Rotaxane-Based Molecular Machines for Organic Synthesis

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

2016

Malcolm Allan Younger Gall

School of Chemistry
Contents

Abstract.................................................................................................................................................. 7
Declaration and Copyright Statement.................................................................................................... 8
Acknowledgements ................................................................................................................................. 9
Abbreviations ........................................................................................................................................ 10

Chapter 1: Beyond the One-Pot Soup: Programming Molecular Systems for Non-Trivial Synthetic Tasks ................................................................................................................................. 15
1.1 Introduction........................................................................................................................................ 16
1.2 Selective Synthesis ............................................................................................................................ 19
1.3 Multi-Step Synthesis .......................................................................................................................... 26
1.4 Processive Catalysis ........................................................................................................................... 33
1.5 Sequence-Specific Synthesis ............................................................................................................ 37
1.6 Conclusion ......................................................................................................................................... 44
1.7 References ......................................................................................................................................... 45

Chapter 2: Sequence-Specific Beta Peptide Synthesis by an Artificial Small-Molecule Machine .............................................................................................................................................. 49
2.1 Introduction........................................................................................................................................ 50
2.2 Results and Discussion ...................................................................................................................... 52
  2.2.1 B³1-Amino Acids as Effective Barriers to Macrocycle Dethreading ............................................ 52
  2.2.2 Synthesis and Operation of a 3-Barrier Molecular Machine .................................................... 53
  2.2.3 Confirmation of Sequence Specific Peptide Synthesis ............................................................. 57
  2.2.4 Enzymatic Liberation of the B³1-Peptide Region from the Translated αβ³-Peptide .............. 59
2.3 Conclusions....................................................................................................................................... 61
2.4 Supporting Information ..................................................................................................................... 62
  2.4.1 Synthetic Procedures ................................................................................................................ 62
  2.4.2 Mass Spectral Analysis of Product of Operation .................................................................... 91
2.5 References ......................................................................................................................................... 96
Chapter 3: An Artificial Molecular Machine that Builds a Helical Oligomeric Asymmetric Catalyst

3.1 Introduction .................................................................................................................. 101
3.2 Results and Discussion .................................................................................................. 102
   3.2.1 System Design ........................................................................................................ 102
   3.2.2 Molecular Machine Synthesis ................................................................................ 103
   3.2.3 Compositional Analysis ....................................................................................... 104
   3.2.4 Operation and Analysis of Information Translation ............................................. 105
   3.2.5 Oligoleucine-Catalysed Juliá-Colonna Epoxidation ........................................... 106
3.3 Conclusion .................................................................................................................... 109
3.4 Future Work .................................................................................................................... 109
3.5 Supporting Information ................................................................................................. 111
   3.5.1 Synthetic Procedures ............................................................................................ 111
   3.5.2 Deprotection and Operation of Machine 3.1 ....................................................... 129
   3.5.3 Juliá-Colonna Epoxidation .................................................................................. 131
3.6 References ..................................................................................................................... 133

Chapter 4: A Rotaxane-Based Molecular Machine with a Trans-Acylation Catalyst for Sequence-Specific Peptide Synthesis .................................................................................. 135

4.1 An Overview of the Intended Molecular Machine Design, its Mode of Operation and Preliminary Results ......................................................................................................... 136
   4.1.1 The Intended Mode of Operation .......................................................................... 137
   4.1.2 Advantages of an Acyl-Transfer Machine over an NCL Machine ....................... 138
   4.1.3 An Antibiotic End-Goal ...................................................................................... 140
   4.1.4 Finding a Suitable Transacylation Catalyst .......................................................... 143
   4.1.5 Preliminary Model Study Results of 1,2,4-Triazoles ............................................ 148
   4.1.6 Operation of a One-Barrier Molecular Machine with an Intermolecular Nucleophile .................................................................................................................. 149
   4.1.7 Operation of a Two-Barrier Molecular Machine .................................................. 150
4.2 Results and Discussion: Redesigning and Synthesising a Second-Generation Transacylation Molecular Machine .......................................................................................... 153
   4.2.1 A Stoppered Terminal Barrier ............................................................................. 153
   4.2.2 Synthesis of a Stoppered Terminal Barrier .......................................................... 154
   4.2.3 Optimising the Synthesis of Aldehyde Macrocycle 4.59..................................... 155
4.2.4 Two-Barrier Rotaxane Assembly ................................................................. 160
4.2.5 Investigation and Development of Catalytic Unit Loading via Hydrazone
  Formation ............................................................................................................. 162

4.3 Operation of the Redesigned Second-Generation Molecular Machine .......... 168
  4.3.1 Two-Barrier Molecular Machine Deprotection ........................................ 168
  4.3.2 A Two-Barrier Machine Operation with DBU ....................................... 169
  4.3.3 Investigating Triethylamine as a Base to Activate the 1,2,4-Triazole Catalyst .....171
  4.3.4 A Two-Barrier Machine Operation with Triethylamine ............................ 172

4.4 Trouble-Shooting Unsuccessful Machine Operations ............................... 176
  4.4.1 Investigation into the Fate of the First Barrier in a Two-Barrier Machine Operation
        with Triethylamine ......................................................................................... 176
  4.4.2 Stability Studies on a 4-Nitrophenylacetic Acid Barrier Mimic ................. 177
  4.4.3 Preventing the Formation of Ketenes and Associated Degradation Products .....178
  4.4.4 Investigating Phenylacetic Acid as a First-Barrier Building Block ............. 179

4.5 Redesigning and Synthesising a Third-Generation Transacylation Molecular Machine
................................................................................................................................. 181
  4.5.1 Reloading the First Barrier ........................................................................ 181
  4.5.2 An Amide Macrocyle: Changed Connectivity between the Macrocyle and
        Catalyst ............................................................................................................. 181
  4.5.3 AMT Rotaxane Formation with an Amide Macrocyle .............................. 184
  4.5.4 Proof of Concept: Proving the Catalytic Mode of Operation .................... 186
  4.5.5 Assembling the Two-Barrier Rotaxane by Appending the Tail-Piece to the One-
        Barrier Rotaxane .............................................................................................. 189
  4.5.6 Analysis of Deprotection Conditions ......................................................... 189

4.6 Operation of the Third-Generation Molecular Machine ......................... 192
  4.6.1 Deprotection of the Third-Generation Molecular Machine .................... 192
  4.6.2 Operation of the Third-Generation Molecular Machine .......................... 194

4.7 A Basic Choice: DBU or NEt3 .................................................................... 196

4.8 Outlook .......................................................................................................... 200

4.9 Summary and Conclusions .......................................................................... 201

4.10 Experimental Section .................................................................................. 204
  4.10.1 Synthetic Procedures ................................................................................ 204
  4.10.2 Supporting Information for Section 4.3 .................................................. 240
  4.10.3 Supporting Information for Section 4.4 .................................................. 245
  4.10.4 Supporting Information for Section 4.5 .................................................. 248
4.10.5 Supporting Information for Section 4.6 ............................................................... 249
4.10.6 Supporting Information for Section 4.7 ............................................................... 257
4.11 References .............................................................................................................. 262

5. Conclusions and Outlook .......................................................................................... 265

6. General Experimental Details .................................................................................. 269
Rotaxane-Based Molecular Machines for Organic Synthesis

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy in the Faculty of Science and Engineering (School of Chemistry) in 2016.

Malcolm Allan Younger Gall

Abstract: Within living organisms in the natural world, highly complex systems have evolved over billions of years to carry out the specific synthetic functions required to support and propagate life. Nature’s use of biological machines for the synthesis of functional molecules has inspired synthetic chemists from a broad range of specialisms to design artificial molecular machines and systems capable of facilitating non-trivial synthetic tasks. A core strategy employed in attempting to emulate biological machines for synthesis has been to mimic Nature’s ability to compartmentalise discrete aspects of a synthetic process. Rotaxanes are favourable architectures around which to design molecular machines as their mechanically-interlocked nature provides the chemist with a unique means by which to achieve compartmentalisation and to control the effective molarity of non-covalently linked components. The research presented in this thesis investigates the design, synthesis and operation of novel, rotaxane-based molecular machines for the non-trivial assembly of individual amino acid building blocks into information-rich oligopeptides. The artificial devices described herein each endeavour to emulate (in a primitive manner) one of Nature’s most remarkable machines for synthesis: the ribosome. Information is programmed into these ‘synthetic ribosomes’ through their careful design and modular assembly; upon operation of the artificial molecular machine, this transcribed information is translated into a pre-defined oligopeptide product. The research presented in this thesis is laid out as follows:

Chapter 1 reviews the current state of the art in biomimetic molecular machines and systems capable of promoting non-trivial synthetic tasks;

Chapter 2 describes a molecular machine capable of non-proteinogenic oligopeptide synthesis via the sequence-specific assembly of beta-homo amino acid building blocks;

Chapter 3 presents a device which operates upon a polymer to assemble individual leucine units into a homo-oligopeptide. This product forms a secondary α-helical structure capable of asymmetric organocatalysis in the Juliá-Colonna epoxidation of chalcone derivatives;

Chapter 4 details a novel mode of amide-bond-forming catalysis for rotaxane-based molecular machines with a view to assembling an advanced peptidic precursor to Penicillin G.

Chapters 2 and 3 are presented as manuscripts which have been compiled for peer-review publication and which represent the collaborative efforts of the Author and the researchers indicated at the beginning of each chapter. The Author’s contributions are also outlined at the beginning of each chapter. These manuscripts have been modified only to ensure consistency with the other chapters contained in this thesis.
Declaration and Copyright Statement

Except where stated at the beginning of each chapter, the work referred to in the thesis has not 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 s/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, trade marks 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=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on Presentation of Theses.
It is a marathon, not a sprint, and it is far from a solo effort.

In the past 4 years, I have worked with an extraordinary group of people in an environment which encourages ambitious and collaborative research, driven by the open-minded pursuit of far-fetched, almost child-like visions. We are afforded this freedom by David Leigh: thank you, Dave. Thanks also go to The University of Manchester for hosting and funding my research.

Guillaume De Bo has been a mentor and a friend for the duration of my time in the Leigh group; I owe much of the content of this thesis to him. The De Bo group is in very safe hands: good luck!

It has been a privilege to work alongside as good a friend as Dan Tetlow. Dan is a fantastic chemist who has taught me to stay grounded and to always look at the bigger picture. For Dan and me, that bigger picture might mean cycling up hills that we’re not quite fit enough for, sharing great food, sneaking in an extra one at the pub or, in all likelihood, all of the above at once.

I am proud to say that Jon Danon is an utterly brilliant man who I absolutely adore (with atrocious alliteration). Jon has been fantastic to work alongside and was the most fun, colourful, understanding and patient flatmate that I could ever have hoped to share Tuscany with. Above all, and to borrow a phrase, Jon will always be a fierce, fierce friend. Stay daft, Brother!

A few of the postdocs that I’ve worked alongside have stood out in helping to shape my journey as a postgraduate in Manchester. I have learnt on the experience and intelligence of Matt, Tom, Alan and John time and time again in the lab and have enjoyed/endured far more hours in Manchester’s watering holes than I should have/can remember. Thank you, Gentlemen.

To all of the others who have helped me along the way – Miriam (my fellow Scot), Salma (Slama), Sonja, Jack, Jeff, Steffen (I gave you the Secret Santa book), Simone, Leoni, Adrian, Javier, our matriarchs, Sau Yin and Valerie, and my pun-dropping, chicken-loving, jam-sandwich-munching write-up buddy, Chris – thank you. Also those from further afield: Catherine, Clare, Donald and Sandy (from the Shire); Katie (and her maple syrup); and Ste, Penny and James.

Outside of the lab (there is an outside of the lab), my family have continued to support me as they have since I was knee-high to a haggis in the Ochils. Alison, Allan, Jim, Lewis, Finlay, Vicki, Aliya and Lucas: you are the most remarkably talented, ambitious and above all inspiring people that I know. I love you all and cannot thank you enough for all of your patience and support: however much of this saga you understand, I hope you’re half as proud of me as I am of you all.

I cannot thank family without thanking Joan, Brian and David (and Sue, Dave, Pat, Peter and the rest of the Eyrie too). I have been completely taken-in under the Eagles’ wing and I am so grateful to you all for the warmth and generosity you have shown me. Here’s to the next chapter!

Lastly, and most importantly, thank you to my truly wonderful girlfriend, Catherine. You have stuck by my side through every step of this adventure and the journey has been as much yours as mine. You have flown with me through the ups and you have carried me (sometimes kicking and screaming) through the darker days. You could not have been more patient, supportive, compassionate, inspiring and uplifting in the past 6 years. I love you and I cannot wait to see where our adventure will take us next...
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMT</td>
<td>Active Metal Template</td>
</tr>
<tr>
<td>APCI</td>
<td>Atmospheric Pressure Chemical Ionisation</td>
</tr>
<tr>
<td>aq.</td>
<td>Aqueous</td>
</tr>
<tr>
<td>Ar</td>
<td>Aryl</td>
</tr>
<tr>
<td>ATRP</td>
<td>Atom Transfer Radical Polymerisation</td>
</tr>
<tr>
<td>9-BBN</td>
<td>9-Borabicyclo[3.3.1]nonane</td>
</tr>
<tr>
<td>β</td>
<td>Beta-\textit{homo} (amino acid)</td>
</tr>
<tr>
<td>Boc</td>
<td>\textit{tert}-Butylcarbamate</td>
</tr>
<tr>
<td>calc.</td>
<td>Calculated</td>
</tr>
<tr>
<td>cat.</td>
<td>Catalyst/Catalytic</td>
</tr>
<tr>
<td>CD</td>
<td>Circular Dichroism or Cyclodextrin</td>
</tr>
<tr>
<td>δ</td>
<td>Chemical Shift</td>
</tr>
<tr>
<td>Conv.</td>
<td>Conversion</td>
</tr>
<tr>
<td>CPK</td>
<td>Corey-Pauling-Koltun</td>
</tr>
<tr>
<td>CRP</td>
<td>Controlled Radical Polymerisation</td>
</tr>
<tr>
<td>CuAAC</td>
<td>Copper(I) Catalyzed Azide-Alkyne Huisgen 1,3-Cycloaddtion</td>
</tr>
<tr>
<td>DCC</td>
<td>\textit{N,N}'-Dicyclohexylcarbodiimide</td>
</tr>
<tr>
<td>DIBAL</td>
<td>Diisobutylaluminium Hydride</td>
</tr>
<tr>
<td>DIPEA</td>
<td>\textit{N,N}'-Diisopropylethylamine (Hünig’s Base)</td>
</tr>
<tr>
<td>DKP</td>
<td>Diketopiperazine</td>
</tr>
<tr>
<td>DMAP</td>
<td>4-{\textit{N,N}'-Dimethylamino}pyridine</td>
</tr>
<tr>
<td>DMF</td>
<td>\textit{N,N}'-Dimethylformamide</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethylsulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>dNbpy</td>
<td>4,4’-Dinonyl-2,2’-dipyridyl</td>
</tr>
<tr>
<td>EDCI(HCl)</td>
<td>1-Ethyl-3-{3\textit{-dimethylaminopropyl}carbodiimide Hydrochloride</td>
</tr>
<tr>
<td>EDTA(4Na)</td>
<td>Tetrasodium Ethylenediaminetetraacetate</td>
</tr>
<tr>
<td>eq./equiv.</td>
<td>Equivalents</td>
</tr>
<tr>
<td>e.r.</td>
<td>Enantiomeric Ratio</td>
</tr>
<tr>
<td>ESI</td>
<td>Electrospray Ionization</td>
</tr>
<tr>
<td>Et₂O</td>
<td>Diethyl Ether</td>
</tr>
<tr>
<td>EtOAc</td>
<td>Ethyl Acetate</td>
</tr>
<tr>
<td>EtOH</td>
<td>Ethanol</td>
</tr>
<tr>
<td>Fmoc</td>
<td>Fluorenylethoxycarbonyl</td>
</tr>
<tr>
<td>HOBt(\cdot\text{H}_2\text{O})</td>
<td>Hydroxybenzotriazole Hydrate</td>
</tr>
<tr>
<td>HPLC</td>
<td>High Performance Liquid Chromatography</td>
</tr>
<tr>
<td>HR/LR</td>
<td>High Resolution/Low Resolution</td>
</tr>
<tr>
<td>IPA</td>
<td>2-Propanol</td>
</tr>
<tr>
<td>J</td>
<td>Coupling Constant</td>
</tr>
<tr>
<td>LC-MS</td>
<td>Liquid-Chromatography Mass Spectrometry</td>
</tr>
<tr>
<td>MALDI</td>
<td>Matrix Assisted LASER Desorption Ionisation</td>
</tr>
<tr>
<td>MeCN</td>
<td>Acetonitrile</td>
</tr>
<tr>
<td>MeNO₂</td>
<td>Nitromethane</td>
</tr>
</tbody>
</table>
MeOH  Methanol
Mₙ  Number Average Molecular Weight
Mₐ  Weight Average Molecular Weight
MW  Molecular Weight
Mₐ/Mₙ (PDI)  Polydispersity Index
m.p.  Melting Point
(MS-)MS  (Tandem) Mass Spectrometry
m/z  Mass:Charge Ratio for Mass Spectrometry
NCL  Native Chemical Ligation
NEt₃  Triethylamine
NMR  Nuclear Magnetic Resonance
PCR  Polymerase Chain Reaction
PDI  Polydispersity Index
PEG  Polyethylene Glycol
Pet.  Petroleum
PhMe  Toluene
Piv  Pivaloyl
Ppm  Parts per Million
Prep.  Preparative Scale
PS  Polystyrene
PyBop  (Benzotriazol-1-yl)trispyrrolidinophosphonium Hexafluorophosphate
quant.  Quantitative
(t)/(m)RNA  (Transfer)/(Messenger)Ribonucleic Acid
ran.  Random Distribution
r.t.  Room Temperature (also ‘rt’)
sat.  Saturated
SEC  Size-Exclusion Chromatography
TBAF  Tetra-n-butylammonium Fluoride
TBTA  tris[(1-Benzyl-1H-1,2,3-triazol-4-yl)methyl]amine
't'BuOH  tert-Butanol
TMS  Trimethylsilyl
rₑ  Retention Time
TSAF  tris(Dimethylamino)sulfonium Difluorotrimethylsilicate
TCEP(HCl)  tris(2-Carboxyethyl)phosphine (hydrochloride)
TFA  Trifluoroacetic acid
THF  Tetrahydrofuran
TIPS  Trisopropylsilane
TLC  Thin Layer Chromatography
TOF  Time of Flight
Trt  Trityl
VT  Variable Temperature
V-50  2,2'-Azobis(2-methylpropanimidamide) Dihydrochloride
Xxx  An Arbitrary Amino Acid

Conventional abbreviations for units, constants, physical quantities and chemical functional groups are used. Conventional 3-letter codes are used to describe amino acids.
“I can foresee also a great deal of work on the synthesis of molecules which will imitate more closely the chemical synthesizing powers of enzymes. I do not believe that you have to have a great big protein in order to get this kind of effect. I suspect a much smaller molecule with the right kind of three dimensional structure will do the same sort of thing, but these kinds of molecules are quite unknown yet in organic chemistry, and obviously we have to make them.”

– Sir Derek H. R. Barton

From an Interview with Derek Barton by Peter Farago

Chapter 1

Beyond the One-Pot Soup: Programming Molecular Systems for Non-Trivial Synthetic Tasks

Foreword

This chapter offers an overview of the multi-disciplinary strategies that synthetic chemists from diverse specialisms have developed in an attempt to mimic elegantly evolved biological systems in the modern laboratory. The complexity of biological life has long been an inspiration to synthetic chemists and as such, an innovative array of solutions has been offered to the common challenges encountered in designing and synthesising biomimetic systems. Key to the success of Nature's functional systems and biological machinery is their ability to compartmentalise individual aspects of their operation. For the synthetic chemist, this distils down to one fundamental goal: developing a handle by which to control effective molarity so as to guide the reactivity of a complex, multi-component system towards a pre-programmed outcome and away from a statistical average. The state of the art in programmable systems able to carry out non-trivial synthetic tasks is discussed herein from a supramolecular chemistry perspective.

Acknowledgements: Background research for this chapter was carried out by the author with the aid of Dr Guillaume De Bo and Dr Sonja Kuschel. Catherine J. Eagle, Dr Thomas A. Singleton and Dr Jonathan J. Danon are gratefully acknowledged for proof-reading this chapter.
Chapter 1

1.1 Introduction

Many of the fundamental aspects of our daily lives in the 21st century revolve around the operation of machines. We rely on machines that transport: the bicycle that we ride, with its well-oiled sprockets and chain, or the elevator that saves its rider’s tired legs from the stairs. We rely on machines to harness and process energy, to build and deconstruct, to gather, retain and transfer information and to carry out whatever function we are able to design a machine to fulfil. Biological life is not so dissimilar.

“Architecture is what ultimately distinguishes a living cell from a soup of the chemicals of which it is composed.” – Franklin M. Harold[1]

Nature has been evolving its own biological machines and architectures over billions of years; within living organisms, on a microscopic scale within the cell, Nature’s machines are responsible for sustaining life.[1,2] Biological machines stabilise the in vivo climate, transport components within an organism and synthesise the materials required to maintain life. They transform energy into useable forms and recycle used or unwanted species on top of a multitude of other essential biological processes, all occurring in continuous, harmonious concert.

Nature’s machines continue to evolve and adapt as they have done since they first came to be. While we can begin to appreciate the various discrete and diverse roles that different biological systems play in the propagation of life, the inner workings of Nature’s machines remain dauntingly complex. A key consideration in understanding biological machines and our own mimics thereof, is that the rules that govern life in our macroscopic world are not the same as the rules governing systems on a molecular level. Take our bicycle for example: without the macroscopic forces of gravity and friction, we would be unable to propel ourselves, no matter how powerfully we turned the pedals. Yet these macroscopic forces have little effect on life within the cell: it is a different set of laws that must instead be considered. Remote from the bounds of gravity and friction, life on the molecular level is dominated by Brownian motion and viscous forces. This impels components within the cell to come up with ingenious mechanisms by which to recruit specific substrates from the bulk, to translocate to specific cellular domains and to ultimately succeed in fulfilling their designated function. Biological machines do have clearly defined roles: each individual enzyme has been optimally shaped over the course of its evolution to carry out specific transformations on a particular class of substrate and for a distinct purpose. Each machine is optimised to work within its particular in vivo climate, either independently or in unison with other machines in a biological ‘assembly line’.
What particularly captivates the synthetic chemist about specific biological machines, and biological systems as a whole, is how they manage to function with such high fidelity, with so few mistakes, with remarkable selectivity, and at such efficient rates, in their hugely complex *in vivo* environment. Synthetic chemists are highly adept in applying the various tools in their arsenal to transform readily-available substrates into high-value products under the confines of the laboratory. We most commonly use iterative, batch-type processes involving isolating and purifying intermediate species which are then reacted further in a controlled manner to reach a desired molecular target. Nature approaches synthesis quite differently. Nature employs its biological machines, utilising their ability to carry out hugely complex series of molecular transformations in order to produce highly complex molecules for function. These systems are able to generate, stabilise and manipulate highly reactive and often incompatible intermediates, channelling their reactivity towards the formation of pre-programmed products. Key to the success in fulfilling their intended function is the ability of biological systems and machines to compartmentalise key stages of their operation such that substrates do not stray from the intended reaction pathway. The elegance of Nature’s compartmentalisation strategies fascinates synthetic chemists, and such fascination comes hand-in-hand with inspiration.

Chemists from a broad range of specialisms have approached the challenge of mimicking Nature’s use of compartmentalisation techniques to carry out non-trivial synthetic tasks. Due to their diverse backgrounds and areas of expertise, the challenge has been met from an array of different angles. Standout examples of molecular machines which can perform non-synthetic function include chemically fuelled molecular motors, transporters, muscles, walkers, logic gates and pumps. Research into supramolecular systems and molecular machines designed to carry out synthetic tasks remains in its infancy and, as a result, the ‘non-trivial’ syntheses involved lack the level complexity of those that occur seamlessly *in vivo*. Although the solutions offered hitherto are only very primitive analogues of biological processes, they represent the hallmarks of a budding field of research, important to and encapsulating all aspects of the physical sciences. The non-trivial tasks discussed herein are therefore defined as high-fidelity synthetic processes that are carried out with the aid of artificial molecular machines or functional supramolecular systems which could not occur in the absence of the systems’ compartmentalising strategy.

The examples presented in this chapter are grouped according to the function or synthetic task that each system fulfils. These sub-categories are summarised in Scheme 1 as Selective Synthesis, Multi-Step Synthesis, Processive Synthesis and Sequence-Specific Synthesis. Common to each
subset is that the reaction outcome is driven away from a statistical mixture of products and towards distinct, programmed species dictated by an “input” command.

Scheme 1: Categorisations of non-trivial synthetic tasks achieved by functional supramolecular systems.

**Selective Synthesis**: Systems which select particular species for reaction from a pool of available reactants, regulated by some form of information input.

**Multi-Step Synthesis**: Functional systems in which a specific multi-step synthesis can be carried out in one pot containing multiple reactants. The process would not be possible without a functional system to compartmentalise each step of the tandem process.

**Processive Catalysis**: Systems which can iteratively carry out a cascade process multiple times without dissociation of the catalytic device from the substrate.

**Sequence-Specific Synthesis**: Systems which can iteratively and autonomously carry out a synthetic step to assemble chemically-similar building blocks into sequence-specific oligomers according to sequence instructions embedded within a programmable template.
1.2 Selective Synthesis

Scheme 2: Selective Synthesis: Systems which select particular species for reaction from a pool of available reactants, regulated by some form of information input.

The beginning of programmable selective synthesis emerged when molecular machines were developed which could moderate the rate of a reaction in response to a stimulus. This is a strategy that Nature uses to allow complex processes to operate in parallel without non-programmed crossover. The field of switchable catalysis\[17\] strives to mimic the control over biological synthesis that is governed by feedback-loops and trigger-induced effects present within the cell.\[18\] Chemists have built activity switches into catalysts mounted upon supramolecular architectures which can, in response to an external stimulus, tune the catalytic activity. These stimuli are themselves varied with examples responding to changes induced by light, coordination events, pH, redox events, mechanical forces and alteration of reaction environment.\[17\]

In its infancy, the field of switchable catalysis was able to modulate the progress of a reaction by switching the activity of a catalyst on and off. Seminal examples from this well-reviewed field\[4,17,19–22\] include Mirkin’s allosterically moderated polymerisation catalyst,\[23\] Schmittel’s nanoswitch catalysts\[24–26\] and Leigh’s rotaxane-based switchable catalysts.\[27,28\] As systems developed, catalysts with more diverse responses to external stimuli emerged. Enatioselective switchable catalysts\[17\] were developed where application of a stimulus could alter the chiral outcome of a reaction, such as Canary’s redox-switchable catalyst for stereospecific Michael additions.\[29\] Feringa’s 4-state light- and temperature-triggered molecular motor-bound catalysts are particularly elegant examples which can, depending on which of the four possible states the system is in, carry out racemic catalysis, $R$-favoured or $S$-favoured enantioselective catalysis, or turn off catalytic activity of Michael additions\[30\] and the Henry reaction.\[31\]

In Nature, very few enzymes are able to catalyse multiple processes: those that can tend to have more than one discrete catalytic domain available.\[32,33\] Notable exceptions include moonlighting enzymes,\[34\] which have the ability to alter the characteristics of a single domain, and exceptional cases where a single enzyme domain is capable of carrying out more than one transformation.
such as the fructose-1,6-bisphosphate aldolase/phosphatase. Artificial switchable catalysts benefit from a less restrictive scope than their biological relatives and systems were soon developed that were able to catalyse more than one process via contrasting activation pathways.

State-of-the-art switchable catalysts are able to go beyond the abilities of their synthetic predecessors to carry out non-trivial synthetic tasks; they can be programmed to select specific substrates from an available pool of reagents. Diaconescu and co-workers have designed a family of redox-switchable polymerisation catalysts consisting of ferrocene-appended polydentate ligands complexed with either Zr(IV) or Ti(IV) which can selectively polymerise specific lactones from a mixture with a monomer dependence based on the ferrocene’s oxidation state. In the example shown (Scheme 3), the reduced form of the catalyst (1.1\text{red}, Switch State 1) is selective for the ring-opening polymerisation of \(\text{L-lactide} 1.2\) while the oxidised form (1.1\text{ox}, Switch State 2) is selective for \(\epsilon\text{-caprolactone} 1.3\). Both 1.1\text{red} and 1.1\text{ox} show excellent substrate selectivity: from a mixture of monomers 1.2 and 1.3, incorporation of the non-selected monomer in the resulting polymer (1.4 or 1.5) is minimal (<5%).

Scheme 3: Diaconescu’s redox-switchable polymerisation catalyst which selects either monomeric lactone substrate 1.2 or 1.3 depending on the oxidation state of the pendant ferrocene complex.

Diaconescu examined the fidelity of the redox switch by taking each monomer and the catalyst which is selective for its polymerisation and subjecting them to redox cycles. It was shown that when the catalyst was switched off (by the respective addition of either a reducing or oxidising
reagent), catalysis all but stopped. Upon re-oxidation/-reduction, catalysis resumed at a virtually identical rate. This sequence was carried out with high fidelity over three iterative cycles.

A natural progression to Diaconescu’s system was to use the redox switch in order to copolymerise monomers 1.2 and 1.3 into a region-defined block-copolymer. Where catalysts 1.1_red/ox had shown the highest activity and selectivity in the polymerisation of each individual monomer relative to its oxidation state, it failed to promote block-copolymerisation of monomers 1.2 and 1.3 in response to in situ redox switching. This shortcoming was attributed to a strong coordination between L-lactide 1.2 and the oxidised form of the catalyst 1.1_ox and was circumvented by exchanging the catalyst for another member of the reported family bearing Ti(IV) in place of Zr(IV) and a slightly modified ligand structure. Upon substituting the catalyst, a redox-moderated block-copolymerisation of L-lactide 1.2 and ε-caprolactone 1.3 was realised in an elegant process displaying the non-trivial selectivity and in situ manipulation of a switchable catalytic system in response to an external stimulus. Although the block copolymerisation shown here is rudimentary, it could lay the groundwork for more elaborate systems which can undergo repeated switching cycles to build block-copolymers with a high degree of block-control governed by a pulse-width-modulation manipulation of the redox-responsive polymerisation catalyst.

Schmittel has contributed to the field of switchable catalysis with the development of a family of catalysts based around rigid, two-state switches. Schmittel first reported a system where the addition and removal of Cu(I) ions toggled a switch to turn on and off a piperidine-catalysed Knoevenagel reaction. Expansion of this switch systems led to a Zn(II)-porphyrin catalyst capable of promoting the cis-trans photoisomerisation of a stilbene derivative when activated and a heterometallic Cu(I)/Fe(II) system which, when activated, mediates the cyclopropanation of Z-cyclooctene.

An example more significant to selective synthesis mediated by synthetic molecular machines came from the same group in 2014. Schmittel and co-workers combined two of their previous designs to produce a catalytic system capable of alternately controlling not only rate but also chemoselectivity using two different catalytic cycles (1.6-(0–2), Scheme 4). Although dual-control of two catalytic processes has been achieved by combining biological and synthetic catalysts, full control over two catalyses mediated entirely by artificial molecular machinery was unprecedented until this point.

Notably in Schmittel’s design, it is not the switch itself that carries out catalysis: rather the switch serves as a high-fidelity, programmable docking station for two discrete catalysts. Docking to the
switch renders the catalysts inactive; the catalysts are active upon their dissociation from the dock. The two labile catalysts are piperidine (1.7), which catalyses a Knoevenagel addition via iminium catalysis,[41] and secondly a Cu(I) ion complexed with 2-ferrocenyl-1,10-phenanthroline ([Cu(I)-phen], 1.8), which catalyses a Huisgen copper-catalysed azide-alkyne cycloaddition (CuAAC).[42,43]

Scheme 4: Schmittel’s toggle nanoswitch that can alternately control two different catalytic reactions: a piperidine catalysed Knoevenagel addition and a Cu(I) catalysed CuAAC reaction. In Switch State 0, both catalyses are completely inhibited. In Switch State 1, only iminium catalysis is active and in Switch State 2, only Cu(I) catalysis is active. No background reaction of inhibited catalysis is observed in any state.

The cycle begins in Switch State 0 with 1.6-(0) (Scheme 4) where both catalysts (piperidine 1.7 and Cu(I)) are docked to the nanoswitch and are inhibited. Switch State 1 (1.6-(1)) is realised on addition of phen (1.9, navy) to 1.6-(0). 1.9 displaces the pyridyl-pyrimidine ligand (red) from the Cu(I) station (orange) as it forms a new heteroleptic complex with the device-bound Cu(I) ion
Chapter 1

Consequently, the pyridyl-pyrimidine ligand switches to associate with the porphyrin-bound Zn(II) cation and, in doing so, releases piperidine 1.7 from the Zn(II) complex into the bulk, activating it for catalysis. Iminium catalysis of the Knoevenagel addition between diethyl malonate (1.10) and p-nitrobenzaldehyde (1.11) is now able to proceed while coordinative saturation renders Cu(I) catalysis inactive with no detectable background CuAAC reaction.

Switch State 2 (1.6-(2)) is achieved on addition of Cu(MeCN)_4PF_6 to 1.6-(1). The added copper strips the phen ligand (navy) from the copper nanoswitch station to form a [Cu(I)-phen]^+ complex (1.8) which catalyses the CuAAC reaction of p-nitrophenylacetylene (1.13) and benzyl azide (1.14) to form triazole 1.15. In turn, the pyridyl-pyrimidine ligand returns to the copper station to satisfy the ion’s coordination sphere by forming its original heteroleptic complex in 1.6-(2). Upon dissociation of the pyrimidine from the Zn(II)-porphyrin station, piperidine (1.7) resumes its Zn(II) association and is sequestered from the bulk, fully inhibiting its catalytic ability.

Finally, Schmittel is able to restore the system to its original Switch State 0 (1.6-(0)) upon addition of one further equivalent of phen (1.9). This addition changes nothing on the nanoswitch itself; however it coordinatively saturates the bulk-residing copper, resulting in catalytically-inactive [Cu·(1.9)_2]^+ species (1.16) which remains in solution as an inert by-product.

In this system, Schmittel demonstrates an elegant level of molecular control in order to turn on and off two alternating catalytic processes with absolute fidelity. The degree of orthogonality in both the catalyses (with their respective substrates and products) and the switching stimuli used raises the complexity of this system far beyond that which had been achieved before in non-trivial synthesis moderated by artificial molecular architectures.

Leigh followed on from Schmittel’s system with a [2]rotaxane-based chemoselective dual-switchable catalyst (1.17, Scheme 5). Based on his earlier multi-functional and chiral switchable catalysts, Leigh again employed mechanically-interlocked molecules as a means of moderating catalyst activity through compartmentalisation. Contrastingly to Schmittel’s nanoswitch system, Leigh’s molecular machine has its two catalytic units built into the rotaxane thread and therefore it is the molecular machine itself that is responsible for catalysis as well as stimuli-responsive selectivity. Additionally, Leigh uses his system to explore two very closely related organo-catalytic processes: the selective Michael addition of diethyl malonate (1.10) to two different electrophiles (1.18 and 1.19) via iminium or hydrogen-bond activation respectively.
Scheme 5: Leigh’s pH-switchable two-station rotaxane based dual catalyst. When the macrocycle (red) resides over a station, it inhibits the catalysis of that station while the exposed station is able to perform catalysis.

Leigh’s two-state system responds to changes in pH in solution. Under basic conditions, the squaramide (orange) is the favoured binding station for the macrocycle (red) and the system resides in *Switch State 1* \((1.17)\). When the macrocycle resides over the squaramide, catalysis through the H-bond activation of *trans*-β-nitrostyrene \((1.18)\) is blocked. Meanwhile, the dibenzylamine station (green) is exposed, permitting it to carry out iminium activation of crotonaldehyde \((1.19)\), thus promoting the formation of Michael addition product \((1.20)\). In *Switching State 1*, disfavoured addition product \((1.21)\) is formed in only trace amounts.

Upon acidification, the system adopts *Switch State 2* \((1.17-H^+)\). The benzylamine station becomes protonated and the resulting ammonium station (blue) is now the favoured binding site for the macrocycle. Iminium catalysis is thereby inhibited due to the protonation and macrocycle-encapsulation of the secondary amine. Concomitantly, hydrogen-bond activation of nitrostyrene \((1.18)\) by the revealed squaramide can proceed, leading to the Michael addition of diethyl malonate \((1.10)\) to form product \((1.21)\). In *Switch State 2*, only a few percent of disfavoured addition product
are formed. When the free thread of the macrocycle is employed in the reaction with 1.10, 1.18 and 1.19, products 1.20 and 1.21 form in a 1:1 ratio with significantly reduced conversion compared to the [2]rotaxane.

Using a molecular machine programmed to respond to an external stimulus, Leigh demonstrated for the first time an ability to control the reaction outcome of a single reagent by altering its reaction partner. This [2]rotaxane-based system utilises a two-state switching process compared to Schmittel’s three-state nanoswitch discussed previously. As a result, Leigh’s system is unable to turn both processes off at the same time. It does, however, have the attractive advantage of being able to switch with high fidelity between two closely related yet distinct catalytic processes using a common nucleophile. Development of Leigh’s system to contain a Switch Off state is an attractive target alongside the inclusion of his approaches for chiral catalysis,\textsuperscript{[29]} enamine\textsuperscript{[45]} and trienamine\textsuperscript{[46]} activation pathways and tandem catalytic processes.\textsuperscript{[35]}

The discussed examples from Diaconescu, Schmittel and Leigh are all linked in employing supramolecular architectures capable of existing in more than one electrochemical, geometrical or conformational environment. In each of these examples, the action of an external stimulus facilitates a system alteration resulting in a change in catalysis. This stimuli-responsive characteristic allows the systems to be programmed to carry out non-trivial synthesis by selecting specific reagents from an available pool, promoting the formation of a defined product over a statistical mixture.
1.3 Multi-Step Synthesis

Scheme 6: Multi-Step Synthesis: Systems in which a specific one-pot multi-step synthesis can be carried out that would not be possible without a functional system to compartmentalise each step of the tandem process.

In vivo synthesis often occurs with multiple biological machines working sequentially along an assembly line to carry out otherwise unachievable syntheses.[47,48] Natural multi-step synthesis progresses with a high degree of compartmentalisation which serves to stabilise intermediates that would be otherwise incompatible with their host environment.[49–51] The ability to carry out a series of discrete reaction steps in one pot has been a long-standing goal in the laboratory, especially in the arena of natural product synthesis where chemists test their mettle against Nature’s own elaborate synthetic targets.[52,53] Continued development in this area can be attributed to increased process simplicity and efficiency alongside enhanced energy- and time-economy which result from the removal of individual workup, purification and isolation steps in a synthesis.

The one-pot syntheses discussed in this section involve orthogonal multi-step (tandem) catalytic processes[54] wherein each different catalytic cycle occurs independently to the others in one pot to facilitate sequential transformations.[55] Domino (or cascade) processes,[55,56] where sequential transformations occur iteratively via the same reaction mechanism, are discussed in Section 1.4.

Often the problem faced by tandem processes is orthogonality between the required components (reagents, catalysts, intermediates, products, co- and by-products) and the reaction environment itself. The use of solid-supported reagents over the past 35 years[57] and the development of flow chemistry[58–60] have advanced our capabilities for the spatial segregation of incompatible reaction components allowing the invention of complex reactor systems. These macroscopic strategies for the physical isolation of opposing reagents[61,62] will not be further discussed herein.

Supramolecular constructs can play a critical role in the development of biomimetic multi-step syntheses due to their ability to compartmentalise key intermediates or reactive environments. A classic example of stabilisation of intermediates through compartmentalisation is Fujita’s molecular flask approach[63–65] in which reactions are carried out inside the protective cavity of molecular cages consisting of metal ion vertices linked by organic ligand edges.[66] These cages can themselves act like synthetic enzymes[67] by stabilising reactive intermediates and transition states within their cavities. This characteristic was exploited by Tiefenbacher in the synthesis of
terpenes through cage-stabilised cationic cascade cyclisations.\textsuperscript{[68]} A similar strategy has been augmented and elaborated by Nitschke who described an entirely abiological self-organising chemical assembly line making use of an \textit{in situ} self-assembling molecular cage.\textsuperscript{[69]} The design is capable of promoting the controlled tandem-synthesis of an information-rich product from three simple stock reagents: Nitschke combines furan, molecular oxygen and nitromethane through three catalytic processes to form \textbf{1.22} in 30\% overall yield (Scheme 7).

\begin{figure}[h!]
\begin{center}
\includegraphics[width=\textwidth]{scheme7.png}
\end{center}
\caption{Scheme 7: Nitschke’s self-organising chemical assembly line\textsuperscript{[69]} which transforms furan, molecular oxygen and nitromethane into information-rich product \textbf{1.22} through three catalytic processes. Cycle A: furan undergoes hetero-Diels-Alder cyclisation with singlet oxygen generated by methylene blue \textbf{1.23} to give intermediate \textbf{1.24}. Cycle B: Intermediate \textbf{1.24} enters the cavity of spontaneously assembled cage \textbf{1.25} and undergoes controlled rearrangement to fumaraldehydic acid \textbf{1.26}. Cycle C: Nitromethane undergoes an \textit{L}-proline (\textbf{1.27}) catalysed 1,4-addition into \textbf{1.26} followed by rearrangement to give product \textbf{1.22} in 30\% yield.}
\end{figure}

In the first of the three processes (\textit{Cycle A}, Scheme 7), action of visible light on photosensitiser methylene blue (\textbf{1.23}) promoted the formation of singlet oxygen from molecular (triplet) oxygen. The formed singlet oxygen underwent hetero-Diels-Alder cyclisation with furan to generate high-energy endoperoxide \textbf{1.24}.\textsuperscript{[70]} This reactive intermediate was a structurally and electronically suitable guest for concomitantly self-assembled cage \textbf{1.25}\textsuperscript{[71]} which it occupied upon formation. Upon encapsulation of endoperoxide \textbf{1.24} within cage \textbf{1.25}, the host catalysed a controlled rearrangement of its guest into fumaraldehydic acid (\textbf{1.26}) (\textit{Cycle B}). In the final catalytic cycle (\textit{Cycle C}), organocatalyst \textit{L}-proline (\textbf{1.27}) activated fumaraldehydic acid \textbf{1.26} to a directed 1,4-addition by nitromethane with chiral control which, upon rearrangement of the addition product, formed reaction product \textbf{1.22} in an overall 30\% yield (\textit{er} = 84:16) over a 48 h reaction period.

The efficiency of the system is noteworthy: the 30\% overall yield (67\% average per cycle) with high stereocontrol is a particular achievement. Catalyst loadings as low as 3.5\% and 0.5\% for photooxidiser \textbf{1.23} and cage \textbf{1.25} respectively, highlight the elegant level of control achieved. The supramolecular cage architecture was shown to be critical to the success of the process; in the
absence of any sub-component of the cage, the rearrangement of 1.24 to 1.26 was not controlled. Instead underwent non-selective, degradative rearrangement which did not lead to desired product 1.22. Nitschke’s assembly line operates under remarkably mild and environmentally benign conditions in water (buffered to pH 4) at room temperature with visible light as the only external energy source. Such mild conditions have the distinct advantage of maximising compatibility of the various components of the system, thereby focussing reactants and intermediates onto the designed reaction pathway. Indeed, key to the success of the system is the non-interfering catalytic systems whose individual components do not cross-react under the compartmentalised reaction environment.

Where Nitschke has built on the catalytic opportunities provided by encapsulation within supramolecular complexes, other branches of chemistry have explored contrasting compartmentalisation methods to permit non-trivial multi-step syntheses. Inspired by the ability of cellular components to control specific micro-climates within its membrane, chemists have turned to colloidal systems to physically segregate conflicting reactive environments.

Yang and co-workers have developed a versatile approach to reagent and reaction-environment segregation using lamination techniques of opposing Pickering emulsions (Figure 1). In their system, two separately prepared Pickering emulsions are alternately layered atop one another. The emulsions have a common continuous oil phase and contrasting dispersed aqueous phases stabilised by emulsifying silica nanospheres. The dispersed aqueous phases host conflicting chemical environments (for example acidic and basic) which remain segregated as non-interacting colloidal microdroplets. Lamination and agitation-free use help to maintain a stable colloidal environment by minimising the contact area between the opposing emulsions. When one or more reactants are added to the continuous phase, they may move freely via autodiffusion throughout the system and between both types of dispersed phase in which they may react according to the encapsulated microenvironment. The laminated system is primed for carrying out non-trivial, multi-step, sequential synthetic processes in one pot due to the compartmentalisation of otherwise incompatible reaction climates. Yang successfully carried out a range of one-pot tandem reactions with impressive overall conversions in addition to a more advanced four-step multi-cascade process.
Figure 1: Yang’s laminated Pickering emulsion with continuous oil phase and alternating layers of silica microsphere-stabilised dispersed aqueous phase hosting conflicting chemical environments (red and blue).

Scheme 8: Yang’s non-trivial one-pot tandem reaction sequences (Reactions I–IV) and an example of a four-step Multi-Cascade Process carried out in a laminated Pickering emulsion system with dispersion-segregated contrasting reaction environments.
Chapter 1

Within the aqueous acidic microdroplets loaded with HCl (red, Scheme 8), the efficient deacetalisation of a range of aromatic and aliphatic cyclic and acyclic acetals (1.28) afforded the parent aldehyde (1.29) in Reactions I–III. The aldehyde intermediates (1.29) were then able to autodiffuse within the system, and upon uptake into basic droplets (blue, Scheme 8), they could undergo further reaction. In Reaction I, aldehyde reduction to its corresponding alcohol (1.30) with NaBH₄ proceeded smoothly with aromatic and aliphatic substrates with near-quantitative overall yields. Reactions II and III explored the reaction of benzaldehyde intermediates in the ethanolamine-mediated Knoevenagel condensation with malonitrile (H₂C(CN)₂) and the 1,2-ethanediamine-catalysed Henry reaction with nitromethane respectively. The deacetalation-Knoevenagel condensation reactions (Reaction II, Scheme 8) all gave good to excellent overall yields of 1.31, as did the deacetalation-Henry processes (Reaction III, Scheme 8), albeit with little selectivity over the mono (1.32) and double addition (1.33) of the nucleophile. The one-pot Sandmeyer-type conversion of various electron rich and electron deficient anilines (1.34) into aryl iodides (1.36) was also explored. A range of anilines firstly underwent diazotisation in acidic droplets containing HCl and NaNO₂: the diazo species were subsequently trapped by NaI within the basic droplets to give the corresponding aryl iodides (1.36) in excellent yields.

More impressive still is the tolerance of the laminated Pickering emulsion systems to a four-step multi-cascade process with two starting materials (Scheme 8). Firstly, benzaldehyde dimethylacetal (1.37) underwent deacetalisation to benzaldehyde (1.38) in an HCl-containing acidic droplet and nitrobenzene (1.39) was concomitantly reduced to aniline (1.40) in a basic droplet containing NaBH₄. These two intermediates were then able to undergo a reductive amination cycle whereby they combined and condensed to their corresponding Schiff base (1.41) which was subsequently reduced within a basic NaBH₄ droplet to give secondary amine 1.42 in 66% yield over four steps in a one-pot process.

Through the innovative lamination of Pickering emulsions where the dispersed phase in each layer contains mutually destructive reaction environments, Yang and co-workers have augmented our ability to perform non-trivial multi-step syntheses. What is most impressive about this work⁷³ and its subsequent development⁷⁴ is its compatibility with a broad range of chemistries and its tolerance to variation in reactants, reagents and conditions, without alteration of the parent colloidal architecture from which it is comprised.

More recently, Weck described a contrasting use of hydrophobic and hydrophilic characteristics to separate otherwise incompatible active sites, priming them for one-pot multi-step sequential catalysis.⁷⁵ Weck combined two incongruous transition-metal catalysts in an amphiphilic cross-
linked core-shell micellar support (1.43) where each catalyst was hosted exclusively in either the hydrophilic or hydrophobic domain. A Co-porphyrin catalyst (red, Scheme 9) was mounted in the hydrophobic cross-linked core of the micelle to catalyse alkyne (1.44) hydration to methyl ketone derivatives (1.45). A chiral Rh catalyst (yellow) was appended to the hydrophilic region of the micelle for the subsequent asymmetric reduction of the methyl ketone intermediates (1.45) to chiral secondary alcohols (1.46).

Scheme 9: Weck’s tandem incompatible-transition-metal catalysis using Co and Rh catalysts grafted onto hydrophobic and hydrophilic regions of a tri-block copolymer, respectively. The copolymer is covalently cross-linked to form a micellar core-shell construct which, in aqueous media, affects the conversion of terminal alkynes to chiral secondary alcohols via methyl-ketone intermediates. [75]

Micelle 1.43 is comprised of cross-linked amphiphilic A8C-tri-block copolymer strands of poly(2-oxazoline)s. Each of the three regions of the tri-block copolymer has a different side chain to give the polymer amphiphilic character: one terminus is hydrophobic and forms the micellar core (red sphere, Scheme 9) in the aqueous reaction medium. The opposite terminus is hydrophilic (blue termini, Scheme 9) with pendant carboxylic acid side-chains through which a catalyst can be appended via amide coupling. The central block (included in the red sphere) has terminal alkene moieties on its side-chains so as to cross-link the polymer chains via thiol-ene chemistry with a tetra-thiol linker. Where two of the linker’s four thiols are responsible for cross-linking the polymer strands, the remaining two are attachment points for the Co-porphyrin catalyst (yellow, Scheme 9) using the same thiol-ene chemistry. The result is a cross-linked core-shell micelle (1.43) which spatially segregates two mutually incompatible transition-metal catalysts such that they may work cooperatively in one-pot in a non-trivial synthetic sequence.

Using this intricate construct, Weck is able to carry out the tandem sequence of alkyne-hydration of 1.44 followed by asymmetric transfer hydrogenation of the resulting methyl ketone (1.45) into the corresponding chiral secondary alcohol (1.46) for a number of electron-neutral and -rich aryl-
and alkyl-acetylenes. The achieved yields were excellent (67–95%) as was the degree of chiral control in the process (ee 94–97%). Notably, internal alkynes (1.44, R’ = Me) underwent no conversion in the process and electron-deficient aryl-alkynes (1.44, R = (4-NO₂)Ph, R’ = H) were not efficiently transformed.

The examples discussed in this section each display their own innovative approaches to a common challenge: creating a biomimetic system capable of controlling multi-step reaction syntheses requiring conflicting reaction conditions. Biological inspiration has stimulated development of our own synthetic capabilities through the invention of these new techniques and technologies. Further development in this field will undoubtedly follow with more advanced systems on the horizon, capable of hosting more and more individual and mutually opposing reaction environments. Broader substrate scopes, higher efficiencies, increased scalability and the ability to respond to external stimuli will lead to programmable biomimetic systems which can carry out diverse multi-step sequences with near-biological sophistication.
1.4 Processive Catalysis

Scheme 10: *Processive Catalysis*: Cascade systems which can iteratively carry out the same process multiple times without dissociation of the catalytic device from the substrate.

Processive catalysis describes cascade (domino) catalytic processes where the same catalytic step occurs iteratively upon a single substrate without its dissociation from the catalyst. Such processes are abundant in Nature: a prime example is the action of λ-exonuclease,[76,77] an exodeoxyribonuclease which locks onto and selectively degrades certain double-stranded DNA to form single-stranded DNA. The T4 DNA polymerase haloenzyme,[78] an enzyme responsible for constructing DNA from deoxyribonucleotides, is another example. A characteristic that is common to both of these natural examples has led to a wide range of research into artificial processive catalysts: both the λ-exonuclease and the T4 DNA polymerase haloenzyme exist as toroidal (doughnut-shaped) structures which function by threading their substrate through the toroid cavity. Chemists have been inspired to use analogous interlocked structures in the form of rotaxanes and *pseudo*-rotaxanes to achieve synthetic processive catalysis with great success.[4,19,20,79]

Some of the best-developed examples of rotaxane and *pseudo*-rotaxane architectures as platforms for processive catalysis have been reported by Nolte, Rowan and co-workers who have approached the field from both top-down and bottom-up angles.[80–86] They have developed a biohybrid catalyst by appending a Mn(III)-porphyrin oxidation catalyst to a T4 sliding clamp (gp45) protein[78] which can associate with, encircle and oxidise DNA substrates.[86] The authors have also developed a fully synthetic system for the processive alkene epoxidation of polybutadienes (Figure 2a, red) using a Mn(III)-porphyrin catalyst built into a toroidal (or macrocyclic) framework (Figure 2a, blue). The processive epoxidation of the polymer to its polybutadieneepoxide derivative occurs *via* the threading of the substrate through the cavity of the macrocyclic catalyst to form a *pseudo*-rotaxane (Figure 2a). The tert-butyl pyridine ligand serves to activate the catalyst as well as to block one of the planar Mn(III)-porphyrin faces. Due to the steric bulk of the ligand, it is unable to coordinate to the *endo*-face of the catalyst within the toroidal cavity: it therefore has to coordinate to the *exo*-face, thus blocking this face from catalysing epoxidation outside of the cavity. The nature of this ligand and alternative ligand-candidates is something that the authors later elaborated upon to increase the overall activity of the catalyst and efficiency of the process.[83,85]
Nolte and Rowan were able to mechanically interlock their *pseudo*-rotaxane architecture by capping each end of the polymer with bulky stoppers to generate [3]rotaxane 1.47 (Figure 2b). The intended epoxidation proceeded with enhanced activity relative to its *pseudo*-rotaxane counterpart as a result of its mechanically-interlocked topology. The authors cannot say conclusively whether the catalyst operates in a sequentially or randomly processive manner, meaning that the catalyst epoxidises each alkene in turn as it encounters them, as their biological inspiration does, or in a random, non-sequential order. They do however quantify that the catalyst’s dynamics upon its threaded substrate are comparable to the speeds of RNA polymerase[81] yet slower than other rotaxane-based molecular shuttles.[87] Furthermore, the activity of their catalyst is comparable to that of naturally-occurring oxidising enzymes such as Cytochrome P450.[81]

Another functional rotaxane architecture for processive catalysis is described by Takata *et al.*[88] who perform the processive hydroamination of an oligomer mediated by a Pd-carrying macrocycle through which the substrate is threaded (1.48, Scheme 11). The threaded oligomer contains four allyl-carbamate groups (red) which, upon operation, undergo a processive intramolecular hydroamination-cyclisation isomerisation to their 2-oxazolidinone derivatives (green, 1.49).
Takata uses variable-temperature (VT) $^1$H NMR techniques to highlight the efficient translocation of the macrocycle along the thread, both in rotaxanes 1.48 and 1.49, resulting in the movement of the Pd(II) catalyst between thread pyridine ligands. Furthermore, the authors detail the significance of the interlocked topology in the described process using model studies where the reaction components in their non-interlocked forms fail to yield the 2-oxazolidinone thread via hydroamination-cyclisation. Extensions to this strategy could be forecast to make use of longer oligomeric substrates and alternative chemical transformations with regard the plethora of chemistries that Pd(II) catalysis can accelerate.

Also inspired by DNA polymerases which utilise clamp-type toroidal architectures, Harada and co-workers have recently elaborated upon their previous application$^{[89]}$ of cyclodextrin (CD) in the polymerisation of lactones with a synthetic polymerase mimic.$^{[90]}$ Their recent design (1.50, Scheme 12) links an $\alpha$-CD (6-membered cyclic oligosaccharide) and a $\beta$-CD (7-membered cyclic oligosaccharide) by a short terephthalamide linker which separates the two sugar cycles by 7.0 Å. The resulting fused $\alpha$$\beta$-CD system (1.50) is able to catalyse and moderate the ring-opening polymerisation of $\delta$-valerolactone (1.51) to produce high molecular weight ($M_n = 11,000$) polyesters (1.52, Scheme 12).
The β-CD domain (green) hosts and activates 1.51 towards ring opening, thereby catalysing the polymerisation. Upon ring opening of the lactone, the α-CD domain (gold) acts as a molecular clamp (or chaperone) and is threaded by the linear growing polymer chain (1.50·1.52). This interaction stabilises the growing polymer 1.52, ensuring that subsequent monomers are only added at the carbonyl terminus held by its β-CD partner. Harada investigated multiple modifications to this design including varying the length of the linker between αβ-CD domains: too short and the catalytic β-CD could no longer recognise the cyclic substrate; too long and the pendant ring-open chain of the polymer did not thread through the clasp α-CD resulting in low conversions. Alternative CD size combinations were investigated highlighting optimal sizes of linker depending on the size of the clasp CD. In addition, control studies were performed to elucidate the operation mechanism proposed in Scheme 12. Most notably, an untethered mixture of α- and β-CD was unable to affect the efficient formation of poly-δ-valerolactone 1.52, highlighting the synergistic nature of the individual CDs in this synthetic enzyme-chaperone analogue. The result of this synergy is a synthetic tool whose efficacy far outweighs the sum of its constituent parts.

Compared to examples of selective, multi-step and sequence-specific synthesis, the small-molecule biomimetic processive catalysis systems described in this section are arguably the most similar in core design to the biological precedent that they strive to emulate. This characteristic can be largely accredited to Nature’s use of rotaxane-like structures to ensure processivity and our own, mature understanding of mechanically interlocked molecules. If methods for the formation and manipulation of complex interlocked molecules continue to advance as they have done over the past decades, the complexity of new technologies designed around these topologies will surely continue to expand to more extravagant and useful non-trivial biomimetic applications.
1.5 Sequence-Specific Synthesis

Scheme 13: Sequence-Specific Synthesis: Systems which can iteratively and autonomously carry out a synthetic step to assemble chemically similar building blocks into sequence-specific oligomers according to sequence instructions embedded within a template.

The vast majority of biological inspiration that leads to synthetic chemists designing programmable molecular systems for non-trivial synthetic tasks stems from the action of enzymes. Enzymes are mainly constructed of highly controlled arrangements of one or more proteins. The key to the sophisticated ability of enzymes to carry out particular synthetic functions lies in their higher order structure. An enzyme’s quaternary structure describes how different protein subunits bind together and, in turn, the 3-dimensional arrangement of each protein constituent is described by its tertiary structure. Where the secondary structure of each protein describes local arrangements of amino acids in α-helices or β-sheets for example, it is the core primary structure of the protein that provides the framework for the overall 3-dimensional form that the secondary, tertiary and quaternary structures produced when folded into place.

The critical primary structure of each protein is a product of the central dogma of molecular biology. Crick’s central dogma describes the flow of genetic information via the transcription of genetic code from DNA into RNA and the subsequent translation of mRNA sequences into polypeptides. The translation of mRNA sequence information into peptides of stringently ordered amino acids is the responsibility of arguably the most remarkable biological machine of all: the ribosome.

The ribosome can bind mRNA strands in a manner similar to the clamping mechanisms required for processive catalysis. The ribosome then ‘reads’ the sequence information within. According to the instructions built into each codon in mRNA, the ribosome recruits the required amino acid building blocks from the cellular medium and assembles them, one-by-one, into a built-to-order, sequence-specific polypeptide.
The fascinating ability of the ribosome has inspired the development of synthetic systems which are capable of sequence-specific synthesis. Since the seminal discoveries of Merrifield\cite{98} and the advent of solid-phase peptide synthesis half a century ago, the possibilities within sequence-controlled synthesis have undergone a monumental paradigm shift with the invention of innovative new technologies with programmable and even autonomous modes of operation.

In the realm of polymer chemistry, sequence defined polymers have wide-reaching attraction.\cite{99,100} New techniques have given access to diverse new polymer sequences. Particular examples include the use of multi-component reactions to iteratively add different monomers to a polymer such as the polymers described by Meier utilising the 4-component Ugi-reaction\cite{101} and the controlled radical polymerisations (CRP) described by Lutz to control the sequence distribution of monomers within a polymer.\cite{102} As advanced as these technologies are, they do require the iterative addition of substrates to introduce a specific sequence, while the ribosome is able to build peptides autonomously by selecting particular building-blocks from the bulk according to the template provided by mRNA.

Over the past two decades, technologies utilising DNA templates as a strategy to direct the sequence-specific assembly of building blocks have been described, furthering the application of such templates in DNA walkers.\cite{103–109} Liu,\cite{110–114} Turberfield and O’Reilly,\cite{115–117} and Seeman\cite{118,119} have described several systems which use single-strand DNA templates to program and direct the formation of prescribed bonds between substrates appended to DNA strands which are complimentary to the template. Association of the complimentary template and substrate strands serves to compartmentalise the components of the bond forming reaction through organised spatial proximity, thus directing an ordered outcome. Using this method, Seeman has carried out the programmed recruitment of gold-nanoparticle cargos upon operation of a DNA walker grafted to a DNA origami assembly line.\cite{118} The DNA walker can pick-up any or all of three cargos during operation according to a programmed set of instructions; all eight possible combinations are achieved. The collaborative efforts of Turberfield and O’Reilly have seen the development of DNA template systems for the sequence-specific assembly of building blocks through Wittig reactions of phosphonium ylides upon aldehydes to generate oligo-olefins.\cite{115,116} Liu on the other hand has explored sequence-specific peptide synthesis using DNA templates in addition to olefin assembly using Wittig chemistry.\cite{110–114} All of these examples lack autonomy: they gain control over one-pot sequential assembly through the step-wise addition of reagents, with the exception of Liu’s thermally-controlled system.\cite{112} In this example, controlled temperature increases dictate which template is actively bound to the growing oligomer chain in
line with the increasing melt-temperatures of longer DNA complexes. These systems do not therefore operate in an autonomous fashion (without external influence), which would be an attractive attribute in mimicking the sequence specific operation of the ribosome.

Autonomous multi-step organic synthesis mediated by a DNA walking molecule was first achieved by Liu.\cite{Liu2011} Liu used a multi-component self-assembly strategy to form a DNA walker system in the ready-state \( (B, \text{Scheme 14}) \) from the individual components shown in \( A, \text{Scheme 14} \).

Scheme 14: Liu’s autonomous DNA walker-mediated synthesiser for the sequence specific assembly of a tripeptide. (W) = walker; (S0) = Initiator codon; (S1–3) = substrates 1–3; (R’–R”) = transferable building blocks; (T) = template thread; (I) and (C1–3) = assembly codons; (I’) and (C1’–3’) = assembly anti-codons; (D1) and (D2) = Walker binding codons; (D1’) and (D2’) = Walker binding anti-codons; (yellow dot) = DNAzyme-cleavable linker; (black hollow dot) = Cleaved linker.\cite{Liu2011}
The individual components comprising the walker are shown in A (Scheme 14): they include a template track (T) bearing four specific codon regions (I and C1–3) responsible for associating with the complimentary anti-codon regions (I′ and C1′–3′) of individual substrates (S1–3); an initiator (S0, akin to the start-codon in mRNA); and a walker (W). The track T and substrates S1–3 are initially self-assembled, followed by introduction of the initiator and walker (S0 and W) to form the ready-to-operate device (B). Upon operation, the walker is designed to ratchet along the track-substrate complex, assembling moieties R1–R3 in sequence. The translocation of the walker is driven by the association of its anti-codon regions (D1′ and D2′) with complimentary codons on the substrates (D1 and D2). Key to the success of this system and also to its autonomous operation is an in-built mechanism by which to cleave the Rn-carrying end and D1 codon of each substrate after it has transferred its cargo to the walker. This allows the walker to associate with the following substrate unit (step (iv), Scheme 14). The mechanism, described by Joyce,[121] involves the cleaving of a linker within the substrate comprising of two RNA nucleotides (yellow dots on S1–S3, Scheme 14) by a specific ‘RNA-cleaving DNAzyme’ domain within the walker (‘loop’ region).

Through the concatenation of a DNA-templated inchworm walker and the autonomous cleaving of codon regions by a DNAzyme, Lui’s system is able to recruit and assemble building blocks without external influence in a manner reminiscent of the ribosome. Via the sequence-specific acyl transfer of N-hydroxysuccinimidyl esters onto a pendant amine of a propagating elongation site carried by the walker, a tri-amide was assembled in ca. 45% overall yield in a matter of hours. This DNA mechanical device has succeeded in achieving the first autonomous, non-trivial synthesis of a sequence-defined oligomer by a biomimetic invention.

More recently, Turberfield has reported a reconfigurable DNA-template molecular assembler for the autonomous and programmable synthesis of oligomers linked by either olefin or amide bonds.[122] Operating again through strand displacement, this technology not only utilised Wittig chemistry for the first time in an autonomous system, but also made use of the otherwise-vestigial record of used DNA instruction sequences which could be sequenced and amplified by PCR. This feature unearths an opportunity to use this technology in combinatorial synthesis because of the DNA record of the synthesised oligomer. Other advantages of this design are its modular nature and that it pulls each building block from the bulk in turn in a manner akin to the ribosome’s ordered recruitment of tRNA.

The described DNA-template devices show a remarkable aptitude for autonomous, programmable synthesis of sequence-specific oligomers which mimic ribosomal function from a
top-down perspective. They exploit the high levels of control that can be gained by the recognition of complimentary DNA sequences and use this tool of Nature’s to control the effective molarity of reagents, directing the formation of programmed substrates where, in the absence of the nano-device, a random product mixture would prevail.

The challenge of emulating the ribosome has also been tackled by Leigh\textsuperscript{[123]} from a bottom-up approach with artificial small-molecule machines for autonomous and sequence-specific peptide synthesis. Leigh again employed the previously-encountered rotaxane strategy to gain control over effective molarity in his system. His design (1.53\textsuperscript{a}, Scheme 15) consists of a [2]rotaxane where the macrocycle is compartmentalised between a bulky stopper and the first of three successive barriers bearing labile amino acid building blocks attached to the track through phenolic esters. Upon the macrocycle is a GlyGlyCys tripeptide which has two roles: firstly, the free amine of the N-terminal glycine is the elongation site upon which the machine will assemble the peptide sequence; secondly, the primary thiol on the cysteine residue is the oligo-peptide assembly catalyst which operates using native chemical ligation (NCL).\textsuperscript{[124,125]}
Upon operation in the presence of base, the thiol catalyst on the macrocycle-bound cysteine residue (1.53b) cleaves the first amino acid building block (phenylalanine, green, 1.53b') from the track to form thioester 1.53c in an [O-S]-acyl shift (Scheme 15). 1.53c subsequently undergoes an [S-N] transacylation step where phenylalanine is passed onto the pendant primary amine of glycine via a [1,11]-acyl shift to form 1.53d, concomitantly regenerating the cysteine thiolate.

Upon removing the amino acid from the track in front of it, the macrocycle is able to slip over the unloaded tyrosine residue on the track and thus advance to the second barrier where leucine blocks its path (III, Scheme 16). The native chemical ligation sequence then removes leucine from its tyrosine foothold upon the track and transfers it as a thioester onto the end of the propagation chain by attaching it to phenylalanine through a slightly larger [1,14] transition state to form IV (Scheme 16). A third iteration of this NCL process transfers the third barrier, alanine, to the growing peptide chain through a [1,17] transition state to form 1.54 (Scheme 16), which can then dethread from the operated track 1.55.
Leigh has since elaborated upon this system where a larger machine is operated with four substrate-bearing barriers to assemble a tetra-peptide in a sequence-controlled manner in 53% yield. One could imagine that following on from this system, new chemistries could be utilised such as the carbon-carbon bond formation achieved by Wittig DNA synthesisers (vide supra). The technology could be applied to larger systems capable of assembling longer peptides, although this may be restricted by the limitations of NCL with regards the increasing ligation transition state explored in each successive NCL cycle.

The described examples of molecular systems designed to carry out sequence-specific synthesis remain primitive analogues of the highly evolved ribosome that they are inspired by. They do, however, mark a significant advance in the capabilities of man-made molecular machines, be they large DNA-based devices with pseudo-biological components or small artificial molecular machines born out of the development of topologically complex supramolecular architectures. To comprehend the potential value of systems which can build oligomers and polymers constituting precisely assembled building blocks, one need not look further than the central dogma of molecular biology and its reliance on sequence-defined macromolecules. Accordingly, further developments in our existing technologies will surely be sought to improve upon versatility, efficacy and reusability which will in turn open the door to real-world applications.
1.6 Conclusion

The fascination held by chemists for the intricate mechanisms that Nature has evolved to support the propagation of life has been a long-standing source of inspiration. Synthetic chemists from diverse specialisations have risen to the challenge of emulating examples of biological ingenuity in synthetic systems with the common goal of achieving non-trivial synthesis. Nature has the ability to channel the fate of a pool of reactants towards the formation of distinct, pre-programmed products. The ability of biological machines, engineered over billions of years of evolution, to compartmentalise various stages of biosynthesis is key to their success. The theme of compartmentalisation is one in particular that synthetic chemists have grasped in developing their own biomimetic technologies.

Examples of machines that can respond to external stimuli in order to change the course of a reaction have been developed. This programmable, selective synthesis allows us to choose which components from an available pool will react with one another and via which mechanism. Inventions facilitating the tandem reaction of substrates through multiple synthetic steps in one pot have been described where the individual transformations, required components or intermediates would be mutually incompatible without compartmentalisation technologies. Examples of molecular machines have been reported which can carry out iterative cascade reactions on a substrate where processivity is maintained using mechanically interlocked architectures as analogues of Nature’s substrate-clamping mechanisms. Finally, systems capable of the autonomous, sequence-specific synthesis of oligomers have been developed which operate in a manner comparable to the action of the ribosome.

This review has introduced the state of the art in a new generation of sophisticated chemical technologies which carry the potential to fundamentally change the way that we approach molecular synthesis. The discussed inventions are innovative and programmable devices and systems, capable of facilitating non-trivial synthetic tasks, but they represent only the very beginning of a new age of synthesis. Further research and investment are critical to the continued success of this young field with foreseeable rewards lying beyond the current proof-of-concept phase and into functional value-adding real-world applications.
1.7 References

[37] X. Wang, A. Thevenon, J. L. Brosmer, I. Yu, S. I. Khan, P. Mehrkhodavandi, P. L. Diaconescu,
Chapter 1


Chapter 1

1601–1605.

Chapter 2

Sequence-Specific Beta Peptide Synthesis by an Artificial Small-Molecule Machine

Abstract

The synthesis and operation of a three barrier, rotaxane-based, artificial molecular machine capable of sequence specific beta-homo ($\beta^3$) peptide synthesis is reported. The machine utilizes non-proteinogenic $\beta^3$-amino acids, a class of amino acids not generally accepted by the natural ribosome. Successful operation of the machine via native chemical ligation (NCL) demonstrates that the challenging 15- and 19-membered ligation transition states, which have been previously unexplored in machina, are suitable for information translation via machine operation. Furthermore, by employing peptidases, a class of naturally occurring biological machines, we demonstrate that artificial and natural machines can work in concert to generate products currently inaccessible by either strategy alone.

Acknowledgements: The research discussed in the following chapter represents a collaborative effort between the author and the following researchers: Dr Guillaume De Bo, Dr Matthew O. Kitching, Dr Sonja Kuschel, Dr Daniel J. Tetlow and Dr John W. Ward under the supervision of Prof. David A. Leigh. This work is presented as a manuscript prepared for peer-review publication by the author and the aforementioned researchers. The author performed the synthesis, analysis and characterisation of all compounds herein at least once with the exception of those indicated at the beginning of Section 2.4.1: Synthetic Procedures.

Dr Alan T. L. Lee and Dr Matthew O. Kitching are gratefully acknowledged for proof-reading this chapter.
2.1 Introduction

Nature utilizes highly evolved molecular machines to achieve an array of complex transformations with remarkable selectivity.\(^1\) For example, the ribosome is capable of translating genetic information held within a polymeric mRNA strand by facilitating the sequence specific recruitment and assembly of amino acid building blocks from the cellular medium.\(^2\)-\(^5\) Key to this ability is the ribosome’s highly evolved 3-dimensional structure\(^6\)-\(^9\) and its capacity to selectively engage tRNA anticodons under the direction of mRNA. However, seemingly minor modifications to the ribosome’s tRNA substrates may cause them to be rejected by the complex. For instance, the ribosome is incapable of iteratively translating non-natural D-amino acids\(^10\)-\(^13\) or homologated \(\beta^3\)-amino acids,\(^14\) even when they are appropriately functionalized with natural tRNA codons. Although bioengineering can go some way to addressing these limitations,\(^15\)-\(^20\) minor changes to substrate or machine can lead to complete loss of function in an unpredictable manner.\(^21\)

An artificial rotaxane-based molecular machine which mimics the sequence-specific peptide synthesis of the ribosome by sequentially adding successive \(\alpha\)-amino acid building blocks to a growing peptide chain was recently reported by our group.\(^22\),\(^23\) This design takes advantage of the well understood native chemical ligation (NCL)\(^24\),\(^25\) process to translate sequence information from the molecular machine into the peptide undergoing synthesis. We reasoned that with minimal change to the architecture of the molecular machine, extension of this methodology to promote the synthesis of \(\beta^3\)-peptides would be possible. The translation of \(\beta^3\)-amino acids by our ‘synthetic ribosome’ would highlight its versatility, but more critically its tolerance to non-proteinogenic amino acid substrates. The \textit{in machina} synthesis of non-proteinogenic \(\beta^3\)-peptides has emerged as an attractive target as such substrates have become increasingly prevalent in pharmaceutical and peptidomimetic fields\(^26\)-\(^28\) due to the enhanced biostability over their bioactive \(\alpha\)-peptidic counterparts.

In order to achieve our goal, two key challenges had to be addressed. Firstly, in a rotaxane-based machine, can \(\beta^3\)-amino acids serve as competent barriers to macrocycle motion, thereby ensuring rigorous compartmentalization of the macrocycle within the machine? Secondly, due to the homologated transition state sizes required, can NCL still allow the \(\beta^3\)-amino acid sequences to be efficiently translated in a purely \textit{intra-machina} and sequence-controlled fashion? Additionally, we were eager to develop a method whereby the translated product of machine operation could be selectively removed from the NCL cysteine catalyst to generate a product free from its translational machinery. Here we report the modular synthesis and sequence-specific operation
of a rotaxane-based molecular machine (2.1) bearing three \( \beta^3 \)-amino acid building blocks. Upon successful operation, the device assembles a mixed \( \alpha\beta^3 \)-hexapeptide of defined sequence. This peptide sequence is shown to be selectively digested by an endopeptidase at the C-terminal \( \alpha \)-amino acid region to liberate the translated \( \beta^3 \)-peptide fragment from the NCL catalyst.
2.2 Results and Discussion

2.2.1 β³-Amino Acids as Effective Barriers to Macrocycle Dethreading

In order to explore our objectives, we first investigated the integrity of β³-amino acids as barriers to macrocyclic shuttling along the thread of a rotaxane. We have shown previously that α-phenylalanine serves as an effective rotaxane-stabilizing barrier,[23] however by homologating the barrier to β³-phenylalanine, it becomes more flexible. This increased conformational freedom amplifies the risk of the barrier folding into a conformation where it no longer serves as an effective blocking group to macrocyclic shuttling. Copper-catalyzed active metal template (AMT)[29,30] construction of [2]rotaxane 2.2 bearing a β³-Phe amino acid barrier was thereby pursued using minor modifications of our previously reported Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC) protocol (Scheme 1).[23] Pretreatment of macrocycle 2.3 (see Section 4.10.1.2 for synthetic details), with CuPF₆·4MeCN, followed by addition of bulky azide stopper 2.4 and β³-amino acid-bearing barrier 2.5 afforded rotaxane 2.2 in 40% yield along with free thread 2.6 (Scheme 1).

Scheme 1: AMT synthesis of a one-barrier rotaxane 2.2 using an N-Boc-β-homo-phenylalanine functionalized barrier 2.5, azide stopper 2.4 and macrocycle 2.3. Reagents and conditions: (i) macrocycle 2.3 (1.5 equiv.), CuPF₆·4MeCN (0.5 equiv.), barrier 2.5 (1 equiv.) and stopper 2.4 (6 equiv.) CH₂Cl₂·BuOH (6:1), r.t., 16 h, 40% of 2.2, 7% of 2.6.
Figure 1: $^1$H NMR stackplot of (i) macrocycle 2.3; (ii) rotaxane 2.2; and (iii) free thread 2.6 showing proton shifts indicative of rotaxane formation (600 MHz, $d_6$-acetone, 298 K). The assignments correspond to the lettering shown in Scheme 1.

Confirmation of mechanical bond formation was established through a combination of mass-spectrometric and $^1$H NMR spectroscopic analysis (Figure 1). In comparison to the non-interlocked components 2.3 and 2.6, rotaxane 2.2 demonstrates significant up-field shifts of both the triazole proton ($H_e$) and the protons on the n-propyl tether (e.g. $H_c$ and $H_d$) which is consistent with shielding by the phenolic group of the macrocycle. Similar shielding effects are apparent for several macrocycle proton environments (e.g. $H_D$, $H_E$ and $H_F$) in addition to diastereotopic splitting of $H_E$ and $H_D$ caused by the non-$C_2$-symmetric thread occupying the macrocycle cavity. Furthermore, remote protons ($H_h$, $H_i$ and $H_j$) demonstrate negligible shifts indicating that macrocycle resides over a region of the thread distal to these environments. $^1$H NMR spectroscopic evidence strongly suggests that $\beta^3$-phenylalanine indeed serves as an effective barrier to uphold the rotaxane’s mechanical bond. Analysis by high resolution mass spectrometry was consistent with the formation of [2]rotaxane 2.2.

2.2.2 Synthesis and Operation of a 3-Barrier Molecular Machine

Confident of the stability of one barrier rotaxane 2.2, our attention turned towards the synthesis and operation of a three-barrier machine (2.1, Scheme 2). Combination of rotaxane 2.2 and 2-
barrier extension piece 2.7 bearing β<sup>3</sup>-Leu- and β<sup>3</sup>-Ala-loaded barriers was achieved by employing our previously reported CuAAC “rotaxane-capping” strategy<sup>[23]</sup> with excellent efficiency, leading to isolation of 3-barrier rotaxane 2.8 in 97% yield. The final synthetic step involved hydrazone formation via the direct condensation of BocGlyGlyCys(Trt)NHNH<sub>2</sub> α-peptide catalytic unit 2.9 with aldehyde-functionalised rotaxane 2.8 in the presence of catalytic aniline.<sup>[31]</sup> This final step completed the highly modular assembly of 2.1, a rotaxane-based molecular machine bearing three reactive barriers, each offering a different non-proteinogenic amino acid building block for the sequence specific assembly of a β<sup>3</sup>-peptide.

Scheme 2: Synthesis of the 3-barrier machine 2.1 bearing non-proteinogenic amino acids. Reagents and conditions: i) CuPF<sub>4</sub>·4MeCN (1.5 equiv.), Tentagel™ TBTA resin, CH<sub>2</sub>Cl<sub>2</sub>:tBuOH (15:2), r.t., 48 h, 97%; ii) BocGlyGlyCys(Trt)NHNH<sub>2</sub> 2.9, aniline, CH<sub>2</sub>Cl<sub>2</sub>, r.t. 48 h, 72%.

Following a global deprotection of 2.1 with trifluoroacetic acid and triisopropylsilane (TIPS), amino acid Boc groups were cleaved in addition to removal and reduction of the cysteine catalyst’s trityl protection. The freshly deprotected molecular machine (2.10) was then operated for 7 days
allowing the iterative transfer of barrier-loaded \(\beta^3\)-amino acids on to the \(N\)-terminus of the catalytic unit \textit{via} NCL as outlined in Scheme 3. Under the basic operation conditions, the freshly revealed nucleophilic thiolate (2.10a) present on the macrocycle-bound cysteine residue is primed to attack the phenolic ester (2.10a') which links the first \(\beta^3\)-amino acid building block to the thread. The first building block is translocated from the thread to the catalytic unit by an \([O-S]\) trans-acyl shift, generating thioester 2.10b. Reaction of the cysteine thiolate at out-of-sequence phenolic esters is prevented in two ways. First, the ability of \(\beta^3\)-amino acids to act as effective barriers to macrocycle shuttling, twinned with the modular AMT rotaxane assembly and subsequent “rotaxane-capping”, serves to completely compartmentalize the catalytic unit-bearing macrocycle between the stopper and first barrier. Second, the inbuilt rigidity of the thread restricts its conformational freedom, reducing the possibility of the catalytic arm reaching beyond its closest barrier to more distal reactive sites on the thread. In addition, operation at high dilution \([0.6 \text{ mM}]\) ensures only \textit{intra-machina} transacylation can prevail.
Scheme 3: Deprotection and operation of molecular machine 2.1. Reagents and conditions: i) TFA, TIPS, CH$_2$Cl$_2$, r.t., 2 h; ii) Et$_3$N (44 equiv.), PPh$_3$ (3.5 equiv.), DMF, 60 °C, 7 d, 29% over 2 steps.

Once generated, thioester 2.10b is poised to undergo intra-molecular [S-N] acyl shift via an 11-membered transition state, transferring the β$^3$-amino acid to the N-terminus of the macrocycle bound GlyGlyCys residue to form 2.10c. A new peptide bond is formed whilst the nucleophilic thiolate of the catalytic unit is concomitantly regenerated. With the first barrier removed, the
macrocycle is free to explore the length of the thread between the stopper and the second barrier. On encountering the second barrier, cysteine can again translocate the β³-amino acid from the thread onto the propagating end of the growing peptide via NCL with a 15-membered [S-N] transacylation transition state. The catalyst may thereafter reach the third sequential barrier for a final NCL iteration through a 19-membered [S-N] transacylation transition state. With the final barrier removed, the macrocycle bearing the fused αβ³-peptide is free to dethread from the operated track and is released into the bulk medium.

After the 7 day operation period, mass spectral analysis revealed complete consumption of deprotected molecular machine 2.10 and the formation of 2.11a. Following solvent removal and purification of the crude operation mixture by preparative TLC, 2.11a had converted to S,N-acetal 2.11b; operation product 2.11b was isolated in 29% yield and displayed a mass spectrometry (MS) profile comparable with an authentic standard. In solution and especially under the mildly acidic conditions of both purification and MS analysis, 2.11a and 2.11b appear to be in exchange, presumably under dynamic equilibrium, where formation of the [S,N] acetyl species 2.11b is favoured (see Section 2.4.2 for further details).

Importantly, due to the additional methylene units present in the homologated β³-amino acid building blocks, the size of the [S-N] acyl transfer transition states vary in a critical way from our previously reported machines.22,23 Extensive investigation by Katritzky32–41 into the relative rates of intra-molecular NCL versus competing ligation pathways (including inter-molecular ligations) has revealed that NCLs by the 15- and 19-membered transition states explored herein are significantly harder to achieve than NCLs through the 14- and 17-membered transition states exploited in our previous work.22,23 We attribute the increase in required operation time and marginally diminished yield of product compared to our previously reported devices22,23 to this more demanding reaction pathway. Pursuing these more challenging transition states increases the risk of undesired inter-machina ligations which would erode sequence specificity. As such, the sequence-specific operation of a machine utilizing these demanding transition states would serve to demonstrate the robustness of our design.

### 2.2.3 Confirmation of Sequence Specific Peptide Synthesis

The sequence specificity of the operation of molecular machine 2.1 was confirmed by tandem ESI mass spectrometry (MS-MS) of the isolated product 2.11b (Figure 2). The most reliable and intense MS-MS fragmentation profile was achieved by examining the doubly charged \([2.11b+2H]^2+\) mass-ion (m/z 898).
Figure 2: Tandem mass-spectrometric analysis of doubly charged \([\text{2.11b}+2\text{H}]^{2+}\) mass ions corresponding to \([S,N]-\text{cyclic acetal derivatized macrocycles bearing the translated Piv-β^3\text{-Ala-β^3\text{-Leu-β^3\text{-Phe-αGly-αGly-αCys peptide.}}}]\)

(i) Fragmentation modes of α- and β^3-peptides; (ii) Independently synthesised standard of the operation product following its exposure to operation conditions and subsequent purification; (iii) Purified operation product. Orange peaks indicated amide bond \(\text{C-N} \) fragmentation; green peaks indicate β^3-peptide retro-Michael fragmentation; red peaks indicate β^3-peptide retro-Mannich fragmentation; and blue peaks indicate addition or loss of water.

As anticipated, fragmentation showed iterative cleavage at the amide bonds of the assembled αβ^3-peptide backbone down to the parent macrocycle, consistent with the expected sequence transcribed into the molecular machine by our modular assembly strategy. This result was corroborated by comparison with authentic material independently prepared by conventional peptide synthesis. Critically, no fragments to indicate formation of a peptide of incorrect sequence were observed. In addition, Seebach has shown that oligomeric peptides containing β^3-amino acids undergo fragmentation under several further pathways when compared to their α-
amino acid congeners as a result of their homologation. Analysis of the MS-MS spectrum showed peaks consistent with hydration and dehydration (Figure 2, shown in blue), retro-Michael (Figure 2, shown in green, Y*) and retro-Mannich (Figure 2, shown in red, Y⁰) fragmentation pathways, providing further confirmation of the successful translation of the sequence (for more details, see Section 2.4.2).

2.2.4 Enzymatic Liberation of the β³-Peptide Region from the Translated αβ³-Peptide Product

After the successful operation of molecular machine 2.1 to form 2.11a and isolated product 2.11b, we sought to develop a method by which the translated β³-sequence could be liberated from its parental machine architecture. We proposed that the proteolytic action of an endotopic protease enzyme may result in the selective catabolism of the α-region of the operation product, thereby liberating the translated β³-sequence from both the cysteine catalyst and the attached macrocycle. The basis for this anticipated selectivity revolves around the remarkably high stability of β-peptides to both in vitro and in vivo biodegradation, primarily investigated by Seebach.

We initially investigated the enzymatic digestion of model αβ³-peptide 2.14 which serves as a mimic for product of operation 2.11a, following thiol acylation to ameliorate handling. The enzyme investigated was subtilisin Carlsberg, a broad scope serine endopeptidase isolated from *bacillus licheniformis* which shows high stability and activity at elevated temperatures and optimal activity at pH 7–8. As such, mixed αβ³-peptide 2.14 was treated with subtilisin Carlsberg in phosphate buffer (pH 8) at 50 °C for 5 days after which analysis by electrospray ionisation (ESI) MS was carried out (Scheme 4). The result indicated not only the somewhat predictable hydrolysis of both the thio- and oxy-esters to give 2.15, but also the efficient hydrolysis of the [αGly-αCys] amide bond to furnish Piv-β³Ala-β³Leu-β³Phe-αGly-αGly-OH (2.16) as the major product visible by ESI(-)-MS. In addition, a small amount of tetrapeptide Piv-β³Ala-β³Leu-β³Phe-αGly-OH (2.17) resulting from cleavage of the [αGly-αGly] amide bond was observed. Following longer incubation times and refreshment of the enzyme, no significant further digestion was observed.
Scheme 4: Crude LR(ESI)-(--)MS analysis of the selective proteolysis of the α-peptide region of model mixed αβ3-peptide 2.14 by subtilisin Carlsberg resulting in the liberation of the β3-peptide region from cysteine, the NCL catalyst used to assemble the β3-peptide during machine operation. No breakdown of the β-peptide region was observed. Reagents and conditions: i) Mixed αβ3-peptide 2.14 (2 mg/mL), subtilisin Carlsberg (1 mg/mL), phosphate buffer (pH 8), 50 °C, 5 days.

The inability of subtilisin Carlsberg to digest the bridging αβ3-peptide bond (α-Gly-β3-Ala) is in line with previous observations of the actions of peptidases (except pronase) on such bonds.\[28,51\] This is postulated to be due to the protecting nature of the β-peptide fragment on the αβ3-peptide junction, presumably brought about by discordance between the β-component and the active site of the enzyme. This theory also explains the lower propensity of the enzyme to cleave the [αGly-αGly] amide adjacent to the αβ3-junction. It must be appreciated however, that the vast majority of prior research into the proteolytic cleavage of amide bonds in mixed αβ-peptides has been carried out on systems with α-peptides at the N-terminus of the β-peptide fragment,\[51\] not at the C-terminus as exhibited here.

Treatment of model αβ3-peptide 2.14 with subtilisin Carlsberg has hereby shown that enzymatic cleavage of the molecular machine-translated β3-peptide sequence from the device’s catalytic cysteine residue and attached macrocycle is a viable target. The proteolysis of 2.11a and 2.12 are currently being explored in our laboratory. This highly selective cleavage of the cysteine NCL catalyst from the translated operation product sequence is highly reminiscent of the excision of the N-terminal methionine residue from ribosomal proteins (installed by the start-codon) by methionine aminopeptidases in eukaryotic post translational modification.\[53,54\]
2.3 Conclusions

In conclusion we report the synthesis and operation of an artificial molecular machine capable of the autonomous synthesis of non-natural, fused αβ^3^-peptides by translating β^3^-amino acid building blocks. We demonstrate the adaptability of our designed molecular machine architecture which is tolerant to variation in operational demands and is compatible with non-proteinogenic substrates. Operating through 11- and particularly challenging 15- and 19-membered NCL transition states, sequence information transcribed into the molecular machine is translated into a product of operation in a sequence-controlled manner without detectable error. Furthermore, by linking our synthetic molecular machine with a naturally occurring enzyme, a natural and an artificial molecular machine can work in concert to produce αβ^3^-peptide products currently inaccessible by either strategy alone. We hope that future work will further demonstrate the potential benefits of natural and artificial molecular machines working in synergy.
2.4 Supporting Information

2.4.1 Synthetic Procedures

Barrier precursor 2.18\textsuperscript{[22]} was prepared according to literature procedures by the author. Azide stopper 2.4\textsuperscript{[55]} was prepared by Dr Matthew O. Kitching (MOK) according to a literature procedure. Aldehyde macrocycle 2.3 was prepared by the author according to the procedure described in Chapter 4 of this thesis. Machine operations were carried out by Dr Guillaume De Bo (GDB), MOK and the author. Enzyme studies were carried out and analyzed by Dr Daniel J. Tetlow (DJT), MOK and the author. All described compounds have been synthesized, analyzed and characterized by the author with the following exceptions: compounds 2.23, 2.25, 2.26 were synthesized by MOK and DJT and were purified, analyzed and characterized by the author; 2.24 was synthesized by MOK and characterized by the author; compounds 2.14, 2.28, 2.29 and 2.30 were synthesized by DJT and characterized by the author and DJT.

2.4.1.1 General Procedures

General Procedure 1: Barrier Loading.
To a solution of barrier (1.0 eq.) in CHCl\textsubscript{3}:THF (1:3, 0.05 M) was added Boc-\(\beta\)-homo-amino acid (1.2 eq.), DMAP (0.1 eq.) and EDCI-HCl (1.1 eq.). The reaction mixture was stirred for 16 h at room temperature. The solvent was removed under reduced pressure. Purification by flash column chromatography (SiO\textsubscript{2}, EtOAc/CH\textsubscript{2}Cl\textsubscript{2} or MeOH/CH\textsubscript{2}Cl\textsubscript{2}) yielded the product.

To a solution of Boc protected peptide (1.0 eq.) and triphenylmethanol (0.5 eq.) in CH\textsubscript{2}Cl\textsubscript{2} (0.05 M) was added TFA (20% by volume). The reaction mixture was stirred at room temperature for 2 hours. The reaction mixture was diluted by 50% with toluene then the solvent was removed under reduced pressure. The crude residue was azeotroped twice with toluene to remove excess TFA. Trituration with Et\textsubscript{2}O:n-hexane (1:1, to remove triphenylmethanol) left the TFA salt of the product which was used without further purification.

General Procedure 3: Amide Coupling of TFA Salts.
To a solution of the TFA salt (1.0 eq.) in CH\textsubscript{2}Cl\textsubscript{2} (0.05 M) was added NE\textsubscript{t}\textsubscript{3} (1.2 eq.) followed by EDCI-HCl (1.2 eq.), HOBt-H\textsubscript{2}O (1.2 eq.) and Boc or Piv protected amino acid (1.2 eq.). The reaction mixture was stirred at room temperature for 16 h. The mixture was diluted with CHCl\textsubscript{3} and washed sequentially with 2 M NaHCO\textsubscript{3}, 1 M HCl and brine. The organic phase was dried (MgSO\textsubscript{4})
and the solvent was removed under reduced pressure. Purification by flash column chromatography (SiO$_2$, MeOH/CH$_2$Cl$_2$) yielded the product.

### 2.4.1.2 Synthesis of Barriers 2.19, 2.20 and 2.21

**Scheme 5: Synthesis of barriers.** Reagents and conditions: (i) 4-(TMS-ethynyl)benzoic acid, HOBt·H$_2$O, EDCI·HCl, CHCl$_3$, THF, r.t., 16 h, 62%. (ii) K$_2$CO$_3$, MeOH, r.t., 16 h, 97%. (iii) Cu(MeCN)$_4$PF$_6$, 3,3'-dimethylbut-1-yne, CH$_2$Cl$_2$, $^t$BuOH, r.t., 4 days. (iv) KOH, THF, MeOH, r.t., 2 h, 76% over 2 steps.

#### 2.19: TMS Unloaded Barrier

To a solution of 2.18$^{[22]}$ (0.256 g, 1.17 mmol, 1.0 eq.) in CHCl$_3$:THF (20 mL, 2:1, 0.05 M) was added 4-(TMS-ethynyl)benzoic acid (1.2 eq.), HOBt·H$_2$O (1.1 eq.) and EDCI·HCl (1.1 eq.). The reaction mixture was stirred at room temperature for 16 hours. The solvent was removed under reduced pressure. Purification by flash column chromatography (SiO$_2$, pet. ether:EtOAc, 2:1) yielded the title compound as a pale yellow foam (0.444 g, 0.718 mmol, 62%).

**m.p.** 129–132 °C; $^1$H-NMR (600 MHz, Acetone-$d_6$) δ 9.36 (s, 1H, r), 8.15 (s, 1H, p), 7.94 (d, J = 7.8 Hz, 1H, i), 7.87 (d, J = 8.1 Hz, 2H, f), 7.55 (dd, J = 8.6, 1.8 Hz, 2H, t), 7.52 (d, J = 8.3 Hz, 2H, e), 7.28 (d, J = 8.6 Hz, 2H, z), 7.19 – 7.14 (m, 4H, m+u), 6.99 (d, J = 8.6 Hz, 2H, aa), 6.73 (d, J = 8.4 Hz, 2H, n), 4.90 (q, J = 7.7 Hz, 1H, j), 3.23 (dd, J = 13.9, 5.9 Hz, 1H, k), 3.07 (dd, J = 13.9, 8.5 Hz, 1H, k'), 1.65 (s, 6H, x) and 0.24 (s, 9H, a). $^{13}$C-NMR (151 MHz, Acetone) δ 170.58 (q), 166.63 (h), 156.99 (o), 148.75 (y), 146.46 (v), 138.10 (ab), 137.59 (s), 135.13 (g), 132.48 (e), 131.21 (m), 129.16 (z+),
128.40 (f), 127.76 (u), 126.83 (d), 120.13 (t), 119.41 (aa), 115.99 (n), 105.07 (c), 96.95 (b), 57.14 (j), 42.88 (w), 37.76 (k), 30.98 (x), -0.15 (a). LRMS (ES+) m/z (rel. intensity) 616 (100%, [M+H]+); HRMS m/z [M+Na]+ calculated for [C36H37N5O3SiNa]+: 638.2558, found: 638.2553. [α]D20 +10.6 (c = 1.01, CH2Cl2). Matches previously reported data.\textsuperscript{[22]}

**2.20: Unloaded Barrier**

To 2.19 (0.267 g, 0.434 mmol, 1.0 eq.) in MeOH (4.3 mL, 0.1 M) was added K2CO3 (0.060 g, 0.43 mmol, 1.0 eq.). The reaction mixture was stirred at room temperature for 16 hours. The solvent was removed under reduced pressure and the crude residue was purified by flash column chromatography (SiO\textsubscript{2}, pet ether:EtOAc, 2:1) to give the title compound as a solid (0.228 g, 0.421 mmol, 97%).

m.p. 149–152 °C; \textsuperscript{1}H-NMR (600 MHz, Acetone-d\textsubscript{6}) \textsuperscript{[p]} δ 9.36 (s, 1H, q), 8.15 (s, 1H, o), 7.94 (d, J = 7.9 Hz, 1H, h), 7.88 (d, J = 7.9 Hz, 2H, e), 7.58 – 7.52 (m, 4H, d+s), 7.28 (d, J = 7.8 Hz, 2H, y), 7.20 – 7.14 (m, 4H, l+t), 6.99 (d, J = 7.7 Hz, 2H, z), 6.73 (d, J = 7.6 Hz, 2H, m), 4.91 (q, J = 7.4 Hz, 1H, i), 3.83 (s, 1H, a), 3.23 (dd, J = 13.9, 5.8 Hz, 1H, j), 3.07 (dd, J = 13.9, 8.5 Hz, 1H, j') and 1.65 (s, 6H, w); \textsuperscript{13}C-NMR (151 MHz, Acetone) δ 170.57 (p), 166.65 (g), 156.99 (n), 148.75 (x), 146.47 (u), 138.10 (aa), 137.57 (r), 135.36 (f), 132.69 (d), 131.21 (l), 129.16 (y+k), 128.42 (e), 127.76 (t), 126.12 (c), 120.14 (s), 119.41 (z), 116.00 (m), 83.47 (b), 81.36 (a), 57.12 (i), 42.88 (v), 37.77 (j), 30.98 (w); LRMS (ES) m/z (rel. intensity) 542 (100%, [M-H]); HRMS m/z [M+Na]+ calculated for [C33H29N5O3Na]+: 566.2163, found: 566.2157. [α]D20 +58.9 (c = 1.00, CH2Cl2).

**2.21: Unloaded Terminal Barrier**

To Cu(MeCN)\textsubscript{4}PF\textsubscript{6} (0.035 g, 0.094 mmol, 0.2 eq.) in a sealed, inerted\textsuperscript{i} microwave vial was added 2.19\textsuperscript{ii} (0.290 g, 0.471 mmol 1.0 eq.) dissolved in a mixture of degassed\textsuperscript{iii} CH\textsubscript{2}Cl\textsubscript{2} and \textsuperscript{7}BuOH (3.5 mL, 4:1). To the resulting black mixture was added 3,3-dimethylbut-1-yne (0.174 mL, 1.413 mmol, 3.0 eq.). The reaction was stirred at room temperature for four days. The mixture was diluted with CH\textsubscript{2}Cl\textsubscript{2} (10 mL) and washed with saturated aqueous EDTA (2 × 15 mL). The organic phase was dried

\textsuperscript{i} Flask was inerted by repeated vacuum evacuation/nitrogen purges.
\textsuperscript{ii} 2.19 was freshly purified by dry flash column chromatography prior to use in this reaction.
\textsuperscript{iii} Solvent was degassed via freeze-pump-thaw degassing.
(MgSO₄) and the solvent removed under reduced pressure. The resulting light red solid was used directly without further purification.

To the crude triazole (assumed 0.471 mmol) in THF (10 mL) was added KOH (0.079 g, 1.413 mmol, 3 eq.) in MeOH (1 mL). The reaction mixture was stirred at room temperature for 2 hours. The reaction was concentrated to an oil under reduced pressure, re-dissolved in EtOAc (10 mL) and acidified with 1M HCl (10 mL). The organic layer was separated and the aqueous phase was extracted with EtOAc (10 mL). The organic layers were combined, dried (MgSO₄) and the solvent removed under reduced pressure. The crude residue was purified by flash column chromatography (SiO₂, 40–45% EtOAc/pet. ether) to give the title compound as a yellow solid (0.225 g, 0.360 mmol, 76%).

m.p. 137–140 °C; ¹H-NMR (600 MHz, Acetone- d₆) δ 9.41 (s, 1H, q), 8.26 (s, 1H, ab), 8.02 (s, 1H, o), 7.97 (d, J = 7.9 Hz, 1H, h), 7.88 (d, J = 8.2 Hz, 2H, e), 7.77 (d, J = 8.7 Hz, 2H, z), 7.59 – 7.53 (m, 4H, s+d), 7.43 (d, J = 8.7 Hz, 2H, y), 7.21 (d, J = 8.7 Hz, 2H, t), 7.18 (d, J = 8.4 Hz, 2H, l), 6.73 (d, J = 8.4 Hz, 2H, m), 4.92 (q, J = 7.8 Hz, 1H, i), 3.83 (s, 1H, a), 3.23 (dd, J = 13.9, 5.9 Hz, 1H, j'), 3.07 (dd, J = 13.9, 8.5 Hz, 1H, j), 1.71 (s, 6H, w), 1.37 (s, 9H, ae); ¹³C-NMR (151 MHz, Acetone) δ 170.63 (p), 170.54 (p'), 166.67 (g), 166.62 (g'), 158.54 (ac), 156.92 (n), 151.83 (x), 146.24 (u), 146.23 (u'), 137.66 (r), 137.57 (r'), 136.17 (aa), 135.33 (f), 135.30 (f'), 132.68 (d), 131.21 (l), 129.13 (k), 128.84 (y), 128.42 (e), 127.81 (t), 126.12 (c), 120.42 (z), 120.20 (s), 118.17 (ab), 115.94 (m), 83.47 (b), 81.36 (a), 57.16 (l), 57.12 (l'), 57.08 (l''), 57.04 (l'''), 43.13 (v), 37.78 (j), 37.74 (j'), 31.43 (ad), 30.94 (w), 30.62 (ae); LRMS (ES⁺) m/z (rel. intensity) 648 (100%, [M+Na⁺]); HRMS m/z [M+H]⁺ calculated for [C₃₉H₃₉N₅O₃H]⁺: 626.3126, found: 626.3125.
2.4.1.3 Loading of 1st Barrier 2.5

Scheme 6: Loading of 1st Barrier. Reagents and conditions: (i) Boc-β-homoPheOH, EDCI·HCl, DMAP, THF, CHCl₃, r.t., 36 h, 74%.

2.5: β-Homophenylalanine-Loaded 1st Barrier

Prepared from 2.20 (0.270 g, 0.497 mmol) and Boc-β-homophenylalanine-OH (0.166 g, 0.596 mmol, 1.2 eq.) using General Procedure 1. Purification by flash chromatography (SiO₂, 0.5% MeOH/CH₂Cl₂) yielded the title compound as a white solid (0.295 g, 0.366 mmol, 74%).

1H-NMR (600 MHz, Acetone-d₆) δ 9.42 (s, 1H, ab), 8.03 (d, J = 7.9 Hz, 1H, h), 7.87 (d, J = 8.0 Hz, 2H, e), 7.57 – 7.52 (m, 4H, d+ad), 7.37 (d, J = 8.1 Hz, 2H, l), 7.32 – 7.26 (m, 6H, t+u+aj), 7.21 (t, J = 6.7 Hz, 1H, v), 7.17 (d, J = 8.6 Hz, 2H, ae), 7.03 (d, J = 8.1 Hz, 2H, m), 6.99 (d, J = 8.6 Hz, 2H, ak), 6.12 (d, J = 8.7 Hz, 1H, w), 4.98 (q, J = 8.3 Hz, 1H, i), 4.31 (sex, J = 7.7 Hz, 1H, q), 3.83 (s, 1H, a), 3.34 (dd, J = 14.0, 5.7 Hz, 1H, j), 3.17 (dd, J = 14.0, 8.7 Hz, 1H, j'), 2.99 – 2.88 (m, 2H, r), 2.76 (dd, J = 14.9, 5.3 Hz, 1H, p), 2.69 (dd, J = 14.9, 8.2 Hz, 1H, p'), 1.65 (s, 6H, ah) and 1.33 (s, 9H, z). 13C-NMR (151 MHz, Acetone) δ 170.41 (o), 170.32 (aa), 166.77 (b), 155.95 (x), 150.60 (n), 148.74 (ai), 146.55 (af), 139.45 (s), 138.11 (al), 137.50 (ac), 136.09 (k), 135.27 (f), 132.69 (d), 131.00 (l), 130.23 (t), 129.16 (u+aj), 128.44 (e), 127.78 (ae), 127.17 (v), 126.16 (c), 122.53 (m), 120.19 (ad), 119.42 (ak), 83.46 (b), 81.38 (a), 78.77 (y), 56.77 (i), 50.59 (q), 42.88 (ag), 41.48 (r), 40.07 (p), 37.82 (j), 30.98 (ah), 28.59 (z); LRMS (ES⁺) m/z (rel. intensity) 805 (100%, [M+H⁺]); HRMS m/z [M+Na⁺] calculated for [C₄₈H₄₈N₆O₆Na⁺]: 827.3474, found: 827.3502.
2.4.1.4 Loading of 2nd Barrier 2.23

Scheme 7: Loading of 2nd Barrier. Reagents and conditions: (i) Boc-β-homoLeu-OH, HOBt·H₂O, EDCI-HCl, CHCl₃, THF, r.t., 16 h, 82%.

2.23: β-Homoleucine-Loaded 2nd Barrier

Prepared from 2.19 (0.104 g, 0.171 mmol) and Boc-β-homoleucine-OH (0.050 g, 0.205 mmol, 1.2 eq.) using General Procedure 1. Purification by flash chromatography (SiO₂, 0.5% MeOH/CH₂Cl₂) yielded the title compound as a pale yellow foam (0.119 g, 0.141 mmol, 82%).

¹H-NMR (600 MHz, Acetone-d₆) δ 9.42 (s, 1H, aα), 8.03 (d, J = 7.9 Hz, 1H, i), 7.86 (d, J = 8.2 Hz, 2H, f), 7.55 (d, J = 8.6 Hz, 2H ac), 7.51 (d, J = 8.2 Hz, 2H, e), 7.37 (d, J = 8.2 Hz, 2H, m), 7.28 (d, J = 8.5 Hz, 2H, ai), 7.17 (d, J = 8.6 Hz, 2H, ad), 7.05 (d, J = 8.2 Hz, 2H, n), 6.99 (d, J = 8.5 Hz, 2H, aj), 5.98 (d, J = 9.1 Hz, 1H, v), 4.99 (q, J = 8.1 Hz, 1H, j), 4.17 (m, 1H, r), 3.34 (dd, J = 14.0, 5.8 Hz, 1H, k), 3.18 (dd, J = 13.9, 8.7 Hz, 1H, k'), 2.70 (dd, J = 14.4, 5.7 Hz, 1H, q), 2.63 (dd, J = 14.4, 7.8 Hz, 1H, q'), 1.77 – 1.69 (m, 1H, t), 1.65 (s, 6H, ag), 1.56 (ddd, J = 14.4, 10.2, 4.5 Hz, 1H, s), 1.37 (m, 10H, γ+s'), 0.94 (d, J = 6.6 Hz, 3H, u), 0.92 (d, J = 6.7 Hz, 3H, u') and 0.24 (s, 9H, a). ¹³C-NMR (151 MHz, Acetone) δ 170.42 (p/z), 170.35 (p/z), 170.26 (p/z), 166.77 (h), 156.18 (w), 150.66 (o), 148.74 (ah), 146.54 (ae), 138.11 (ak), 137.51 (ab), 136.06 (l), 135.05 (g), 132.48 (e), 130.99 (m), 129.16 (ai), 128.42 (f), 127.78 (ad), 126.87 (d), 122.55 (n), 120.19 (ac), 119.42 (aj), 105.07 (c), 96.96 (b), 78.64 (x), 56.78 (j), 47.20 (r), 44.91 (s), 42.89 (af), 41.73 (q), 37.79 (k), 30.99 (ag), 28.64 (y), 25.59 (t), 23.53 (u), 22.13 (u'), -0.14 (a); LRMS (ES') m/z (rel. intensity) 865 (95%, [M+Na]⁺); HRMS m/z [M+Na]⁺ calculated for [C₄₈H₈₉N₆O₆SiNa]⁺: 865.4079, found: 865.4070.
2.4.1.5 Loading of Terminal Barrier 2.25

Scheme 8: Loading of Terminal Barrier. Reagents and conditions: (i) Piv-β-homoAla-OH (2.24), HOBT·H₂O, EDCI·HCl, CHCl₃, THF, r.t., 16 h, 60%.

2.24: Piv-β-homoAla-OH

The following compound was synthesised by Dr Matthew O. Kitching and analysis was carried out by the author. To L-β-homoalanine hydrochloride (1.55 g, 11.1 mmol) in THF (20 mL) and aq. NaOH (1 M, 20 mL) was added pivaloyl chloride (1.7 mL, 13.8 mmol). The reaction mixture was stirred at room temperature for 4 hours. The reaction mixture was concentrated by approximately half in vacuo and diluted with water (10 mL). The mixture was acidified to pH1 using KHSO₄ and extracted with EtOAc (7 × 50 mL). The organic phase was dried (MgSO₄) and the solvent removed under reduced pressure. The resulting solid was dried further by azeotroping with toluene to give the title compound as an analytically pure white powder which was used without further purification (1.24 g, 6.63 mmol, 60%).

m.p. 100–104 °C; ¹H-NMR (600 MHz, DMSO-d₆) δ 12.09 (br. s, 1H, a), 7.22 (d, J = 8.4 Hz, 1H, f), 4.09 (sept., J = 7.2 Hz, 1H, d), 2.41 (dd, J = 15.0 and 6.6 Hz, 1H, c), 2.28 (dd, J = 15.0 and 7.2 Hz, c'), 1.06 (s, 12H, e+i); ¹³C-NMR (151 MHz, DMSO) δ 176.49 (g), 172.65 (b), 41.77 (d), 40.57 (c), 37.90 (h), 27.37 (i), 20.16 (e); LRMS (ES) m/z (rel. intensity) 186 (100%, [M-H]-); HRMS m/z [2M-H]- calculated for [C₁₈H₃₃N₂O₆]: 373.2344, found: 373.2338; [α]D²⁰−18.9 (c = 1.00, MeOH).

2.25: β-Homoalanine-Loaded Terminal Barrier

Prepared from 2.21 (0.110 g, 0.176 mmol) and Piv-β-homoAla-OH 2.24 (0.039 g, 0.211 mmol, 1.2 eq.) using General Procedure 1. Purification by flash chromatography (SiO₂, 1% MeOH/CH₂Cl₂) yielded the title compound as a colourless foam (0.104 g, 0.131 mmol, 74%).
$^1$H-NMR (600 MHz, Acetone-$d_6$) δ 9.44 (s, 1H, x), 8.27 (s, 1H, ai), 8.05 (d, $J = 7.8$ Hz, 1H, h), 7.86 (d, $J = 8.4$ Hz, 2H, e), 7.77 (d, $J = 8.4$ Hz, 2H, ag), 7.56 (d, $J = 9.0$ Hz, 2H, z), 7.54 (d, $J = 8.4$ Hz, 2H, d), 7.43 (d, $J = 8.4$ Hz, 2H, af), 7.37 (d, $J = 8.4$ Hz, 2H, l), 7.21 (d, $J = 9.0$ Hz, 2H, aa), 7.03 (d, $J = 8.4$ Hz, 2H, m), 6.80 (d, $J = 7.8$ Hz, 1H, s), 5.02-4.97 (m, 1H, i), 4.44 (sept., $J = 6.6$ Hz, 1H, q), 3.83 (s, 1H, a), 3.34 (dd, $J = 13.8$ and 6.0 Hz, 1H, j), 3.18 (dd, $J = 13.8$ and 8.4 Hz, 1H, j'), 2.72 (dd, $J = 15.0$ and 7.2 Hz, 1H, p), 2.68 (dd, $J = 15.0$ and 6.6 Hz, 1H, p'), 1.71 (s, 6H, ad), 1.37 (s, 9H, al), 1.24 (d, $J = 6.6$ Hz, 3H, r) and 1.12 (s, 9H, v); $^{13}$C-NMR (151 MHz, Acetone) δ 177.64 (t), 170.45 (o/w), 170.41 (o/w), 166.78 (g), 158.55 (aj), 151.83 (ae), 150.58 (n), 146.29 (ab), 137.64 (y), 136.19 (k), 136.11 (ah), 135.27 (f), 132.68 (d), 131.04 (l), 128.85 (af), 128.47 (e), 127.82 (aa), 126.15 (c), 122.49 (m), 120.42 (ag), 120.25 (z), 118.17 (ai), 83.47 (b), 81.37 (a), 56.88 (i), 43.23 (q/ac), 43.14 (q/ac), 41.68 (p), 38.98 (u), 37.84 (j), 31.43 (ak), 30.95 (ad), 30.62 (al), 27.82 (v), 20.62 (r); LRMS (ES') $m/z$ (rel. intensity) 795 (75%, [M+H]) and 817 (100%, [M+Na]); HRMS $m/z$ [M+H]$^+$ calculated for [C$_{48}$H$_{54}$N$_6$O$_5$H]: 795.4228, found: 795.4221.

### 2.4.1.6 Synthesis of 2-Barrier Extension Piece.

![Scheme 9: Synthesis of 2-Barrier Extension Piece. Reagents and conditions (i) Cu(MeCN)$_4$PF$_6$, Tentalgel$^{TM}$ TBTA resin, CH$_2$Cl$_2$, $^t$BuOH, r.t., 16 h. (ii) TSAF, DMF, r.t., 2 mins, 59% over 2 steps.](image-url)
Tentagel™ TBTA resin (0.17 mmol/g, 220 mg, 37.5 μmol, 0.5 eq.) and Cu(MeCN)_4PF_6 (13 mg, 34.1 μmol, 0.45 eq.) were inerted iv in a sealed microwave vial and dissolved in degassed CH_2Cl_2 (2.0 mL). The mixture was stirred for 10 minutes. To the mixture was added a solution of 2.23 (63 mg, 0.075 mmol, 1.0 eq.) and 2.25 (59 mg, 0.075 mol, 1.0 eq.) in degassed CH_2Cl_2: tBuOH (2.5 mL, 4:1). The reaction was stirred for 16 hours at room temperature. The reaction was filtered (to recover and recycle TBTA resin) and the solvent was removed under reduced pressure. The crude product was used directly without further purification.

The crude product was dissolved in DMF (1.5 mL, 0.05M) and TSAF (40 mg, 0.15 mmol, 2 eq.) was added. The reaction mixture was stirred at room temperature for 2 minutes. The reaction was poured into a stirred mixture of CH_2Cl_2 (10 mL) and saturated aqueous NH_4Cl (10 mL). The organic phase was isolated, dried (Na_2SO_4) and the solvent removed under reduced pressure. The crude residue was purified by prep. TLC (SiO_2 (Merck 1000), 5% IPA/CH_2Cl_2) to give the title compound as a pale yellow glass (69 mg, 44.1 μmol, 59% over 2 steps).

^1H-NMR (600 MHz, Acetone-d_6) δ 9.46 (s, 2H, z+bh), 9.04 (s, 1H, ak), 8.26 (s, 1H, bs), 8.07 – 8.01 (m, 4H, h+an+ar), 7.98 (d, J = 8.4 Hz, 2H, ao), 7.88 (d, J = 8.5 Hz, 2H, e), 7.86 (d, J = 8.8 Hz, 2H, ai), 7.77 (d, J = 8.7 Hz, 2H, bq), 7.60 – 7.56 (m, 4H, ab+bj)), 7.55 (d, J = 8.3 Hz, 2H, d), 7.49 (d, J = 8.7 Hz, 2H, ah), 7.43 (d, J = 8.7 Hz, 2H, bp), 7.40 (d, J = 8.5 Hz, 2H, l/av), 7.38 (d, J = 8.4 Hz, 2H, l/av), 7.24 (d, J = 8.9 Hz, 2H, ac/bk), 7.22 (d, J = 8.9 Hz, 2H, ac/bk), 7.07 – 7.02 (m, 4H, m+aw), 6.79 (d, J = 8.2 Hz, 1H, bc), 5.97 (d, J = 9.1 Hz, 1H, u), 5.06 – 4.98 (m, 2H, i+as), 4.44 (dp, J = 13.7, 6.8 Hz, 1H, ba), 4.22 – 4.13 (m, 1H, q), 3.82 (s, 1H, a), 3.39 – 3.33 (m, 2H, j+at), 3.24 – 3.16 (m, 2H, j+at‘), 2.75 – 2.66 (m, 3H, p+az+az’), 2.63 (dd, J = 14.4, 7.8 Hz, 1H, p’), 1.78 – 1.67 (m, 13H, s+af+bn), 1.56 (ddd, J = 14.4, 10.2, 4.8 Hz, 1H, r), 1.44 – 1.32 (m, 19H, r’+x+bv), 1.24 (d, J = 6.7 Hz, 3H, bb), 1.11 (s, 9H, bf), 0.94 (d, J = 6.5 Hz, 3H, t) and 0.92 (d, J = 6.7 Hz, 3H, t’);

^13C-NMR (151 MHz, Acetone) δ 177.68

iv Flask was inerted by repeated vacuum evacuation/argon purges.

v Solvent was degassed by sparging with argon for 30 minutes before use.
2.4.1.7 Synthesis of 1-Barrier Rotaxane 2.2 and Free Thread 2.6

Scheme 10: Synthesis of 1-Barrier Rotaxane 2.2 and free thread 2.6. Reagents and conditions: (i) Cu(MeCN)$_4$PF$_6$, CH$_2$Cl$_2$, tBuOH, r.t., 48 h, 40% (rotaxane 2.2), 7% (free thread 2.6).
2.2: 1-Barrier Rotaxane

Aldehyde macrocycle 2.3
(0.147 g, 0.262 mmol, 1.9 eq.)
and Cu(MeCN)₄PF₆ (0.029 g,
0.068 mmol, 0.5 eq.) were
dissolved in degassed⁵
CH₂Cl₂:BuOH (3 mL, 6:1).
The mixture was stirred for
30 minutes. To the mixture
was added a solution of stopper 2.4⁵⁵ (0.401 g, 0.683 mmol, 5.0 eq.) and barrier 2.5 (0.110 g,
0.137 mmol, 1.0 eq.) in degassed CH₂Cl₂:BuOH (2 mL, 6:1). The reaction mixture was stirred at room temperature for 16 h. The solvent was removed under reduced pressure. The crude residue
was purified by flash column chromatography (SiO₂, 3% MeOH/CH₂Cl₂) and subsequent prep. TLC
(SiO₂ (Analtech 1000), 40% EtOAc/n-hexane) to give the title compound as a glass (0.101 g,
0.054 mmol, 40%).

¹H-NMR (600 MHz, Acetone-d₆) δ 9.94 (s, 1H, ms), 9.44 (s, 1H, ap), 8.01 (s, 1H, o), 7.95 (d, J = 7.9
Hz, 1H, v), 7.84 (d, J = 8.3 Hz, 2H, s), 7.80 (d, J = 8.3 Hz, 2H, r), 7.60 – 7.51 (m, 6H, ma+mq+mo+ar),
7.38 (d, J = 8.3 Hz, 2H, z), 7.33 – 7.24 (m, 12H, d+ah+ai+ax), 7.22 – 7.18 (m, 1H, aj), 7.17 (d, J = 8.7
Hz, 2H, as), 7.11 (d, J = 8.5 Hz, 6H, e), 7.05 – 7.01 (m, 4H, mb+aa), 6.98 (d, J = 8.5 Hz, 2H, ay), 6.93
(d, J = 8.8 Hz, 2H, l), 6.76 (d, J = 8.3 Hz, 4H, mh), 6.48 – 6.41 (m, 6H, mi+j), 6.11 (d, J = 8.7 Hz, 1H,
ak), 5.00 (q, J = 7.9 Hz, 1H, w), 4.30 (dq, J = 13.2, 6.8 Hz, 1H, ae), 4.05 (t, J = 7.0 Hz, 2H, n), 3.83 –
3.72 (m, 4H, mk), 3.46 (t, J = 6.0 Hz, 2H, l), 3.33 (dd, J = 14.0, 5.9 Hz, 1H, x), 3.16 (dd, J = 13.9, 8.4
Hz, 1H, x'), 2.97 – 2.88 (m, 2H, af), 2.78 – 2.65 (m, 6H, mn+ad), 2.63 – 2.56 (m, 4H, mcl), 2.50 (t, J =
6.7 Hz, 4H, ml), 1.91 – 1.75 (m, 10H, me+mm+m), 1.73 (dt, J = 13.7, 6.0 Hz, 4H, m1l), 1.65 (s, 6H,
av), 1.33 (s, 9H, an) and 1.30 (s, 27H, a); ¹³C-NMR (151 MHz, Acetone) δ 193.15 (ms), 170.42
(ac/ao), 170.40 (ac/ao), 170.34 (ac/ao), 167.21 (u), 162.50 (mc), 162.49 (mc'), 157.99 (mj), 157.24
(k), 155.95 (al), 150.58 (ab), 149.06 (c), 148.74 (aw), 146.87 (p), 146.51 (at), 145.31 (f),
144.71 (mo), 140.12 (h), 139.43 (ag), 138.09 (az), 137.93 (mr), 137.61 (ma/aq), 137.55 (ma/aq),
137.46 (ma/aq), 136.18 (y), 135.56 (mp), 135.20 (q), 134.24 (mg), 133.88 (t), 132.56 (l), 131.41 (e),
131.02 (z), 130.23 (ah), 129.98 (mh), 129.15 (ai+ax), 128.73 (s), 127.94 (mq), 127.91 (mq'),
127.78 (as), 127.16 (aj), 125.93 (r), 125.05 (d), 122.53 (aa), 122.23 (o), 120.61 (mb), 120.17 (ar),

⁵ Solvent was degassed by sparging with Argon for 30 minutes before use.
Free thread 2.6 was also recovered as a colourless glass (18 mg, 9.6 μmol, 7%).

\[ \text{\textsuperscript{1}H-NMR (600 MHz, Acetone-\text{d}_6)} \]
\[ \delta \text{ 9.43 (s, 1H, ap), 8.48 (s, 1H, o), 7.98 (dd, } J = 7.8, 3.1 \text{ Hz, 1H, v), 7.93 (s, 4H, r+s), 7.55 (d, } J = 8.7 \text{ Hz, 2H, ar), 7.38 (d, } J = 8.1 \text{ Hz, 2H, z), 7.31 (d, } J = 8.6 \text{ Hz, 6H, d), 7.28 (m, 6H, ah+ai+ax), 7.20 (t, } J = 6.5 \text{ Hz, 1H, aj), 7.17 (d, } J = 8.7 \text{ Hz, 2H, as), 7.12 (d, } J = 8.6 \text{ Hz, 6H, e), 7.09 (d, } J = 8.9 \text{ Hz, 2H, i), 7.03 (d, } J = 8.1 \text{ Hz, 2H, aa), 6.98 (d, } J = 8.6 \text{ Hz, 2H, ay), 6.84 (d, } J = 8.9 \text{ Hz, 2H, j), 6.12 (d, } J = 8.6 \text{ Hz, 1H, ak), 5.00 (q, } J = 7.8 \text{ Hz, 1H, w), 4.69 (t, } J = 6.9 \text{ Hz, 2H, n), 4.30 (sex., } J = 7.7 \text{ Hz, 1H, ae), 4.07 (t, } J = 5.9 \text{ Hz, 2H, l), 3.35 (dd, } J = 14.0, 5.8 \text{ Hz, 1H, x'), 3.19 (dd, } J = 13.9, 8.6 \text{ Hz, 1H, x}, 2.92 (h, } J = 6.7, 6.3 \text{ Hz, 2H, af), 2.76 (dd, } J = 15.0, 5.4 \text{ Hz, 1H, ad'), 2.69 (dd, } J = 14.8, 8.2 \text{ Hz, 1H, ad'}, 2.45 (p, } J = 6.4 \text{ Hz, 2H, m), 1.65 (s, 6H, av), 1.32 (d, } J = 2.3 \text{ Hz, 9H, an) and 1.29 (s, 27H, a);} \text{\textsuperscript{13}C-NMR (151 MHz, Acetone)} \]
\[ \delta 170.46 (ac/ao), 170.40 (ac/ao), 167.18 (u), 157.64 (k), 155.95 (al), 150.58 (ab), 149.15 (c), 148.74 (aw), 146.96 (q), 146.51 (at), 145.23 (f), 140.47 (h), 139.43 (ag), 138.09 (az), 137.52 (aq), 136.16 (y), 135.26 (p/t), 134.14 (p/t), 132.80 (l), 131.40 (e), 131.02 (z), 130.23 (ah), 129.15 (al+ax), 128.86 (r/s), 127.78 (as), 127.16 (aj), 125.91 (r/s), 125.08 (d), 122.52 (aa), 122.49 (o), 120.10 (ar), 119.41 (ay), 114.09 (j), 78.77 (am), 65.23 (l), 63.88 (g), 56.75 (w), 50.57 (ae), 47.91 (n), 42.88 (au), 41.47 (af), 40.05 (ad), 37.81 (x), 34.85 (b), 31.64 (a), 30.99 (av), 30.81 (m), 28.60 (an); LRMS (ES') m/z (rel. intensity) 1392 (100%, [M+H]+) and 1414 (85%, [M+Na]+); HRMS m/z [M+Na]+ calculated for [C_{88}H_{37}N_{9}O_{7}Na]: 1414.7403, found: 1414.7360.
2.4.1.8 Synthesis of 3-Barrier Molecular Machine 2.1

Scheme 11: Synthesis of 3-Barrier Molecular Machine 2.1. Reagents and conditions: (i) Cu(MeCN)$_4$PF$_6$, Tentalgel™ TBTA resin, CH$_2$Cl$_2$, tBuOH, r.t., 48 h, 97%; (ii) Boc-Gly-Gly-OH, EDCI·HCl, HOBT·H$_2$O, DMF, r.t., 16 h, quant.; (iii) NH$_2$NH$_2$ (50-60% in H$_2$O), EtOH, r.t., 48 h, 65%; (iv) Aniline, CH$_2$Cl$_2$, r.t., 2 days, 72%.
2.8: 3-Barrier Aldehyde Rotaxane

Tentagel™ TBTA resin (350 mg, 60 μmol, 3.0 eq.) and Cu(MeCN)$_2$PF$_6$ (12 mg, 30 μmol, 1.5 eq.) were inerted$^{ii}$ in a sealed microwave vial and dissolved in degassed$^{vii}$ CH$_2$Cl$_2$ (2.5 mL). The mixture was stirred for 10 minutes. To the mixture was added a solution of 2.2 (40 mg, 20.5 μmol, 1.0 eq.) and 2.7 (45 mg, 24.6 μmol, 1.4 eq.) in degassed CH$_2$Cl$_2$:$^6$BuOH (1.5 mL, 1:1). The reaction was stirred for 48 hours at room temperature. The reaction was filtered (to recover and recycle TBTA resin) and the solvent was removed under reduced pressure. The crude residue was purified by prep. TLC (SiO$_2$ (Merck 1000), 3% MeOH/CH$_2$Cl$_2$) to give the title compound as a glass (70 mg, 19.8 μmol, 97%).

$^1$H-NMR (600 MHz, Acetone-$d_6$) δ 9.94 (s, 1H, ms), 9.47-9.42 (m, 3H, ap+bz+di), 9.07-9.03 (s, 2H, ba+ck), 8.26 (s, 1H, dt), 8.07-7.94 (m, 12H, o+v+be+bd+bh+cn+co+cr), 7.89-7.83 (m, 6H, s+ay+ci), 7.82-7.74 (m, 4H, r+dr), 7.61-7.47 (m, 14H, ar+ax+cb+ch+dk+ma+mp+mq), 7.46-7.36 (m, 8H, z+bl+cw+dq), 7.31-7.17 (m, 17H, d+ah+ai+aj+as+cc+dl), 7.10 (d, J = 8.4 Hz, 6H, e), 7.08-6.99 (m, 8H, aa+bm+cx+mh), 6.93 (d, J = 9.0 Hz, 2H, l), 6.79 (d, J = 7.8 Hz, 1H, dd), 6.75 (d, J = 7.8 Hz, 4H, mh), 6.48-6.40 (m, 6H, j+mi), 6.11 (d, J = 9.0 Hz, 1H, ak), 5.98 (d, J = 9.0 Hz, 1H, bu), 5.05-4.97 (m, 3H, w+bi+cs), 4.44 (sept., J = 6.6 Hz, 1H, db), 4.33-4.27 (m, 1H, ae), 4.21-4.13 (m, 1H, bq), 4.05 (t, J = 7.2 Hz, 2H, n), 3.82-3.71 (m, 4H, mk), 3.46 (t, J = 6.0 Hz, 2H, l), 3.38-3.31 (m, 3H, x+bj+cu), 3.23-3.14 (m, 3H, x'+bj'+cu'), 2.95-2.87 (m, 2H, af), 2.78-2.66 (m, 9H, ad+bp+da+mn), 2.65-2.56 (m, 5H, md+bp'), 2.49 (t, J = 7.2 H, 4H, mf), 1.91-1.69 (m, 33H, m+av+bs+cf+do+me+ml+mn), 1.58-1.52 (m, 1H, br), 1.41-1.26 (m, 55H, a+an+br'+bx+dw), 1.24 (d, J = 7.2 Hz, 3H, dc), 1.11 (s, 9H, dg), 0.92 (d, J = 6.0 Hz, 3H, bt) and 0.91 (d, J = 6.6 Hz, 3H, bt'); $^{13}$C-NMR (151 MHz, DMSO) δ 193.54, 177.14, 170.77, 170.73, 170.18, 170.13, 170.06, 166.49, 161.47, 157.75, 156.87, 156.00, 155.57, 155.35, 151.78, 151.09, 149.47, 149.43, 148.08, 146.89, 145.97, 145.28, 145.18, 144.44, 143.78, 139.42,

$^{ii}$Flask was inerted by repeated vacuum evacuation/argon purges.

$^{vii}$Solvent was degassed by sparging with argon for 30 minutes before use.
139.05, 138.92, 137.16, 137.10, 135.08, 134.80, 134.74, 133.73, 133.56, 133.51, 133.20, 131.47, 130.57, 130.53, 130.50, 129.64, 129.34, 128.75, 128.64, 128.47, 128.27, 127.42, 127.33, 127.29, 126.62, 125.41, 125.05, 124.67, 121.83, 120.90, 120.27, 120.13, 120.07, 119.66, 118.68, 114.37, 113.44, 78.09, 78.02, 67.23, 64.00, 62.91, 56.22, 55.40, 49.84, 46.76, 46.43, 44.13, 42.65, 42.56, 42.44, 41.27, 41.22, 39.97, 38.37, 37.55, 34.81, 34.75, 34.48, 32.04, 31.59, 30.99, 30.77, 30.67, 29.35, 28.67, 28.62, 28.09, 27.80, 24.79, 23.51, 22.13, 20.80.

**LRMS (ES) m/z (rel. intensity) 3520 (100%, [M+H]+); HRMS m/z (100) [M+2H]²⁺ calculated for [C₂₁₉H₂₄₄N₂₂O₂₁H₂]²⁺: 1760.9455, found: 1760.9682; m/z (100) [M+H+Na]²⁺ calculated for [C₂₁₉H₂₄₄N₂₂O₂₁HNa]²⁺: 1771.9365, found: 1771.9540.

Figure 3: Observed (upper) and calculated (lower) HRMS isotope patterns of 3-barrier rotaxane 2.8. Species are doubly charged. The isotopic distribution centred at m/z 1760.9682 corresponds to [M+2H]²⁺: [C₂₁₉H₂₄₄N₂₂O₂₁H₂]²⁺. The isotopic distribution centred at m/z 1771.9540 corresponds to [M+H+Na]²⁺: [C₂₁₉H₂₄₄N₂₂O₂₁HNa]²⁺.
2.26: Trityl-Cysteine Ethyl Ester

L-Cysteine ethyl ester hydrochloride (7.45 g, 40.1 mmol, 1.1 eq.) and triphenylmethanol (9.50 g, 36.5 mmol, 1 eq.) were dissolved in TFA (25 mL). The orange solution was stirred at room temperature for 75 min the solvent was removed in vacuo. The resulting brown oil was dissolved in CH₂Cl₂ (200 mL) and water (200 mL) was added upon which the brown organics turned colourless. The mixture was neutralised by careful portion-wise addition of K₂CO₃ (ca. 15 g). The organic layer was separated, dried over MgSO₄ and concentrated to a light purple oil. The crude oil was triturated with n-hexane which was then decanted to give the title compound as a yellow oil (12.9 g, 33.0 mmol, 82%).

¹H-NMR (600 MHz, DMSO-d₆) δ 7.37 – 7.29 (m, 12H, h+i), 7.27 – 7.21 (m, 3H, j), 4.02 (q, J = 7.1 Hz, 2H, b), 3.14 (t, J = 6.4 Hz, 1H, d), 2.36 (dd, J = 11.9, 6.2 Hz, 1H, e), 2.28 (dd, J = 11.9, 6.7 Hz, 1H, e'), 1.82 (s, 2H, k), 1.13 (t, J = 7.1 Hz, 3H, a); ¹³C-NMR (151 MHz, DMSO) δ 173.7 (c), 144.4 (g), 129.1 (i), 128.0 (h), 126.7 (j), 65.9 (f), 60.2 (b), 53.7 (d), 36.4 (e), 14.1 (a); LRMS (ES⁺) m/z (rel. intensity) 414 (100%, [M+Na]⁺); 805 (73%, [2M+Na]⁺); HRMS m/z [M+H]⁺ calculated for [C₂₄H₂₅N₁O₂SH]: 392.1679, found: 392.1673; [α]D²⁰ +80.0 (c = 1.00, MeOH).

2.27: Boc-Gly-Gly-Cys(Trt)-OEt

To 2.26 (2.53 g, 6.46 mmol, 1.0 eq.) in DMF (130 mL, 0.05 M) was added Boc-Gly-Gly-OH (1.50 g, 6.46 mmol, 1.0 eq.), EDCI-HCl (1.92 g, 9.69 mmol, 1.5 eq.) and HOBt·H₂O (0.20 g, 1.29 mmol, 0.2 eq.) The reaction mixture was stirred at room temperature for 16 hours. The solvent was removed under reduced pressure and the crude residue azeotroped three times with toluene (to remove residual DMF). The crude residue was dissolved in CH₂Cl₂ (100 mL), washed with 0.5 M HCl (100 mL), saturated aqueous NaHCO₃ and brine. The organic phase was dried (Na₂SO₄) and the solvent removed under reduced pressure. The resulting pale yellow solid was analytically pure and was used without further purification (3.90 g, 6.44 mmol, quant.).

m.p. 74–79 °C; ¹H-NMR (600 MHz, DMSO-d₆) δ 8.36 (d, J = 7.8 Hz, 1H, k), 7.96 (t, J = 5.4 Hz, 1H, n), 7.34 (t, J = 7.8 Hz, 6H, i), 7.29 (dd, J = 7.2 6H, h), 7.26 (t, J = 7.2 Hz, 3H, j), 7.00 (t, J = 6.0 Hz, 1H, q), 4.12-4.07 (m, 1H, d), 4.03-3.95 (m, 2H, b), 3.73 (d, J = 6.0 Hz, 2H, m), 3.55 (d, J = 6.0 Hz, 2H, p), 2.53 (dd, J = 12.6 and 8.4 Hz, 1H, e), 2.36 (dd, J = 12.0 and 5.4 Hz, 1H, e'), 1.37 (s, 9H, t) and 1.09 (t, J = 6.6 Hz, 3H, a); ¹³C-NMR (151 MHz, DMSO) δ 169.96 (c), 169.53 (o), 168.81 (l), 155.80 (r),
144.06 (g), 129.07 (h), 128.14 (i), 126.88 (j), 78.08 (s), 66.39 (f), 60.83 (b), 51.55 (d), 43.23 (m), 41.41 (p), 32.83 (e), 28.21 (t), 13.94 (a);
LRMS (ES') m/z (rel. intensity) 628 (100%, [M+Na]'),
HRMS m/z [M+Na]^+ calculated for [C_{33}H_{39}N_{3}O_{5}SNa]^+: 628.2452, found: 628.2441; [α]_{D}^{20} +208.7 (c = 1.03, CH₂Cl₂).

2.9: Catalytic Unit Hydrazide

To a stirred solution of 2.27 (1.50 g, 2.48 mmol) in EtOH (2 mL) at room temperature was added hydrazine monohydrate (0.3 mL, 50-60% in H₂O, 1.2 eq.). After 48 hours the formed precipitate was filtered, dried over azeotrope with anhydrous toluene (3 x 4 mL) and dried under vacuum to afford the title compound (0.95 g, 1.61 mmol, 65%) as a white amorphous solid, which was used without further purification.

m.p. 155–157 °C; ¹H NMR 1H NMR (600 MHz, Chloroform-d) δ 7.89 (br. s, 1H, b), 7.39 (d, J = 7.6 Hz, 6H, h), 7.28 (t, J = 7.7 Hz, 6H, i), 7.21 (t, J = 7.3 Hz, 3H, j), 7.08 (ap. s, 1H, s), 6.73 (d, J = 7.2 Hz, 1H, k), 5.42 (ap. s, 1H, q), 4.15 – 4.05 (m, 1H, d), 3.88 (dd, J = 16.7, 5.5 Hz, 1H, m), 3.84 – 3.71 (m, 3H, m'+p), 2.72 (dd, J = 13.0, 7.6 Hz, 1H, e), 2.58 (dd, J = 12.9, 5.5 Hz, 1H, e'), 1.43 (s, 9H, t); ¹³C NMR (151 MHz, Chloroform) δ 170.67 (o), 170.25 (c), 168.93 (l), 156.47 (r), 144.41 (g), 129.67 (h), 128.24 (l), 127.11 (j), 80.69 (s), 67.41 (f), 51.38 (d), 44.34 (p), 43.00 (m), 33.46 (e), 28.48 (t); LRMS (ES') m/z (rel. intensity) 592 (100%, [M+H]'); HRMS m/z [M+H]^+ calculated for [C_{31}H_{37}N_{5}O_{5}S]^+: 592.2584, found: 592.2583; [α]_{D}^{20} + 6.2 (c = 1.08, CH₂Cl₂).
2.1: 3-BARRIER CATALYST-LOADED β-PEPTIDE SYNTHESISER

To 2.8 (52 mg, 14.8 μmol) and hydrazide 2.9 (10.5 mg, 17.7 μmol, 1.2 eq.) in CH₂Cl₂ (2 mL) was added aniline (5 drops). The reaction mixture was stirred at room temperature for 2 days. The solvent was removed under reduced pressure. Purification by prep. TLC (SiO₂ (Merck 1000), 4% MeOH/CH₂Cl₂) and precipitation with n-hexane gave the title compound as a white powder (44 mg, 10.7 μmol, 72%).

¹H NMR (600 MHz, DMF-d₇) δ 11.39 – 11.30 (m, 1H, xa), 10.37 – 10.30 (m, 3H, ap+bz+di), 8.40 (s, 2H, ba+ck), 8.84 (d, J = 7.7 Hz, 2H), 8.78 (d, J = 7.1 Hz, 1H), 8.58 (s, 1H, dt), 8.41 – 8.31 (m, 1.5H), 8.26 – 8.22 (m, 1H), 8.19 (s, 0.5H), 8.09 (ap. d, J = 2.2 Hz, 8H), 7.97 – 7.92 (m, 4H), 7.91 – 7.86 (m, 4H), 7.72 – 7.65 (m, 6H), 7.57 (t, J = 7.8 Hz, 1H, aj/ma), 7.56 – 7.52 (m, 4H), 7.52 – 7.43 (m, 10H), 7.43 – 7.35 (m, 10H), 7.33 (d, J = 8.3 Hz, 6H, d/e), 7.31 – 7.19 (m, 17H), 7.17 (t, J = 7.2 Hz, 1H, aj/ma), 7.15 – 7.03 (m, 15H), 6.98 (t, J = 5.8 Hz, 0.5H), 6.96 – 6.90 (m, 3H), 6.84 – 6.76 (m, 5H), 6.52 – 6.47 (m, 4H), 6.45 (d, J = 8.7 Hz, 2H), 5.57 (td, J = 8.4, 4.3 Hz, 0.5H, xc), 5.08 – 5.00 (m, 3H, w+bi+cs), 4.60 – 4.53 (m, 0.5H, xc'), 4.43 (sept., J = 7.0 Hz, 1H, db), 4.28 (dq, J = 15.7, 8.5, 8.0 Hz, 1H, ae), 4.19 – 4.07 (m, 3H, bq+n), 4.01 (t, J = 4.6 Hz, 1H), 3.94 (t, J = 5.3 Hz, 1H), 3.87 – 3.72 (m, 6H, mk+others), 3.48 – 3.42 (m, 4H, l+others), 3.40 – 3.30 (m, 4H), 3.29 – 3.19 (m, 4H), 2.69 – 2.56 (m, 11H), 2.51 (s, 5H), 1.95 – 1.66 (m, 37H, av+cf+do+others), 1.55 (ddd, J = 14.6, 10.4, 4.9 Hz, 1H, br), 1.39 (m, 30H), 1.32 (m, 9H, 1Bu), 1.29 – 1.27 (m, 32H, 1Bu+others), 1.23 (d, J = 6.7 Hz, 3H, dc), 1.12 (s, 9H, dg), 0.91 (d, J = 6.6 Hz, 3H, bt), 0.89 (d, J = 6.7 Hz, 3H, bt'); The ¹H NMR spectrum could not be fully assigned due to its high complexity leading to overlapping peaks in both 1D and 2D experiments in addition to spectra being run in a different solvent to preceding structures (partial assignment is given); ¹³C NMR (151 MHz, DMF) δ 177.12, 170.42, 169.98, 169.96, 169.92, 166.39, 161.61, 157.81, 157.10, 156.35, 155.66, 155.45, 153.79, 151.75, 151.07, 149.68, 149.64, 148.34, 147.01, 146.09, 145.21, 145.10, 144.76, 144.71, 144.51, 143.29, 143.27, 139.13, 138.89, 137.38, 137.34, 136.86, 136.02, 135.23, 134.88, 133.88, 133.66, 133.57, 131.59, 130.50, 130.28, 130.25,
Chapter 2

129.48, 129.43, 129.43, 129.24, 128.33, 128.28, 128.18, 128.12, 128.07, 127.02, 126.99, 126.93, 126.83, 126.32, 125.14, 124.43, 121.79, 121.65, 121.61, 120.26, 119.86, 119.52, 119.29, 118.08, 114.24, 113.42, 78.45, 78.33, 77.81, 77.71, 67.25, 67.23, 66.39, 66.11, 64.17, 63.07, 56.37, 56.33, 51.50, 49.93, 46.83, 46.45, 44.07, 42.47, 42.45, 42.36, 41.02, 40.64, 40.13, 39.54, 38.18, 37.58, 37.00, 32.07, 30.92, 30.70, 30.17, 28.81, 28.76, 28.17, 27.93, 27.90, 27.14, 24.74, 22.81, 21.39, 20.02; LRMS (ES') m/z (rel. intensity) 1379 (100%, [M+H+2Na]3+) and 2057 (45%, [M+H+Na]2+); HRMS m/z (100) [M+H+Na] calculated for \( [\text{C}_{250}\text{H}_{279}\text{N}_{27}\text{O}_{25}\text{S}\text{HNa}]^{2+} \): 2058.5561, found: 2058.5544; m/z (100) [M+2H+Na]2+ calculated for \( [\text{C}_{250}\text{H}_{279}\text{N}_{27}\text{O}_{25}\text{SH}_{2}\text{Na}]^{3+} \): 1372.7068, found: 1372.7098.

Figure 4: Observed (upper) and calculated (lower) HRMS isotope patterns of 3-barrier rotaxane 2.1. Species are triply charged. The isotopic distribution centred at m/z 1372.7098 corresponds to [M+2H+Na]2+: \( [\text{C}_{250}\text{H}_{279}\text{N}_{27}\text{O}_{25}\text{H}_{2}\text{Na}]^{3+} \).
2.4.1.9 Operation of artificial β-Peptide Synthesiser 2.1

Scheme 12: Operation of Artificial β-3-Peptide Synthesiser. Reagents and conditions: (i) TIPS, TFA, CH₂Cl₂, r.t., 1 h; (ii) NEt₃, PPh₃, DMF, 65 °C, 7 days, 29% of 2.11b over 2 steps.

2.11b: Machine Operation Product

Molecular machine 2.1 (8 mg, 1.95 μmol, 1 eq.) was dissolved in CH₂Cl₂ (0.20 mL) and TFA (0.05 mL) was added followed immediately by TIPS (1.68 μL, 8.21 μmol, 4.2 eq.) upon which the yellow solution became colourless in ca. 5 min. The resulting solution was stirred for 1 h at room temperature. Toluene (5 mL) was added and the solution concentrated to dryness in vacuo. The resulting solid was twice further dried by azeotrope with toluene (2 x 5 mL) in vacuo. The resulting solid was triturated with Et₂O:n-hexane (1:1, 3 x 5 mL) to remove triphenylmethane. The remaining solid was dried thoroughly in vacuo to give the crude deprotected machine in assumed quantitative yield.

The crude deprotected machine (assumed 1.95 μmol, 1 eq.) and PPh₃ (1.8 mg, 6.86 μmol, 3.5 eq.) were dissolved in DMF (3.2 mL, 0.6 mM). NEt₃ (12 μL, 86.3 μmol, 44 eq.) was added and the
resulting solution was heated at 60 °C for 7 days. The operation mixture was cooled to room temperature and solvent was removed in vacuo. The crude operation mixture was analysed by LRMS indicating the formation of 2.11a:

![LRMS spectrum of 2.11a](image)

Figure 5: Observed (top) and predicted (bottom) LR(ESI)-(+)-MS of product of operation 2.11a recorded from the crude operation mixture. The predicted isotope pattern is for [C\textsubscript{71}H\textsubscript{95}N\textsubscript{9}O\textsubscript{9}SH]\textsuperscript{+} corresponding to [2.11a+H]\textsuperscript{+}. Double macrocycle S,N-acetal 2.11b was not observed on the crude mass spectrum but was formed upon purification.

The crude operation mixture was purified by prep. TLC (SiO\textsubscript{2} (Merck 500), 10% MeOH/CH\textsubscript{2}Cl\textsubscript{2}) to give product material consistent with the title compound 2.11b (1.0 mg, 29%) as a complex mixture of isomers.
2.4.1.10 Synthesis of Authentic Standard mixed αβ3-Peptide 2.31

Scheme 13: Synthesis of reference hexa-peptide 2.31: Reagents and conditions: (i) Ph₃COH, TFA, CH₂Cl₂, r.t., 2 h; (ii) Boc-β-homophenylalanine-OH, HOBt·H₂O, EDCI·HCl, CH₂Cl₂, r.t., 16 h, 54% over 2 steps; (iii) Ph₃COH, TFA, CH₂Cl₂, r.t., 2 h; (iv) Boc-β-homoleucine-OH, HOBt·H₂O, EDCI·HCl, CH₂Cl₂, r.t., 16 h, 94% over 2 steps; (v) Ph₃COH, TFA, CH₂Cl₂, r.t., 2 h; (vi) Piv-β-homoalanine-OH, HOBt·H₂O, EDCI·HCl, CH₂Cl₂, r.t., 16 h, 74% over 2 steps; (vii) NH₂NH₂, EtOH, 90 °C, 48 h; (viii) 2.3, aniline, CH₂Cl₂, DMF, r.t., 16 h, 48% over 2 steps.

2.28: Boc-β³Phe-Gly-Gly-Cys(Trt)-OEt

Prepared from 2.27 (0.784 g, 1.66 mmol) using General Procedure 2 and then General Procedure 3 on 50% of the crude deprotected material. Purification by flash chromatography (SiO₂, 5% MeOH/CHCl₃) yielded the title compound as a white solid (0.344 g, 0.449 mmol, 54% over 2 steps).

m.p. 89–91 °C; ¹H-NMR (600 MHz, DMSO-d₆)  δ 8.36 (d, J = 7.8 Hz, 1H, k), 8.12 (t, J = 6.0 Hz, 1H, NH, q), 8.05 (t, J = 6.0 Hz, 1H, NH, n), 7.37-7.32 (t, J = 7.7 Hz, 6H, i), 7.31-7.27 (d, J = 7.1 Hz, 6H, h), 7.27 – 7.23 (m, 5H, j+x), 7.19-7.13 (m, 3H, w+y), 6.72 (d, J = 8.4 Hz, 1H, z), 4.09-4.05 (m, 1H, d), 4.01-3.99 (m, 3H, b+t), 3.78-3.69 (m, 3H, m+p), 3.65 (dd, J = 16.8 and 6.0 Hz, 1H, p'), 2.73 (dd, J = 13.2 and 5.4 Hz, 1H, u'), 2.64 (d, J = 13.2 and 7.8 Hz, 1H, u'), 2.55 (dd, J = 12.6 and 8.4 Hz, 1H, e), 2.36 (dd, J = 12.6 and 5.4 Hz, 1H, e'), 2.32-2.24 (m, 2H, s), 1.30 (s, 9H, ac) and 1.08 (t, J = 7.2 Hz, 3H, a); ¹³C-NMR (151 MHz, DMSO)  δ 170.59 (r), 169.97 (c), 169.16 (o), 168.78 (l), 154.83 (aa),
Chapter 2

144.07 (g), 138.95 (v), 129.23 (w), 129.07 (h), 128.14 (i), 128.02 (x), 126.89 (j), 125.91 (y), 77.49 (ab), 66.39 (f), 60.81 (b), 51.58 (d), 49.29 (t), 42.07 (p), 41.38 (m), 40.28 (s), 32.82 (e), 28.23 (ac) and 13.94 (a); The $^{13}$C signal for (u) is obscured by DMSO at ca. δ 39.4 ppm; LRMS (ES$^+$) m/z (rel. intensity) 789 (100%, [M+Na]$^+$); HRMS m/z [M+H]$^+$ calculated for [C$_{43}$H$_{60}$N$_4$O$_7$SNa]$^+$: 789.3292, found: 789.3264.

2.29: Boc-β-Leu-β-Phe-Gly-Gly-Cys(Trt)-OEt

Prepared from 2.28 (0.314 g, 0.447 mmol) using General Procedures 2 and 3. Purification by flash chromatography (SiO$_2$, 5% MeOH/CHCl$_3$) yielded the title compound as a pale yellow solid (0.377 g, 0.421 mmol, 94% over 2 steps).

m.p. 142–146 °C; $^1$H-NMR (600 MHz, DMSO-d$_6$) δ 8.35 (d, J = 7.8 Hz, 1H, NH, k), 8.15 (t, J = 5.4 Hz, 1H, q), 8.07 (t, J = 5.4 Hz, 1H, n), 7.75 (d, J = 8.4 Hz, 1H, z), 7.36-7.31 (m, 6H, i), 7.29 (d, J = 7.1 Hz, 6H, h), 7.27 – 7.23 (m, 5H, j+x), 7.19 – 7.14 (m, 3H, w+y), 6.52 (d, J = 9.0 Hz, 1H, ag), 4.32-4.24 (m, 1H, t), 4.09-4.06 (m, 1H, d), 4.01-3.95 (m, 2H, b), 3.77-3.69 (m, 4H, m+p+ac), 3.66 (dd, J = 12.0 and 8.1 Hz, 1H, p'), 2.76 (dd, J = 13.8 and 8.3 Hz, 1H, u), 2.64 (dd, J = 13.8 and 8.4 Hz, 1H, u'), 2.55 (dd, J = 12.8 and 8.4 Hz, 1H, e), 2.37 (dd, J = 12.5, 5.3 Hz, 1H, e'), 2.33-2.23 (d, J = 6.9 Hz, 2H, s), 2.11 (dd, J = 13.8 and 5.4 Hz, 1H, ab), 2.00 (dd, J = 13.8 and 8.4 Hz, 1H, ab'), 1.51-1.44 (m, 1H, ae), 1.34 (s, 9H, aj), 1.17-1.10 (m, 1H, ad), 1.08 (t, J = 7.2 Hz, 3H, a), 0.88-0.82 (m, 1H, ad'), 0.76 (d, J = 6.6 Hz, 3H, af) and 0.74 (d, J = 6.6 Hz, 3H, af'); $^{13}$C-NMR (151 MHz, DMSO) δ 170.38 (r), 169.97 (c), 169.36 (aa), 169.21 (l), 168.81 (o), 154.91 (ah), 144.06 (g), 138.87 (v), 129.12 (w), 129.07 (h), 128.14 (l), 128.02 (x), 126.88 (j), 125.95 (y), 77.30 (ai), 66.39 (f), 60.81 (b), 51.59 (d), 47.44 (t), 45.76 (ac), 42.91 (ad), 42.09 (p), 42.04 (p'), 41.41 (m), 40.36 (s), 32.81 (e), 28.27 (aj), 24.24 (ae), 23.31 (af), 21.63 (af') and 13.94 (a); The $^{13}$C signal for (u) is obscured by DMSO at ca. δ 39.4 ppm; LRMS (ES$^+$) m/z (rel. intensity) 916 (100%, [M+Na]$^+$); HRMS m/z [M+H]$^+$ calculated for [C$_{50}$H$_{60}$N$_4$O$_7$SH]$^+$: 894.4470, found: 894.4438.
2.30: Boc-β³Ala-β⁴Leu-β⁵Phe-Gly-Gly-Cys(Trt)-OEt

Prepared from 2.29 (0.377 g, 0.435 mmol) using General Procedures 2 and 3. Purification by flash chromatography (SiO₂, 5–6% MeOH/CHCl₃) yielded the title compound as a white solid (0.310 g, 0.321 mmol, 74% over 2 steps).

**m.p.** decomposes at >180 °C; ¹H NMR (600 MHz, DMSO-d₆) δ 8.36 (d, J = 7.7 Hz, 1H, k), 8.17 (t, J = 5.7 Hz, 2H, q), 8.07 (t, J = 5.6 Hz, 1H, n), 7.76 (d, J = 8.4 Hz, 1H, z), 7.60 (d, J = 8.7 Hz, 1H, ag), 7.34 (t, J = 7.6 Hz, 6H, i), 7.29 (m, 7H, h+ai), 7.27 – 7.21 (m, 5H, j+x), 7.19 – 7.14 (m, 3H, w+y), 4.28 (dq, J = 15.0, 7.6 Hz, 1H, t), 4.11 – 3.93 (m, 5H, b+d+ac+aj), 3.77 – 3.70 (m, 3H, m+p), 3.66 (dd, J = 16.6, 5.6 Hz, 1H, p'), 2.77 (dd, J = 13.7, 4.9 Hz, 1H, u), 2.64 (dd, J = 13.7, 8.6 Hz, 1H, u'), 2.55 (dd, J = 12.4, 8.5 Hz, 1H, e), 2.37 (dd, J = 12.5, 5.3 Hz, 1H, e'), 2.33 – 2.25 (m, 2H, s), 2.21 (dd, J = 13.8, 5.9 Hz, 1H, ab/ai), 2.17 – 2.09 (m, 2H, (ab/ab'+ai'/ai')), 2.00 (dd, J = 13.8, 8.5 Hz, 1H, ab'/ai'), 1.50 – 1.42 (m, 1H, ae), 1.16 (ddd, J = 14.1, 10.7, 4.2 Hz, 1H, ad), 1.08 (t, J = 7.1 Hz, 3H, a), 1.05 (s, 9H, ao), 1.00 (d, J = 6.6 Hz, 3H, ak), 0.88 (ddd, J = 13.4, 9.9, 3.7 Hz, 1H, ad'), 0.77 (d, J = 6.6 Hz, 3H, af), 0.71 (d, J = 6.5 Hz, 3H, af'); ¹³C NMR (151 MHz, DMSO) δ 176.35 (am), 170.41 (r), 169.98 (c), 169.57 (ah), 169.22 (aa), 169.17 (o), 168.82 (l), 144.07 (g), 138.89 (v), 129.14 (w), 129.08 (h), 128.15 (i), 128.03 (x), 126.89 (j), 125.97 (y), 66.40 (f), 60.81 (b), 51.59 (d), 47.45 (t), 44.12 (ac/aj), 42.50 (ad), 42.42 (ac/aj), 42.04 (ab/ai), 41.95 (ab/ai), 41.75 (p), 41.40 (m), 40.32 (s), 37.89 (an), 32.82 (e), 27.34 (ao), 24.16 (ae), 23.41 (af), 21.41 (af'), 19.93 (ak), 13.94 (a); The ¹³C signal for (u) is obscured by DMSO at ca. δ 39.4 ppm; LRMS (ES⁺) m/z (rel. intensity) 985 (100%, [M+Na⁺]); HRMS m/z [M+Na⁺] calculated for [C₅₄H₇₀N₆O₈SNa]: 985.4868, found: 985.4824.

2.31: Trityl-Protected Operation Product Standard

Hydrazine hydrate (64% w/v in H₂O, 29 μL, 4.5 eq.) was added to a solution of 2.30 (0.115 g, 0.119 μmol, 1.0 eq.) in ethanol (2 mL, 50 mM) in a sealed microwave vial and heated at 90 °C for 48 hours. The reaction mixture was cooled to room temperature, upon which a precipitate
formed. The solvent was removed under reduced pressure and the solid azeotroped three times with toluene. The resulting orange solid was used directly without further purification.

To crude hydrazide (assumed 0.119 mmol, 1.1 eq.) in CH₂Cl₂ (0.5 mL) and DMF (0.5 mL) was added Aldehyde Macrocycle 2.3 (0.061 g, 0.180 mmol, 1.0 eq.) and aniline (3 drops, cat.). The resulting cloudy suspension was stirred overnight at room temperature. The solvent was removed under reduced pressure and the crude residue purified by flash column chromatography (SiO₂, (0.5% NH₃, 4–5% MeOH)/CH₂Cl₂) to give the title compound as glass (0.085 g, 56.9 μmol, 48% over 2 steps). Obtained as a 1:1 mixture of isomers which complicated the analysis of the NMR spectra; full ¹H NMR assignment is given where isomeric positions are marked (’) where appropriate, however due to the complexity of the ¹³C NMR spectrum, observed peaks are reported but are unassigned.

¹H NMR (600 MHz, DMSO-d₆) δ 11.45 (s, 1H, b), 8.40 (d, J = 8.0 Hz, 0.5H, k), 8.32 (d, J = 8.3 Hz, 0.5H, k’), 8.21 – 8.11 (m, 1H+2(0.5H), q+a+n), 8.06 (t, J = 5.7 Hz, 0.5H, n’), 7.91 (s, 0.5H, a’), 7.76 (d, J = 8.3 Hz, 1H, z), 7.63 – 7.55 (m, 2H, ag+ma), 7.37 – 7.10 (m, 23H, h+i+j+w+x+y+mp+mq), 7.07 – 7.04 (m, 2H, mb+mb’), 7.01 (d, J = 8.5 Hz, 2H, mh), 6.99 (d, J = 8.5 Hz, 2H, mh’), 6.73 (d, J = 8.4 Hz, 2H, mi), 6.67 (d, J = 8.4 Hz, 2H, mi’), 5.37 (td, J = 8.3, 4.2 Hz, 0.5H, d), 4.45 (q, J = 7.4 Hz, 0.5H, d’), 4.27 (dq, J = 13.8, 6.9 Hz, 1H, t), 4.07 – 3.99 (m, 2H, ac+aj), 3.89 (t, J = 5.8 Hz, 2H, mk), 3.86 – 3.81 (m, 2H, mk’), 3.81 – 3.63 (m, 4H, m+p), 2.76 (dt, J = 12.9, 6.3 Hz, 2H, u), 2.71 – 2.59 (m, 9H, u’+md+mn), 2.49 – 2.34 (m, 6H, e+e’+mf), 2.34 – 2.25 (m, 2H, s), 2.21 (dd, J = 13.8, 5.8 Hz, 1H, ab/ai), 2.17 – 2.06 (m, 2H, (ab/ab’)+(ai/ai’)), 2.00 (ddd, J = 13.3, 8.5, 4.4 Hz, 1H, ab’/ai’), 1.92 (p, J = 7.5 Hz, 4H, me), 1.75 – 1.60 (m, 8H, ml+mn), 1.49 – 1.41 (m, 1H, ae), 1.19 – 1.12 (m, 1H, ad), 1.04 (s, 9H, ao), 0.99 (d, J = 6.5 Hz, 3H, ak), 0.93 – 0.83 (m, 1H, ad’), 0.78 – 0.73 (t, J = 6.0 Hz, 3H, af), 0.71 (t, J = 6.0 Hz, 3H, af’); ¹³C-NMR (151 MHz, DMSO) δ 176.35, 170.97, 170.44, 170.33, 169.61, 169.58, 169.23, 169.18, 168.72, 168.70, 165.97, 160.55, 156.76, 156.74, 148.15, 144.22, 144.15, 144.09, 142.59, 142.56, 138.90, 138.84, 138.04, 139.09, 133.99, 133.94, 133.91, 130.38, 129.16, 129.11, 129.08, 128.97, 128.15, 128.03, 127.94, 126.86, 126.67, 125.96, 124.95, 124.73, 120.18, 114.11, 67.03, 66.04, 65.61, 50.95, 48.47, 47.49, 47.45, 44.14, 42.52, 42.50, 42.42, 42.07, 42.01, 41.92, 41.81, 41.75, 41.59, 37.88, 36.56, 36.54, 34.32, 34.25, 33.69, 33.67, 31.39, 27.89, 27.84, 27.66, 27.64, 27.34, 24.16, 23.39, 21.44, 21.41 and 19.94; LRMS (ES⁺) m/z (rel. intensity) 758 (100%, [M+H+Na]²⁺) and 1493 (10%, [M+H]+). HRMS m/z [M+Na]⁺ calculated for [C₉₀H₁₀₃N₉O₉SNa]⁺: 1514.7961, found: 1514.7929.
2.11a/b (and 2.12): Deprotected Operation Product Standard and Operated Standard

To Trityl-Protected Operation Product Standard 2.31 (20 mg, 13.4 μmol, 1 eq.) in CH₂Cl₂ (0.216 mL) and TFA (0.054 mL) was added trisopropylsilane (13.7 μL, 67 μmol, 5 eq.). The resulting yellow reaction mixture turned colourless and was stirred at room temperature for 2 hours. Toluene (2 mL) was added and the resulting solution was concentrated to dryness under reduced pressure. The mixture was azeotroped twice with toluene (5 mL) to remove residual TFA and CH₂Cl₂. Trituration with a Et₂O:pet. ether (1:1) gave the title species as a white powder (assumed quantitative) which was analysed by LRMS (see Section 2.4.2) and used without further purification.

PPh₃ (11 mg, 67 μmol, 5 eq.) was added to the deprotected standard (assumed 13.4 μmol, 1 eq.) along with DMF (22 mL). NEt₃ (18.6 μL, 134 μmol, 10 eq.) was added and the resulting solution was heated at 60 °C for 6 days then cooled to room temperature and analysed by LRMS. The reaction mixture was concentrated in vacuo and triturated with Et₂O (3 x 5 mL) then n-pentane

Ⅹ Total reaction volume 0.27 mL, 0.050 M
(5 mL) to remove phosphorus species. The solid residue was purified by prep. TLC (SiO$_2$ (2 x Merck 250), (0.5% NH$_3$ (35% aq.) + 5% MeOH)/CH$_2$Cl$_2$, 2 elutions) to give the title species as a colourless glass (4.4 mg, 26% based on 2.11b).

LRMS (ES') m/z (rel. intensity) 899 (100%, [2.11b+2H]$^{2+}$); 910 (41%, [2.11b+H+Na]$^{2+}$); 1275 (7%, [2.11a+Na]$^+$. For spectra, see Figure 8 and Figure 9 (Section 2.4.2). For MS-MS analysis, see Figure 10 and Figure 11 (Section 2.4.2).

The product is isolated as a complex mixture of isomers which convolute characterisation by NMR spectroscopy. An example $^1$H NMR spectrum in deuterated DMF is shown in Figure 6.

![Figure 6: $^1$H NMR (d$_7$-DMF, 298 K) spectrum of the complex mixture obtained on removing the trityl protection from cysteine-protected authentic standard 2.31.](image)

### 2.4.1.11 Synthesis of Model Mixed αβ$_3$-Peptide 2.14 with Product of Operation Sequence

![Scheme 14: Synthesis of model αβ$_3$-peptide with product of operation sequence: Reagents and conditions: (i) TIPS, TFA, CH$_3$Cl$_2$, r.t., 1 h; (ii) AcCl, NET$_3$, CH$_3$Cl$_2$, DMF, r.t., 2 h, 92% over 2 steps.](image)
2.14: Model Mixed αβ³-Peptide

To 2.30 (0.142 g, 0.152 mmol, 1 eq.) in CH₂Cl₂ (3 mL) was added triisopropylsilane (61 μL, 0.298 mmol, 2 eq.) and TFA (0.60 mL). The resulting orange reaction mixture was stirred at room temperature for 1 hour, after which the mixture had turned colourless. Toluene (5 mL) was added and the solvent was removed under reduced pressure and the crude residue was azeotroped twice with toluene (2 × 5 mL) to remove residual TFA. Trituration with a mixture of diethyl ether and pet. ether (1:1, 3 × 5 mL) gave the crude deprotected thiol as a white film (0.106 mg, 0.153 mmol, quant.) which was used without further purification.

To the crude deprotection product (0.050 g, 0.072 mmol) in CH₂Cl₂ (1 mL) and DMF (1 mL) was added NEt₃ (12 μL, 0.0864 mmol, 1.2 eq.) and acetyl chloride (6 μL, 0.0864 mmol, 1.2 eq.). The reaction was stirred at room temperature for 2 hours. The solvent was removed under reduced pressure. The crude residue was redissolved in a mixture of CHCl₃ and IPA (20 mL, 10:1), washed with water (10 mL), dried (Na₂SO₄) and the solvent removed under reduced pressure. The resulting white solid (0.050 g, 0.066 mmol, 92%) was used without further purification.

m.p. 180–183 °C; ¹H NMR (600 MHz, DMSO-d₆) δ 8.34 (d, J = 7.9 Hz, 1H, h), 8.19 (t, J = 5.7 Hz, 1H, n), 8.11 (t, J = 5.8 Hz, 1H, k), 7.77 (d, J = 8.4 Hz, 1H, w), 7.60 (d, J = 8.7 Hz, 1H, ad), 7.29 (d, J = 7.9 Hz, 1H, ai), 7.27 – 7.20 (m, 2H, u), 7.20 – 7.15 (m, 3H, t+v), 4.39 (td, J = 7.9, 5.6 Hz, 1H, d), 4.28 (h, J = 7.7 Hz, 1H, q), 4.13 – 3.98 (m, 4H, b+z+ag), 3.78 – 3.70 (m, 3H, j+m), 3.67 (dd, J = 16.6, 5.7 Hz, 1H, m'), 3.31 (dd, J = 13.8, 5.5 Hz, 1H, e), 3.09 (dd, J = 13.8, 8.1 Hz, 1H, e'), 2.78 (dd, J = 13.7, 4.9 Hz, 1H, r), 2.65 (dd, J = 13.7, 8.6 Hz, 1H, r'), 2.35 – 2.25 (m, 5H, g+p), 2.21 (dd, J = 13.8, 5.9 Hz, 1H, y/af), 2.16 – 2.09 (m, 2H, y'/af'/y/af), 2.00 (dd, J = 13.8, 8.5 Hz, 1H, y'/af'), 1.50 – 1.40 (m, 1H, ab), 1.21 – 1.13 (m, 3H, a+aa), 1.05 (s, 9H, al), 1.00 (d, J = 6.6 Hz, 3H, ah), 0.88 (ddt, J = 15.9, 9.8, 4.9 Hz, 1H, a+a'), 0.77 (d, J = 6.6 Hz, 3H, ac), 0.72 (d, J = 6.5 Hz, 3H, ac'); ¹³C NMR (151 MHz, DMSO) δ 194.70 (f), 176.35 (aj), 170.48 (o), 169.91 (c), 169.56 (ae), 169.29 (l), 169.19 (x), 169.01 (i), 138.88 (s), 129.14 (t), 128.03 (u), 125.98 (v), 61.03 (b), 51.69 (d), 47.46 (q), 45.71 (q), 44.12 (z/ag), 42.49 (aa), 42.42 (z/ag), 42.10 (m), 41.95 (y/af), 41.75 (y/af), 41.54 (j), 40.35 (p), 37.88 (ak), 30.44 (g), 29.69 (e), 27.34 (al), 24.16 (ab), 23.41 (ac), 21.41 (ac'), 19.93 (ah), 14.00 (a); The ¹³C signal for (r) is obscured by DMSO at ca. δ 39.3 ppm; LRMS (ES⁺) m/z (rel. intensity) 786 (100%, [M+Na⁺]); 801 (48%, [M+K⁺]); 1563 (31%, [2M+K⁺⁺]); HRMS m/z [M+H]+ calculated for [C₁₃H₁₉N₅O₃SH]: 763.4059, found: 763.4054.
2.4.1.12 Proteolysis of 2.14 with Subtilisin Carlsberg

\(\alpha^\beta^3\)-Peptide 2.14 (2 mg, 2.62 \(\mu\)mol, 1 eq.) and subtilisin Carlsberg (1 mg lyophilised powder) were suspended in aqueous phosphate buffer (pH 8, 1 mL) and were stirred at 50 °C for 5 days. The reaction mixture was analysed crude by LRMS only. 0.1 mL aliquots of the crude reaction mixture were prepared for analysis by diluting with MeCN (1 mL) to precipitate insoluble materials and then subsequently filtered (0.45 µm pore size PTFE membrane). The filtrate was then analysed by LRMS (ES)  \(m/z\) (rel. intensity) 531 (11%, [2.17-H]); 588 (100%, [2.16-H]) and 691 (18%, [2.15-H]): spectrum shown in Figure 7.

![Figure 7: LR(ESI)-(-)-MS analysis of the selective proteolytic digestion of the \(\alpha\)-peptide region of model mixed \(\alpha^\beta^3\)-peptide 2.14 by subtilisin Carlsberg resulting in the liberation of the \(\beta^3\)-peptide region from cysteine, the NCL catalyst used to assemble the \(\beta^3\)-peptide during machine operation. No breakdown of the \(\beta\)-peptide region was observed.](image-url)
2.4.2 Mass Spectral Analysis of Product of Operation 2.11a/b

Operation product 2.11a spontaneously converts in solution to 2.11b (Scheme 15), which is in line with previously reported results,[22,23] leading to convolution of NMR spectra. For this reason, we have used mass spectrometry techniques to characterise the product of operation.

![Scheme 15: Interconversion of hydrazone 2.11a and S,N-acetal species 2.11b under presumed dynamic equilibrium.](image)

The conversion of 2.11a to 2.11b is observed during analysis of a reference sample (2.11a/b) synthesised independently by conventional peptide synthesis. Upon generation of 2.11a via trityl removal from its thiol-protected precursor 2.31, LRESI-(+)-MS analysis shows mass-ions correlating to both 2.11a and 2.11b as adducts with combinations of proton, sodium and potassium counterions (Figure 8).
Figure 8: LRESI-(+)-MS spectrum of ‘non-operated’ reference sample 2.11a/b following trityl protecting group removal facilitated by TIPS and TFA. Sample is infused directly into the spectrometer as a solution in methanol with formic acid (to aid ionisation). \textit{m}/\textit{z} labels correspond to the most abundant peak in each isotopic distribution.
Figure 9: LRESI-(+)-MS spectrum of ‘operated’ reference sample 2.11a/b following exposure to operation conditions and purification by preparative TLC. Sample is infused directly into the spectrometer as a solution in methanol with formic acid (to aid ionisation). m/z labels correspond to the most abundant peak in each isotopic distribution.

After submitting the reference sample 2.11a/b to operation conditions and subsequent purification by preparative TLC, equivalent to the operation and purification procedure experienced by deprotected molecular machine 2.10, mass spectral analysis again indicated the interconversion of 2.11a and 2.11b (Figure 9).

Due to the particularly high abundance of the base peak (m/z 899.1, corresponding to [2.11b+2H]^2+, Figure 9) in the ‘operated’ standard, this mass-ion was selected for fragmentation analysis by tandem mass spectrometry in both the ‘operated’ standard and the actual product of operation.

Tandem mass spectrometry (MS-MS) was used to determine the sequence-specificity of the product of operation of molecular machine 2.1 compared to classically prepared reference standard 2.11b. The tandem mass spectrometry and associated fragmentations of operation product and authentic standard are compared in Figure 10.

The sequence-specific translation of information from molecular machine 2.10 (deprotected 2.1) into product of operation 2.11b can be seen by the progressive amide fragmentation of the product of operation ((iii), Figure 10). No mass-ions correlating to an incorrect sequence of the three β3 amino acids were identified within the fragmentation indicating that deprotected molecular machine 2.10 is unable to operate out-of-sequence. This conclusion is further corroborated by the more intricate retro-Michael and retro-Mannich fragmentation modes that β3-peptides undergo described by Seebach.[42]

The molecular species responsible for each described fragmentation are outlined in Figure 11.
Figure 10: Comparative MS-MS fragmentation analysis of authentic reference standard and isolated product of operation. (i) Observed amide fragmentation modes of α- and β-peptide bonds (Yn) and additional retro-Michael (Yn*) and retro-Mannich (Yn*) fragmentation modes observed in beta peptides; (ii) Observed fragmentations of reference standard annotated with the responsible fragmentation mode; (iii) Observed fragmentations of isolated product of operation of deprotected molecular machine 2.10 indicating the progressive loss of amino acid residues according precisely to the sequence information transcribed into molecular machine 2.1.
Figure 11: Potential and observed molecular fragments from MS-MS analysis of authentic reference standard and isolated product of operation. All fragments are visible as either singly or double charged mass-ions with one or two proton counterions respectively.
2.5 References


Chapter 3

An Artificial Molecular Machine that Builds a Helical Oligomeric Asymmetric Catalyst

“It may be that the chemical ligation tool kit will find its ‘killer application’ in nanoscience, for the bottom-up fabrication of molecular arrays and protein-inspired molecular devices.”

– Stephen B. H. Kent

From Total Chemical Synthesis of Proteins

Chapter 3

Abstract

An artificial, rotaxane-based molecular machine that operates in a processive manner on a polymeric track to assemble a catalytically active oligopeptide is reported herein. The produced oligoleucine species spontaneously forms a secondary structure with α-helical character, allowing it to promote the enantioselective epoxidation of a chalcone derivative with excellent conversion and chiral control. These results demonstrate that, in a manner reminiscent of the ribosome assembling the primary structure of an enzyme, it is possible to use artificial molecular machines for the non-trivial synthesis of functional molecules.

Acknowledgements: The research discussed in the following chapter represents the collaborative efforts of the author and Dr Guillaume De Bo (GDB), Dr Julien De Winter (JDW) and Dr Sonja Kuschel (SK) under the supervision of Prof. David A. Leigh. This work is presented as a manuscript prepared for peer-review publication by the author and the aforementioned researchers. The author performed the synthesis, analysis and characterisation of all compounds indicated in Section 3.5.1. Calculations based on polymeric species, distributions and ratios were performed GDB, who also performed all styrene polymerisations. JDW performed mass spectrometry on rotaxane-bound polymers.

Dr Jonathan J. Danon, Dr Thomas A. Singleton and Catherine J. Eagle are gratefully acknowledged for proof-reading this chapter.
3.1 Introduction

Molecular machines have been prevalent throughout Nature for billions of years and are employed to carry out a broad range of complex, life-essential processes. This biological precedent has inspired synthetic chemists to design artificial molecular machines capable of carrying out non-trivial tasks. Over the past two decades, synthetic molecular machines have been developed with increasingly complex functionality: examples include molecular robotic transporters, autonomous and chemically-fuelled molecular motors, molecular pumps, walkers and muscles. The level of dexterity achieved in synthetic molecular machines to date has enabled chemists to design biomimetic devices capable of performing non-trivial synthetic tasks; the most sophisticated of these small-molecule synthesisers can operate autonomously to facilitate programmed, multi-step syntheses.

Other synthetic systems have mimicked the ability of biological machines to operate repeatedly upon oligomeric substrates in a processive manner, whereby the device binds to a substrate which is not then released until the iterative operation is complete. Examples include Nolte’s system capable of the processive epoxidation of an unsaturated polymer strand, and his successive oxidation of a DNA strand using a bio-hybrid catalyst, both of which utilise rotaxane or pseudo-rotaxane frameworks to maintain processivity. The concatenation of such processive synthetic devices with rotaxane-architectures, composed of a macrocycle encircling a polymeric thread, could provide a platform for the autonomous synthesis of functional oligomeric species, such as catalytically active oligopeptides. A growing interest in the field of small, unnatural peptide-mediated catalysis is driven by the likeness of such catalysts to primitive stereospecific enzymes which has given birth to the hypothesis that they may have been prevalent in the early developments of prebiotic life.

Here we report a rotaxane-based molecular machine that operates autonomously and processively upon a polymeric track to assemble a homo-oligopeptide by joining together amino acid building blocks distributed along the polymer. Furthermore, the assembled oligopeptide can adopt an α-helical secondary structure, rendering it a potent asymmetric catalyst for the Juliá-Colonna epoxidation of a chalcone derivative. The operation of the artificial molecular machine can be likened to the ribosome’s ability to assemble oligopeptides according to the template provided by messenger RNA; upon folding into a specific, higher-order structure, these peptides go on to form functional biological machines such as enzymes.
3.2 Results and Discussion

3.2.1 System Design

In designing a rotaxane-based molecular machine for the synthesis of a catalytic homo-oligopeptide, we reasoned that the synthesis of consecutive rotaxane barriers could be achieved more efficiently than in our previous stepwise approaches\textsuperscript{[12,13]} as, in this case, only one type of amino acid building block would be required. Modern controlled radical polymerisation (CRP) techniques, such as atom transfer radical polymerisation (ATRP), provide facile access to well-defined functional polymers with controlled architectures.\textsuperscript{[26]} An approach where a building block-bearing monomer would be copolymerised with another ‘blank’ monomer to generate a random co-polymer track was investigated, where the amino acid-loaded building blocks would be distributed randomly along the length of the polymer backbone. We opted for a track with an atactic polystyrene (PS) spine, as PS is expected to be inert under machine operation conditions and to show reduced flexibility in solution relative to other common polyolefins. However, an atactic PS track would be significantly ‘thicker’ than the oligomeric track employed in our original systems, potentially restricting the macrocycle from shuttling freely along the track. CPK modelling suggested that the macrocycle could indeed shuttle along a PS track.

In terms of the amino acid building block, leucine proved to be an ideal candidate for several reasons. Firstly, it is one of the amino acids most prone to forming α-helical secondary structures.\textsuperscript{[23]} Secondly, its α-helical oligomers are amongst the most active and selective catalysts in the Juliá-Colonna epoxidation.\textsuperscript{[27]} Thirdly, in our experience, leucine is large enough to act as a barrier to macrocycle shuttling and it has reacted reliably in previous machine operations.\textsuperscript{[12,13]} In addition, leucine has characteristic ^1H NMR signals which are distinct from the PS backbone fingerprint which can simplify machine operation monitoring and product analysis.

The comprehensive mechanistic insights into the Juliá-Colonna epoxidation reported by Roberts\textsuperscript{[28,29]} suggest that the 4 to 5 N-terminal residues in an α-helical leucine oligomer are responsible for asymmetric catalysis. These residues help to stabilise the α-helix secondary structure at the N-terminus through hydrogen bonding, thereby forming a catalytic pocket. Specifically, Roberts highlights the importance of the three terminal amide N–H groups (which are not involved in intra-residue hydrogen bonding) in the formation of an ‘oxy-anion hole’, the catalyst’s active site. It thus follows that short oligo-leucine chains containing at least 4 to 5 N-terminal leucine residues could adopt an α-helical secondary structure able to catalyse the Juliá-Colonna epoxidation of chalcones. With this in mind, we set out to synthesise molecular machine
3.1 with 6 leucine residues on average per machine (where one leucine comes from the rotaxane component 3.2 and the remainder are from the copolymer track 3.3, Figure 1a).

3.2.2 Molecular Machine Synthesis

Machine 3.1 was assembled in a modular fashion using our recently described rotaxane capping strategy\textsuperscript{[13]} where one-barrier rotaxane 3.2 was extended with leucine-bearing polymeric track 3.3 (Figure 1a). The synthesis of alkyne-terminated polymeric track 3.3 began with the copolymerisation of styrene (3.5) and \( \rho \)-leucyloxystyrene (3.6) by ATRP (Figure 1b). Because of

![Figure 1: Assembly of rotaxane-based molecular machine 3.1 by elongation of rotaxane 3.2 with copolymer 3.3. a) Synthesis of 3.1: i) CuPF\(_6\)·(CH\(_3\)CN)\(_4\), CH\(_2\)Cl\(_2\)::BuOH (4:1), r.t., 36 h, 96%. b) Synthesis of copolymer 3.3 by ATRP: ii) CuBr (1 eq.), 4,4′-dinyonyl-2,2′-dipyridyl (2.1 eq.), anisole (20 wt%), 90 °C, 24 h, 22% conversion; iii) Bu\(_3\)SnH (3 eq.), toluene, 85 °C, 1 h, 80%. c) SEC traces of copolymer 3.3 in dark blue and machine 3.1 in light blue (THF, 1 mL/min, refractive index detector, normalised intensities are shown). d) +MALDI-TOF-MS of machine 3.1 ([M+Na]\(^+\)) indicating overlapping distributions of machines with increasing numbers of leucine units (n). e) Gaussian fit of the main peak of the GPC trace of machine 3.1 (light blue trace) represented in terms of molecular weight where green bars indicate the relative intensity of chains bearing n leucine units.](image)
their similarity in structural and electronic properties, the assumption was made that the relative copolymerisation reactivity of styrene and \( p \)-leucyloxy styrene would be similar to the relative reactivity of the styrene/\( p \)-acetoxystyrene couple. Since styrene and \( p \)-acetoxystyrene have reactivity ratios close to 1.00 (\( r_1 = 0.89 \) and \( r_2 = 1.22 \)),\(^{[30]}\) it was expected that the styrene/\( p \)-leucyloxy styrene combination would lead to the desired random copolymer, incorporating sufficiently well-dispersed leucine-bearing monomers into the PS backbone. A monomer ratio of 11:1 (styrene:\( p \)-leucyloxy styrene) was targeted, representing an average distance of ca. 22 Å (linear, unfolded distance)\(^{[31]}\) between two leucine-bearing monomers. This value is comparable to the analogous distance (ca. 20 Å) between substrates in our previous oligopeptide-forming molecular machines, which proved to be a good balance between the inter-building blocks distance and the overall length of the track. ATRP polymerisation of monomers \( 3.5 \) (240 eq.) and \( 3.6 \) (20 eq.) was carried out from TMS-alkyne-terminated initiator \( 3.4 \) to afford polymer \( 3.7 \) in a process that concomitantly cleaved the silyl group. The composition of copolymer \( 3.7 \) was confirmed to be in close agreement with the polymerisation feeding ratio of co-monomers \( 3.5 \) and \( 3.6 \) by \(^{1}H\) NMR spectroscopy. So as to prevent any undesired reaction of the vestigial benzylic bromide during operation of the machine, polymer \( 3.7 \) was subjected to radical debromination mediated by \( \text{Bu}_3\text{SnH} \) to form polymer extension piece \( 3.3 \) (Figure 1b).

Finally, polymer track \( 3.3 \) was appended to [2]rotaxane \( 3.2 \) via a copper(I) catalysed azide-alkyne cycloaddition (CuAAC) reaction in presence of \( \text{CuPF}_6\cdot(\text{CH}_3\text{CN})_4 \) to afford the elongated rotaxane \( 3.1 \) in 96\% yield (Figure 1a).

### 3.2.3 Compositional Analysis

Comparison between the compositions of the machine and the product resulting from its operation should provide an insight into the transfer of information from the track to the product. Due to the polydispersity of the PS track, the machine exists as a collection of chains of different lengths, hence containing different numbers of leucine barriers. The average leucine content can be inferred directly from the \(^{1}H\) NMR spectrum (6 units/machine on average), but a more detailed picture of the chain compositions can be gained from the MALDI spectrum and the SEC trace (Figure 1c–e). The MALDI profile, which follows a Gaussian distribution with a median value around 9.6 kDa, reveals the presence of various species highlighted in Figure 1d, the most abundant of which contain 6–7 units. Evaluation of their relative abundances is complicated by the overlap of populations separated by 5 \( p \)-leucyloxy styrene units (see Section 3.5.1.4 for details). A Gaussian fit of the main peak of the SEC trace of machine \( 3.1 \) (Figure 1c) is represented in terms of molecular weight in Figure 1e. Integrating the area below the curve around masses
corresponding to chains containing \( n \) leucine units reveals the relative abundance of each population. Molecular machines bearing 3 to 13 leucine residues can be inferred with the most abundant populations carrying 6–7 units in line with the MALDI data.

This extrapolated information, largely concordant between SEC and MALDI data analysis, gives a detailed overall picture of the information that has been transcribed into the molecular machine in the form of a distribution of incorporated leucine residues across machines of different mass, stemming from the inherent polydispersity of the ATRP copolymerisation process.

### 3.2.4 Operation and Analysis of Information Translation

Confident in our understanding of the information contained within molecular machine 3.1, we turned towards operation of the device to test its competency in translating this information into an oligoleucine product. Fully-assembled machine 3.1 was deprotected under acidic conditions in the presence of TIPS to reveal the amines and thiol necessary for operation via native chemical ligation (NCL).\textsuperscript{[32]} Under the basic operation conditions, the freshly-revealed cysteine thiolate on the macrocycle-bound catalytic unit abstracts a leucine from the track via an [O-S] acyl-transfer reaction. The acyl group of the transient thioester is then quickly captured by the terminal amine on the tripeptide catalytic unit to form a new amide bond and to regenerate the catalytic thiolate. The macrocycle can then shuttle beyond the unloaded barrier residue to perform NCL on distal leucine barriers, further along the polymer tail piece. Deprotected machine 3.8 was operated at 65 °C in DMF-\( d_7 \) for 4 days in presence of triethylamine and triphenylphosphine (to reverse any disulfide formation) to afford oligoleucine operation product 3.9 (Figure 2a).

The +ESI MS profile of the operation product revealed the presence of narrowly distributed leucine oligomers with between 3 and 12 units incorporated: the distribution was centred around 6 leucine units (Figure 2c). Critically, these values are in good agreement with what was anticipated from the composition of machine 3.1 (vide supra). Not only does the obtained oligomer distribution fall within the extrapolated boundaries, but also the relative abundance of each member of the distribution (Figure 2c) closely resembles the relative abundance of the corresponding machines at the peak of the distribution of 3.1 (Figure 2b). Finally, the \( M_w/M_n \)
value observed for 3.9 (1.01) is in agreement with the theoretical value derived from the SEC trace of track 3.3 (1.01, calculated by Dr Guillaume De Bo). Taken together, these data strongly suggest that machine 3.1 has undergone operation in a predominantly *intramachina* fashion, successfully translating information, in the form of polydispersity, from the track to the product.

### 3.2.5 Oligoleucine-Catalysed Juliá-Colonna Epoxidation

For 3.9 to be an effective catalyst in the Juliá-Colonna epoxidation of chalcones, it must be able to fold into an α-helix of the type indicated by the energy-minimised structure of the product shown in Figure 3b. To probe the catalytic capabilities of 3.9, operation product-type oligoleucine 3.11 was synthesised with 6 leucine residues on average per oligomer (Figure 3a, n = 6, see Section 3.5.1.5 for details). The presence of a maximum around 190 nm, and two minima at around 205 and 220 nm respectively in the circular dichroism (CD) spectrum of 3.11, is characteristic of α-helicity (Figure 3d). Although the overall shape of the CD spectrum suggests that this α-helical conformation exists alongside a random coil arrangement (which is common for short peptides), it appeared that 3.11 might display sufficient helicity to catalyse the Juliá-Colonna epoxidation.
Leucine oligomer 3.11 was successfully employed in the homogeneous Juliá-Colonna epoxidation of furyl chalcone 3.13 to furnish epoxide 3.14 with 93% conversion (Figure 3c). In addition to the excellent conversion, operation product-type catalyst 3.11 was able to affect the chiral outcome of the epoxidation, achieving a 63:37 e.r. of 3.14.

Some of the best reported Juliá-Colonna epoxidations have achieved up to 99:1 e.r. using oligoleucines with as few as 10 units on average. Furthermore, Roberts reports exceptional chiral control using shorter oligomers still with an average of only 3.9 leucine residues terminating each end of a PEG linker. Using his system, with only enough residues on average to complete one full turn of an α-helix, Roberts achieved chalcone epoxidation with 80% conversion and 99:1 e.r. With evidence that we were forming machine operation products that could show catalytic activity in the Juliá-Colonna epoxidation, we sought to better our attained level of chiral control.

Following analysis of operation product 3.9, it became apparent that oxidation of the catalyst’s cysteine thiol to its corresponding sulfonic acid would likely occur under the oxidative conditions of the Juliá-Colonna epoxidation. Sulfonic acid formation could lead to disruption of the α-helix’s
hydrogen-bonding network under the basic epoxidation conditions. To circumvent any potential derogation of the catalyst’s secondary structure during the epoxidation, the cysteine residue of 3.11 was subjected to a reductive radical desulfurisation to form its corresponding alanine derivative 3.12 (Figure 3a).\textsuperscript{[36]} As anticipated, the CD spectrum of the post operationally-modified oligoleucine 3.12 also suggested a partial α-helical secondary structure (Figure 3d).

Post-operationally modified catalyst 3.12 was employed in the Juliá-Colonna epoxidation under the previously explored conditions to once again afford epoxide 3.14, only now with complete conversion (by \textsuperscript{1}H NMR) and excellent enantioselectivity (96:4 e.r. by HPLC) in 18 h at room temperature (Figure 3c). This result shows comparable asymmetric control to that achieved with similarly short leucine oligomers.\textsuperscript{[35]} The excellent conversion and degree of chiral control achieved by this small oligoleucine catalyst designed with an average of only 6 leucine residues at its N-terminus, as well as an achiral macrocycle bound to an Ala-Gly-Gly linker at its C-terminus, was beyond our best expectations.
3.3 Conclusion

We have designed and synthesised a small, rotaxane-based molecular machine which assembles oligopeptides by native chemical ligation as it operates autonomously along a polymeric strand bearing leucine building blocks. The information transcribed into the polymeric track during its synthesis is translated into the machine’s oligopeptide operation product. Furthermore, an oligomer analogous to the operation product is able to adopt an α-helical secondary structure which, after a chemical modification, renders it a potent asymmetric catalyst in the Juliá-Colonna epoxidation of chalcones with both excellent efficiency and enantioselectivity.

The described process can be likened to the ribosome’s ability to assemble peptide sequences according to the genetic information transcribed into messenger RNA from DNA. Moreover, a molecular machine operation-type product undergoes a post-translational modification before assuming a catalytically active secondary structure in a process akin to the modification and folding of ribosomal protein products which transforms them into biologically active molecules.

3.4 Future Work

In moving away from the rigid barrier architecture used in previous generations of rotaxane-based molecular machines,\textsuperscript{[12,13]} the PS rotaxane track has become both longer and more flexible. As the described device requires only one type of amino acid building block to construct a homo-oligopeptide, it is impossible to determine whether the molecular machine is operating in a sequence specific manner. Alternatively, the catalytic unit could ‘pick up’ amino acid barriers out of sequence by ‘over-reaching’. It is equally difficult to infer with confidence whether or not intermachina ligations are occurring during operation, although the successful translation of polydispersity from the copolymer track into the distribution of oligoleucine peptide products does suggest a predominantly intramachina operation. In order to probe these mechanistic intricacies, two experiments have been conceived which are summarised in Figure 4.
Figure 4: Control studies to test for: (top) out-of-sequence `pick up' of barriers in the rotaxane via `over-reaching' of the catalyst; (bottom) a one-pot operation of two different molecular machines, each loaded exclusively with a different amino acid barrier, to test for intermachina operation.

In the first control, the polymer tail has two discrete regions, each loaded with different amino acid building blocks (of different mass). Such a tail piece could be realised by joining two copolymers, separately prepared by ATRP using barrier units with different amino acid building blocks. Sequence-specific operation of this machine would give oligomeric products, each with two discrete peptide regions, whereas out-of-sequence operation would yield oligomers with an undefined mixture of the two amino acids. These contrasting products would be differentiable by tandem mass spectrometry (MS-MS).

In the second control, two separately-prepared machines, each exclusively loaded with different amino acids (of different mass), could be operated together in one pot. A purely intramachina operation would give two separate distributions of homo-oligopeptides, whereas a scrambled mixture of oligomers containing both amino acids would indicate an intramachina operation. Again, MS-MS could differentiate the two operation modes.

Additional project targets include employing the actual product of operation 3.9 in the Juliá-Colonna epoxidation and attempting to form the interlocked architecture of the machine in a direct rotaxane formation using the alkyne-terminated polymer unit along with macrocycle (3.18) and stopper 3.23. Initial results in both of these areas have been promising and they are currently under ongoing investigation.
3.5 Supporting Information

3.5.1 Synthetic Procedures

4-[(4-azidophenyl)propan-2-yl]aniline 3.25\textsuperscript{[12]} and azide stopper 3.23\textsuperscript{[37]} were prepared according to literature procedures. The author synthesised and analysed all listed compounds at least once with the following exceptions: compounds 3.1 and 3.3 were synthesised by the author but characterised jointly by the author, Dr Guillaume De Bo (GDB) and Dr Julien De Winter (JDW); compound 3.6 was synthesised and characterised by Dr Sonja Kuschel (SK); compound 3.7 was synthesised and analysed by GDB. SK and GDB carried out and analysed the Juliá-Colonna epoxidations of 3.13. The author carried out the deprotection, operation of machine 3.1 and the operation product (3.9) was analysed jointly by the author and GDB.

3.5.1.1 Synthesis of Macrocycle 3.18

Scheme 1: Amide macrocycle 3.18 synthesis from nitrile macrocycle 3.15. Reagents and conditions: i) LiAlH\textsubscript{4}, THF, 0 °C – r.t., 15 h, 99%; ii) Boc-Cys(Trt)-OH, EDCI·HCl, HOBt·H\textsubscript{2}O, DIPEA, DMF, r.t., 3 h, quant.; iii) TFA, Ph\textsubscript{3}COH, CH\textsubscript{2}Cl\textsubscript{2}, r.t., 2 h; iv) Boc-GlyGly-OH, EDCI-HCl, HOBt-H\textsubscript{2}O, THF, CHCl\textsubscript{3}, r.t., 14 h, quant..

3.16: Amine Macrocycle

To a solution of nitrile macrocycle 3.15 (400 mg, 716 µmol, 1 eq.) in THF (10 mL, 72 mM) cooled to 0 °C was added a solution of LiAlH\textsubscript{4} (1 M (THF)), 1.79 mL, 1.79 mmol, 2.5 eq.). The resulting solution was stirred and allowed to warm over 15 h. The reaction mixture was cooled to 0 °C and quenched cautiously with water (2.5 mL). Solvent was removed in vacuo and the residue was dissolved in CH\textsubscript{2}Cl\textsubscript{2} (20 mL) before being washed with aqueous potassium sodium tartrate (sat., 30 mL). The organics were dried over MgSO\textsubscript{4} and
solvent removed in vacuo to give the title compound as a white solid (399 mg, 99%) which was used in subsequent steps without further purification.

m.p. 107–110 °C. 1H NMR (600 MHz, Acetone-d6) δ 7.56 (t, J = 7.6 Hz, 1H, t), 7.07 (d, J = 8.3 Hz, 4H, m), 7.05 (d, J = 7.8 Hz, 2H, s), 7.02 (s, 3H, d+f), 6.75 (d, J = 8.3 Hz, 4H, l), 4.38 (s, 2H, b), 3.93 (t, J = 5.9 Hz, 4H, j), 2.76 (t, J = 7.3 Hz, 4H, q), 2.66 (tt, J = 7.2 Hz, 4H, g), 2.54 (t, J = 7.61 Hz, 4H, o), 2.05 – 1.99 (m, 4H, p), 1.81 (p, J = 7.1 Hz, 4H, h), 1.74 (p, J = 6.0 Hz, 4H, i). 13C NMR (151 MHz, Acetone) δ 162.0 (r), 158.3 (k), 142.9 (e), 142.0 (c), 137.3 (t), 135.2 (n), 130.1 (m), 127.1 (f), 126.5 (d), 120.9 (s), 115.0 (l), 68.2 (j), 55.6 (b), 38.0 (q), 35.8 (g), 35.1 (o), 32.6 (p), 29.3 (l), 28.9 (h). LRESI-(+)-MS: m/z 563.5 (100, [M+H]+). HRFT-(+)-MS: m/z 563.3618 [M+H]+; 563.3632 calc. for [C38H46N2O2H]⁺.

3.17: Mac-Cys(Trt)-Boc

Amine macrocycle 3.16 (196 mg, 348 µmol, 1 eq.), Boc-Cys(Trt)-OH (169 mg, 365 µmol, 1.05 eq.), EDCI-HCl (100 mg, 522 µmol, 1.5 eq.) and HOBt·H2O (78 mg, 522 µmol, 1.5 eq.) were dissolved in DMF (7 mL, 50 mM). DIPEA (182 µL, 135 mg, 1.04 mmol, 3 eq.) was added dropwise and the reaction mixture was stirred for 3 h before solvent was removed in vacuo. The residue was dissolved in ETOAc (20 mL) and washed successively with water (20 mL), aqueous NaHCO3 (sat., 20 mL), aqueous NH4Cl (sat., 2 × 20 mL), aqueous NaHCO3 (sat., 20 mL) and brine (20 mL). The organics were dried over MgSO4 and solvent was removed in vacuo to give the title compound as a white foam (350 mg, quant.) requiring no further purification.

m.p. 70–72 °C. [α]D20 = +13.1 (c = 1.00, CH2Cl2). 1H NMR (600 MHz, Acetone-d6) δ 7.58 (t, J = 7.6 Hz, 1H, a), 7.50 (t, J = 5.2 Hz, 1H, t), 7.43 (d, J = 7.9 Hz, 6H, z), 7.33 (t, J = 7.7 Hz, 6H, aa), 7.25 (t, J = 7.2 Hz, 3H, ab), 7.09 (d, J = 8.3 Hz, 4H, h), 7.07 – 7.04 (m, 3H, b+p), 6.96 (s, 2H, q), 6.76 (d, J = 8.3 Hz, 4H, i), 6.06 (d, J = 8.0 Hz, 1H, ac), 4.40 (dd, J = 14.9, 5.7 Hz, 1H, s), 4.31 (dd, J = 15.0, 5.7 Hz, 1H, s’), 4.13 (q, J = 6.6 Hz, 1H, v), 3.91 (t, J = 5.9 Hz, 4H, k), 2.78 (t, J = 7.3 Hz, 4H, d), 2.66 – 2.59 (m, 6H, n+w+w’), 2.59 – 2.53 (m, 4H, f), 2.09 – 2.00 (m, 4H, e), 1.78 (p, J = 7.1 Hz, 4H, m), 1.71 (dt, J = 12.8, 6.1 Hz, 4H, l), 1.41 (s, 9H, af). 13C NMR (151 MHz, Acetone) δ 170.8 (u), 161.9 (c), 158.2 (j), 156.0 (ad), 145.7 (y), 143.2 (o), 139.9 (r), 137.2 (a), 135.2 (g), 130.4 (z), 130.1 (h), 128.8 (aa), 127.8 (p), 127.6 (ab), 126.0 (q), 120.9 (b), 115.0 (i), 79.5 (ae), 68.1 (k), 67.3 (x), 54.7 (v), 43.5 (s), 37.9 (d), 35.7 (n), 35.2 (w), 35.0 (f), 32.6 (e), 29.2 (l), 28.8 (m), 28.5 (af). LRESI-(+)-MS: m/z 1008.8 (100 [M+H]+). HRFT-(+)-MS: m/z 1008.5340 [M+H]+; 1008.5344 calc. for [C38H46N2O2H]⁺.
3.18: Mac-Cys(Trt)-Gly-Gly-Boc

Mac-Cys(Trt)-Boc 3.17 (140 mg, 139 µmol, 1 eq.) and triphenylmethanol (7.5 mg, 28 µmol, 0.2 eq.) were dissolved in a mixture of CH$_2$Cl$_2$:TFA (4:1, 2.75 mL, 50 mM). The solution was stirred at room temperature for 2 h before toluene (5 mL) was added and the mixture concentrated to dryness in vacuo at 40 °C. The resulting solid was twice further taken up in toluene (5 mL) and concentrated in vacuo. The crude solid was triturated with a mixture of Et$_2$O:n-hexane (1:1, 3 × 5 mL) before being thoroughly dried in vacuo.

Boc-GlyGly-OH (48 mg, 208 µmol, 1.5 eq.), EDCI·HCl (41 mg, 208 µmol, 1.5 eq.) and HOBt·H$_2$O (32 mg, 208 µmol, 1.5 eq.) were added to the crude residue and the mixture was dissolved in THF:CH$_2$Cl$_2$ (3:1, 2.8 mL, 50 mM) and stirred at room temperature for 14 h. Solvent was removed in vacuo and the residue purified by flash column chromatography (SiO$_2$, 2–4% MeOH/CH$_2$Cl$_2$) to give the title compound as a white solid (155 mg, quant.).

m.p. 160–163 °C. [α]$_D^{10}$ = +9.1 (c = 1.00, CH$_2$Cl$_2$). $^1$H NMR (600 MHz, Chloroform-d) δ 7.49 (t, J = 7.6 Hz, 1H, a), 7.40 (d, J = 7.9 Hz, 6H, z), 7.25 (t, J = 7.9 Hz, 6H, aa), 7.20 (t, J = 7.2 Hz, 3H, ab), 7.03 (d, J = 8.4 Hz, 4H, h), 6.96 (d, J = 7.6 Hz, 2H, b), 6.94 (s, 1H, p), 6.85 (s, 2H, q), 6.74 (d, J = 8.4 Hz, 4H, i), 6.38 – 6.31 (m, 2H, t+ac), 5.05 (s, 1H, ai), 4.35 (dd, J = 14.7, 5.9 Hz, 1H, s), 4.24 (dd, J = 14.7, 5.3 Hz, 1H, s'), 4.15 – 4.08 (m, 1H, v), 3.90 (t, J = 5.1 Hz, 4H, k), 3.82 (d, J = 5.2 Hz, 2H, ae), 3.71 (d, J = 5.8 Hz, 2H, ah), 2.86 – 2.76 (m, 5H, d+w), 2.62 – 2.51 (m, 9H, f+n+w'), 2.00 (p, J = 7.8 Hz, 4H, e), 1.79 – 1.68 (m, 8H, l+m), 1.41 (s, 9H, al). $^{13}$C NMR (151 MHz, Chloroform-d) δ 170.1 (ag), 169.4 (u), 168.3 (ad), 161.5 (c), 157.3 (j), 144.5 (y), 142.9 (o), 137.7 (r), 136.7 (a), 134.6 (g), 129.7 (z), 129.4 (h), 128.2 (aa), 127.5 (p), 127.1 (ab), 125.7 (q), 120.3 (b), 114.4 (i), 80.7 (ak), 67.7 (k), 67.4 (x), 52.6 (v), 44.4 (ah), 43.7 (s), 43.2 (ae), 38.0 (d), 35.2 (n), 34.8 (f), 33.9 (w), 32.6 (e), 28.7 (l), 28.4 (al), 28.1 (m). 156.2 (a) is not detected by $^{13}$C NMR but a cross-peak with (ah) is visible in HMBC. LRESI-(+)-MS: m/z 572.7 (100, [M+Na+H]$^+$), 1122.8 (50, [M+H]$^+$), 1144.8 (9, [M+Na]$^+$). HRFT-(+)-MS: m/z 1122.5763 [M+H]$^+$; 1122.5773 calc. for [C$_{69}$H$_{79}$N$_5$O$_7$SH]$^+$. 

![Diagram of Mac-Cys(Trt)-Gly-Gly-Boc](image-url)
3.5.1.2 Synthesis of Leucine Barrier 3.22

![Chemical Structures](attachment:image.png)

Scheme 2: Leucine barrier synthesis. Reagents and conditions: i) 4-ethynylbenzoic acid, EDCI·HCl, HOBt·H₂O, NET₃, CH₂Cl₂, r.t., 26 h, 79%; ii) LiOH, MeOH, H₂O, r.t., 5.5 h, quant.; iii) 4-(2-(4-azidophenyl)propan-2-yl)aniline 3.25 EDCI-HCl, HOBt-H₂O, THF, CHCl₃, r.t., 5.25 h, 73%; iv) Boc-Leu-OH, EDCI-HCl, DMAP, r.t., 6 h, 80%.

3.19: N-(4-Ethynyl)-benzoyltyrosine methyl ester

4-Ethynyl benzoic acid (1.96 g, 13.4 mmol, 1 eq.), tyrosine methyl ester hydrochloride (3.42 g, 14.8 mmol, 1.1 eq.), EDCI-HCl (3.09 g, 16.1 mmol, 1.2 eq.) and HOBt·H₂O (205 mg, 1.34 mmol, 0.1 eq.) were suspended in CH₂Cl₂ (60 mL, 225 mM) and NET₃ (2.1 mL, 1.5 g, 14.8 mmol, 1.1 eq.) was added dropwise. The resulting solution was stirred for 22 h before further EDCI-HCl (1.50 g, 8.0 mmol, 0.6 eq.) was added. The solution was stirred for a further 4 h before being diluted with CH₂Cl₂ (100 mL) and washed with aqueous NaHCO₃ (sat. 100 mL) then NH₄Cl (sat. 100 mL). The organics were dried over MgSO₄ and solvent was removed in vacuo. The residue was purified by flash column chromatography (SiO₂, 2–2.5% MeOH/CHCl₃) to afford the title compound as a yellow foam (3.42 g, 79%).

m.p. 126–127 °C. [α]D₂⁰ = +90.5 (c = 1.00, CH₂Cl₂). ¹H NMR (600 MHz, Acetone-d₆) δ 8.20 (s, 1H, o), 7.90 (d, J = 7.7 Hz, 1H, h), 7.84 (d, J = 8.5 Hz, 2H, e), 7.56 (d, J = 8.4 Hz, 2H, d), 7.13 (d, J = 8.4 Hz, 2H, l), 6.75 (d, J = 8.5 Hz, 2H, m), 4.82 (td, J = 8.7, 5.5 Hz, 1H, i), 3.83 (s, 1H, a), 3.68 (s, 3H, q), 3.16 (dd, J = 14.0, 5.5 Hz, 1H, j), 3.04 (dd, J = 14.0, 8.9 Hz, 1H, j'). ¹³C NMR (151 MHz, Acetone) δ 172.9 (p), 166.5 (g), 157.1 (n), 135.3 (l), 132.7 (d), 131.1 (l), 128.8 (k), 128.3 (e), 126.1 (c), 116.1 (m), 83.5
(b), 81.4 (a), 55.5 (i), 52.3 (q), 37.2 (j). LRESI-(+)-MS: m/z 346.1 (18, [M+Na]+), 669.2 (100, [2M+Na]+). HRFT-(+)-MS: m/z 346.1061 [M+Na]+; 346.1055 calc. for \([\text{C}_{19}\text{H}_{17}\text{NO}_{4}\text{Na}]^+\).

### 3.20: \(N\)-(4-Ethynyl)-benzoyltyrosine

\(N\)-(4-Ethynyl)-benzoyltyrosine methyl ester 3.19 (3.41 g, 10.55 mmol, 1 eq.) was dissolved in a mixture of MeOH (10 mL) and water (5 mL). LiOH (1.27 g, 52.7 mmol, 5 eq.) was added and the resulting solution was stirred at room temperature for 5.5 h. The reaction mixture was concentrated to ca. 5 mL to remove MeOH and aqueous HCl (1 M) was added to neutralise the resulting aqueous solution upon which a white solid precipitated. The precipitate was collected by vacuum filtration and then azeotroped as a suspension in toluene (30 mL) and dried in vacuo to give the title compound as an off-white amorphous solid (3.26 g, quant.).

m.p. 87–89 °C. \([\alpha]_{D}^{20} = +75.9\) (c = 1.00, 1% MeOH/CH₂Cl₂).

\(^1\)H NMR (600 MHz, DMSO-d₆) δ 12.74 (br. s, 1H, q), 9.19 (s, 1H, o), 8.75 (d, \(J = 8.1\) Hz, 1H, h), 7.81 (d, \(J = 8.3\) Hz, 2H, e), 7.57 (d, \(J = 8.3\) Hz, 2H, d), 7.09 (d, \(J = 8.4\) Hz, 2H, l), 6.64 (d, \(J = 8.4\) Hz, 2H, m), 4.52 (ddd, \(J = 10.7, 8.2, 4.4\) Hz, 1H, i), 4.39 (s, 1H, a), 3.06 (dd, \(J = 13.9, 4.3\) Hz, 1H, j), 2.94 (dd, \(J = 13.8, 10.7\) Hz, 1H, j'). \(^{13}\)C NMR (151 MHz, DMSO) δ 173.2 (p), 165.5 (h), 155.8 (n), 134.0 (f), 131.6 (d), 130.0 (l), 128.1 (k), 127.7 (e), 124.6 (c), 115.0 (m), 82.9 (a), 82.9 (b), 54.7 (i), 35.5 (j). LRESI-(−)-MS: m/z 308.1 (100, \([\text{M} - \text{H}]^-\)); 617.2 (58, [2M-H]). HRFT-(−)-MS: m/z 308.0921 [M−H]; 308.0917 calc. for \([\text{C}_{18}\text{H}_{14}\text{NO}_{4}]^-\).

### 3.21: Unloaded Azide-Alkyne Barrier

\(N\)-(4-Ethynyl)-benzoyltyrosine 3.20 (1.11 g, 3.60 mmol, 1 eq.), 4-(2-(4-azidophenyl)propan-2-yl)aniline 3.25 \(^{[12]}\) (1 g, 3.96 mmol, 1.1 eq.), EDCI·HCl (1.07 g, 5.40 mmol, 1.5 eq.) and HOBt·H₂O (606 mg, 3.96 mmol, 1.1 eq.) were dissolved in THF:CHCl₃ (3:1, 75 mL, 48 mmol) and stirred at room temperature for 5.25 h. The reaction mixture was then diluted with CHCl₃ (100 mL) and washed with aqueous HCl (1 M, 100 mL) followed by brine (100 mL). The organics were dried over MgSO₄ and solvent was removed in vacuo to give a crude yellow foam. The crude solid was purified by flash column chromatography (SiO₂, 5–8% EtOAc/CH₂Cl₂) to give the title compound as an off-white amorphous solid (1.43 g, 73%).

m.p. >250 °C (charring). \([\alpha]_{D}^{20} = +59.0\) (c = 1.00, CH₂Cl₂).

\(^1\)H NMR (600 MHz, Acetone-d₆) δ 9.36 (s, 1H, q), 8.15 (s, 1H, o), 7.94 (d, \(J = 7.9\) Hz, 1H, h), 7.88 (d, \(J = 7.9\) Hz, 2H, e), 7.58 – 7.52 (m, 4H, d+s),
7.28 (d, $J = 7.8$ Hz, 2H, γ), 7.20 – 7.14 (m, 4H, I+I), 6.99 (d, $J = 7.7$ Hz, 2H, γ), 6.73 (d, $J = 7.6$ Hz, 2H, m), 4.91 (q, $J = 7.4$ Hz, 1H, i), 3.83 (s, 1H, a), 3.23 (dd, $J = 13.9, 5.8$ Hz, 1H, j), 3.07 (dd, $J = 13.9, 8.5$ Hz, 1H, j), 1.65 (s, 6H, w). $^{13}$C NMR (151 MHz, Acetone) δ 170.57 (p), 166.65 (g), 156.99 (n), 148.75 (x), 146.47 (u), 138.10 (aa), 137.57 (r), 135.36 (f), 132.69 (d), 131.21 (l), 129.16 (y+k), 128.42 (e), 127.76 (t), 126.12 (c), 120.14 (s), 119.41 (z), 116.00 (m), 83.47 (b), 81.36 (a), 57.12 (i), 42.88 (v), 37.77 (j+γ'), 30.98 (w). LRESI-(−)-MS: m/z 542.3 (100, [M−H]−), 578.3 (38, [M+Cl]−). LRESI-(+)-MS: m/z 566.3 (25, [M+Na]+), 1108.9 (100, [2M+Na]+). HRFT-(+)-MS: m/z 566.2157 [M+Na]+; 566.2163 calc. for [C$_{33}$H$_{58}$N$_{6}$O$_{3}$Na]$^{+}$.

3.22: Boc-Leucine-Loaded Azide Alkyne Barrier

Unloaded Azide-Alkyne Barrier 3.21 (50 mg, 92 µmol, 1 eq.), Boc-Leu-OH (23.4 mg, 101 µmol, 1.1 eq.), EDCI-HCl (20 mg, 106 µmol, 1.15 eq.) and DMAP (5.6 mg, 46 µmol, 0.5 eq.) were dissolved in THF:CHCl$_3$ (3:1, 2 mL, 50 mmol) and stirred at room temperature for 6 h. Solvent was then removed in vacuo and the solid residue was purified by flash column chromatography (SiO$_2$, 5% EtOAc/CH$_2$Cl$_2$) to give the title compound as an amorphous white solid (56 mg, 80%).

m.p. decomposes at 95 °C. [α]$_{50}$ = $^{13}$H NMR (600 MHz, Acetone-$d_6$) δ 9.41 (s, 1H, x), 8.03 (d, $J = 7.7$ Hz, 1H, h), 7.87 (d, $J = 8.0$ Hz, 2H, e), 7.60 – 7.51 (m, 4H, d+z), 7.40 (d, $J = 8.1$ Hz, 2H, l), 7.28 (d, $J = 8.6$ Hz, 2H, a)), 7.17 (d, $J = 8.6$ Hz, 2H, aa), 7.03 (d, $J = 8.2$ Hz, 2H, m), 6.99 (d, $J = 8.6$ Hz, 2H, ag), 6.45 (d, $J = 7.8$ Hz, 1H, t), 5.02 – 4.96 (m, 1H, i), 4.39 – 4.31 (m, 1H, p), 3.83 (s, 1H, a), 3.35 (dd, $J = 13.9, 5.7$ Hz, 1H, j), 3.18 (dd, $J = 13.9, 8.7$ Hz, 1H, j'), 1.84 (hept, $J = 6.7$ Hz, 1H, r), 1.81 – 1.68 (m, 2H, q+q'), 1.65 (s, 6H, ad), 1.40 (s, 9H, v), 1.01 – 0.95 (m, 6H, s+s'). $^{13}$C NMR (151 MHz, Acetone) δ 172.7 (o), 170.3 (w), 166.8 (g), 156.6 (t), 150.6 (n), 148.7 (ae), 146.6 (ab), 138.1 (ah), 137.5 (y), 136.3 (k), 135.3 (f), 132.7 (d), 131.2 (l), 129.2 (af), 128.4 (e), 127.8 (aa), 126.2 (c), 122.2 (m), 120.2 (z), 119.4 (ag), 83.5 (b), 81.4 (a), 79.4 (u), 56.8 (i), 53.4 (p), 42.9 (ac), 41.0 (q), 37.8 (j), 31.0 (ad), 28.5 (v), 25.6 (r), 23.2 (s), 21.8 (s'). LRESI-(−)-MS: m/z 779.6 (33, [M+Na]−), 795.5 (100, [M+K]−). HRFT-(+)-MS: m/z 774.3971 [M+NH$_4$]+; 774.3974 calc. for [C$_{44}$H$_{40}$N$_{6}$O$_{3}$NH$_4$]$^{+}$. 116
3.5.1.3 Synthesis of Copolymer 3.3

Scheme 3: Synthesis of copolymer tailpiece 3.3. Reagents and conditions: i) NaOH\textsubscript{aq} (5 M), THF, r.t., 15 h; ii) Boc-Leu-OH, DCC, DMAP, CH\textsubscript{2}Cl\textsubscript{2}, r.t., 120 h; 86% over 2 steps; iii) CuBr, dNbpy, anisole, 90 °C, 24 h, 22% conv.; iv) Bu\textsubscript{3}SnH (1 M in cyclohexane), toluene, 85 °C, 1 h, 95%. ran. = random distribution.

3.6: \textit{p-}{(N-Boc)}leucyloxystyrene

The synthesis and characterisation of 3.6 was carried out by Dr Sonja Kuschel. NaOH\textsubscript{aq} (5 M, 20 ml) was added to a solution of 4-acetoxystyrene (3.06 ml, 20 mmol, 1.0 eq.) in THF (25 ml). The resulting biphasic mixture was stirred at room temperature for 15 hours, then the solution was neutralised by addition of HCl\textsubscript{aq} (1 M). The aqueous phase was extracted with EtOAc (3 x 30 ml) and the combined organic fractions were dried (MgSO\textsubscript{4}) and solvent was removed under reduced pressure. The residue was dried under high vacuum 48 h to afford 2.40 g of a white solid which was used without further purification.

To the crude hydrolysis product (2.40 g, assumed 20 mmol, 1.0 eq.) in CH\textsubscript{2}Cl\textsubscript{2} (400 ml), Boc-Leu-OH (6.94g, 30 mmol, 1.5 eq.), DCC (6.2g, 30 mmol, 1.5 eq.) and DMAP (244 mg, 2 mmol, 0.1 eq.) were added. The resulting suspension was stirred at room temperature for 5 days and then filtered. The filtrate was concentrated under reduced pressure and the residue was purified by
flash column chromatography (SiO₂, 1% MeOH/CH₂Cl₂) to afford the title compound as an amorphous white solid (5.50 g, 86% yield).

m.p. 70–71 °C. [α]D²⁰ = -30.9 (c = 0.99, CH₂Cl₂). ¹H NMR (600 MHz, Chloroform-d) δ 7.41 (d, J = 8.2 Hz, 2H, c), 7.06 (d, J = 8.5 Hz, 2H, d), 6.70 (dd, J = 17.6, 10.9 Hz, 1H, b), 5.71 (d, J = 17.5 Hz, 1H, aₜran), 5.25 (d, J = 10.9 Hz, 1H, a₁cis), 4.95 (d, J = 8.7 Hz, 1H, i), 4.56 – 4.50 (m, 1H, e), 1.87 – 1.75 (m, 2H, f + g), 2.69 (ddd, J = 13.1, 9.4, 5.1 Hz, 1H, f'), 1.46 (s, 9H, j), 1.02 (d, J = 6.2 Hz, 6H, h). ¹³C NMR (151 MHz, Chloroform) δ 172.3, 155.6, 150.2, 136.0, 135.7, 127.4, 121.6, 114.3, 80.2, 52.5, 41.8, 28.5, 23.1, 22.1. LRESI-(+)MS: m/z 334.2 (100, [M+H]⁺). HRESI(+)MS: m/z 334.2013 [M+H]⁺; 334.2013 calc. for [C₁₉H₂₇N₁O₄H]⁺.

3.7: Polystyrene Track (Bromo)

The synthesis and characterisation of 3.7 was carried out by Dr Guillaume De Bo. A solution, degassed by three freeze-pump-thaw cycles, of initiator 3.4 (27.7 mg, 0.10 mmol, 1.0 eq.), styrene 3.5 (2.750 mL, 24.00 mmol, 240.0 eq.), p-(N-Boc)leucyloxystyrene 3.6 (667 mg, 2.00 mmol, 20.0 eq.) and 4,4’-Dinonyl-2,2’-dipyridyl (dNbpy, 86 mg, 0.21 mmol, 2.1 eq.) in anisole (0.83 mL) was introduced to a sealed flask, purged with nitrogen (5×), containing CuBr (99.999%, 14.3 mg, 0.10 mmol, 1.0 eq.). The resulting solution was stirred for 15 min at room temperature and then stirred for 24 hours in an oil bath at 90 °C. The reaction was quenched by dipping the reaction flask in liquid nitrogen for 1 minute (crude ¹H NMR indicated 22% conversion). The crude reaction solution was diluted with CH₂Cl₂ (2 mL) and precipitated in a vortex of rapidly stirring MeOH (500 mL) to afford, after filtration, the title compound as an amorphous white powder (570 mg).

¹H NMR (600 MHz, Chloroform-d) δ 7.25 – 6.12 (m, 220H, j+k+l+m+n), 4.97 (br. s, 3.7H, s), 4.54 (br. s, 4.3H, o), 4.45 – 4.26 (m, 1.2H, o+i), 4.18 – 3.78 (m, 2H, b), 2.55 – 1.17 (m, 295H, d+e+f+g+h+p+q+t), 1.02 (s, 30H, r), 0.99 – 0.94 (m, 3H, c), 0.93 – 0.88 (m, 3H, c'). Mᵣn = 7100 (SEC), PDI = 1.10.

3.3: Polystyrene Track (des-Bromo)

A solution of polymer 3.7 (100 mg, 13.9 μmol, 1 eq.) in toluene (1.40 ml) was degassed via nitrogen sparging for 5 min, then Bu₃SnH (1M in cyclohexane, 42 μL, 42 μmol, 3 eq.) was added and the solution was stirred at 85 °C for
1h. The crude mixture was cooled to room temperature and precipitated in a vortex of rapidly stirring MeOH (150 mL) to afford, after filtration, the title compound as an amorphous white powder (94 mg, 95%).

$^1$H NMR (600 MHz, Chloroform-$d$) $\delta$ 7.24 – 6.16 (m, 285H, $j+k+l+m+n$), 4.97 (br. s, 3.9H, s), 4.52 (br. s, 3.9H, o), 4.20 – 3.73 (m, 2H, b), 2.55 – 1.17 (m, 282H, $d+e+f+g+h+p+q+t$), 1.02 (s, 30H, r), 0.98 – 0.86 (m, 9H, $c+c^{'}$+$\text{'grease'}$). PDI = 1.10.
3.5.1.4  Synthesis of Molecular Machine 3.1

Scheme 4: Rotaxane synthesis and elongation. Reagents and conditions: i) CuPF₆·4MeCN, CH₂Cl₂, tBuOH, r.t., 17 h, 46%; ii) CuPF₆·4MeCN, Tenta-Gel™-TBTA, CH₂Cl₂, tBuOH, r.t., 19 h, 96%.
3.2: Rotaxane Formation

Macrocyle 3.18 (111 mg, 99 µmol, 1.5 eq.) and CuPF$_6$·4MeCN (12.3 mg, 33 µmol, 0.5 eq.) were dissolved in CH$_2$Cl$_2$ (2 mL) and the resulting solution was both degassed and concentrated by approximately half via nitrogen sparging. The resulting solution was added to azide stopper 3.23 [37] (233 mg, 396 µmol, 6 eq.) and Boc-Leu-loaded barrier 3.22 (50 mg, 66 mmol, 1 eq.) in CH$_2$Cl$_2$ (1 mL) with CH$_2$Cl$_2$ (2 × 0.5 mL) being used as a transfer wash. The resulting mixture was further degassed via nitrogen sparging to give a total reaction volume of ca. 2 mL (c ≈ 33 mM w.r.t. the barrier/limiting reagent). tBuOH (2 drops) was added and the solution stirred at room temperature for 17 h. The resulting solution was concentrated to dryness in vacuo. The crude residue was purified by flash column chromatography (SiO$_2$, EtOAc/pet. ether gradient: 20% [stopper] – 50% [free thread] – 55–100% [rotaxane product and macrocycle]). The product-containing fractions were collected and further purified by preparative scale TLC (SiO$_2$, 4 × 1000 µm Merck plates, 4% MeOH/CH$_2$Cl$_2$) to give the title rotaxane as a light yellow amorphous solid (74 mg, 46%).

$^1$H NMR (600 MHz, Acetone-$d_6$) δ 9.68 (s, 1H, Bp), 8.10 (d, J = 8.1 Hz, 1H, Bh), 7.97 (br. t, J = 4.9 Hz, 1H, Mt), 7.83 (s, 1H, Ba), 7.77 (d, J = 8.4 Hz, 1H, Cl), 7.73 (br. t, J = 5.4 Hz, 1H, Cl), 7.61 (d, J = 7.8 Hz, 2H, Be), 7.58 (d, J = 8.7 Hz, 2H, Br), 7.56 – 7.50 (m, 3H, Bd+Ma), 7.39 (d, J = 8.3 Hz, 2H, Bl), 7.33 (d, J = 7.7 Hz, 6H, Cf), 7.30 – 7.26 (m, 9H, Sd+Bx+Ch), 7.20 (t, J = 7.6 Hz, 6H, Cg), 7.17 (d, J = 8.8 Hz, 2H, Bs), 7.15 (s, 1H, Mp), 7.12 (d, J = 8.6 Hz, 6H, Se), 7.04 – 6.97 (m, 6H, Bm+By+Mb), 6.96 (s, 2H, Mq), 6.81 (d, J = 8.9 Hz, 2H, Mh), 6.79 (d, J = 8.9 Hz, 2H, Mh'), 6.58 (d, J = 8.7 Hz, 2H, Sj), 6.46 (d, J = 8.1 Hz, 1H, Lf), 6.40 (d, J = 8.5 Hz, 2H, Mi), 6.38 (d, J = 8.5 Hz, 2H, Mi'), 6.29 (s, 1H, Co), 5.19 – 5.12 (m, 1H, Bi), 4.70 – 4.64 (m, 1H, Cb), 4.53 (dd, J = 14.9, 6.0 Hz, 1H, Ms), 4.41 – 4.35 (m, 1H), 4.14 (dd, J = 15.2, 4.1 Hz, 1H, Ms'), 4.07 (t, J = 6.9 Hz, 2H, Sn), 4.02 (br. s, 2H, Ck), 3.76 – 3.72 (m, 2H, Cn), 3.72 – 3.67 (m, 2H, Mk), 3.66 – 3.61 (m, 2H, Mk'), 3.50 (t, J = 5.8 Hz, 2H, Sl), 3.31 (dd, J = 13.7, 5.5 Hz, 1H, Bj), 3.14 (dd, J = 13.7, 8.8 Hz, 1H, Bj'), 2.69 (dd, J = 12.0, 5.6 Hz, 1H, Cc), 2.67 –
Chapter 3

2.60 (m, 4H, Mf), 2.60 – 2.46 (m, 8H, Md+Mn), 2.46 – 2.39 (m, 1H, Cc'), 1.94 – 1.81 (m, 7H, Sm+Ld+Me), 1.80 – 1.70 (m, 4H, Mm), 1.69 – 1.56 (m, 10H, Ml+Ml'+Bv), 1.40 (s, 18H, Li+Cr), 1.29 (s, 27H, Sa), 0.98 (d, J = 6.7 Hz, 3H, Le), 0.96 (d, J = 6.6 Hz, 3H, Le'). \(^{13}\)C NMR (151 MHz, Acetone) \(\delta\) 172.8 (La), 171.2 (Cm), 170.5 (Bo), 170.3 (Ca), 170.0 (Cj), 167.2 (Bg), 162.5 (Mg), 162.5 (Mc'), 158.0 (Mj), 158.0 (Mj'), 157.3 (Sk), 157.1 (Cp), 156.6 (Lg), 150.6 (Bn), 149.1 (Sc), 148.8 (Bw), 146.8 (Bb), 146.5 (Bt), 145.6 (Ce), 145.3 (Sf), 140.3 (Sh), 138.1 (Bz), 137.6 (Bq+Ma), 136.3 (Bk), 135.0 (Bc), 134.3 (Mg), 134.2 (Mg'), 133.2 (Bf), 132.7 (Si), 131.4 (Se), 131.3 (Bl), 130.3 (Cf), 130.0 (Mh), 130.0 (Mh'), 129.2 (Bx+Ch), 128.8 (Cg), 128.6 (Be), 127.8 (Bs), 127.7 (Mp), 126.1 (Mq), 125.9 (Bd), 125.0 (Sd), 122.2 (Bm), 122.0 (Ba), 120.6 (Mb), 120.6 (Mb'), 120.2 (Br), 119.4 (By), 114.9 (Mj), 114.9 (Mj'), 114.1 (Sj), 79.6 (Lh/Cq), 79.4 (Lh/Cq) 68.2 (Mk), 68.1 (Mk'), 67.3 (Cd), 64.7 (Si), 63.9 (Sg), 56.6 (Bi), 53.4 (Lb), 53.3 (Cb), 47.4 (Sn), 44.8 (Cn), 43.6 (Ms), 43.5 (Ck), 42.9 (Bu), 41.1 (Lc), 38.4 (Bj+Md), 36.3 (Mn), 36.2 (Mn'), 35.6 (Mf), 35.2 (Cc), 34.9 (Sb), 33.0 (Me), 31.7 (Sa), 31.0 (Bv), 30.3 (Sm), 28.6 (Cr+Li), 25.6 (Ld), 23.3 (Le), 21.8 (Le'). \(\textit{Mo and Mr} \quad ^{13}\text{C signals are not observed}. \) \(\textit{Ml and Mn} \quad ^{13}\text{C signals are beneath the residual solvent signal at } \delta \text{ ca. 29.5 ppm}. \) LRESI(+)-MS: m/z 2466.1 (55, [M+H]^+), 2467.1 (94, [M+H]^+), 2468.1 (100, [M+H]^+), 2469.0 (57, [M+H]^+), 2470.0 (48, [M+H]^+), 2471.0 (13, [M+H]^+), 2488.1 (50, [M+Na]^+), 2489.1 (100, [M+Na]^+), 2490.1 (96, [M+Na]^+), 2491.1 (55, [M+Na]^+), 2492.1 (37, [M+Na]^+), 2493.1 (26, [M+Na]^+), 2494.1 (13, [M+Na]^+). HRFSI(+)-MS: m/z 2466.3287 [M+H]^+; 2466.3284 calculated for [C\textsubscript{153}H\textsubscript{176}N\textsubscript{14}O\textsubscript{14}SH]^+. 

Figure 5: Left hand side shows LRESI(+)-MS: observed (top left); predicted (bottom left) for [C\textsubscript{153}H\textsubscript{176}N\textsubscript{14}O\textsubscript{14}SHNa]^+ corresponding to [M+H+Na]^+. Centre shows observed LRESI(+)-MS: observed (top right); predicted (bottom right) for [C\textsubscript{153}H\textsubscript{176}N\textsubscript{14}O\textsubscript{14}SH]^+ corresponding to [M+Na]^+. Right hand side shows LRESI(+)-MS: observed (top right); predicted (bottom right) for [C\textsubscript{153}H\textsubscript{176}N\textsubscript{14}O\textsubscript{14}SH]^+ corresponding to [M+H]^+ and [C\textsubscript{153}H\textsubscript{176}N\textsubscript{14}O\textsubscript{14}SNa]^+ corresponding to [M+Na]^+. The observed (top right) spectrum also shows a weak signal corresponding to [2M+H+Na]^+ centred around m/z 2478.7 where \( z = 2 \).
CuPF₆·4MeCN (1.6 mg, 4.35 µmol, 2 eq.) and Tenta-Gel™-TBTA resin (0.17 mmol/g loading, 38 mg, 6.53 µmol, 3 eq.) were suspended in CH₂Cl₂ (1.00 mL) and suspension was both degassed and concentrated by approximately half via nitrogen sparging. 1-Barrier rotaxane 3.2 (5.4 mg, 2.18 µmol, 1 eq.) and alkyne terminated polymer 3.3 (31 mg, 4.35 µmol, 2 eq.) were dissolved in CH₂Cl₂ (0.50 mL) and tBuOH (0.13 mL) and this solution was added to the suspension of resin-bound copper with CH₂Cl₂ (2 × 0.25 mL) being used as a transfer wash. The resulting suspension was concentrated to approximately 0.75 mL (c ≈ 2.9 mM) via further nitrogen sparging. The resulting suspension was stirred at room temperature for 19 h before the resin was removed by filtration and the filtrate was purified by preparative scale TLC (SiO₂, 2 × 500 µm Analtech plates, 4% MeOH/CH₂Cl₂) to give the title rotaxane as a white amorphous solid (20 mg, 96%).

$^1$H NMR (600 MHz, Acetone-$d_6$) δ 9.66 (br. s, 1H, Bp), 8.32 (br. s, 1H, Ta), 8.05 (d, $J = 8.0$ Hz, 1H, Bh), 7.95 (d, $J = 10.0$ Hz, 1H, Mt), 7.83 (br. s, 1H, Ba), 7.80 – 7.69 (m, 4H, By+Cl+Cl), 7.63 – 7.57 (m, 4H, Be, Br), 7.56 – 7.50 (m, 3H, Bd, Ma), 7.48 – 7.42 (m, 2H, Bx), 7.39 (d, $J = 8.5$ Hz, 2H, Bl), 7.33 (d, $J = 8.0$ Hz, 6H, Cf), 7.28 (d, $J = 8.7$ Hz, 6H, Sd), 7.26 – 6.44 (m, 341H, CH₆), 6.40 (d, $J = 8.2$ Hz, 2H, Mi), 6.38 (d, $J = 7.8$ Hz, 2H, Mi'), 6.28 (br. t, 1H, Co), 5.15 (br. s, 1H, Bi), 4.67 (br. s, 1H,Cb), 4.57 – 4.50 (m, 1H, Ms), 4.49 – 4.33 (m, 5H, Lb, Td), 4.15 (d, $J = 15.1$ Hz, 1H, Ms'), 4.07 (br. s, 2H, Sn), 4.03 (br. s, 2H, Ck), 3.74 (d, $J = 6.0$ Hz, 2H, Cn), 3.69 (br. s, 2H, Mk), 3.64 (br. s, 2H, Mk'), 3.51 (dd, $J = 9.3$, 4.7 Hz, 2H, Sl), 3.31 (d, $J = 13.3$, 4.5 Hz, 1H, Bj), 3.17 – 3.10 (m, 1H, Bj'), 2.69 (dd, $J = 12.1$, 4.7 Hz, 1H, Cc), 2.64 (q, $J = 6.0$ Hz, 4H, Mf), 2.59 – 2.47 (m, 8H, Md, Mn), 2.46 – 2.36 (m, 1H, Cc'), 2.01 – 1.41 (m, 188H, CH), 1.45 (br. s, 36H, Ti), 1.40 (br. s, 18H, Cr, Li), 1.29 (s, 27H, Sa), 1.01 (s, 40H, Tg), 1.00 – 0.94 (m, 6H, Le). Mₙ = 10200 (SEC), PDI = 1.08.
Figure 6: ESI-(+)-MS of machine 3.1 as a triply charged adduct $[\text{M+H+2Na}]^{3+}$.

Figure 7: ESI-(+)-MS zoom of machine 3.1. Top: predicted isotope pattern for polymer with 50 styrene monomers and 5 Leu-barriers (triply charged species as a $[\text{M+H+2Na}]^{3+}$ adduct corresponding to $[\text{C}_{655}\text{H}_{721}\text{N}_{19}\text{O}_{36}\text{S}\text{Na}_{2}]^{3+}$). Bottom: zoom of the measured ESI-(+)-MS of Machine 3.1 showing perfect overlap of the observed and predicted isotope pattern.

Figure 8: MALDI-MS of machine 3.1 as a singly charged adduct $[\text{M+Na}]^{+}$. The MALDI and ESI data and analyses are in agreement with one another.
Overlapping populations in MALDI of 3.1

The +MALDI-TOF-MS spectrum of molecular machine 3.1 (Section 3.2.2, Figure 1d) shows multiple distributions which can be attributed to populations containing a constant number of leucine units (n) and a varying number of styrene spacer units (m). On visual inspection, the most abundant populations are where the distribution peaks at n = 6, followed closely by n = 7. Absolute deconvolution of the relative abundances of each population is complicated by the overlap of peaks separated by 5 leucine residues (5n). This overlap results from the mass of 5 \( p \)-leucyloxystyrene monomers (3.6) \([5 \times 333.4 \text{ Da} = 1667.0 \text{ Da}]\) being almost identical to the mass of 16 monomers of styrene (3.5) \([16m = 16 \times 104.2 \text{ Da} = 1667.2 \text{ Da}]\).

It follows that machine 3.1 with a total of 6 leucine units (1 Leu from rotaxane 3.2 and 5 Leu from copolymer 3.3) will overlap with a machine with a total of only 1 leucine barrier (and an additional 16 monomers of styrene 3.5). Additionally, these machines would overlap with an 11-barrier machine (containing 16 fewer monomers of styrene 3.5, relative to the 6-barrier machine). This is outlined in an expansion of the +ESI-TOF-MS which shows machine 3.1 as a triply charged adduct \([\text{[3.1+H+2Na]}^{3+}]\) where peaks have been assigned according to their constituent numbers of styrene monomer (m) and the total number of leucine barriers (n, Figure 9). For example, the highlighted peak (3212.5 m/z, z = 3) is produced by machine 3.1 with an m:n ratio of 48:6 which overlaps with machine 3.1 with m:n ratios of 32:11 and 64:1).

![Figure 9: Expansion of the +ESI-TOF-MS spectrum of machine 3.1 ([M+H+2Na]^{3+}) with assignments highlighting the overlap of peaks correlating to machines separated by 5 leucine units (5n) incorporated into the polymeric track and a corresponding number of styrene monomers (m).](image-url)
3.5.1.5 Synthesis of Model Catalysts 3.11 and 3.12

Scheme 5: Model polymerisation and thiol reduction. Reagents and conditions: i) p-cresol, EDCI- HCl, DMAP, CH$_2$Cl$_2$, r.t., 17 h, 63%; ii) 3.24, TFA, TIPS, CH$_2$Cl$_2$, r.t., 2.5 h; iii) NEt$_3$, PPh$_3$, DMF, 65 °C, 74.5 h, 65% over 2 steps; iv) V-50, tBuSH, NEt$_3$, TCEP·HCl, DMF, 55 °C, 24 h, 70%. $n = 6$ on average.

3.24: Boc-Leu-OCresol

Boc-Leu-OH (500 mg, 2.16 mmol, 1 eq.), p-cresol (350 mg, 3.24 mmol, 1.5 eq.), EDCI-HCl (497 mg, 2.59 mmol, 0.2 eq.) and DMAP (396 mg, 3.24 mmol, 1.5 eq.) were dissolved in CH$_2$Cl$_2$ (20 mL, 108 mM) and stirred at room temperature for 17 h. The resulting solution was washed with aqueous NaOH (2 M, 20 mL) followed by aqueous NH$_4$Cl (sat., 20 mL). The organics were dried over MgSO$_4$ and concentrated in vacuo to a colourless oil. The crude oil was purified by flash column chromatography (SiO$_2$, 8% EtOAc/n-hexane) to give the title compound as a white crystalline solid (439 mg, 63%).
m.p. 52 °C. [α]_D^20 = -24.2 (c = 0.91, CH₂Cl₂). \(^1\)H NMR (600 MHz, Chloroform-d) δ 7.17 (d, J = 8.0 Hz, 2H, c), 6.97 (d, J = 8.2 Hz, 2H, d), 4.95 (d, J = 8.0 Hz, 1H, k), 4.59 – 4.45 (m, 2H, g), 2.34 (s, 3H, a), 1.85 – 1.76 (m, 2H, h+i), 1.67 – 1.60 (m, 1H, h'), 1.01 (d, J = 6.3 Hz, 6H, j). \(^13\)C NMR (151 MHz, CDCl₃) δ 172.5 (f), 155.6 (l), 148.4 (e), 135.8 (b), 130.1 (c), 121.2 (d), 80.1 (m), 52.4 (g), 41.9 (h), 28.5 (n), 25.1 (i), 23.1 (j'), 21.0 (a). LRESI(+)–MS: m/z 344.3 (53, [M+Na]⁺), 360.2 (100, [M+K]⁺).

3.11 Model Cysteine Catalyst

Macrocycle 3.18 (100 mg, 89 µmol, 1 eq.) and Boc-Leu-O-Cresol ester 3.24 (172 mg, 534 µmol, 6 eq.) were dissolved in CH₂Cl₂:TFA (4:1, 3.56 mL, 25 mM) and TIPS (91 µL, 70 mg, 5 eq.) was added to produce a fluorescent yellow solution. The reaction mixture was stirred for 2.5 h then toluene (10 mL) was added. The resulting solution was concentrated to dryness in vacuo. The residue was twice further taken up in toluene (2 × 10 mL) and concentrated in vacuo. The residue was then triturated with Et₂O:n-hexane (1:1, 3 × 10 mL) before being dried in vacuo to give the deprotected mixture as an amorphous solid (in assumed quant. yield) which was operated without further purification.

To the deprotected crude mixture was added PPh₃ (51 mg, 311 µmol, 3.5 eq.) and DMF (8.9 mL, 10 mM) and the solution was heated to 65 °C. Once at temperature, NEt₃ (186 mL, 135 mg, 1.34 mmol, 15 eq.) was added and the solution stirred at 65 °C for 74.5 h. Solvent was removed from the reaction mixture in vacuo and the residue was dissolved in CH₂Cl₂ (15 mL) and washed with aqueous NaHCO₃ (sat., 2 × 15 mL) and brine (15 mL). The organics were dried over Na₂SO₄ and concentrated in vacuo to dryness. The crude residue was purified by prep. TLC (SiO₂, 4 × Analtech 1500 µm plates, 7% MeOH/CH₂Cl₂) to give the title compound as an amorphous solid (85 mg, 65%).

\(^1\)H NMR (600 MHz, DMF-d₇) δ 8.62 – 7.79 (m, ca. 10H, NH), 7.63 (t, J = 7.6 Hz, 1H, a), 7.13 – 7.07 (m, 6H, b+h), 7.06 (br. s, 1H, p), 7.02 (br. s, 2H, q), 6.80 (d, J = 8.5 Hz, 4H, i), 4.63 – 4.51 (m, 1H, v), 4.50 – 4.26 (m, 7H, s+af), 3.98 – 3.80 (m, 9H, k+z+ac+al), 3.05 – 2.97 (m, 1H, w), ca. 2.75* (5H, d+w'), 2.65 (t, J = 7.0 Hz, 4H, n), 2.58 – 2.53 (m, 4H, f), 2.00 (p, J = 7.7 Hz, 4H, e), 1.96 – 1.89 (m, 2H, ah+an), 1.82 – 1.56 (m, 24H, l+m+ag+ah'+am), 0.96 – 0.84 (m, 36H, ai+ao). \(^13\)C NMR (151 MHz,
Chapter 3

DMF-d$_7$ δ 173.5 (br. C=O), 170.1 (br. C=O), 169.0 (C=O), 161.4 (c), 157.6 (j), 142.6 (o), 139.7 (r), 137.0 (a), 134.7 (g), 129.6 (h), 127.0 (p), 125.4 (q), 120.5 (b), 114.5 (i), 67.7 (k), 56.4 (br. v), 53.7 (al), 52.4 (br. af), 44.7 (am), 43.0 (br. z+ac), 42.9 (br. s), 40.6 (ag), 37.4 (d), ca. 35.4* (n), ca. 34.5* (f) 32.2 (e), 28.8 (l), 28.3 (m), 26.7 (w), 24.8 (br. ah), 24.3 (an), 23.1(ai/ao), 21.6 (ai/ao). The majority of the carbonyl $^{13}$C signals are broad and indistinct from the baseline. *signals are covered by solvent peak. LRESI(+)-MS: m/z 560.3 (25, n=3, [M+2H]$^{2+}$), 616.8 (54, n=4, [M+2H]$^{2+}$), 673.4 (82, n=5, [M+2H]$^{2+}$), 729.9 (100, n=6, [M+2H]$^{2+}$), 786.9 (93, n=7, [M+2H]$^{2+}$), 843.4 (68, n=8, [M+2H]$^{2+}$), 900.0 (43, n=9, [M+2H]$^{2+}$), 956.4 (20, n=10, [M+2H]$^{2+}$).

![Figure 10: CD spectrum (298 K, 190–260 nm) of 3.11 (1.37 × 10$^{-4}$ M) showing partial α-helix character.](image)

3.12: Model Alanine Catalyst

To polymer macrocycle 3.11 (20 mg, 13.7 µmol, 1 eq.) was added V-50 (0.38 mg, 1.4 µmol, 0.1 eq.), TCEP-HCl (20 mg, 69 µmol, 5 eq.), DMF (274 µmol, 50 mM), NEt$_3$ (28.5 µL, 20.7 mg, 205 µmol, 15 eq.) and $^t$BuSH (15 µL, 12 mg, 137 µmol, 10 eq.). The resulting solution was heated to 55 °C and stirred for 24 h before being cooled to room temperature and solvent was removed in vacuo. The crude residue was purified by prep. TLC (SiO$_2$, 1 × Analtech 1000 µm plate, 7% MeOH/CH$_2$Cl$_2$) to give the title compound as a thin solid film (14 mg, 70%).
\[ \text{H NMR (600 MHz, DMF-d7) } \delta 8.59 - 7.82 \text{ (m, ca. 10H, NH)}, 7.63 \text{ (t, } J = 7.6 \text{ Hz, 1H, a)}, 7.12 - 7.07 \text{ (m, 6H, b+h)}, 7.05 \text{ (s, 1H, p)}, 6.99 \text{ (s, 2H, q)}, 6.80 \text{ (d, } J = 8.4 \text{ Hz, 4H, i)}, 4.52 - 4.26 \text{ (m, 7H, s+af)}, 4.03 - 3.78 \text{ (m, 9H, k+z+ac+al)}, 2.75* \text{ (4H, d)}, 2.66 \text{ (t, } J = 7.6 \text{ Hz, 1H, a)}), 2.00 \text{ (p, } J = 7.7 \text{ Hz, 4H, e)}, 1.96 - 1.89 \text{ (m, 2H, ah+an)}, 1.88 - 1.55 \text{ (m, 24H, l+m+ag+ah'+am)}, 1.43 - 1.34 \text{ (m, 3H, w)}, 1.02 - 0.80 \text{ (m, 36H, ai+ao)}.

\[ \text{13C NMR (151 MHz, DMF-d7) } \delta 173.4 \text{ (br. C=O)}, 172.6 \text{ (br. C=O)}, 169.0 \text{ (br. C=O)}, 161.4 \text{ (c)}, 157.6 \text{ (j)}, 142.6 \text{ (o)}, 139.9 \text{ (r)}, 137.0 \text{ (a)}, 134.7 \text{ (g)}, 129.6 \text{ (h)}, 127.0 \text{ (p), 125.3 (q), 120.5 (b), 114.5 (l), 67.7 (k), 56.3 (v), 53.7 (al), 52.5 (br. af)}, 49.3 \text{ (br. af')}, 44.7 \text{ (am)}, 43.1 \text{ (br. z+ac)}, 42.7 \text{ (br. s)}, 40.6 \text{ (ag)}, 37.4 \text{ (d)}, 35.4* \text{ (n)}, 34.6* \text{ (f)}, 32.2 \text{ (e)}, 28.8 \text{ (l)}, 28.3 \text{ (m)}, 24.9 \text{ (br. ah)}, 24.3 \text{ (an)}, 23.0 \text{ (ai/ao)}, 21.6 \text{ (ai/ao)}, 18.0 \text{ (w)}.

The majority of the carbonyl 13C signals are broad and indistinct from the baseline. *signals are covered by solvent peak. LRESI(+) - MS: m/z 544.3 (17, n=3, [M+2H]^{2+}), 600.9 (51, n=4, [M+2H]^{2+}), 657.9 (86, n=5, [M+2H]^{2+}), 714.4 (100, n=6, [M+2H]^{2+}), 770.9 (78, n=7, [M+2H]^{2+}), 827.5 (56, n=8, [M+2H]^{2+}), 884.0 (31, n=9, [M+2H]^{2+}), 940.5 (16, n=10, [M+2H]^{2+}).

![Figure 11: CD spectrum (298 K, 190 – 260 nm) of 3.12 (1.40 × 10^-4 M) showing partial α-helix character.]

### 3.5.2 Deprotection and Operation of Machine 3.1

**Deprotection of machine 3.1:** A solution of machine 3.1 (10 mg, 1.06 μmol, 1 eq.) and TIPS (5.5 μL, 26.40 μmol, 25 eq.) in CH₂Cl₂ (500 μL) and TFA (100 μL) was stirred for 2 hours at room temperature. Then the solvent was removed by azeotropic distillation with toluene (2 × 2mL). The solid residue was dried for 30 min under high vacuum, then washed with Et₂O (5 × 1mL). The residual crude solid (3.8) was dried for 30 min under high vacuum, dissolved in DMF-d7 for analysis and operated without further purification.
**Operation of machine 3.8:** A solution of freshly deprotected machine 3.8 (1.06 μmol, 1 eq.) in DMF-\textit{d}_7 (1 mL) was added to a capped 5mL Biotage® vial loaded with PPh₃ (0.8 mg, 3.2 μmol, 3 eq.) and purged with 5 vacuum/nitrogen cycles. The resulting solution was degassed via nitrogen sparging for 5 min. NEt₃ (7.5 μL, 53.0 μmol, 50 eq.) was added and the solution was stirred at 65 °C for 96 h.

An aliquot was removed for analysis (100 μL) and diluted with LC-MS grade MeOH (1 mL). A precipitate formed. The solution was filtered over a 0.45 μm PTFE membrane. MS analysis of the filtrate confirmed the production of the expected product and NMR of the precipitate shows the formation of the free-thread.

The crude reaction mixture was concentrated to dryness and purified by prep. TLC (Merck, 500 μm, DCM:MeOH:NH₃[aq] [100/10/0.5], 1 elution). The fraction at \( R_f = 0.5 \) was extracted. The recovered solid was washed with MeOH (5 × 1mL). The combined MeOH fractions were concentrated \textit{in vacuo} to afford 3.9 as a white solid (0.7 mg, 50%). HRFT(+)-MS: \( m/z \) (n = 6) 729.9695 [M+2H]^{2+}; 729.9635 calculated for \([C_{45}H_{122}N_{12}O_{11}SH_2]^2+\) where 6 leucine units are incorporated (n = 6). ESI(+)-MS:

![Figure 12: ESI(+)-MS of crude operation product 3.9. The labelled distribution shows the doubly charged adduct of oligoleucine 3.9 corresponding to [M+2H]^{2+} which is centred around the incorporation of 6 leucine units.](image)
3.5.3 Juliá-Colonna Epoxidation

\[
\begin{align*}
&\text{Scheme 6: Synthesis of chalcone 3.13 and Juliá-Colonna epoxidation. Reagents and conditions: i) 2 M NaOH}_aq, \text{EtOH, r.t., 1 h, 58%; ii) Catalyst 3.11 or 3.12, urea·H}_2O_2, \text{DBU, THF, r.t., 4 days.}

\end{align*}
\]

3.13: Furly Chalcone

2-acetyl furan (0.50 g, 4.50 mmol, 1 eq.) and benzaldehyde (0.46 mL, 0.48 g, 1 eq.) were stirred in EtOH (20 mL) and NaOH\(_{aq}\) (2 M, 10 mL) for 1 h at room temperature. The reaction mixture was acidified with HCl\(_{aq}\) (1 M, 22 mL) and a white solid precipitated. The solid was collected by vacuum filtration and washed with H\(_2\)O (15 mL) then ice cold Et\(_2\)O (15 mL) and dried in vacuo to give the title compound as an off-white crystalline solid (0.52 g, 58%).

m.p. 87–88 °C. \(^1\)H NMR (600 MHz, Chloroform-\(d\)) \(\delta\) 7.89 (d, \(J = 15.8\) Hz, 1H, e), 7.69 – 7.63 (m, 3H, c+i), 7.46 (d, \(J = 15.8\) Hz, 1H, f), 7.44 – 7.40 (m, 3H, a+b), 7.34 (d, \(J = 3.5\) Hz, 1H, k), 6.60 (dd, \(J = 3.4, 1.5\) Hz, 1H, j). \(^1\)C NMR (151 MHz, Chloroform) \(\delta\) 178.2 (g), 153.8 (h), 146.7 (i), 144.2 (e), 134.9 (d), 130.8 (a), 129.1 (b), 128.7 (c), 121.3 (f), 117.7 (k), 112.7 (j). LRESI(+) -MS: \(m/z\) 198.8 (100, [M+H])**.

3.14: Juliá-Colonna Epoxidation Product

Procedure 1 (Catalyst 3.11): Peptide 3.11 (2.2 mg, 1.5 μmol, 0.15 eq.) was added to a solution of DBU (2.4 mg, 15.4 μmol, 1.5 eq.) and urea-H\(_2\)O\(_2\) complex (1.2 mg, 12.4 μmol, 1.2 eq.) in THF (250 μL). The mixture was stirred at room temperature for 30 min, before furly chalcone 3.13 (2.0 mg, 10.0 μmol, 1 eq.) was added. After 15h at room temperature, another portion of DBU (2.4 mg, 15.4 μmol, 1.5 eq.) and urea-H\(_2\)O\(_2\) complex (1.2 mg, 12.4 μmol, 1.2 eq.) were added. The reaction was stirred for a total of 4 days after which, the conversion was determined by \(^1\)H NMR (93%) and the e.r. by chiral HPLC (63:37).

Chiral HPLC (Chiralpak IC column, eluent: n-hexane:IPA, 95:5, 1 mL/min):
Chapter 3

Figure 13: Chiral HPLC trace of the crude reaction mixture of the Juliá-Colonna epoxidation of furyl chalcone 3.13 catalysed by 3.11 after 4 days. e.r. = 63:37.

Procedure 2 (Catalyst 3.12): Peptide 3.12 (6.0 mg, 4.2 μmol, 0.1 eq.) was added to a solution of DBU (9.6 mg, 63.0 μmol, 1.5 eq.) and urea-H₂O₂ complex (4.7 mg, 50.4 μmol, 1.2 eq.) in THF (1 mL). The mixture was stirred at room temperature for 30 min then the substrate 3.13 was added (8.3 mg, 10.0 μmol, 1 eq.). After 2 h at room temperature, another portion of DBU (9.6 mg, 63.0 μmol, 1.5 eq.) and urea-H₂O₂ complex (4.7 mg, 50.4 μmol, 1.2 eq.) were added. The reaction was stirred for a total of 18 h after which, the conversion was determined by ¹H NMR (100%) and the e.r. by chiral HPLC (96:4). The mixture was purified by prep. TLC (Merck, 500 μm, 20% EtOAc/pet. ether, 1 elution) to afford 3.14 as a white solid (8.0 mg, 89%).

[α]£D²⁰ = -108.4 (c = 0.13, CH₂Cl₂). ¹H NMR (600 MHz, Chloroform-d) δ 7.68 (dd, J = 1.7, 0.8 Hz, 1H, i), 7.46 (dd, J = 3.6, 0.7 Hz, 1H, k), 7.41 – 7.37 (m, 3H, a+b), 7.35 (dd, J = 7.7, 1.9 Hz, 2H, c), 6.60 (dd, J = 3.7, 1.7 Hz, 1H, j), 4.15 (d, J = 1.8 Hz, 1H, e/f), 4.14 (d, J = 1.8 Hz, 1H, e/f). ¹³C NMR (151 MHz, Chloroform) δ 182.2, 147.8, 135.5, 129.2, 128.9, 126.0, 119.7, 112.8, 77.2, 60.8, 59.8, 29.9.

Chiral HPLC (Chiralpak IC column, eluent: n-hexane:IPA, 95:5, 1 mL/min):

Figure 14: Chiral HPLC trace of purified epoxide 3.14. e.r. = 96:4.

Absolute stereochemistry was determined by comparison with literature values. [38,39]
3.6 References

Chapter 4

A Rotaxane-Based Molecular Machine with a Trans-Acylation Catalyst for Sequence-Specific Peptide Synthesis

Overview

This chapter describes the Author’s contributions towards a rotaxane-based molecular machine for sequence-specific oligopeptide synthesis which operates using an alternative mode of catalysis from those discussed previously in this thesis. Where NCL was used in previously discussed projects to facilitate amide bond formation, this work examines bond formation via acyl-transfer catalysis. Prior to the Author’s involvement, a range of acyl-transfer catalysts were investigated over a period of ca. 5 years with varying degrees of success: this work is summarised in Section 4.1. The sections following this introduction and summary of previous work describe the Author’s contribution to the project using a 1,2,4-triazole moiety as an acyl-transfer catalyst. Multiple 1- and 2-barrier molecular machines were designed and successfully synthesised, however upon operation, these did not yield significant quantities of the desired oligopeptides. After multiple redesign processes, the decision was taken to refocus efforts on the research reported in Chapters 2 and 3, whose success is partially attributed to the experience gained in carrying out the work reported in this chapter.

Acknowledgements: The author worked closely with Dr Sonja Kuschel and Dr Guillaume De Bo while carrying out this research: their input in both the work described in this chapter and the research leading to it is greatly appreciated. In addition, the author is grateful to all Leigh group members who have contributed to Peptide Synthesiser projects in the past. The author performed the synthesis, analysis and characterisation of all compounds herein with the exception of those indicated at the beginning of Section 4.10.1: Synthetic Procedures.

Dr Jonathan J. Danon, Dr Daniel J. Tetlow and Dr Thomas A. Singleton are gratefully acknowledged for proof-reading this chapter.
Background

The first two examples of rotaxane-based molecular machines capable of sequence-defined peptide synthesis were reported by Leigh in 2013\cite{1} and 2014.\cite{2} In these systems, the Leigh group utilised cysteine-mediated native chemical ligation (NCL) to autonomously form amide bonds between amino acid building blocks ‘loaded’ upon a track in a pre-specified order. Although NCL is an efficient and reliable method for the synthesis of peptides,\cite{3,4} the Leigh group has long sought alternative chemistries by which to form sequence-specific oligopeptides \textit{via} acyl-transfer catalysis. An acyl-transfer molecular machine would display key advantages over its NCL predecessors which can be elucidated by examining how such a machine might function.

4.1 An Overview of the Intended Molecular Machine Design, its Mode of Operation and Preliminary Results

In designing an acyl-transfer molecular machine, Leigh’s core rotaxane architecture would be employed. In such a rotaxane, a threaded macrocycle is locked in a compartment between a bulky stopper and the first of three cleavable barriers bearing amino acid building blocks. Leigh’s rotaxane capping strategy\cite{2} would ensure correct initial macrocycle compartmentalisation between the stopper and first barrier (green, Figure 1). The macrocycle would again bear a catalytic unit but in this example, the Cys-Gly-Gly tripeptide linker catalyst required for NCL would be replaced by a nucleophilic acyl-transfer catalyst (arbitrarily labelled as Nu:). A generic structure of the transacylation synthesiser (4.1) is shown in Figure 1.

![Figure 1: A generic design for a rotaxane-based molecular synthesiser (4.1) bearing a nucleophilic acyl-transfer catalyst (black/red Nu:) on a macrocycle (blue) and amino acid building blocks (green, pink and red) loaded onto a linear track (black). Between the stopper and green barrier, the track threads through the cavity of the macrocycle to give an interlocked rotaxane architecture.](image-url)
4.1.1 The Intended Mode of Operation

The intended mode of operation of molecular machine 4.1 is outlined in Scheme 1. Following an activation step, the nucleophilic catalyst will be primed to react with the electrophilic carbonyls of the phenolic esters by which the building blocks are attached (through tyrosine residues on the track).

As the macrocycle and catalyst are compartmentalised between the stopper and the green barrier, it is this barrier that the catalyst will first encounter and cleave (Scheme 1, Step a). As the catalyst ‘picks up’ the amino acid it removes that barrier from its path. The movement of the macrocycle with respect the thread is governed by Brownian motion; upon removal of the first barrier, the macrocycle is free to explore the track between the stopper and the second (pink) barrier. The first building block remains attached to the catalytic unit as an activated ester during this time (4.1a). Upon encountering the second (pink) barrier, the activated ester can be trapped by the second barrier’s primary amine (Scheme 1, Step b), thus completing one full transacylation cycle. The result of this cycle is the formation of one new amide bond and the net movement of the building block from the first barrier onto the second barrier building block to give 4.1b.

The nucleophilic catalyst is then able to begin a second cycle by picking up the newly extended building block attached to the second tyrosine residue to form another activated ester (4.1c, Scheme 1, Step c). On removing the second barrier from the track the macrocycle can now explore the area between the stopper and the third barrier. On meeting the third barrier, the growing peptide chain can be deposited onto the free amine of the third (red) building block (4.1d, Scheme 1, Step d) which completes a second transacylation cycle and forms a second amide bond.

Finally, the catalytic unit can then cleave the growing peptide from the third tyrosine residue (Scheme 1, Step e). This step removes the final barrier on the rotaxane, allowing the macrocycle (4.2) to dethread from the track (4.3). The operation would be completed by trapping the activated ester of the growing peptide with a nucleophilic quench.
4.1.2 Advantages of an Acyl-Transfer Machine over an NCL Machine

There are fundamental differences to using the described nucleophilic acyl-transfer catalysis in a peptide-forming molecular machine in place of NCL. These features and their benefits are described in the following section and can be summarised as follows:

- Peptides would be grown in the ribosomal direction, from the N-terminus to the C-terminus, instead of the non-ribosomal C- to N-direction facilitated by NCL;
- Amide bond formation would occur over a regularly-sized repeating transition state, not the growing cyclic transition state exhibited by sequential NCLs;
- The product peptide would not be restricted to contain the fingerprint Cys-Xxx-Xxx NCL catalytic unit at its C-terminus;
• The product would be readily cleavable from the machine architecture post-operation to give a peptide containing only the information embedded on the track and the nucleophile used to cleave it from the macrocycle post-operation;

• The catalysis should be more suitable to a broader range of amino acids including those not well-tolerated by NCL.

4.1.2.1 Direction of Peptide Growth

A key difference between an NCL machine and an acyl-transfer machine is that the latter would build a growing peptide chain in the natural ribosomal direction from the N- to the C-terminus[5–7] whereas the former grows peptides from the C- to the N-terminus[6,8]. As one of our overall goals is to mimic, albeit in a very primitive sense, a specific function of the ribosome, it is an attractive progression to use molecular machines to build peptides in the ribosomal orientation.

4.1.2.2 Regular-Sized Amide Bond Forming Transition State

Another advantage of the described transacylation machine is that acyl-transfer would always occur over a defined and consistent size of transition state, unlike the NCL designs where the transition state size increases upon each iterative operation cycle. When this project was conceived, significantly less was known about the limitations of NCL with regard to molecular machines, and specifically about how large an NCL trans-acyl shift transition state could be tolerated by a molecular machine without compromising its functionality. Concurrently with the development of the initial Leigh group systems, Katritzky and co-workers were investigating long-range chemical ligations with acyl-transfer from a C-terminal cysteine thioester to the N-terminus of the growing peptide chain. They demonstrated NCL occurring over 17-membered cyclic transition states (corresponding to ligation between the first [cysteine] and fifth peptide in a chain of α-amino acids) and have investigated up to 19-membered transition states using combinations of α-, β- and γ-amino acids.[9] Beyond Katritzky’s comprehensive experimental and computational studies into the relative rates of NCL versus intermolecular ligation occurring over varying sized cyclic transition states (between 5- and 19-membered),[10–13] little exploration has been done into the efficacy of ligations across larger transition states. One notable exception from Payne and co-workers does demonstrate ligation over a 29-membered transition state, however.[14]

A natural development of the Leigh group’s published systems was to design molecular machines capable of growing longer peptide chains; for this purpose, amide bond formation over a regular-sized transition state would be highly advantageous.
4.1.2.3 No Cys-Xxx-Xxx Fingerprint Region in the Product Peptide

A requirement of an NCL catalyst is the inclusion of a cysteine residue to facilitate pick-up and transfer of amino acids from the track to the growing peptide chain. The catalytic unit also requires two further amino acids (-Xxx-Xxx-NH₂) so as to bypass the unfavourable 8-membered cyclic transition state¹¹,¹² which would be encountered mid-way through the machine operation in its absence. The growing peptide is assembled onto the N-terminus of the Cys-Xxx-Xxx-NH₂ catalyst; thus, the final product has to contain this tripeptide fingerprint which would be very difficult to selectively remove. Due to the nature of the acyl-transfer catalysis via activated-ester intermediates discussed in Section 4.1.1, the product peptide can be readily cleaved from the catalyst by addition of a nucleophilic quench reagent. The product would therefore contain only the information installed on the barriers along the track and would not bare any vestigial scar-features from the molecular machine on which it was assembled.

4.1.2.4 Amino Acids Unsuitable for NCL Should be Tolerated

Not all amino acids are amenable to NCL. β-branched amino acids such as valine, isoleucine and threonine have been shown to perform intramolecular ligation at a prohibitively slower rate than comparable non-β-branched amino acids.¹⁵⁻¹⁷ Proline is also known to be problematic when present at the C-terminus of a growing peptide chain due to its propensity for side-reactions including elimination as a diketopiperizine with another amino acid from the propagating chain.¹⁵ We would anticipate an acyl-transfer catalyst to be amenable to a broader range of amino acids which would in turn grant access to more varied peptide sequences.

4.1.3 An Antibiotic End-Goal

In 1988, Baldwin and Abraham reported on the biosynthesis of penicillins and cephalosporins.¹⁸ In this seminal account, the authors discuss in particular the enzyme-mediated biosynthesis of Penicillin G (4.4, Scheme 2).
Scheme 2: The biosynthesis of penicillin G (4.4) from amino acid building blocks (4.5–4.7) via the precursors (LLD)-ACV (4.8) and isopenicillin N (4.9).

Amino acids $L$-$\alpha$-amino adipic acid (4.5), $L$-cysteine (4.6) and $D$-valine (4.7) are initially combined in a specified order by the ACV synthetase enzyme to give (LLD)-ACV (4.8). This ACV tripeptide intermediate undergoes cyclisation mediated by the isopenicillin N synthase (IPNS) enzyme in the presence of molecular oxygen and iron(II), closing to form the typical penicillin $\beta$-lactam architecture (4.9). Baldwin and Abraham go on to show that isopenicillin N (4.9) is further modified to form penicillin G (4.4).

Enzymes are naturally substrate-sensitive, however Baldwin explains that changes at the $L$-$\alpha$-amino adipic acid side of (LLD)-ACV (4.8) can be tolerated by IPNS enzyme. If $L$-$\alpha$-amino adipic acid in (LLD)-ACV is replaced with phenylacetic acid, the resulting PCV trimer 4.11 (Scheme 3) can be cyclised directly into penicillin G under the same IPNS enzymatic conditions.

Scheme 3: IPNS-mediated cyclisation of PCV (4.11) to form penicillin G (4.4) directly from a three-component building block comprised of phenylacetic acid (4.10), $L$-cysteine (4.6) and $D$-valine (4.7).

The difficulty presented by this modification is that the enzyme responsible for assembling the trimer primed for cyclisation, ACV synthetase, will not readily accept phenylacetic acid in place of $L$-$\alpha$-amino adipic acid. Thus combining phenylacetic acid, $L$-cysteine and $D$-valine in the correct PCV order must be achieved by other means. Of course one could simply employ traditional peptide couple/protecting group chemistry to form the two required amide bonds...
over a few synthetic steps, however there is also the opportunity to use a molecular machine which could carry out the synthesis autonomously in one biomimetic step.

The proposed design and mode of operation of a transacylation-based molecular machine would be an ideal architecture on which to synthesise PCV. The three building blocks (4.10, 4.6 and 4.7) would be loaded as barriers on machine 4.12 as depicted in Scheme 4.

Scheme 4: Proposed design for a molecular machine (which mimics the ACV synthetase enzyme) capable of forming PCV (4.11) and the subsequent IPNS-mediated cyclisation of 4.11 to form penicillin G (4.4) in a biomimetic synthesis combining biological and man-made molecular machines. PG = protecting group; Nu: = transacylation catalyst.

The machine would operate with the catalytic unit initially cleaving the green phenylacetic acid barrier (4.12) and transferring this onto the cysteine amine loaded onto the second barrier, thus forming the first amide bond. The second barrier would then be cleaved by the catalytic unit and deposited onto the third barrier forming the second amide bond and completing the sequence-specified combination of the three components. The catalytic unit can then remove the PCV sequence from the track and, upon hydrolysis, release the trimer (4.11) to the bulk.

The proposed use of a molecular machine to assemble 4.11 would not be possible using an established NCL-machine, as the product requires elongation to occur in the N- to C-direction. Furthermore, valine (4.7) is known to be a poor candidate for NCL due to its β-branching. The use of a transacylation molecular machine is the ideal choice by which to form 4.11 which can be transformed directly into penicillin G in a sequence which combines man-made molecular machines with Nature’s own biological machines from which our inspiration stems.
4.1.4 Finding a Suitable Transacylation Catalyst

Successfully designing a molecular machine capable of fulfilling the requirements and goals outlined above relies on first identifying a suitable catalyst. Prior to the author’s involvement in this project, numerous transacylation catalysts were investigated within the Leigh group dating back to 2007. The catalysts, investigated chronologically, include:

- 4-N,N-Dimethylaminopyridine (DMAP)-type catalysts;
- Aliphatic thiol-appended DMAP-type catalysts;
- Aliphatic thiol catalysts;
- Thiophenol catalysts;
- Hydroxybenzotriazole catalysts;
- 1,2,4-Triazole catalysts.

The range of catalysts investigated showed varying levels of success. In some cases, model studies were promising enough for molecular machines to be synthesised and operations to be tested. The outcomes of the investigations into each catalyst are summarised in the following sections.

4.1.4.1 DMAP-Type Catalysts

Initially, catalysts inspired by DMAP were investigated. For example, the macrocycle-appended catalyst 4.13 shown in Scheme 5 (which uses a previous incarnation of the macrocycle with an oxime junction) was explored.

![Scheme 5: Transacylation catalysis of a DMAP-inspired catalytic unit (4.13, black/red) going through an activated acyl-pyridinium intermediate (4.14). An acyl building block (4.15, pink) undergoes net transfer from a tyrosine residue (a barrier mimic, black) to an intermolecular nucleophile (4.16, benzylamine, green) to form an amide bond in 4.17.](image)

1 The macrocycle was changed to its current design (aldehyde macrocycle, 4.59) following development of the AMT rotaxane formation carried out previously by Leigh group members.
The catalytic unit in this system displayed poor nucleophilicity towards barrier-mimic 4.15, a phenolic ester of tyrosine, as indicated by a slow reaction rate. Additionally, acyl-pyridinium intermediate 4.14 appeared too reactive, showing side reactions (mainly hydrolysis) competing with the desired condensation with the benzylamine nucleophile (4.16). This highlights a particular balance point in operating such a molecular machine: the barrier ester must be readily cleaved by the catalytic unit and thereafter the formed activated ester intermediate should be long-lived enough to remain intact until it comes into contact with the subsequent barrier bearing the intended nucleophile. Ideally there would be a particular selectivity for the acyl group to be transferred to the desired nucleophile over competing nucleophiles in the bulk or a tyrosine phenolate (a cleaved barrier). Passing the acyl motif back onto the track is a particular concern as it could be deposited either in front of or behind the shuttling macrocycle, thus compromising the sequence-specificity of the operation.

In an attempt to activate the phenolic ester further, chlorinated tyrosine derivatives were synthesised. Although these did increase the rate of barrier pick-up, the background reaction rate was equally enhanced. A more nucleophilic catalyst with a more stable activated intermediate was required.

4.1.4.2 Aliphatic Thiol-Appended DMAP-Type Catalysts

Attempts were made to append DMAP-type catalyst with an aliphatic thiol arm. This appendage was designed to trap the acyl-pyridinium to give the more stable activated thioester. The thioester would then transfer the acyl unit to the intended nucleophile. Numerous such catalysts, condensed with a macrocycle (4.18a–d, Figure 2), were investigated in the manner outlined in Scheme 5.

![Figure 2: Aliphatic thiol-appended DMAP-type catalytic units intended to generate a more stable activated thioester intermediate.](image)

Catalysts 4.18a–c (Figure 2) did give better results than the initial DMAP-type catalyst but these improvements were matched by 4.18d, a simple aliphatic thiol with no pyridine. The use of a
DMAP-type motif in the catalyst was rendered redundant by this result as the thiol alone appeared capable of cleaving the tyrosine ester. Research into aliphatic thiol catalysts ensued.

### 4.1.4.3 Aliphatic Thiol Catalysts

Aliphatic thiols have been shown to facilitate in vivo non-ribosomal peptide synthesis by acting as acyl-carriers.\(^{[19]}\) Aliphatic thiol catalysts connected to the macrocycle via a number of different chemistries were investigated. These catalysts (4.19a–d) are shown in Figure 3 and required stoichiometric DBU (1,8-diazabicyclo[5.4.0]undec-7-ene) to form the catalytically active thiolate in situ.

![Figure 3: Aliphatic thiol catalytic units bearing a range of macrocycle connectivities.](image)

Although catalysts 4.19a–d (Figure 3) were effective in cleaving tyrosine phenolic esters, the formed thioester intermediates were more stable than anticipated. The desired thioester cleavage could only be achieved using Lewis-acid activation with Hg\(^{+}\) and Ag\(^{+}\) salts. The subsequent removal of these heavy metals was problematic, as was the susceptibility of the thiols to oxidise to their corresponding disulfides, prompting further investigation into this branch of aliphatic thiol catalysts to be halted.

It was at this point that the Leigh group moved towards cysteine-mediated NCL as a way to form peptide bonds on a rotaxane-based molecular machine. After evolving though multiple designs and carrying out multiple machine operations of varying success, the Leigh group arrived at a system which has seen three- and four-barrier molecular machines successfully operated.\(^{[1,2]}\) The NCL designs will not be discussed further in their own right in this chapter as they utilise a distinctly different form of catalysis.

### 4.1.4.4 Thiophenol Catalysts

Thiophenol catalysts were investigated as a back-up to the NCL system with a view to overcoming the aliphatic thioester stability. Thiophenolic esters were envisaged to be more labile, therefore precluding the requirement for heavy-metal Lewis-acids. Moreover, weaker
bases such as Hüning’s base (N,N-diisopropylethylamine, DIPEA) could be used to generate the thiophenolate nucleophile in place of DBU. Following promising model studies with this catalytic system, a two-barrier molecular machine (4.20, Scheme 6) was synthesised and operated. Unfortunately the operation conditions (80 °C for 18 h under microwave irradiation) proved to be too harsh for the machine and yielded mainly hydrolysed barriers and none of the intended dipeptide (4.21).

Scheme 6: Unsuccessful operation of two-barrier thiophenol machine 4.20.

The combined failure of this machine and concurrent success of the NCL system ended exploration into thiophenol catalysts.

4.1.4.5 Hydroxybenzotriazole Catalysts

Hydroxybenzotriazole (HOBt) was investigated as a transesterification-transacylation catalyst but in model studies, this required even harsher conditions than the thiophenol machine (100 °C for 18 h under microwave irradiation). Although this approach was successful in model operations and could be used to produce peptide products in ca. 65% yield, the operation conditions would be far too harsh for a molecular machine and milder conditions were unproductive. HOBt would not be a suitable catalyst.

4.1.4.6 1,2,4-Triazole Catalysts

The catalyst that was under investigation when the author joined the team working on this project was based on transacylation catalysis reported by Birman using 1,2,4-triazolium moieties as acyl carriers. Birman compared a number of acyl-transfer catalysts (including
DMAP, HOBr, benzotriazole and 1,2,3-triazole amongst others) and reported that 1,2,4-triazole (4.22) was the most effective catalyst for the acyl transfer shown in Scheme 7.

Scheme 7: Birman’s acyl-transfer test reaction.

Birman postulated that 1,2,4-triazole (4.22a) was only catalytically active as its triazolium anion (4.22a⁻) and proposed the catalytic cycle shown in Scheme 8.

Scheme 8: Birman’s suggested catalytic cycle for 1,2,4-triazolium-mediated transacylation.[20] R²O⁻ = alkoxide leaving group; Nu-H = protic nucleophile

1,2,4-Triazole (4.22a, pKₐ ≈ 10.2) is deprotonated by DBU (pKₐ ≈ 13.3) to form the catalytically active triazolium conjugate base (4.22a⁻). The nucleophilic triazolium can then cleave an unactivated ester (4.26) to form an acyl-triazolium activated ester (4.27) which will facilitate acyl-transfer. The acyl group of the activated ester is then trapped out by a nucleophile forming a new, transacylated species (4.28) and regenerating catalyst 4.22a upon reprotonation.

Background reactions in the absence of 4.22a and DBU respectively were shown to be effectively non-productive and the substitution of DBU for a weaker base, triethylamine (TEA, pKₐ ≈ 10.6), resulted in a ten-fold decrease in reaction progress over a fixed duration.

Preliminary investigations and model studies using a 1,2,4-triazole catalyst were carried out by Dr Guillaume De Bo and Dr Sonja Kuschel. The key milestones contributing to the project landscape that the author joined will be summarised in the following section. A full recount of this development can be found in Dr Sonja Kuschel’s PhD thesis.[21]

[20] Calculated values for 1,2,4-triazole and DBU pKₐ were generated using Advanced Chemistry Development (ACD/Labs) Software V11.02 (© 1994-2016 ACD/Labs)
4.1.5 Preliminary Model Study Results of 1,2,4-Triazoles

Initially, a variety of 1,2,4-triazoles (4.22a–d) were screened to assess their catalytic activity in the model reaction in Scheme 9. Concerning the substrates, valine coupling partner 4.29 was appended with a highly UV-visible fluorenyl moiety to aid chromatographic detection. It was present as its cresol (4-methylphenol) ester which mimics connection to a tyrosine residue on the track of a molecular machine. Conversion of starting materials 4.29 and 4.30 to 4.31 was monitored by a combination of reverse-phase HPLC and \( ^3 \)H-NMR spectroscopy.

Scheme 9: The model system used to investigate catalysts and conditions for use with a molecular machine. Variations on conditions investigated are summarised in the main text. Reactions were carried out at room temperature (r.t.) under an atmosphere of N\(_2\) and at \([c] = 33.3 \text{ mM}\). Substrates and catalyst were present in equimolar amounts.

The effect of catalyst functionalisation at position 3 was assessed: this would be the most convenient handle through which to attach the catalyst to the macrocycle of a rotaxane. The results showed a modest decrease in activity of catalysts 4.22b–d compared to 4.22a where conversions over 3 days dropped from 70% with 4.22a to ca. 50% with the catalysts functionalised at position 3. Notably, no product was observed in the absence of a catalyst which ruled out direct aminolysis of 4.29.

Deuterated and non-deuterated solvents including acetonitrile (MeCN), \(N,N\)-dimethylformamide (DMF) and dimethylsulfoxide (DMSO) were tested. Solvents were either used pure or in two-component mixtures. The best results were observed using neat MeCN, however this solvent alone was known from previous experience to be incapable of dissolving rotaxane-based molecular machines, so DMF was used as a co-solvent. MeCN:DMF (2:1, anhydrous) was found to be the optimal mixture.

Different bases were screened including DBU, potassium tert-butoxide, 1,8-bis(dimethylamino)napthalene (Proton-sponge\(^\text{\textregistered}\)) and DIPEA but only DBU was effective in activating the catalyst. Results were best with 1.8 eq. DBU (where 1 eq. was required to neutralise the hydrochloride salt of the leucine substrate).
Production of amide 4.31 under mild conditions using 1,2,4-triazoles functionalised in the 3-position as catalysts (4.22b–d) was achieved in ca. 50% yields with low side- and background-reactions for an intermolecular model reaction. Based on these results, a 1,2,4-triazole-bearing molecular machine was synthesised.

### 4.1.6 Operation of a One-Barrier Molecular Machine with an Intermolecular Nucleophile

Promising results were obtained when a one-barrier transacylation machine 4.32 was operated in the presence of an excess of an intermolecular nucleophile 4.33 by Dr Sonja Kuschel.

![Image of 4.32 machine with 4.33 nucleophile](image)

Scheme 10: One-barrier machine (4.32) operation with an intermolecular nucleophile (4.33) and the crude HPLC trace of the reaction mixture showing excess 4.33 (tR = 8.2 min), desired product 4.34 (13.2 min), one-barrier rotaxane molecular machine 4.32 and free macrocycle 4.35 (both at tR = 21.8 min) and operated thread 4.36 (tR = 35.4 min). Operation conditions: DBU (1.2 eq.), MeCN:DMF (2:1, 3.3 mM), r.t., 48 h, 47% conversion by HPLC.

After operation of machine 4.32 for 48 h, the crude reaction mixture was analysed by reverse-phase HPLC. The obtained trace (Scheme 10) showed that the reaction had proceeded cleanly to the desired product amide 4.34 and operated thread 4.36. The employed HPLC method was unable to separate rotaxane 4.32 and the free macrocycle 4.35 but conversion to 4.34 at 48 h
was calculated using a prepared calibration curve to be 47%. The reaction progressed to 56% conversion in 72 h where progress stalled. Notably, 4-nitrophenylacetic acid (4.37) was not detected in the HPLC trace (t_R = 4.9 min, see section 4.10.2.2) indicating no hydrolysis of the rotaxane barrier or of the acyltriazolium-activated intermediate. The production of desired product 4.34 was corroborated using LRESI(-)MS.

### 4.1.7 Operation of a Two-Barrier Molecular Machine

After the successful operation of a one-barrier transacylation machine, a two barrier machine was synthesised (4.38, Scheme 11). Operation of this machine would test the ability of the acyl transfer catalyst to not only pick up the first barrier from the track, but to also deposit this building block onto the second barrier, thereby generating the first new amide bond. Subsequently, the catalyst should be able to cleave the newly extended second barrier from the track allowing the macrocycle to dethread bearing the product. Hydrolysis of the product-macrocycle connection would occur upon quenching.

Molecular machine 4.38 was synthesised and then deprotected and operated under the conditions outlined in Scheme 11.
Scheme 11: Deprotection and operation of two-barrier molecular machine 4.36. Deprotection conditions: CH$_2$Cl$_2$:TFA (4:1), TIPS, r.t., 0.5 h. Operation conditions: DBU (1.2 eq), MeCN:DMF (2:1, 6.5 mM), r.t., 72 h, 8% conversion by HPLC.ii

Unfortunately, the success of the one-barrier machine operation was not extended to the operation of two-barrier machine 4.38. The reaction yielded many unidentified side products observed by both ESI(±)-MS and HPLC. Although the desired product was visible in both of these analyses, conversion to 4.39 was <10%. The crude HPLC trace is shown in Figure 4.

One dominant side product, which eluted at t$_R$ = 13.9 min (Figure 4), was isolated and assigned as 4.40. It was postulated that this product might have been formed by attack of a 4-nitrophenylalanine derivative on the activated ester of the desired product attached to the macrocycle prior to the quench. The 4-nitrophenylalanine derivative could either have been the second barrier of a machine which had not yet undergone complete operation, or else it may have been free in the bulk if the barrier had been hydrolysed by trace water in the system.

---

ii The conversion by HPLC in this case can be interpreted only as a rough approximation due to the presence of overlapping peaks in the HPLC trace.
Figure 4: HPLC trace of crude operation mixture of two-barrier rotaxane 4.38 showing desired product 4.39 ($t_R = 9.5\, \text{min}$), side product 4.40 ($t_R = 13.9\, \text{min}$), macrocycle 4.35 ($t_R = 21.7\, \text{min}$) and operated track 4.41 ($t_R = 32.5\, \text{min}$) along with numerous unidentified impurities.

Undesired cleavage of the terminal barrier via hydrolysis is a problem not unknown to this family of projects: experience within the Leigh group recounts numerous cases where the terminal barriers of rotaxane-based molecular machines are lost. Interestingly, it is almost exclusively the terminal barrier that is lost and not internal barriers, regardless of the total number of barriers. The terminal barrier is also known to be more labile after its deprotection, however its higher kinetic lability compared to internal barriers is something that the Leigh group has not yet been able to explain.

Although the operation of two-barrier molecular machine 4.38 was significantly less successful than the preceding one-barrier operation, some positive points could be drawn from the result. Firstly, the desired product (4.39) was formed, albeit in low quantities. Secondly, the major side product (4.40) is a close derivative of the desired product which is likely to have formed by overreaction (whilst attached to the catalyst, prior to quench) within the operation mixture. Indeed, if the macrocycle-product intermediate had not reacted further, the outcome of operation would look substantially more positive.

On these grounds it was decided that the project team would continue working towards the project goals of successfully operating a two- and three-barrier rotaxane-based molecular machine using a 1,2,4-triazole transacylation catalyst. In order to give the project the best possible chance of success, changes would clearly need to be made to some aspects of the design.

It was at this stage that the author joined and began working with the project team on a second-generation design for a transacylation-based molecular machine.
4.2 Results and Discussion: Redesigning and Synthesising a Second-Generation Transacylation Molecular Machine

After the disappointing results of the first-generation machine operations on two-barrier rotaxane systems, certain issues needed to be addressed with the machine design: these were the initial goals of the author on this project. Firstly, the apparent instability of the final barrier, especially after it is deprotected, had to be addressed. Secondly, the poorly-yielding and unreliable loading of the catalytic unit onto the rotaxane via hydrazone formation required development. The results of these investigations and the conclusions drawn are outlined in the following sections.

4.2.1 A Stoppered Terminal Barrier

The terminal amino acid building block appears to hydrolyse readily from the track, especially after its Boc deprotection. This issue has been encountered previously in the NCL peptide synthesisers, but to a seemingly lesser degree. Hydrolysis of internal barriers has not been an issue in other projects, nor has it been detected here. With a view of trying to make the terminal barrier ‘more internal’ the design of the machine was changed to a doubly stoppered rotaxane which will increase the steric bulk around the terminal barrier.

It was anticipated that the bulk of the new stopper might help to stabilise the barrier both before and after its Boc deprotection. Adding this second stopper would also mean that the macrocycle is unable to dethread from the track after the machine has operated. This change should not be problematic as the product can be readily cleaved from the catalyst and therefore removed from the machine, unlike with NCL machines where it remains bound to the Cys-Gly-Gly catalytic unit. As a result, the machine’s interlocked architecture will be preserved during the operation, which represents a step towards reloadable and reusable systems.

The modified structure of a stoppered terminal barrier is indicated in Figure 5 along with the intended disconnections required to form it from known building blocks.
Figure 5: The original design of the terminal barrier and a new, stoppered terminal barrier with disconnections numbered in the order of their intended forward-synthesis.

4.2.2 Synthesis of a Stoppered Terminal Barrier

The previously described propargyl stopper 4.46\textsuperscript{[22]} required for the terminal barrier could be accessed from readily available phenol stopper intermediate 4.45 (Scheme 12), which is synthesised on large scale according to literature procedures\textsuperscript{[23]} and available within the Leigh group on gram-scale. Williamson ether synthesis of 4.46 was carried out with quantitative yield from phenol 4.45 and propargyl bromide.

![Propargyl stopper](image)

Scheme 12: Propargyl stopper 4.46 formation from propargyl bromide and phenol stopper 4.45 which is synthesised in bulk and is available for use on multi-gram scale.

Propargyl stopper 4.46 was coupled to trimethylsilyl (TMS)-protected barrier 4.47 (for synthetic details, see Section 2.4.1.2) in a Huisgen CuAAC (copper catalysed azide-alkyne cycloadditon) reaction to afford stoppered barrier 4.48 after basic silyl protecting group removal in good overall yield (Scheme 13). \(N\)-Boc-4-nitrophenylalanine was then loaded onto the barrier under
standard EDCI·HCl (N-(3-dimethylaminopropyl)-N’-ethylcarbodiimide hydrochloride) coupling conditions with N,N-dimethylaminopyridine (DMAP) in excellent yield to afford the loaded and stoppered terminal barrier 4.49.

4.2.3 Optimising the Synthesis of Aldehyde Macrocyle 4.59

The macrocycle common to all discussed projects in this thesis was originally reported by the Leigh group.\(^1\) Within its synthesis there were key steps which were low yielding and unoptimised: addressing these bottlenecks would be advantageous to all projects as it would ease pressure on synthesis and allow for larger quantities of macrocycle to be prepared. The first building block to be assessed was the lower half (or southern hemisphere) of the macrocycle.

In the original route (Scheme 14), commercial 4-(3-bromopropyl)anisole 4.50 was metallated with zinc in the presence of iodine. The formed alkylzinc species then underwent a palladium-catalysed Negishi cross-coupling with 2,6-dibromopyridine to afford methylated southern hemisphere 4.51 in 80% combined yield. 4.51 was demethylated at −78 °C using boron...
tribromide to afford 4.52 as a difficult to handle viscous oil in 63% yield for the step and 50% combined yield over three steps.

Scheme 14: Original synthetic route to macrocycle southern hemisphere 4.52 with 50% yield over three steps.

Regarding the metallation and Negishi coupling, the reported 80% yield is excellent, however the process was practically arduous and somewhat capricious, prompting a more reliable alternative to be sought. The demethylation suffered from yield-eroding side-reactions, therefore a more controlled removal of the methyl groups was investigated. In place of the original approach, the following optimised process (Scheme 15) was developed.

Scheme 15: Optimised synthetic route to macrocycle southern hemisphere 4.54 with 82% yield over three steps.

In an associated synthesis to the original, estragole 4.53 underwent regioselective Brown hydroboration with 9-BBN (9-borabicyclo[3.3.1]nonane). The resulting borane was submitted to Suzuki-Miyaura cross-coupling conditions with 2,6-dibromopyridine to give diether 4.51 in 86% yield over two steps. The overall yield for this borylation and coupling is marginally higher than the original Negishi procedure however it is operationally more reliable and less demanding. Additionally, the estragole starting material 4.53 in the optimised route is significantly less expensive (ca. 70 times cheaper)\textsuperscript{iv} than the alkyl bromide (4.50) used in the original route.

The demethylation of diether 4.51 proceeded smoothly and cleanly in refluxing concentrated hydrobromic acid with complete conversion in 14 h. Conveniently, the hydrobromide salt of the

\textsuperscript{iv} Prices for estragole (4.53, £116/mol) and 4-(3-bromopropyl)anisole (4.50, £8362/mol) are taken from SigmaAldrich\textsuperscript{®} and are correct as of May 2016 based on products A29208-25G (98% purity) and 696137-1G (97% purity) respectively.
product (4.54) precipitated from the reaction medium upon cooling and could be isolated in excellent yield simply by filtration, precluding the need for chromatographic purification which was required after the original boron tribromide demethylation.

The overall synthetic yield of the southern hemisphere was increased from 50% over 3 steps to 82% over three steps. The optimised synthesis replaced unreliable processes with more robust and operationally simple ones, all the while using less expensive starting materials. Additionally, the formed product was more trivial to isolate and easier to handle as a solid instead of a oil.

The synthesis of the northern hemisphere of the macrocycle is shown in Scheme 16.

Scheme 16: Optimised synthetic route to the macrocycle northern hemisphere 4.57.

Sonogashira cross-coupling of 4-butyn-1-ol with 3,5-dibromobenzonitrile afforded diol 4.55 which was reduced over palladium on carbon to afford aliphatic diol 4.56. Both of these steps were carried out by Jason Hui using published procedures: both processes were robust and high yielding so were not investigated further.

The third transformation is a double Appel bromination of 4.56, where the procedure was subtly modified from the published protocol. The literature procedure used 4 eq. carbon tetrabromide and 3 eq. triphenylphosphine relative to the substrate to give 4.57 in 52% yield. By reducing the stoichiometries to 2.4 eq. carbon tetrabromide and 2.2 eq. triphenylphosphine, an optimised 81% yield was obtained on gram-scale.

With both northern (4.57) and southern (4.54) macrocycle hemispheres prepared, macrocycle formation via Williamson ether synthesis was investigated (Scheme 17). Macrocyclisation processes are carried out at high dilution (4.2 mM in this case) with a precise 1:1 ratio of each hemisphere to favour the formation of [1+1]macrocycles. The high dilution acts to increase the effective molarity of the pro-macrocycle coupling partners in solution and to disfavour the formation of common side-products: larger cyclic or linear oligomers.
Scheme 17: Formation of macrocycle 4.58 via Williamson ether synthesis at high dilution followed by nitrile reduction with DIBAL to give aldehyde macrocycle 4.59.

The previous macrocyclisation procedure used an excess of caesium carbonate (5 eq.) as base and the free base of pyridine *southern hemisphere* (4.52). Because the hydrobromide salt of this coupling partner (4.54) is used in Scheme 17, a total of 6 eq. of caesium carbonate were used in this modified process.

Where high dilution is one technique employed to favour macrocycle formation, slow parallel addition of each hemisphere to the reaction medium is another useful strategy and often, both are used concomitantly. This was the case in the literature procedure: each hemisphere was separately dissolved in DMF and it added over 20 h to a suspension of caesium carbonate in DMF at 50 °C. On complete addition, the reaction temperature was raised to 60 °C and held for a further 48 h to ensure complete reaction.

The 48 hour hold period post-addition is somewhat counter-intuitive as the point of such a slow addition is that the reaction occurs in solution at a rate faster than substrates are added. Neither substrate is therefore allowed to accumulate in solution, creating a much lower instantaneous concentration of substrates at any one time in the reaction medium compared to an *all-in* method. This should result in the reaction being complete shortly after complete substrate addition and hence further reaction time post-addition should not be required. With this in mind, the macrocycle formation was investigated under the conditions outlined in Table 1 beginning with a repeat of the literature procedure (Table 1, Entry 1).

<table>
<thead>
<tr>
<th>Entry</th>
<th>Addition time (h)</th>
<th>Addition temperature (°C)</th>
<th>Post-addition time at 60 °C (h)</th>
<th>Isolated yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>50</td>
<td>48</td>
<td>56</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>50</td>
<td>72</td>
<td>38</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>50</td>
<td>20</td>
<td>54</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>50</td>
<td>5</td>
<td>54</td>
</tr>
<tr>
<td>5</td>
<td>16.5</td>
<td>60</td>
<td>6.5</td>
<td>55</td>
</tr>
</tbody>
</table>

*Table 1: Conditions screen for formation of macrocycle 4.58.*
On repeating the literature procedure, albeit with the HBr salt 4.54 and an associated extra equivalent of base (Table 1, Entry 1), the 56% yield achieved was significantly higher than the reported 40%. The exact source of this yield enhancement is unknown. The yield achieved when the post-addition hold time was increased and decreased respectively was investigated (Entries 2 and 3). Interestingly, increasing the hold time by 50% to 72 hours (Entry 2) resulted in erosion of the yield, indicating instability of the macrocycle under the reaction conditions. This instability could be attributed to the fragility of phenolic ethers in the structure under the basic reaction conditions in the presence of bromide. Shortening the hold time by over 50% to 20 hours (Entry 3) gave virtually no change in yield and the same yield was obtained with only a 5-hour hold time (Entry 4). The preservation of the yield over decreasing post-addition hold times indicates that the reaction is indeed complete, or close to completion, on complete addition of reagents, suggesting that the slow-addition strategy is effective. A final experiment was carried out where reagents were added at a higher temperature (60 °C) over a shorter period of time (16.5 h) and held at 60 °C post-addition for only 6.5 hours, representing the shortest overall reaction: in this final case, the yield was again maintained (Entry 5). It can be concluded that slow addition of the macrocycle hemispheres is effective in promoting the desired macrocyclisation reaction. Also, a long post-addition hold time is unnecessary and can be detrimental to the reaction yield. This macrocyclisation has been carried out on a multi-gram scale; scale does not appear to affect the reaction outcome.

The final stage of the synthesis of aldehyde macrocycle 4.59 is the reduction of the benzonitrile moiety (4.58) to its corresponding benzaldehyde (Scheme 18). This transformation is achieved using disobutylaluminium hydride (DIBAL) in the published protocol to give 4.59 in 72% yield after workup and purification. This yield was said by those responsible for its synthesis to be unreliable: often yields of <30% were obtained.

Scheme 18: DIBAL reduction of nitrile macrocycle 4.58 to aldehyde macrocycle 4.59 where the remainder of the macrocycle is omitted for clarity.

The published protocol indicates the use of 2 molar equivalents of DIBAL with respect the benzonitrile starting material. This ratio is counterintuitive as the benefit of DIBAL is that it will only carry out a reduction over a single oxidation state if only one equivalent is used and
temperature control is maintained. With this in mind, the reduction of nitrile macrocycle 4.58 was carried out using the stoichiometries outlined in Table 2.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Equivalents of DIBAL (eq.)</th>
<th>Isolated Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0</td>
<td>78%</td>
</tr>
<tr>
<td>2</td>
<td>1.1</td>
<td>95%</td>
</tr>
<tr>
<td>3</td>
<td>1.1 + additional 0.3*</td>
<td>79%</td>
</tr>
</tbody>
</table>

Table 2: The observed effect of varying the number of equivalents of DIBAL on the isolated yield in the conversion of nitrile 4.58 to aldehyde 4.59. *Additional DIBAL (0.3 eq.) was added after TLC analysis indicated incomplete reaction after the same reaction time used in Entry 2 indicating a lower strength solution of DIBAL.

When the reaction was initially carried out with a stoichiometric quantity of DIBAL (1 eq., Entry 1), the reported reaction yield was marginally improved, however analysis of the crude reaction mixture by $^1$H NMR after quenching indicated unreacted starting material. Failure to reach complete conversion could be explained by one of two likely scenarios: either the reaction mixture contained moisture which partially consumed the DIBAL, or the concentration of the commercial DIBAL solution was lower than stated. To ensure that the accurate concentration of DIBAL used was known, the solution could have been titrated.\[24\]

The reaction was repeated with a slight excess of DIBAL (1.1 eq., Entry 2) which led to the yield increasing to 95%. In a subsequent repeat of the reaction, again using 1.1 eq. DIBAL (Entry 3), TLC analysis of the reaction mixture prior to quench suggested incomplete conversion, however after an additional 0.3 eq. was added, the reaction yield was eroded. It was concluded that 1.1 eq. of DIBAL was the optimal quantity; any unreacted starting material present after the quench could be reisolated by chromatographic purification for further use. This completed optimisation of the synthesis of aldehyde macrocycle 4.59 where multiple steps were improved upon in terms of both yield and practical efficiency. The result was that 4.59 was synthesised in 8 steps with 30% overall yield (5-step longest linear sequence).

### 4.2.4 Two-Barrier Rotaxane Assembly

Leigh’s rotaxane capping strategy\[2\] was again employed to assemble the two-barrier machine (4.60, Scheme 19). Stoppered barrier 4.49 was coupled to aldehyde rotaxane 4.61 (synthesised by Dr Sonja Kuschel)\[21\] under copper(I) catalysis. TentaGel®-TBTA (polymer-bound tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine) was used to ligate the copper to disfavour its undesired coordination to either substrate. An example of an undesired coordination site is the macrocycle’s pyridine moiety which is designed to bind copper to facilitate AMT\[25,26\] (active metal template) rotaxane formation. The use of TBTA in its TentaGel® polymer-bound form allows efficient removal of both copper and the ligand after the reaction which minimises work-
up processes such as aqueous washes, which could lead to hydrolysis of barriers and loss of advanced material.

Scheme 19: CuAAC rotaxane capping to form two-barrier aldehyde rotaxane 4.60 under copper(I) catalysis with a polymer resin-bound TBTA ligand.

Key considerations in this step are that phenolic ester barrier linkages must remain intact and that the isolated material is free from fully or partially hydrolysed material. The presence of these impurities can be easily detected by mass spectrometry as well as by $^1$H-NMR where a characteristic upfield shift of tyrosine aryl C-H proton environments results from their hydrolysis.

Two-barrier rotaxane 4.60 was isolated in high yield (85%) to give >100 mg of molecular machine requiring only the catalytic-unit (4.62) to be loaded via hydrazone formation before its operation could be carried out.
4.2.5 Investigation and Development of Catalytic Unit Loading via Hydrazone Formation

Loading the hydrazide catalytic unit (4.62) onto the aldehyde macrocycle of previous two-barrier rotaxane 4.63 via direct hydrazone formation (Scheme 20) had been carried out under acid catalysis with trifluoroacetic acid (TFA). The yields achieved in this step were variable and large proportions of advanced material were often lost: the loading underwent optimisation using a series of models described in the following section.

\[ \text{Scheme 20: Unreliable loading of the catalytic unit (4.62) onto a previous generation two-barrier aldehyde rotaxane 4.63 via acid-catalysed direct hydrazone formation to give first generation machine 4.38.} \]

4.2.5.1 Hydrazone Formation Using a 3,5-Dimethylbenzaldehyde Model

So as to investigate and optimise this process in the most efficient manner possible, and to preserve the limited supplies of the second generation two-barrier rotaxane (4.60), investigation began with a model system where the rotaxane was substituted for 3,5-dimethylbenzaldehyde (Scheme 21). Initially acidic and basic catalyses were compared with the catalyst-free reaction: solvent composition was investigated thereafter. Reactions were carried out using deuterated
solvents (with the exception of tetrahydrofuran, THF) and were monitored by $^1$H-NMR spectroscopy. Solubility of the highly polar hydrazide 4.62 in neat dichloromethane (CH$_2$Cl$_2$) was poor, therefore methanol was used as a polar co-solvent to ensure a homogeneous reaction mixture.

Scheme 21: 3,5-Dimethyl benzaldehyde model reaction to investigate hydrazone formation conditions for the catalytic unit.

The rate of formation of hydrazone 4.65 under acid (TFA) and base (aniline) catalysis was compared with the corresponding catalyst-free reaction in a mixture of $d_2$-dichloromethane and $d_4$-methanol. In this initial test, no product formation was observed in either the catalyst-free reaction nor in the aniline catalysed reaction: only under TFA catalysis was product formed.

With TFA identified as the catalyst of choice, optimal solvent conditions were sought. The reaction progress in mixtures of dichloromethane and methanol was slow and analysis highlighted that acid-catalysed formation of acetal 4.64 presented a competing reaction pathway which slowed the rate of hydrazone formation. Observations on this process are outlined in Table 3.
Table 3: Observations regarding the competing acetal formation during hydrazone-forming reactions in $d_2$-dichloromethane in the presence of $d_4$-methanol as co-solvent and catalytic TFA. Observations are made by analysis of $^1H$-NMR spectra.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Co-solvent</th>
<th>Reaction Time</th>
<th>Comments and Observations from $^1H$-NMR Spectra</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>$d_4$-methanol</td>
<td>18 h</td>
<td>0.8:1 of total hydrazone:acetal</td>
</tr>
<tr>
<td>1b</td>
<td>$d_4$-methanol</td>
<td>48 h</td>
<td>Reaction progresses from acetal to hydrazone</td>
</tr>
<tr>
<td>1c</td>
<td>$d_4$-methanol</td>
<td>72 h</td>
<td>1.1:1 of total hydrazone:acetal</td>
</tr>
<tr>
<td>2</td>
<td>None (then $d_4$-methanol)</td>
<td>18 h (18 h)</td>
<td>No reaction in the absence of $d_4$-methanol</td>
</tr>
</tbody>
</table>

These experiments show that the presence of TFA and methanol results in the rapid formation of dimethylacetal 4.64 from which hydrazone 4.65 still formed, but at an impractically slow rate (Entries 1a-c, Table 3). Formation of the desired hydrazone did not occur in the absence of methanol (Entry 2, Table 3), presumably due to insolubility of hydrazide 4.62 in neat dichloromethane; on addition of methanol, both acetal and hydrazone formations proceeded. These studies indicated that, while a polar co-solvent is required for the reaction to progress, methanol is not an ideal choice due to acetal formation. A substitute co-solvent was therefore investigated.

The solubility of 4.62 in various polar solvents was explored to find a non-alcoholic replacement for methanol. As the product (4.65) was soluble in dichloromethane, only partial solubility of 4.62 in the co-solvent was necessary, although complete solubility was desired. Solubility was assessed qualitatively by looking for signals corresponding to 4.62 in the $^1H$-NMR spectra in each of the solvents in Table 4.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Solvent</th>
<th>Solubility (by $^1H$-NMR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MeCN-$d_3$</td>
<td>Insoluble</td>
</tr>
<tr>
<td>2</td>
<td>MeNO$_2$-$d_3$</td>
<td>Insoluble</td>
</tr>
<tr>
<td>3</td>
<td>DMSO-$d_6$</td>
<td>Fully soluble</td>
</tr>
<tr>
<td>4</td>
<td>THF</td>
<td>Insoluble*</td>
</tr>
<tr>
<td>5</td>
<td>THF: $D_2$O (7:3)</td>
<td>Partially soluble*</td>
</tr>
<tr>
<td>6</td>
<td>$D_2$O</td>
<td>Fully soluble</td>
</tr>
</tbody>
</table>

Table 4: Observed solubility of 4.62 in polar solvents. *By TLC of filtered supernatant as non-deuterated THF was used.
4.62 was fully soluble in deuterated DMSO and deuterated water as well as partially soluble in a 7:3 THF:water mixture. The hydrazide was insoluble in acetonitrile, nitromethane and tetrahydrofuran. Further hydrazone formations were then carried out using the suitable co-solvents from Table 4 in a 1:1 ratio with CH$_2$Cl$_2$.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Co-solvent</th>
<th>Reaction Time</th>
<th>Comments and Observations from $^1$H-NMR Spectra</th>
</tr>
</thead>
</table>
| 1     | THF:D$_2$O (7:3) | 18 h | • Biphasic: only organic layer sampled for NMR  
   |               |     | • No reaction of aldehyde                      |
| 2     | D$_2$O     | 18 h | • Biphasic: only organic layer sampled for NMR  
   |               |     | • No reaction of aldehyde                      |
| 3     | DMSO-$d_6$ | 5 min | • Complete consumption of aldehyde and hydrazide  
   |               |     | • $Cis$ and $trans$ hydrazones formed          
   |               |     | • Most promising result                        |

Table 5: Attempted hydrazone formation conditions in dichloromethane in the presence of catalytic TFA using various co-solvent mixtures to ensure solubility of hydrazide 4.62. Observations are made from analysis of $^1$H-NMR spectra.

When D$_2$O and a mixture of THF and D$_2$O were used, the co-solvent was immiscible with the primary solvent, dichloromethane (Entries 1 and 2, Table 5). No reaction was observed in either case due to low concentrations of the hydrazide 4.62 in the organic phase.

DMSO-$d_6$ was miscible with the primary solvent and the result was more promising (Entry 3, Table 5); the aldehyde was completely consumed and converted cleanly to the desired hydrazone (4.65) in 18 hours. On setting up the reaction, the aldehyde was dissolved in a suspension of the partially soluble hydrazide in CD$_2$Cl$_2$ and DMSO-$d_6$. On addition of a solution of TFA in deuterated dichloromethane, the reaction mixture appeared initially unchanged but became clear (monophasic) after <5 mins at which point the conversion was complete by $^1$H-NMR. Both the $cis$- and $trans$-hydrazone isomers of 4.65 were formed in a 1:1 ratio. 4.65 could be precipitated by adding water to the DMSO solution after removal of dichloromethane in vacuo. $^1$H-NMR and mass spectrometry analyses encouraged the use of these conditions in forming the macrocycle hydrazone.
4.2.5.2 Hydrazone Formation on the Aldehyde Macrocycle

Optimal conditions from the model reaction with 3,5-dimethylbenzaldehyde were applied to the loading of aldehyde macrocycle 4.59.

![Scheme 22: Test hydrazone loading on aldehyde macrocycle 4.59.]

The results were initially promising with the reaction proceeding as anticipated with regards the previous model, however the crude product was not as easily precipitated from the reaction mixture. In order to simplify product isolation, the amount of DMSO used was reduced by half to aid its removal from the crude reaction mixture in vacuo. The reaction continued to proceed smoothly with less DMSO, albeit more slowly (3 hours as opposed to minutes for complete conversion). Progress of the hydrazone formation was judged by crude $^1$H-NMR analysis: 4.66 was not isolated due to poor solubility of the polar catalytic unit in organic solvents. This subtly modified loading procedure was deemed suitable to test on the newly-designed two-barrier rotaxane 4.60.

4.2.5.3 Hydrazone Formation on the Doubly-Stoppered Two-Barrier Aldehyde Rotaxane 4.60

Loading of the two-barrier machine was initially carried out in a test-scale reaction (5 mg, 1.6 µmol). The reaction was worked up after 5 h (a sufficient duration on all model systems) but preliminary analysis of the crude reaction mixture by ESI-(-)MS and TLC showed the reaction to be incomplete. The crude mixture was resubmitted to reaction conditions and gave spot-to-spot conversion to 4.67 by TLC in a total of 20 h reaction time. This decrease in reaction rate can most likely be attributed to the lower reaction concentration; due to constraints of scale, the reaction could not be carried out at the precise concentration established during optimisation.
Scheme 23: Hydrazone loading of catalytic unit 4.62 onto two-barrier aldehyde rotaxane 4.60 to give two-barrier, doubly-stoppered machine 4.67, ready for deprotection and operation.

The $^1$H-NMR spectrum of the product after aqueous extraction and preparative-scale (prep.) TLC purification showed 4.67 as the major species but also indicated the presence of an impurity resulting from hydrolysis of the second (4-nitropheyl[N-Boc]alanine, pink) barrier. Based on suspicions that this hydrolysis had occurred during workup, aqueous washes were avoided in the workup when the reaction was repeated on larger scale (25 mg, 7.7 µmol, Scheme 23) which yielded pure 4.67 in 62% isolated yield (16 mg, 4.8 µmol) with no indication of barrier hydrolysis.
4.3 Operation of the Redesigned Second-Generation Molecular Machine

Following the successful isolation of two-barrier molecular machine 4.67, attention turned to its deprotection and operation. Removal of Boc protecting groups from amino acid barriers is generally carried out immediately before operating as the deprotection renders the machine less stable. The deprotection and operation of the second-generation molecular machine is summarised in the following sections.

4.3.1 Two-Barrier Molecular Machine Deprotection

The deprotection of machine 4.67 involves removal of the Boc group from the second barrier amino acid. This transformation is carried out under acidic conditions using TFA (Scheme 24).

![Scheme 24: Boc deprotection of machine 4.67 in the presence of TFA to give 4.68, the TFA salt of the deprotected machine.](image)

The deprotection of 4.67 proceeded smoothly to complete conversion in 0.5 h (monitored by the disappearance of Boc 'Bu signals in its $^1$H-NMR spectrum). The workup procedure involved diluting the reaction mixture by half with toluene and subsequently concentrating in vacuo at 40 °C to dryness. The addition of toluene ensures that the concentration of TFA is not increased with respect its reaction concentration as the more volatile CH$_2$Cl$_2$ evaporates before the residual TFA. The residue was then dissolved in toluene and concentrated twice more to ensure complete removal of excess TFA; the only TFA remaining in the deprotected crude material should therefore be present as a salt of the machine. The crude mixture was analysed by ESI-(±)-MS to check for complete deprotection and to determine whether any barrier hydrolysis had occurred prior to operation.
4.3.2 A Two-Barrier Machine Operation with DBU

The freshly deprotected molecular machine 4.68 was subject to operation conditions with DBU on a scale of 0.7 µmol (Scheme 25).

![Diagram of molecular machine](image)

Scheme 25: Operation of two-barrier molecular machine 4.68 with 2.0 eq. DBU.

On addition of a standard solution of DBU in acetonitrile to the substrate in DMF, the reaction mixture immediately turned from light yellow to an intense pink-purple colour. The operation was stirred at room temperature for 25 h at which point it was quenched by the addition of water. Upon quenching, the intense pink-purple reaction colour reverted to yellow and the formation of a white precipitate turned the reaction mixture opaque. The operation mixture was filtered over a 0.45 µm pore-size PTFE membrane to give a clear solution. The precipitate removed by the membrane filtration was shown by ESI(−)-MS to be rotaxane with no intact barriers (operated rotaxane) and derivatives thereof, although this material was never purified or analysed further.

The filtered solution of crude operation mixture was analysed by reverse-phase HPLC: the resulting spectrum is shown in Figure 6.
Figure 6: HPLC trace of machine operation using 2 eq. DBU showing the peak corresponding to desired product 4.39 ($t_R = 9.4 \text{ min}$) and numerous other products. Note that the conversion to 4.39 can only be viewed as a rough approximation due to the presence of overlapping peaks on either side of the product peak.

Although a peak was visible at $t_R = 9.4 \text{ min}$ where the desired product 4.39 is expected, its integration is complicated by overlapping peaks on either side. An integration of ca. 9% was approximated with reference to a calibration standard. Alongside the disappointingly low conversion, the operation HPLC trace shows that many side products have formed in significant quantities, despite the fact that the operated two-barrier rotaxane co-product 4.69 is removed from the operation mixture by the precipitation and filtration process prior to analysis.

Investigation into the rich impurity profile was continued with the aid of LC-MS, which confirmed the presence of 4.39 at $t_R = 9.3 \text{ min}$. Disappointingly, many of the other peaks indicated degradation of the machine’s rotaxane architecture as a whole. For example, free macrocycle 4.66 was detected, but this should remain threaded on the doubly-stoppered rotaxane thread. Various cleaved stopper species were also detected explaining the presence of non-interlocked macrocycle.
### 4.3.3 Investigating Triethylamine as a Base to Activate the 1,2,4-Triazole Catalyst

From the results of this machine operation, concern arose over the strength of the base, DBU, that was being used to deprotonate the catalytic unit. In particular, concern over how DBU may interfere with other sensitive sites on the machine prompted investigation into alternative bases. DBU has a conjugate acid pKₐ of ca. 13.3 and 1,2,4-triazole has a pKₐ typically ca. 10.2 (although the pKₐ of the acyl triazole catalyst may be expected to be slightly lower). Encouraged by Birman’s observation[^20] that triethylamine (conjugate acid pKₐ ca. 10.6), a weaker base than DBU, could also activate 1,2,4-triazole to catalyse the desired transacylation (albeit at a slower rate) model studies were conceived to test triethylamine’s efficacy. Birman reports that triethylamine is able to activate the catalyst to some degree but the outcomes are poor, although only one equivalent of triethylamine was reported to have been used. It seemed logical that an excess of triethylamine would accelerate this rate.

The employed model reaction is indicated in Scheme 26 and was carried out by Dr Sonja Kuschel. Cresol ester 4.70 was used as a mimic of the first barrier and the hydrochloride salt of (s)-4-nitrophenylalanine methyl ester (4.71) as a mimic of the second barrier. The substrates and catalyst (4.72) were present in equimolar quantities. As the nucleophilic amine was present as a hydrochloride salt (4.71), one equivalent of triethylamine would be consumed in neutralising it.

![Scheme 26: Model reaction using one equivalent each of 4.71, 4.70 and model catalyst 4.72 with NEt₃ as base.](image)

The model studies showed that a small excess of triethylamine (2.5 eq.) was able to produce a modest amount of product 4.34 (20% isolated yield) but pleasingly the yield increased to 50% when a larger excess of triethylamine was used (10 eq.). Although the reactions did not reach completion in the 18 h time period, it was encouraging to see that their progress had been almost entirely free of side reactions. Following on from this model study, the operation of two barrier machine 4.68 was attempted with an excess of triethylamine in place of DBU.
4.3.4 A Two-Barrier Machine Operation with Triethylamine

Initially, a batch of two-barrier molecular machine (4.67, 2.2 µmol) was deprotected according to the protocol outlined in Section 4.3.1 to give 4.68. This machine was then operated using a standard solution of the triethylamine base in acetonitrile (Scheme 27).

Scheme 27: Operation of two-barrier molecular machine 4.68 with 2 additions of 10 eq. triethylamine.

After 24 h operation time, the reaction mixture was sampled and analysed. Although the analysis showed a relatively clean reaction mixture, no desired product (4.39) had formed. A further 10 eq. of triethylamine was added and the operation allowed to proceed for a further 72 h. Upon subsequent sampling after a total of 96 h reaction time, the obtained HPLC trace (Figure 7) was generally clean showing one major peak (t_R = 11.5 min) with high UV absorption and λ_max at 272–274 nm consistent with nitro-aryl species.\(^v\)

Figure 7: HPLC trace of a 96 h time point of the operation of two-barrier molecular machine 4.68 using a total of 20 eq. NEt$_3$. Scale = 2.2 µmol (7.5 mg); conc. = 4.88 mM.

The retention time of this new species ($t_R = 11.5$ min) was not in agreement with the desired product, 4.39 ($t_R = 9.5$ min): the reaction mixture was therefore analysed further by LC-MS. The mass spectrum obtained for the peak eluting at ($t_R = 11.5$ min) indicated the formation of a 2,5-diketopiperazine (DKP) derivative of two units of 4-nitrophenylalanine (4.74, Figure 8).
Although all of the major peaks observed in this mass spectrum were assigned to adducts of the suspected DKP product 4.74 (or adducts of TFA): further proof was obtained by carrying out an unambiguous synthesis of 4.74.

DKP 4.74 was synthesised using a modified procedure from Govindaraju\textsuperscript{[27]} outlined in Scheme 28. Fmoc-4-nitro-\textit{l}-phenylalanine (4.75) and 4-nitro-\textit{l}-phenylalanine methyl ester hydrochloride (4.71) were coupled under standard EDCI/HOBt coupling conditions to give dipeptide 4.76 in very good yield. Dipeptide 4.76 was subject to Fmoc (fluorenylmethyloxycarbonyl) deprotection conditions with diethylamine. Upon Fmoc removal, the dipeptide spontaneously cyclised and, through liberation of methoxide, formed DKP 4.74.
Scheme 28: Synthesis of DKP 4.74 from an N- and an O- protected 4-nitrophenylalanine.

The unambiguously prepared DKP 4.74 was shown to have a retention time consistent with the major operation product. It is postulated that 4.74 must be formed by an intermachina dimerisation of the second barrier (4.77), most likely through the activation of one barrier by the catalytic unit (4.78, Scheme 29).

Scheme 29: Postulated mechanism of formation of DKP 4.74 via an intermachina dipeptide formation and rapid cyclisation. MAC = macrocycle; TRACK = connection to the track via tyrosine α-carbon.

Alternatively, 4.74 may be formed between two track-loaded barriers (4.77, phenolic-type esters of the barrier on tyrosine’s side chain), however this is less likely as these esters are less activated. Cyclisation to form the 6-membered DKP ring is known to proceed very efficiently with little required activation of the electrophilic carbonyl. [28]

The operations of the molecular machines reported above were largely unsuccessful. Although some evidence indicated the generation of the intended product of operation (4.39) when DBU and triethylamine were used to activate the catalyst, a large number of undesired side-products were formed. The side-product profile suggested that the molecular machines did not operate with the intended level of control and that further investigations and modifications would be required if a successful operation were to be realised.
Chapter 4

4.4 Trouble-Shooting Unsuccessful Machine Operations

In order to gain control of future operations, the means by which side-products were formed, and control was lost, had to be probed. What was most concerning about the operation described in Section 4.3.4 was that no derivatives of the first barrier could be detected in the operation mixture. In order for the system to form DKP 4.74, the first barrier must somehow be lost from the track to allow the macrocycle and catalytic unit to reach and react with the second barrier, thereby losing the sequentiality of the molecular machine operation. The stability and possible fate of the first barrier was therefore investigated. The results of this investigation are summarised in the following sections.

4.4.1 Investigation into the Fate of the First Barrier in a Two-Barrier Machine Operation with Triethylamine

Concern was raised over the acidity of the methylene protons on the first barrier (4.80, Scheme 30) due to its highly activating α-neighbouring groups: a carbonyl and a 4-nitrobenzene moiety. It was proposed that under the basic operation conditions, the 4-nitrophenylacetic acid barrier may be eliminating from the molecular machine as a ketene (4.82). This could either occur when the unit is connected to the track in its starting position or when it is picked up and further activated by the catalytic unit (Scheme 30).

Scheme 30: Proposed mechanism for the base-induced loss of the first barrier as a ketene 4.82. Mac = macrocycle; Track = connection to the track via tyrosine α-carbon.

The eliminated ketene (4.82) would then be expected to react rapidly with any nucleophile present; this could result in reattachment to the track or catalytic unit, coupling with the primary amine of the second barrier amino acid, or reaction with an external nucleophile. Each of these possibilities would result in a loss of sequentiality of the product through lack of processivity within the machine operation.

Critically, such degradation could have even been occurring unbeknown in model reactions, as nucleophilic attack of an intermolecular nucleophile present in excess into ketene 4.82 would
still give the desired product, only through a different reaction pathway (for example see Scheme 10, Section 4.1.6).

Another observation which may support this hypothesis is the intense colour change that reaction mixtures exhibit in both model operations and machine operations bearing phenolic-esters of 4-nitrophenylacetic acid. Upon the addition of base, the reaction mixtures invariably turn a vibrant pink-purple colour. This could indicate the presence of the conjugated enolate intermediate 4.81 and furthermore, when a solution of the cresol ester of 4-nitrophenylacetic acid (4.70) is treated with triethylamine alone, the described colour change occurs, suggesting that it is this component that is responsible.

### 4.4.2 Stability Studies on a 4-Nitrophenylacetic Acid Barrier Mimic

With this information in mind, stability studies on a model cresol ester of the first barrier (4.70) were carried out using $^1$H-NMR and HPLC to monitor any degradation. Model barrier 4.70 was shown to be stable under Boc-deprotection conditions, however degradation of the cresol ester was observed under operation conditions in the absence of the catalytic unit (Scheme 31). The degradation product was characterised as diethyl amide 4.83 by comparison of its NMR, MS and HPLC data with an unambiguously prepared reference sample.

![Scheme 31: The proposed formation of diethylamide 4.83 from cresol ester 4.70 and excess triethylamine via ketene intermediate 4.82.](image)

The reaction of ketene species with tertiary amines to generate tertiary amides is described by Tidwell. Triethylamine is able to generate the enolate of 4.70 which subsequently eliminates cresol to form ketene 4.82. The highly electrophilic ketene is subject to attack from the weakly nucleophilic triethylamine to generate zwitterionic intermediate 4.84. Amide 4.83 is eliminated upon removal of an ethyl group from the initial triethylamine by another equivalent of the base.
Ketenes are able to undergo thermal [2+2] cycloaddition reactions with alkenes.\[^{30}\] In an attempt to confirm the presence of ketene intermediates, a ketene trap experiment was carried out (Scheme 32).

Scheme 32: Failed attempted trapping of ketene intermediate 4.82 with cyclopentene via a thermal [2+2] cycloaddition.

Cresol ester 4.70 was mixed under operation conditions with an equivalent each of model catalyst 4.72 and cyclopentene. Over the course of 72 h, no cycloaddition product (4.85) was detected by $^1$H-NMR, although cresol formation and formation of diethylamide 4.83 was observed. The lack of formation of cyclobutanone 4.85 could be due to the low reactivity of the ketene with cyclopentene at ambient temperature. Therefore as trapping of the ketene with an alkene appeared unfavourable, the ketene was slowly dragged towards the formation of diethylamide 4.83. It is noteworthy that the actual concentration of ketene in solution at any one time is likely to be very low based on the equilibria present. In the absence of a formal nucleophile, the ketene is likely to recombine with cresol to reform the starting material (4.70) in this study.

### 4.4.3 Preventing the Formation of Ketenes and Associated Degradation Products

Without conclusive evidence of their formation, steps were taken to minimise the possibility of losing the first barrier as a ketene. A number of possible alterations could be explored. Substitution of the two $\alpha$-protons with methyl groups may be effective but the introduction of a geminal dimethyl moiety alpha to the carbonyl would drastically increase the steric bulk around the electrophile and therefore slow the rate of barrier pick-up by the catalyst. Extending the alkyl chain between the carbonyl and aryl ring with an extra methylene would break conjugation between the electron deficient ring and the enol of the carbonyl, thereby increasing the $pK_a$ of the protons, but this would also make the barrier more flexible. Making the barrier more flexible is undesirable as it increases the likelihood of the macrocycle being able to slip over the
loaded barrier. Somewhat counter intuitively, making a barrier larger does not necessarily make it a more bulky or larger kinetic barrier. Instead, adding a methylene unit would give the barrier more flexibility. It therefore possesses more degrees of freedom and thus more chance of folding into an orientation with the track where the macrocycle can slip over it.

Removal of the para-nitro group from the barrier to give phenylacetic acid (4.10) as the first building block was instead investigated. The function of the nitro group was to act as a fluorescent handle for analysis and seeing as the second barrier also bears a nitro group, the product of the operation should still be readily detectable by HPLC. An added attraction of using phenylacetic acid (4.10) as the first barrier is that it is the required first barrier for a molecular machine capable of synthesising PCV (4.11), a precursor to penicillin G (4.4, see Section 4.1.3).

### 4.4.4 Investigating Phenylacetic Acid as a First-Barrier Building Block

Before a change of building block loaded onto the first barrier was initiated, studies into the comparative stability to side reactions of 4-nitrophenylacetic and phenylacetic model barriers (cresol esters 4.70 and 4.86) were carried out (Scheme 33). Non-anhydrous deuterated solvents were used whereby trace amounts of residual water would indicate the comparative propensity of each ester to hydrolyse. Reactions were monitored by $^1$H NMR.

![Scheme 33: Stability studies of cresol esters 4.70 and 4.86 under operation conditions with a model catalyst 4.65 in the absence of any formal nucleophile.](image)

Pleasingly, the results of this study supported our hypothesis that stability of the barrier would be increased on removal of the nitro group. The des-nitro ester (4.86) was more stable to hydrolysis than nitro species 4.70 as indicated by the formation of only minimal cresol (4.88, ca. 7% by $^1$H-NMR). The nitro species exhibited ca. 40% conversion to cresol. In addition, 4.86 showed a significantly reduced conversion (<3% conversion) to its diethylamide counterpart 4.87 compared to 4.70 which formed 4.83 with ca. 14% conversion. An additional observation was that the des-nitro substrate produced far fewer unknown side-products over the course of the study compared to the nitro substrate. This suggested that by changing to phenylacetic acid
as the building block to load onto the first barrier, an increased level of control would be gained.

Critically, CPK modelling confirmed that replacing the first cleavable barrier unit with its des-nitro analogue would result in a barrier that was still large enough to prevent the macrocycle from passing over it before it is removed.

The combination of these observations encouraged the replacement of 4-nitrophenylacetic acid with phenylacetic acid as the first barrier of a two-barrier molecular machine; this change would require resynthesis of a third-generation of the machine from advanced intermediates.
4.5 Redesigning and Synthesising a Third-Generation Transacylation Molecular Machine

The following sections discuss the changes implemented in the third-generation design involving adapting the building block on the first barrier and changing the connectivity between the macrocycle and the catalytic unit. The results of the AMT rotaxane formation of these altered components and the rotaxane’s subsequent capping are also discussed herein.

### 4.5.1 Reloading the First Barrier

Reloading the first barrier with phenylacetic acid could be conveniently realised by esterifying unloaded barrier 4.89 with phenylacetyl chloride (4.90) in the presence of triethylamine (Scheme 34).

![Scheme 34: Acylation-loading of unloaded barrier 4.89 with phenylacetyl chloride to yield phenylacetic acid-loaded barrier 4.91.](image)

The reaction yield was lower than desired (65%) due to problematic chromatographic separation and removal of an unidentified by-product. However, with a sacrifice in yield, 4.91 could be isolated in high purity ready to be used in an AMT rotaxane formation.

### 4.5.2 An Amide Macrocycle: Changed Connectivity between the Macrocycle and Catalyst

Although loading of the catalytic unit via hydrazone formation had been investigated and optimised, it remained a capricious step with a risk of material loss. The decision was therefore taken to investigate the viability of a permanent amide link between the macrocycle and the 1,2,4-triazole catalytic unit.

Previous experience from within the Leigh group dating back to the development of the first molecular machine capable of sequence-specific peptide synthesis[1] showed that the AMT rotaxane assembly of such machines was sensitive to the synthetic handle installed on the macrocycle through which the catalytic unit was attached.
It was shown that having the Cys(Trt)-Gly-Gly-NHBoc NCL catalytic unit attached to the macrocycle via an amide bond at cysteine resulted in very poor conversions to the interlocked rotaxane. At the time, it was believed that having the catalyst attached to the macrocycle by a benzylic amide bond on the tripeptide’s O-terminus shut down the AMT templation. The tripeptide unit was thought to be competing with the pyridine ligand in the macrocycle cavity for the ligation of copper, explaining the low yields of the interlocked architecture. For this reason, an acyl hydrazone was used to join the catalytic unit to the molecular machine once the interlocked structure had been assembled, where the AMT rotaxane formation proceeded efficiently with aldehyde macrocycle 4.59. A benefit of this connection strategy is that the dynamic covalent bond could foreseeably be cleaved after machine operation to detach the product of operation from the macrocycle.

Considering the differences between the transacylation and NCL catalytic units, it was proposed that the 1,2,4-triazole catalytic unit may behave differently from the tripeptide in the AMT rotaxane formation. Furthermore, the different nature of the transacylation mechanism of operation precludes the requirement for a cleavable junction between the catalyst and the product for release of the peptide post-operation. Thus, an investigation began into a rotaxane with the catalytic unit attached to the macrocycle by a permanent amide bond with the assumption that such a change would not affect the triazole’s catalytic activity. Rotaxane 4.92 (Figure 9) was the third-generation molecular machine target.

![Figure 9: The target structure of a third-generation molecular machine (4.92) capable of sequence-specific peptide synthesis via a 1,2,4-triazole catalyst. The design has a benzylic amide junction between the macrocycle and catalytic unit, the first barrier has been changed to phenylacetic acid and both ends of the rotaxane are stoppered. The 1,2,4-triazole may require a protecting group (R = PG) if its nucleophilicity is deemed problematic in AMT rotaxane formation.](image-url)

The synthesis of amide macrocycle 4.93 (Scheme 35) was carried out to allow testing of its templating efficacy in an AMT rotaxane formation. β-Alanine methyl ester hydrochloride was coupled to 1,2,4-triazole-3-carboxylic acid to give methyl ester 4.72. This coupling was attempted with milder coupling conditions (such as EDCI/HOBt) but these did not yield 4.72,
thus the more aggressive PyBOP ((benzotriazol-1-xyloxy)tripyrrolidinophosphonium hexafluorophosphate) coupling reagent was used and an acceptable yield obtained.

![Scheme 35: The synthesis of a modified catalytic unit and its connection to amine macrocycle 4.97 accessed from nitrile macrocycle 4.58.]

With concern that the unprotected 1,2,4-triazole might compromise the coupling of the carboxylate of methyl ester 4.72 and amine 4.97, and that it may subsequently compromise the AMT CuAAC rotaxane formation, the triazole was protected with a trityl (triphenylmethyl) protecting group (PG) using a modified literature procedure.\[^{[31]}\] Trityl protection of the catalyst proceeded smoothly to give 4.95 in excellent yield.

Multiple efforts of hydrolysing 4.72 to its corresponding carboxylic acid prior to the decision to add a trityl protecting group had failed. The saponification would progress but the product was insoluble in the majority of organic solvents and was difficult to purify. The presence of the trityl protecting group simplified the saponification of ester 4.95 to acid 4.96 as the solubility of the resulting carboxylic acid in organic solvents was greatly improved. 4.96 was thereby isolated in excellent yield.

Where aldehyde macrocycle 4.59 had been formed via a DIBAL reduction of nitrile macrocycle 4.58 (see Section 4.2.3), benzylic amine macrocycle 4.97 could be accessed using lithium aluminium hydride as reductant. Interestingly, the amine product underwent partial decomposition during chromatographic purification on silica, but by avoiding flash column
chromatography and purifying by other means, 4.97 could be isolated pure with near quantitative yield.

Pleasingly, after making efforts to trityl-protect the catalytic unit, 4.96 was coupled under mild EDCI/HOBt coupling conditions to the benzylic amine macrocycle 4.97 in near quantitative yield to give the target amide macrocycle 4.93.

Before attempting AMT rotaxane formation with amide macrocycle 4.93, a trial deprotection was carried out (Scheme 36) to simulate its proposed deprotection as part of a global deprotection (removing Boc and trityl groups) prior to operation of the molecular machine.

The deprotection using TFA to cleave the trityl group as its carbocation and TIPS (1.1 eq.) to reduce this carbocation to triphenylmethane was carried out in CH$_2$Cl$_2$ at room temperature over 4.5 h. The deprotection proceeded cleanly and, after azeotropic TFA removal with anhydrous toluene, and trituration of the residue with a diethyl ether/pet. ether mixture, gave the trityl-deprotected amide macrocycle 4.98 with excess TIPS and the formed triphenylmethane removed. This is the deprotection strategy that would be used prior to machine operation to cleave the trityl group from the catalyst and also the Boc group on the second barrier’s amino acid.

**4.5.3 AMT Rotaxane Formation with an Amide Macrocycle**

Confident that the trityl group could be efficiently removed from the molecular machine prior to operation, CuAAC rotaxane formation of the desired interlocked structure was investigated (Scheme 37).
Scheme 37: AMT rotaxane formation using azide stopper 4.99 (6 eq.), amide macrocycle 4.93 (1.5 eq.) and phenylacetic acid-loaded barrier 4.91 (1 eq.) to give rotaxane 4.100 in 68% isolated yield.

Phenylacetic acid-loaded barrier 4.91 was reacted with an excess of azide stopper 4.99 (6 eq.) and a small excess of amide macrocycle 4.93 (1.5 eq.) with a copper(II) catalyst (50 mol%) for 22 h at room temperature. Fortunately the change in connectivity between the catalytic unit and the macrocycle appeared to have a positive effect on the ability of the macrocycle to template AMT rotaxane formation as rotaxane 4.100 was isolated in a 68% yield: a sizable increase on the commonly achieved average yield of ca. 40% for such a rotaxane formation with aldehyde macrocycle 4.59. Other commonly produced side-products from such AMT rotaxane formations are free thread (4.101) and over-reacted rotaxanes with more than one barrier. Free thread (4.101) forms when the stopper and barrier couple outside of the macrocycle cavity and thus form non-interlocked species. Over-reacted rotaxanes form from the reaction of one barrier’s aryl azide with the alkyne of another barrier, either before or after the barrier is incorporated into a rotaxane. Free thread and over-reacted rotaxanes were isolated in 5% and 12% respectively, however the 12% of over-reacted rotaxanes was not characterised. Additionally, up to 98% of the excess of 4.99 is recovered and the entire excess of 4.93 is reclaimed.
Even more pleasing than the efficient conversion to interlocked species 4.100 was that the rotaxane purification transpired to be significantly simpler than with previous rotaxanes. In general, the chromatographic characteristics of a rotaxane will be governed either by the polarity characteristics of its constituent thread or by the macrocycle. The rotaxane will therefore be difficult to separate chromatographically from whichever of the free thread or macrocycle has the dominant polarity. These tedious separations are usually achieved by firstly removing all other by-products or excess reagents by flash column chromatography on silica to give a mixture of the rotaxane and its close-running impurity and then subsequently purifying by prep. TLC. This preparative-scale purification represents a synthetic bottleneck because of the low quantities of crude material that can be purified at any one time. Fortunately rotaxane 4.100 could be readily separated from both the free thread (4.101, lower polarity) and the excess macrocycle (4.93, higher polarity) by flash column chromatography on silica to give pure 4.100 without the need for scale-limiting prep. TLC.

The success of the rotaxane formation using amide macrocycle 4.93 paired with its seemingly trivial deprotection cemented the decision to move forward with this third-generation design.

4.5.4 Proof of Concept: Proving the Catalytic Mode of Operation

One fragile aspect of this project has been that few control studies have been carried out on model systems to investigate rates of potential background reactions, or to prove that the catalytic unit itself is responsible for the bond-forming reactions. One benefit of having a protecting group on the catalytic unit is that it allows for control experiments to be performed which can highlight the critical role of the catalyst in the bond-forming pathway. Previously when model studies had been carried out, alternative reaction pathways could not always be conclusively ruled out.

As such, it was decided to carry out contrasting one-barrier rotaxane operations to compare the outcomes when the catalyst was protected and deprotected (Scheme 38). This investigation was carried out in an NMR tube using periodic $^1$H-NMR monitoring. Secondary amine diethylamine (HNEt$_2$) was chosen to act as both the base for catalyst deprotonation and as the external nucleophile for amide bond formation due to its analogous pK$_a$ to triethylamine and the characteristic $^1$H-NMR profiles of both it and its amide derivative 4.87.
Scheme 38: Controlled model operations of a one-barrier rotaxane with an external secondary amine nucleophile (HNEt₂), which also acts as a base for activation of the 1,2,4-triazole catalytic unit, were carried out in duplicate. The catalytic unit is protected and therefore inactive in one case (the control, using 4.100) and is deprotected and active in the other (using 4.102). Chemical environments that are labelled and highlighted by colour indicate characteristic proton environments in the operation ¹H-NMR spectra (see Figure 10).

In line with the hypothesised mode of operation, the rotaxane remained unchanged by ¹H-NMR in the control experiment (R = Trt, 4.100) and no amide product was generated above trace levels. Furthermore, upon operation of the deprotected rotaxane (R = H, 4.102) clean conversion to the desired diethyl amide product 4.87, operated thread 4.103 and free macrocycle 4.98 was observed.

Signals characteristic of deprotected rotaxane 4.102 shown in red (C and D, Scheme 38) decrease over time. Operated thread (4.103) signals increase over time (A and E, Scheme 38), as do signals corresponding to the uninterlocked amide macrocycle 4.98 (B and F, Scheme 38) indicating removal of the barrier and dethreading of the rotaxane. Catalytic unit signal ‘H’ sharpens as the reaction proceeds: H is broad in the ¹H-NMR of rotaxanes 4.102 and 4.100 but is
sharp in the spectrum of free amide macrocycle 4.98. Most importantly, the generation of peaks G and I, corresponding to the desired product 4.87, are observed to increase over time. The signals are confirmed to be those of 4.87 by doping a small sample of this product into the NMR tube post-operation resulting in perfect signal overlap. At the final 74.5 h time point, conversion to 4.87 was approximated to 75%.

Figure 10: Periodic $^1$H-NMR analysis of a deprotected one-barrier machine operation using HNEt$_2$ as base and nucleophile. Labels refer to the environments highlighted in Scheme 38. Red arrows indicate a decrease in the signals corresponding to the starting material (4.102) and blue, purple and green arrows indicate growth of the signals corresponding to products (4.87, 4.98 and 4.103) over time. Signal ‘H’ is shown to sharpen as the reaction proceeds, consistent with the dethreading of the macrocycle.

This study provided the first conclusive evidence that the deprotected 1,2,4-triazole catalytic unit was responsible for the formation of amide bonds. Additionally, the rate of the background reaction has been shown by the control to be negligible over the 74.5 h operation duration where 75% conversion to product was achieved in the model operation. With renewed confidence in the design and concept, efforts progressed towards the formation of a two-barrier machine.
4.5.5 Assembling the Two-Barrier Rotaxane by Appending the Tail-Piece to the One-Barrier Rotaxane

In order to assemble the two-barrier rotaxane, the rotaxane capping strategy was employed as before (see Section 4.2.4) whereby rotaxane 4.100 is coupled to tail-piece 4.49 using a CuAAC reaction (Scheme 39).

![Scheme 39: Synthesis of two-barrier rotaxane 4.104 using the rotaxane capping strategy of coupling one-barrier rotaxane 4.100 and tail-piece 4.49.](image)

Two-barrier rotaxane 4.104 was isolated in high yield and pleasingly, the isolated product showed no indication of barrier hydrolysis by either ESI-(±)-MS (see Section 4.10.4.1) or 1H-NMR, suggesting that this third-generation design may have a lower propensity for undesired barrier hydrolysis.

4.5.6 Analysis of Deprotection Conditions

Further to investigating the deprotection of the catalytic unit described in Section 4.5.2, the successful conditions were used on a sample of one-barrier rotaxane 4.100. The exact deprotection conditions were repeated on 4.100 but on analysis after 4 h reaction time, the trityl deprotection was incomplete. A possible explanation is that the very small quantities of TIPS (1.1 eq., 1.5 µmol, 0.3 µL) used were not measured with sufficient accuracy, even though
stock solutions were used. This deprotection was repeated using a large excess of TIPS (1 drop, ca. 10 µL, ca. 35 eq.) which, over the same time period, showed complete deprotection of the trityl group but worryingly it also showed complete hydrolysis of the phenolic ester barrier. By MS, the only species visible was a perched/pseudo-rotaxane: a complex of the hydrolysed track and the deprotected macrocycle.

It was clear that the deprotection required optimisation work before further molecular machine material was committed. It is poignant to note that the barrier hydrolysis problems during deprotection illustrated here are reminiscent of those seen in the only successful two-barrier operation to date (generating 5–8% of product) reported in Section 4.1.7 (Scheme 11). It could have also been the case that the second barrier was cleaved during the deprotection, either partially or fully, leading to the low observed yields.

A thorough investigation into the deprotection began with stability studies on the cresol ester of phenylacetic acid (4.86, a barrier-1 mimic, Scheme 40). This ester was subject to a harsh set of deprotection conditions using an extreme excess of TIPS and high concentration. The ester proved entirely stable to these conditions indicating a discrepancy between the model and the rotaxane. Additionally, in a repeat study also containing a mimic of the catalytic unit, the cresol ester was again stable and the deprotection of the trityl group was complete and clean.

![Scheme 40: Stability studies on a cresol ester mimic of barrier-1 (4.86) using a large excess of TIPS both in the absence and presence of a catalytic unit mimic (4.95) bearing the trityl group to be deprotected.](image)

During the deprotection study in the presence of the trityl-protected catalytic unit, it was noted that the fluorescent yellow colour produced when the catalyst is dissolved in a mixture of dichloromethane and TFA (from the dissociated trityl cation) disappeared to leave a colourless solution within 10 s of TIPS addition indicating complete reduction of the trityl cation.

The deprotection was carried out using conditions developed by Dr M. O. Kitching: TIPS (5 eq.), CH₂Cl₂:TFA (2.5:23, 16 mM), r.t., 4.5 h. This deprotection gave a colourless reaction mixture (by eye) in <60 s. Upon workup of the reaction mixture after 60 s, the deprotection was shown to be complete by ¹H-NMR. These conditions were used on trityl-protected amide macrocycle 4.93 to give complete deprotection in <150 s (both visually and subsequently by ¹H-NMR analysis after
workup). Following this, one-barrier rotaxane 4.100 was deprotected to give complete deprotection in <180 s by $^1$H-NMR and ESI-MS. Confidence in the complete deprotection of rotaxane 4.100 in <180 s under these conditions was important as the rotaxane itself is yellow in solution and its deprotection cannot therefore be as readily followed by colour-change.

The optimised deprotection procedure was subsequently carried out on the cresol ester of N-Boc-(4-nitro)phenylalanine (4.110) and no ester hydrolysis or other side reactions were observed. These results showed that a robust and rapid deprotection strategy had been established that ensured complete global deprotection without side reactions, specifically barrier cleavage.
4.6 Operation of the Third-Generation Molecular Machine

The operation of molecular machine 4.104 was preceded by its global deprotection using the successful conditions developed in Section 4.5.6. The deprotection of the molecular machine, followed by the results and discussion of its operation are described in the following sections.

4.6.1 Deprotection of the Third-Generation Molecular Machine

The global deprotection of molecular machine 4.104 was carried out with TIPS and TFA to remove the Boc protecting group on the 4-nitrophenylalanine barrier and the trityl protecting group on the catalytic unit (Scheme 41). Optimised conditions (see Section 4.5.6) were again employed only at higher dilution for the practical reason of working with small quantities of molecular machine (<3 µmol, 10 mg of 4.104) and therefore over a longer duration: an alteration which could compromise barrier stability.

The resulting TFA salt of the deprotected machine (4.105) underwent the previously used purification process: dilution and azeotrope with anhydrous toluene, followed by trituration with a mixture of diethyl ether and n-hexane (see Section 4.3.1 and Section 4.5.2).

Scheme 41: Trityl and Boc global deprotection of 4.104 to give 4.105, ready for operation.
The deprotected machine (4.105) was not fully characterised before its operation, however an analytical sample was retained to record its mass spectrum using ESI-(+)-MS. The obtained spectrum is shown in Figure 11.

![Figure 11: ESI-(+)-MS analysis of the TFA salt of deprotected two-barrier rotaxane 4.105. Masses are shifted by ca. 0.5 m/z units due to a calibration discrepancy in the spectrometer and in the case of the [M+H]+ base peak, the labelled signal does not correspond to the lowest mass isotope but to the base peak of the isotopic distribution.

The base peak of the mass spectrum corresponds to the proton adduct of the expected product and the sodium adduct can also be assigned. Unfortunately, the spectrum also showed some trityl-protected machine indicating an incomplete reduction of the trityl carbocation, most likely as a result of the reduced concentration of TIPS. A signal corresponding to the loss of the second barrier was also visible. Although it could be argued that this barrier cleavage is a result of fragmentation upon ionisation in the spectrometer, it would be more realistic to assume that some hydrolysis has occurred during the deprotection which was carried out over a longer duration due to higher dilution. These problems could be realistically overcome by employing exactly the optimised deprotection conditions when working on a larger scale.

Regardless of the comparatively small amounts of trityl protected and partially hydrolysed rotaxane material, the largely clean deprotected mixture containing 4.105 was operated with triethylamine.
4.6.2 Operation of the Third-Generation Molecular Machine

The machine operation was attempted under the following conditions: 10 mg (2.8 µmol) of machine 4.105; 10 eq. NEt₃; MeCN:DMF-d₇ (2:1); 4.0 mM; r.t.. The reaction mixture was sampled after 18 h, 36 h and 72 h. Each aliquot was quenched with 10% H₂O/MeCN and analysed by HPLC, LC-MS (m/z range 200–1000) and ESI-(±)-MS (m/z range 50–4000). The reaction mixture was also analysed by DOSY NMR at 72 h by Dr Guillaume De Bo. The results and analysis of this operation are summarised below.

**HPLC** analysis indicated the formation of many species:

- No desired product 4.106 was detected by HPLC;
- DKP 4.74 was formed and was confirmed by LC-MS;
- Various signals assignable to machine-type architectures were visible;
  - LC-MS showed only amide macrocycle 4.98 for these peaks due to the available m/z range: thread-containing components would be outside of this range;
- One dominant peak with unknown identity was visible;
  - LC-MS indicated singly charged peaks with m/z of +280 and -317/318, most likely corresponding to [M+H]⁺ and [M+Cl]⁻ molecular ions respectively and therefore a compound of ca. 279–281 Da. Poor accuracy in this extrapolated mass is a result of a discrepancy between positive and negative tune calibration of the LC-MS spectrometer;
- Two aliquots were taken at the second time point (36 h). One aliquot was quenched with 10% water in acetonitrile and the other with 10% methanol in acetonitrile. One would expect these quenches to form the product acid (4.106) and methyl ester (4.108) respectively. No differences were observed in the analysis of these divergent quenches indicating that the quench was not cleaving an activated ester from the catalytic unit as anticipated.

**ESI-(±)-MS (m/z range 50–4000)** showed machine based species:

- Uninterlocked macrocycle 4.98 seen in LC-MS could not be detected;
- Of note in the high-mass region (m/z > 2000) are 4.105 missing its second barrier but with its first barrier attached (4.107a);
- and fully ‘operated’/unloaded rotaxane (4.107b).
**DOSY NMR** of the quenched operation mixture at 72 h was carried out by Dr Guillaume De Bo and indicated that residual operated and semi-unloaded thread species diffuse at the same rate as the macrocycle. It can therefore be assumed that the macrocycle and thread remain interlocked throughout the operation and quench and that the amide macrocycle 4.98 visible by LC-MS is formed via decomposition or fragmentation of rotaxane in the MS chamber. This assumption is in concordance with the detection of rotaxane-species by ESI-(±)-MS.

After analysis had been carried out on the crude operation mixture, it was purified by prep. TLC. Two major (highly UV-active) species were present in the prep. TLC which contained 14 visible fraction bands. These two major species were confirmed by ESI-(±)-MS to be fully unloaded rotaxane 4.107b and rotaxane 4.107a where only the second barrier had been lost. No further information could be gleaned from the remaining 12 fraction bands due to the minute quantities on material that could be isolated. It is pertinent to note that operation of a 10 mg sample of molecular machine theoretically yields ca. 0.9 mg of the desired product. The quantities of the multiple side-products formed are therefore very low, commonly <100 µg, rendering them difficult to handle and analyse.

From this operation, unsettling conclusions could be drawn. Firstly, the problematic formation of DKP 4.74 was persisting. DKP must be formed via an intermolecular reaction and therefore higher dilution would be required to circumvent this. Secondly, ‘partially-operated’ or ‘partially-cleaved’ machines were detected having barrier one intact but barrier two cleaved (4.107a). This is particularly concerning as it indicates a reluctance of the catalyst to pick-up barrier one compared to the background rate of barrier two cleavage (which may be induced by all or any combination of: hydrolysis; catalyst over-reaching; or intermolecular pick-up).
4.7 A Basic Choice: DBU or NEt₃

Progressing from the disappointing operation of the third-generation design, the use of DBU as the base for operation was reassessed. Compared to NEt₃, the more basic DBU was envisaged to increase the rate of barrier-1 pick-up by increasing the ratio of deprotonated to protonated triazole catalyst.

Further one-barrier machine operation studies based on those carried out with diethylamine (Section 4.5.4) were implemented. Differences included neutralised 4-nitrophenylalanine methyl ester hydrochloride acting as the intermolecular nucleophile to mimic the prospective second barrier of a two-barrier machine. In these further studies DBU and NEt₃ would be directly compared as operation bases. Additionally, each operation would be carried out in parallel with a control study where the catalytic unit would remain trityl protected, and hence inactive, to gain insight into background processes. The compositions of these experiments and the summarised results are outlined in Scheme 42. The studies were followed and analysed by HPLC and ¹H-NMR.

In the cases where DBU was used as base, the amide product 4.108 was readily formed regardless of whether the catalytic unit had been deprotected or not. When the catalytic unit was protected (R = Trt), the one-barrier rotaxane 4.100 was completely consumed after 17 h. Non-interlocked trityl-protected amide macrocycle 4.93 and free unloaded thread 4.101 were identified in the reaction mixture alongside the amide product 4.108. When the catalytic unit was deprotected (R = H, 4.102), the same results were observed after 17 h: complete conversion to amide 4.108 with non-interlocked co-products 4.98 and 4.101. In both the protected and deprotected cases, conversion to the product 4.108 appears to be clean with no discernible by-products detected by HPLC or by ¹H-NMR.

In contrast, the control reaction (with 4.100, R = Trt) carried out using NEt₃ as base showed no formation of amide 4.108 as desired. In the case of the deprotected machine (4.102, R = H), small quantities of unidentified side products were generated over the 7-day reaction period, but only a trace amount of the desired amide 4.108 could be detected by HPLC. One-barrier rotaxanes 4.100/4.102 were again identified in both NEt₃ reaction residues after the 7-day operation period with no evidence of barrier loss.

This combination of results reinforced an underlying suspicion that DBU itself was acting as an acyl-transfer catalyst preferentially to the 1,2,4-triazole catalytic unit. This explains the formation of product 4.108 regardless of whether or not the triazole was protected and the
Scheme 42: One-barrier control studies and machine operations using 4-nitrophenylalanine methyl ester as a nucleophile for acyl pickup and comparing DBU and NEt$_3$ as bases. Reactions are carried out in MeCN:DMF (3:1, 0.4 mM) at room temperature.
observation that when DBU is removed from the system and replaced with NEt₃, no product amide **4.108** is formed from the machine in either its protected or deprotected form.

DBU is known to facilitate acyl transfer as described by Birman[^32] and others[^33,34] although it was later reported by Birman[^20] (in his publication describing the 1,2,4-triazole catalysis) that amide bond formation was shut down when DBU was present in the absence of the triazole catalyst. Additionally, Dr Sonja Kuschel showed this to also be the case with model cresol esters[^21] when initially investigating this system.

To affirm the theory of DBU acting as an acyl-transfer catalyst, the formation of amide **4.108** from cresol ester **4.86** and neutralised 4-nitrophenylalanine methyl ester hydrochloride in the presence of DBU was monitored by periodic ¹H-NMR. Test **A** was carried out as a control with 1 eq. of DBU to neutralise the hydrochloride salt of the amine **4.71**; test **B** was carried out using 2 eq. of DBU. The ¹H-NMR results of test **B** only are shown in Scheme 43.

The results shown in Scheme 43 confirm the ability of DBU to act as a trans-acylation catalyst: conversion to amide **4.108** at 51 h was 55–60%.[^vi] As such, the combination of evidence described here and previously led the project team to abandon DBU as a suitable base for this system. Furthermore, all previous operation results that were obtained in the presence of DBU must be re-evaluated and in many cases, discounted. Sadly this includes the successful one- and two-barrier transacylation molecular machine operations which have both inspired and justified the continued research and investment in this project.

[^vi]: In the control experiment (condition set A), conversion to **4.108** was 10–15% over the same time period, however would likely be negligible if there had been no hydrochloride salt to neutralise **in situ** and the study was in complete absence of DBU. Further elucidation of this was deemed unnecessary due to the nature of the result under condition set B.
Scheme 4.3: Stability study of p-tolyl 2-(4-nitrophenyl)acetate 4.86 and 4-nitrophenylalanine methyl ester hydrochloride (4.71) with: (A) 1 eq. and; (B) 2 eq. of DBU. In each case, 1 eq. of DBU is required to neutralise the amine hydrochloride salt. Reactions are in MeCN-d$_3$:DMF-d$_7$ (3:1, 4.4 mM). Selected $^1$H-NMR spectra of B only are shown which highlight the formation of amide 4.108 (red highlighted peaks) and cresol 4.88 (purple highlighted peaks). Blue and green highlighted peaks indicate the consumption of cresol ester 4.86 and amine 4.71FB starting materials respectively. Reactions were carried out in NMR tubes at r.t. with no agitation. Time points are recorded at 12, 25, 38, 49 and 51 hours. The top reference spectrum is of pure amide 4.108 in MeCN-d$_3$:DMF-d$_7$ (3:1).

In light of DBU’s incompatibility with the designed system and the poor performance of operations using NEt$_3$, the decision was taken to cease investigations into this transacylation molecular machine. Instead, the other projects included in this thesis would be prioritised.
4.8 Outlook

Although research into this project was ceased by the author, the project goal remains a valuable prize and there are many unexplored avenues which could lead to successful functional systems able to achieve it. Within the realms of 1,2,4-triazole catalysis, many different activating bases could be explored. One such base is Hünig’s base (DIPEA) which has shown some promise in preliminary studies. Equally, success may be found in moving away from ambient temperature operations; model studies using Hünig’s base at up to 60 °C have also shown promise in model reactions and barrier mimics have shown good stability, although conversions are slow, typically requiring over 7 days. Non-nitrogen bases have yet to be explored in any depth; they may also show promise.

Further possibilities clearly lie beyond 1,2,4-triazole catalysis: there is a wide variety of other transacylation architectures which remain unexplored in terms of their compatibility with a transacylation molecular machine. Such candidates could include many of the broad range of available acyl transfer catalysts, for example using histidine in an imidazole-based transacylation.

In addition to changes in the catalytic unit and the general mode of catalysis, different amino acid barriers could be investigated, as could a simplified rotaxane architecture, specifically in the thread. It is advised, however, that caution be taken with any new design such that model reactions and CPK modelling are carried out in an ab initio manner without prejudice or preconception from previous systems.
4.9 Summary and Conclusions

The proposed design and development of a rotaxane-based molecular machine which can assemble short peptide chains of three or more components in a sequence-specific manner via the operation of a transacylation catalyst has been a prominent target within the Leigh group for almost a decade. The use of a transacylation-based machine represents a significant progression and diversion from the related NCL-based molecular machines previously published by Leigh.[1,2]

A transacylation design could display many advantageous features compared to its NCL predecessor. These include a catalysis which assembles peptide chains in the ribosomal N-to-C direction, opening doors to a wide range of new potential products including precursors to important drug molecules such as Penicillin G. The catalysis would operate iteratively via a constant transition state size and could display broader amino acid substrate compatibility than NCL catalysts. Furthermore, the presence of a Cys-Xxx-Xxx catalytic unit fingerprint left behind on an NCL-operation product would be avoided to give products with no traces of the machines by which they were synthesised.

Various candidates for transacylation catalysts were previously investigated, including derivatives of dimethylamino pyridines, aliphatic thiols, thiophenols and hydroxybenzotriazoles. All of these catalysts showed varying degrees of promise, but were all eventually let down by the diverse demands that their use placed on rotaxane-based systems.

More success was achieved when 1,2,4-triazoles were investigated as the transacylation catalyst which is active as a triazolium anion when deprotonated. This catalyst successfully facilitated amide bond formation when activated with DBU in model studies using mimics of amino acid building blocks loaded onto barriers via phenolic (cresol) esters and nucleophilic N-unprotected amino acids. Furthermore, this catalyst could be incorporated into the architecture of a one-barrier rotaxane molecular machine with a modified design inspired by its NCL-ancestor. Operation of such a one-barrier machine was successful in giving clean, if incomplete, conversion to an amide product using an intermolecular nucleophile. Extension of this system to a two-barrier rotaxane model was less successful, highlighting many competing processes and complex side-reaction profiles.

Many synthetic steps were investigated and optimised in order to increase the overall efficiency of machine synthesis, deprotection and operation. These include developing the process for loading the catalytic unit, the final synthetic step prior to operation, and exploring different bases which could activate the catalytic unit during operation. The system underwent a series of design
alteration, operations, operation analyses and subsequent redesign cycles. Some of the more prominent outcomes of these redesigns involved changing barriers to suppress side-reactions and synthesising a rotaxane with stoppers on each end of the track to maintain the interlocked architecture after complete operation in a step towards a reloadable machine.

A new strategy for connecting the catalytic unit to the molecular machine was instated. The novel approach involved the successful application of AMT rotaxane formation on a macrocycle bearing the catalytic unit attached through a non-dynamic link. Attempts of AMT rotaxane formations with such macrocycles had been previously unsuccessful in the Leigh group but since their development within this project, they have become the go-to option in all new related designs. This strategy had an additional, serendipitous advantage in that purification of rotaxane species from AMT co-products was also simplified, thus removing a significant synthetic bottleneck.

The developed AMT process with a modified macrocycle bearing the catalytic unit also opened up new opportunities as the catalyst bore a trityl protecting group which, when intact, rendered the catalyst inactive. This allowed investigations to be carried out into the actual mechanism of peptide bond formation in a one-barrier operation. Diethylamine was used, acting both as the base and as the intermolecular nucleophile. Parallel operations were successfully carried out to compare the results when the catalytic unit was protected (inactive) and deprotected (active), giving the first direct evidence of the catalytic unit working as designed where the result could be directly compared to a reference control.

Sadly, further operations of one- and two-barrier molecular machines did not result in progress towards the overall goal of successfully operating a three-barrier transacylation-based molecular machine. Each iterative operation brought with it more problems, riddled with more challenging analyses. This led to regressive steps to reinvestigate the operation bases being employed, DBU and NEt₃, but with the benefit of having a reference control from which to measure the success of operations. It became quickly evident that when DBU was employed, the outcome was the same regardless of the protection-state of the catalyst indicating that DBU itself could catalyse the reaction. Additionally, NEt₃ was shown to be unsuccessful in generating significant amounts of intended products. With clarified insight into the actions of each base during operations, the difficult decision was taken to cease to investigate and develop this project further at this time.

Arriving at this decision was made possible only by analysing model reactions alongside control references. Previously, it was assumed that the catalytic unit was working as intended and indeed some model studies where the catalyst was omitted supported this assumption. Only by installing
a protecting group on the catalyst could the real problems with DBU background reactions be elucidated, however.

Model studies are a means of simplifying a system in order to focus-in on a specific aspect of it. This permits the chemist to quickly and conveniently learn as much as possible about an aspect of the system without investing in lengthy syntheses. Model studies are one of the most valuable tools to chemists developing molecular machines but they must always be treated with caution. Model studies are simplifications: with each simplification that the chemist employs, the larger the risk becomes of a result being non-representative of the system it was designed to probe. In this project, many model studies were carried out where there was simply no way to compare results to a true control, and the turbulent trajectory by which the project developed highlights the dangers associated with this. The importance of carrying out robust, watertight model studies where one understands the possible implications and caveats associated with their interpretations is one of the most important pieces of learning to be extracted from this chapter.

Many thousands of hours of research have been invested in this project by many different people and, although the final goal has not yet been achieved, a huge amount of valuable learning has been obtained. This learning has since become critical to the success of related projects reported both in this thesis and elsewhere within the Leigh group: these achievements are to be celebrated just as much as the lessons learned on designing and developing projects are to be remembered.
4.10 Experimental Section

4.10.1 Synthetic Procedures

General notes on synthesis:

- Barrier amine 4.109,\(^1\) phenol stopper 4.45,\(^{[23]}\) azide stopper 4.99\(^{[22]}\) and \(p\)-tolyl 2-(4-nitrophenoxy)acetate 4.70\(^{[29]}\) were prepared according to literature procedures.
- Propargyl stopper 4.46 was prepared according to a modified literature procedure\(^{[22]}\).
- Compounds 4.55 and 4.56 were prepared in the Leigh group by Jason Hui.
- 4-Nitrophenoxy acetate one-barrier aldehyde rotaxane 4.61 was prepared by Dr Sonja Kuschel: its presented data is extracted from Kuschel’s PhD. thesis\(^{[21]}\).
- The initial investigation into the use of \(\text{NEt}_3\) as a base to activate the 1,2,4-triazole catalyst (Section 4.3.3) was also carried out by Kuschel.
- Benzylic amine macrocycle 4.97 was prepared as outlined in Chapter 3.
- Phenylacetic acid cresol ester (\(p\)-tolyl phenylacetate, 4.86) was purchased from Sigma-Aldrich.
- TMS-unloaded barrier 4.47 and unloaded barrier 4.89 were prepared according to the modified literature procedure\(^{[1]}\) outlined in Chapter 2.
4.10.1.1 Synthesis of Stoppered Terminal Barrier (Loaded) 4.49

Scheme 44: Synthesis of stoppered terminal barrier (loaded) 4.49. Reagents and conditions: a) 4-(TMS-ethynyl)benzoic acid, EDCI·HCl, HOBt·H₂O, THF, CH₂Cl₂, r.t., 16 h, 70%; b) propargyl bromide, K₂CO₃, DMF, 80 °C, 22 h, quant.; c) CuPF₆·4MeCN, NEt₃, CH₂Cl₂, tBuOH, r.t., 21 h, then KOH, MeOH, CH₂Cl₂, r.t., 15 h, 63%; d) Boc-(4-NO₂)Phe-OH, EDCI·HCl, DMAP, CH₂Cl₂, r.t., 19 h, 96%.
4.47: TMS-Barrier (Unloaded)

To a solution of amine 4.109 (773 mg, 1.86 mmol, 1 eq.), 4-[(trimethylsilyl)ethynyl]benzoic acid (486 mg, 2.23 mmol, 1.2 eq.) and EDCI·HCl (425 mg, 2.22 mmol, 1.2 eq.) in a mixture of THF:CH₂Cl₂ (1:3, 40 mL, 47 mM) at room temperature was added HOBT·H₂O (301 mg, 1.97 mmol, 1.06 eq.). The resulting mixture was stirred for 16 h at room temperature before solvent was removed in vacuo and the residue was partitioned between CH₂Cl₂ (250 mL) and H₂O (150 mL). The organic layer was separated, dried with MgSO₄ and then concentrated to dryness in vacuo. The residue was purified by flash column chromatography (SiO₂, 14–50% EtOAc/pet. ether) to afford the title compound as a pale yellow amorphous solid (801 mg, 70%).

m.p. 132–134 °C. [α]D²⁰ = +10.6 ° (c = 1.01, CH₂Cl₂). ¹H NMR (600 MHz, Acetone-d₆) δ 9.36 (s, 1H, r), 8.15 (s, 1H, p), 7.94 (d, J = 7.8 Hz, 1H, i), 7.87 (d, J = 8.1 Hz, 2H, f), 7.55 (dd, J = 8.6, 1.8 Hz, 2H, t), 7.52 (d, J = 8.3 Hz, 2H, e), 7.28 (d, J = 8.6 Hz, 2H, z), 7.19 – 7.14 (m, 4H, m+u), 6.99 (d, J = 8.6 Hz, 2H, aa), 6.73 (d, J = 8.4 Hz, 2H, n), 4.90 (q, J = 7.7 Hz, 1H, j), 3.23 (dd, J = 13.9, 5.9 Hz, 1H, k), 3.07 (dd, J = 13.9, 8.5 Hz, 1H, k'), 1.65 (s, 6H, x), 0.24 (s, 9H, a). ¹³C NMR (151 MHz, Acetone) δ 170.6 (q), 166.6 (h), 157.0 (o), 148.8 (y), 146.5 (v), 138.1 (ab), 137.6 (s), 135.1 (g), 132.5 (e), 131.2 (m), 129.2 (z), 128.4 (f), 127.8 (u), 126.8 (d), 120.1 (t), 119.4 (aa), 116.0 (n), 105.1 (c), 97.0 (b), 57.1 (j), 42.9 (w), 37.8 (k), 31.0 (x), -0.2 (a). LR(ESI)-(−)-MS: m/z -614.3 [M-H]⁻. HR(ESI)-(−)-MS: m/z +638.2553 [M+Na]⁺; +638.2558 calculated for [C₃₆H₃₇O₃N₅SiNa]⁺. Data in agreement with literature.[¹]

4.46: Propargyl Stopper

To a solution of 4-tris[4-(tert-butyl)phenyl]methyl phenol (4.45, 1.60 g, 3.17 mmol, 1 eq.) and K₂CO₃ (2.21 g, 16 mmol, 5 eq.) in DMF (30 mL, 106 mM) was added propargyl bromide (80% solution in toluene, 0.53 mL, 4.75 mmol, 1.5 eq.). The resulting solution was heated at 80 °C for 21.5 h before being cooled to ambient temperature and concentrated in vacuo to a brown solid. The residue was partitioned between EtOAc (100 mL) and aq. LiCl (5% w/v, 100 mL) and passed over a glass sinter. The organics were separated and the aqueous layer was extracted further with EtOAc (2 x 100 mL). The combined organics were dried over MgSO₄ and concentrated in vacuo to dryness to give a crude solid. The crude solid was triturated with refluxing MeCN (80 mL) for 1 h.
then upon cooling to room temperature, the slurry was filtered. The filtrate was dried in vacuo to give the title compound as a white amorphous solid (1.72 g, quant.).

m.p. 252–256 °C. \textsuperscript{1}H NMR (600 MHz, Chloroform-\textit{d}) \( \delta \) 7.23 (d, \( J = 8.5 \) Hz, 6H, \textit{d}), 7.10 (d, \( J = 8.9 \) Hz, 2H, \textit{i}), 7.07 (d, \( J = 8.5 \) Hz, 6H, \textit{e}), 6.84 (d, \( J = 8.9 \) Hz, 2H, \textit{j}), 4.66 (d, \( J = 2.3 \) Hz, 2H, \textit{l}), 2.52 (t, \( J = 2.3 \) Hz, 1H, \textit{n}), 1.30 (s, 27H, \textit{a}). \textsuperscript{13}C NMR (151 MHz, Chloroform) \( \delta \) 155.6 (\textit{k}), 148.5 (\textit{c}), 144.2 (\textit{f}), 140.6 (\textit{h}), 132.4 (\textit{i}), 130.9 (\textit{m}), 124.2 (\textit{d}), 113.4 (\textit{j}), 78.9 (\textit{m}), 75.5 (\textit{n}), 63.2 (\textit{g}), 55.9 (\textit{l}), 34.4 (\textit{b}), 31.5 (\textit{a}).

\textsuperscript{LR(APCI)}-(+)MS: m/z +560.4 [M+NH\textsubscript{4}]\textsuperscript{+}. HR(APCI)-(+)MS: m/z +542.3545 [M]\textsuperscript{+}; +542.3543 calculated for [C\textsubscript{40}H\textsubscript{46}O]\textsuperscript{+}. Data in agreement with literature.\textsuperscript{[22]}

### 4.48: Stoppered Terminal Barrier (Unloaded)

![Chemical structure diagram]

TMS barrier 4.47 (756 mg, 1.23 mmol, 1 eq.), propargyl stopper 4.46 (1.00 g, 1.84 mmol, 1.5 eq.) and CuPF\textsubscript{6}·4MeCN (92 mg, 0.25 mmol, 0.2 eq.) were dissolved under an inert atmosphere in a thoroughly degassed mixture of CH\textsubscript{2}Cl\textsubscript{2} (49 mL), \textsuperscript{1}BuOH (1 mL) and NEt\textsubscript{3} (0.2 mL). The resulting mixture (conc. = 25 mM) was stirred at room temperature for 21 h. Upon completion, the reaction mixture was diluted with CH\textsubscript{2}Cl\textsubscript{2} (50 mL) and washed with EDTA·4Na (sat. aq., 2 \times 50 mL) before the combined organics were dried over Na\textsubscript{2}SO\textsubscript{4} and concentrated by approximately half in vacuo to leave a CH\textsubscript{2}Cl\textsubscript{2} solution (ca. 25 mL) of the crude product.

To the CH\textsubscript{2}Cl\textsubscript{2} solution of the crude product (ca. 20 mL) was added a solution of KOH (200 mg, 3.57 mmol, 2.9 eq.) in MeOH (20 mL). The mixture was stirred at room temperature for 15 h before being washed with aq. HCl (1 M, 2 \times 50 mL). The combined organics were then dried over Na\textsubscript{2}SO\textsubscript{4} and concentrated to dryness in vacuo. The crude residue was purified by flash column chromatography (SiO\textsubscript{2}, 35–50% EtOAc/pet. ether) to afford the title compound as an off-white amorphous solid (847 mg, 63%).

m.p. 177–182 °C. \([\alpha]\textsubscript{D}\textsuperscript{20} = -1.18 \degree \) (c = 1.00, CH\textsubscript{2}Cl\textsubscript{2}). \textsuperscript{1}H NMR (600 MHz, Acetone-\textit{d}) \( \delta \) 9.40 (s, 1H, Bn), 8.60 (s, 1H, Bc), 8.17 (s, 1H, Bu), 7.96 (d, \( J = 8.0 \) Hz, 1H, Bv), 7.87 (d, \( J = 8.3 \) Hz, 2H, By), 7.78 (d, \( J = 8.6 \) Hz, 2H, Be), 7.57 (d, \( J = 8.7 \) Hz, 2H, Bl), 7.54 (d, \( J = 8.3 \) Hz, 2H, Bz), 7.45 (d, \( J = 8.6 \) Hz, 2H, Bf), 7.31 (d, \( J = 8.6 \) Hz, 6H, Sd), 7.21 (d, \( J = 8.7 \) Hz, 2H, Bk), 7.18 (d, \( J = 8.3 \) Hz, 2H, Bs), 7.15 – 7.09 (m, 8H, Se+Si), 6.97 (d, \( J = 8.9 \) Hz, 2H, Sj), 6.73 (d, \( J = 8.3 \) Hz, 2H, Bt), 5.24 (s, 2H, Ba), 4.93 (ap. td, \( J = 8.3, 8.0, 5.8 \) Hz, 1H, Bp), 3.82 (s, 1H, Bac), 3.24 (dd, \( J = 13.9, 5.8 \) Hz, 1H, Bq), 3.07 (dd, \( J = 13.9, 8.3, 8.0, 5.8 \) Hz, 1H, Bp).
4.49: Stoppered Terminal Barrier (Loaded)

To a solution of stoppered unloaded barrier 4.48 (200 mg, 184 µmol, 1 eq.), Boc-4-nitro-L-phenylalanine (63 mg, 203 µmol, 1.1 eq.) and EDCI·HCl (41 mg, 212 µmol, 1.15 eq.) in CH₂Cl₂ (10 mL, 18 mM) was added DMAP (34 mg, 276 µmol, 1.5 eq.) at room temperature. The resulting mixture was stirred for 19 h before a further portion of EDCI·HCl (18 mg, 92 µmol, 0.5 eq.) was added. The mixture was stirred for a further 5 h before it was diluted with CH₂Cl₂ (20 mL) and washed with NH₄Cl (sat. aq., 2 × 20 mL). The organics were dried over Na₂SO₄ and then concentrated to dryness in vacuo. The resulting crude solid was purified by flash column chromatography (SiO₂, 10–20% EtOAc/CH₂Cl₂) to afford the title compound as an off-white amorphous solid (243 mg, 96%).

m.p. 165–167 °C. [α]D 20 = −5.18 ° (c = 1.01, CH₂Cl₂). 1H NMR (600 MHz, Acetone-d₆) δ 9.43 (s, 1H, ba), 8.62 (s, 1H, bl), 8.20 (d, J = 8.5 Hz, 2H, at), 8.04 (d, J = 8.0 Hz, 1H, ah), 7.87 (d, J = 8.3 Hz, 2H, ae), 7.79 (d, J = 8.7 Hz, 2H, bj), 7.65 (d, J = 8.4 Hz, 2H, as), 7.59 - 7.54 (m, 4H, bc+ad), 7.46 (d, J = 8.7 Hz, 2H, bi), 7.40 (d, J = 8.3 Hz, 2H, al), 7.32 (d, J = 8.6 Hz, 6H, bv), 7.23 (d, J = 8.6 Hz, 2H, bd), 7.15 - 7.10 (m, 8H, bu+bq), 7.02 (d, J = 8.4 Hz, 2H, am), 6.98 (d, J = 8.9 Hz, 2H, bp), 6.63 (d, J = 8.2 Hz, 1H, av), 5.25 (s, 2H, bn), 5.00 (q, J = 8.1 Hz, 1H, ai), 4.74 - 4.64 (m, 1H, ap), 3.83 (s, 1H, aa), 3.47 (dd, J = 13.9, 5.3 Hz, 1H, aq), 3.36 (dd, J = 13.9, 5.8 Hz, 1H, aj), 3.31 (dd, J = 13.9, 9.7 Hz, 1H, aq'), 3.18 (dd, J = 13.9, 8.7 Hz, 1H, aj'), 1.72 (s, 6H, bg), 1.35 (s, 9H, ay), 1.30 (s, 27H, by). 13C NMR (151 MHz, Acetone) δ 171.1 (ao), 170.3 (az), 166.8 (ag), 157.4 (bo), 156.4 (aw), 152.4 (bh), 150.5 (an), 149.2 (bw), 147.9 (au), 146.5 (ar), 146.2 (be), 145.3 (bm), 145.2 (bt), 140.7 (br), 137.6 (bb), 136.5 (ak), 135.8 (bk), 135.2 (af), 132.8 (bq), 132.7 (ad), 131.5 (as), 131.4 (bu), 131.2 (al), 129.0
Optimised Synthesis of Aldehyde Macrocycle 4.59

4.10.1.2 Optimised Synthesis of Aldehyde Macrocycle 4.59

Figure 12: Optimised synthesis of aldehyde macrocycle 4.59. Reagents and conditions: a) Pd(PPh₃)₂Cl₂, Cul, NEt₃, THF, 70 °C, 16 h, 87%; b) H₂, Pd/C, THF, r.t., 48 h, 98%; c) PPh₃, CBr₄, CH₂Cl₂, 0 °C – r.t., 18 h, 81%; d) 9-BBN, THF, r.t., 2 h; then 2,6-dibromopyridine, PdCl₂, PPh₃, K₂CO₃, H₂O, 60 °C, 16 h, 86% over 2 steps; e) conc. HBr, reflux, 14 h, 95%; f) C₂O₂, DMF, 4.2 mM, 50 – 60 °C, 68 h, 56%; g) DIBAL, CH₂Cl₂, 0 °C – r.t., 18 h, 95%.

4.55: 3,5-bis(4-hydroxybut-1-yn-1-yl)benzonitrile

The synthesis and characterisation of 4.55 was carried out by Mr Jason Hui.

3-butyn-1-ol (33 mL, 0.65 mol, 8.7 eq.) and 3,5-dibromobenzonitrile (13.86 g, 50 mmol, 1 eq.) were dissolved in triethylamine (104 mL) and THF (260 mL) and stirred under argon for 20 mins. Pd(PPh₃)₂Cl₂ (3.53 g, 5 mmol, 0.1 eq.) and
CuI (1.9 g, 10 mmol, 0.2 eq.) were added and the mixture was heated at reflux (70 °C) for 16 h. The reaction was then cooled, EtOAc (300 mL) was added, and the mixture was filtered over Celite®. The resulting solution was washed with NH₄Cl (sat. aq., 3 × 100 mL), dried over Na₂SO₄ and concentrated in vacuo. Purification by flash column chromatography (SiO₂, 70–80% EtOAc/n-hexane) afforded the title compound as an off-white amorphous solid (10.5 g, 87%).

m.p. 64–66 °C. ¹H NMR (600 MHz, Acetone-d₆) δ 7.71 (d, J = 1.5 Hz, 2H, c), 7.66 (t, J = 1.6 Hz, 1H, e), 4.10 (br. s, 2H, j), 3.73 (td, J = 6.5, 2.0 Hz, 4H, i), 2.63 (t, J = 6.6 Hz, 4H, h). ¹³C NMR (151 MHz, Acetone-d₆) δ 139.1 (e), 134.4 (c), 126.6 (d), 118.1 (a), 114.1 (b), 92.2 (g), 79.4 (f), 61.0 (i), 24.3 (h).

LRESI-(+)-MS: m/z 262.1 m/z [M+Na]⁺. HRFT-(+)-MS: 240.1016 m/z [M+H]⁺, 240.1019 calc. for [C₁₅H₁₃NO₂H]⁺. Data is in accordance with literature precedent.[1]

4.56: 3,5-bis(4-hydroxybutyl)benzonitrile

The synthesis and characterisation of 4.56 was carried out by Mr Jason Hui. Hydrogen was bubbled through a solution of 4.55 (10.9 g, 45.6 mmol, 1 eq.) in THF (200 mL) for 1 h. Palladium on carbon (10 wt.% loading, 1.0 g) was added to the solution and the resulting mixture stirred at room temperature under an atmosphere of hydrogen for 48 h. The mixture was filtered over Celite® and concentrated in vacuo to give the title compound as a colourless oil (8.9 g, 98%).

¹H NMR (600 MHz, Chloroform-d) δ 7.30 (d, J = 1.6 Hz, 2H, c), 7.24 (d, J = 1.6 Hz, 1H, e), 3.67 (t, J = 6.4 Hz, 4H, i), 2.65 (t, J = 7.7 Hz, 4H, f), 1.87 – 1.84 (m, 2H, j), 1.72 – 1.68 (m, 4H, h), 1.59 (dd, J = 8.8, 6.4 Hz, 4H, g). ¹³C NMR (151 MHz, Chloroform) δ 143.6 (d), 133.3 (e), 129.5 (c), 119.3 (a), 112.2 (b), 62.6 (i), 35.2 (f), 32.1 (h), 27.3 (g). LRESI-(+)-MS: m/z 270.1 m/z [M+Na]⁺. HRFT(+)−MS: 248.1643 m/z [M+H]⁺, 248.1645 calc. for [C₁₅H₂₁NO₂H]⁺. Data is in accordance with literature precedent.[1]

4.57: 3,5-bis(4-bromobutyl)benzonitrile

Multi-gram scale procedure: Tetrabromomethane (27.36 g, 82.5 mmol, 2.4 eq.) was added to a solution of 4.56 (8.50 g, 34.4 mmol, 1 eq.) in CH₂Cl₂ (850 mL, 40 mM) at -10 °C. Triphenylphosphine (19.83 g, 75.6 mmol, 2.2 eq.) was added in ca. 2 g portions over 5 min before the resulting mixture was allowed to warm to room temperature and left to stir for 24 h. The reaction mixture was concentrated in vacuo to a thick slurry (ca. 100 mL) which was subsequently triturated with pet. ether (500 mL). The liquors were decanted and trituration of the remaining solid was repeated twice further (2 × 500 mL). The combined liquors were concentrated in vacuo to an oil (ca. 20 g). The resulting oil
was dissolved in CH$_2$Cl$_2$ (200 mL) and dry-loaded onto silica (SiO$_2$, 120 g, 6 mass eq.). The crude-loaded silica was loaded onto a sintered filter and washed with pet. ether (2 L) which was discarded. The crude loaded silica was then eluted with EtOAc/pet. ether (10%, 1 L) to elute the product. The product containing eluent was concentrated in vacuo to give the title compound as a colourless oil (9.40 g, 73%).

Gram-scale procedure: Tetrabromomethane (3.22 g, 9.7 mmol, 2.4 eq.) was added to a solution of 4.56 (1.00 g, 4.04 mmol, 1 eq.) in CH$_2$Cl$_2$ (15 mL, 270 mM) at -10 °C. Triphenylphosphine (2.33 g, 8.90 mmol, 2.2 eq.) was added in ca. 0.5 g portions over 5 min before the resulting mixture was allowed to warm to room temperature and left to stir for 24 h. The reaction mixture was concentrated in vacuo to dryness and the crude residue dry-loaded onto silica (8 mass eq.) and purified by flash column chromatography (SiO$_2$, 6–10% EtOAc/pet. ether) to give the title compound as an analytically pure colourless oil (1.23 g, 81%).

$^1$H NMR (600 MHz, Chloroform-d) δ 7.31 (d, $J$ = 1.7 Hz, 2H, c), 7.23 (t, $J$ = 1.7 Hz, 1H, e), 3.43 (t, $J$ = 6.6 Hz, 4H, i), 2.65 (t, $J$ = 7.7 Hz, 4H, f), 1.91 – 1.85 (m, 4H, h), 1.78 (dq, $J$ = 9.6, 7.5, 7.0 Hz, 4H, g).

$^{13}$C NMR (151 MHz, Chloroform) δ 143.4 (d), 133.3 (e), 129.8 (c), 119.2 (a), 112.5 (b), 34.7 (i), 33.4 (f), 32.2 (h), 29.6 (g). Note: no CBr$_4$ was detected by $^{13}$C NMR at ca. δ -30 ppm. LRESI-(+) MS: m/z 371.0, 372.9 [M+H]$^+$, HRFT(+)-MS: 370.9870 m/z [M+H]$^+$, 370.9878 calc. for [C$_{15}$H$_{19}$NBr$_2$H]$^+$. Data is in accordance with literature precedent.$[^1]$

**4.51: 2,6-bis(3-(4-methoxyphenyl)propyl)pyridine**

9-BBN (0.5 M in THF, 13.5 mL, 6.75 mmol) was added via cannula to a dry 50 mL three-neck round bottom flask containing estragole (4.53, 4-allyl anisole, 1 g, 6.75 mmol) under an inert atmosphere at room temperature. The resulting clear, light yellow solution was stirred at room temperature for 2 h. The reaction mixture was then transferred via cannula to a dry 100 mL three-neck round bottom flask containing 2,6-dibromopyridine (40 mg, 1.69 mmol), palladium(II) dichloride (60 mg, 0.34 mmol), triphenylphosphine (178 mg, 0.68 mmol), potassium carbonate (2.80 g, 20.0 mmol) under an inert atmosphere at room temperature followed immediately by water (7 mL). The reaction mixture was heated to 55 °C and stirred at this temperature for 22 h before being cooled to ambient temperature. The reaction mixture was partitioned between water (50 mL) and ethyl acetate (50 mL), the phases were separated and the aqueous phase was extracted with ethyl acetate (2 × 30 mL). The combined organics were dried over MgSO$_4$ and solvents were removed in vacuo. The crude residue was stirred in HCl (1M in MeOH, 25 mL) for
one hour before solvents were removed in vacuo. The residual oil was tritreated vigorously with pet. ether (3 x 25 mL; on the third trituration, the mixture was stirred vigorously overnight before the pet. ether was decanted). Ethyl acetate (30 mL) was added to the methanolic phase before KHCO$_3$ (sat. aq., ca. 50 mL) was cautiously added to neutralise the mixture. Once a stable basic pH had been reached, the organic phase was separated and washed with brine (25 mL) then dried over MgSO$_4$ and solvents were removed in vacuo. Purification of the crude mixture by flash column chromatography (SiO$_2$, 5–10% ethyl acetate/pet. ether) afforded the title compound as a white crystalline solid (554 mg, 86%).

m.p. 70–75 °C. $^1$H NMR (600 MHz, DMSO-$d_6$) δ 7.76 – 7.46 (m, 1H, k), 7.16 – 6.99 (m, 6H, d+j), 6.82 (d, $J = 8.5$ Hz, 4H, c), 3.71 (s, 6H, a), 2.75 – 2.64 (m, 4H, h), 2.53 (t, $J = 7.5$ Hz, 4H, f), 1.92 (p, $J = 7.5$ Hz, 4H, g). $^{13}$C NMR (151 MHz, DMSO) δ 161.0 (i), 157.8 (b), 137.3 (k), 134.2 (e), 129.7 (d), 120.6 (j), 114.1 (c), 55.4 (a), 37.2 (h), 34.3 (f), 31.7 (g). LRESI-(+)-MS: 376.33 m/z [M+H]$^+$; 376.2256 m/z [M+H]$^+$; 376.2271 calc. for [C$_{25}$H$_{29}$NO$_2$H]$^+$. Data is in accordance with literature precedent.$^{[1]}$

4.54: 4,4'-(pyridine-2,6-diylbis(propane-3,1-diyl))diphenol hydrobromide

2,6-bis(3-(4-methoxyphenyl)propyl)pyridine (8.90 g, 23.7 mmol) was refluxed in aq. hydrobromic acid (48%, 60 mL) for 14 hours. After cooling to ambient temperature, the precipitate was filtered in vacuo and washed with water (4 x 100 mL), followed by pet. ether (4 x 100 mL) to give the title compound as an off-white solid (9.62 g, 95%).

m.p. 162–166 °C. $^1$H NMR (600 MHz, DMSO-$d_6$) δ 15.09 (br. s, 1H, l), 9.18 (br. s, 2H, a), 8.36 (br. s, 1H, k), 7.75 (br. s, 2H, j), 6.99 (d, $J = 7.7$ Hz, 4H, d), 6.67 (d, $J = 7.7$ Hz, 4H, c), 2.94 (t, $J = 7.7$ Hz, 4H, h), 2.54 (t, $J = 7.7$ Hz, 4H, f), 1.95 (p, $J = 7.7$ Hz, 4H, g). $^{13}$C NMR (151 MHz, DMSO) δ 156.7 (i), 155.6 (b), 146.0 (k), 131.0 (e), 129.2 (d), 124.0 (j), 115.1 (c), 33.7 (h), 32.4 (f), 30.6 (g). LRESI-(+)-MS: 348.25 m/z [M-Br]$. HRESI-(+)-MS: 348.1945 m/z [M+H]+; 348.1968 calc. for [C$_{23}$H$_{22}$NO$_2$]+.
### 4.58: Nitrile Macrocycle

Caesium carbonate (21.54 g, 66 mmol) was suspended in dimethylformamide (DMF, 2.5 L) and the mixture was heated to 50 °C under an inert atmosphere. **4.57** (4.11 g, 11 mmol) and **4.54** (4.72 g, 11 mmol) were separately dissolved in DMF (50 mL each) and these solutions were added simultaneously to the warmed reaction mixture, dropwise, over a period of 20 h (rate = 2.5 mL/min, final conc. = 4.2 mM). On complete addition the reaction mixture was heated to 60 °C and stirred at this temperature for 48 h before being cooled to ambient. The reaction mixture was filtered and the filtrate was concentrated to dryness in vacuo.

The residue was extracted with CH$_2$Cl$_2$ (300 mL) and filtered. The filtrand was washed with CH$_2$Cl$_2$ (2 × 200 mL) and the combined filtrates were concentrated in vacuo. The residue was purified by flash column chromatography (SiO$_2$, 20% ethyl acetate/pet. ether) to afford the title compound as white solid (3.42 g, 56%) which can be further purified by recrystallisation (ethyl acetate/hexanes, 92% recovery).

m.p. 141–142 °C. $^1$H NMR (600 MHz, DMSO-$_d_6$) δ 7.58 (t, 1H, $J = 7.7$ Hz, s), 7.51 (s, 2H, c), 7.47 (s, 1H, e), 7.06 (d, 2H, $J = 7.7$ Hz, r), 7.02 (d, 4H, $J = 8.1$ Hz, l), 6.73 (d, 4H, $J = 8.1$ Hz, k), 3.89 (t, 4H, $J = 6.0$ Hz, i), 2.71 – 2.65 (m, 8H, f+p), 2.48 (t, 4H, $J = 7.7$ Hz, n), 1.92 (dt, 4H, $J = 7.6$, 7.6 Hz, o), 1.72 (dt, 4H, $J = 7.5$, 7.6 Hz, g), 1.74 (dt, 4H, $J = 6.7$, 7.3 Hz, h). $^{13}$C NMR (151 MHz, DMSO) δ 160.6 (q), 156.7 (j), 143.7 (d), 136.7 (s), 134.0 (m), 133.5 (e), 129.6 (c), 129.2 (l), 120.2 (r), 119.2 (b), 114.1 (k), 111.1 (a), 66.9 (i), 36.6 (p), 33.8 (f), 33.7 (n), 31.4 (o), 27.8 (h), 27.4 (g). LRESI-(+)-MS: m/z 559.3 [M+H]$^+$.

**HRFT(+)-MS:** 559.3323 $m/z$ [M+H]$^+$, 559.3319 calc. for [C$_{38}$H$_{42}$N$_2$O$_2$H]$^+$.

*Data is in accordance with literature precedent.*[1]

### 4.59: Aldehyde Macrocycle

Nitrile macrocycle **4.58** (2.13 g, 3.82 mmol) was dissolved in degassed CH$_2$Cl$_2$ (35 mL) under an inert atmosphere and the solution cooled to 0 °C. DIBAL (1 M in CH$_2$Cl$_2$, 4.2 mL, 4.2 mmol, 1.1 eq.) was added dropwise over 10 minutes and the resulting clear colourless solution (final conc. = 0.1 M) was allowed to warm to ambient temperature over 13 h before being stirred for a further 5 h at room temperature. The reaction mixture was then cooled to 0 °C and quenched cautiously with MeOH (11 mL). After warming to ambient temperature, the resulting suspension was filtered over Celite® and the
filtrand was washed with CH$_2$Cl$_2$ (3 × 25 mL). The combined filtrates were washed with disodium tartrate (sat. aq., 30 mL), the organics were separated and the aqueous phase was washed with CH$_2$Cl$_2$ (3 × 50 mL). The combined CH$_2$Cl$_2$ phases were dried over MgSO$_4$ and concentrated in vacuo to give a white foam (2.12 g crude mass). The crude product was purified by flash column chromatography (SiO$_2$, 10–20% ethyl acetate/pet. ether) to afford the title compound as white solid (2.04 g, 95%).

m.p. 109–110 °C. $^1$H NMR (600 MHz, DMSO-$d_6$) δ 9.96 (s, 1H, a), 7.60 – 7.56 (m, 3H, c+s), 7.46 (s, 1H, e), 7.06 (d, 2H, J = 7.7 Hz, r), 7.02 (d, 4H, J = 8.5 Hz, l), 6.73 (d, 4H, J = 8.6 Hz, k), 3.90 (t, 4H, J = 6.0 Hz, i), 2.72 (t, 4H, J = 7.4 Hz, f), 2.68 (t, 4H, J = 7.4 Hz, p), 2.47 (t, 4H, J = 7.8 Hz, n), 1.92 (dt, 4H, J = 7.4, 7.5 Hz, o), 1.79 – 1.72 (m, 4H, g), 1.70 – 1.63 (m, 4H, h). $^{13}$C NMR (151 MHz, DMSO) δ 193.3 (a), 160.6 (q), 156.7 (j), 143.2 (d), 136.7 (s), 136.5 (b), 134.7 (e), 134.0 (m), 129.2 (l), 127.2 (c), 120.2 (r), 114.1 (k), 67.0 (l), 36.6 (p), 34.0 (f), 33.7 (n), 31.4 (o), 27.8 (h), 27.6 (g). LRESI-(+)-MS: m/z 562.6 [M+H]$^+$ HRFT-(+)-MS: 562.3310 m/z [M+H]$^+$, 562.3316 calc. for [C$_{38}$H$_{43}$NO$_3$H]$^+$. Data is in accordance with literature precedent.$^{[1]}$

4.10.1.3 Synthesis of Catalytic Unit Hydrazide 4.62 and Test-Hydrazone Formations

Scheme 45: Synthesis and trial-loading of 1,2,4-triazole catalytic unit. Reagents and conditions: a) β-alanine, PyBOP, NEt$_3$, MeCN, 0 °C – r.t., 16 h, 37%; b) H$_4$N$_2$·H$_2$O, EtOH, 25 °C, 12 h, quant.; c) 3,5-dimethylbenzaldehyde, TFA, CD$_2$Cl$_2$, DMSO-$d_6$, r.t., 5 min, 80%; d) Aldehyde macrocycle 4.59, TFA, CH$_2$Cl$_2$, DMSO-$d_6$, r.t., 3 h, 57%.
4.72: Catalytic Unit Methyl Ester

To a suspension of 1,2,4-triazole-3-carboxylic acid (2.00 g, 17.7 mmol, 1.0 eq.) in acetonitrile (100 mL) was added triethylamine (4 mL, 2.90 g, 28.5 mmol, 1.6 eq.) and the resulting mixture was cooled to 0 °C. PyBOP (9.20 g, 17.7 mmol, 1 eq.) was added and the resulting solution was warmed to room temperature before β-alanine methyl ester hydrochloride (2.72 g, 19.5 mmol, 1.1 eq.) was added. The reaction was stirred at room temperature for 16 h. The formed precipitate was filtered and the filtrate was washed with NaHCO₃ (sat. aq., 50 mL), water (50 mL), MeCN (50 mL) and finally Et₂O (50 mL) before being dried in vacuo to give the title compound as a white solid (1.30 g, 6.56 mmol, 37%). m.p. 161–163 °C. ¹H NMR (600 MHz, DMSO-d₆) δ 14.63 (br. s, 1H, h), 8.63 (s, 1H, e), 8.43 (s, 1H, i), 3.60 (s, 3H, a), 3.48 (q, J = 6.9 Hz, 2H, d), 2.59 (t, J = 7.1 Hz, 2H, c). ¹³C NMR (151 MHz, DMSO-d₆) δ 171.7 (b), 51.4 (a), 34.8 (d), 33.4 (c); ¹³C signals for (f), (g) and (i) are not observed. LRESI-(±)-MS: m/z -197.1 [M-H]; +221.1 [M+Na]; +419.2 [2M+Na]. HRFT(+)-MS: m/z +199.0824 [M+H]⁺, +199.0826 calc. for [C₇H₁₀N₃O₃H]⁺.

4.65: Macrocycle Mimic Model Catalyst

A solution of 3,5-dimethylbenzaldehyde (20 mg, 0.149 mmol, 1 eq.) in CD₂Cl₂ (1 mL) was added to a suspension of hydrazide 4.62 (32.5 mg, 0.164 mmol, 1.1 eq.) in DMSO-d₆ (1 mL). To the resulting suspension was added TFA in CD₂Cl₂ (0.1 mL, 0.74 M, 74 µmol, 0.5 eq.). The suspension cleared to give a colourless solution after 5 min at which
point TFA and CD$_2$Cl$_2$ were removed in vacuo. Water (2 mL) was added to the residual DMSO-$d_6$ solution to precipitate the product. The product was isolated as a white amorphous solid upon filtration (37 mg, 80%).

m.p. 261–263 °C. $^1$H NMR (600 MHz, DMSO-$d_6$) δ 15.00 (s, 1H, $\alpha_{\text{min}}$), 14.65 (s, 1H, $\alpha_{\text{maj}}$), 11.63 (s, 1H, $\gamma_{\text{min}}$), 11.34 (s, 1H, $\gamma_{\text{maj}}$), 8.86 (s, 1H, $\kappa''_{\text{min}}$), 8.81 (s, 1H, $\kappa'_{\text{min}}$), 8.66 (s, 1H, $\gamma_{\text{min}}$), 8.50 (s, 1H, $\kappa''_{\text{maj}}$), 8.46 (s, 1H, $\kappa'_{\text{maj}}$), 8.11 (s, 1H, $\delta_{\text{min}}$), 8.08 (s, 1H, $\delta_{\text{maj}}$), 7.94 (s, 1H, $\delta_{\text{maj}}$), 7.27 (s, 1H, $\delta_{\text{min}}$), 7.04 (s, 1H, $\delta_{\text{maj}}$), 7.02 (s, 1H, $\delta_{\text{min}}$), 3.60 – 3.50 (m, 4H, $j_{\text{maj}+\text{min}}$), 2.91 (t, $J = 7.0$ Hz, 2H, $i_{\text{maj}}$), 2.81 (t, $J = 7.0$ Hz, 2H, $i_{\text{min}}$), 2.29 (s, 3H, $a_{\text{min}}$), 2.28 (s, 3H, $a_{\text{maj}}$).

$^{13}$C NMR (151 MHz, DMSO) δ 172.7 (h$_{\text{maj}}$), 166.9 (h$_{\text{min}}$), 146.2 (f$_{\text{min}}$), 143.4 (f$_{\text{maj}}$), 137.9 (b$_{\text{min}+\text{maj}}$), 134.3 (e$_{\text{min}/\text{maj}}$), 134.1 (e$_{\text{min}/\text{maj}}$), 131.5 (c$_{\text{min}}$), 131.3 (c$_{\text{maj}}$), 124.8 (d$_{\text{min}}$), 124.5 (d$_{\text{maj}}$), 35.2 (i$_{\text{min}+\text{maj}}$), 34.6 (i$_{\text{min}+\text{maj}}$), 33.9 (i$_{\text{min}}$), 32.1 (i$_{\text{maj}}$), 20.8 (a), 20.8 (a').

NMR Notes: $^{4.65}$ presents as a pair of hydrazone rotamers under slow rotation in DMSO-$d_6$ with a major:minor ratio of 0.56:0.44; hydrazone rotamers are labelled ‘$\alpha_{\text{min}}$’ and ‘$\alpha_{\text{maj}}$’ for example for the minor and major components respectively. Additionally, amide NH (k) exhibits further rotameric character where its two rotamers (major:minor ratio 0.63:0.37) show different NH chemical shifts (k’ for the major and k’’ for the minor); these rotamers do not permeate further through the spectrum. Rotameric carbon environments that cannot be unambiguously assigned to the major or minor rotamer are assigned (x$_{\text{min}/\text{maj}}$) for example. The proton signal for environment ‘n$_{\text{maj}}$’ shows a cross-peak in HSQC to δ 144.4 ppm but a signal for carbon ‘n’ is not visible here in $^{13}$C NMR. Signals corresponding to environments ‘l’, ‘m’ and ‘n$_{\text{min}}$’ are also not visible in $^{13}$C NMR.

LRESI-{$^+$}-MS: m/z +337.1 [M+Na]$^+$. HRESI-{$^+$}-MS: m/z +337.1383 [M+Na]$^+$; +337.1383 calculated for [C$_{15}$H$_{18}$N$_6$O$_2$Na]$^+$.

4.66: Macrocycle Mimic Model Catalyst

Aldehyde macrocycle $^{4.59}$ (20 mg, 36 µmol, 1 eq.) and DMSO-$d_6$ (0.5 mL) were added to a suspension of hydrazide $^{4.62}$ (7.8 mg, 39 µmol, 1.1 eq.) in CH$_2$Cl$_2$ (1 mL). A solution of TFA in CH$_2$Cl$_2$ (0.1 mL, 0.18 M, 18 µmol, 0.5 eq.) and the resulting mixture was stirred for 3 h at which point the TFA and CH$_2$Cl$_2$ were removed in vacuo. Water (2 mL) was added to the residual DMSO-$d_6$ solution to precipitate the product; the product was isolated by vacuum filtration as a white
amorphous solid (15 mg, 57%).

m.p. 182–184 °C. \(^1\)H NMR (600 MHz, DMF-d$_7$) \(\delta\) 15.24 (s, 0.5H, \(ab\)), 14.66 (s, 0.5H, \(ab'\)), 11.55 (s, 0.5H, \(t\)), 11.24 (s, 0.5H, \(t'\)), 8.84 – 8.64 (br. m, 1H, \(aa+aa'\)), 8.52 – 8.35 (br. m, 0.6H, \(x+x'+x''\)), 8.25 (s, 1H, \(s\)), 8.14 (s, 1H, \(s'\)), 11.55 (s, 0.5H, \(t\)), 11.24 (s, 0.5H, \(t'\)), 8.84 – 8.64 (br. m, 1H, \(aa+aa'\)), 8.52 – 8.35 (br. m, 0.6H, \(x+x'+x''\)), 8.25 (s, 1H, \(p\)), 7.14 (br. s, 2H, \(b\)), 7.09 (d, \(J = 8.4\) Hz, 4H, \(h\)), 6.80 (d, \(J = 8.4\) Hz, 4H, \(i\)), 3.96 (t, \(J = 5.8\) Hz, 4H, \(k\)), 3.81 – 3.71 (m, 2H, \(v\)), 3.11 (br. t, 1H, \(w\)), 2.80 – 2.76 (m, 4H, \(d\)), 2.74 – 2.67 (m, 5H, \(n+w')\), 2.56 (t, \(J = 7.6\) Hz, 4H, \(f\)), 2.00 (p, \(J = 7.7\) Hz, 4H, \(e\)), 1.88 – 1.78 (m, 4H, \(m\)), 1.78 – 1.72 (m, 4H, \(l\)). \(^{13}\)C NMR (151 MHz, DMF) \(\delta\) 174.2 (\(u\)), 162.0 (\(c\)), 158.4 (\(j\)), 147.6 (\(s\)), 144.8 (\(s'\)), 144.1 (\(o\)), 136.0 (\(r\)), 135.9 (\(r'\)), 135.3 (\(g\)), 131.2 (\(p\)), 131.0 (\(p'\)), 130.4 (\(a+h\)), 126.1 (\(q\)), 125.9 (\(q'\)), 121.4 (\(b\)), 115.2 (\(i\)), 68.5 (\(k\)), 38.0 (\(d\)), 36.6 (\(w\)), 33.7 (\(v\)), 32.9 (\(e\)), 29.5 (\(l\)), 29.1 (\(m\)).

**NMR Notes:** \(4.66\) presents as a pair of hydrazone rotamers under slow exchange in DMF-d$_7$ with a major:minor ratio of 0.65:0.35; hydrazone rotamers are labelled \(\alpha\) and \(\alpha'\) for example. Additionally, the amide NH (\(x\)) exhibits further rotameric character where its two rotamers (major:minor ratio cannot be determined due to peak broadening) show different NH chemical shifts (\(x\), \(x'\), \(x''\), \(x'''\)); these rotamers do not permeate further through the spectrum. HSQC indicates \(^{13}\)C signals at 35.2 ppm (\(f\)) and 35.7 ppm (\(n\)) but these signals fall beneath the DMF-d$_7$ peaks. HMBC indicates \(^{13}\)C signals at 174.2 ppm (\(u\)) and 168.5 (\(u'\)) but these signals are not visible in \(^{13}\)C NMR. Signals corresponding to environments ‘\(y'\)’, ‘\(z'\)’ and ‘\(aa'\)’ are also not visible in \(^{13}\)C NMR.

LRESI-(+)-MS: +742.3 [M+H]$^+$. HRESI-(+)-MS: \(m/z\) +742.4064 [M+H]$^+$; +742.4075 calculated for [C$_{46}$H$_{51}$N$_7$O$_4$H]$^+$. 

Chapter 4

217
4.10.1.4 Synthesis of Two-Barrier Rotaxane 4.67

Scheme 46: Synthesis of catalyst-loaded two-barrier rotaxane 4.67. One-barrier aldehyde rotaxane 4.61 was prepared by Dr Sonja Kuschel.\textsuperscript{[21]} Reagents and conditions: a) CuPF\textsubscript{6}-4MeCN, CH\textsubscript{2}Cl\textsubscript{2}, tBuOH, 48 h, r.t., 35%; b) Tail Piece 4.49, CuPF\textsubscript{6}-4MeCN, Tenta-Gel\textsuperscript{TM}-TBTA, CH\textsubscript{2}Cl\textsubscript{2}, tBuOH, r.t., 24 h, 85%; c) Hydrazide 4.62, TFA, CH\textsubscript{2}Cl\textsubscript{2}, DMSO-d\textsubscript{6}, r.t., 3 h, 61%.
4.61: 4-Nitrophenyl acetic acetate-one-barrier aldehyde rotaxane

The synthesis and characterisation of 4.61 was carried out by Dr Sonja Kuschel. To a degassed solution of aldehyde macrocycle 4.59 (100 mg, 0.18 mmol, 1.8 eq.) in CH₂Cl₂ (4 mL), CuPF₆·4MeCN (18.6 mg, 0.05 mmol, 0.5 eq.) is added. The solution is stirred for 30 min at room temperature under N₂ atmosphere. A degassed solution of azide 4.99 (280 mg, 0.5 mmol, 5.0 eq.) and 4.111 (70 mg, 0.10 mmol, 1 eq.) in a mixture of CH₂Cl₂:tBuOH (4:1, 6 mL) was added. The mixture was stirred for 2 d at room temperature under N₂ atmosphere and then diluted with CH₂Cl₂ (20 mL). The organic phase was washed with EDTA·4Na (50% sat. aq., 3 × 15 mL) solution, dried over Na₂SO₄ and the solvent was evaporated under reduced pressure. The crude residue was purified by prep. TLC (SiO₂, 50% EtOAc/pet. ether) to give rotaxane 4.61 (66 mg, 35%) as light yellow film.

1H NMR (600 MHz, Acetone-d₆) δ 9.94 (s, 1Hₐ), 9.47 (s, 1H₁₄), 8.21 (d, J = 8.7 Hz, 2H₃), 8.01 – 7.95 (m, 2H₁₁₂), 7.83 (d, J = 8.4 Hz, 2H₁₇), 7.79 (d, J = 8.3 Hz, 2H₁₀), 7.68 (d, J = 8.6 Hz, 2H₁₆), 7.62 – 7.49 (m, 6H₁₅b,c,n), 7.39 (d, J = 8.1 Hz, 2H₂₂), 7.29 (d, J = 8.4 Hz, 6H₂), 7.26 (d, J = 8.5 Hz, 2H₁₉), 7.16 (d, J = 8.5 Hz, 2H₂₆), 7.11 (d, J = 8.5 Hz, 6H₉), 7.04 (d, J = 8.5 Hz, 2H₂₁), 7.01 (dd, J = 7.8, 2.9 Hz, 2Hₘ), 6.97 (d, J = 8.6 Hz, 2H₁₈), 6.94 (d, J = 8.7 Hz, 2H₄), 6.75 (d, J = 8.2 Hz, 4H₈), 6.46 (d, J = 8.9 Hz, 2H₅), 6.43 (dd, J = 8.6, 2.9 Hz, 4H₉), 5.02 (td, J = 8.1, 5.8 Hz, 2H₁₃), 4.11 (s, 2H₄), 4.04 (t, J = 7.1 Hz, 2H₆), 3.83 – 3.70 (m, 4H₆), 3.46 (t, J = 6.1 Hz, 2H₈), 3.34 (dd, J = 14.1, 5.9 Hz, 1Hₙ), 3.17 (dd, J = 14.0, 8.5 Hz, 1H₁₀), 2.74 – 2.69 (m, 4H₉), 2.62 – 2.57 (m, 4H₉), 2.50 (t, J = 7.1 Hz, 4H₂), 1.91 – 1.75 (m, 10Hₜ₂ₕ), 1.74 – 1.70 (m, 4H₆), 1.64 (s, 6H₂₁), 1.30 (s, 27H₁). 13C NMR (151 MHz, Acetone) δ 193.1, 170.4, 169.7, 167.2, 162.5, 158.0, 157.2, 150.5, 149.0, 148.7, 146.8, 146.5, 145.3, 145.2, 144.7, 142.7, 140.1, 138.1, 137.9, 137.6, 136.5, 135.5, 134.2, 133.8, 132.6, 131.7, 131.4, 131.2, 130.0, 129.1, 128.7, 127.9, 127.9, 127.8, 125.9, 125.1, 125.0, 124.3, 122.2, 122.2, 120.6, 120.2, 119.4, 114.9, 114.1, 68.1, 67.9, 64.8, 63.9, 56.7, 47.5, 40.9, 38.4, 37.8, 35.7, 35.5, 34.9, 32.9, 31.7, 31.0, 29.2, 29.0, 28.8. LRMS(+)-MS: m/z +1878.7 [M+Na]+, HRMS(+)–MS: m/z +1855.9723 [M+H]⁺, +1855.9723 calc. for [C₁₁₉H₁₂₇N₁₀O₁₀] Isotopic distribution for [C₁₁₉H₁₂₇N₁₀O₁₀] ([M+H]⁺):
Chapter 4

+1855.9723 (74%), +1856.9749 (100%), +1857.9784 (72%), +1858.9793 (22%), +1859.9902 (10%), +1860.9586 (4%).

4.60: Two-Barrier, Double Stopper Aldehyde Rotaxane

Aldehyde rotaxane 4.61 (70 mg, 37.7 µmol, 1 eq.) and loaded stoppered terminal barrier 4.49 (62 mg, 45.2 µmol, 1.2 eq.) were dissolved in a degassed mixture of CH$_2$Cl$_2$:tBuOH (19:1, 2 mL). CuPF$_6$·4MeCN (28 mg, 75 µmol, 2 eq.) and Tenta-Gel™-TBTA resin (0.17 mmol/g loading, 665 mg, 113 µmol, 3 eq.) were suspended in a degassed mixture of CH$_2$Cl$_2$:tBuOH (19:1, 4 mL) which was then concentrated by half via nitrogen sparging over 15 min. The rotaxane and terminal barrier solution was added to the catalyst solution with CH$_2$Cl$_2$ (3 × 0.5 mL) being used as a transfer wash (ca. 5 mL total reaction volume, ca. 7.5 mM). The reaction mixture was stirred for 24 h before the catalyst-bearing resin was removed by filtration. The solution was dried over Na$_2$SO$_4$ and concentrated in vacuo. The crude residue was purified by prep. TLC (SiO$_2$, 3 × 1500 µm plate, 4% MeOH/CH$_2$Cl$_2$) to give the title compound as a yellow solid (104 mg, 85%).

$^1$H NMR (600 MHz, Acetone-$d_6$) δ 9.94 (s, 1H, ma), 9.46 (s, 2H, aj+bu), 9.03 (s, 1H, au), 8.60 (d, J = 8.6 Hz, 2H, ag), 8.19 (d, J = 8.4 Hz, 2H, bn), 8.06 – 7.92 (m, 7H, o+v+ax+ay+bb), 7.85 (d, J = 8.7 Hz, 2H, as), 7.82 (d, J = 8.0 Hz, 1H, s), 7.78 (d, J = 8.6 Hz, 4H, r+cd), 7.68 (d, J = 8.6 Hz, 2H, af), 7.64 (d, J = 8.4 Hz, 2H, bm), 7.60 – 7.56 (m, 5H, al+bw+me), 7.56 – 7.51 (m, 3H, mc+ms), 7.48 (d, J = 8.8 Hz, 2H, ar), 7.46 (d, J = 8.7 Hz, 2H, cc), 7.43 (d, J = 8.2 Hz, 2H, bf), 7.40 (d, J = 8.6 Hz, 2H, z), 7.33 – 7.27 (m, 12H, d+cp), 7.23 (d, J = 8.4 Hz, 4H, am+bx), 7.14 – 7.08 (m, 14H, e+ck+co), 7.07 – 7.00 (m, 6H, aa+bg+mr), 6.97 (d, J = 8.9 Hz, 2H, cj), 6.93 (d, J = 8.8 Hz, 2H, i), 6.75 (d, J = 8.4 Hz, 4H, ml), 6.62 (d, J = 8.2 Hz, 1H, bp), 6.46 (d, J = 8.8 Hz, 2H, j), 6.44 – 6.40 (m, 4H, mk), 5.24 (s, 2H, ch), 5.06 – 4.97 (m, 2H, w+bc), 4.68 (p, J = 7.7, 7.1 Hz, 1H, bj), 4.12 (s, 2H, ad), 4.06 (d, J = 7.0 Hz, 2H, n), 3.82 – 3.71 (m, 4H, mi), 3.50 – 3.42 (m, 3H, i+bk), 3.41 – 3.26 (m, 3H, x+bd+bk'), 3.24 – 3.14 (m, 2H, x'+bd'), 2.75 – 2.65 (m, 4H, mf), 2.66 – 2.56 (m, 4H, mp), 2.49 (t, J = 7.1 Hz, 4H, mn), 1.91 – 1.80 (m, 6H, m+mo), 1.80 – 1.75 (m, 4H, mg), 1.75 – 1.68 (m, 16H, ap+ca+mh), 1.34 (s, 9H, bs),
1.29 (ap. s, 54H, a+cs). $^{13}$C NMR (151 MHz, Acetone) δ 192.3, 170.2, 169.6, 168.8, 166.3, 161.6, 157.1, 156.9, 156.5, 156.4, 155.5, 151.7, 151.6, 149.6, 148.3, 148.2, 147.2, 147.0, 146.9, 146.0, 145.6, 145.3, 144.4, 144.3, 143.8, 141.9, 139.8, 139.3, 137.0, 136.8, 135.6, 135.0, 134.7, 133.8, 133.3, 132.0, 131.7, 130.9, 130.7, 130.5, 130.4, 130.3, 129.1, 128.1, 128.0, 127.8, 127.1, 127.0, 125.2, 125.1, 125.0, 124.2, 124.2, 123.4, 123.3, 122.0, 121.4, 121.3, 119.9, 119.8, 119.5, 119.4, 119.3, 114.1, 113.5, 113.2, 78.8, 78.3, 67.0, 63.9, 63.0, 61.3, 55.9, 55.8, 55.0, 46.7, 42.3, 40.1, 37.5, 37.0, 36.8, 34.9, 34.7, 34.0, 32.1, 30.8, 30.0, 29.5, 28.4, 28.2, 27.6. LR(ESI)-(+)-MS: m/z +3233.6 [M+H]$^+$. HR(ESI)-(+)-MS:

![HR(ESI)-(+)-MS of 4.60](image)

Figure 13: HR(ESI)-(+)-MS of 4.60: Observed [M+2H]$^{2+}$ (top); predicted for [C$_{206}$H$_{217}$N$_{17}$O$_{19}$H$_2$]$^{2+}$ (bottom)

4.67: Catalyst-Loaded Two-Barrier, Double Stopper Rotaxane

To two-barrier double stopper aldehyde rotaxane 4.60 (25 mg, 7.7 µmol, 1 eq.) and hydrazide 4.62 (2.30 mg, 11.6 µmol, 1.5 eq.) was added CH$_2$Cl$_2$ (0.5 mL) and DMSO-d$_6$ (0.25 mL). To the resulting suspension was added TFA in CH$_2$Cl$_2$ (0.1 mL, 39 mM, 3.9 µmol, 0.5 eq.) and the resulting solution was stirred for 20 h at room temperature at which point the TFA, CH$_2$Cl$_2$ and DMSO-d$_6$
were removed in vacuo. The dry crude reaction mixture was inerted and dissolved in CH₂Cl₂ (2 mL) before pre-washed Amberlyst® A21 beads (ca. 50 mg) were added. The mixture was left without agitation for 3 h before the beads were removed by filtration and the filtrand concentrated in vacuo. The crude reaction mixture was purified by prep. TLC (SiO₂, 2 × 500 µm Analtech, 4.5% MeOH/CH₂Cl₂) to give the title compound as a light yellow film (16 mg, 4.7 µmol, 61%).

¹H NMR (600 MHz, Acetone-d₆) δ 9.53 (s, 1H, aj/bu), 9.50 (s, 1H, aj/bu), 9.04 (s, 1H, au), 8.60 (s, 1H, cf), 8.23 – 8.17 (m, 4H, ag+bn), 8.11 – 8.06 (m, 1H, ma), 8.06 – 8.01 (m, 3H, ax/ay+vb/bb), 8.00 (s, 1H, o), 7.98 (d, J = 4.9 Hz, 2H, ax/ay), 7.90 (s, 1H, v/bb), 7.87 – 7.81 (m, 6H, r+s+as), 7.79 (d, J = 8.7 Hz, 2H, cd), 7.74 (s, 1H, mc), 7.73 (s, 1H, mc'), 7.68 (d, J = 8.6 Hz, 2H, af), 7.64 (d, J = 8.5 Hz, 2H, bm), 7.61 – 7.58 (m, 4H, al+bw), 7.52 (t, J = 7.7 Hz, 1H, ms), 7.48 (d, J = 8.6 Hz, 2H, ar), 7.46 (d, J = 8.7 Hz, 2H, cc), 7.43 (d, J = 8.5 Hz, 2H, bf), 7.41 (d, J = 8.5 Hz, 2H, z), 7.35 (s, 1H, me), 7.31 (d, J = 8.7 Hz, 6H, d/cp), 7.29 (d, J = 8.7 Hz, 6H, d/cp), 7.25 – 7.21 (m, 4H, am+bx), 7.14 – 7.10 (m, 14H, e+ck+co), 7.07 – 6.99 (m, 6H, aa+bq+mr), 6.97 (d, J = 8.9 Hz, 2H, cj), 6.95 (d, J = 8.8 Hz, 2H, i), 6.75 (d, J = 8.4 Hz, 4H, ml), 6.64 – 6.57 (m, 1H, bp), 6.47 (d, J = 8.8 Hz, 2H, j), 6.44 – 6.41 (m, 4H, mk), 5.24 (s, 2H, ch), 5.07 – 4.98 (m, 2H, w+bc), 4.68 (s, 1H, bj), 4.12 (s, 2H, ad), 4.03 (t, J = 7.0 Hz, 2H, n), 3.81 – 3.75 (m, 4H, mi), 3.75 – 3.69 (m, 2H, dd), 3.47 (dd, J = 14.0, 5.3 Hz, 1H, bk), 3.42 (t, J = 5.9 Hz, 2H, l), 3.40 – 3.27 (m, 3H, x+bd+bk'), 3.24 – 3.16 (m, 2H, x'+bd'), 3.07 (t, J = 6.5 Hz, 2H, dc), 2.64 – 2.56 (m, 8H, mf+mp), 2.50 (t, J = 7.1 Hz, 4H, mn), 2.08 – 2.06 (m, 4H, mo), 1.89 – 1.79 (m, 6H, m+mg), 1.73 – 1.71 (m, 16H, ap+ca+mh), 1.34 (s, 9H, bs), 1.29 (s, 27H, a/cs), 1.28 (s, 27H, a/cs). ¹³C NMR (151 MHz, Acetone) δ 174.4, 173.5, 169.6, 167.4, 167.1, 162.5, 158.0, 157.4, 157.2, 152.4, 150.4, 149.2, 149.0, 147.7, 146.6, 146.0, 145.3, 145.2, 144.5, 144.1, 142.6, 140.6, 140.1, 137.6, 137.3, 136.4, 135.7, 135.0, 134.3, 132.8, 132.6, 131.7, 131.6, 131.5, 131.4, 131.2, 130.0, 129.8, 129.0, 128.8, 128.4, 127.8, 127.6, 126.1, 125.9, 125.1, 124.3, 124.2, 123.9, 122.8, 122.3, 122.1, 120.8, 120.7, 120.6, 120.3, 120.0, 115.7, 115.0, 114.4, 114.2, 113.9, 79.6, 70.8, 68.0, 67.7, 64.7, 63.9, 62.2, 56.9, 55.5, 47.3, 43.2, 43.0, 40.7, 38.4, 37.7, 37.6, 35.6, 35.2, 34.9, 32.9, 32.6, 31.7, 31.6, 30.9, 30.3, 29.7, 29.3, 29.1, 28.4, 25.2, 23.3, 14.4.

NMR Notes: ¹H signals for NH environments ‘da’, ‘de’ and ‘di’ as well as position ‘dh’ were not detected; they were therefore not unambiguously assigned due to their weak, broad resonances and overlapping signals. Not all ¹³C NMR resonances were visible.
LR(ESI)-{+}-MS: m/z +3413.2 [M+H]^+;

Figure 14: LR(ESI)-{+}-MS of 4.67 [M+H]^+: Observed (top); Predicted for [C_{212}H_{225}N_{23}O_{20}H]^+ (bottom)

4.10.1.5 Synthesis of Trityl-Protected Amide Macrocycle 4.93

Scheme 47: Synthesis of trityl-protected catalytic amide macrocycle 4.93. Reagents and conditions: a) Catalyst ester 4.72, Trt-Cl, NEt, DMF, -10 °C – r.t., 18 h, 94%; b) LiOH, H_2O, THF, 45 °C, 15 h, 92%; c) LiAlH_4, THF, -10 °C – r.t., 20 h, 99%, see Chapter 3; d) EDCI-HCl, HOBT-H_2O, DIPEA, DMF, r.t., 12 h, 99%.
4.95: Trityl-Catalyst Methyl Ester

To a solution of catalyst methyl ester 4.72 (200 mg, 1.01 mmol, 1 eq.) in DMF (10 mL) at -10 °C was added NEt₃ (1 mL, 726 mg, 7.2 mmol, 7 eq.) followed by trityl chloride (335 mg, 1.2 mmol, 1.2 eq.). The resulting mixture was stirred at -10 °C for 30 min before being warmed to room temperature and stirred for a further 18 h. The reaction mixture was concentrated in vacuo to dryness and then taken up in EtOAc (25 mL). The organic solution was washed with aq. LiCl (5%, 20 mL), NaHCO₃ (sat. aq., 20 mL), brine (20 mL) before the organics were dried over MgSO₄. Solvent was removed in vacuo to give a crude white solid foam which was purified by flash column chromatography (SiO₂, 0–2% MeOH/CH₂Cl₂) to give the title compound as a white solid (417 mg, 94 %) which can be recrystallized from n-hexane/EtOAc (78% recryst. recovery).

m.p. 140–143 °C. ¹H NMR (600 MHz, DMSO-d₆) δ 8.60 (t, J = 5.9 Hz, 1H, e), 8.30 (s, 1H, i), 7.44 – 7.37 (m, 9H, m+n), 7.10 – 7.04 (m, 6H, l), 3.59 (s, 3H, a), 3.46 (q, J = 7.0 Hz, 2H, d), 2.58 (t, J = 7.1 Hz, 2H, c). ¹³C NMR (151 MHz, DMSO) δ 171.8 (b), 158.6 (f), 156.7 (g), 146.7 (i), 141.4 (k), 129.5 (l), 128.3 (m), 78.1 (j), 51.4 (a), 34.9 (d), 33.4 (c). LR(ESI)-(-)-MS: m/z +243.2 [M+2Na]⁺; +463.2 [M+Na]⁺; +903.3 [2M+Na]⁺. HRFT(+) -MS: m/z +463.1736 [M+Na]⁺; +463.1741 calc. for [C₂₆H₂₄N₄O₃Na]⁺.

4.96: Trityl-Catalyst Carboxylic Acid

Trityl-catalyst methyl ester 4.95 (150 mg, 340 µmol, 1 eq.) and lithium hydroxide (41 mg, 1.70 mmol, 5 eq.) were dissolved in a mixture of THF (10 mL) and water (0.1 mL). The solution was stirred at 45 °C for 15 h before being neutralised with Dowex® 50WX8-100 ion exchange resin which was subsequently removed via filtration over Celite®. The filtrate dried over Na₂SO₄ and then concentrated to dryness. The crude residue was purified by flash column chromatography (SiO₂, 0–10% MeOH/CH₂Cl₂) to give the title compound as a white solid (134 mg, 92%).

m.p. decomposes above 83 °C. ¹H NMR (600 MHz, Chloroform-d) δ 10.49 (br. s, 1H, a), 7.99 (s, 1H, h), 7.98 (br. t, J = 6.0 Hz, 1H, e) 7.38 – 7.30 (m, 9H, l+m), 7.11 (d, J = 6.8 Hz, 6H, k), 3.74 (q, J = 6.0 Hz, 2H, d), 2.69 (t, J = 5.8 Hz, 2H, c). ¹³C NMR (151 MHz, Chloroform) δ 175.0 (b), 159.2 (f), 156.5 (g), 146.5 (h), 141.4 (j), 130.2 (k), 128.7 (m), 128.3 (l), 89.5 (i), 35.0 (d), 33.8 (c). LR(ESI)-(+) -MS: m/z -425.2 [M-H]. HRFT(+) -MS: m/z +449.1582 [M+Na]⁺; +449.1584 calc. for [C₂₅H₂₂N₄O₃Na]⁺.
4.93: Trityl-Catalyst Amide Macrocyle

Amine macrocycle 4.97 (144 mg, 256 µmol, 1 eq.), trityl-catalyst carboxylic acid 4.96 (109 mg, 256 µmol, 1 eq.), EDCI·HCl (73 mg, 383 µmol, 1.5 eq.) and HOBt·H₂O (59 mg, 383 µmol, 1.5 eq.) were suspended in DMF (10 mL) before DIPEA (133 µL, 99 mg, 767 µmol, 3 eq.) was added and the reaction mixture stirred for 12 h. The reaction mixture was then concentrated to dryness in vacuo and the residue was taken up in CH₂Cl₂ (30 mL) and washed with aq. LiCl (5%, 2 × 20 mL). The organics were dried over Na₂SO₄ and concentrated to dryness in vacuo and the crude residue was purified by flash column chromatography (SiO₂, 0–4% MeOH/CH₂Cl₂) to give the title compound as a white solid foam (246 mg, 99%).

m.p. 72–75 °C. ¹H NMR (600 MHz, Acetone-d₆) δ 8.14 (s, 1H, aa), 8.13 (t, J = 5.7 Hz, 1H, x), 7.60 (t, J = 5.5 Hz, 1H, t), 7.56 (t, J = 7.6 Hz, 1H, a), 7.43 – 7.36 (m, 9H, ae+af), 7.20 – 7.14 (m, 6H, ad), 7.07 (d, J = 8.6 Hz, 4H, h), 7.05 (d, J = 7.7 Hz, 2H, b), 7.03 (s, 1H, p), 6.96 (s, 2H, q), 6.74 (d, J = 8.6 Hz, 4H, i), 4.37 (d, J = 5.8 Hz, 2H, s), 3.90 (t, J = 6.0 Hz, 4H, k), 3.65 (q, J = 6.3 Hz, 2H, w), 2.76 (t, J = 7.4 Hz, 4H, d), 2.62 (t, J = 7.3 Hz, 4H, n), 2.58 – 2.51 (m, 6H, f+v), 2.04 – 1.99 (m, 4H, e), 1.82 – 1.73 (m, 4H, m), 1.73 – 1.67 (m, 4H, l). ¹³C NMR (151 MHz, Acetone) δ 171.6 (u), 161.9 (c), 159.3 (y), 158.2 (j+z), 147.4 (aa), 143.3 (o), 142.8 (ac), 140.3 (r), 137.3 (a), 135.2 (g), 130.8 (ad), 130.1 (h), 129.2 (af), 129.0 (ae), 127.8 (p), 126.2 (q), 120.9 (b), 115.0 (i), 79.5 (ab), 68.1 (k), 43.5 (s), 37.9 (d), 36.3 (w), 35.9 (n), 35.7 (v), 35.1 (f), 32.6 (e), 29.2 (l), 28.8 (m). LR(ESI)-(+)-MS: m/z +971.5 [M+H]+. HRFT(+)-MS: m/z +971.5211 [M+H]+; +971.5218 calc. for [C₆₃H₆₇N₆O₄H]+.
4.10.1.6 Synthesis of Two-Barrier Rotaxane 4.104

Unloaded azide-alkyne barrier 4.89 (220 mg, 405 μmol, 1 eq.) was dissolved in CH₂Cl₂ (10 mL) and cooled to -10 °C. Phenylacetyl chloride (107 μL, 125 mg, 810 μmol, 2 eq.) was added followed by NEt₃ (167 μL, 121 mg, 1.20 mmol, 3 eq.) and the resulting solution warmed to room temperature and stirred for 3 h. The resulting solution was...
diluted with CH$_2$Cl$_2$ (20 mL) and washed with NaHCO$_3$ (sat. aq., 20 mL) followed by NH$_4$Cl (sat. aq., 20 mL). The organics were dried over MgSO$_4$ and solvent was removed in vacuo. The crude residue was purified by flash column chromatography (SiO$_2$, EtOAc/pet. ether) to give the title compound as yellow amorphous solid (173 mg, 65%).

m.p. decomposes above 150 °C. 1H NMR (600 MHz, Acetone-$d_6$) δ 9.40 (s, 1H, v), 8.02 (d, $J = 8.0$ Hz, 1H, h), 7.86 (d, $J = 8.4$ Hz, 2H, e), 7.57 – 7.50 (m, 4H, d+x), 7.40 – 7.33 (m, 6H, l+r+s), 7.31 – 7.26 (m, 3H, t+ad), 7.17 (d, $J = 8.7$ Hz, 2H, y), 7.01 (d, $J = 8.5$ Hz, 2H, m), 6.99 (d, $J = 8.7$ Hz, 2H, ae), 4.98 (td, $J = 8.2$, 6.3 Hz, 1H, i), 3.89 (s, 2H, p), 3.83 (s, 1H, a), 3.34 (dd, $J = 14.0$, 5.7 Hz, 1H, j), 3.16 (dd, $J = 14.0$, 8.7 Hz, 1H, j’), 1.65 (s, 6H, ab). 13C NMR (151 MHz, Acetone) δ 170.5 (o), 170.4 (u), 166.8 (g), 150.7 (n), 148.7 (ac), 146.6 (z), 138.1 (af), 137.4 (w), 136.2 (p), 135.2 (q), 132.7 (d), 131.1 (l), 129.3 (s), 129.1 (ad), 128.4 (e), 127.9 (t), 127.8 (y), 126.2 (c), 122.3 (m), 120.2 (x), 119.4 (ae), 83.5 (b), 81.4 (a), 56.8 (l), 42.9 (ac), 41.4 (p), 37.8 (j), 31.0 (ab). LR(ESI)-(+)-MS: m/z +684.3 [M+Na$^+$]. HRFT-(+)-MS: m/z +662.2760 [M+H$^+$]; +662.2762 calc. for [C$_{41}$H$_{35}$N$_5$O$_4$H]$^+$. 4.100 (+ 4.101): Phenylacetic Acetate One-Barrier Rotaxane (and Free Thread)

Trityl-catalyst amide

macrocyle 4.93 (88 mg, 90.7 µmol, 1.5 eq.) and CuPF$_6$·4MeCN (11.3 mg, 30.2 µmol, 0.5 eq.) were dissolved in CH$_2$Cl$_2$ (1.5 mL) and the resulting solution was both degassed and concentrated by half via nitrogen sparging. Azide stopper 4.99 (213 mg, 363 µmol, 6 eq.) and phenylacetic acid-loaded barrier 4.91 (40 mg, 60.4 µmol, 1 eq.) were dissolved in CH$_2$Cl$_2$ (1.5 mL) and the resulting solution was also degassed and concentrated by half via nitrogen sparging. The solution containing the macrocycle and copper catalyst was added to the solution containing barrier and stopper with CH$_2$Cl$_2$ (0.5 mL) being used as a transfer wash. The resulting mixture was degassed via nitrogen sparging for 5 min further to give a total reaction volume of ca. 2 mL (c = 30 mM w.r.t. the barrier/limiting reagent) and stirred at room temperature for 22 h. The resulting solution was diluted with CH$_2$Cl$_2$ (10 mL) and washed with EDTA-4Na (50% sat. aq.,
3 × 10 mL), dried over Na₂SO₄ and concentrated to dryness in vacuo. The crude residue was purified by flash column chromatography (SiO₂, EtOAc/pet. ether gradient: 10% [stopper 4.99] – 40% [free thread 4.101] – 55–70% [rotaxane product 4.100] – 100% [macrocycle 4.93]) to give the title rotaxane (91 mg, 68%) as a light yellow amorphous solid.

Free thread 4.101 (4 mg, 5% w.r.t. barrier) was also isolated as a thin film alongside recovered excess stopper 4.99 (160 mg, 90% of excess).

This procedure was initially carried out with trityl-catalyst amide macrocycle 4.93 (211 mg, 218 µmol, 1.8 eq.), CuPF₆·4MeCN (23 mg, 61 µmol, 0.5 eq.), azide stopper 4.99 (285 mg, 484 µmol, 4 eq.) and phenylacetic acid-loaded barrier 4.91 (80 mg, 121 µmol, 1 eq.) over 18 h at the same concentration. Rotaxane 4.100 (153 mg, 57%) was isolated alongside free thread 4.101 (7.5 mg, 5%), excess stopper 4.99 (210 mg, 98% of excess) and macrocycle 4.93 (60 mg, 64% of excess).

Data for rotaxane 4.100: ¹H NMR (600 MHz, Acetone-d₆) δ 9.94 (s, 1H, aj), 8.41 (t, J = 6.3 Hz, 1H, br), 8.17 (s, 1H, bu), 7.93 – 7.87 (m, 2H, v+bn), 7.86 (s, 1H, o), 7.63 (d, J = 8.7 Hz, 2H, al), 7.54 (t, J = 7.7 Hz, 1H, au), 7.49 (ap. s, 4H, r+s), 7.38 (d, J = 7.6 Hz, 2H, af), 7.37 – 7.26 (m, 19H, d+aq+ar+by+bx), 7.25 (t, J = 8.4 Hz, 2H, z), 7.14 – 7.07 (m, 15H, e+ah+am+bx), 7.03 (d, J = 7.7 Hz, 2H, av), 7.00 (d, J = 8.9 Hz, 2H, i), 6.99 – 6.96 (m, 4H, as+bk), 6.95 (d, J = 8.3 Hz, 2H, aa), 6.92 (s, 1H, bj), 6.78, (d, J = 8.0 Hz, 2H, bb), 6.77 (d, J = 8.3 Hz, 2H, bb’), 6.58 (d, J = 8.6 Hz, 2H, j), 6.35 (d, J = 8.3 Hz, 2H, bc), 6.34 (d, J = 8.3 Hz, 2H, bc’), 5.09 (td, J = 8.6, 5.3 Hz, 1H, w), 4.41 (dd, J = 14.9, 5.8 Hz, 1H, bm), 4.35 (dd, J = 14.8, 5.7 Hz, 1H, bm’), 4.11 (t, J = 7.0 Hz, 2H, n), 3.89 (s, 2H, ad), 3.79 – 3.72 (m, 1H, bq), 3.72 – 3.60 (m, 5H, be+bq’), 3.54 (t, J = 6.1 Hz, 2H, i), 3.17 (dd, J = 14.0, 5.3 Hz, 1H, x), 3.00 (dd, J = 14.0, 9.1 Hz, 1H, x’), 2.74 – 2.65 (m, 2H, bp+bp’), 2.65 – 2.60 (m, 4H, ax), 2.59 – 2.44 (m, 8H, az+bh), 1.97 – 1.90 (m, 2H, m), 1.90 – 1.83 (m, 4H, ay), 1.70 – 1.59 (m, 14H, ap+bf+bg), 1.29 (s, 27H, a). ¹³C NMR (151 MHz, Acetone) δ 171.5 (bo), 170.5 (ac), 170.3 (ai), 167.1 (u), 162.5 (aw), 159.4 (bs), 158.6 (bt), 158.0 (bd), 157.4 (k), 150.6 (ab), 149.1 (c), 147.7 (bu), 146.8 (p), 146.3 (an), 145.3 (f), 143.7 (bi), 142.7 (bw), 140.3 (h), 138.1 (ak+at), 137.6 (au), 136.2 (y), 135.1 (ae), 134.9 (q/t), 134.2 (ba), 133.3 (q/t), 132.7 (i), 131.4 (e), 131.2 (z), 130.8 (bx), 130.3 (af), 129.9 (bb), 129.9 (bb’), 129.4 (ag), 129.2 (ar+bx), 129.0 (by), 128.4 (r/s), 127.9 (ah), 127.7 (am), 126.1 (bk), 125.9 (r/s), 125.1 (d), 122.2 (aa), 122.0 (o), 120.6 (av), 120.1 (al), 119.4 (as), 114.9 (bc), 114.9 (bc’), 114.1 (j), 68.2 (be), 64.8 (l), 63.4 (g), 56.1 (w), 47.4 (n), 43.6 (bm), 42.9 (ao), 41.5 (ad), 38.4 (x+ax), 36.7 (bq), 36.3 (bh), 36.1 (bp), 35.6 (az), 34.9 (b), 33.0 (ay), 31.7 (a), 31.0 (ap), 30.3 (m), 29.4 (bf/bg), 29.3 (bf/bg); ¹³C NMR signals for carbons aq, bj, bl and bv were not observed.

LR(ESI)-(+)MS:
Figure 15: LR(ESI)-(+)-MS: Observed (top); Predicted (bottom) for \([C_{144}H_{150}N_{14}O_9H]^+\) corresponding to \([M+H]^+\) and \([C_{144}H_{150}N_{14}O_9Na]^+\) corresponding to \([M+Na]^+\). The observed (top) spectrum also shows a weak signal corresponding to \([2M+H+Na]^{2+}\) centred around \(m/z = 2232\) where \(z = 2\).

HRFT(+)-MS: \(m/z +2220.1738\) (60%), +2221.1809 (100%) \([M+H]^+\); +2220.1783 (60%), +2221.1813 (100%) calc. for \(C_{144}H_{150}N_{14}O_9H^+\).

**Data for free thread 4.101:**

\(^1\)H NMR (600 MHz, Acetone-\(d_6\))
δ 9.42 (s, 1H, aj), 8.48 (s, 1H, o), 7.97 (d, \(J = 8.0\) Hz, 1H, v), 7.93 (ap. s, 4H, r+s), 7.54 (d, \(J = 8.6\) Hz, 2H, al), 7.40 – 7.36 (m, 4H, z+af), 7.34 (t, \(J = 7.7\) Hz, 2H, ag), 7.31 (d, \(J = 8.6\) Hz, 6H, d), 7.29 – 7.25 (m, 3H, ah+ar), 7.16 (d, \(J = 8.6\) Hz, 2H, am), 7.12 (d, \(J = 8.5\) Hz, 6H, e), 7.09 (d, \(J = 8.8\) Hz, 2H, i), 7.01 (d, \(J = 8.5\) Hz, 2H, aa), 6.98 (d, \(J = 8.6\) Hz, 2H, as), 6.84 (d, \(J = 8.8\) Hz, 2H, j), 4.99 (td, \(J = 8.4, 5.9\) Hz, 1H, w), 4.69 (t, \(J = 6.9\) Hz, 2H, n), 4.07 (t, \(J = 5.9\) Hz, 2H, l), 3.89 (s, 2H, ad), 3.34 (dd, \(J = 14.0, 5.8\) Hz, 1H, x), 3.18 (dd, \(J = 14.0, 8.6\) Hz, 1H, x'), 2.45 (p, \(J = 6.4\) Hz, 2H, m), 1.65 (s, 6H, ap), 1.29 (s, 27H, a). \(^{13}\)C NMR

229
(151 MHz, Acetone) $\delta$ 170.5 (ac), 170.4 (ai), 167.2 (u), 157.6 (k), 150.7 (ab), 149.2 (c), 148.7 (aq), 147.0 (q), 146.5 (an), 145.2 (f), 140.5 (h), 138.1 (at), 137.5 (ak), 136.3 (y), 135.3 (p/t), 135.1 (ae), 134.1 (p/t), 132.8 (l), 131.4 (e), 131.1 (z), 130.3 (af), 129.3 (ag), 129.2 (ar), 128.9 (r/s), 127.9 (ah), 127.8 (am), 125.9 (r/s), 125.1 (d), 122.5 (o), 122.3 (aa), 120.2 (al), 119.4 (as), 114.1 (j), 65.2 (l), 63.9 (g), 56.7 (w), 47.9 (n), 42.9 (ao), 41.4 (ad), 37.8 (x), 34.9 (b), 31.6 (a), 31.0 (ap), 30.8 (m).

$\mu$H NMR (600 MHz, Acetone-d$_6$) $\delta$ 9.96 (s, 1H, bj), 9.48 (s, 1H, ct), 9.02 (s, 1H, bu), 8.60 (s, 1H, de), 8.42 (t, $J = 6.1$ Hz, 1H, ep), 8.19 (d, $J = 8.3$ Hz, 2H, cn), 8.16 (s, 1H, es), 8.06 – 8.02 (m, 3H, bx/by+cb), 7.97 (d, $J = 8.2$ Hz, 2H, bx/by), 7.95 – 7.89 (m, 2H, av+el), 7.86 – 7.82 (m, 3H, ao+bs), 7.78 (d, $J = 8.7$ Hz, 2H, dc), 7.68 – 7.62 (m, 4H, bl+cm), 7.59 (d, $J = 8.8$ Hz, 2H, cv), 7.53 (t, $J = 7.7$

4.104: Trityl Protected Catalyst-Loaded Two-Barrier, Double Stopper Rotaxane

CuPF$_6$-4MeCN (21.8 mg, 58.5 µmol, 2 eq.) and Tenta-Gel™-TBTA resin (0.17 mmol/g loading, 516 mg, 88 µmol, 3 eq.) were suspended in CH$_2$Cl$_2$ (3 mL) and were both degassed and concentrated by a third via nitrogen sparging over 15 min. A solution of phenylacetic acetate one-barrier rotaxane 4.100 (65 mg, 29.3 µmol, 1 eq.) and loaded stoppered terminal barrier 4.49 (48 mg, 35.1 µmol, 1.2 eq.) dissolved in CH$_2$Cl$_2$ (1 mL) was added to the suspension of resin-supported catalyst with CH$_2$Cl$_2$ (2 × 0.5 mL) being used as a transfer wash (4 mL total reaction volume, ca. 7.5 mM). The reaction mixture was stirred for 26 h at room temperature before the catalyst-bearing resin was removed by filtration. The resulting CH$_2$Cl$_2$ solution was washed with EDTA-4Na (sat. aq., 2 × 5 mL) before the organics were dried over Na$_2$SO$_4$ and concentrated in vacuo. The crude residue was purified by prep. TLC (SiO$_2$, 2 × 1000 µm plate, 25% EtOAc/CH$_2$Cl$_2$; then SiO$_2$, 2 × 1000 µm plate, 4% EtOAc/CH$_2$Cl$_2$) to give the title compound as a yellow glass (82 mg, 78%).
Hz, 1H, ds), 7.51 – 7.44 (m, 8H, ar+as+br+db), 7.42 (d, J = 8.1 Hz, 2H, cf), 7.38 (d, J = 7.5 Hz, 2H, bf),
7.35 – 7.24 (m, 25H, ad+az+bg+do+ew+ex), 7.22 (d, J = 8.7 Hz, 2H, cw), 7.16 (d, J = 8.6 Hz, 2H, bm),
7.14 – 7.07 (m, 21H, ae+ah+dj+dn+ev), 7.04 – 6.93 (m, 10H, ai+ba+cg+di+ei), 6.91 (s, 1H, eh), 6.79 – 6.74 (m, 4H, dz+dz’), 6.63 (d, J = 8.3 Hz, 1H, cp), 6.58 (d, J = 8.7 Hz, 2H, aj), 6.37 – 6.31 (m, 4H, ea+ea’), 5.24 (s, 2H, dg), 5.09 (q, J = 7.8 Hz, 1H, aw), 5.03 (q, J = 7.7 Hz, 1H, cc), 4.67 (q, J = 8.4, 7.8 Hz, 1H, ej), 4.40 (dd, J = 15.1, 5.7 Hz, 2H, ek), 4.34 (dd, J = 15.0, 5.7 Hz, 1H, ek’), 4.10 (t, J = 7.0 Hz, 2H, an), 3.88 (s, 2H, bd), 3.78 – 3.71 (m, 1H, eo), 3.71 – 3.59 (m, 5H, ec+eo’), 3.54 (t, J = 6.1 Hz, 2H, al), 3.46 (dd, J = 14.0, 5.4 Hz, 1H, ck), 3.37 (dd, J = 14.0, 5.8 Hz, 1H, cd), 3.30 (dd, J = 13.9, 9.6 Hz, 1H, ck’), 3.24 – 3.14 (m, 2H, ax+cd’), 3.01 (dd, J = 14.1, 9.1 Hz, 1H, ax’), 2.74 – 2.59 (m, 6H, dv+en+en’), 2.57 – 2.46 (m, 8H, dx+ef), 1.95 – 1.90 (m, 2H, am), 1.90 – 1.82 (m, 4H, dw), 1.72 (ap.
s, 12H, bp+cz), 1.68 – 1.58 (m, 8H, ed+ee), 1.34 (s, 9H, cr), 1.29 (s, 27H, aa/dr), 1.28 (s, 27H, aa/dr). 13C NMR (151 MHz, Acetone) δ 171.6, 171.1, 170.9, 170.5, 170.4, 167.2, 162.5, 160.3, 158.0, 157.4, 156.1, 152.7, 152.4, 150.5, 149.2, 149.1, 147.9, 147.8, 146.8, 146.5, 146.2, 145.9, 145.3, 145.2, 143.7, 142.6, 140.7, 140.3, 140.0, 137.6, 136.6, 136.2, 135.8, 135.7, 135.1, 134.9, 134.7, 134.2, 132.8, 132.7, 131.5, 131.4, 131.3, 131.1, 130.8, 130.3, 129.9, 129.4, 129.2, 129.0, 128.4, 127.9, 127.8, 126.1, 125.9, 125.1, 124.2, 122.8, 122.2, 122.0, 120.8, 120.7, 120.6, 120.4, 120.3, 120.2, 114.9, 114.4, 114.1, 79.7, 68.1, 64.8, 63.9, 62.2, 60.5, 47.4, 43.6, 43.2, 41.5, 38.4, 36.3, 36.1, 35.6, 34.9, 33.0, 31.7, 30.9, 28.5. This molecule contains 122 carbon environments: 103 environments only were observed due to overlapping peaks of similar environments. It is also assumed that some environments will not be observed by 13C NMR as has been the case in precursors to this molecule (vide supra). LR(ESI)−(+)−MS: m/z +3597.7 [M+H]+, +3620.1 [M+Na]+ and +3609.9 (centre) [2M+H+Na]+2. HR(ESI)−(+)−MS:

Figure 16: HR(ESI)−(+)−MS: Observed (top); Predicted for [C233H245N21O21H2]2+ (bottom).
4.10.1.7 Unambiguous Synthesis of Analysis Standards

Scheme 49: Unambiguous synthesis of analysis standards for operation and model operation monitoring. Reagents and conditions: a) SOCl₂, MeOH, 0 °C – r.t., 16 h, 95%; b) 4-nitophenylacetic acid, EDCI·HCl, HOBt·H₂O, NEt₃, CH₂Cl₂, r.t., 18.5 h, quant.; c) LiOH, H₂O, MeOH, r.t., 2 h, then HCl (1.25 M in MeOH), 33%; d) Phenylacetyl chloride, NEt₃, CH₂Cl₂, -10 °C – r.t., 14 h, 99%; e) LiOH, H₂O, THF, 40 °C, 18 h, then DOWEX®50WX8-100, 72%; f) Fmoc-(4-NO₂)Phe-OH (4.75), EDCI-HCl, HOBt-H₂O, NEt₃, CH₂Cl₂, 0 °C – r.t., 18 h, 88%; g) HNET₂, CH₂Cl₂, MeOH, r.t., 18 h, 36%; h) HNET₂, EDCI-HCl, HOBt-H₂O, DIPEA, DMF, r.t., 16 h, 99%; i) HNET₂, NEt₃, CH₂Cl₂, 0 °C – r.t., 15 h, 99%.

4.71: 4-Nitrophenylalanine Methyl Ester Hydrochloride

N-Boc-4-nitro-L-phenylalanine (0.50 g, 1.61 mmol) was dissolved in anhydrous MeOH (20 mL, 80 mM) and cooled to 0 °C. Thionyl chloride (0.7 mL, 1.14 g, 9.59 mmol, 6 eq.) was added dropwise and the resulting solution was allowed to warm to room temperature and stirred for 16 h. The reaction mixture was then concentrated in vacuo and the resulting solid was suspended in NaHCO₃ (sat. aq., 50 mL). The aqueous suspension was extracted with CH₂Cl₂ (5 × 10 mL) and the combined organics were dried over MgSO₄ before solvent was removed in vacuo. The residue was dissolved in Et₂O (10 mL) and HCl (1M in Et₂O, 2 mL) was added to precipitate the product which was isolated by vacuum filtration as a white solid (0.40 g, 95%).
m.p. 204–206 °C. [α]_D^{20} = +5.42 ° (c = 1.00, MeOH). ¹H NMR (600 MHz, DMSO-d_6) δ 8.73 – 8.56 (m, 3H, a), 8.21 (d, J = 8.6 Hz, 2H, f), 7.57 (d, J = 8.6 Hz, 2H, e), 4.41 (t, J = 6.7 Hz, 2H, b), 3.69 (s, 3H, i), 3.30 (dd, J = 14.3, 6.5 Hz, 1H, c), 3.26 (dd, J = 14.1, 7.3 Hz, 1H, c'). ¹³C NMR (151 MHz, DMSO) δ 169.1 (h), 146.9 (g), 143.0 (d), 131.0 (e), 123.6 (f), 52.8 (l), 52.7 (b), 35.4 (c). LR(ESI)-(+) -MS: m/z +225.1 [M-Cl]^+. HR(ESI)-(+) -MS: m/z +225.0873 [M-Cl]^+; +225.0870 calculated for [C₁₀H₁₃O₃N₃]^+.

4.34: Operation Reference Product – Methyl Ester

4-Nitrophenylalanine methyl ester hydrochloride 4.71 (101 mg, 387 µmol, 1 eq.), 4-nitrophenylacetic acid (84 mg, 465 µmol, 1.2 eq.), EDCI·HCl (96 mg, 503 µmol, 1.3 eq.) and HOBT·H₂O (12 mg, 77 µmol, 0.2 eq.) were suspended in CH₂Cl₂ (3.8 mL, 100 mM) and NEt₃ (0.27 mL, 195 mg, 1.94 mmol, 5 eq.) was added dropwise. The reaction mixture was stirred for 2.5 h before a further portion of EDCI·HCl (96 mg, 503 µmol, 1.3 eq.) was added and the reaction mixture stirred for 16 h. The solution reaction mixture was diluted with CHCl₃:IPA (3:1, 20 mL) and washed with aq. HCl (0.1 M, 20 mL). The organics were dried over MgSO₄ and solvent was removed under reduced pressure. The residue was purified by flash column chromatography (SiO₂, 1% MeOH/CH₂Cl₂) to give the title compound as a white solid (150 mg, quant.).

m.p. 192–194 °C. [α]_D^{20} = +74.54 ° (c = 1.00, CH₂Cl₂). ¹H NMR (600 MHz, Chloroform-d) δ 8.21 (d, J = 8.7 Hz, 2H, b), 8.10 (d, J = 8.7 Hz, 2H, l), 7.41 (d, J = 8.7 Hz, 2H, c), 7.19 (d, J = 8.7 Hz, 2H, k), 6.03 (d, J = 7.2 Hz, 1H, g), 4.93 – 4.88 (m, 1H, h), 3.76 (s, 3H, o), 3.65 (d, J = 15.2 Hz, 1H, e), 3.62 (d, J = 15.2 Hz, 1H, e'). ¹³C NMR (151 MHz, Chloroform) δ 171.4 (l), 168.9 (f), 147.5 (a/m), 147.4 (a/m), 143.5 (j), 141.6 (d), 130.3 (c), 130.2 (k), 124.1 (b), 123.8 (l), 53.2 (h), 53.0 (o), 43.1 (e), 37.8 (i). LR(ESI)-(--) -MS: m/z -386.1 [M-H]^-. HRFT(--)-MS: m/z -386.0998 [M-H], -386.0994 calc. for [C₁₃H₁₅N₃O₃]^-

4.39: Operation Reference Product – Carboxylic Acid

A solution of LiOH (26 mg, 1.08 mmol, 5 eq.) in water (0.7 mL) was added to a solution of methyl ester 4.71 (82.5 mg, 213 µmol, 1 eq.) dissolved in MeOH (1.3 mL). The resulting purple mixture was stirred at room temperature for 2 h before being neutralised with HCl (1.25 M, MeOH) resulting in a clear, light yellow solution. The residue was concentrated to dryness in vacuo and the residue triturated with MeOH (2 mL). The resulting slurry was filtered and the white precipitate was further purified by flash column chromatography (SiO₂, 0.5% formic acid + 5% MeOH/CH₂Cl₂) to give the title compound as a white solid (26 mg, 33%).

233
m.p. 176–179 °C. [α]$_D^{20}$ = +66.81 ° (c = 1.00, 1% MeOH/CH$_2$Cl$_2$). $^1$H NMR (600 MHz, Acetone-$d_6$) δ 8.13 (d, J = 8.5 Hz, 2H, b), 8.07 (d, J = 8.5 Hz, 2H, l), 7.71 (d, J = 7.6 Hz, 1H, g), 7.51 (d, J = 8.5 Hz, 2H, c), 7.48 (d, J = 8.5 Hz, 2H, k), 4.85 (td, J = 8.6, 5.2 Hz, 1H, h), 3.71 (d, J = 14.4 Hz, 1H, e), 3.67 (d, J = 14.4 Hz, 1H, e'), 3.38 (dd, J = 13.9, 5.0 Hz, 1H, i), 3.14 (dd, J = 13.8, 8.9 Hz, 1H, i'). $^{13}$C NMR (151 MHz, Acetone) δ 172.5 (n), 169.8 (f), 147.8 (a+m), 146.3 (j), 144.7 (d), 131.4 (k), 131.2 (c), 124.0 (d+j), 53.6 (h), 42.9 (e), 37.9 (l). LR(ESI)-(+) -MS: m/z -372.0 [M-H]$^+$; HRFT(-)-MS: m/z -372.0823 [M-H]$^+$; -372.0837 calc. for [C$_{18}$H$_{14}$N$_2$O$_7$].

4.108: Operation Product Methyl Ester

4-Nitrophenylalanine methyl ester hydrochloride 4.71 (1.00 g, 3.84 mmol, 1 eq.) was suspended in CH$_2$Cl$_2$ (30 mL, 128 mM) and NEt$_3$ (1.65 ml, 1.20 g, 11.55 mmol, 3 eq.) was added. The resulting clear solution was cooled to -10 °C and phenylacetyl chloride (0.61 mL, 0.71 g, 4.60 mmol, 1.2 eq.) was added dropwise over 5 min. The resulting solution was stirred at -10 °C for 5 min before being allowed to warm to room temperature and stirred for 14 h. The resulting slurry was diluted with CH$_2$Cl$_2$ (50 mL) and washed with water (50 mL), NaHCO$_3$ (sat. aq., 50 mL) and aq. HCl (1 M, 50 mL). The organics were then dried over MgSO$_4$ and concentrated to dryness in vacuo. The crude yellow solid was purified by flash column chromatography (SiO$_2$, 5–10% EtOAc/CH$_2$Cl$_2$) to furnish the title compound as a white solid (1.30 g, 99%).

m.p. 127–129 °C. [α]$_D^{20}$ = +39.75 ° (c = 1.00, CH$_2$Cl$_2$). $^1$H NMR (600 MHz, Chloroform-$d$) δ 8.01 (d, J = 8.7 Hz, 2H, l), 7.38 – 7.30 (m, 3H, a+b), 7.22 – 7.18 (m, 2H, c), 7.03 (d, J = 8.6 Hz, 2H, k), 5.85 (d, J = 7.3 Hz, 1H, g), 4.90 (dt, J = 7.5, 5.8 Hz, 1H, h), 3.72 (s, 3H, o), 3.58 (d, J = 15.5 Hz, 1H, e), 3.53 (d, J = 15.5 Hz, 1H, e'), 3.22 (dd, J = 13.8, 5.7 Hz, 1H, i), 3.08 (dd, J = 13.8, 5.9 Hz, 1H, i'). $^{13}$C NMR (151 MHz, Chloroform) δ 171.3 (n), 170.6 (f), 147.2 (m), 143.5 (j), 134.4 (d), 130.2 (k), 129.4 (c), 129.3 (b), 127.8 (a), 123.8 (l), 52.8 (h), 52.7 (o), 43.9 (e), 37.6 (l). LR(ESI)-(+) -MS: m/z +343.1 [M+H]$^+$; +365.1 [M+Na]$^+$. HRFT(+) -MS: m/z +343.1291 [M+H]$^+$; +343.1288 calc. for [C$_{18}$H$_{14}$N$_2$O$_7$Na]$^+$. 4.106: Operation Product Carboxylic Acid

Operation product methyl ester 4.108 (0.70 g, 2.04 mmol, 1 eq.) was dissolved in a solution of LiOH (245 mg, 10 mmol, 5 eq.) in water (0.5 mL) and THF (4.5 mL) and stirred at 40 °C for 18 h. The reaction mixture was neutralised with Dowex® 50WX8-100 ion exchange resin which was subsequently removed via filtration over Celite®. The filtrate was dried over Na$_2$SO$_4$ and then concentrated to dryness in vacuo to give a crude orange solid. The solid was triturated...
with Et₂O (2 × 10 mL) to remove impurities and leave the title compound as an off-white amorphous solid (0.48 g, 72%).

m.p. 181–185 °C. [α]D²⁰ = +52.39 ° (c = 1.00, 1% MeOH/CH₂Cl₂). ¹H NMR (600 MHz, Acetone-d₆) δ 8.06 (d, J = 8.7 Hz, 2H, l), 7.42 (d, J = 8.6 Hz, 2H, k), 7.40 (d, J = 7.9 Hz, 1H, g), 7.28 – 7.19 (m, 5H, a+b+c), 4.82 (td, J = 8.4, 5.1 Hz, 1H, h), 3.52 (d, J = 14.2 Hz, 1H, e), 3.48 (d, J = 14.2 Hz, 1H, e'), 3.35 (dd, J = 13.8, 5.1 Hz, 1H, i), 3.14 (dd, J = 13.8, 8.6 Hz, 1H, i'). ¹³C NMR (151 MHz, Acetone) δ 172.6 (n), 171.0 (f), 147.7 (m), 146.3 (j), 136.8 (d), 131.4 (k), 130.0 (c), 129.1 (b), 127.4 (a), 124.0 (l), 53.6 (h), 43.5 (e), 37.8 (i). LR(ESI)-(−)-MS: m/z -327.1 [M-H]⁻; -655.2 [2M-H]⁻; -677.2 [2M-2H+Na]⁻. HRFT(−)-MS: m/z -327.0972 [M-H]⁻; -327.0986 calc. for [C₁₁H₁₅N₂O₃]⁻.

4.76: Fmoc-(4-nitro)Phe-(4-nitro)Phe-OMe

Fmoc-(4-nitro)phenylalanine 4.75 (100 mg, 0.23 mmol, 1 eq.), (4-nitro)phenylalanine methyl ester hydrochloride 4.71 (66 mg, 0.25 mmol, 1.1 eq.), EDCI-HCl (49 mg, 0.25 mmol, 1.1 eq.) and HOBt·H₂O (7 mg, 46 µmol, 0.2 eq.) were suspended in CH₂Cl₂ (4.6 mL, 50 mM) and cooled to 0 °C. NEt₃ (35 µl, 26 mg, 0.25 mmol, 1.1 eq.) was added and the resulting solution warmed to room temperature and stirred for 18 h. The reaction mixture was diluted with CHCl₃:IPA (3:1, 20 mL) and washed with aq. HCl (0.1 M, 20 mL). The organics were dried over MgSO₄ and solvent was removed under reduced pressure. The residue was purified by flash column chromatography (SiO₂, 1% MeOH/CH₂Cl₂) to give the title compound as a white solid (130 mg, 88%).

m.p. decomposes above 220 °C. [α]D²⁰ = +6.28 ° (c = 1.00, 1% MeOH/CH₂Cl₂). ¹H NMR (600 MHz, Chloroform-d) δ 8.11 (d, J = 8.1 Hz, 2H, o/w), 8.08 (d, J = 8.6 Hz, 2H, o/w), 7.78 (d, J = 7.5 Hz, 2H, b), 7.54 (d, J = 7.5 Hz, 2H, e), 7.42 (td, J = 7.1, 3.1 Hz, 2H, c), 7.35 – 7.30 (m, 2H, d), 7.27 (m, 2H, n/v), 7.21 (d, J = 7.8 Hz, 2H, n/v), 6.26 (d, J = 6.9 Hz, 1H, r), 5.14 (d, J = 7.7 Hz, 1H, j), 4.80 (ap. d, J = 6.3 Hz, 1H, s), 4.53 (dd, J = 10.0, 6.9 Hz, 1H, h), 4.41 (dd, J = 9.7, 6.8 Hz, 1H, h'), 4.36 (q, J = 6.7 Hz, 1H, k), 4.19 (t, J = 6.4 Hz, 1H, g), 3.70 (s, 3H, z), 3.23 (dd, J = 13.8, 4.9 Hz, 1H, t), 3.16 (dd, J = 13.1, 7.5 Hz, 1H, l), 3.13 – 3.04 (m, 2H, i'±t'). ¹³C NMR (151 MHz, Chloroform) δ 170.8 (y), 147.4 (m+u / p+x), 143.6 (m+u / p+x), 143.3 (f), 141.5 (a), 130.3 (n+v), 128.0 (c), 127.3 (d), 124.9 (e), 124.0 (o/w), 123.9 (o/w), 120.3 (b), 67.2 (h), 55.8 (k), 53.0 (s+z), 47.2 (g), 37.9 (l+t); carbamate carbon (i) and amide carbon (q) were not detected. LR(ESI)-(−)-MS: m/z +661.1 [M+Na]⁺. HRFT(−)-MS: m/z +639.2083 [M+H]⁺; +639.2086 calc. for [C₂₃H₂₅N₅O₆]⁺.
4.74: Diketopiperazine Reference

To a solution of Fmoc-(4-nitro)Phe-(4-nitro)Phe-OMe 4.76 (51 mg, 79.8 µmol, 1 eq.) in CH₂Cl₂ (2 mL and MeOH (0.5 mL) was added HNEt₂ (250 µL, 177 mg, 2.42 mmol, 30 eq.). The resulting solution was stirred at room temperature for 18 h, by which point a solid had precipitated. The precipitate was collected by vacuum filtration and dried in vacuo to give the title compound as a white solid (11 mg, 36%).

m.p. >250 °C. [α]D²⁵ = -28.324 ° (c = 1.10, DMSO-d₆). ¹H NMR (600 MHz, DMSO-d₆) δ 8.27 (s, 2H, h), 8.02 (d, J = 8.5 Hz, 4H, b), 7.26 (d, J = 8.5 Hz, 4H, c), 4.23 (ap. br. s, 2H, f), 2.84 (dd, J = 13.7, 4.5 Hz, 2H, e), 2.78 (dd, J = 13.6, 5.1 Hz, 2H, e'). ¹³C NMR (151 MHz, DMSO) δ 166.3 (g), 146.1 (a), 144.7 (d), 131.2 (c), 122.9 (b), 54.7 (f), 37.5 (e). LR(ESI)-(+-)MS: m/z +385.1 [M+H]+. HRFT(+-)MS: m/z +385.1145 [M+H]+; +385.1143 calc. for [C₁₈H₁₆N₄O₆H].

4.83: N,N-diethyl-2-(4-nitrophenyl)acetamide

2-(4-Nitrophenyl)acetamide (100 mg, 0.55 mmol, 1 eq.), EDCI·HCl (159 mg, 0.83 mmol, 1.5 eq.) and HOBt·H₂O (127 mg, 0.83 µmol, 1.5 eq.) were dissolved in DMF (4 mL). A solution of diethylamine (86 µL, 60.5 mg, 0.83 mmol, 1.5 eq.) in DMF (1 mL) was added followed by DIPEA (192 µL, 142 mg, 1.10 mmol, 2 eq.) and the resulting solution stirred for 16 h at room temperature. Solvent was then removed in vacuo and the residue was redissolved in CH₂Cl₂ (15 mL) and washed with NH₄Cl (sat. aq., 2 × 10 mL) followed by NaHCO₃ (sat. aq., 2 × 10 mL). The organics were dried over MgSO₄ and solvent was removed to give the title compound (128 mg, 99%) as a yellow oil requiring no subsequent purification.

¹H NMR (600 MHz, Acetone-d₆) δ 8.18 (d, J = 8.7 Hz, 2H, h), 7.57 (d, J = 8.7 Hz, 2H, c), 3.90 (s, 2H, e), 3.46 (q, J = 7.1 Hz, 2H, g), 3.37 (q, J = 7.1 Hz, 2H, g'), 1.17 (t, J = 7.1 Hz, 3H, h), 1.07 (t, J = 7.1 Hz, 3H, h'). ¹³C NMR (151 MHz, Acetone) δ 169.0 (f), 147.6 (a), 145.5 (d), 131.4 (c), 124.0 (b), 42.8 (g), 40.7 (g'), 40.2 (e), 14.7 (h), 13.3 (h'). LR(ESI)-(+-)MS: m/z +237.0 [M+H]+. HRFT(+-)MS: m/z +237.1223 [M+H]+; +237.1234 calc. for [C₁₂H₁₂N₂O₃H].

4.87: N,N-Diethyl-2-phenylacetamide

Phenylacetyl chloride (0.43 mL, 0.50 g, 3.23 mmol, 1 eq.) was dissolved in CH₂Cl₂ (20 mL) before diethylamine (0.67 mL, 0.47 g, 6.46 mmol, 2 eq.) was added at 0 °C followed by NEt₃ (5 mL). The reaction mixture was warmed to
room temperature and stirred for 15 h. The reaction mixture was then diluted with CH$_2$Cl$_2$ (30 mL) and washed with aq. HCl (1 M, 2 × 50 mL) followed by NaHCO$_3$ (sat. aq., 2 × 50 mL). The organics were dried over MgSO$_4$ and solvent was removed to give the title compound (615 mg, 99%) as a colourless oil requiring no subsequent purification.

$^1$H NMR (600 MHz, Acetone-d$_6$) $\delta$ 7.32 – 7.26 (m, 4H, b+c), 7.24 – 7.18 (m, 1H, a), 3.69 (s, 2H, e), 3.38 (q, $J = 7.1$ Hz, 2H, g), 3.34 (q, $J = 7.1$ Hz, 2H, $g'$), 1.08 (t, $J = 7.1$ Hz, 3H, h), 1.05 (t, $J = 7.1$ Hz, 3H, h'). $^{13}$C NMR (151 MHz, Acetone) $\delta$ 170.0 (f), 137.4 (d), 129.7 (c), 129.1 (b), 127.1 (a), 42.8 ($g'$), 41.0 (e), 40.4 (g), 14.6 (h'), 13.3 (h). LR(ESI)-(+-)MS: $m/z$ +405.4 [2M+Na]$^+$. HRFT-(+-)MS: $m/z$ +214.1208 [M+Na]$^+$; +214.1202 calc. for [C$_{12}$H$_{17}$NONa]$^+$.

4.10.1.8 Synthesis of Reference Standards of Molecular Machine Components

![Scheme 50: Synthesis of reference standards of molecular machine components for operation and model operation monitoring. Reagents and conditions: a) LiOH, MeOH, r.t., 18 h, then HCl$_{(aq)}$ (1M), 65%; b) TFA, TIPS, CH$_2$Cl$_2$, r.t., 3 min, 93%.](image)
4.36: Operated/Unloaded One-Barrier Track

To a solution of 4-nitrophenyl acetic acetate one-barrier aldehyde rotaxane 4.61 (10 mg, 5.4 µmol, 1 eq.) in MeOH (5 mL, 1.1 mM) was added LiOH (1.3 mg, 54 µmol, 10 eq.) at room temperature. The resulting red solution was stirred for 18 h before being quenched with aq. HCl (1.25 M, 150 µL, ca. 35 eq.) upon which the solution turned light yellow. The solvent was removed in vacuo and the residue was purified by prep. TLC (SiO₂, 250 µm plate, 5% MeOH/CH₂Cl₂) to give the title compound as a colourless film (4 mg, 65%).

¹H NMR (600 MHz, Acetone-d₆) δ 9.38 (s, 1H, ae), 8.50 (s, 1H, o), 8.15 (s, 1H, ac), 7.94 (ap. s, 4H, r+s), 7.89 (d, J = 7.6 Hz, 1H, v), 7.61 – 7.51 (m, 2H, ag), 7.31 (d, J = 8.5 Hz, 6H, d), 7.28 (d, J = 8.7 Hz, 2H, am), 7.20 – 7.15 (m, 4H, z+ah), 7.12 (d, J = 8.6 Hz, 6H, e), 7.09 (d, J = 8.9 Hz, 2H, i), 6.99 (d, J = 8.7 Hz, 2H, an), 6.85 (d, J = 8.9 Hz, 2H, j), 6.74 (d, J = 8.4 Hz, 2H, aa), 4.91 (q, J = 7.6 Hz, 1H, w), 4.70 (t, J = 6.9 Hz, 2H, n), 4.08 (t, J = 5.9 Hz, 2H, i), 3.24 (dd, J = 14.0, 5.9 Hz, 1H, x), 3.09 (dd, J = 13.9, 8.4 Hz, 1H, x'), 2.45 (p, J = 6.4 Hz, 2H, m), 1.65 (s, 6H, ak), 1.30 (s, 27H, a). ¹³C NMR (151 MHz, Acetone) δ 170.7 (ad), 167.0 (u), 157.7 (k), 149.2 (c), 148.8 (al), 147.0 (q), 146.4 (ai), 145.2 (f), 140.5 (h), 138.1 (ao), 137.6 (af), 135.2 (p), 134.6 (t), 132.8 (i), 131.4 (e), 131.2 (x), 129.2 (y), 129.2 (am), 128.8 (r/s), 127.8 (ah), 125.9 (r/s), 125.1 (d), 122.5 (o), 120.1 (ag), 119.4 (an), 116.0 (aa), 114.1 (j), 65.3 (l), 63.9 (g), 57.1 (w), 47.9 (n), 42.9 (aij), 37.8 (x), 34.9 (b), 31.6 (a), 31.0 (ak), 30.8 (m). A ¹³C NMR signal for carbon (ab) could only be detected via a 2D NMR cross peak at ca. 157.0 ppm. LR(MALDI)-(+)/MS: m/z +1131.8 [M+H]⁺. HRFT(+)−MS: m/z +1148.6495 [M+NH₄]⁺; +1148.6484 calc. for [C₇₃H₇₀N₈O₄NH₄]⁺.
4.98: Deprotected Catalyst Amide Macrocycle

TFA (515 µL) was added to a solution of trityl-catalyst amide macrocycle 4.93 (10 mg, 10.3 µmol, 1 eq.) in CH₂Cl₂ (2.06 mL, c_total = 4 mM) to give an intense yellow solution. TIPS (10.5 µL, 8.15 mg, 51.5 µmol, 5 eq.) was added and the solution became colourless after 2.5 min. After 3 min, toluene (4 mL) was added and the resulting solution was concentrated to dryness in vacuo at 40 °C. The resulting solid was twice further dissolved in toluene (4 mL) and concentrated in vacuo to give the trityl-deprotected catalyst macrocycle title compound contaminated with Ph₃CH. The crude residue was purified by prep. TLC (SiO₂, 1 × 250 µm, 5% MeOH/CH₂Cl₂) to give the title compound (7 mg, 93%) as a white solid.

m.p. 200–203 °C. ¹H NMR (600 MHz, DMSO-d₆) δ 15.17 – 14.17 (br. m, 1 H, ab), 8.85 – 8.42 (br. m, 1H, x), 8.39 (t, J = 5.8 Hz, 1H, t), 7.58 (t, J = 7.6 Hz, 1H, a), 7.06 (d, J = 7.6 Hz, 2H, b), 7.02 (d, J = 8.4 Hz, 4H, h), 6.94 (s, 1H, p), 6.87 (s, 2H, q), 6.73 (d, J = 8.5 Hz, 4H, i), 4.23 (d, J = 5.8 Hz, 2H, s), 3.88 (t, J = 5.7 Hz, 4H, k), 3.49 (q, J = 6.7, 5.4 Hz, 2H, w), 2.69 (t, J = 7.3 Hz, 4H, d), 2.55 (t, J = 6.9 Hz, 4H, n), 2.50 – 2.42 (m, 6H, f+v), 1.92 (p, J = 7.5 Hz, 4H, e), 1.72 – 1.59 (m, 8H, l+m). A ¹H NMR signal for proton aa was not observed. ¹³C NMR (151 MHz, DMSO) δ 170.4 (u), 160.6 (c), 156.8 (j), 141.9 (o), 139.2 (r), 136.7 (a), 133.9 (g), 129.2 (h), 126.6 (p), 125.0 (q), 120.2 (b), 114.1 (l), 67.0 (k), 42.0 (s), 36.6 (d), 35.5 (w), 34.8 (v), 34.4 (n), 33.7 (f), 31.4 (e), 27.9 (l/m), 27.6 (l/m). ¹³C NMR signals for carbons y, z and aa were not observed. LR(ESI)-(+)–MS: m/z +751.4 [M+Na]+. HRFT(+)–MS: m/z +729.4130 [M+H]+; +729.4123 calc. for [C₄₄H₅₂N₆O₄H]+.
4.10.2 Supporting Information for Section 4.3

4.10.2.1  HPLC/LC-MS Method 1

HPLC was performed on an Agilent Technologies Series 2100 instrument using an Agilent Eclipse XDB-C18 (5 µm, 4.6 × 150 mm) column and the following conditions:

- Injection volume: 20 µL
- Detection wavelength: 254 nm
- Flowrate: 1 mL/min
- Solvents: MeCN (solvent A) and 0.05 M NH₄CHO₂(aq) (solvent B)
- Gradient: 30% A in B to 50% A in B (0–10 min); 50% A in B to 95% A in B (10–15 min), 95% A in B (15–40 min); end.

Analysis by LC-MS was carried out using reverse phase HPLC (Method 1) and an Agilent Technologies 1200 LC system with 6130 single quadrupole MS detector.

4.10.2.2  Reference Standards using HPLC Method 1

Figure 17: HPLC (Method 1) reference standard of commercially available 4-nitrophénylacetic acid.
Figure 18: HPLC (Method 1) reference standard of carboxylic acid product 4.39 from water quench.

Figure 19: HPLC (Method 1) reference standard of methyl ester product 4.34 from MeOH quench.
Figure 20: HPLC (Method 1) reference standard of diketopiperazine 4.74.

4.10.2.3 LC-MS Mass spectrum of 9.4 min LC-MS peak from DBU operation (Section 4.3.2) after 25 h

Figure 21: MS trace (negative m/z) of 9.4 min LC-MS peak from DBU operation (Scheme 25) after 25 h.

4.10.2.4 Boc-Deprotection Protocol for Catalyst-Loaded Two-Barrier, Doubly Stoppered Rotaxane 4.67 (Section 4.3.1)

TFA (0.6 mL, 0.89 g, 7.8 mmol) was added to a solution of molecular machine 4.67 (7.5 mg, 2.20 µmol) in CH₂Cl₂ (2.4 mL). The resulting solution (c = 0.73 mM) was stirred at room temperature for 0.5 h. Toluene (5 mL) was added and the mixture concentrated to dryness in vacuo at 40 °C. The resulting solid was twice taken up in toluene (5 mL) and concentrated in
vacuo to give the TFA salt of Boc-deprotected molecular machine 4.68 (assumed quant.) which was operated without further purification or full analysis.

LR[(ESI)-(+)]-MS: Observed (upper); Predicted (lower) for \((C_{207}H_{217}N_{23}O_{18}H)^{2+}\) corresponding to \([M+2H]^{2+}\) and \((C_{207}H_{217}N_{23}O_{18}HNa)^{2+}\) corresponding to \([M+H+Na]^{2+}\);

![Graph showing isotope patterns](image)

Figure 22: Observed (upper) and calculated (lower) isotope patterns of deprotected catalyst-loaded two-barrier, double stopper rotaxane 4.68. Species are doubly charged. The left hand signal centred at \(m/z\) 1658 corresponds to \([M+2H]^{2+}\) \((C_{207}H_{217}N_{23}O_{18}H)^{2+}\); the right hand signal centred at \(m/z\) 1669 corresponds to \([M+H+Na]^{2+}\) \((C_{207}H_{217}N_{23}O_{18}HNa)^{2+}\).

### 4.10.2.5 Operation Protocol for Deprotected Catalyst-Loaded Two-Barrier, Double Stopper Rotaxane 4.68 with DBU (Section 0)

DMF (50 µL) was added to the crude Boc-deprotected molecular machine 4.68 (0.73 µmol) followed by a solution of DBU in MeCN (0.1 mL, 14.6 mM, 2 eq.) at which point the operation mixture \((c = 4.9 \text{ mM})\) turned intensely pink. The operation mixture was stirred at room temperature and sampled for analysis after 6 h and 25 h.
Chapter 4

Analysis by HPLC: Water (1 drop) was added to a 50 µL aliquot of the operation mixture which was then diluted with 150 µL of MeOH. The resulting solution was filtered over a PTFE filter (0.45 µm pore size) to remove the formed precipitate and analysed by reverse phase HPLC (General Method 1).

Analyses by ESI-(±)-MS and LC-MS were carried out using the prepared HPLC sample. The filtered precipitate was separately analysed by dissolving it in a mixture of CH$_2$Cl$_2$ and MeOH.

### 4.10.2.6 Model Operation Protocol to Assess the Ability of NEt$_3$ to Activate the 1,2,4-Triazole Catalyst (Scheme 26)

The following process was carried out by Dr Sonja Kuschel:

4-Nitrophenylalanine methyl ester hydrochloride 4.71 (48.0 mg, 0.18 mmol), p-tolyl 2-(4-nitrophenyl)acetate 4.70 (50.0 mg, 0.18 mmol) and model catalyst 4.72 (36.5 mg, 0.18 mmol) were dissolved in MeCN (3 mL) and DMF (1 mL). NEt$_3$ (252 µL, 183 mg, 1.8 mmol, 10 eq.) was added and the resulting solution (c = 45 mM) stirred at room temperature for 18 h. The reaction mixture was then concentrated in vacuo and the residue purified by prep. TLC (SiO$_2$, 2% MeOH/CH$_2$Cl$_2$) to give the title compound as a white solid (34.8 mg, 90 µmol, 50%).

The above process repeated by Kuschel with fewer equivalents of NEt$_3$ (63 µL, 46.5 mg, 0.46 mmol, 2.5 eq.) where the product was isolated in 20% yield.

### 4.10.2.7 Operation Protocol for Deprotected Catalyst-Loaded Two-Barrier, Double Stopper Rotaxane 4.68 with NEt$_3$ (Section 4.3.4)

DMF (150 µL) and MeCN (200 µL) were added to the crude Boc-deprotected molecular machine 4.68 (2.20 µmol) followed by a solution of NEt$_3$ in MeCN (0.1 mL, 22.0 mM, 10 eq.). The operation mixture (c = 4.8 mM) was stirred at room temperature for 24 h when an aliquot was taken for analysis. Subsequently, a further aliquot of NEt$_3$ in MeCN (0.1 mL, 22.0 mM, 10 eq.) was added and the operation mixture stirred for a further 72 h at room temperature after which it was again analysed.

Analysis by HPLC: Water (1 drop) was added to a 50 µL aliquot of the operation mixture which was then diluted with 150 µL of MeOH. The resulting solution was filtered over a PTFE filter
(0.45 µm pore size) to remove the formed precipitate and analysed by reverse phase HPLC (General Method 1).

Analyses by ESI-(±)-MS and LC-MS were carried out using the prepared HPLC sample.

4.10.3 Supporting Information for Section 4.4

4.10.3.1 HPLC/LC-MS Method 2

HPLC was performed on an Agilent Technologies Series 2100 instrument using an Agilent Eclipse XDB-C18 (5 µm, 4.6 × 150 mm) column and the following conditions:

- Injection volume: 20 µL
- Detection wavelength: 254 nm
- Flowrate: 1 mL/min
- Solvents: 10% IPA/MeOH (solvent A) and 0.05 M NH₄CH₂O₂(aq) (solvent B)
- Gradient: 20% A in B to 95% A in B (0–20 min); 95% A in B (20–35 min), 95% A in B to 20% A in B (35–45 min); end.

Analysis by LC-MS was carried out using reverse phase HPLC (Method 2) and an Agilent Technologies 1200 LC system with 6130 single quadrupole MS detector.
4.10.3.2 Reference Standards using HPLC Method 2

Figure 23: HPLC (Method 2) reference standard of cresol ester p-tolyl 2-(4-nitrophenyl)acetate 4.70 prepared according to Duddy and Broxton.\textsuperscript{[39]}

Figure 24: HPLC (Method 2) reference standard of 4-nitrophenylacetic acid diethylamide 4.83.
4.10.3.3 Protocol and HPLC result for \( p \)-Tolyl 2-(4-Nitrophenyl) Acetate

4.70 Stability Study with \( \text{NEt}_3 \) (Scheme 31)

\( p \)-Tolyl 2-(4-nitrophenyl)acetate 4.70 (11.9 mg, 44.0 µmol, 1 eq.) was dissolved in a mixture of MeCN (7.88 mL) and DMF (3 mL). \( \text{NEt}_3 \) (0.12 mL, 89 mg, 880 µmol, 20 eq.) was added and the reaction mixture turned bright pink. The reaction mixture was stirred at room temperature for 16 h and analysed by HPLC and \(^1\)H NMR (Method 1).

![HPLC trace](image)

Figure 25: HPLC trace (Method 2) for cresol ester \( p \)-tolyl 2-(4-nitrophenyl)acetate 4.70 stability study with \( \text{NEt}_3 \) (Scheme 31) showing stating material and diethylamide degradation product.

4.10.3.4 Protocol for Comparative Stability Study of Cresol Ester \( p \)-Tolyl 2-(4-Nitrophenyl)acetate 4.70 and \( p \)-Tolyl 2-Phenylacetate 4.86 with \( \text{NEt}_3 \) and Catalyst Mimic 4.65 (Scheme 33)

The appropriate cresol ester (4.70 or 4.86, 2.2 µmol, 1 eq.) and catalyst 4.65 (2.2 µmol, 1 eq.) were dissolved in MeCN-\( d_3 \): DMF-\( d_7 \) (8:3, 0.55 mL, 4.0 mM). \( \text{NEt}_3 \) (12.2 µL, 8.9 mg, 88 µmol, 20 eq.) was added. Note: the 4-nitro bearing substrate turned bright pink on addition of \( \text{NEt}_3 \); the des-nitro substrate remained colourless. Reactions were monitored by \(^1\)H NMR at 6 h intervals for 3 days.
4.10.4 Supporting Information for Section 4.5

4.10.4.1 Mass Spectrum of Third-Generation Two-Barrier Rotaxane

4.104 Formed via Rotaxane Capping (Section 4.5.5)

Of interest: no peaks corresponding to loss of pink barrier two (3304.8 Da), green barrier one (3478.8 Da) or barriers one and two (3186.7 Da).

![Mass Spectrum of Third-Generation Two-Barrier Rotaxane](image)

**Figure 26:** LR-†(ESI)-MS of trityl protected catalyst-loaded two-barrier, double-stopper rotaxane 4.104.

4.10.4.2 Generic Trityl or Trityl and Boc deprotection Protocol for Molecular Machines and Models Thereof (Section 4.5.6)

The following conditions were optimised jointly by the author and Dr Matthew Kitching: Protected molecular machine (1 eq.) is dissolved in CH$_2$Cl$_2$:TFA (4:1, 4 mM, 1 volume) and TIPS (5 eq.) is added to produce a fluorescent yellow solution. **Note: for particularly small scale deprotections with total volume below ca. 500 µL, stock solutions of TFA and TIPS are used for accuracy.** The reaction mixture is stirred until the fluorescent yellow colour has dissipated (typically <5 min, never more than 30 min) indicating complete reduction of the trityl cation by TIPS. Toluene (1 volume) is added and the resulting solution concentrated to dryness in vacuo. The resulting solid is twice further taken up in toluene (1–2 volumes) and concentrated in vacuo. The residue is then triturated with Et$_2$O:n-hexane (1:1, 3 × ca. 4 volumes) before being dried in...
vacuo to give the deprotected molecular machine as an amorphous solid (in assumed quant. yield) which is operated without further purification.

4.10.4.3 Operation Protocol for One-Barrier Rotaxane 4.102 with HNEt$_2$ with 4.100 as a Control (Section 4.5.4)

_Control Study (4.100, R = Trt, Scheme 38):_

Phenylacetic acetate one-barrier rotaxane 4.100 (3 mg, 1.35 µmol, 1 eq.) was dissolved in MeCN-$d_3$: DMF-$d_7$ (2:1, 0.625 mL, 2 mM) and the mixture was filtered over a PFTE filter (0.45 µm pore size) before HNEt$_2$ (0.69 mg, 9.5 µmol, 7 eq.) was added.

_HNEt$_2$ Operation (4.102, R = H, Scheme 38):_

Phenylacetic acetate one-barrier rotaxane 4.100 (3 mg, 1.35 µmol, 1 eq.) was deprotected according to the deprotection protocol described in Section 4.10.4.2. The crude deprotected residue (4.102) was dissolved in MeCN-$d_3$:DMF-$d_7$ (2:1, 0.625 mL, 2 mM) upon which some insoluble material precipitated. The mixture was filtered over a PFTE filter (0.45 µm pore size) before HNEt$_2$ (0.69 mg, 9.5 µmol, 7 eq.) was added.

The reactions were monitored by $^1$H NMR from t = 0 h (both prior to and post HNEt$_2$ addition). Spectra were recorded over a period of 3 days. After 3 days, the control study showed only trace diethylamide 4.87 and rotaxane starting material (4.100); the HNEt$_2$ operation of 4.102 indicated ca. 75% conversion of rotaxane to diethylamide 4.87.

4.10.5 Supporting Information for Section 4.6

4.10.5.1 Deprotection procedure for Trityl Protected Catalyst-Loaded Two-Barrier, Double Stopper Rotaxane 4.104

Trityl protected catalyst-loaded two-barrier, double-stopper rotaxane 4.104 (10 mg, 2.78 µmol, 1 eq.) was deprotected using generic conditions (see Section 4.10.4.2) except for two deviations: 1) the deprotection was carried out at 1.11 mM concentration; and 2) the deprotection was carried out over 3 h.
4.10.5.2 $\text{NEt}_3$ Operation procedure for Trityl Protected Catalyst-Loaded Two-Barrier, Double Stopper Rotaxane 4.105

A solution of $\text{NEt}_3$ in MeCN (0.46 mL, 60.4 mM, 27.8 µmol, 10 eq.) was added to the crude deprotected rotaxane (4.105, assumed 2.78 µmol, 1 eq.) dissolved in DMF-$d_7$ (0.23 mL) to give an operation solution of 4 mM concentration. The operation mixture was stirred at room temperature and sampled after 18 h, 36 h and 72 h. Each aliquot was quenched with 10% $\text{H}_2\text{O}/\text{MeCN}$, filtered over a PTFE filter (0.45 µm pore size) and analysed by HPLC, LC-MS (m/z range 200–1000) and ESI-(±)-MS (m/z range 50–4000). The reaction mixture was also analysed by DOSY NMR at 72 h. Two aliquots were taken at the second time point (36 h) where the second aliquot was quenched with 10% methanol in acetonitrile.

4.10.5.3 Reference Standards using HPLC Method 2

![Phenylacetic Acid](image)

Figure 27: HPLC (Method 2) reference standard of commercially available phenylacetic acid.
Chapter 4

Figure 28: HPLC (Method 2) reference standard of commercially available 4-nitro-L-phenylalanine.

Figure 29: HPLC (Method 2) reference standard of commercially available Boc-4-nitro-L-phenylalanine.
Figure 30: HPLC (Method 2) reference standard of desired operation product acid 4.106.

Figure 31: HPLC (Method 2) reference standard of desired operation product methyl ester 4.108.
Figure 32: HPLC (Method 2) reference standard of diketopiperazine cyclised dimer 4.74 of barrier 2.

Figure 33: HPLC (Method 2) reference standard of diethylamide 4.87.
Figure 34: HPLC (Method 2) reference standard of uninterlocked deprotected macrocycle 4.98.

4.10.5.4 NEt₃ operation HPLC traces from Section 4.6.2

Figure 35: HPLC (Method 2) of NEt₃ operation (Section 4.6.2) sampled at 18 h and quenched with H₂O.
Figure 36: HPLC (Method 2) of NEt₃ operation (Section 4.6.2) sampled at 36 h and quenched with H₂O.

Figure 37: HPLC (Method 2) of NEt₃ operation (Section 4.6.2) sampled at 36 h and quenched with MeOH.
4.10.5.5 Mass Spectra of Partially and Fully Operated Machines Isolated by Prep. TLC

Figure 38: LR-(+ESI)-MS of fully-operated catalyst-loaded two-barrier, double-stopper rotaxane 4.107b.

Figure 39: LR-(+ESI)-MS of partially cleaved catalyst-loaded two-barrier, double-stopper rotaxane 4.107a where barrier 1 is intact and barrier two has been cleaved.
4.10.6 Supporting Information for Section 4.7

4.10.6.1 Procedure for Stability Study of Trityl-Protected One-BARRIER Rotaxane 4.100 in the Presence of 4-Nitrophenylalanine Methyl Ester (4.71FB) and either NEt₃ or DBU (Scheme 42)

This procedure was carried out in duplicate: Procedure A used 7.5 eq. NEt₃; Procedure B used 7.5 eq. DBU.

In duplicate for A and B, phenylacetic acetate one-barrier rotaxane 4.100 (4 mg, 1.80 µmol, 1 eq.) and 4-nitrophenylalanine methyl ester hydrochloride 4.71 (2.35 mg, 9.01 µmol, 5 eq.) were dissolved in MeCN (3.38 mL). To A was added a solution of NEt₃ in DMF (1.12 mL, 12.06 mM, 13.51 µmol, 7.5 eq.). To B was added a solution of DBU in DMF (1.12 mL, 12.06 mM, 13.51 µmol, 7.5 eq.). Reactions A and B were stirred at room temperature for 7 days with sampling at 17, 40 and 88 h.

4.10.6.2 Reference Standards using HPLC Method 2

![Image of HPLC reference standard of 4-nitrophenylalanine methyl ester 4.71FB (free base).](image_url)

Figure 40: HPLC (Method 2) reference standard of 4-nitrophenylalanine methyl ester 4.71FB (free base).
Figure 41: HPLC (Method 2) reference standard of uninterlocked trityl protected macrocycle 4.93.

Figure 42: HPLC (Method 2) reference standard of operated/cleaved free track 4.36 from a one-barrier rotaxane. Signal intensity is low due to poor solubility.
Figure 43: HPLC (Method 2) reference standard of trityl protected one-barrier rotaxane 4.100. Signal intensity is low due to poor solubility.

4.10.6.3 17 h Analyses of NEt₃ and DBU Control Operations of Trityl-Protected One-Barrier Rotaxane 4.100 (Scheme 42)

Figure 44: HPLC (Method 2) of NEt₃ operation (Scheme 42) sampled at 17 h showing the amine 4.71FB only.
**Figure 45:** HPLC (Method 2) of DBU operation (Scheme 42) sampled at 17 h showing amine nucleophile 4.71FB, amide product 4.108 and operated rotaxane components including protected macrocycle 4.93 indicating undesired product formation via a competing mechanism with DBU.

### 4.10.6.4 Procedure for Model Operation of Deprotected One-Barrier Rotaxane 4.102 in the Presence of 4-Nitrophenylalanine Methyl Ester 4.71FB and either NEt₃ or DBU (Scheme 42)

This procedure was carried out in duplicate: **Procedure A** used 7.5 eq. NEt₃; **Procedure B** used 7.5 eq. DBU.

In duplicate for A and B, phenylacetic acetate one-barrier rotaxane 4.100 (4 mg, 1.80 µmol, 1 eq.) was deprotected according to the generic trityl deprotection protocol 4.10.4.2.

To each of the crude deprotected rotaxanes A and B was added 4-nitrophenylalanine methyl ester hydrochloride 4.71 (2.35 mg, 9.01 µmol, 5 eq.) followed by MeCN (3.38 mL). To A was added a solution of NEt₃ in DMF (1.12 mL, 12.06 mM, 13.51 µmol, 7.5 eq.) i. To B was added a solution of DBU in DMF (1.12 mL, 12.06 mM, 13.51 µmol, 7.5 eq.). Reactions A and B were stirred at room temperature for 7 days with sampling at 17, 40 and 88 h.
4.10.6.5 Procedure for Stability Study of p-Tolyl 2-Phenylacetate 4.86 and 4-Nitrophenylalanine Methyl Ester 4.71FB in the Presence of DBU (Scheme 43)

This procedure was carried out in duplicate: Procedure A used 1 eq. DBU; Procedure B used 2 eq. DBU. Procedure A is a control where DBU only neutralises the HCl salt of 4-nitrophenylalanine methyl ester hydrochloride 4.71. Procedure B has an additional equivalent of DBU to assess its ability to act as a transacylation catalyst.

In duplicate for A and B, phenylacetic acid cresol ester 4.86 (2.5 mg, 11.1 µmol, 1 eq.) and 4-nitrophenylalanine methyl ester hydrochloride 4.71 (2.88 mg, 11.1 µmol, 1 eq.) were dissolved in MeCN-\textit{d}3: DMF-\textit{d}7 (3:1, 1 mL, 11 mM). DBU (1.65 µL, 1.68 mg, 11.1 µmol, 1 eq.) was added to reaction A; DBU (3.29 µL, 3.36 mg, 22.1 µmol, 2 eq.) was added to reaction B.

Reactions A and B were held at room temperature with no agitation and were monitored at 12, 25, 38, 49 and 51 h by \textsuperscript{1}H NMR.
4.11 References


5. Conclusions and Outlook

This thesis has described research carried out in the pursuit of advancing our capabilities in the field of artificial molecular machines. The molecular machines that have been presented here are synthetic, small-molecule mimics of the fascinating, highly-evolved architectures used by nature within the so-called ‘central dogma’ of molecular biology. These biological machines, captained by the ribosome, facilitate the constant flow of information between DNA, the blueprint to our existence, and the ever-refreshing pool of built-for-purpose functional molecules that keep us alive.

Our artificial molecular machines are primordial mimics of Nature’s information-translators but, in a very rudimentary way, they each aim to re-enact a fundamental characteristic of the ribosome. That is, that they are built to read information that has been transcribed into some form of template, and to then translate this information into a pre-programmed product.

To design systems capable of achieving these goals, we are faced with one principal question: In order to avoid statistical synthetic outcomes and to go ‘beyond the one-pot soup’, how can we control effective molarity and achieve programmed compartmentalisation as Nature does? There is no single correct answer.

The devices discussed in this thesis have centred on the use of rotaxanes as interlocked structures by which to control the effective molarity between two mechanically (but not covalently) linked molecules. In each case the machines have been built in a modular fashion whereby the interlocked architecture was achieved by an AMT rotaxane formation and onto this rotaxane was attached the separately prepared, information-rich template. The rotaxane-formation strategies in all of the machines reported herein were very similar, however the information templates that these rotaxanes were appended to varied hugely. This demonstrates the potential for diversity in designing and building molecular machines using this strategy.

The molecular machines discussed were designed to form oligopeptides. In Chapter 2, the information-synthon consisted of a template bearing non-proteinogenic beta-homo ($\beta^3$) amino acids. The resulting molecular machine was able to translate the sequence of amino acids stationed along the rotaxane thread into a sequence-defined oligopeptide. The successful translation of this sequence information is reminiscent of the ribosome translating information
from mRNA into sequence-defined polypeptides; however, the ribosome is unable to iteratively translate multiple $\beta^3$ amino acids, which demonstrates the adaptability of our design. Furthermore, a link was established between this artificial machine and one of Nature’s own machines. A model species with the machine-translated peptide sequence was selectively catabolised by a protease enzyme to remove the peptide bond-forming catalyst from the translated sequence in a manner comparable to natural post-translational peptide modifications.

In Chapter 3, the goal of machine operation was not to translate sequence information but instead to translate the polydispersity characteristics of a distribution of one type of amino acid, leucine, dispersed along a polymeric track. The successful operation of this machine translated this information into an associated distribution of oligoleucine products of varying length. The formed oligoleucine moieties were able to adopt partial $\alpha$-helical secondary structures in solution and, as such, were efficient asymmetric catalysts for the Juliá-Colonna epoxidation of chalcones. In a similar fashion to the ribosome printing the primary structure of molecules that go on to fold into the specific, higher-order structures of bioactive enzymes, we have shown that we too can print potent, catalytically active molecules using an artificial molecular machine.

The final project discussed in this thesis explored an alternative mode of peptide bond formation compared with its predecessors. Where cysteine-mediated native chemical ligation (NCL) was previously used to assemble oligopeptides, machines based on peptide bond formation via an acyl-transfer catalyst were explored with a view of accessing new classes of oligopeptide products and removing some of the obstacles presented by NCL. Multiple potential acyl-transfer catalysts were investigated and deemed unsuitable prior to the research outlined in Chapter 4 of this thesis which focused on 1,2,4-triazoles as catalysts. The application of acyl-transfer catalysis with 1,2,4-triazoles in molecular machinery was unsuccessful due to low catalyst selectivity and an accompanying propensity for side reactions. Although this project did not produce a functional artificial molecular machine, extensive investigations into the system did provide invaluable information which contributed significantly to the successful completion of the projects discussed in Chapters 2 and 3.

The prospects offered by artificial molecular machines are diverse and are very exciting indeed. Their prevalence in the research fields has been growing apace over the past decades, but the biggest breakthroughs will emerge when proof-of-principle research shifts up a gear into functional applicability. This is when we will begin to see artificial molecular machines impacting positively on our daily lives, be it in manufacturing, in medicine or in increasingly complex intellectually-responsive technologies. At present, we have only begun to scratch the surface of
what is possible with artificial molecular machines. With Nature’s abundance of biological machines as inspiration, the future landscape of this field will not be defined by our current capabilities but by the scope of our curious imagination.
6. General Experimental Details

Unless stated otherwise, reagents were obtained from commercial sources and used without purification. Unless otherwise stated, all reactions were carried out in anhydrous solvents and under an N\textsubscript{2} atmosphere. Room temperature (r.t.) reactions were carried out between 16 – 26 °C. Anhydrous THF (HPLC grade, Fischer scientific), CHCl\textsubscript{3} (99.8+%, Fischer scientific), CH\textsubscript{2}Cl\textsubscript{2} (HPLC grade, Fischer scientific), and PhMe (>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). DMF (Peptide synthesis grade, Merck) was used unless otherwise stated. \textsuperscript{1}H NMR spectra were recorded on a Bruker Avance III instrument with an Oxford AS600 magnet equipped with a cryoprobe [5mm CPDCH 13C\textsubscript{-1 H/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 ((CD\textsubscript{3})\textsubscript{2}SO = 2.50 ppm, CDCl\textsubscript{3} = 7.26 ppm, CD\textsubscript{2}Cl\textsubscript{2} = 5.32 ppm, (CD\textsubscript{3})\textsubscript{2}NCDO = 2.75 ppm, CD\textsubscript{3}OD = 3.31, CD\textsubscript{2}CN = 1.94 and (CD\textsubscript{3})\textsubscript{2}CO = 2.05 ppm). All \textsuperscript{1}H resonances are reported to the nearest 0.01 ppm. The multiplicity of \textsuperscript{1}H signals are indicated as: s = singlet; d = doublet; t = triplet; quint. = quintet; sex. = sextet; sept. = septet; 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. \textsuperscript{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 ((CD\textsubscript{3})\textsubscript{2}SO = 39.52 ppm, CDCl\textsubscript{3} = 77.16 ppm, CD\textsubscript{2}Cl\textsubscript{2} = 54.00 ppm, (CD\textsubscript{3})\textsubscript{2}NCDO = 29.76 ppm, CD\textsubscript{3}OD = 49.00, CD\textsubscript{2}CN = 118.26 and (CD\textsubscript{3})\textsubscript{2}CO = 29.84 ppm). All \textsuperscript{13}C resonances are reported to the nearest 0.1 ppm in general, or to 0.01 ppm to aid in the differentiation of closely resolved signals. 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 20x20 cm, 60 F254 prep. plates (Merck) or Silica Gel GF 20x20 cm, U254 prep. 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 wave 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 either an Agilent 6130 single quadrupole MS detector or an Advion Expression CMS L
single quadrupole MS detector. High-resolution mass spectrometry was carried out by the EPSRC National Mass Spectrometry Service Centre (Swansea, UK) or by staff at the Mass Spectrometry Service, School of Chemistry, The University of Manchester. Melting points (m.p.) were determined using a Büchi M-565 apparatus and are corrected. Optical rotations were measured using a Rudolph Research Analytical Autopol I polarimeter with both AP Accuracy (±0.004°) and resolution upgrades with a built in thermoprobe for temperature measurement/control. Measurements were conducted using a sodium lamp (\( \lambda \) 589 nm, D- line) at 20 °C; \([\alpha]_D^{20}\) values were reported in \(10^{-\circ} \text{cm}^2 \text{g}^{-1}\) (abbreviated to °), concentration (c) in g per 100 ml. Enantiomeric ratios were determined by chiral HPLC on an Agilent 1260 Infinity system (UV detection, 210 nm). A Chiralpak IC column (5 \( \mu \text{m} \) particle size, 250×4.6 mm, Diacel Corporation) with \( n\)-hexane:IPA (95:5) as eluent (1 ml/min flow-rate) was used for separations unless otherwise indicated.