the assembled state of perylene diimide units.
64. Field, R. J., Körös, E. & Noyes, R. M. Oscillations in chemical systems. II. Thorough analysis of temporal oscillation in the bromate–cerium–malonic acid system. *J. Am. Chem. Soc.* **94**, 8649–8664 (1972).
65. Epstein, I. R. & Showalter, K. Nonlinear chemical dynamics: oscillations, patterns, and chaos. *J. Phys. Chem.* **100**, 13132–13147 (1996).
66. Leira-Iglesias, J., Sorrenti, A., Sato, A., Dunne, P. A. & Hermans, T. M. Supramolecular pathway selection of perylenediimides mediated by chemical fuels. *Chem. Commun.* **52**, 9009–9012 (2016).
67. Kaya, M., Tani, Y., Washio, T., Hisada, T. & Higuchi, H. Coordinated force generation of skeletal myosins in myofilaments through motor coupling. *Nat. Commun.* **8**, 16036 (2017).
68. Biagini, C. & Di Stefano, S. Abiotic chemical fuels for the operation of molecular machines. *Angew. Chem. Int. Ed.* **59**, 8344–8354 (2020).
69. Carnall, J. M. A. et al. Mechanosensitive self-replication driven by self-organization. *Science* **327**, 1502–1506 (2010).
70. Weißenfels, M., Gemen, J. & Klajn, R. Dissipative self-assembly: fueling with chemicals versus light. *Chem*. **7**, 23–37 (2021).
71. Esposito, M. Open questions on nonequilibrium thermodynamics of chemical reaction networks. *Commun. Chem.* **3**, 107 (2020).
72. Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* **75**, 126001 (2012).
73. Rao, R. & Esposito, M. Nonequilibrium thermodynamics of chemical reaction networks: wisdom from stochastic thermodynamics. *Phys. Rev. X* **6**, 041064 (2016).
74. Penocchio, E., Rao, R. & Esposito, M. Thermodynamic efficiency in dissipative chemistry. *Nat. Commun.* **10**, 3865 (2019).
75. Parrondo, J. M. R., Horowitz, J. M. & Sagawa, T. Thermodynamics of information. *Nat. Phys.* **11**, 131–139 (2015).
76. Heard, A. W. & Goldup, S. M. Simplicity in the design, operation, and applications of mechanically interlocked molecular machines. *ACS Cent. Sci.* **6**, 117–128 (2020).
77. Aprahamian, I. The future of molecular machines. *ACS Cent. Sci.* **6**, 347–358 (2020).
78. Seeman, J. I. Effect of conformational change on reactivity in organic chemistry. Evaluations, application, and extensions of Curtin–Hammett/ Winstein–Holness kinetics. *Chem. Rev.* **83**, 83–134 (1983).
79. Boekhoven, J., Hendriksen, W. E., Koper, G. J. M., Eelkema, R. & van Esch, J. H. Transient assembly of active materials fueled by a chemical reaction. *Science* **349**, 1075–1079 (2015).
80. Kariyawasam, L. S., Hossain, M. M. & Hartley, C. S. The transient covalent bond in abiotic nonequilibrium systems. *Angew. Chem. Int. Ed.* **60**, 12648–12658 (2020).
This minireview summarizes the various transient covalent bonds exploited in dissipative assembly in the context of the corresponding reaction cycles and makes the connection to molecular machines.

Acknowledgements

We thank the Engineering and Physical Sciences Research Council (EPSRC; EP/P027067/1), the European Research Council (ERC Advanced Grant 786630) and East China Normal University for funding. We thank E. Penocchio, B. M. W. Roberts and C. Tian for insightful discussions. D.A.L. is a Royal Society Research Professor.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence should be addressed to David A. Leigh.