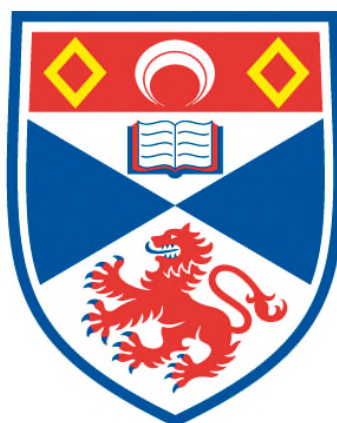


**BUILDING COMPLEX SYSTEMS BASED ON  
SIMPLE MOLECULAR ARCHITECTURES**

**Craig Collumbine Robertson**

**A Thesis Submitted for the Degree of PhD  
at the  
University of St Andrews**



**2011**

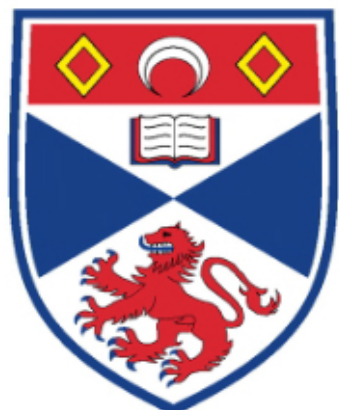
**Full metadata for this item is available in  
St Andrews Research Repository  
at:**

**<http://research-repository.st-andrews.ac.uk/>**

**Please use this identifier to cite or link to this item:**

**<http://hdl.handle.net/10023/2573>**

**This item is protected by original copyright**



University  
of  
St Andrews

# **Building Complex Systems Based on Simple Molecular Architectures**

**By**

**Craig Collumbine Robertson**

**A Thesis Presented for the Degree of**

**Doctor of Philosophy**

**in the**

**School of Chemistry**

**University of St Andrews**

**St Andrews**

**February 2011**



**1. Candidate's declarations:**

I, Craig Collumbine Robertson, hereby certify that this thesis, which is approximately 85,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September 2006 and as a candidate for the degree of Ph.D. in September 2007; the higher study for which this is a record was carried out in the University of St Andrews between 2006 and 2010.

**2. Supervisor's declaration:**

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Ph.D. in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date . . . . . signature of supervisor . . . . .

**3. Permission for electronic publication: (to be signed by both candidate and supervisor)**

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the electronic publication of this thesis:

(iii) Embargo on both [all] of printed copy and electronic copy for the same fixed period of 1 year (*maximum five*) on the following ground(s):

Publication would preclude future publication;

Date . . . . . signature of candidate . . . . . signature of supervisor. . . . .





<b>Table of Contents.....</b>	<b>v–ix</b>
<b>Abstract.....</b>	<b>xi</b>
<b>Acknowledgements.....</b>	<b>xiii–xv</b>
<b>Abbreviations.....</b>	<b>xvii–xviii</b>
<b>Chapter One - Introduction.....</b>	<b>1-95</b>
1.1 Introduction	1
1.2 Non–Covalent Molecular Recognition	2
1.2.1 Hydrogen Bonding Molecular Recognition	3
1.2.2 Secondary Interactions and the Jørgensen Model	6
1.2.3 Two Hydrogen Bond Motifs	7
1.2.4 Three Hydrogen Bond Motifs	8
1.2.5 Four Hydrogen Bond Motifs	8
1.3 Enzyme Catalysis	12
1.3.1 Effective Molarity	15
1.3.2 Intramolecular Reaction Rate Acceleration	15
1.4 Templates	20
1.4.1 Template Directed Synthesis	22
1.4.2 Interweaving Template	23
1.4.3 Linear Templates	24
1.5 Minimal Model of Self–Replication	30
1.5.1 Self–Replication and the Origins of Life	32
1.5.2 Self–Replicating Systems Based on DNA	36
1.5.3 Self–Replicating Systems Based on Polypeptides	40
1.5.4 Self–Replicating Systems Based on RNA	44
1.5.5 Synthetic Minimal Replicators	46
1.6 Systems Chemistry	63
1.6.1 Multiple Replicator Systems	64
1.6.2 Combining Replication Strategies With Dynamic Chemistry	70
1.7 Crystallisation as Method of Selectivity From a Dynamic Library	82
1.7.1 Mechanosensitive Self–Replication	84
1.8 Extended Systems and Molecular Logic	87
1.8.1 DNA Logic	89
1.8.2 Extended Systems and Molecular Logic with Self–Replicators	92
<b>Chapter Two - Directing Systems of Multiple Replicators.....</b>	<b>97-188</b>
2.1 Complex Systems	97
2.1.1 Computational Network Experiment	98

2.1.2	Two Replicators In One System – OR Gate	101
2.2	Aims of Project	104
2.2.1	Synthesis of <i>para</i> -Nitronone <b>52</b>	105
2.2.2	Synthesis of <i>meta</i> -Nitronone <b>58</b>	105
2.2.3	Synthesis of Maleimides <b>29</b> and <b>60</b>	106
2.3	Characterisation of Self-Replicating Systems	106
2.3.1	Analysis of the reaction of Maleimide <b>29</b> with <i>meta</i> -Nitronone <b>58</b>	107
2.3.2	Analysis of the reaction of Maleimide <b>29</b> with <i>para</i> -Nitronone <b>52</b>	109
2.3.3	Isolated Replicator Conclusions	112
2.3.4	Cross-Catalysis Investigations	112
2.4	Competitive Two Replicator System	113
2.4.1	Template Instruction of System	116
2.5	Coupling the Two Replicator System to a Dynamic Process	120
2.5.1	Template Direction of a Dynamic Reagent Pool	125
2.6	Summary of two replicator systems	127
2.7	Change of Maleimide	128
2.7.1	Analysis of the reaction of maleimide <b>72</b> with <i>p</i> -nitronone <b>52</b>	130
2.7.2	Analysis of the reaction of maleimide <b>72</b> with <i>m</i> -nitronone <b>68</b>	133
2.7.3	Investigation for Cross-Catalysis	137
2.8	Two Replicator Kinetic System	138
2.8.1	Template Instruction of Two Replicator System	141
2.9	Coupling Two Replicator System to a Dynamic Process	143
2.9.1	Template Direction of Dynamic Two Replicator System	146
2.10	Direction of a Mixture of Simple Components	148
2.11	Extended System of Dynamic Reagents	154
2.12	Summary of Two Replicator Systems	159
2.13	Design and Characterisation of a Third Self-Replicating System	161
2.13.1	Synthesis of Biphenyl Nitronone <b>82</b>	162
2.13.2	Synthesis of Long Nitronone <b>83</b>	163
2.13.3	Analysis of the Reaction of <i>l</i> -Nitronone <b>83</b> with Maleimide <b>72</b>	163
2.14	Expanding the Replicator Network - Three Replicator System	167
2.14.1	Template Direction of Three Replicator System	171
2.15	Coupling Three Replicator System to a Dynamic Pool of Reagents	174
2.15.1	Template Direction of Three Replicator System with Dynamic Reagent Pool	177
2.16	Three Replicator System From Simple Components	180
2.16.1	Template Direction of Simple Components	183
2.17	Overall Conclusions	186

## Chapter Three - An Orthogonal Reciprocal Replicating System.....189–240

3..1	Replication and the Origin of Life	189
3.1.1	Reciprocal Replication	191
3.1.2	Synthetic Reciprocal Template Effects	194
3.2	Design of an Orthogonal Reciprocal Replicator	198
3.2.1	Molecular Modelling	201
3.2.2	Synthesis of Nitrones <b>91</b> and <b>92</b>	202
3.2.3	Synthesis of Maleimides <b>34</b> and <b>72</b>	203
3.2.4	Synthesis of Aniline <b>88</b>	203
3.2.5	Synthesis of Aldehydes <b>89</b> and <b>90</b>	204
3.2.6	Synthesis of Imines <b>97</b> and <b>98</b>	204
3.2.7	Amine Synthesis	205
3.3	Template Effects in a Reciprocal System	205
3.3.1	Phenylacetic Nitron <b>92</b> and Phenylacetic Maleimide <b>34</b>	207
3.3.2	Carboxy Nitron <b>91</b> and Phenylacetic Maleimide <b>34</b>	208
3.3.3	Phenylacetic Nitron <b>92</b> and Carboxy Maleimide <b>72</b>	209
3.3.4	Carboxy Nitron <b>91</b> and Carboxy Maleimide <b>72</b>	210
3.4	Investigation of the Origin of <i>Cis</i> Selectivity	212
3.5	Acceleration of the Reaction of <i>m</i> -Nitron <b>91</b> with <i>m</i> -Maleimide <b>72</b> by an Imine Template	216
3.5.1	Reciprocal Template Effect	223
3.6	Four Component Reciprocal Replicator	225
3.6.1	Template Instruction of Reciprocal System	229
3.7	Evolution of Reciprocal System	230
3.7.1	Synthesis of Nitron <b>113</b>	231
3.7.2	Reactivity of Nitron <b>113</b> towards Maleimide <b>72</b>	232
3.8	Reciprocal Replicating System	233
3.8.1	Template Instruction of Reciprocal Replicating System	236
3.9	Environmental Dependence of Replication Mechanism	237
3.10	Conclusions	239

## Chapter Four - Rational Design of a Bifunctional Replicator.....241–295

4.1	Introduction	241
4.2	Combining [ <b>A•B</b> ] Complex Channel and Autocatalysis. A Fresh Perspective	245
4.3	Design of the Bifunctional System	247
4.4	Designing an [ <b>A•B</b> ] Complex Pathway	248
4.4.1	Opening Recognition Sites on <b>AB</b> Template	252
4.4.2	Synthesis of Nitron <b>121</b>	253
4.4.3	Synthesis of Nitron <b>52</b>	254

4.4.4	Kinetic Behaviour of [A•B] System	254
4.5	Kinetic Behaviour of the Self–Replicating System	258
4.6	Synthesis of Bifunctional Nitron 120	261
4.6.1	Kinetic Analysis of the Reaction of Nitron 120 with Maleimide 34	261
4.7	Relative Efficiencies of the Three Recognition Mediated Systems	267
4.8	Cross–Catalytic Doping Reactions	269
4.9	Increasing the Solubility of Reagents in Non–Polar Solvents	272
4.10	Determination of Association Constant	275
4.11	Kinetic behaviour of [A•B] system	277
4.12	Kinetic behaviour of self–replicating system	279
4.13	Kinetic Behaviour of the Bifunctional System	282
4.13.1	Further Investigations and Characterisation of Origins of Bifunctional Behaviour	285
4.14	Comparison of Efficiencies of Functionalities	287
4.15	Cross–Catalytic doping Reactions	288
4.16	Conclusions From Bifunctional System	290
4.17	Future Work: Adding Bifunctionality to Existing Replicators.	291
4.18	Overall conclusions	294

## **Chapter Five - Autocatalytic Supramolecular Polymer Synthesis.....297–340**

5.1	Nanoscale Construction – Engineering with Replication	297
5.1.1	Conventional Covalent Polymers	300
5.2	Supramolecular Polymers	301
5.2.1	Applying Self–Replication to Supramolecular Polymer Synthesis	303
5.3	First Generation Monomer Design	305
5.4	Second Generation Monomer Design	308
5.5	Third Generation Monomer Design	309
5.5.1	Model Compound for Analysis: mono Nitron	311
5.5.2	Kinetic Analysis of the Reaction of mono Nitron 155 with Maleimide 34	312
5.5.3	Kinetic Analysis of Bisnitron 152 with Maleimide 34	315
5.5.4	Cross Doping Reactions	321
5.5.5	Conclusions From Kinetic Analysis	323
5.6	Properties of Bistemplate – Evidence for Supramolecular Polymerisation	323
5.7	Concentration Dependence on Polymer Length	327
5.7.1	Diffusion Ordered Spectroscopy	329
5.7.2	Results of DOSY Experiments	332
5.7.3	Extension of Linker	334
5.8	Physical Properties of Supramolecular Polymer	337
5.9	Conclusions	338

**Chapter Six – Overall Conclusions.....341–343**

6.1	Overall Conclusions	341
-----	---------------------	-----

**Chapter Seven – General Experimental and Characterisation of..... 345–410  
Compounds**

7.1	General Experimental Procedures	345
7.2	Preparation and Analysis of Two or Three Replicator Experiments in Chapter 2	348
7.3	Preparation and Analysis of Two or Three Replicators Coupled to a Dynamic Process Experiments in Chapter 2	349
7.4	Preparation and Analysis of Two or Three Replicators from Simple Components Experiments in Chapter 2	350
7.5	Preparation and Kinetic Monitoring of Experiments by <sup>1</sup> H NMR spectroscopy and Deconvolution of Data used in Chapter 2, 3, 4 and 5.	351
7.6	Preparation and Kinetic Monitoring of Experiments by <sup>19</sup> F NMR spectroscopy and Deconvolution of Data used in Chapters 2, 3, 4 and 5.	354
7.7	Kinetic Simulation and Fitting	354
7.8	Characterisation of Compounds	355

**Chapter Eight – References.....411–418**

8.1	References	411
-----	------------	-----

**Chapter Nine – Appendix.....A1–A5**

A1.	Typical SimFit input file for a Bimolecular control reaction with explanation of commands	A1
A2.	Typical SimFit input file for a Self–Replicating reaction	A2
A3.	Four Product System Simulation SimFit Input File (All Rate Equal)	A3
A4.	Four Product System Simulation SimFit Input File (AB Rate Enhanced)	A4
A5.	Publications	A5



## Abstract

Over the past twenty years molecules capable of templating their own synthesis, so called self-replicating molecules have gained prominence in the literature. We show herein that mixing the reagents for replicating molecules can produce a network of self-replicators which coexist and that the networks can be instructed by the addition of preformed template upon initiation of the reaction.

Whilst self-replicating molecules offer the simplest form of replication, nature has evolved to utilise not minimal self-replication but reciprocal replication where one strand templates the formation of not an identical copy of itself but a reciprocal strand. Efforts thus far at producing a synthetic reciprocal replicating system are discussed and an alternative strategy to address the problems encountered is proposed and successfully implemented.

The kinetic behaviour of a self-replicating reaction bears two distinctive time periods. Upon initiation, the reaction proceeds slowly as no template exists to catalyse the reaction. Upon production of the template, the reaction proceeds more rapidly *via* template direction. During this slow reaction period, the system is prone to mistakes as the reaction is slow and unselective. The creation of an [A•B] binary complex through non-covalent recognition of reagents allows for the reaction to proceed at an accelerated rate upon initiation however products of such a reaction are usually catalytically inert and do not promote further template directed reaction. A strategy to combine the desired behaviour of an [A•B] binary complex with the further template directed autocatalytic self-replicating reaction is described and implemented.

Supramolecular polymers consist of repeating monomers which are held together by non-covalent interactions. The strong association of a self-replicating template dimer is comparable to that of supramolecular polymers reported thus far in the literature which are produced by cumbersome standard linear synthetic procedures. Herein the application of self-replication to the field of supramolecular polymer synthesis is discussed. As the autocatalytic reaction to produce the template monomers occurs under the same conditions as required to allow polymerisation to proceed, the polymer is able to spontaneously form *in situ* by self-replicating supramolecular polymerisation.





## Acknowledgements

I would firstly like to thank my supervisor Prof. Douglas Philp for accepting me into his research group, letting me loose in his lab and for all of his enthusiasm, encouragement and guidance during my time in St. Andrews which has made this thesis not only possible but enjoyable to prepare.

I would also like to thank all the members of the Philp research group, past and present, which whom I have enjoyed gazing out of the lab window over the wonderful view of The Old Course, West Sands and the Grampian Mountains. The very knowledgeable postdocs Dr Simon Turega, who helped get me started in the lab and introduced me to Wagamama noodles, and Dr Vicente del Amo who make lab work look easy and whose dedication to Real Madrid and waterpolo could not be faltered! (Dr) Jürgen Huck, guten morgen to you, thank you for too much to mention. You have been a great friend in and out of the lab. It has been a pleasure and a privilege to work with you. Best of luck for your golf in the future, I shall miss you dearly. (Dr) Izzaty Hussan, it has been a pleasure to have met and worked with you. I wish you all the best of luck for your future when you return home to Malaysia. Dr Ilaria Bottero for shining some light (excuse the chemistry pun!) into the lab! Thank you Dr Jan Sadownik for always being fun to work with and Dr Annick Viddone for showing me how to do everything properly! I must also thank the project students who helped make the lab a wonderful place to work for so many years.

I am indebted to the wonderful technical staff at the University of St Andrews. Not only for the incredible services they provide but for teaching me more about how they work and expanding my knowledge and appreciation of the techniques. Thank you Melanja Smith for running a wonderful NMR department and always being up for a good chat. Thank you Dr Tomas Lebl for all your expert advice for everything and anything NMR related. Thank you Mrs Caroline Horsburgh for great chats about running, weddings, dogs and Edinburgh and for your tireless dedication to getting great high res. spectra of compounds most other people probably wouldn't have bothered with. Thank you Mrs Sylvia Williamson for elemental analysis and Prof. Alex Slawin for trying X-ray crystallography on so many crystals.

Thank you Dr Neil Keddie, not only for your help and knowledge in the lab, but for being a great friend, dog sitter and chilli chef! I enjoyed many walks on many beaches with yourself, Katy and Sanna. I do dearly miss them now we have moved although I very much look forwards to joining you again sometime soon with your new addition Joshua Jack Keddie. Many congratulations.

Over the many years I have lived in St. Andrews, I have had the good fortune to make many good friends as an undergraduate, as part of the Ents crew in the union, and as a postgraduate. Dr Craig Campbell and Neil McDowell are two very special friends who did a wonderful job suffering my friendship and a wonderful job as best men at my wedding. I am eternally indebted to them both.

My thanks must also go to my wonderful wife Jennifer, who is actually asleep on my shoulder as I typed this, for being very tolerant and helpful during the lost weekends and evenings she has endured during the preparation of this thesis. I promise to make them up to you. Thank you for all your love and support through this process and for the all the time since we met. You are my best friend and a wonderful wife and I look forwards to the rest of our lives together. Thank you for introducing me to your lovely yellow lab Saffron who definitely made me a dog lover. Saffy sadly has now passed away but she is not forgotten and we are now lovingly tormented by one of her relatives, a black lab called Bella whom I enjoy exploring with just as much.

Last, but by no means least, I am forever thankful to my parents Edith and Nick Robertson (mum and dad!) and my sister and brother Victoria and Graeme for their constant support during my time at University in St Andrews. Ironically, in the home of golf I was never able to match my brothers abilities on the course and I wish him the best of luck making a future in golf. To my sister who typically now I am leaving St. Andrews has recently taken up a course across the Tay, I wish you the best of luck in life. To my mum and dad, I cannot thank you enough for all your love and support from the very start. Bringing me up in such a great home in Torwood, which was so difficult to leave to come to university and for all your help and support whilst at university. Firstly as an undergraduate and then a postgraduate who hopefully phoned home less for money (damn students!) and more for catch up chats. Thank you for warmly welcoming my wife Jennifer into the family and hopefully for welcoming someone else into the family in the near future.

And it is to my parents that I dedicate this thesis.



## Abbreviations

A	Adenosine
A	Hydrogen Bond Acceptor
C	Cytosine
CBPQT	Cyclobis(paraquat- <i>p</i> -phenylene)
D	Hydrogen Bond Donor
D	Diffusion Coefficient
DAN	2,7-Diamino[1,8]naphthyridine
DCC	Dicyclohexyldiimide
DCC	Dynamic Covalent Chemistry
DCL	Dynamic Combinatorial Library
DBU	1,8-Diazabicycloundec-7-ene
DFT	Density Field Theory
DMAP	<i>p</i> -Dimethylaminopyridine
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic Acid
DNP	Dioxynaphthalene
DOSY	Diffusion Ordered Spectroscopy
DP	Degree of Polymerisation
<i>E-coli</i>	<i>Escherichia coli</i>
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
ee	Enantiomeric excess
<i>EM</i>	Effective Molarity
EtOAc	Ethyl Acetate
G	Guanidine
HMDS	Hexamethyldisilazane
HMTA	Hexamethylenetetraamine
HPLC	High Performance Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
$K_a$	Association Constant
$K_d$	Dissociation Constant
MALDI	Matrix-Assisted Laser Desorption/Ionisation
Me	Methyl

## Abbreviations

MS	Mass Spectrometry
MTM	Thiomethyl
NAC	Near Attack Conformation
NMR	Nuclear Magnetic Resonance
PCR	Polymerase Chain Reaction
Pd/C	Palladium on Activated Carbon
<i>p</i> -TSA	<i>p</i> -Toluenesulfonic acid
Rh/C	Rhodium on Activated Carbon
RNA	Ribonucleic Acid
RT	Room Temperature (25 °C)
T	Thymine
TFA	Trifluoroacetic Acid
THF	Tetrahydrofuran
TLC	Thin Layer Chromatography
TOF	Time of flight
TTF	Tetrathiofulvalene
U	Uracil
UG	Urea guanine
UPy	Uridopyrimidinone

# 1

## Introduction

### 1.1 Preamble

The modern day synthesis of complex organic molecules for the purposes of biomedical research, such as in the drug discovery industry, is centred around the exploitation of covalent transformations brought about by centuries of systematic research. Nowadays, complex compounds such as those isolated from nature, for example taxol, can be produced and modified by a wholly synthetic method.<sup>1,2</sup> Whilst the purpose for creating these compounds is biological in nature, their synthesis does not exploit the wealth of knowledge accumulated by biological systems which have had millions of years to evolve into complex and highly efficient machines for catalysis and information storage.

One field of chemistry that is taking inspiration from processes observed in nature is the field of supramolecular chemistry which is described<sup>3-6</sup> as chemistry beyond the molecule. Much like nature uses non-covalent interactions to organise the structure of proteins and the storage of genetic information in DNA, supramolecular chemistry is concerned with the assembly of structures, often from discrete units, associated by non-covalent interactions. The structures produced vary widely in nature with mechanically interlocked molecules,<sup>7</sup> molecular sensors<sup>8,9</sup> and artificial enzymes<sup>10-13</sup> perhaps the most well documented examples. The non-covalent interactions used vary in strength but are usually transient in nature which means they can be conveniently controlled by outside interference or by internal processes.

The study of non-covalent interactions is crucial to the understanding of self-organisation responsible for all cellular processes and life itself. The lessons learned from this understanding also has lead to the development of many novel materials unobtainable from covalent chemistry alone.



## 1.2 Non-Covalent Molecular Recognition

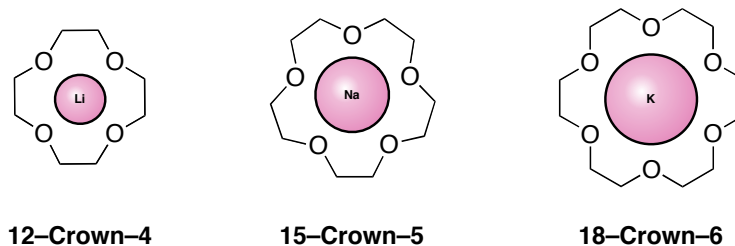
For over a century, organic chemists have investigated covalent chemical reactions in such a way that a wealth of synthetic methodologies are available for countless organic transformations. The covalent bonds formed by such reactions are high in energy with typical carbon-carbon bond strengths of  $\sim 340 \text{ kJmol}^{-1}$  produced.<sup>14,15</sup>

**Table 1.1:** Typical values of covalent and non-covalent interactions between molecules/atoms. Compiled from references 14 and 15.

Interaction	Bond Strength ( $\text{kJmol}^{-1}$ )
Covalent C-C bond	Typically 340
Hydrogen Bonding	Most commonly 1–20 (up to 160)
Salt Bridge	1–12
$\pi$ - $\pi$	< 4

Molecules are also able to transiently interact with each other by comparatively low energy, non-covalent interactions such as hydrogen bonds,<sup>16-23</sup> Van der Waals forces<sup>24,25</sup> or hydrophobic forces.<sup>22,26</sup> Molecules recognise each other on account of the favourable interactions formed from associating with each other as opposed to the bulk solution. The importance of non-covalent interactions for life to exist cannot be over emphasised. These interactions are critical in countless biological processes for example hydrogen bonding between base pairs is utilised to store and transfer genetic information in DNA and give rise to the iconic double helix structure. Enzymes responsible for the regulation of cellular activities use hydrogen bonding, ion pairing and  $\pi$ - $\pi$  forces in concert to exert their exquisite substrate specificity.

The importance of molecular recognition began to intrigue scientists after the accidental discovery of crown ethers (Figure 1.1) by Pedersen<sup>27,28</sup> which lead to subsequent work on synthetic receptors for charged cations and neutral molecules by Cram<sup>29</sup> and Lehn.<sup>30</sup> The three were later awarded a Nobel prize for their work such was its importance.

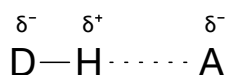


**Figure 1.1:** Variety of crown ethers of different ring sizes. The identity of the cation which is bound is determined by the cavity size of the ring.

A complete discussion of the wide variety of molecular recognition is beyond the scope of this thesis and indeed many journal reviews. As such only the molecular recognition which arises from hydrogen bonding, which is the basis for the work to be described, will be discussed briefly.

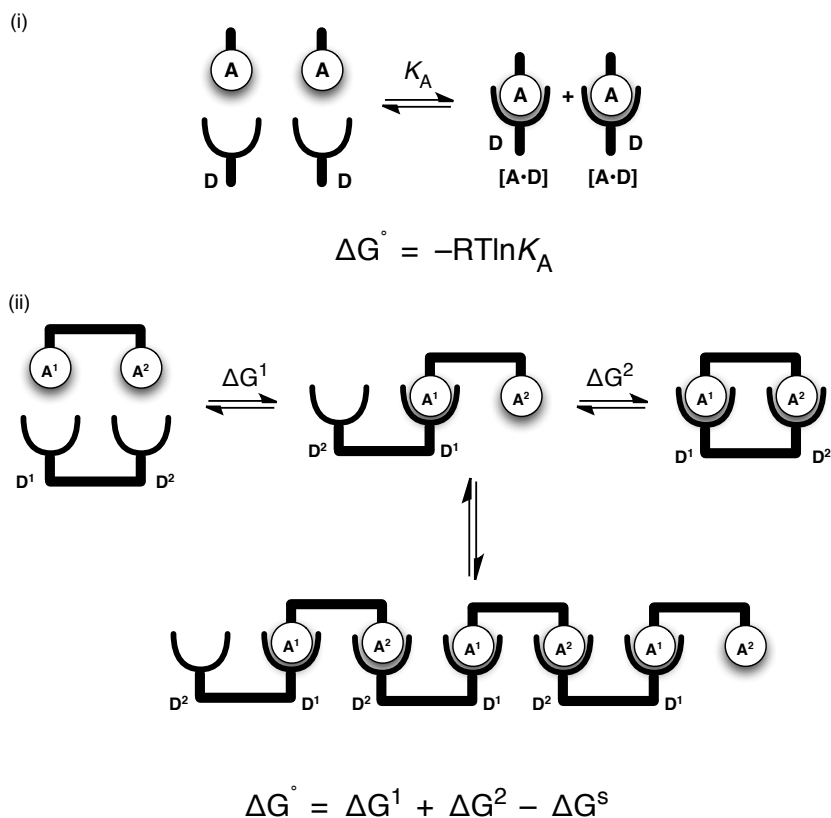
### 1.2.1 Hydrogen Bonding Molecular Recognition

The directionality and relative strength of hydrogen bonds make them the most useful of all non-covalent recognition motifs in self-assembly.<sup>15</sup> A hydrogen bond is formed between molecules containing a polar X–H bond and the non-bonding electron pairs on atom Y. As demonstrated in Figure 1.2, the X–H bond is referred to as the hydrogen bond donor (D) and Y the hydrogen bond acceptor (A) with the hydrogen bond itself denoted by dashed lines. The relative strength of hydrogen bonds is determined by the magnitude of the dipole moment between D and H. Thus with atoms of high electronegativity such as oxygen, nitrogen and fluorine, larger dipoles and hence stronger hydrogen bonds are produced.



**Figure 1.2:** Schematic of a hydrogen bond between a hydrogen atom attached to an electronegative donor atom **D** and the non bonding electron pair of an acceptor atom **A**.

There are a number of functionalities that by the definition given above are able to dimerise through a single hydrogen bond for example phenols and amines. These single hydrogen bonded structures are short lived and consequently of little use for host-guest chemistry. When two or more hydrogen bonds are employed as molecular recognition, straightforwardly thinking the complex should form with a strength equal to the sum of the two hydrogen bonds. In reality, the complex will form with a strength stronger, or weaker than the two hydrogen bonds combined in a phenomenon known as cooperativity.<sup>15,31,32</sup>



**Scheme 1.1:** (i) Reversible association of hydrogen bond donors with two unconnected acceptors. Each association has a unique free energy of association  $\Delta G^\circ$ . (ii) Reversible association of two connected hydrogen bond donors with two connected acceptors. Each association once again has a unique energy of association  $\Delta G^\circ$  with the total energy of binding the two acceptors determined by summing the free energy of each binding interaction and subtracting the connection free energy  $\Delta G^s$ . Figure adapted from reference 31.

If we consider a reversible association of a hydrogen bond donor with a complementary acceptor (Scheme 1.1 (i)), the binding has a free energy of  $\Delta G^\circ$ . Chelate cooperativity arises when two or more donors and acceptors are connected together (Scheme 1.1 (ii)) and the total energy of binding now includes a term to describe the connection free energy,  $\Delta G^s$ .

$$(1) \quad \Delta G^\circ = \Delta G^1 + \Delta G^2 - \Delta G^s$$

$$(2) \quad \Delta G^s = \Delta G^1 + \Delta G^2 - \Delta G^\circ$$

**Equation 1.1:** (1) Equation of free energy from binding the guest to the host with two hydrogen bonds. Each of the two bonds has a unique strength represented for  $D^1-A^1$  by  $\Delta G^1$  and for  $D^2-A^2$  by  $\Delta G^2$ .  $\Delta G^s$  denotes the connection free energy. (2) Equation rearranged to determine  $\Delta G^s$ .

In a positively cooperative system, after the association of the first donor and acceptor, the connection orientates the second acceptor in a spatial arrangement which is more favourable to bind to the second donor. The total energy of binding will

be greater than the sum of the two isolated associations meaning  $\Delta G^s$  is +ve, Equation 1.1 (1). This value represents that after associating the first hydrogen bond, it is easier to associate the second hydrogen bond in comparison to it being unconnected to the first.

Conversely, the first association of donor with an acceptor may make the second association unfavourable, most trivially because of poor orientation of the second association site with respect to the donor. In this case it is less favourable for second association to occur therefore in comparison to if it were unconnected to the first and  $\Delta G^s$  would be -ve. Instead of forming the chelated macrocycle, the thermodynamically more favourable polymer chain will form.

$$(1) \quad K_a = \frac{[Host \bullet Guest]}{[Host][Guest]}$$

$$(2) \quad \Delta G^s = RT \ln \left( \frac{K_{[Host \bullet Guest]}}{K_{D^1A^1} K_{D^2A^2}} \right)$$

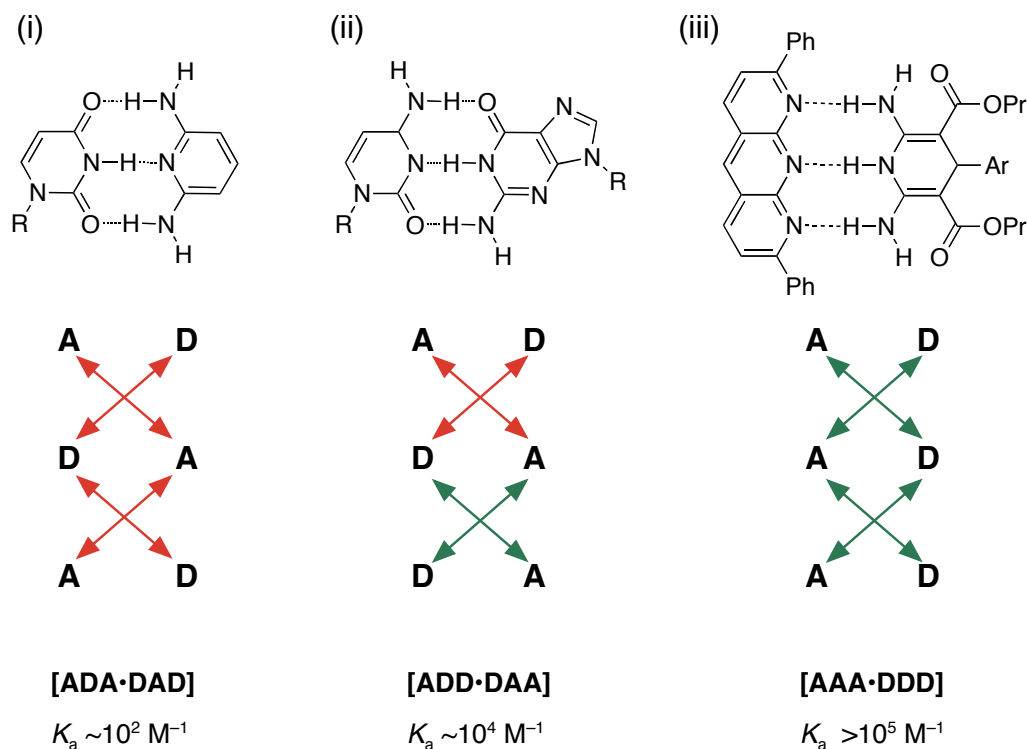
**Equation 1.2:** (1) Equation to determine the association constant of the host-guest complex (2) Gibbs free energy equation to determine  $\Delta G^s$ .

The association constant ( $K_a$ ) can be determined experimentally,<sup>14</sup> Equation 1.2 (1), and from these results, and the equilibrium constants of individual hydrogen bonding interactions (which can be obtained from experimental determination with control compounds), the connection free energy,  $\Delta G^s$  can be determined, Equation 1.2 (2).

It is worth noting at this point that the strength of hydrogen bonds is highly dependent on their environment. For example, a host and guest dissolved in a non-polar solvent such as chloroform will not experience competition for the hydrogen bonding site and a high  $K_a$  could be observed. If that same complex were dissolved in DMSO, a highly polar solvent, the hydrogen bond acceptors on DMSO would compete for those on the host or guest and thus disrupt the desired association. The observed  $K_a$  in this case will be significantly lower. Likewise increasing the temperature of a solution will lower the  $K_a$  of a host guest complex, lowering the temperature will increase the  $K_a$ . Therefore strictly speaking the magnitude of a hydrogen bond is useless unless reported with the solvent and temperature used.

## 1.2.2 Secondary Interactions and the Jørgensen Model

In 1967 Rich and co workers<sup>32</sup> produced triple hydrogen bonding dimeric complexes and discovered that different motifs had surprisingly different association constants which ranged from  $\sim 10^2$  to  $10^5$  M<sup>-1</sup> in CHCl<sub>3</sub> at room temperature. As each structure had three hydrogen bonds, obviously more than just the number of hydrogen bonds were responsible for the stabilities of complexes. This observation inspired Jørgensen and co workers to investigate and more than 20 years later they showed<sup>33,34</sup> that the difference in stability corresponds to attractive or repulsive secondary interactions. When there are two or more hydrogen bonds adjacent to each other, a secondary interaction arises which can either strengthen or weaken the primary hydrogen bond. Stabilisation arises from electrostatic interactions between positively and negatively polarised atoms on adjacent hydrogen bonds and will strengthen the hydrogen bond. Destabilisation arises from repulsive electrostatic interactions between two positively or two negatively polarised atoms on adjacent hydrogen bonds and will weaken the hydrogen bond.



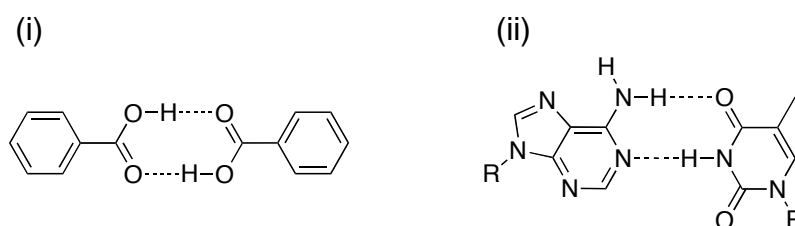
**Figure 1.3:** Demonstration of secondary interactions with three dimerising complexes from Jørgensen (i) and (ii) and Zimmerman (iii). (i) [1-methyluracil·2,6-diaminopyridine] (ii) [1-methylcytosine·9-methylguanine] and (iii) [pyrido[2,3-*b*][1,8]naphthyridine·1,4-dihydropyridine-2,6-diamine].  $K_a$ s reported in CHCl<sub>3</sub> at RT.

Jørgensen used molecular dynamics calculations to demonstrate that [ADD•DAA] complex (ii) (Figure 1.3) was  $10.7 \text{ kcal mol}^{-1}$  ( $45 \text{ kJ mol}^{-1}$ ) lower in energy than [ADA•DAD] complex (i) (Figure 1.3). These results concurred with the experimental data that showed complex (ii) to be more stable than complex (i). The model predicts that the strongest complexes should be formed from [AAA•DDD] complexes which maximise positive secondary interaction.

The next year, Murray and Zimmerman produced<sup>35</sup> a library of triple hydrogen bond complexes to investigate Jørgensen's model including a [AAA•DDD] complex (Figure 1.3 (iii)). The association constants for [ADA•DAD], [DAA•ADD] and [AAA•DDD] complexes were found to be in the order  $10^2 \text{ M}^{-1}$ ,  $10^3\text{--}10^4 \text{ M}^{-1}$  and  $10^5 \text{ M}^{-1}$  in  $\text{CHCl}_3$  at  $25^\circ \text{C}$  respectively, thus concurring with Jørgensen's model. Schneider devised an empirical free energy relationship for the prediction of multiple hydrogen bonded complexes in  $\text{CHCl}_3$ .<sup>36</sup> Primary hydrogen bonds were assigned a value of  $-8 \text{ kJ mol}^{-1}$  with secondary interactions given  $-2.9 \text{ kJ mol}^{-1}$  or  $+2.9 \text{ kJ mol}^{-1}$  for attractive or repulsive interactions respectively.

### 1.2.3 Two Hydrogen Bond Motifs

Common organic functional groups such as carboxylic acids are able to dimerise through two hydrogen bonds (Figure 1.4 (i)). Nature utilises a two hydrogen bond motif as well in the adenine–thymine base pair (Figure 1.4 (ii)).



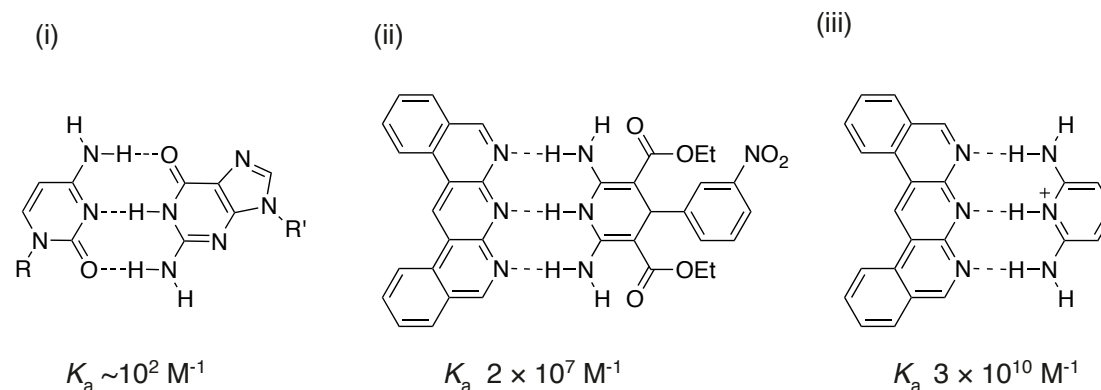
**Figure 1.4:** (i) Benzoic acid dimerises by two hydrogen bonds. (ii) The base pair of adenine and thymine associates with two hydrogen bonds.

Isolated synthetic two hydrogen bonding complexes are generally considered to be weak. The enthalpic gain of associating two molecules is generally outweighed by entropic loss and thus higher concentrations ( $>10 \text{ mM}$ ) are required to form high concentrations of the complex. The relative stability of [AA•DD] motifs is calculated and experimentally demonstrated<sup>23</sup> by Zimmerman *et al* to be stronger than the analogous [AD•DA] complexes on account of the secondary interaction. There are

however, examples of non-covalent assemblies constructed based on a pair of two hydrogen bonding units connected together which will be discussed later in the four hydrogen bond motif Section 1.2.5.

### 1.2.4 Three Hydrogen Bond Motifs

The base pair of cytosine to guanine is nature's greatest example of a triple hydrogen bonding complex (Figure 1.5 (i)). It is interesting to note that in each case, the base pairs do not maximise the number of positive secondary interactions to achieve the strongest possible binding strength. Perhaps nature has evolved this way to avoid base pairs which are too strong and thus unable to dissociate for cellular functions.



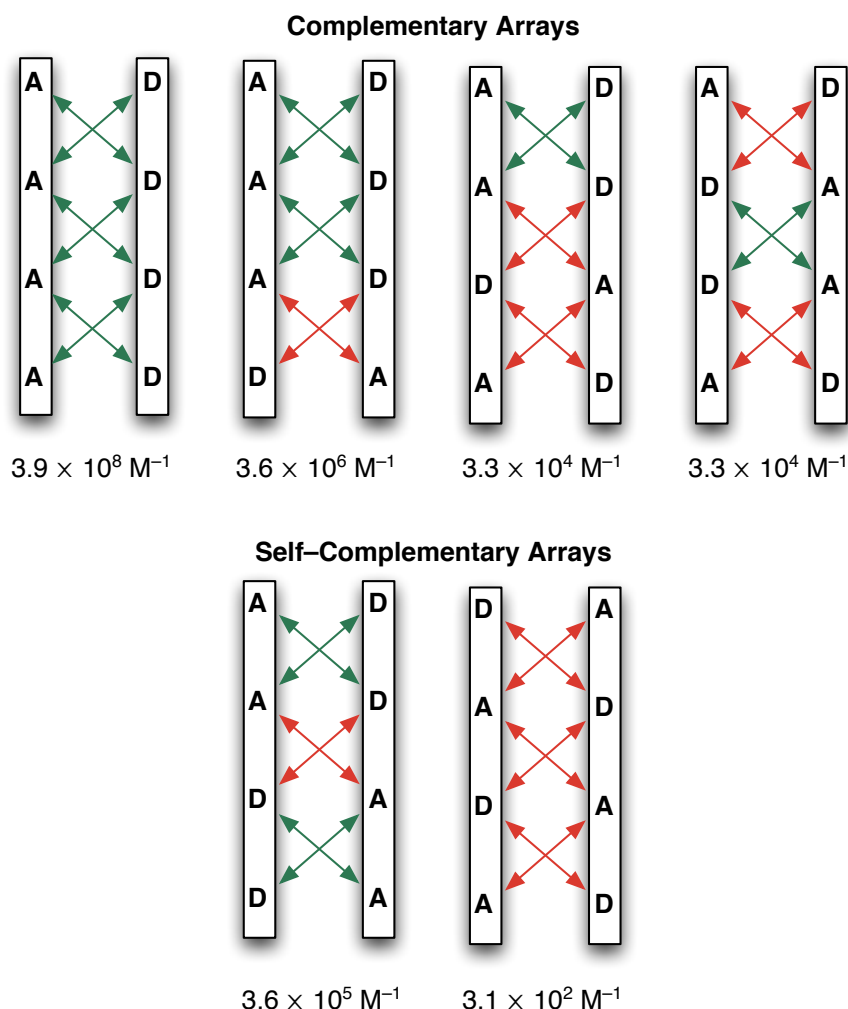
**Figure 1.5:** Triple hydrogen bonding complexes. (i) natural base pair of cytosine to guanine [C•G]. (ii) Very strong neutral triple hydrogen bond complex produced<sup>37</sup> by Leigh *et al.* (iii) Cationic triple hydrogen bond complex described<sup>19</sup> by Leigh *et al* forms a very strong duplex.

As mentioned previously, the work of Jørgensen and Zimmerman independently produced an array of triple hydrogen bond complexes. Zimmerman demonstrated at that time the strongest [AAA•DDD] complex had a  $K_a$  of  $\sim 10^5 \text{ M}^{-1}$ . More recently, Leigh *et al* have produced very strong neutral<sup>37</sup> ( $K_a 2 \times 10^7 \text{ M}^{-1}$ ) (Figure 1.5 (ii)), and cationic<sup>19</sup> ( $K_a 3 \times 10^{10} \text{ M}^{-1}$ ) (Figure 1.5 (iii)), triple hydrogen bond complexes.

### 1.2.5 Four Hydrogen Bond Motifs

Apart from the obvious expected increase in strength by adding a further hydrogen bonding interaction between host and guest, the even number of **D** and **A** atoms in a 4-hydrogen bond array means that structures which are self-complementary are able to be produced. The different arrangement of **D** and **A** atoms can lead to four

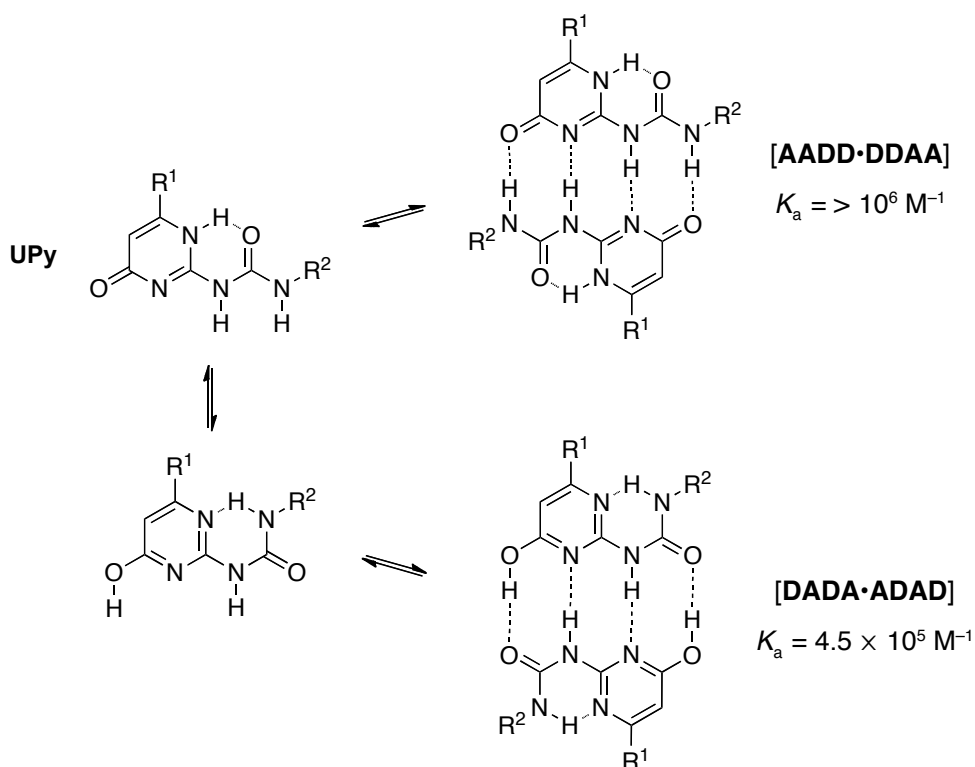
complementary arrays and two self-complementary arrays<sup>38</sup> (Figure 1.6). These self-complementary structures provided the basis for work on supramolecular polymers by Meijer *et al* and indeed for the design of self-replicating templates produced in our laboratory which will be discussed later.



**Figure 1.6:** All six arrangements of D and A sites in a four hydrogen bonding complex. The top four arrangements are complementary and thus require two different molecules. The bottom two are self-complementary and can be achieved by identical molecules. Secondary interactions are shown by red or green arrows as before. Association constants of complexes are theoretically determined by Schneider's rule<sup>36</sup> discussed above. Figure adapted from reference 38.

Meijer and co workers demonstrated<sup>16,17</sup> a four hydrogen bonding array of 2-ureido-4(1*H*)-pyrimidinone (UPy) derivatives. The ability of UPy derivatives to undergo tautomerisation between two self-complementary motifs, **[DADA•ADAD]** and **[AADD•DDAA]**, complicates their analysis however strong  $K_a$ s of  $3.6 \times 10^5 \text{ M}^{-1}$  and  $> \times 10^6 \text{ M}^{-1}$  were determined for each arrangement respectively (Scheme 1.2).





**Scheme 1.2.** UPy 4 hydrogen bond derivatives and their association constants as described by Meijer *et al.*

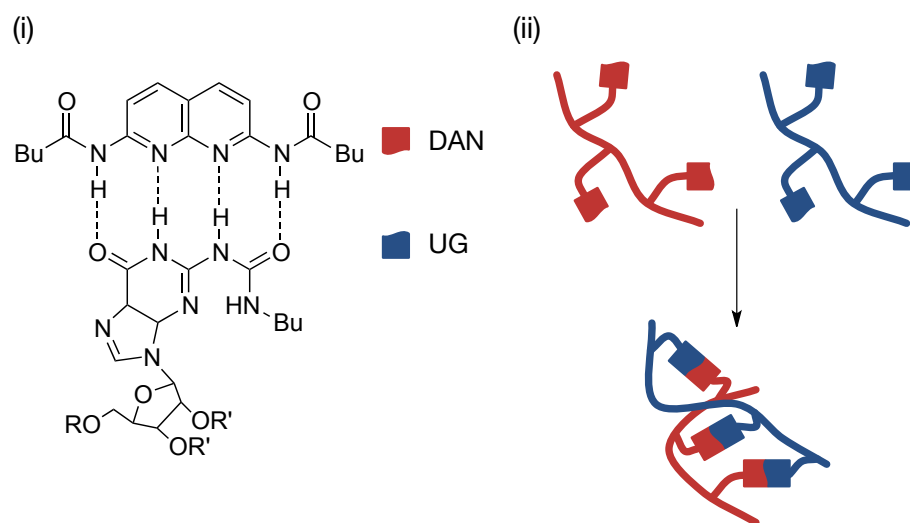
The difference in strength, as expected from Schneider's model, for the **[DADA•ADAD]** tautomer was attributed to the preorganisation of the hydrogen bonding array by the formation of intramolecular hydrogen bonds and differences from **D** atoms used by Schneider's examples. The UPy unit has been successfully used as a non-covalent interaction for supramolecular polymers by linking two UPy units together. This application is discussed further in this thesis, Section 5.2.

Zimmerman *et al* have explored the use of four hydrogen bonding arrays to produce a pair of strongly binding complementary templates<sup>39</sup> and applied them into the creation of polymer blends.<sup>40</sup> When the arrays were capable of tautomerisation, which altered the arrangements of **A** and **D** atoms, such as with UPy, problems were encountered with the formation of undesired complexes. Zimmerman coined the term fidelity<sup>41</sup> to describe the extent to which the desired complex is formed over other possible complexes.

$$\text{Fidelity} = \frac{[\text{Desired Complexes}]}{[\text{All Possible Complexes}]}$$

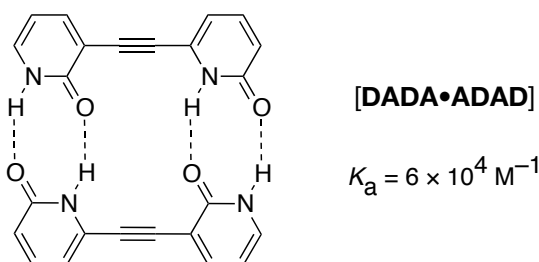
**Equation 1.3:** Fidelity is the mole fraction of the desired complex.

Zimmerman attempted to make a copolymer blend in which one hydrogen bonding array (UPy or UG) was attached to one type of covalent polymer and the complementary template (DAN) was attached to a second covalent polymer normally not known to form a blend with the first. When UPy was used, no polymer blend was seen to be formed as a result of the strong propensity for UPy to self associate rather than associate with DAN, therefore having a low fidelity. The UG array however has high fidelity binding for DAN and a lower propensity for self-association. When UG was attached to the polymer instead of UPy, association of UG with DAN lead to the formation of a polymer blend. (Figure 1.7 (ii)).



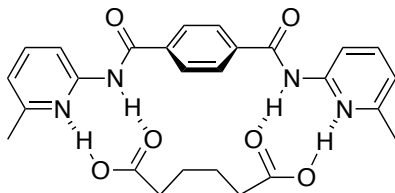
**Figure 1.7:** (i) Example of a four hydrogen bond array between DAN and UG reported<sup>39</sup> by Zimmerman and (ii) the implementation of the hydrogen bonding arrays into a copolymer blend.

An alternative design for a four hydrogen bonding motif as eluded to earlier is to combine a pair of established two hydrogen bonding motifs together. This approach was demonstrated by Wuest and co workers who linked two self-complementary 2-pyridone units together<sup>20,21,42,43</sup> (Figure 1.8).



**Figure 1.8:** Self-complementary structure formed from linking two 2-pyridone units together by Wuest *et al.*

In work which would later be applied by this group and that of von Kiedrowski *et al*, Hamilton described a receptor for adipic acid which used four hydrogen bonds based on linking a pair of two hydrogen bonding motifs<sup>44</sup> (Figure 1.9).



**Figure 1.9:** Receptor for adipic acid designed by Hamilton *et al* operates with 4 hydrogen bonds.

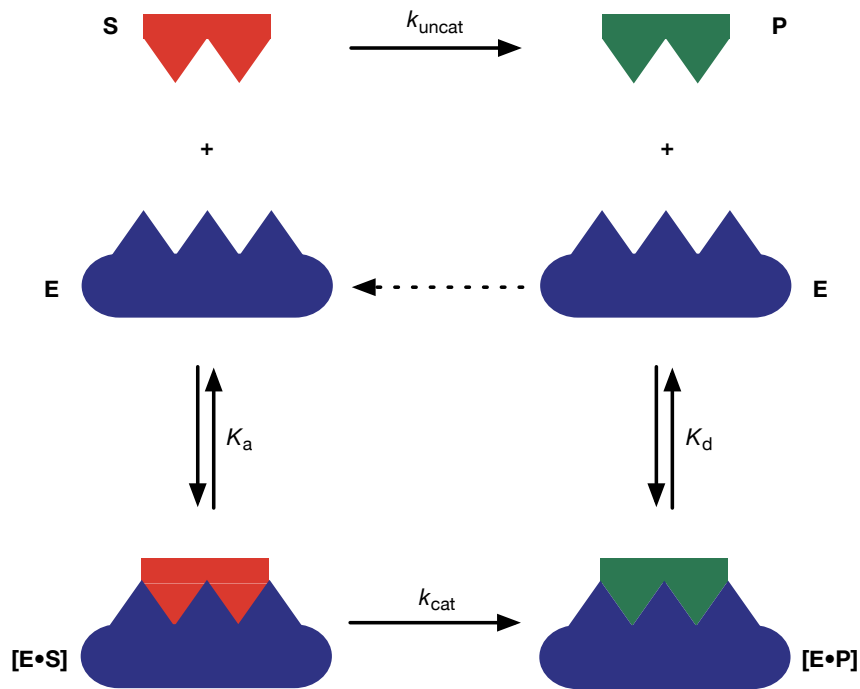
The authors noted that addition of solid adipic acid to a solution of the bisamidopyridine receptor in chloroform lead to rapid dissolution of the normally insoluble diacid. The authors attempted to determine the strength of the association constant for the four hydrogen bonded motif by dilution experiments monitored by <sup>1</sup>H NMR spectroscopy, however no change was seen within NMR spectroscopy detectable concentrations which indicated a  $K_a > 10^5 \text{ M}^{-1}$ .

Whilst hydrogen bonding motifs of greater than four hydrogen bonds have been reported<sup>45</sup> they shall not be discussed here.

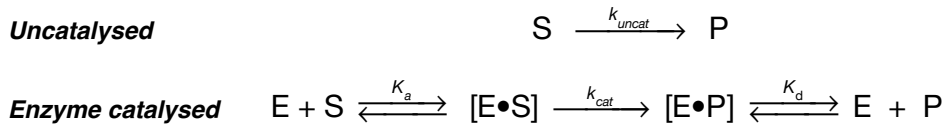
### 1.3 Enzyme Catalysis

Enzymes are the most revered of all catalysts and do more than merely accelerate simple chemical reactions, they also have a high degree of selectivity with the ability to be substrate specific and promote stereospecific reactions with efficiencies most catalytic chemists would marvel at. All of this is also achieved in an aqueous environment at physiological pH whereas many of the transformations they promote require organic solvents or harsh reagents to be reproduced in a laboratory setting.

Enzymes, like traditional catalysts, do not alter the equilibrium for a reaction, nor are they consumed by the reaction they catalyse. However enzymes are unique from many other forms of catalysis as they are very specific for the substrates whose reactions are catalysed. It was noted<sup>46</sup> by Fischer as early as 1894 that enzymes are specific catalysts and this specificity was as a result of the enzyme and substrate possessing a specific geometry of complementary shapes which could fit exactly into each other as depicted in Figure 1.10.



**Figure 1.10:** Cartoon representation of the enzyme **E** in blue with a well defined binding geometry specific for the complementary substrate **S** in red. Association of the substrate forms the enzyme substrate complex **[E•S]** which undergoes a covalent reaction to form the enzyme product complex **[E•P]** which can then dissociate to free the product **P** and release the enzyme to begin catalysis again.



**Scheme 1.3:** **Top:** An uncatyzed reaction from a substrate to a product is slow proceeding at rate  $k_{\text{uncat}}$ . **Bottom:** Enzyme catalysed reaction. Association of the substrate on to the enzyme,  $K_a$  leads to the formation of an enzyme substrate complex **[E•S]**. Reaction within the complex,  $k_{\text{cat}}$ , proceeds at an increased rate in comparison to  $k_{\text{uncat}}$  and produces the enzyme product complex **[E•P]**. Dissociation of the product from the complex,  $K_d$ , releases the product to solution and frees the enzyme to begin another catalytic cycle.

A characteristic of enzyme catalysis is that they follow Michaelis–Menten kinetics<sup>14,47</sup> (Scheme 1.3). The reaction is initiated by association of the substrate **S** on to the enzyme **E**,  $K_a$ , to form an enzyme substrate complex **[E•S]**. Covalent reaction can then proceed at a rate,  $k_{\text{cat}}$ , far greater than the uncatyzed process,  $k_{\text{uncat}}$ , to transform the substrate into the product which remains associated with the enzyme in an enzyme product complex **[E•P]** until dissociation releases the product **P** and frees the enzyme to undertake further catalysis. This hypothesis often referred to as the ‘lock and key’ model provides an oversimplified explanation to the substrate specificity. Whilst the equilibrium leading to the formation of the **[E•S]** complex rationalises the substrate specificity of enzymes, the ‘lock and key’ model implies that

enzymes are rigid structures whilst in actuality enzymes are reasonably flexible structures. In some examples,<sup>48</sup> enzymes failed to catalyse related structures to known substrates or a dramatically reduced catalytic ability was observed. In order to account for these discrepancies, Koshland proposed<sup>49</sup> an alternative model which stated that;

- a) *a precise orientation of catalytic groups is required for enzyme action;*
- b) *‘the substrate may cause an appreciable change in the three–dimensional relationship of the amino acids at the active site; and*
- c) *the changes in protein structure caused by a substrate will bring the catalytic groups into proper orientation for reaction whereas a non–substrate will not.’*

The set of postulates was called the ‘induced–fit’ model and describes how an enzyme is able to change its conformation upon binding a substrate and may be envisaged as a hand in a glove. Described in Scheme 1.4, the enzyme is able to associate with the substrate to form the [E•S] complex which can undergo a reversible conformational change,  $K_C$ , either of the enzyme, the substrate or both, to better orientate the catalytic sites into a more reactive conformation [E•S]<sup>‡</sup>. Covalent reaction in the reactive complex promotes transformation of the substrate to the product which remains associated to the enzyme in the [E•P] complex before dissociation once again releases the enzyme and product.



**Scheme 1.4:** Rate equations and equilibrium constants for induced fit model for an enzyme catalysed reaction. The enzyme can bind the substrate to form a [E•S] complex before a reversible conformational change  $K_C$  produces the catalytically active enzyme conformation [E•S]<sup>‡</sup>. Reaction can then occur at rate  $k_{\text{cat}}$  to form the [E•P] complex before dissociation releases the product and enzyme.

Whilst the model explains the specificity of enzymes for substrates, it does not explain the massive increase in reaction rate constants for catalysed processes. As many as 21 different mechanisms<sup>50</sup> have been proposed in an attempt to rationalise the catalytic efficiencies achieved by enzymes. The general consensus is that enzymes bind their substrates selectively and, while more than one mechanism may be active, the overall effect is to lower the activation barrier<sup>51</sup> by decreasing transition state energy  $\Delta G^{\ddagger}$ , a theory first proposed<sup>52</sup> by Pauling over 60 years ago.

### 1.3.1 Effective Molarity

By associating with its substrate(s), the enzyme effectively increases the local concentration of substrate molecules with respect to the bulk solution. In doing so, the reaction is essentially converted from an intermolecular reaction into a pseudo intramolecular reaction and reaps the rewards of such a process. In order to quantify the efficiency of intramolecular processes in comparison to their corresponding intermolecular reaction the concept of effective molarity<sup>53-57</sup> (*EM*) was conceived. The kinetic *EM* for a reaction is the ratio between the first order intramolecular rate constant and the second order rate constant for the corresponding intermolecular reaction, Equation 1.4, and similarly the thermodynamic *EM* is the ratio of equilibrium constants for an intramolecular reaction and the corresponding intermolecular reaction.

$$EM = \frac{k_{\text{cat}}}{k_{\text{uncat}}} = \frac{k_{\text{intra}}}{k_{\text{inter}}} \quad \text{units} = \frac{\text{s}^{-1}}{\text{M}^{-1}\text{s}^{-1}} = \text{M}$$

**Equation 1.4:** The kinetic effective molarity can be determined by taking the ratio of intramolecular reaction rate to the intermolecular reaction rate. In terms of enzyme catalysis this is the ratio of the catalysed rate against the uncatalysed rate. The unit of the *EM* is a hypothetical concentration in M.

The value obtained for the *EM* of a reaction has the dimensions of concentration. The calculated value of *EM* reflects the hypothetical concentration required by one of the substrates for the bimolecular reaction to proceed at the same rate as the intramolecular reaction. For accurate *EM* determination, the catalysed and uncatalysed reaction should ideally be performed under identical conditions. The values obtained from calculations of *EM* often given physically unobtainable concentrations of reactants with values up to 10<sup>15</sup> M reported,<sup>53</sup> Table 1.2.

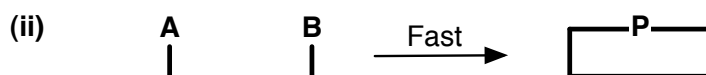
### 1.3.2 Intramolecular Reaction Rate Acceleration

As highlighted in the previous sections, the proximity of reactive and catalytic sites of the substrate and enzyme respectively make an important contribution to the overall rate acceleration achieved by enzymes. Bringing the reactive sites into close proximity renders the reaction an intramolecular process (Scheme 1.5). Intramolecular systems and enzyme substrate complexes both hold their reactive sites in close proximity, the former by covalent bonds and the later *via* transient non-covalent interactions.

**Table 1.2:** Collection of enzymes with the reactions they catalyse and the rate acceleration expressed in terms of effective molarity.

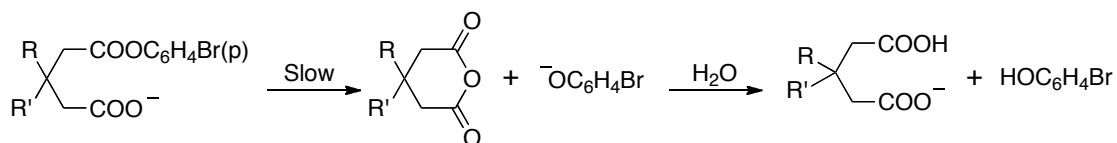
Enzyme	Reaction Catalysed	EM/M
Mandelate racemase	Isomerisation	$1.7 \times 10^{15}$
Carboxypeptidase B	Peptide hydrolysis	$1.3 \times 10^{13}$
Cytidine deaminase	Deamination	$2.2 \times 10^{12}$
Carbonic anhydrase	Bicarbonate formation	$7.7 \times 10^6$
Chorismate mutase	Claisen rearrangement	$1.9 \times 10^{6*}$

\* The Claisen rearrangement is already an intramolecular process and this enhancement is a rate acceleration and not an EM.



**Scheme 1.5:** General scheme for (i) an intermolecular reaction between A and B to slowly produce P and (ii) an intramolecular reaction in which the functional groups A and B are linked together allowing fast reaction to form P.

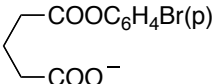
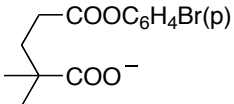
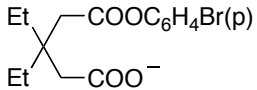
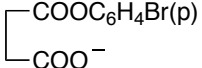
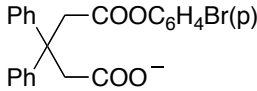
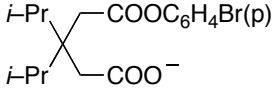
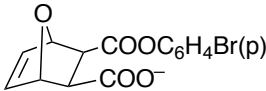
Studies pioneered by Bruice *et al* sought to demonstrate and determine the magnitude of catalysis achieved by rendering an intermolecular reaction into an intramolecular reaction with small organic molecules, using the intramolecular carboxyl group nucleophilic catalysis of the hydrolysis of esters as the reaction<sup>58</sup> (Scheme 1.6).



**Scheme 1.6:** Mechanism of the intramolecular carboxyl group nucleophilic catalysis of the hydrolysis of esters demonstrating the rate determining step to be the cyclic anhydride formation.

The mechanism of the reaction involves an initial, rate determining, step which forms the cyclic anhydride intermediate by ring closure followed by hydrolytic scission to furnish the product. It was expected that the process should therefore be sensitive to factors such as proximity and geometry of the reactive sites. As a result, an array of substrates with whose reactive sites possessed varying degrees of proximity and geometry were investigated relative to an intermolecular equivalent, Table 1.3.

**Table 1.3** List of the monophenyl esters used to monitor the intramolecular reactions (i)–(vii) with the relative rate of intramolecular reaction / intermolecular reaction (control) shown as their *EM*. Table adapted from reference 59.

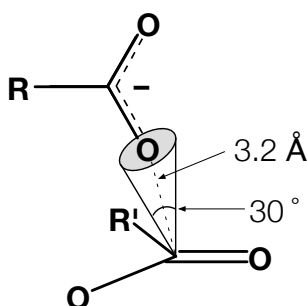
Entry	Substrate	<i>EM</i>
(control)	$\text{CH}_3\text{COO}^- + \text{CH}_3\text{COOC}_6\text{H}_4\text{Br(p)}$	1
(i)		$1 \times 10^3 \text{ M}$
(ii)		$\sim 3.6 \times 10^3 \text{ M}$
(iii)		$1.8 \times 10^5 \text{ M}$
(iv)		$2.3 \times 10^5 \text{ M}$
(v)		$2.7 \times 10^5 \text{ M}$
(vi)		$1.3 \times 10^6 \text{ M}$
(vii)		$\sim 8 \times 10^7 \text{ M}$

The results from the series of experiments demonstrated that, in strain free systems, rate accelerations of up to  $10^8$  could be achieved and it was concluded that this was as a result of the increasing structural rigidity of the substrate and imposition of a steric conformation closely resembling the transition state for the reaction in which the reactive sites were in close proximity.

At the time, the conclusions were not embraced by all with Koshland,<sup>59,60</sup> Jencks<sup>61</sup> and Menger<sup>62</sup> to name but a few to propose alternative theories each of which had their own merits and pitfalls.



The robustness of Bruice's theory was demonstrated 30 years later after a theoretical computational study undertaken by Bruice and Lightstone which focused on the reactions experimentally described in Table 1.3. For each structure, between 10,000 and 40,000 unique conformations were generated and energy minimised to give a unique conformer. In order to determine whether each unique conformer would lead to a reaction, the term near attack conformation (NAC) was introduced<sup>63,64</sup> to describe the steric conformation required by juxtaposed reagents to enter into a transition state.

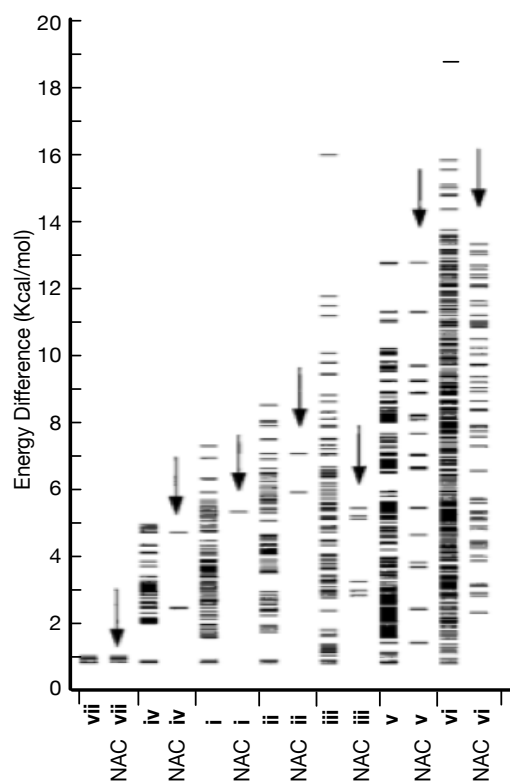


**Figure 1.11 :** The geometry of a near attack conformation (NAC) for intramolecular carboxylate interaction with the ester. The approaching carboxylate ion is no more than 3.2 Å away from the ester reactive site and with an angular window of 30° to perpendicular approach. Figure adapted from reference 64.

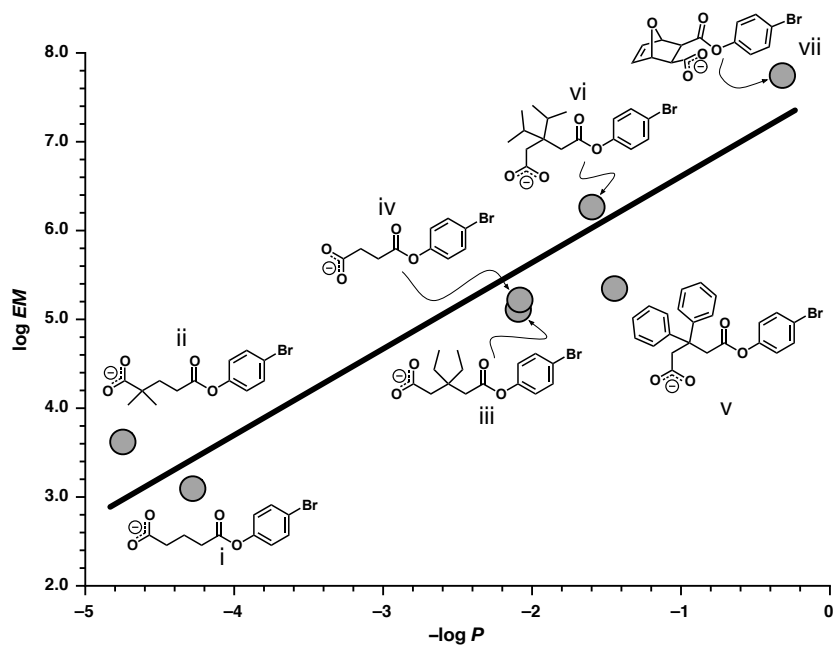
They defined a NAC for the cyclisation reaction as a genuine ground state configuration which with the reactive centres properly orientated for reaction but bond formation has yet to occur. For this criteria to be achieved, the conformer should have the reactive sites separated by a distance of between 2.8 and 3.2 Å and the approach of the nucleophile is within an angular window of 30° thus placing the approaching carboxylate anion into the antibonding orbital of the carbonyl bond (Figure 1.11). In this configuration the van der Waal's radii do not overlap denoting the bond making has not begun and the carbonyl carbon remains  $sp^2$ .

Using the energy of the 10,000–40,000 unique conformers and a Boltzmann distribution profile, the mole fraction  $P$  of conformations present as NACs was determined. As the inherent structural rigidity of the substrate increased, a greater mole fraction was found to be present as NACs (Figure 1.12 (i)). This observation meets the expectation of the original study which sought to induce the proximity effect which would promote rate acceleration. A plot of  $\log P$  vs. the log of the  $EM$  for each of the monophenyl esters revealed a direct linear free energy relationship between the log of the mole fraction present as NACs and  $\Delta G^\ddagger$  (Figure 1.12 (ii)).

(i)



(ii)



**Figure 1.12:** (i) For each ester in Table 1.3, two columns of plots of differences in final energies of local minimum conformations relative to the lowest energy ground state are shown. The column above the structure label (i→vii) shows a line to represent a unique conformation. If a unique conformation is also a near attack conformation, this is shown by a line in the second column (arrows pointing to NACs). Figure adapted from reference <sup>64</sup> (ii) Plot of the log of the *EM* for the anhydride formation from Table 1.3 vs. the log of the mole fraction *P* for NAC formation. Figure adapted from reference 64.

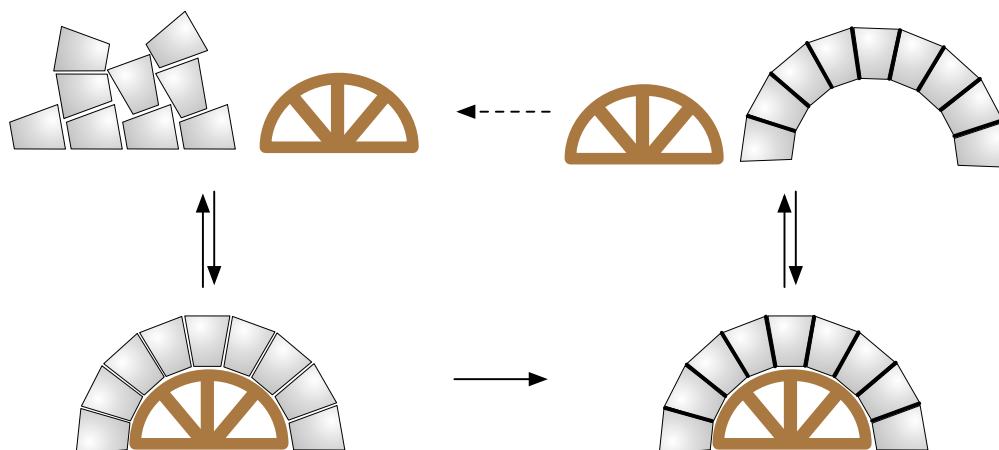
As a result of these calculations, it was concluded that the greater the mole fraction of NACs present in the ground state, the greater the rate constant for that reaction. When the ground state of the substrate contains only NACs, vii, a rate acceleration of up to  $10^8$  can be achieved. Furthermore, the  $\log EM$  was found to be directly related to the ground state enthalpy  $\Delta H^\circ$  and not to the change in entropy  $T\Delta S^\circ$  from which the observed rate accelerations in this reaction could be attributed to enthalpic and not entropic contributions.

These results from Bruice *et al* show that even with a great deal of time and effort devoted to the quest to understand enzyme catalysis, a comprehensive understanding of these rate accelerations is yet to be truly achieved. The ground state proximity effect was only able to account for a rate acceleration of up to  $10^8$  over an uncatalysed process whereas up to  $10^{15}$  has been realised in nature. Given that enzymes have had millions of years to evolve their catalytic processes, it is unsurprising that science lags behind attempting to replicate enzymatic ability in a wholly synthetic system. Naturally given the remarkable efficiency of enzyme catalysis, scientists continue to search for a deeper understanding of how enzymes are able to achieve this efficiency and attempt to replicate it in a laboratory setting.

#### **1.4 Templates**

The practical use of natural enzymes in a synthetic laboratory setting is limited as a result of the substrate specificity displayed by enzymes. However with an ever growing appreciation of enzymatic processes, such as those described above in rendering a bimolecular reaction into a *pseudo* intramolecular reaction to achieve rate acceleration, rationally designed 'artificial enzymes' for a particular reaction of choice have the potential to become a highly rewarding field of research.

An outside body which is able to spatially arrange some form of material is more commonly referred to as a template. A template conveys information in the form of the desired structure of the product and their uses are not confined to the molecular scale. Templates are commonly employed in building and manufacturing and whilst these exist on a macroscopic scale, the principles for macroscopic and molecular templates are identical.



**Figure 1.13:** In the macroscopic construction of for example an archway, a template (in brown) may be used to hold blocks in place while they are fixed by cementing (black). One completed, the template can then be removed to leave behind the finished arch and a template which in principle may be recycled. Figure adapted from reference 65.

For example, when building an archway for a bridge<sup>65</sup> (Figure 1.13) stone blocks may be temporarily held in place by a template whilst they are fixed by cement. Once finished, the template may be removed to leave the free standing structure. With no change in the template throughout construction, it has the potential to be reused to produce another arch. With this analogy it is clear to see a resemblance to enzyme catalysis. On the molecular scale an enzyme, acting as the template, will associate its substrate(s) promote the formation of a particular product over all others. Upon completion, dissociation will liberate the product and enzyme which is unaltered and able to participate in further reactions.

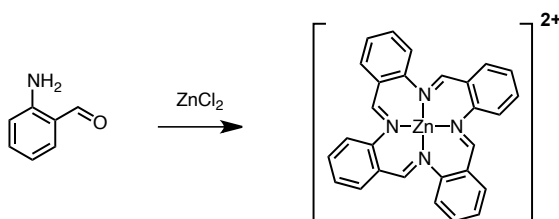
On the molecular stage, templates are common place in nature. Enzymes commonly position substrates into conformations which promote reaction and the synthesis of DNA and RNA is templated by DNA. Upon the discovery<sup>66</sup> of the DNA double helix structure in 1953, Watson and Crick realised that the method of its replication would involve template directed synthesis.<sup>67</sup> Whilst the term template is commonly used in a variety of sciences, it is convenient to understand the distinction between a chemical template to others. The definition of a chemical template as given by Busch<sup>68</sup> is:

“A chemical template organises an assembly of atoms, with respect to one or more geometric loci, in order to achieve a particular linking of atoms.”

From this definition, a number of essential features for a chemical template can be described.<sup>65</sup> Firstly, that a template provides information which organises the substrate(s) into a well defined spatial arrangement and secondly, that the effect of the template is to favour the formation of a single product which otherwise has the potential to assemble and react in a variety of different ways. Changing the structure and thus information of the template will result in a change in the arrangement of substrate(s) and consequently yield a different product.

### 1.4.1 Template Directed Synthesis

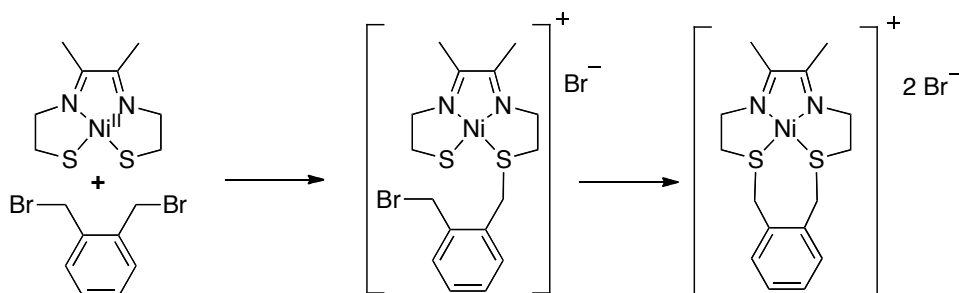
Classically, a template is able to bind reversibly (by non-covalent interactions) to its substrate and promote the formation of a covalent bond. Retrospectively, it can be seen that chemists had made use of templates for organic reactions as far back as 1926, before their significance was realised. Initial work on template directed reactions utilised metal cations as templates for macrocyclisation reactions.



**Scheme 1.7:** Condensation of 2-aminobenzaldehyde in the presence of zinc (II) chloride furnishes the tetrameric macrocycle.

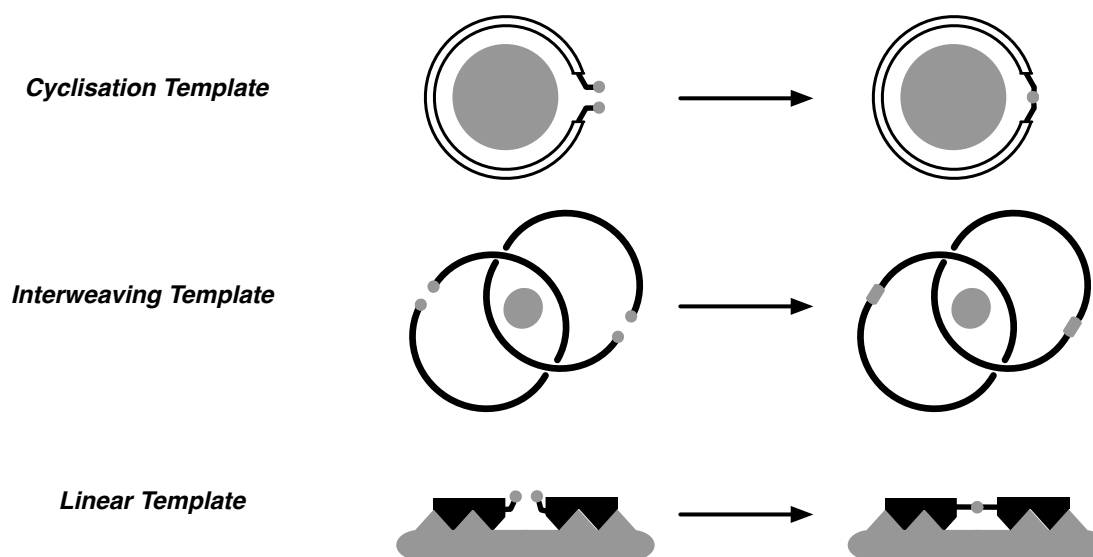
Seidel<sup>69</sup> reacted anhydrous zinc chloride with 2-aminobenzaldehyde but was unaware that the product was the macrocycle (Scheme 1.7). It was a further 40 years before Busch<sup>70,71</sup> reported the structure of the macrocycle in the investigation of metal templated macrocyclisations at the time reporting it as another example of the influence of a metal ion in producing an organic molecule not obtained in its absence. The first example of an intentional use of a template in synthesis came from Busch<sup>72-74</sup> who used a metal cation for a macrocyclisation to demonstrate his concept of a template effect. Depicted in Scheme 1.8 Busch used a nickel (II) dithiolate complex which could react with 1,2-bis(bromomethyl)benzene. After the first reaction of a bromomethyl with a thiolate, the nickel template induces the reactive ends of the intermediate into close proximity therefore promoting cyclisation to form the tetradentate ligand. In the absence of the metal template, the reaction of 1,2-bis

(bromomethyl)benzene with the unbound thiol leads to a mixture of macrocycles of different sizes and polymeric products.



**Scheme 1.8:** Macrocyclisation using nickel (II) as a template by Busch. Initial reaction of 1,2-bis(bromomethyl)benzene with the nickel thiolate complex forms an intermediate in which the reactive sites of the remaining thiolate and bromomethyl groups are positioned into close proximity promoting the macrocyclisation.

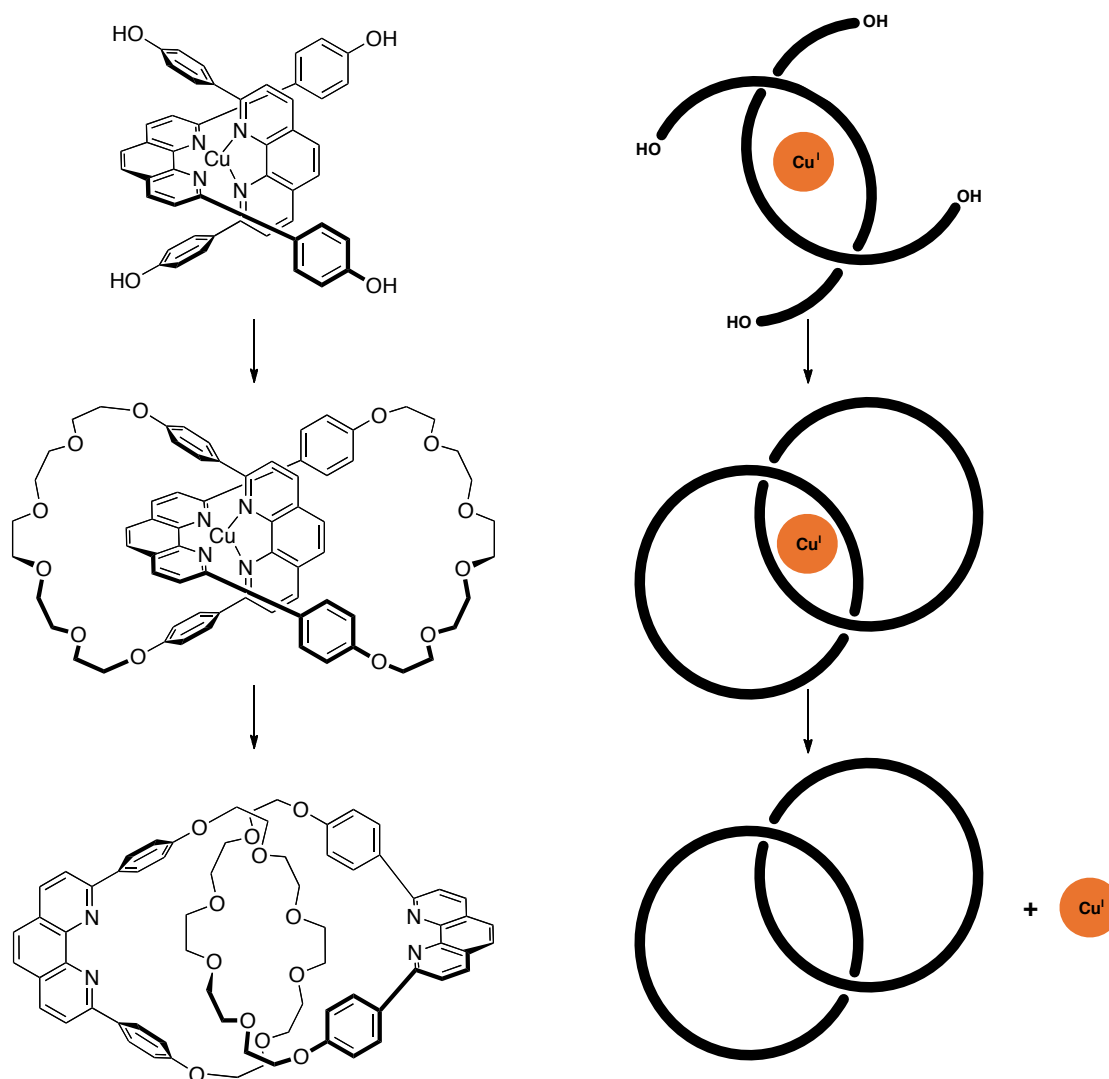
Busch's example of a template directed macrocyclisation was a demonstration of a cyclisation template, one of the three topological classes of template as described by Anderson<sup>75</sup> (Figure 1.14).



**Figure 1.14:** Topological Classifications of templates. Figure adapted from reference 75.

### 1.4.2 Interweaving Template

The archetypal interweaving template was designed by Sauvage and co-workers in their pioneering work using tetrahedral copper (I) to direct the formation of a variety of molecular interlocked architectures.<sup>76,77</sup> The formation of a [2]catenane is displayed in Figure 1.15.



**Figure 1.15:** Copper template directed synthesis of a [2]catenane by Sauvage and co-workers.

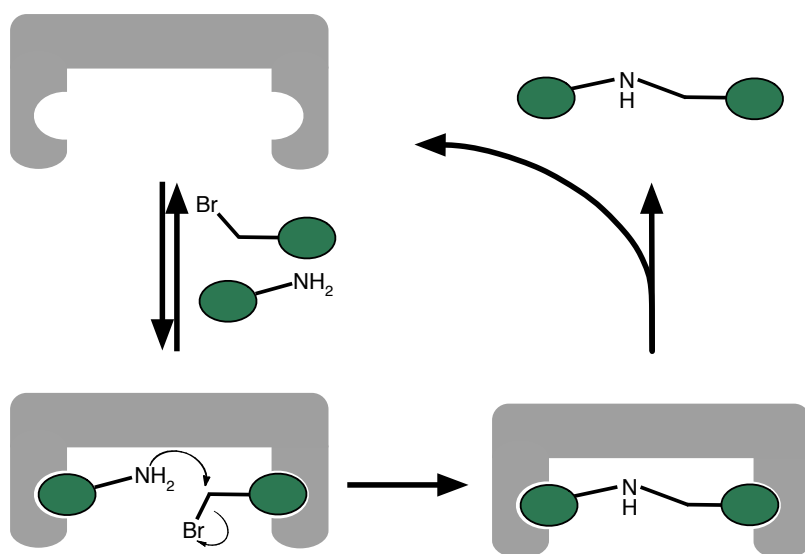
Metal template methodology to create molecular interlocked structures has subsequently been employed by numerous groups including Leigh<sup>78-80</sup> and Beer.<sup>81</sup> An analogous method using  $\pi$ - $\pi$  interactions was developed by Stoddart and co-workers<sup>82-84</sup> to develop rotaxanes and catenanes although strictly speaking the template in their case is both an interweaving and a cyclisation template and is incorporated into the product structure.

### 1.4.3 Linear Templates

A linear template serves to enhance the reaction between two substrates which are associated to the template rather than between two ends of the same substrate. In

nature, this type of template is proficiently applied in the replication of DNA and RNA transcription, *vida infra*, Section 3.1.1.

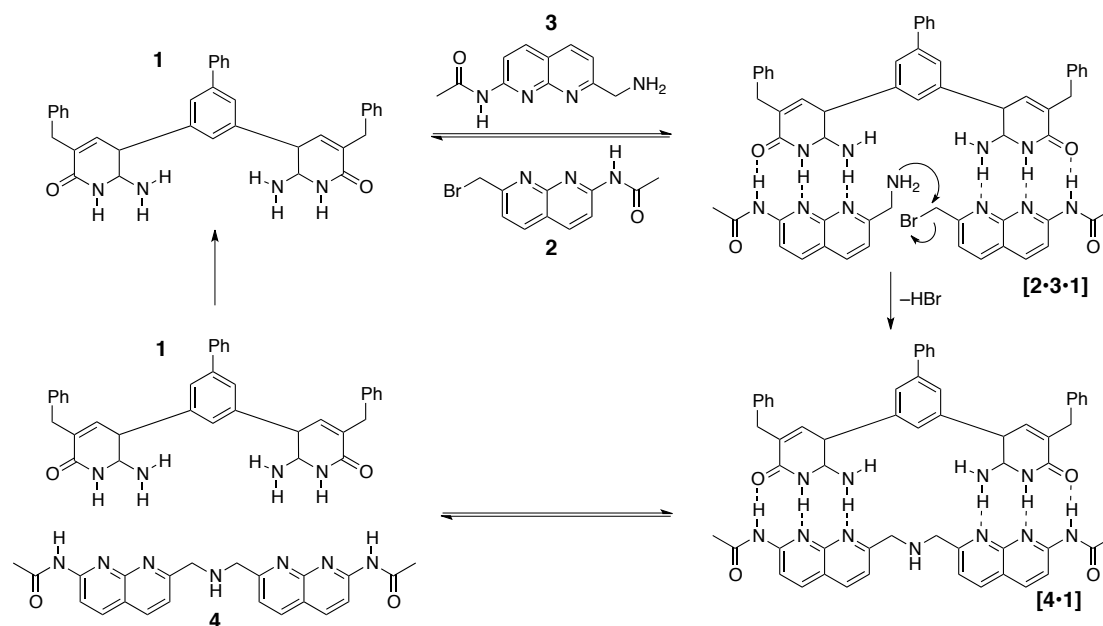
Kelly *et al* were the first<sup>85</sup> to explore a fully synthetic linear template system (Figure 1.16) which was able to bind two substrates *via* complementary hydrogen bonding and promote the mechanistically straightforward  $S_N2$  alkylation of an amine by an alkyl halide as a consequence of the transient intramolecularity of the reaction.



**Figure 1.16:** Representation of Kelly template in grey with two molecular recognition sites which can simultaneously associate the complementary recognition sites possessed by the reagents. Association then brings the reactive site of the amine and bromomethyl group into close proximity promoting their reaction to form the template product complex. Upon dissociation, the product is released from the template which is then able to begin another catalytic cycle.

The template, **1**, is able to associate the two substrates, **2** and **3** by complementary recognition sites which facilitates the formation of the ternary complex **[2•3•1]** which brings the reactive sites into close proximity, promoting the bond formation reaction to furnish **[4•1]**. The subsequent product template complex then dissociates releasing the product amine **4** and returning an unmodified template **1** to complete a catalytic cycle (Scheme 1.9).

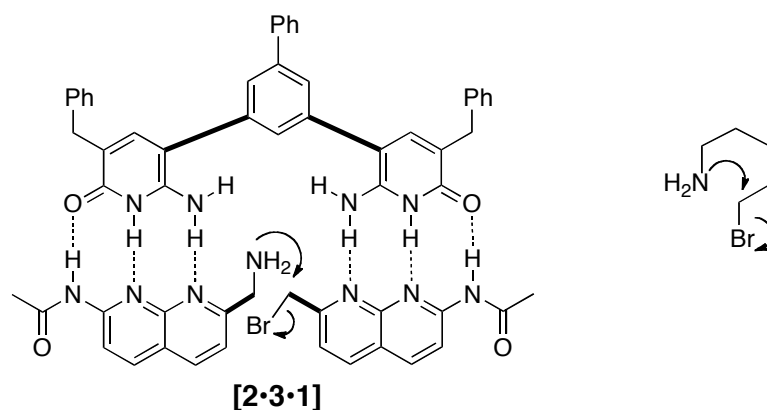




**Scheme 1.9:** Catalytic cycle for the bisubstrate template demonstrated by Kelly.<sup>85</sup> Template **1** is able to reversibly associate **2** and **3** to form the ternary complex **[2•3•1]** which promotes covalent bond formation to produce the complex **[4•1]** before dissociation frees **4** and returns **1** back to solution.

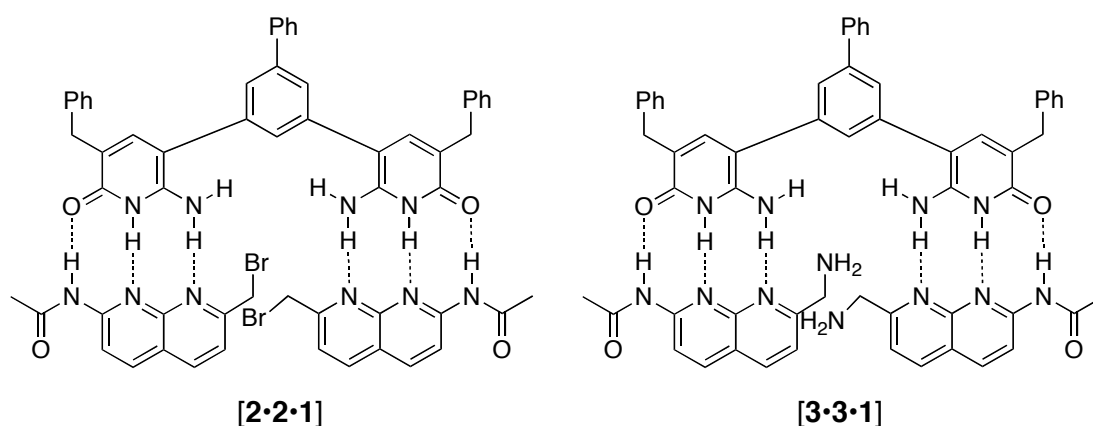
The first system reported by Kelly utilised identical recognition motifs on both substrates for synthetic ease, but it was later shown to decrease the efficiency of the template. A kinetic analysis of the initial system observed a six fold rate increase in comparison to the untemplated bimolecular reaction of only **2** with **3**. Further control reactions demonstrated that the rate acceleration could be attributed to the ternary complex and not from catalysis by a subunit of the template. The results validated the concept for a design of a bisubstrate catalyst although the rate accelerations observed were disappointingly less than expected. Comparison of the template directed reaction to an analogous intramolecular alkylation (Scheme 1.10) in which a rate enhancement of  $5 \times 10^3$  was observed, highlighted that each structure contained four single bonds for which rotation would lead to an unreactive conformer.

Kelly attributed the poor performance of the template to the identical recognition sites used for each substrate. With identical recognition sites, the reactive ternary complex **[2•3•1]** is unlikely to be the only ternary complex present.



**Scheme 1.10:** (i) The *pseudo* intramolecular  $S_N2$  reaction as facilitated by the reaction template with the four single bonds around which rotation yields unreactive conformers highlighted in bold and (ii) Intramolecular cyclisation of 5-bromopentan-1-amine.

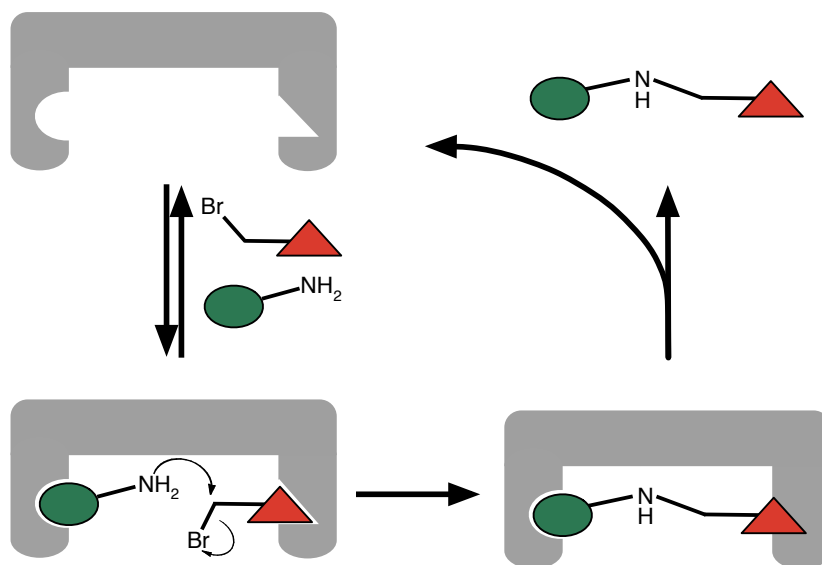
Equally probable are the non-productive ternary complex structures which bear two of the same substrate associated to the template **[2·2·1]** and **[3·3·1]** (Figure 1.17).



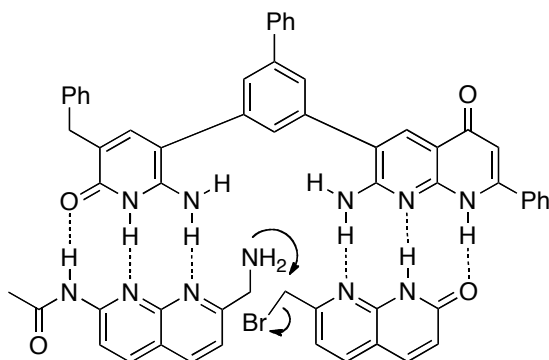
**Figure 1.17:** left **[2·2·1]** and right **[3·3·1]**. Both are equally probable non-productive ternary complexes obtainable in Kelly's original system.

Addressing this problem, Kelly designed a new system bearing non-identical recognition motifs for each of the substrates<sup>86</sup> (Scheme 1.11). Utilising a hydrogen bonding triad once again, careful consideration of the new recognition motif had to be employed to avoid an arrangement of hydrogen bonding donors and acceptors which would lead to the new substrates recognising each other.

Consequently the new recognition site chosen was the association of a 7-amino-1,8-naphthyridin-4-one with a 1,8-naphthyridin-2-one as shown in Figure 1.18.



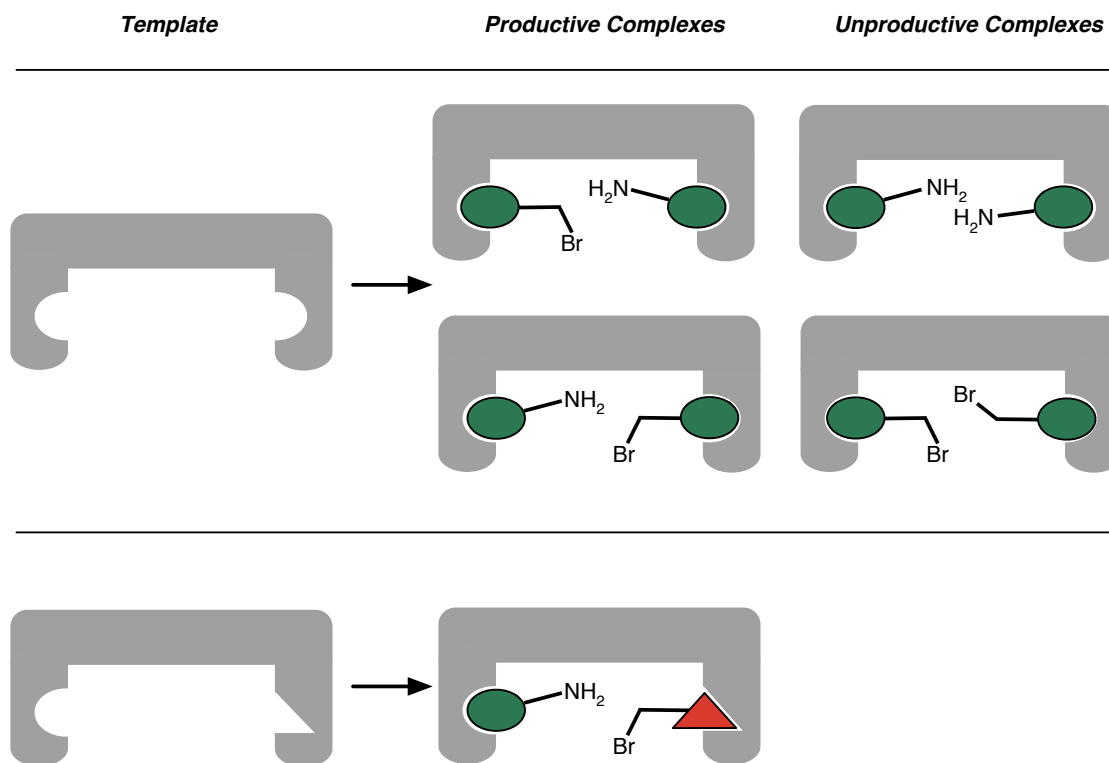
**Scheme 1.11:** Representation of Kelly's template with non-identical binding sites showing specificity for a single reagent at each recognition site.



**Figure 1.18:** Modified ternary complex of Kelly template with non-identical binding sites.

Kelly empirically predicted that the new template would be twice as efficient in comparison to the original template as a result of the change from identical to non-identical recognition motifs. This prediction was rationalised by considering the ternary complexes able to be formed from identical and non-identical recognition motifs (Figure 1.19). With identical recognition motifs, four possible ternary complexes are able to be constructed. Of the four, two are productive placing the alkyl halide in close proximity to the amine whilst the remaining two are unproductive as they associate the same substrate. Thus 50% of the ternary complexes available are reactive ternary complexes. In contrast, with non-identical recognition motifs, each receptor is complementary to only one substrate and as such only a single ternary complex may be formed which is a productive complex. In this case, 100% of the ternary complexes are reactive. With a doubling in the percentage of reactive ternary complexes available, Kelly rather optimistically expected a doubling in the

efficiency of the template from six times the corresponding bimolecular rate to twelve times the rate. Serendipitously a rate enhancement of twelve times the corresponding bimolecular rate was observed.



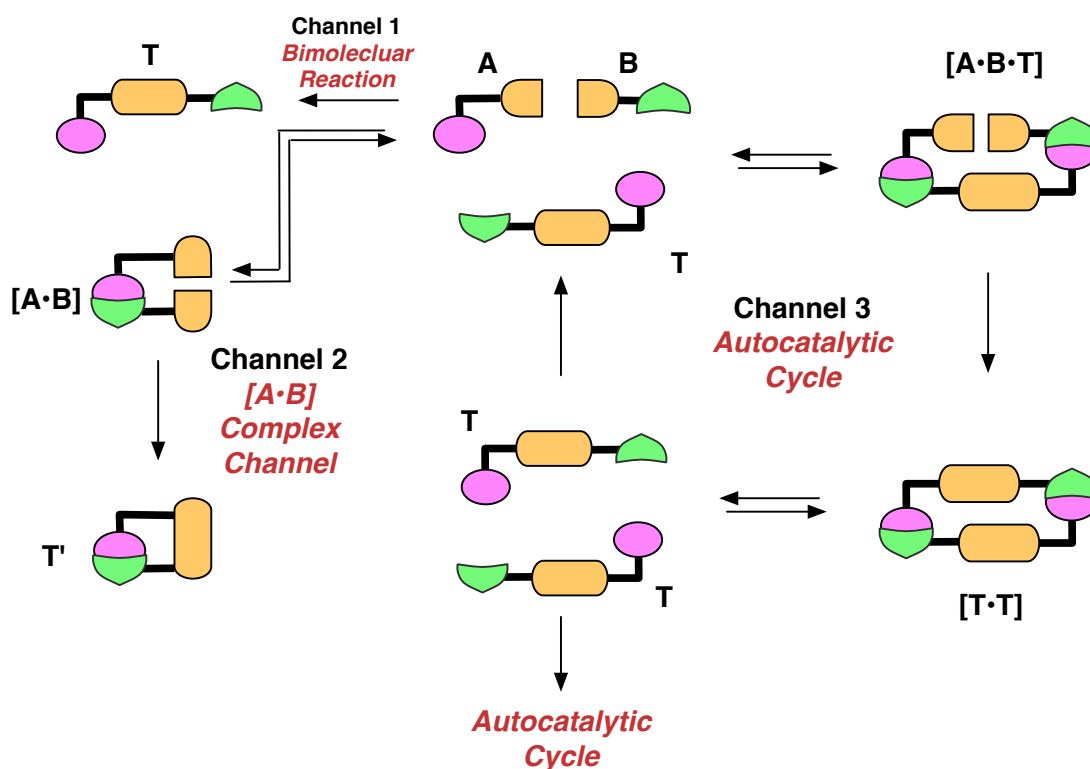
**Figure 1.19:** Possible ternary complexes formed from (i) identical recognition motifs showing four possible ternary complexes and (ii) non-identical recognition motifs showing only one possible ternary complex.

Kelly noted that in his prediction, numerous assumptions were made, one of the most important being that the new recognition site had the same efficacy as the previous site. This assumption was in fact inaccurate. Binding studies showed that the new site had a weaker association by a factor of 27 lower than the original binding site. With this information the mole fraction of ternary complex in the two systems under reaction conditions could be estimated with 84% for the identical recognition system and 41% for the non-identical recognition system. Taking into account that half of the ternary complexes formed in the identical recognition system are non-productive complexes, a more illustrative value of 42% of reactive ternary complexes should perhaps be used. Therefore, with a similar population of reactive ternary complexes in each system, it is clear that the non-identical recognition system is actually twice as efficient.

Kelly's work served to demonstrate the concepts underlying bisubstrate linear template design and that optimisation is possible through rational design. In Kelly's system, the template was designed to be complementary to the substrates and in a sense mimics a form of template directed synthesis exploited by nature in DNA synthesis where, albeit with the aid of enzymes, one strand serves as a template for the production of a complementary strand. Linear templates also have the potential to be designed so that the product is identical to the template. In this case, these templates catalyse the formation of themselves and are known as self-replicators.

### 1.5 Minimal Model of Self-Replication

The minimal model describes the simplest form of non enzymatic self-replication and is the model by which most examples published in this field adhere to.



**Figure 1.20:** The minimal model of self-replication. Two substrates, **A** and **B** are shown with their reactive sites coloured orange and the complementary recognition sites coloured green and purple.

In the minimal model of replication (Figure 1.20) three reaction channels are present. Channel 1 depicts the the simple bimolecular reaction between reagents **A** and **B** in the absence of recognition mediated processes to produce template **T**. In this

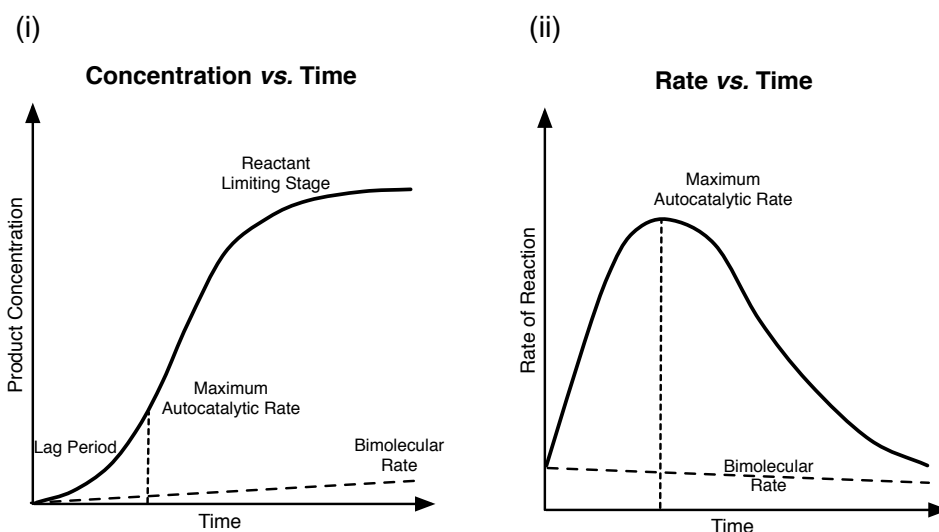
channel, the reaction is facilitated by random collisions and as such is the slow and unselective pathway in the system.

However as **A** and **B** are decorated with complementary recognition sites association of **A** with **B** leads to the formation of the binary complex [**A•B**], Channel 2. Potentially, this complex has the ability to bring the reactive sites into close proximity which can render the reaction *pseudo*-intramolecular and thus is able to proceed at an increased rate relative to the background bimolecular rate. Reaction through this channel leads to the product **T'** in which the recognition sites remain associated together which renders the product catalytically inert. As the approach of the reactive sites has been orchestrated, the formation of **T'** can occur with strict control of stereochemistry in the product. This channel is known as the [**A•B**] Complex Channel.

In contrast to the structure of **T'**, template which has been formed through the Bimolecular Channel 1 to form **T** has its recognition sites open to solution. As displayed in Channel 3, these are able to associate with their complementary substrates to form the ternary complex [**A•B•T**] which orientates the reactive sites of each substrate into close proximity thereby promoting their reaction leading to a template duplex [**T•T**]. In doing so, the template transmits any structural information such as stereo or regio chemistry to the new product as an exact copy is produced. Dissociation of the template duplex then releases two molecules of **T** back into solution, each able to begin another autocatalytic cycle and thus completing an autocatalytic cycle. The autocatalytic cycle can potentially lead to exponential growth of **T** in a closed system until reagents **A** and **B** are consumed.

A characteristic of a self-replicating reaction is the concentration vs. time profile for the evolution of **T** which displays a non-linear sigmoidal growth profile (Figure 1.21 (i)). In the early stages of the reaction, there is no product **T** available to catalyse the reaction and thus product formation is slow through the bimolecular channel only. As product **T** is produced, the autocatalytic cycle is activated and the rate of production of **T** will increase until the maximum autocatalytic rate is reached. Of course in a closed system the reagents will be consumed and thus the rate will dwindle.

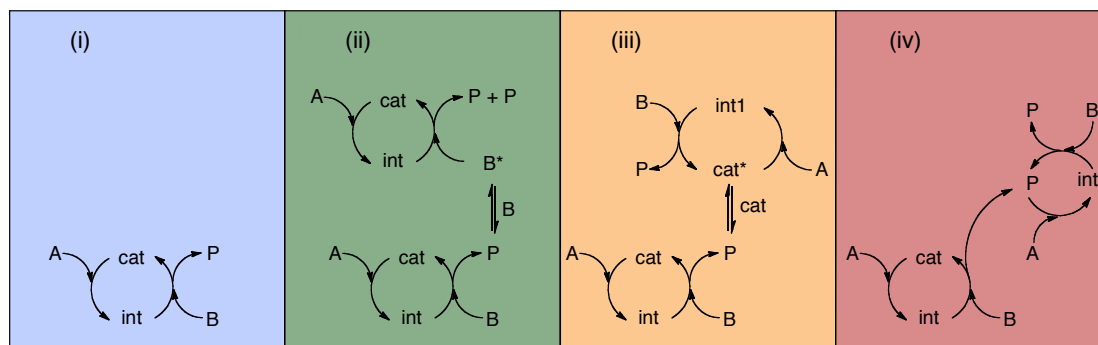
By taking the first derivative of the concentration vs. time profile, a corresponding rate vs. time profile can be constructed (Figure 1.21 (ii)), in which the change in rate with time can be visualised more clearly.



**Figure 1.21** (i) Typical concentration vs. time profile for the formation of a self-replicating product (solid line) and its corresponding bimolecular reaction (dashed) line. (ii) Typical rate vs. time profile of a self-replicating product (solid line) and its corresponding bimolecular reaction (dashed) line. Vertical dashed line on each profile highlights the point at which the maximum autocatalytic rate is reached.

### 1.5.1 Self-Replication and the Origins of Life

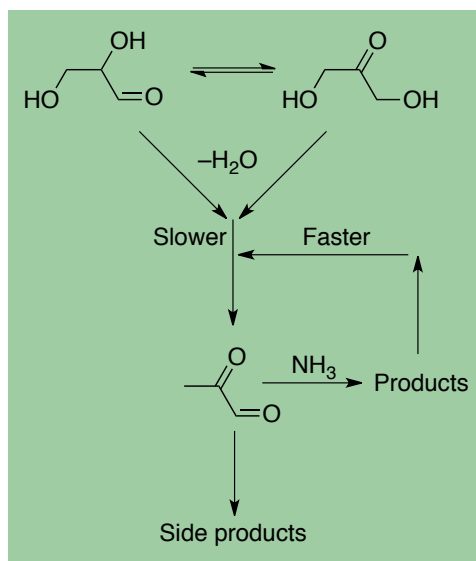
From a synthetic chemist's point of view, self-replication may be seen as an attractive synthetic tool capable of exponential production of a single product over all side reactions. For the biologist, self-replication can be seen as a link to the origins of self-organisation and evolution, key to the origin of life itself. Autocatalytic cycles naturally appear crucial in most discussions on the origins of life as a method to accumulate material in the pre-biotic soup. A critical criteria noted by Orgel<sup>87</sup> was that the success of any autocatalytic system in a prebiotic scenario lies in its long term survival.



**Figure 1.22:** Four catalytic reactions used by Blackmond in theoretical study. All reactions are for a simple  $A + B \rightarrow P$ . (i) Simple catalytic cycle (ii) Simple catalytic cycle plus product enhanced autoinductive cycle (iii) Simple catalytic cycle plus ligand accelerated autoinductive cycle. (iv) Simple catalytic cycle plus product directed autocatalysis. Figure adapted from reference 88.

Blackmond undertook a theoretical study examining the roles of autocatalytic cycles in proposed primordial reactions to examine their feasibility for long term survival.<sup>88</sup> In the study, four cases of catalysis were compared (Figure 1.22).

Case (i) (Figure 1.22 (i)) describes a simple catalytic cycle in which a catalyst intervenes to accelerate the reaction between **A** and **B** to form **P**. This scenario might be visualised as Kelly's bisubstrate template described previously, Section 1.4.3.



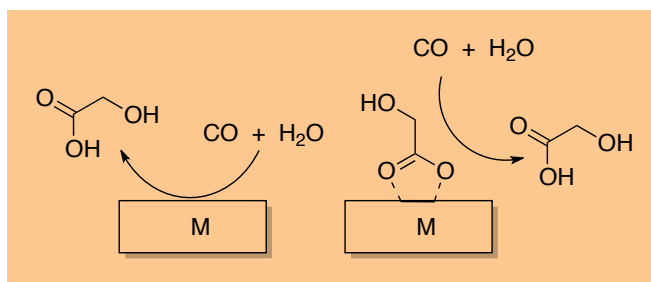
**Figure 1.23:** Weber's sugar model,<sup>89</sup> in which formation of pyruvaldehyde from glyceraldehyde is accelerated by the products of the triose–ammonia reaction.

Case (ii) (Figure 1.22 (ii)), describes a simple catalytic reaction which has been coupled to a second cycle in which interaction of the product **P** with substrate **B** leads to the formation of an activated form of **B** which is more reactive with the same catalytic intermediate to form **P**. This product enhanced network is reminiscent of Weber's autoinductive sugar model<sup>89</sup> in which the formation of pyruvaldehyde from glyceraldehyde is enhanced by a unspecified product of the triose–ammonia reaction (Figure 1.23).

Case (iii) (Figure 1.22 (iii)) describes a simple catalytic reaction which has been coupled to a second cycle in which the product **P** interacts with the catalyst to form a species which is more efficient at promoting the reaction of **A** and **B** though a new catalytic intermediate, int1, to form **P**. This system is reminiscent of ligand accelerated synthesis as described by Wächtershäuser<sup>90</sup> (Figure 1.24) in which the



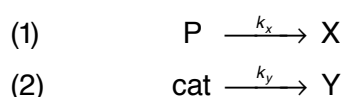
product of a metal catalysed reaction may be able to act as a ligand for the same metal core and increase its catalytic efficiency.



**Figure 1.24:** Proposed autocatalytic reaction by Wächtershäuser in which the product of a reaction is able to bind as a ligand to a catalytic metal core which would render the catalytic core more reactive.

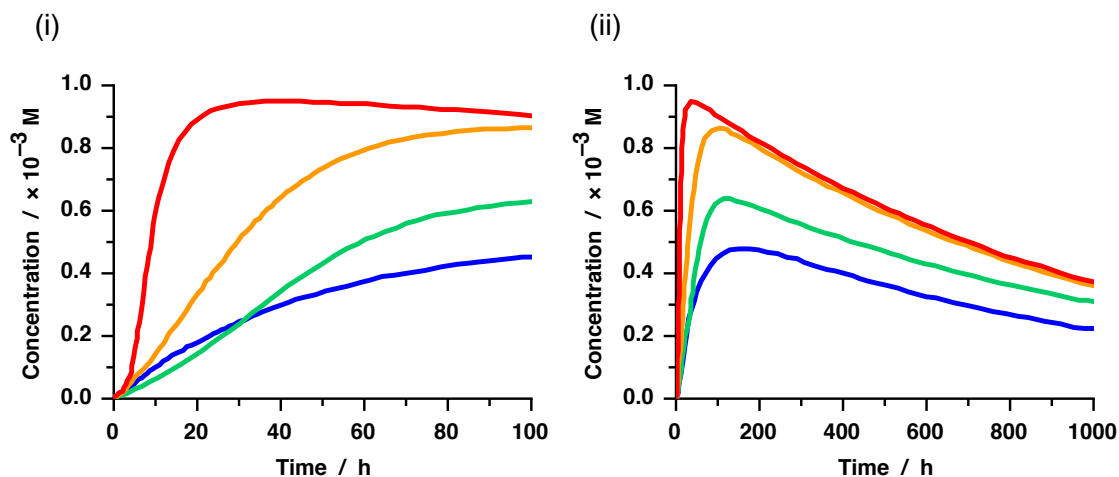
Case (iv) (Figure 1.22 (iv)) describes a truly autocatalytic system in which the product **P** is able to act as a catalyst in a new cycle. Whilst initially **P** is formed by the original catalytic cycle, **P** is then able to independently catalyse the reaction of **A** with **B** and thus the system has evolved away from the requirement for the original catalyst. This system is reminiscent of the minimal model of self-replication (Figure 1.20) however in the minimal model the initial reaction to form **P** is not catalysed instead proceeding through a bimolecular reaction.

In order to assess the longevity of each of the autocatalytic cycles, Blackmond's study assumed firstly, that product **P** is consumed either by reaction or deterioration to **X**, a catalytically inactive breakdown product and secondly, that the catalyst may be subject to irreversible degradation over time, Equation 1.5 (1) and (2) respectively.



**Equation 1.5:** Removal of product **P** to **X** proceeds at rate  $k_x$  and degradation of catalyst **cat** to **Y** proceeds at rate  $k_y$ .

Firstly a kinetic reaction simulation was carried out for each of the four cases in a closed system, which may be envisaged on prebiotic earth as a pond. The systems in the simulation were orchestrated to follow standard Michael-Menten kinetics in which a rapid equilibrium of **A** with **cat** leads to formation of **int** followed by the rate determining reaction with **B** to form **P**. The autoinductive rate constants in case ii and iii were chosen to be ten times greater than that of the original catalytic cycle whilst the autocatalytic cycle of case iv was set to the same rate as the original cycle and thus ten time slower than the autoinductive cycles.

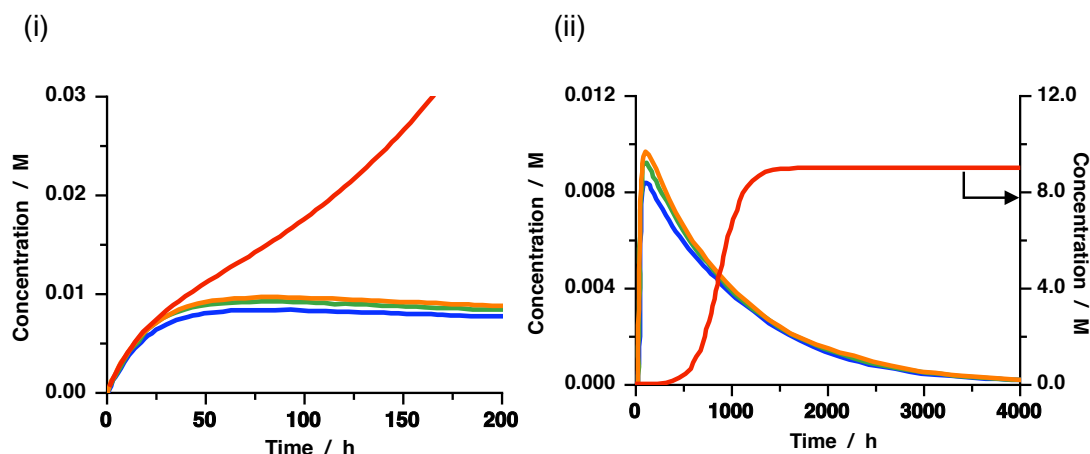


**Figure 1.25:** Results of kinetic reaction simulation by Blackmond of a closed system presented as concentration vs. time profiles for each case with the colours consistent from above descriptions: case (i) in blue, case (ii) in green, case (iii) in orange and case (iv) in red. Concentration vs. time profiles in graph (i) shows the initial stages of the reaction with graph (ii) showing a longer time period of reaction. Data taken from reference 88.

The results of the reaction (Figure 1.25 (i)) showed that all three autocatalytic cycles, cases (ii)–(iv), as well as the original catalytic cycle, case (i), produce the product with high efficiency in the early stages of the reaction but are helpless to stop the slow deterioration of the product concentration as time progresses (Figure 1.25 (ii)) as all of the reagents are consumed. Closer inspection of the initial stages of the reactions highlights that the autocatalytic pathway of case (iv) shows the most significant rate enhancement even though its rate constant for the autocatalytic cycle is ten times slower than in cases (ii) and (iii). This observation serves to demonstrate a characteristic of a true autocatalytic reaction as each turnover increases the concentration of catalyst within the system. In comparison, whilst the the autoinductive cases may increase the catalytic reactivity of the system, the rate of turnover is always limited by the initial concentration of catalyst.

The second scenario investigated by Blackmond replaced the closed system with an open system in which substrates were continually supplied at low contributions to the pool. This scenario has preboitic relevance as it could represent a lake with a river constantly supplying fresh nutrients on one side with another river allowing product and catalyst to flow out.

The results of the open system simulation revealed that in the very initial stages (Figure 1.26 (i)) (up to Time = 40 h) there is very little to distinguish the four catalytic reactions. After this point in cases (i), (ii) and (iii), the production of **P** then reaches a maximum concentration after which point the concentration of **P** decays even though



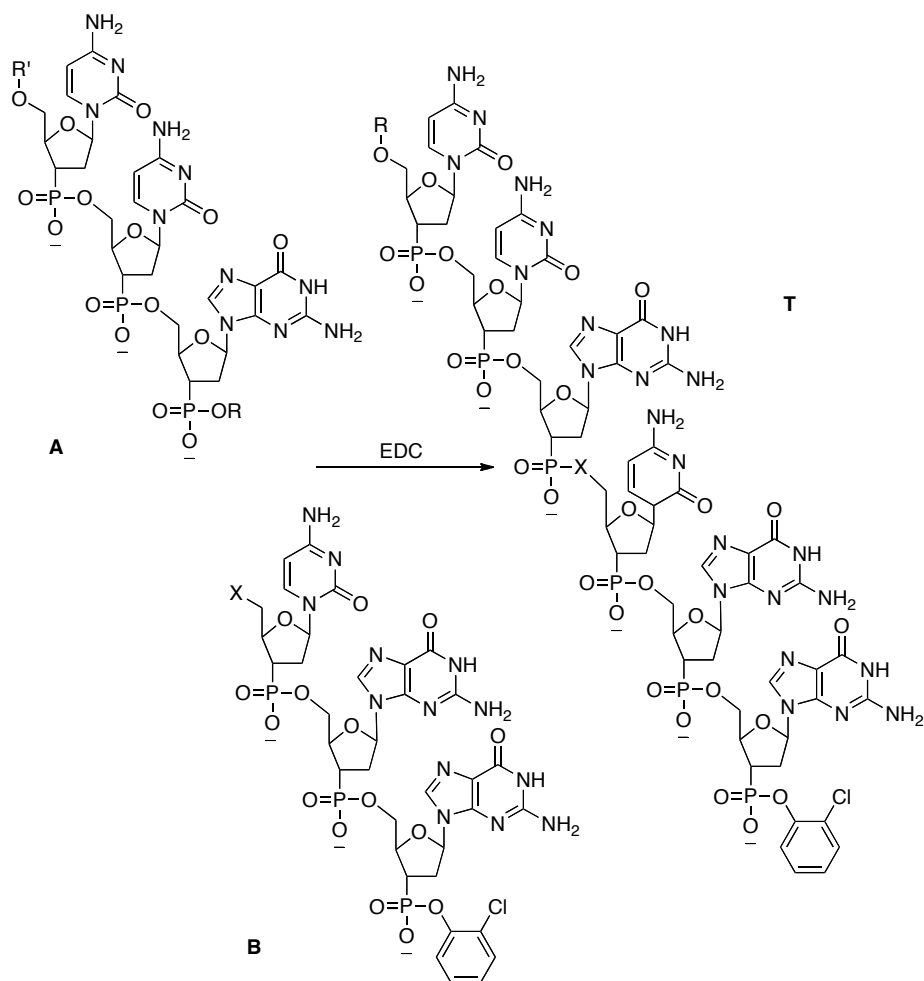
**Figure 1.26:** Results of kinetic reaction simulation by Blackmond of an open system presented a concentration vs. time profiles for: (i) initial stages of the reaction (ii) longer time period of reaction. Note change of scale for case (iv) which uses right hand axis. Data taken from reference 88.

fresh **A** and **B** are being fed to the system (Figure 1.26 (ii)). In contrast, after 40 h, the autocatalytic reaction of case (iv) showed an acceleration in the production of **P**. With a constant supply of reagents, the system eventually reaches a steady state in which the rate of supply of nutrients is equal to the rate of consumption of **P** (Figure 1.26 (ii)).

This study highlighted that the survival of autoinductive reactions (cases (ii) and (iii)) were unfeasible as a result of the degradation of catalyst (**cat**) which was vital to the system. In case (iv), the product itself is able to act as a catalyst for the reaction and thus the system has evolved away from the requirement of the original catalyst. Case (iv) can be experimentally realised in a self-replicating reaction. This study highlights that self-replication is the only viable process of autocatalysis available for the accumulation of material in the prebiotic soup leading to pre-organisation and life itself.

### 1.5.2 Self-Replicating Systems Based on DNA

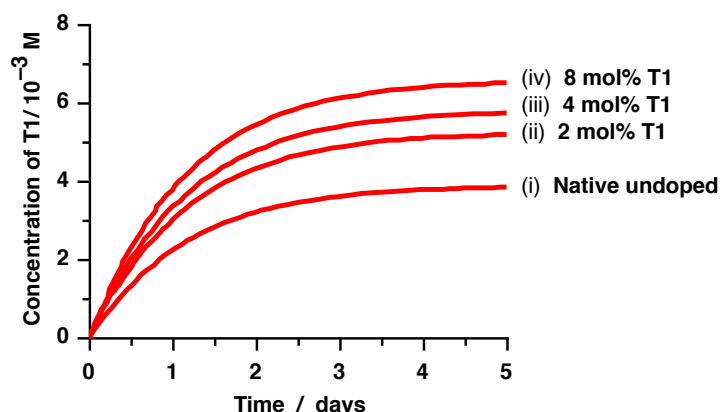
It is not surprising that since nature inspired the field of research into self-replication that the first example of a minimal self-replicating system was based on oligonucleotides. In 1986, von Kiedrowski<sup>91</sup> published the first example of a self-replicating system based on a palindromic sequence of DNA (Scheme 1.12).



A	B	Template
<b>A1:</b> R = CH <sub>3</sub> R' = H	—	—
<b>A1*:</b> R = CH <sub>3</sub> R' =	<b>B1:</b> X = OH	<b>T1:</b> R = CH <sub>3</sub> ,    X = O
<b>A2:</b> R = CH <sub>2</sub> SCH <sub>3</sub> R' = H	—	—
<b>A2*:</b> R = CH <sub>2</sub> SCH <sub>3</sub> R' =	<b>B2:</b> X = NH <sub>2</sub>	<b>T2:</b> R = CH <sub>2</sub> SCH <sub>3</sub> ,    X = NH

**Scheme 1.12:** Generic structure of the self-replicating hexanucleotide by von Kiedrowski with structural modifications described in the table below. The first self-replicating template reported was<sup>91</sup> **T1** produced from the reaction of **A1**(\*) with **B1**. Subsequent modification of the system lead to the the observation of sigmoidal growth<sup>92</sup> form the self-replicating reaction to form template **T2** from the reaction of **A2**(\*) with **B2**.

The substrates were; trinucleotide **A1** with base sequence 5'-CCG-3' with the 5' protected as the methyl ester and trinucleotide **B1**, with complementary sequence 5'-GGC-3' with the 3' end protected with the *o*-chlorophenyl group. The template **T1** formed from **A1** and **B1** was a self-complementary hexanucleotide with base sequence 5'-CCGCGG-3' likewise with the 5' terminus protected as a methyl ester and the 3' terminus protected with the *o*-chlorophenyl group. Recognition between the substrates and template was facilitated by cytosine-guanine base pairing. Reaction of two trimeric nucleotide substrates at 10 mM in aqueous solution is initiated by EDC activation of **A1** to form the intermediate **A1\***. Condensation with the 5' hydroxyl group on **B1** then leads to phosphoester formation and yields the hexanucleotide template **T1**.

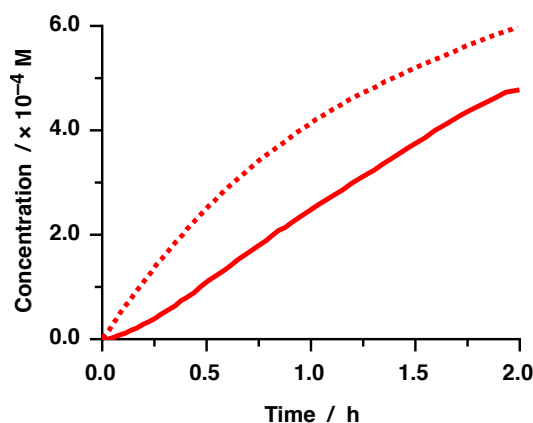


**Figure 1.27:** Concentration vs. time profile for the production of **T1** with varying initial concentrations of **T1**: (i) 0 mM **T1**, (ii) 0.2 mM **T1**, (iii) 0.4 mM **T1** and (iv) 0.8 mM **T1**. Data taken from reference 91.

The progress of the reaction with and without preformed **T1** was monitored by HPLC and the results presented as a concentration vs. time profile (Figure 1.27) which revealed that in the absence of added template, (i), the reaction proceeds slowly. Increasing the initial concentration of **T1**, (ii), (iii) and (iv) had the effect of increasing the rate of formation of **T1** thus demonstrating autocatalysis. Sigmoidal growth was not observed in this system which also displayed a very poor conversion of only 12%. von Kiedrowski explained these observations were as a consequence of a number of reasons. Firstly the CDI activated intermediate **A\*** underwent rapid hydrolysis to unreactive **A** which was itself reactive with **A\*** and could undergo self condensation to form the pyrophosphate 5'-MeCCG<sup>3'</sup>pp<sup>3'</sup>GCC-Me<sup>5'</sup> (**pp**) which was observed by HPLC to be present in similar concentrations to **T1**. The presence of the **pp** template in solution would act as an inhibitor as it will facilitate unproductive binding events with **B1**.

The sequence of base pairs on the added template was of utmost importance.<sup>93</sup> Significantly slower template directed reactions were observed when non-complementary base pairs were introduced. The closer the non-complementary base pairs were to the reactive site the poorer the rate. Only fully complementary templates provided a strong rate enhancement.

von Kiedrowski went on to modify the system by changing the reaction forming event. The change allowed sigmoidal growth, characteristic of an autocatalytic reaction, to be observed for the first time.<sup>92</sup>

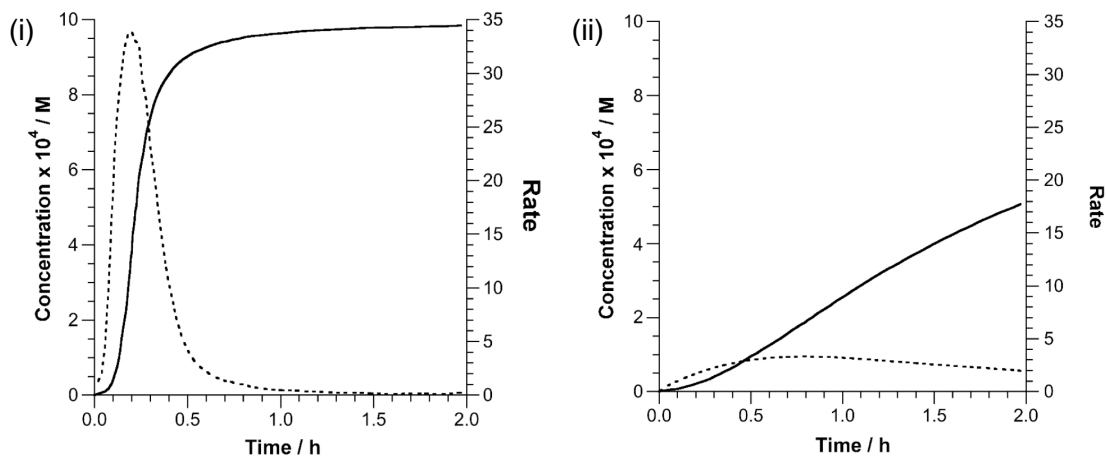


**Figure 1.28:** Concentration vs. time profile for the production of **T2** from **A2** and **B2** in the absence of added **T2** (red line) and in the presence of 0.32 mM (32 mol%) of **T2** (dotted red line). Reaction undertaken at 1 mM concentration of **A** and **B**. Data taken from reference 92.

In changing **B1** from a hydroxyl group into an amino group, **B2**, the ligation with **A\*** proceeded more rapidly. After monitoring the reaction by HPLC, a concentration vs. time profile for the untemplated reaction (Figure 1.28, red line) demonstrated the rate of formation of **T2** is far greater (in comparison to **T1**) with close to 50% conversion after 2 h. **T2** was also the only observed product (no **pp** observed). The rate of formation of **T2** was seen to increase with time, characteristic of an autocatalytic reaction. Autocatalysis was demonstrated by addition of 32 mol% preformed **T2** to a fresh set of reagents for the reaction. The concentration vs. time profile (Figure 1.28, dotted red line) demonstrated an increase in the production of **T2** with a loss of the sigmoidal shape.

In what was termed the square root law, the initial rate of reaction was found to be proportional to the square root of the concentration ( $c^p$  where  $p = 0.5$ ) of added template. This was rationalised on account of a strong product duplex not allowing all of the added template to be catalytically available.  $p$  was later<sup>94</sup> described to be the reaction order which gave an insight into the autocatalytic efficiency and thus extent

of product inhibition of a replicator. For a perfect replicator operating with no product inhibition, exponential growth would be observed (Figure 1.29 (i)) and with the reaction order  $p = 1$ . In a strongly product inhibited replicator, sigmoidal growth would be observed (Figure 1.29 (ii)) and the reaction order  $p = 0.5$ .<sup>95</sup>



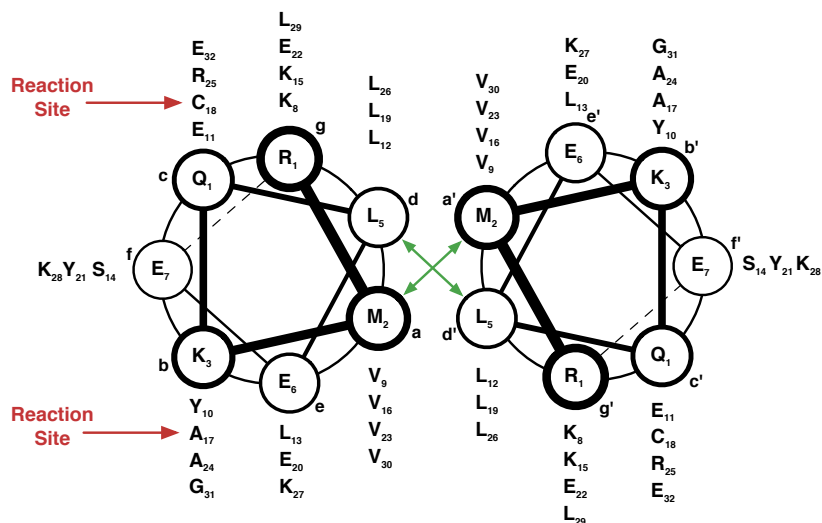
**Figure 1.29.** Simulated concentration vs. time (solid line) and rate vs. time (dashed line) profiles for replicators displaying (i) exponential and (ii) parabolic growth. Figure adapted from reference 95.

The DNA based self-replicating systems described by von Kiedrowski *et al* served as the basis for the construction of extended replicating networks which combined more than one replicator in one system which will be discussed later.

### 1.5.3 Self-Replicating Systems Based on Polypeptides

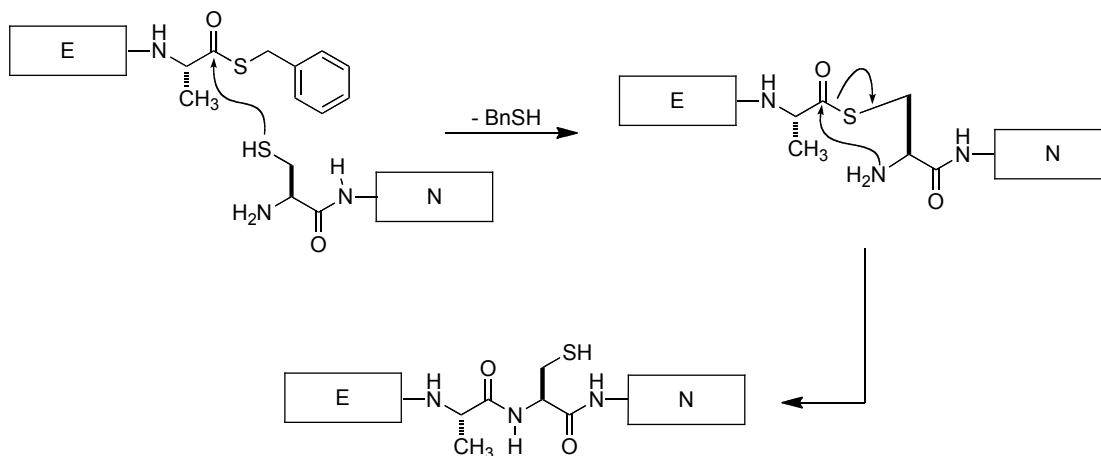
In another biomimetic study, polypeptides have also been successfully utilised to create self-replicating systems. Ghadiri *et al* reported<sup>96,97</sup> a 32-residue  $\alpha$ -helical coiled coil peptide similar to the yeast transcription factor GCN4<sup>98</sup> which acts as a template to assemble its components, a 15 and 17 residue fragment, utilising hydrophobic interactions as recognition to accelerate a native Kent<sup>99</sup> ligation reaction (Figure 1.30).

The  $\alpha$ -helical peptide structure is distinguished by the amphiphilic primary sequence of a repeating heptad (abcdefg)<sup>n</sup>. The interhelical recognition surface is a hydrophobic core derived from the the first (a) and fourth (d) amino acid residues. The reaction site was designed to lie on the solvent exposed surface (ligating between (b) and (c), shown as arrows on Figure 1.30, to avoid interference with the hydrophobic binding site. Interhelical secondary electrostatic interactions of amino acids at the (e) and (g) positions give rise to structural effects within the coiled coil.



**Figure 1.30:** Helical-wheel diagram of the Ghadiri's peptide template in the dimeric  $\alpha$ -helical coiled coil configuration to display the repeating heptad motif. The intermolecular recognition is facilitated by hydrophobic packing interactions at positions *a* and *d* shown as green arrows and electrostatic interactions at *e* and *g*. The ligation site, highlighted by red arrows, is between a cysteine residue 18 at position *c* and alanine residue 17 at position *b* on the solvent exposed face. Figure adapted from reference 96.

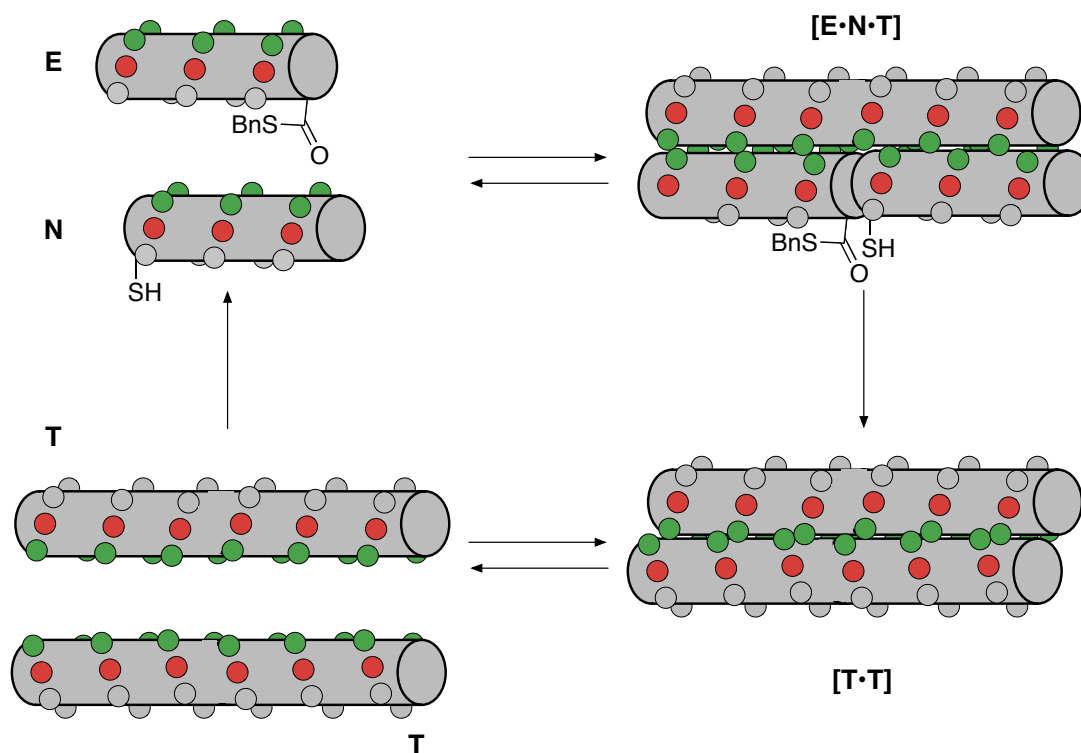
The reaction had to be carried out at a specific pH and ion concentration so that the secondary structure would be maintained. The reaction, a Kent ligation, is a thioester promoted amide bond formation (Figure 1.31).



**Figure 1.31:** Mechanism of Native <sup>99</sup> ligation reaction. The first step is a transthioesterification of the cysteine residue on the nucleophilic fragment with the activated C terminus of the alanine residue on the electrophilic fragment. The intermediate thioester rapidly rearranges to furnish the amide bond.

The first step in the ligation is the transthioesterification of the activated C terminus of the electrophilic peptide fragment **E** with the *N*-terminal cysteine side chain of the nucleophilic peptide fragment **N**. The intermediate then rapidly rearranges *via* an intramolecular S to N acyl transfer to furnish the amide product and an exact copy of the template.

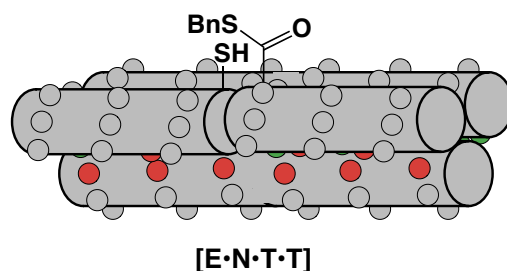




**Figure 1.32:** Autocatalytic cycle for a peptide replicator. Inter-helical recognition sites are shown as green circles and electrostatic interaction residues as red circles.

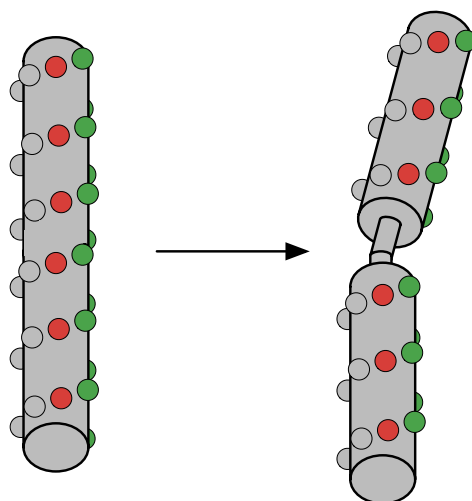
The catalytic cycle (Figure 1.32) is initiated by association of **E** and **N** on to template **T** to form the ternary complex **[E·N·T]**. The ternary complex orientates the reactive sites into close proximity which renders the Kent ligation reaction *pseudo*-intramolecular and thus proceeds at an increased rate relative to the bimolecular ligation to form the template duplex **[T·T]**. Dissociation of the duplex to free two template molecules completes the catalytic cycle. Sigmoidal growth was observed and autocatalysis was demonstrated by addition of preformed **T** into fresh reagents, the result of which was to increase the initial rate of reaction. The reaction order was determined to be 0.63 indicating that product inhibition still reduced the autocatalytic efficiency.

The authors later described that the template duplex, **[T·T]**, was itself capable of reversibly associating the substrates and promoting their reaction through a quaternary complex **[E·N·T·T]** (Figure 1.33) thus adding to the autocatalytic efficiency of the system as the stable template duplex was not completely inhibiting autocatalytic turnover.



**Figure 1.33:** Quaternary structure of a template duplex associating with the substrates.

Chmielewski *et al* were also interested in the design of peptide replicators. Initially peptide replicators sensitive to environmental changes such as pH<sup>100</sup> and ion concentration<sup>101</sup> were produced by exploiting the structural dependence of the coiled coil on electrostatic interactions arising from residues at positions *e* and *g* in the heptad repeat. Their attention was then drawn to increasing the autocatalytic efficiency by reducing inhibition arising from the strong template duplex. Their first approach was to shorten the polypeptide chain from 35 to 26 residues in an attempt to destabilise the coiled coil.<sup>102</sup> The result was a reduction in product inhibition and a reaction order of 0.91.



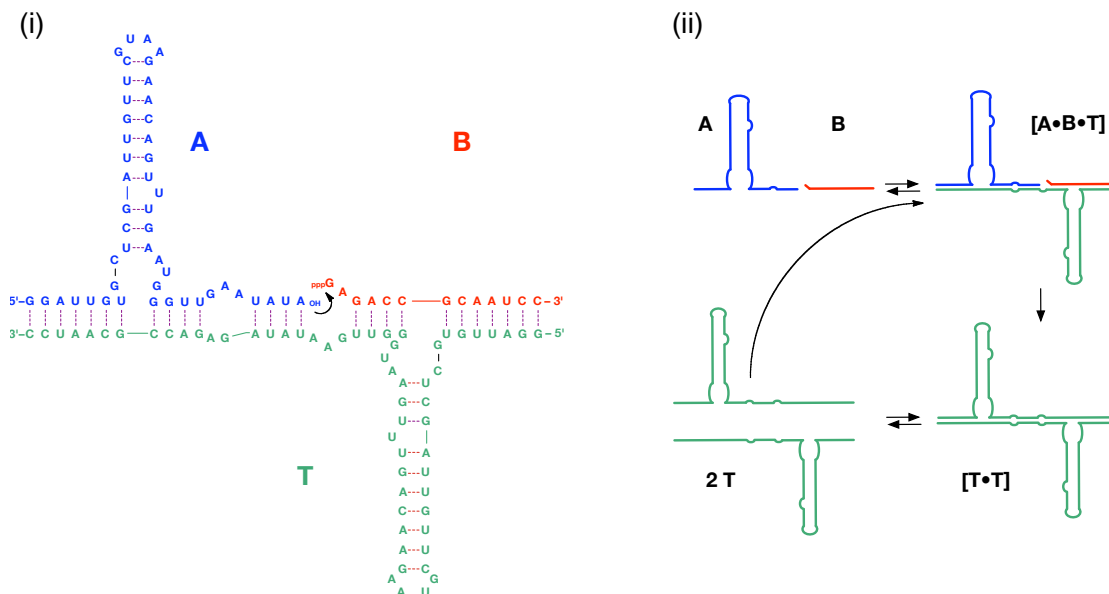
**Figure 1.34:** Modification of the peptide template by introducing a single proline residue leads to a geometrical change in the template structure. The addition of a so called 'kink'.

The most recent report utilised the original 35 peptide template which was modified by introduction of a single proline residue. The result was a 'kink' in the product structure (Figure 1.34) which reduced the stability of the coiled coil product duplex, reducing product inhibition and increasing the reaction order to 0.91.<sup>103</sup> It would be interesting to observe if a further increase of reaction order could be achieved by introducing the proline kink into the shorter template therefore achieving a truly

exponential replicator. The peptide replicators described by Ghadri and Chmielewski were used as the basis of extended replicating systems and logical operations which will be discussed later.

### 1.5.4 Self-Replicating Systems Based on RNA

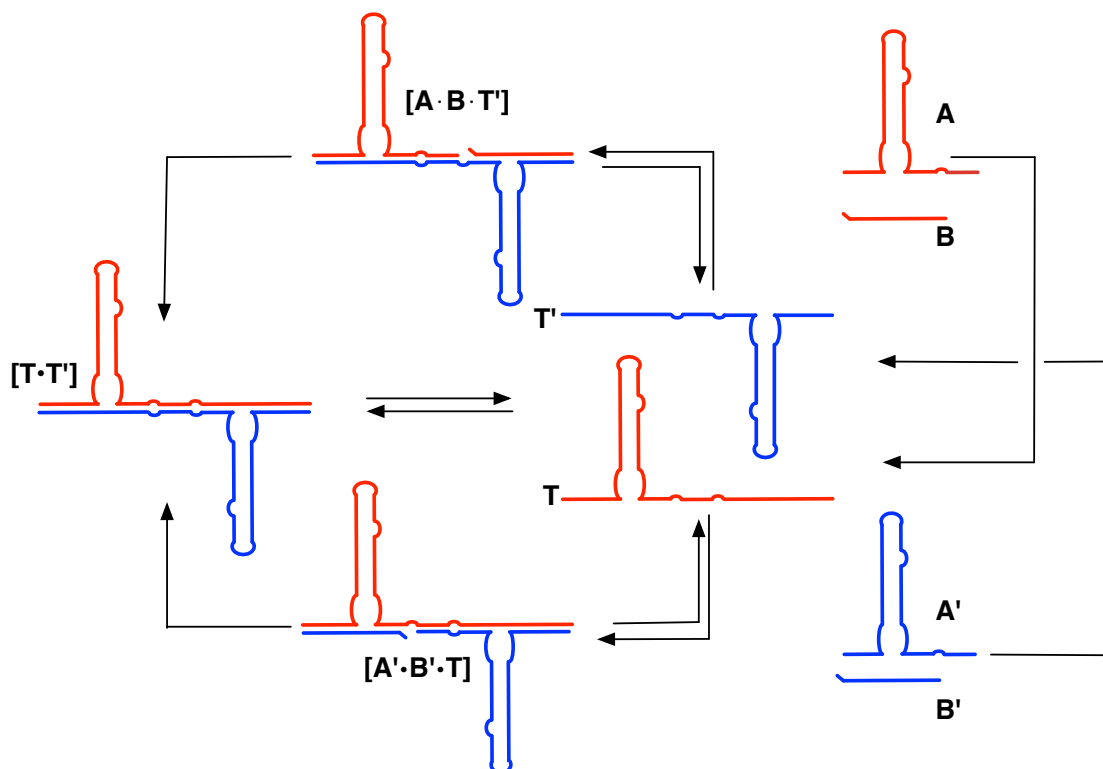
In the search for the origin of life based on DNA as we know it, the existence of an RNA world,<sup>104</sup> one which used the information storage and catalytic<sup>105</sup> properties of RNA, is considered as a plausible scenario. Drawing on inspiration from this theory, Joyce, an advocate of the RNA world scenario,<sup>106</sup> demonstrated<sup>107</sup> a self-replicating ligase ribozyme (Figure 1.35 (i)). Basing the design on the R3C ribozyme, an RNA ligase that binds two oligonucleotides through Watson-Crick base pairing and catalyses the reaction of a 3'-hydroxyl (OH) group on one substrate with the 5'-triphosphate (ppp) group of the other to form a 3',5'-phosphodiester linkage between two RNA molecules and releasing inorganic phosphate in the process. The ribozyme was redesigned so that it would connect two substrates to produce an exact copy of itself (Figure 1.35 (ii)).



**Figure 1.35:** RNA based self-replicating system. (i) Structure of the ternary complex with reagents shown in red and blue, the template in green and base pairing serving as intermolecular recognition shown as purple dashed lines. (ii) The autocatalytic cycle by which T associates the substrates A and B into a ternary complex [A•B•T] before reacting to form an exact copy of itself and the template duplex [T•T]. Subsequent dissociation releases two template molecules. Colours follow the structure in (i).

Although problems were encountered from very strong  $[A \cdot B]$  complex formation Joyce demonstrated that the initial rate of reaction of  $A$  with  $B$  to form  $T$  is accelerated when the reaction is seeded with preformed  $T$ . After this initial acceleration however, the rate then follows the unseeded reaction rate. Joyce commented that when preformed  $T$  is added to the reaction, the ternary complex is immediately formed. As a substoichiometric quantity of  $T$  is added, there will still remain free  $A$  and  $B$  in solution which form the strong  $[A \cdot B]$  complex. When  $T$  becomes available after autocatalytic turnover, it is not able to break up the strong  $[A \cdot B]$  complex allowing for further template directed reaction.

Joyce commented at the time that as the R3C ligase is amenable to sequence modifications<sup>108</sup> and *in vitro* evolution<sup>109,110</sup> might be used to achieve a more efficient self-replicating structure. Alternatively, the structure could be tailored to express a different mechanism of replication, namely reciprocal replication, in which the  $+$  strand would act as template for the formation of the  $-$  strand from its components and *vice versa* (Figure 1.36).

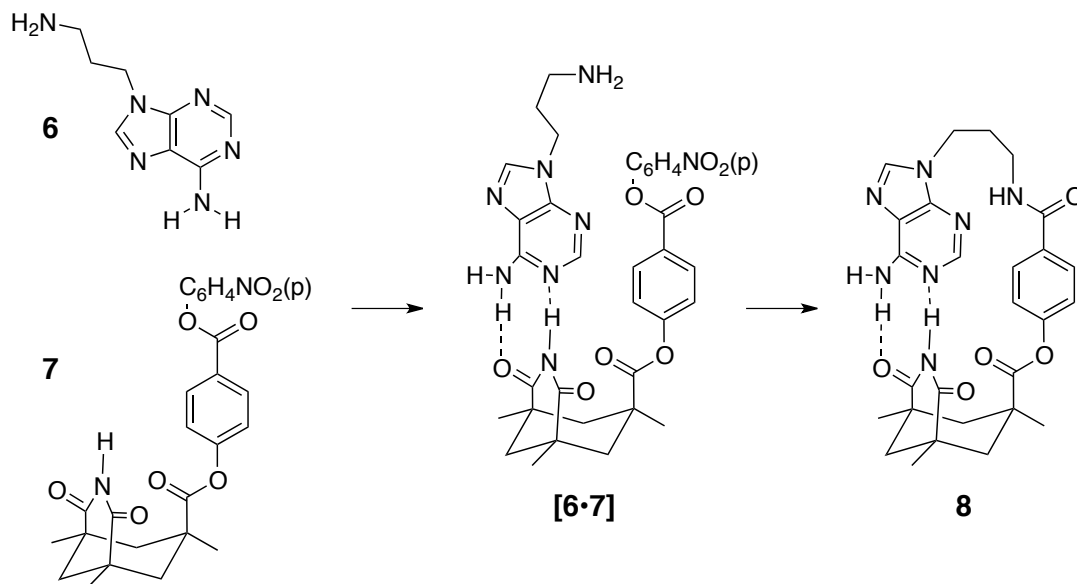


**Figure 1.36:** Reciprocal replication cycle. Ribozyme  $T$  (red) catalysis the ligation of blue substrates  $A'$  and  $B'$  to form the reciprocal template  $T'$  (blue). Likewise,  $T'$  is able to catalyse ligation of red substrates  $A$  and  $B$  to form  $T$ . Figure adapted from reference 111.

This method was realised in subsequent publications<sup>111</sup> which then went on to demonstrate self-sustained replication of an RNA enzyme<sup>112</sup> using serial transfer experiments in which 4% of the product from one system was added to fresh batch of reagents for the next and repeated six times with the products autocatalytically produced in each case.

### 1.5.5 Synthetic Minimal Replicators

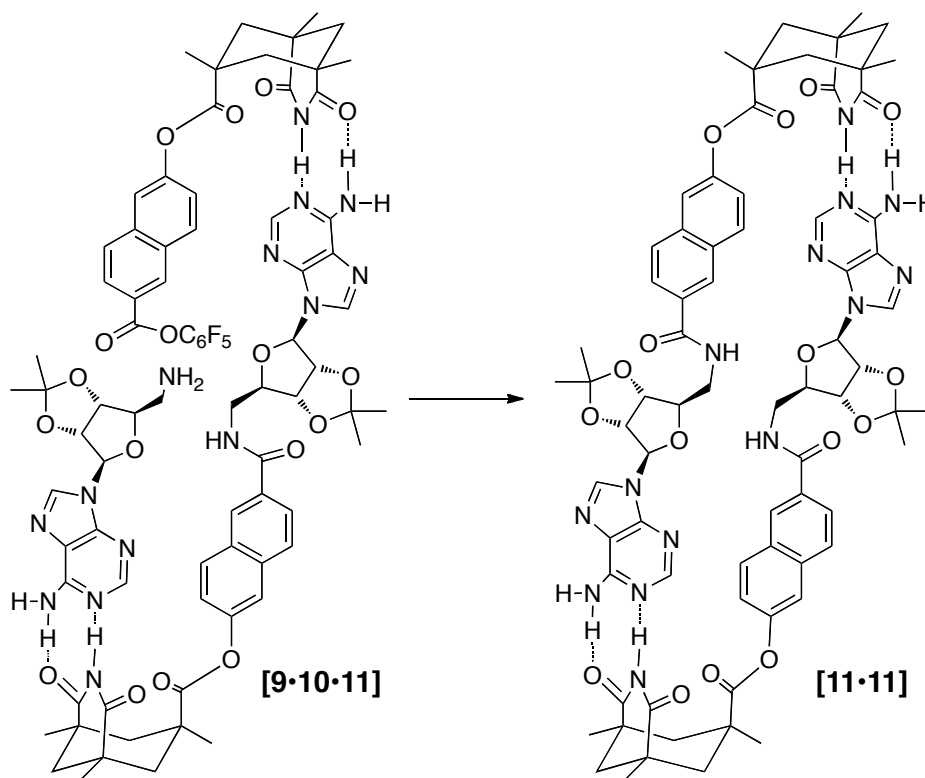
The first wholly synthetic self-replicating system was described by Rebek *et al* in 1990. Still drawing on inspiration from nature, the recognition motif in the system was the association of Kemp's triacid and adenosine previously observed in Rebek's laboratory.<sup>113</sup> The bond forming reaction was the amide bond formation by aminolysis of a pentafluorophenolate ester by a primary amine. Rebek's work served to demonstrate the importance of the design of the group which separates the recognition site from the reactive sites, the so called spacer group. Rebek's initial attempt at the design of a minimal replicator utilised a spacer which was too short as well as a flexible chain which resulted in the formation of a reactive [A•B] complex<sup>114</sup> (Scheme 1.13).



**Scheme 1.13:** Initial attempts to design a minimal replicator from Rebek used substrates which were too flexible and formed closed products through the [A•B] complex channel.

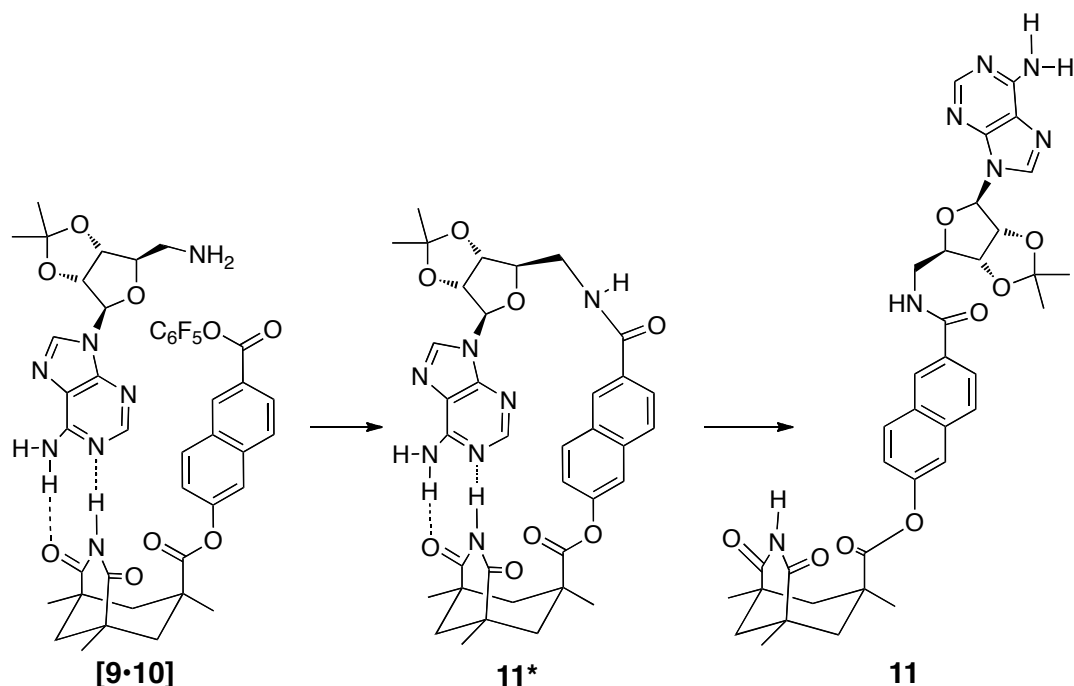
In response to this result, a longer spacer unit was incorporated by changing the phenyl ring into a naphthalene spacer on the Kemp's acid substrate **10**. A ribose ring was also incorporated into the adenine substrate **9** to reduce the flexibility of the

amine and the steric bulk of the ring should serve to destabilise the template duplex [11•11] and thus prevent product inhibition<sup>114</sup> (Scheme 1.14).



**Scheme 1.14:** Ternary complex proposed by Rebek with the two substrates clearly associated by Kemp's triacid – adenosine complexation which brings the reactive sites into close proximity.

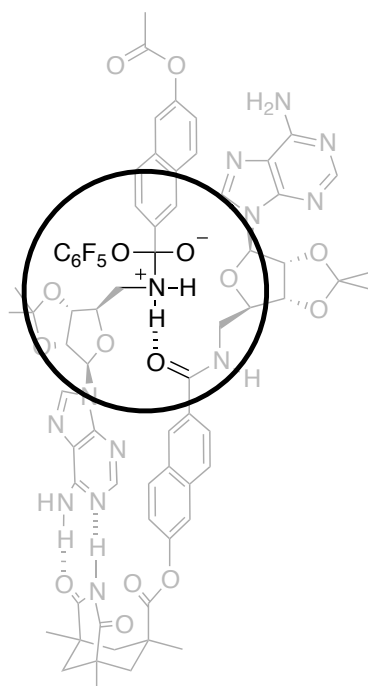
The authors failed to see sigmoidal growth of the product however they did give evidence to support a self-replicating mechanism in the form of template effects. Rebek found that addition of prefabricated template increased the initial rate of formation of the product demonstrating the template directed nature of the reaction. The effect of recognition was demonstrated to be important as the rate of reaction is 10 times faster when using the free *N*-H than with the *N*-methyl derivative which was used in a control reaction to inhibit recognition. The absence of a sigmoidal growth profile was explained by the ability of the substrates to form a reactive [A•B] complex [9•10] which forms the *cis* amide intermediate. Isomerisation of the amide bond to the more favourable *trans* amide then furnishes the desired product (Scheme 1.15).



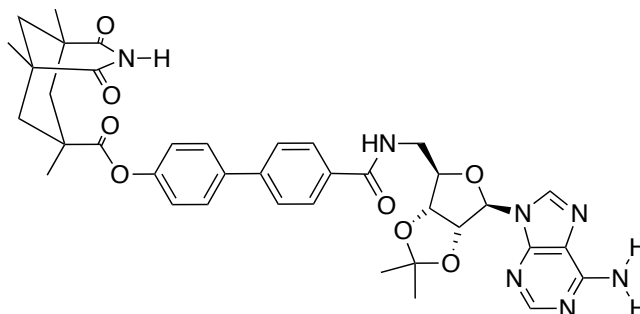
**Scheme 1.15:** Association of the substrates into an **[A•B]** complex **[9•10]** promotes the formation of the *cis* amide **11\*** by intramolecular reaction before *cis* to *trans* amide isomerisation furnished the open template **11**.

The mechanism of replication was the cause of significant controversy. Menger<sup>115</sup> proposed an alternative explanation which was that the proximity of an amide in the template to the bond forming reaction would stabilise the zwitterionic intermediate formed by the aminolysis of the ester (Figure 1.37). Rebek disagreed<sup>116</sup> and the two continued to publish evidence to support their own theories<sup>117,118</sup> until ultimately Reinholdt intervened.<sup>119</sup> To determine to what extent each mechanism was contributing to the overall mechanism, a detailed kinetic analysis was undertaken which concluded that the mechanism of self-replication originally proposed by Rebek was operating in the system however its observation was obscured by other pathways in the system such as those proposed by Menger.

Whilst Rebek's system was eventually shown to behave as a self-replicator, this chapter served to demonstrate the pitfalls easily fallen into in replicator design and the overall inefficiency of the system lead Rebek to subsequently redesign to the system. A redesigned replicator was produced which avoided the original pitfall of the production of the reactive **[A•B]** complex by incorporating a longer still biphenyl spacer to replace the naphthalene spacer and sigmoidal growth was observed (Figure 1.38).<sup>120</sup>



**Figure 1.37:** Menger's control reaction of a control template bearing a single adenosine recognition site with recognition enabled amine **9** and a recognition disabled activated ester. Circle to draw attention to the tetrahedral intermediate demonstrating that the carbonyl amide is in close proximity to the ammonium ion allowing for stabilisation of the developing positive charge.



**Figure 1.38:** Second generation of Rebek replicating template from which a sigmoidal growth profile was observed.

A new generation of replicating structures **12** based on a xanthene scaffold with intermolecular recognition provided by a diaminotriazine and pyrimidine (Figure 1.39 (i)), were demonstrated<sup>121</sup> by Rebek in the same time period. Recently,<sup>122,123</sup> these structures have undergone a renaissance and have been modified to incorporate organocatalyst elements (Figure 1.39 (ii)). The result of the combination is the creation of a self-complementary template which has both self-replication and metabolism behaviour.

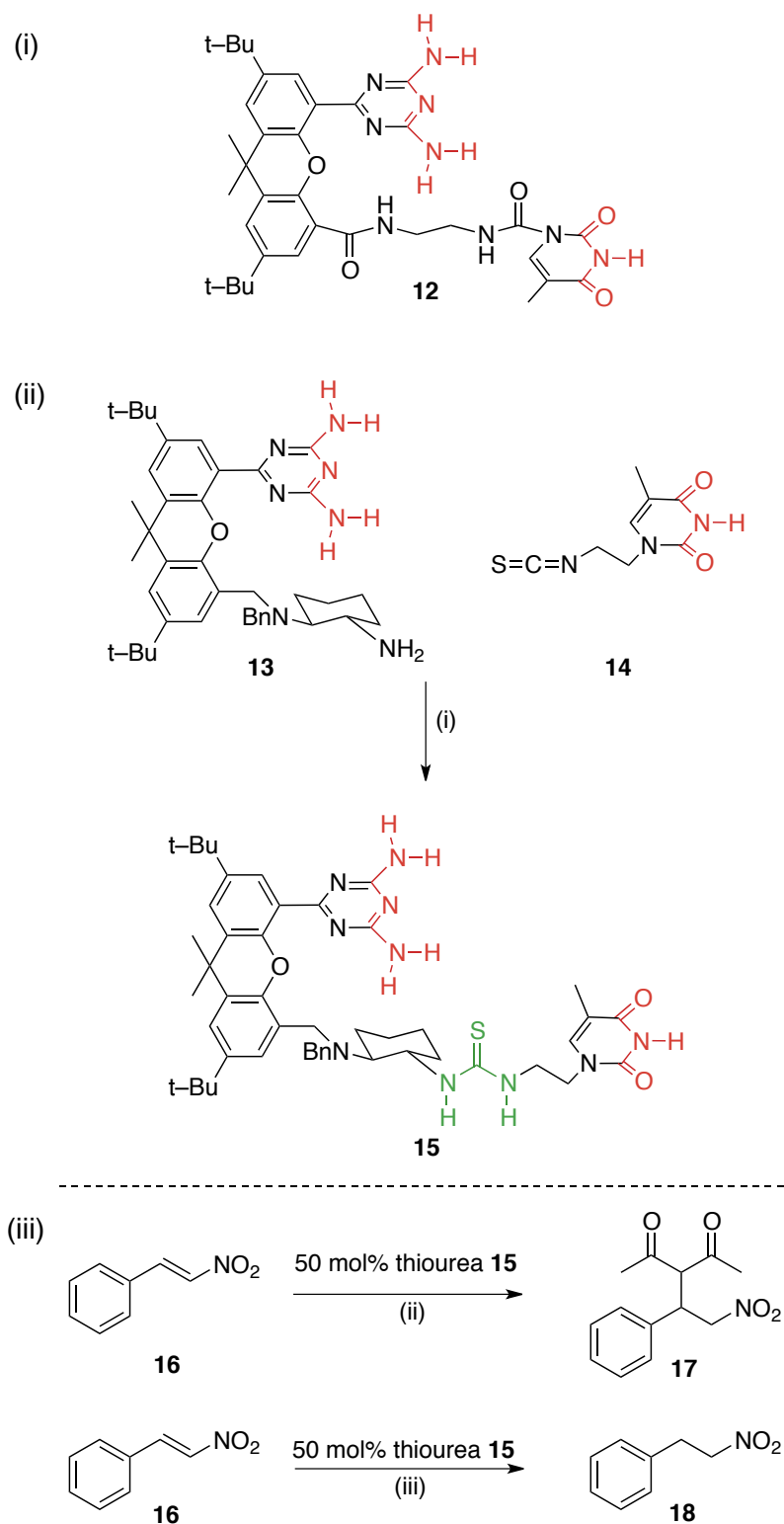


The formation of thiourea **15** from amine **13** and isothiocyanate **14** was shown to be autocatalytic by competition experiments with recognition disabled control compounds which required low temperatures as the thiourea formation was too rapid to monitor at room temperature.

With the autocatalytic behaviour of the template established, the authors demonstrated that the thiourea unit was capable of organocatalysis (Figure 1.39 (iii)) catalysing a Michael addition of 2,5-pentanedione on nitrovinyl benzene **16**. In the absence of thiourea **15**, the Michael addition reaches 60% after 28 h. In the presence of either half of the organocatalyst, 50 mol% amine **13** or 50 mol% isothiocyanate **14**, the conversion was similarly 60%. In the presence of thiourea **15** on the other hand, the conversion had reached 90%. These results demonstrated that the catalytic function arose from the thiourea moiety and not from either half of the template. An interesting experiment also combined amine **13** with isothiocyanate **14** with the Michael addition reactants and allowed formation of the organocatalyst *in situ*, which then went on to catalyse the Michael addition to 90% again, a not altogether unexpected result as the thiourea formation would take place in a matter of minutes whilst the Michael addition required a further 26 h.

A transfer hydrogenation to reduce the same nitroolefin **16** to saturated product **18** was also demonstrated to be catalysed by thiourea **15**. In the absence of thiourea **15**, the reduction of nitroolefin **16** with Hantzsch ester reached about 2% conversion in 6 days. Again, control reactions with either amine **13** or isothiocyanate **14** were carried out to demonstrate that each half of the template did not produce a catalytic effect. These control reactions reached 8% and 10% conversion respectively after 6 days. In the presence of thiourea **15** on the other hand, the reaction reached 94% after 6 days. Once again, the template was produced *in situ* under the transfer hydrogenation conditions and allowing the reduction reached the same overall conversion.

In each case, however, the organocatalyst thiourea **15** showed lower catalytic activity than a recognition disabled control template which was prevented from forming a template duplex by methylating the *NH* on the pyrimidinone. This observation demonstrated that the organocatalytic activity of the thiourea is decreased by dimerisation of the self-complementary template. A further example of an



**Figure 1.39:** (i) Structure of an alternative self-complementary replicating template reported<sup>121</sup> by Rebek in 1992 with the recognition sites highlighted in red. (ii) Reaction of isothiocyanate **14** with amine **13** leads to the formation of thiourea **15**, a self-complementary template capable of autocatalytic formation of itself and organocatalysis from the thiourea site (green) described<sup>123</sup> in 2009. (iii) Reactions catalysed by the presence of thiourea **15**. Reagents and conditions (i) 5% NEt<sub>3</sub>, CH<sub>2</sub>Cl<sub>2</sub>, -78 °C, (ii) 50 mol% **15**, 5 eq. 2,5-pentanedione, 5 eq. NEt<sub>3</sub>, CDCl<sub>3</sub>. (iii) 50 mol% **15**, 10 eq. Hantzsch ester, CDCl<sub>3</sub>.

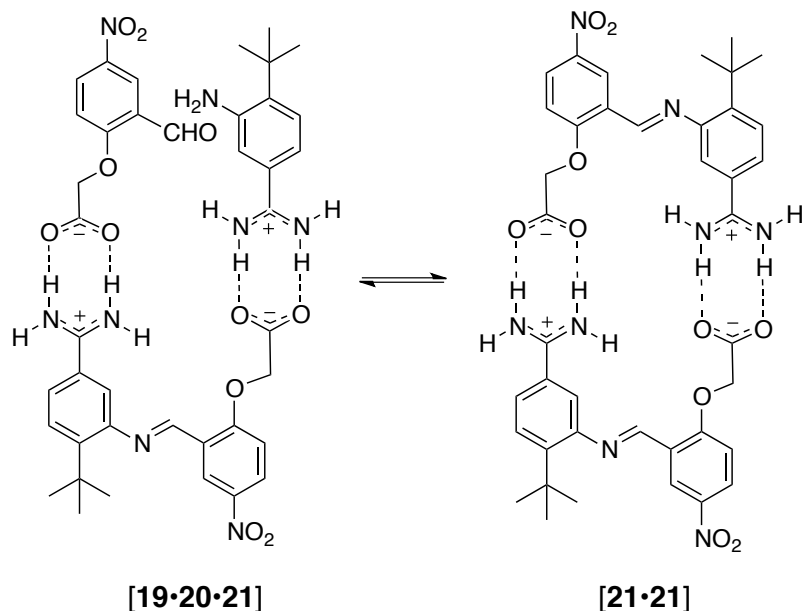
organocatalyst produced by autocatalysis using a similar scaffold was more recently produced<sup>122</sup> by Rebek *et al* based on the same principles, this time producing a MacMillan type<sup>124</sup> imidazolidinone which was shown to be capable of accelerating two reactions: a reduction of cinnamaldehyde with Hantzsch ester and a Friedel–Crafts alkylation of cinnamaldehyde in both cases by activating the  $\alpha,\beta$ -unsaturated aldehyde as an iminium ion.

Whilst Rebek's work is the only example of replicating template which has been demonstrated to act as an organocatalyst for a subsequent reaction, the evidence to support the autocatalytic formation of the template, arguably the most important property of the system, is arguably scarce with no demonstration of a concentration vs. time profile. Given the controversy which arose from Rebek's initial publications of self-replicating templates, we would have hoped a greater emphasis would be given to a comprehensive characterisation of the autocatalytic reaction.

An ideal goal for this work would be to achieve a system in which multiple catalytic processes are able to occur *in situ*. Firstly the covalent bond forming event in the autocatalytic cycle would produce the organocatalytic site and that this catalytic site is more active as part of the template than in an isolated state. In order to achieve a degree of stereochemical control of the organocatalytic action, the recognition motif used for constructing the template may also serve to associate the substrates for organocatalysis, bringing them into close proximity to the catalytic site and thus accelerating the rate of catalysis and achieving a degree of control in the stereochemical nature of the product. Nonetheless, Rebek's work presents a demonstration of what the authors coin 'autocatalytic reproduction of catalytic capabilities'.

The covalent bond forming event in a self-replicating system does not necessarily need to be an irreversible reaction. The first example of a system utilising a reversible covalent reaction was demonstrated by von Kiedrowski<sup>125</sup> (Scheme 1.16). The system was designed as an attempt to control self-replication utilising the smallest possible organic fragments. The system was also incidentally the first to be reported which used a non-nucleotide based recognition motif. The recognition motif in this case was the amidinium–carboxylate salt bridge which was sufficiently strong

in DMSO ( $K_a = 350 \text{ M}^{-1}$  at  $35 \text{ }^\circ\text{C}$ ). The reaction employed was the simple condensation of an amine **19** with an aldehyde **20** to yield an imine **21** and the reaction progress was monitored by  $^1\text{H}$  NMR spectroscopy.

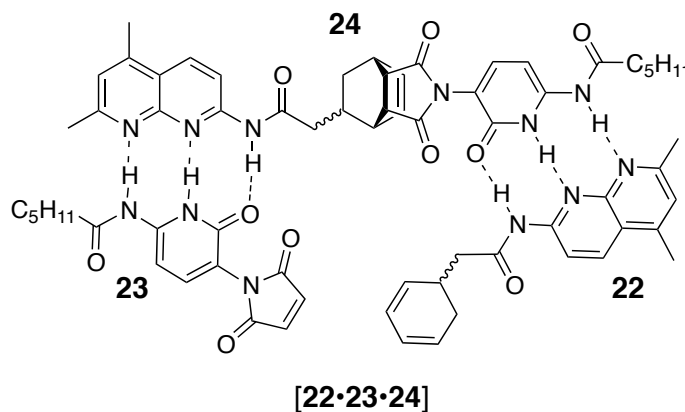


**Scheme 1.16:** Ternary complex [19·20·21] of von Kiedrowski's imine replicator and subsequent template duplex [21·21].

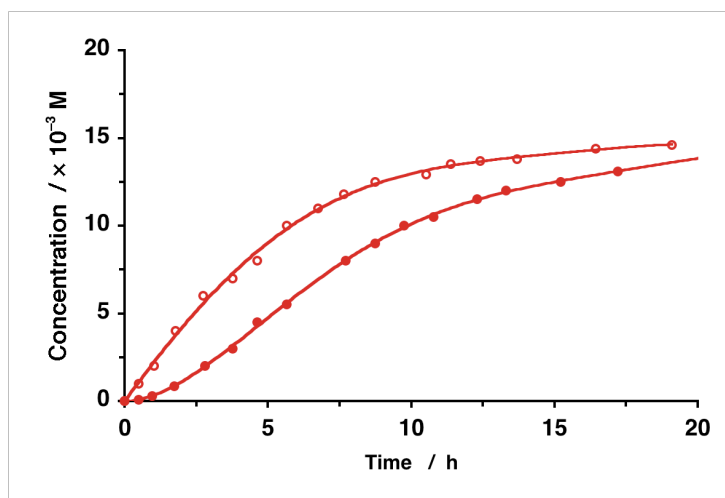
Evidence for self-replication once again was provided in the form of the template effect. The higher the initial concentration of preformed imine added to a fresh reaction, the faster the initial rate of its formation was. The system demonstrated that relatively simple organic molecules were able to be used for self-replication.

Another example of a carefully crafted non-natural self-replicating system was produced by Wang and Sutherland<sup>126</sup> (Figure 1.40). A triple hydrogen bonding motif based on Kelly's bisubstrate template,<sup>85,86</sup> Section 1.4.3, of an aminopyridone and an amidonaphthyridine provided the molecular recognition in the system. The reaction employed was a [4+2] Diels-Alder cycloaddition of a 1,3-cyclohexadiene **22** with a maleimide **23** to form cycloadduct **24** and the reaction progress was monitored by  $^1\text{H}$  NMR spectroscopy. The concentration vs. time profile for the evolution of the product **24** (Figure 1.41 filled circles) for the reaction clearly demonstrated a sigmoidal profile characteristic of autocatalytic self-replication. Autocatalysis was demonstrated with the template effect as addition of preformed template to a fresh set of reactants resulted in the removal of the lag period allowing the reaction to proceed at its

maximum rate upon initiation (Figure 1.41 open circles). With a racemic substrate present there were questions left unanswered by the authors as to the stereochemistry in the process. No comment was made whether one major isomer was produced or whether cross-catalysis between diastereoisomers was present. Later, von Kiedrowski produced<sup>127</sup> a variant of the Wang and Sutherland system to answer the questions posed which will be discussed further on.



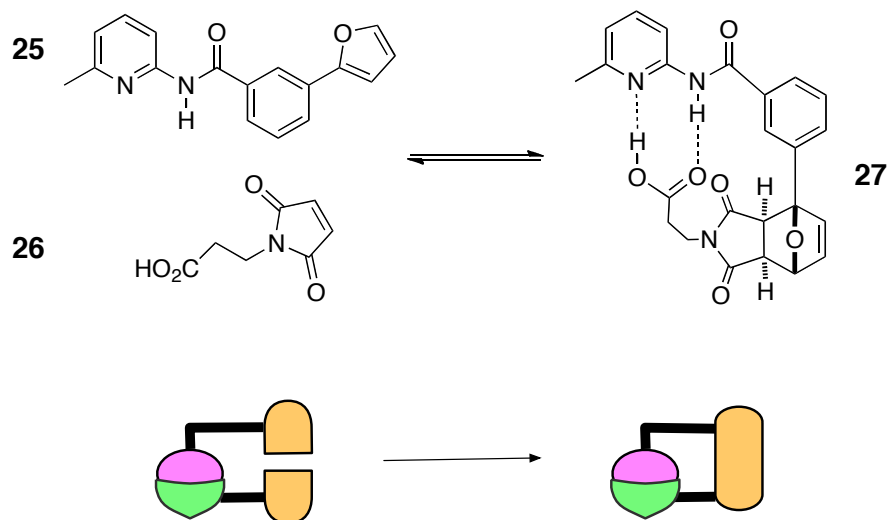
**Figure 1.40:** Ternary complex of Wang and Sutherland replicator.<sup>126</sup>



**Figure 1.41:** Concentration vs. time profile of the Wang and Sutherland replicator. The native recognition mediated reaction is shown as filled circles and red line. The sigmoidal growth profile is clearly present. When preformed template is added, the result (open circles and red line) is an absence of lag period and the reaction proceeds at the maximum rate upon initiation. Data taken from reference 126.

Research into self-replicating systems in our laboratory takes its origins from initial work into recognition mediated acceleration of cycloaddition reactions.<sup>128-133</sup> A successful example of this work is a recognition mediated acceleration of Diels–Alder [4+2] cycloaddition reaction of the dienophile maleimide **26** with diene furan **25**. Normally, the *endo* product is thermodynamically favoured on account of secondary

orbital interactions, in this case however, to selectively was reversed and the disfavoured *exo* product **27**, formed through a *pseudo* intramolecular reaction and stabilised by intramolecular H-bonds, was the major product.<sup>128</sup> (Scheme 1.17).

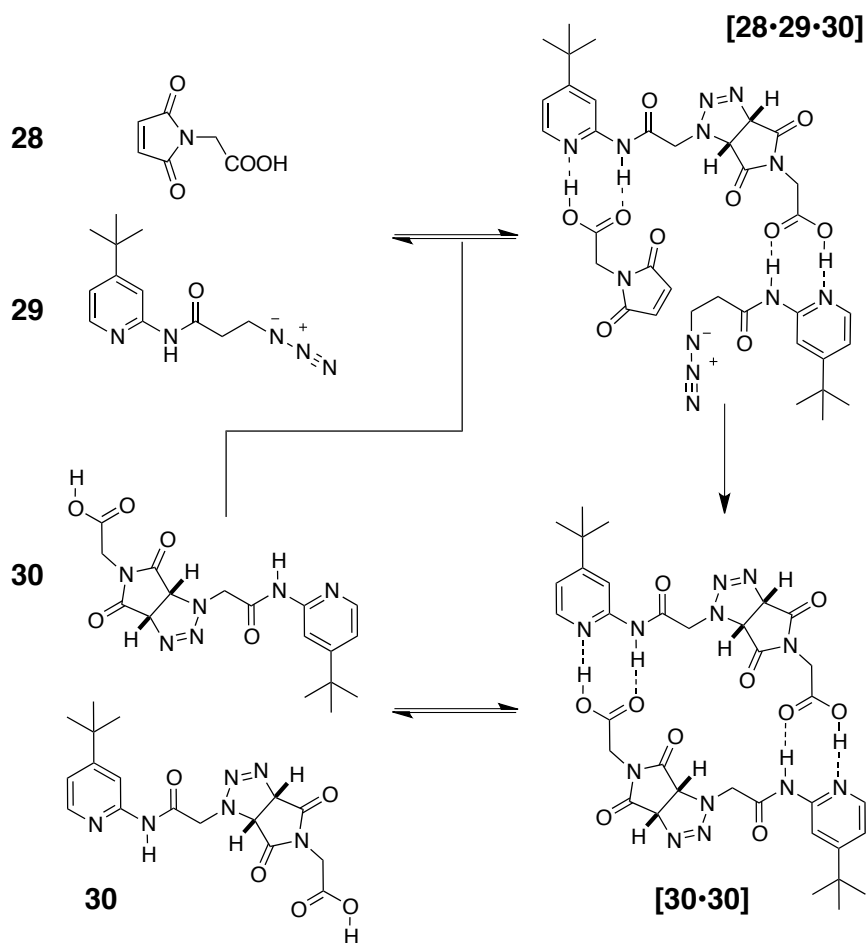


**Scheme 1.17:** Association of reagents into a reactive  $[A \cdot B]$  complex promotes the formation of the thermodynamically unfavoured *exo* product with the hydrogen bonding living on in the product stabilising it from retro reaction.

The recognition motif in the system was the association of a carboxylic acid with an amidopyridine as described by Hamilton<sup>44</sup> and demonstrated previously by von Kiedrowski. This recognition served not only to assemble the substrates into an  $[A \cdot B]$  complex which selectively increased the rate of cycloaddition to form the disfavoured *exo* **27** cycloadduct, but also to stabilise **27** from retro Diels–Alder reaction.

This stabilisation effect was demonstrated after taking **27**, which was stable in  $CDCl_3$  where the hydrogen bonding is permitted, and preparing a solution in DMSO, which disrupted the hydrogen bonding stabilisation, and the product underwent rapid retro Diels–Alder. A similar system was the subject of a computational analysis<sup>134</sup> which confirmed the enthalpic gain from the intramolecular hydrogen bonds was sufficient to stabilise the product from reverse reaction.

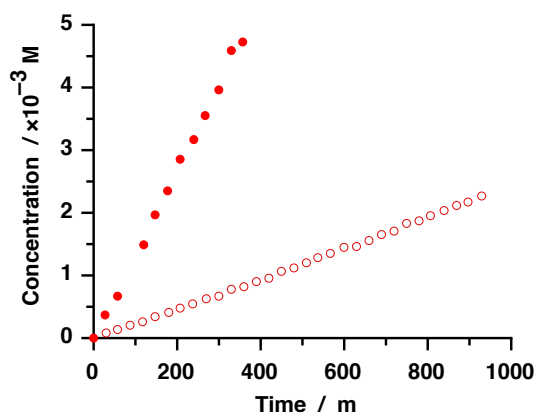
Using a similar recognition motif, one of the early self-replicating systems described<sup>135</sup> by our lab involved the autocatalytic 1,3-dipolar cycloaddition of an amidopyridine bearing azide **28** with an acid maleimide **29** (Scheme 1.18).



**Scheme 1.18:** Autocatalytic cycle of the self-replicator based on a 1,3 dipolar cycloaddition of an azide with a maleimide.<sup>135</sup>

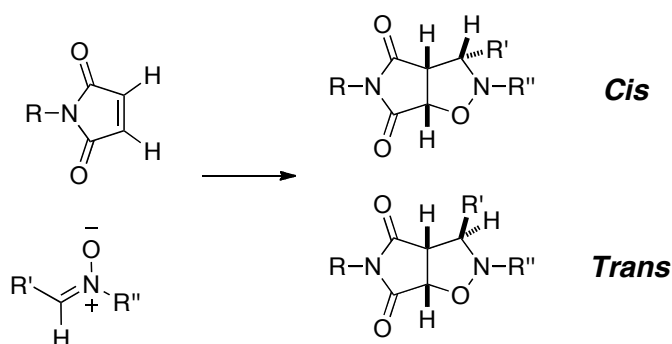
The rate of cycloaddition to form **30** in the recognition mediated reaction, monitored by <sup>1</sup>H NMR spectroscopy, was demonstrated to increase in comparison to a recognition disabled control (using a maleimide methyl ester which blocks association). However the template produced **30** by the recognition mediated reaction was found to undergo inefficient self-replication (Figure 1.42 open circles). A kinetic analysis using von Kiedrowski's method<sup>94</sup> for analysing parabolic replicators determined the reaction order  $p = 0.4$ .

The inefficient replication was therefore a consequence of a very strong product duplex formed which inhibited autocatalytic turnover. Autocatalysis was still able to be demonstrated using the template effect, observing an increase in initial reaction rate with addition of 50 mol% preformed **30** to a fresh batch of reactants **28** and **29**, Figure 1.42 filled circles.



**Figure 1.42:** Concentration vs. time profile for the formation of **30** in the native recognition mediated reaction (open circles) and in the presence of 50 mol% preformed **30**. Data taken from reference 135.

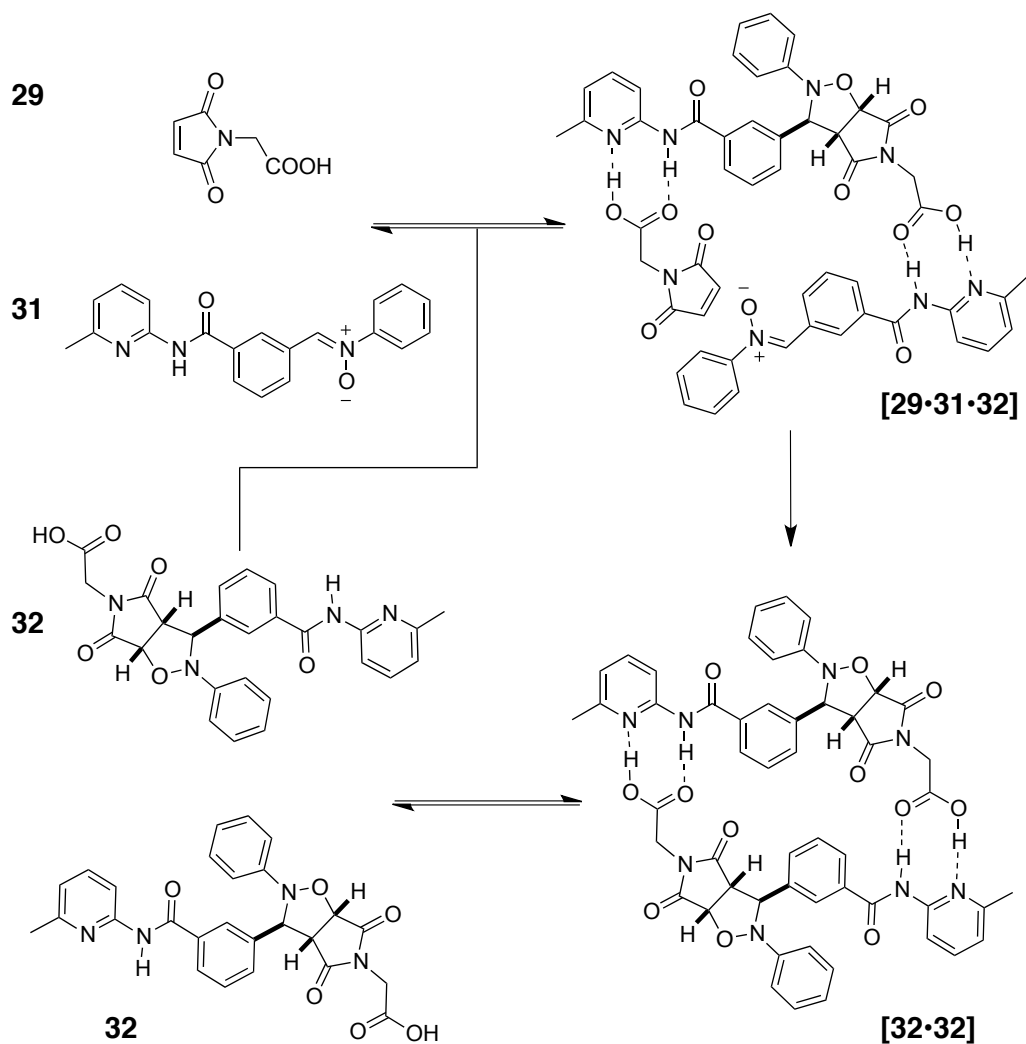
A more efficient system was described<sup>136</sup> after changing the reaction event to a 1,3-dipolar cycloaddition of a nitron with a maleimide. This change has the effect of altering the geometry of the 1,3-dipole in the ternary complex and ultimately more significant rate increases were observed. The change in reaction had the added interest of allowing two potential diastereoisomeric products (Scheme 1.19). The two products were named *trans* and *cis* which arose from the relationship between the protons on the fused ring site and the proton adjacent to the R' group and were easily distinguishable by <sup>1</sup>H NMR spectroscopy, Chapter 7 Section 7.5. The cycloaddition is relatively immune to electronic effects and Brønsted acid catalysis. In the absence of recognition effects, the cycloaddition proceeds to give roughly a 3:1 ratio of *trans* : *cis* isomers in almost all cases.



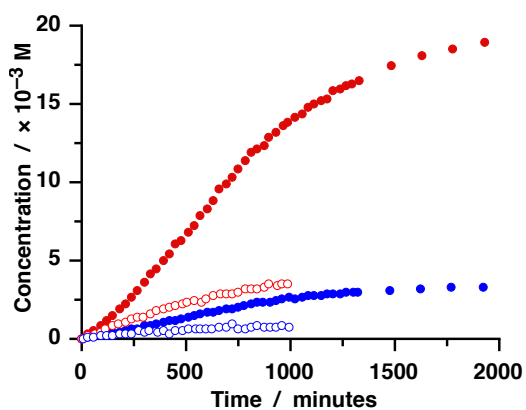
**Scheme 1.19:** The reaction of a nitron with a maleimide produces two diastereoisomeric products which are arbitrarily labelled *trans* and *cis* on account of the relationship between the protons on the fused ring site from the maleimide and the proton originating from the nitron.

The expectation for this new system, backed up by molecular modelling, was that stereochemical information would be passed from parent to daughter template during the replication process (Scheme 1.20).



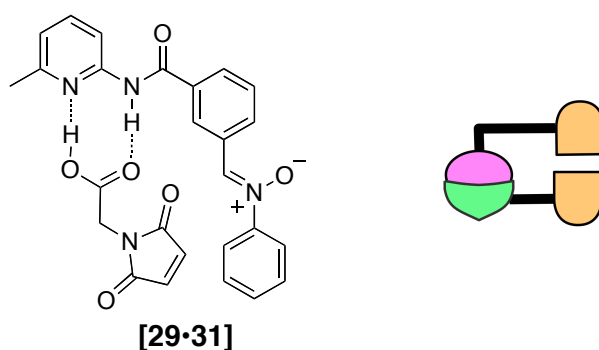


**Scheme 1.20:** Autocatalytic cycle of the self-replicator based on a 1,3 dipolar cycloaddition of nitrone **31** with maleimide **29** to form *trans*-**32**.<sup>136</sup>



**Figure 1.43:** Concentration vs. time profile for the recognition enabled reaction of nitrone **31** with maleimide **29** to form the self-replicating template *trans*-**32** (filled red circles) and diastereoisomeric partner *cis*-**32** (filled blue circles). For comparison, the recognition disabled control reaction to form the *trans* (open red circles) and *cis* (open blue circles) in the absence of recognition processes is shown. Data taken from reference 136.

To investigate this expectation, the reaction of maleimide **29** with nitron **31** was undertaken at 25 mM in CDCl<sub>3</sub> at 10 °C and monitored by <sup>1</sup>H NMR spectroscopy. The results of the reaction, presented as a concentration vs. time profile (Figure 1.43, filled circles) realised our expectations as *trans*-**32** was found to be the major product of the reaction with a *trans* : *cis* ratio of 6:1, and displaying the characteristic sigmoidal growth profile. In comparison to a recognition disabled control reaction (Figure 1.43, open circles) the recognition enabled reaction produced a curious increase in the production of the *cis*-**32** as well. A further template directed reaction was carried out to demonstrate autocatalysis in which preformed *trans*-**32** was added to a fresh set of reagents **29** and **31**. The results of this reaction revealed the absence of the lag period in the production of *trans*-**32** however no template effect was observed in the production of *cis*-**32**. Therefore the slight increase in production of *cis*-**32** in comparison to the bimolecular control reaction was attributed to the formation of an [A•B] complex [29•31] (Figure 1.44) which promoted cycloaddition to form *cis*-**32**.

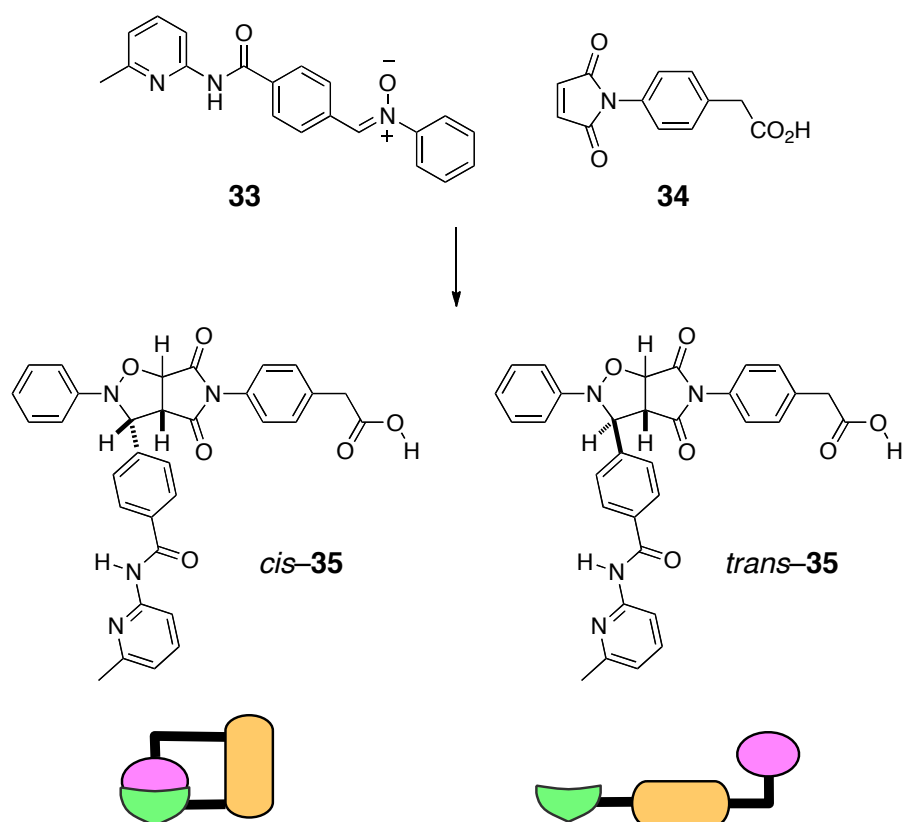


**Figure 1.44:** Production of a reactive [A•B] complex leads to an enhancement in the *cis* cycloadduct in comparison the recognition disabled pathway. This channel served to lower the efficiency of the autocatalytic cycle.

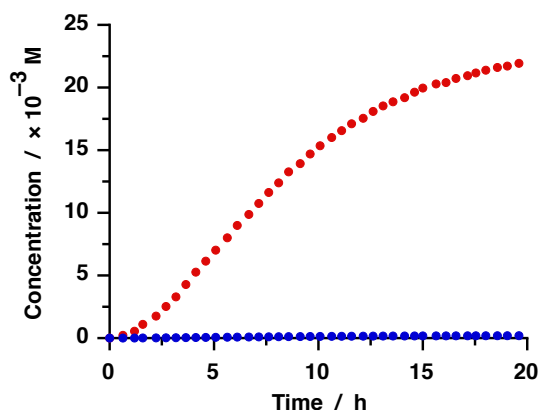
In light of the presence of an [A•B] complex channel increasing the undesired product *cis*-**32** and lowering the overall efficiency of the replicator, studies were undertaken to gain an appreciation of the requirements for efficient replicator design.<sup>131,137,138</sup> The conclusion from these studies was that given enough conformational freedom, the most probable recognition mediated pathway available to a system will be the [A•B] complex channel. In order to produce an efficient self-replicating system, the design must attempt to eliminate the [A•B] complex channel from being active. This design could be achieved by changing spacer lengths (as noted

previously by Rebek's system) or reducing conformational freedom by increasing the rigidity of components used.

Using the lessons learned from this study, a more efficient self-replicating system was devised by Kassianidis and Philp<sup>139</sup> based on the reaction of nitron **33** with maleimide **34**. The 1,3-dipolar cycloaddition reaction was chosen once again as well as the established amidopyridine-carboxylic acid recognition motif. In this case, structural differences such as a phenyl spacer on the maleimide along with a change in substitution position of the nitron reactive site to the *para* position. Once again, two possible diastereoisomeric products were able to be produced (Scheme 1.21) *trans*-**35** and *cis*-**35**. In the absence of recognition, achieved by using the methyl ester of the maleimide carboxylic acid thus preventing association of the recognition motif, the reaction was slow reaching 9% conversion at  $-10\text{ }^{\circ}\text{C}$  in  $\text{CDCl}_3$  after 16 h with a *trans* : *cis* ratio of 3:1.



**Scheme 1.21:** Reagents and products of the highly efficient self-replicating system reported by Kassianidis and Philp. The *trans* and *cis* cycloadducts once again are the two possible products with their open or closed structure depicted by cartoon below.

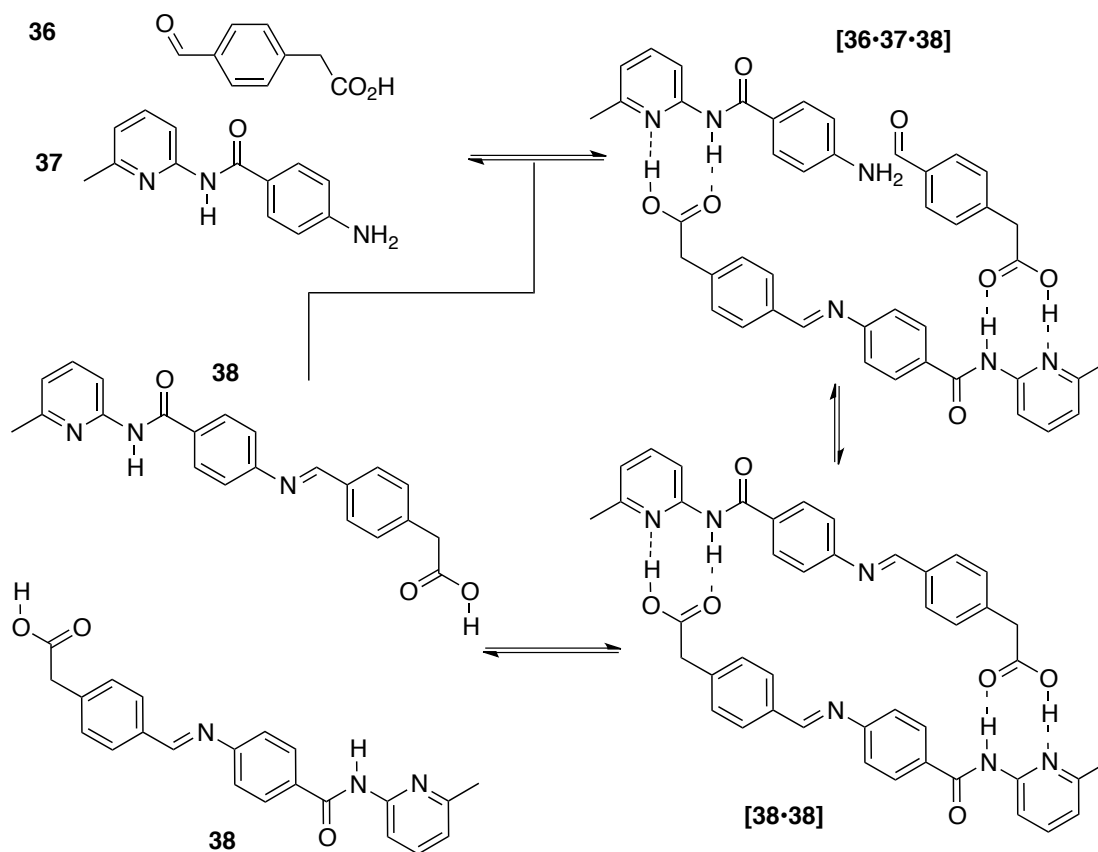


**Figure 1.45:** Concentration vs. time profile for the reaction of nitrone **33** with maleimide **34** demonstrating the sigmoidal growth of the *trans*-**35** (red circles) and of *cis*-**35** (blue circles) during the reaction when performed at  $-10$  °C in  $\text{CDCl}_3$ . Data taken from reference 139.

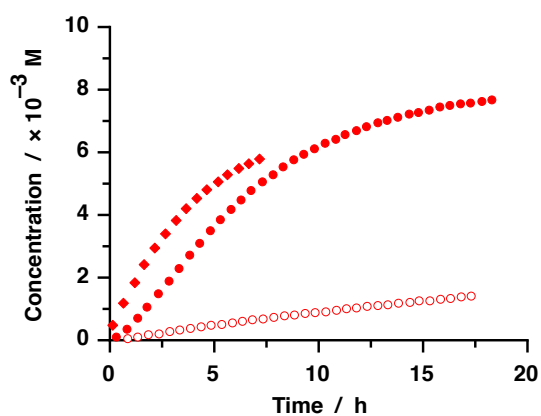
Under the same conditions, the recognition mediated reaction between nitrone **33** with maleimide **34** demonstrated a remarkable acceleration in the production of *trans*-**35** reaching 85% conversion after 16 h with a *trans* : *cis* ratio of 115:1 observed. The concentration vs. time profile for the production of *trans*-**35** (Figure 1.45) demonstrated a sigmoidal profile characteristic of an autocatalytic reaction. The template directed nature of the reaction was demonstrated by addition of preformed *trans*-**35** to a fresh batch of reactants **33** and **34** which removed the lag period and allowed the reaction to proceed at the maximum autocatalytic rate upon initiation. This system from Kassianidis and Philp was the basis for an extended replicating system which will be discussed later and for some work in this thesis.

More recently, our laboratory successfully deployed<sup>140</sup> a replicating strategy to accelerate the reversible condensation reaction between aldehyde **36** and amine **37** to form imine **38** and subsequently stabilise the dynamic product by the formation of a stable template duplex (Scheme 1.22).

In the absence of recognition effects (Figure 1.46, open circles) the imine condensation reaction was demonstrated to be slow reaching 9% conversion after 16 h at a reaction concentration of 15 mM in  $\text{CDCl}_3$  at 25 °C. In the recognition mediation system on the other hand (Figure 1.46, filled circles) the overall conversion was enhanced reaching over 50% conversion after the same time under the same conditions. The concentration vs. time profile for the production of imine **38** displayed a sigmoidal profile and autocatalysis was demonstrated by addition of preformed imine **38** to a fresh batch of reactants, Figure 1.46 diamonds. The concentration vs.



**Scheme 1.22:** Autocatalytic cycle for the production of imine **38** by self-replication.



**Figure 1.46:** Concentration vs. time profile for; (open circles) the recognition disabled control reaction, (filled circles) the recognition enabled autocatalytic reaction to form imine **38** and (filled diamonds) preformed template doped reaction. Information is not shown for the entire time period for the template doped reaction on account of precipitation from the reaction mixture. Data taken from reference 140.

time profile in this case revealed the absence of a sigmoidal shape with the reaction proceeding at the maximum rate upon initiation demonstrating that imine **38** was indeed a catalyst for its own production.

## 1.6 Systems Chemistry

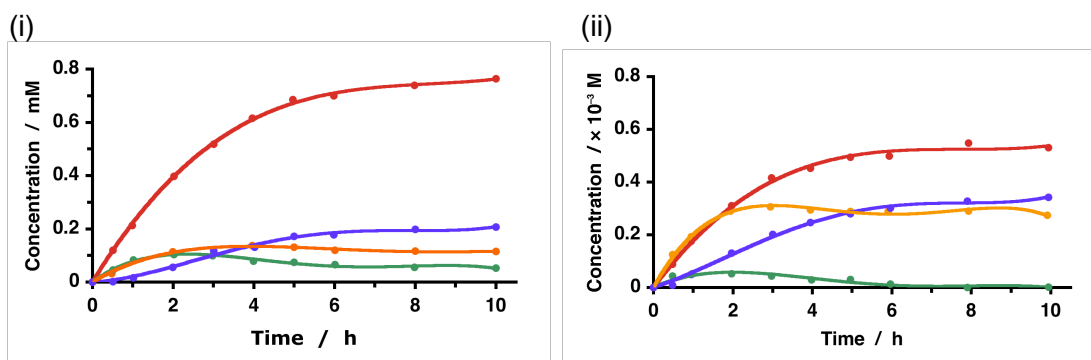
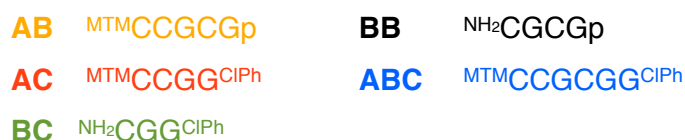
Orgel claimed<sup>87</sup> that for a true example of information transfer in a self-replicator, more complicated systems are required. In such a system, three or more starting materials would be able to react to produce a number of possible self-replicating template products. The ability to transfer information efficiently could then be measured by the ability of an added template (which instructs the system to produce itself) to increase the selectivity of itself from mixture of products. In theory the system could be instructed to produce a number of results based on the instructions given to it. Investigation of self-replicating systems have gained prominence as a consequence of their possible link to prebiotic behaviour. We are not for a second suggesting that highly purified and somewhat water sensitive reagents demonstrated in some of the replicating systems above were present on prebiotic earth, but the behaviour which can be expressed by them is what is important. In isolation, each replicator is capable of a single behaviour which is to amplify itself. However life today does not exist using single replicating molecules. It is through the orchestrated cooperation of complex networks which proceed with exquisite selectivity that allows life to exist. These complex networks are built on interactions and reactions between molecules which allow biological systems to adapt and respond rapidly to external stimuli creating a complex range of signalling and metabolic pathways with checkpoint controls connecting members of the network. The emerging field of systems chemistry attempts to capture the complexity of such biological networks in a wholly synthetic framework. Using a chemical approach, a group of synthetic molecules are designed to interact (non-covalently) and react (covalently) with multiple partners within a network in programmed ways. The outcome is the creation of a system which may have system-level properties which are not simply the the sum of the properties of individual components, so called emergent properties.

Combining multiple self-replicators into a single system may therefore be an appropriate application of self-replication to the field of systems chemistry. When multiple replicators are combined, structural similarities could allow non-covalent interactions between replicators, leading to an enhancement or inhibition in their production. When replicators are forced to compete for common material, Darwinian selection or 'survival of the fittest' behaviour might be observed. In either way, system-level behaviour leading to self-organisation may be observed. Ultimately this

could lead to an understanding behaviour of potential replicators on prebiotic earth which lead to the existence of life.

### 1.6.1 Multiple Replicator Systems

von Kiedrowski devised an extended replicating system from three starting materials<sup>141</sup> based on work from their successful oligonucleotide based replicators. The three oligonucleotide components are  $\text{MTMCCGp}$  (**A**),  $\text{NH}_2\text{CGp}$  (**B**) and  $\text{NH}_2\text{Gp}^{\text{CIPh}}$  (**C**) where MTM and CIPh represent the thiomethyl and *o*-chlorophenyl terminating groups respectively and  $\text{NH}_2$  and p represent the amine and phosphate reactive sites respectively. The three components were mixed together in equal concentrations and the reaction monitored for 10 h. From the three reagents, five possible condensation products are able to be formed: **AB**, **AC**, **BC**, **BB** and **ABC**. Their identities are described below.

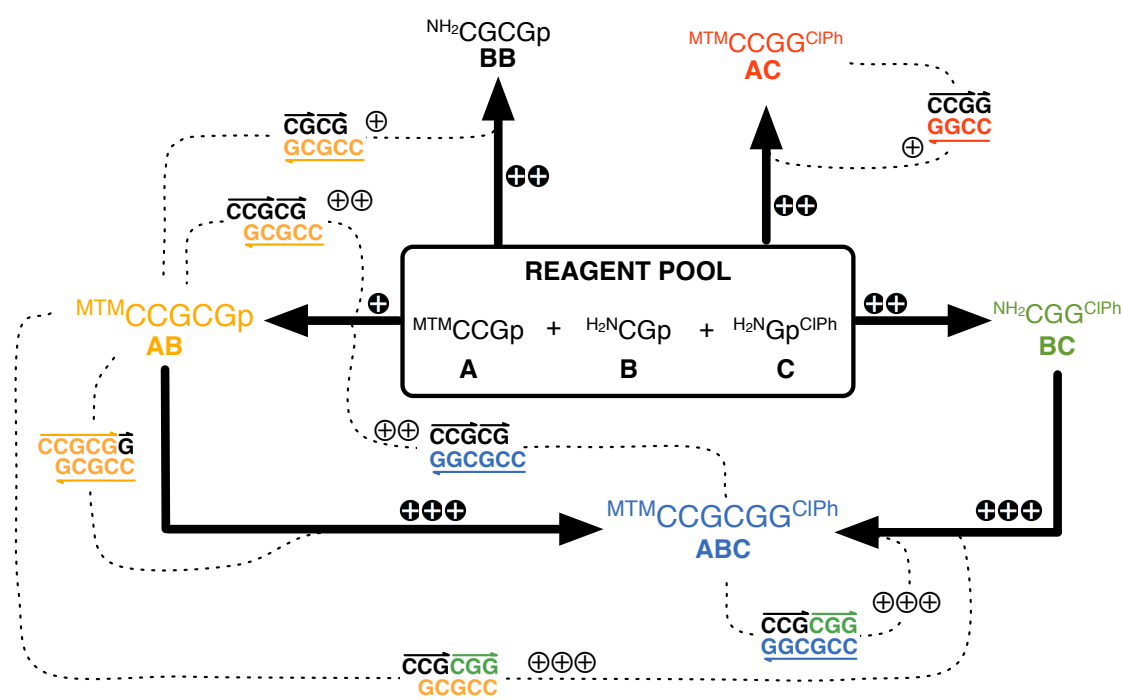


**Figure 1.47:** Concentration vs. time profiles for the evolution of products during the course of (a) the native reaction and (b) the **ABC** template doped reaction. **BC** in green, **AB** in orange, **ABC** in blue and **AC** in red. Data taken from reference 141.

The results of monitoring the three component reaction are displayed in a concentration vs. time profile in Figure 1.47 (i). From the results, it is clear that the major product from the system was **AC**, a self-complementary four oligonucleotide strand which is produced from a single condensation reaction of **A** with **C**. Once formed, the self complementary **AC** may then reproduce itself autocatalytically from **A** and **C**. Another self-complementary template, **ABC**, is produced in two manners, the reaction of **AB** with **C** or from the reaction of **A** with **BC**. As neither **AB** or **BC** is present upon initiation of the reaction, these must first be formed which means two

condensation reactions are required to form **ABC** with the concentrations of **AB** and **BC** passing through a concentration maxima. **BB** was not able to be separated from **B** and thus its concentration was not reported.

To demonstrate the connectivity of this simple network, preformed **ABC** was added to a fresh batch of reagents for the system (Figure 1.47 (ii)). The result of adding the template was that the initial rate in the formation of **AB** increases as a consequence of their partial complementary nature allowing **ABC** to template the formation of **AB**. As there is now significantly more **AB** in the system, the subsequent reaction through to **ABC** is now more prominent and a 65% enhancement† in the concentration of **ABC** is seen after 10 h with concurrent 31% inhibition of **AC**. From these results, the authors were able to describe the overall network topology in a reaction scheme depicted in Figure 1.48.



**Figure 1.48:** Reaction network topology for the condensation reactions of **A**, **B** and **C** to form products **AB**, **AC**, **BB**, **BC** and **ABC**. Starting from the reagent pool, **BOLD ARROWS** represent bimolecular reaction pathways from the reagent to the product. When a product is able to act as a template for a reaction, a dashed line is employed originating *from* the template *to* the bold arrow of the reaction being templated. The ternary complex for each templated reaction is shown along the dashed line with the reagents on top and the template on below. The greater the number of ⊕ or ⊕ symbols (representing the relative efficiency of the template independent and template dependent pathways respectively) the greater the relative reaction rate. Figure adapted from reference 141.

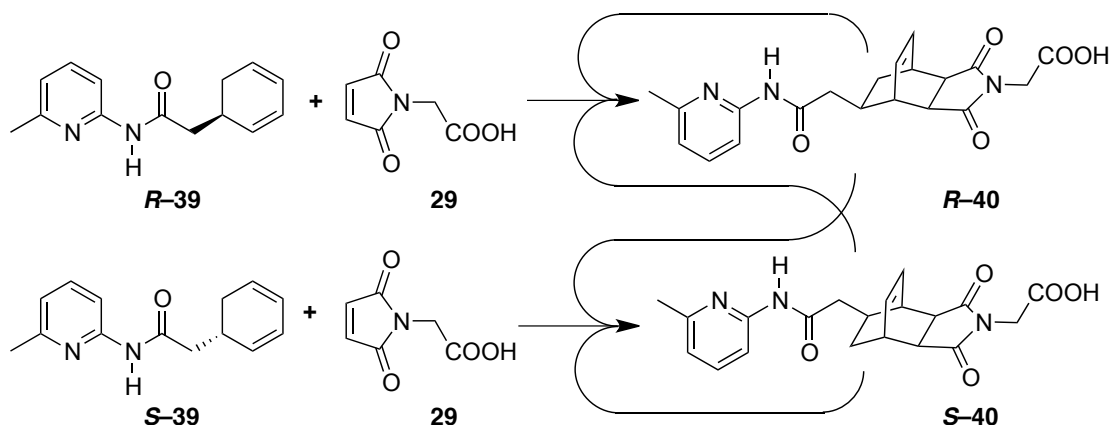
† The calculation of enhancement is explained later in Section 2.1.1



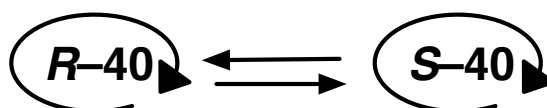
Each condensation reaction pathway is represented as a bold arrow from the reagent to the product. A dashed line is employed to demonstrate when a product is able to act as a template for a reaction, originating *from* the template employed going *to* the bold arrow of the reaction being templated. The ternary complex of the template directed reactions are shown along the line with the reagents on top and template on the bottom. The number of ⊕ and ⊕ symbols on a reaction or template directed reaction line represent the relative efficiency of the template independent and template dependent pathways respectively. (⊕ → ⊕⊕ → ⊕⊕⊕ = increasing efficiency). Starting from the reagent pool, two molecules of **B** may react to form **BB**, **A** and **C** may form **AC** which may then act as a template for its own formation (autocatalysis), **B** and **C** may react to form intermediate **BC** and **A** and **B** may react to form intermediate **AB**. **AB** may then react once more to form **ABC**. In addition to this reaction, **AB** is rich in cross catalytic behaviour. **AB** may act as a template for the formation of **BB** from **B** and **B**, **ABC** from **A** and **BC**, **ABC** from **AB** and **C** and itself from **A** and **B**. Finally **BC** will react with **A** to form **ABC** and this reaction may be catalysed by **AB** as described above or **ABC**.

The results were the first to demonstrate rudimentary information transfer and that a system relatively rich in crosscatalytic pathways can be produced from a system of only a few starting materials and is an excellent example of a synthetic chemical network.

In an attempt to answer the questions of stereochemistry posed by Wang and Sutherland's Diels–Alder replicator (Figure 1.40) von Kiedrowski designed a structural variant of the system to conduct a series of kinetic, computational and theoretical mechanistic studies<sup>127</sup> (Scheme 1.23). The authors concluded that only a single racemic diastereoisomer was produced in the reaction above the detection limits of <sup>1</sup>H NMR spectroscopy and its identity was confirmed by X-ray crystallography. The *R* and *S* enantiomers of the diastereoisomer diene, **R-39** and **S-39**, were synthesised and their reactions with maleimide **29** investigated in isolation, and then in the presence of either the *R* or *S* templates, **R-40** or **S-42**. The result was that both enantiomers of template successfully increased the initial rates of reaction when either of the dienes were used thus demonstrating crosscatalysis between them (Figure 1.49).



**Scheme 1.23:** von Kiedrowski's results of the variant of the Wang and Sutherland system. Each diastereoisomeric template is autocatalytic in its own production and an active crosscatalyst to produce the isomeric partner.

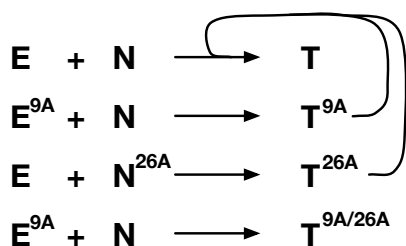


**Figure 1.49:** Representation of the results observed by von Kiedrowski as a graph architecture. The two products of the system, **R-40** and **S-40** are represented as nodes with curved arrows to denote autocatalysis. The directed edges (straight black arrows) represent the crosscatalysis.

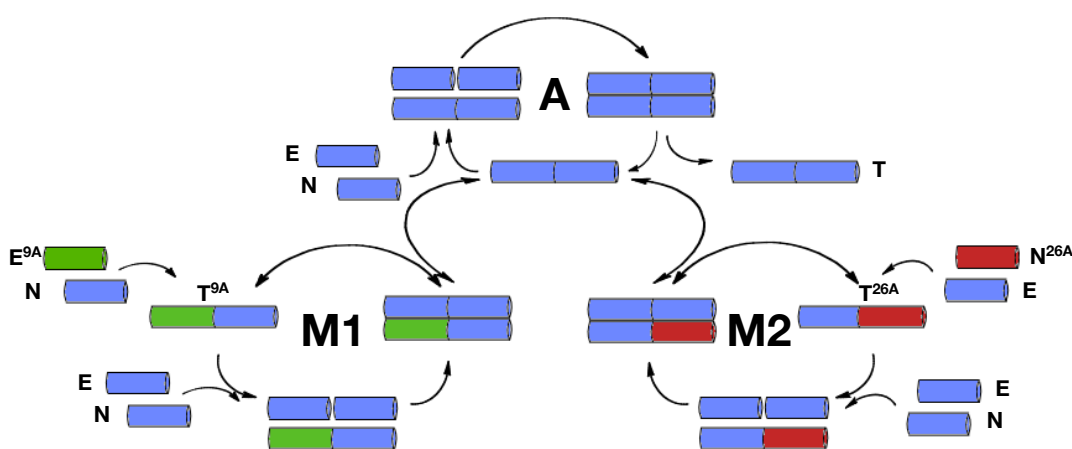
von Kiedrowski later used a variant of the same system to demonstrate<sup>142</sup> that thermodynamic information such as the template duplex association strength could be extracted *in situ* from the change in shift of integrals in <sup>1</sup>H NMR spectroscopy instead of independent NMR spectroscopy titrations.<sup>143,144</sup>

In a series of papers, Ghadiri applied the successful peptide replicator described above into ever more complicated systems which were able to demonstrate behaviour similar to many processes essential to life and later akin to molecular logic.

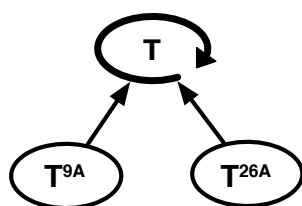
The ability of replicating systems to conserve their information against the accumulation of errors is crucial in maintaining life.<sup>145</sup> DNA today has error checking mechanisms which ensure information is efficiently transferred from parent to daughter strand. Error checking mechanisms now are sophisticated however at some point in the evolution of life, simpler mechanisms must have existed. Ghadiri demonstrated a peptide replicator system which the original sequence is conserved through a simple error checking mechanism.<sup>146</sup> The original replicator fragments **E** and **N** were subject to a conservative mutation of a single residue to form two new fragments **E<sup>9A</sup>** and **N<sup>26A</sup>**. All combinations of the four fragments gave rise four possible templates; **T**, **T<sup>9A</sup>**, **T<sup>26A</sup>**, **T<sup>9A/26A</sup>** as shown in Scheme 1.24.



**Scheme 1.24:** The reaction of each electrophilic fragment **E** or **E<sup>9A</sup>** with each nucleophilic fragment **N** or **N<sup>26A</sup>** was performed in isolation to produce their respective templates **T**, **T<sup>9A</sup>**, **T<sup>26A</sup>** or **T<sup>9A/26A</sup>**. The original template **T** is produced autocatalytically represented by the thin line. Each of the single mutant templates are not produced autocatalytically however are cross catalysts for the formation of **T**. The double mutant template is catalytically infertile.



**Figure 1.50:** Cartoon representation of the error correcting network. The original template **T** from **E** and **N** is produced through an autocatalytic cycle **A**. Mutant templates, **T<sup>9A</sup>** and **T<sup>26A</sup>**, which appear in the system are able to direct the formation of original template **T** from **E** and **N** thus correcting the original error.

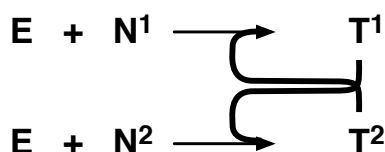


**Figure 1.51:** Representation of the error correcting system as a graph architecture. Each node represents a unique template with the curved arrow around **T** denoting autocatalysis. The directed edges (straight black arrows) represent the crosscatalysis.

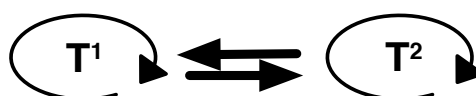
In isolated experiments, the authors found that each single mutated template, **T<sup>9A</sup>** and **T<sup>26A</sup>**, could not act as a template for their own formation but were crosscatalysts for the production of the unmutated template **T**. The double mutant template and **T<sup>9A/26A</sup>** was both autocatalytically and crosscatalytically infertile. When all four components were mixed in one system, the major product was the unmutated

template **T** as a result of the combination of its own autocatalytic cycle **A** and crosscatalytic cycles from mutated templates **M1** and **M2** (Figures 1.50 and 1.51).

Ghadiri *et al* also went on to demonstrate a symbiotic relationship between two peptide replicators.<sup>147,148</sup> Symbiosis is a mutually beneficial relationship between two species in which one enhances the formation the other which in turn enhances the production of the former. Using the established peptide replicator, two replicating templates were produced from the combination of a common electrophilic fragment **E** with either of two nucleophilic fragments **N<sup>1</sup>** or **N<sup>2</sup>** to produce **T<sup>1</sup>** or **T<sup>2</sup>** respectively (Scheme 1.25). Comparison of the two replicators revealed that **T<sup>2</sup>** was more efficient in isolation than **T<sup>1</sup>**. Therefore in a system combining the equimolar quantities of the three reagents, **E**, **N<sup>1</sup>** and **N<sup>2</sup>**, a ‘survival of the fittest’<sup>149</sup> type of behaviour might be expected in which **T<sup>2</sup>** production would overwhelm **T<sup>1</sup>**. This expectation was not realised however as in the three component system, both templates co–existed. It was subsequently found that each template was a better crosscatalyst than an autocatalyst (Figure 1.52).



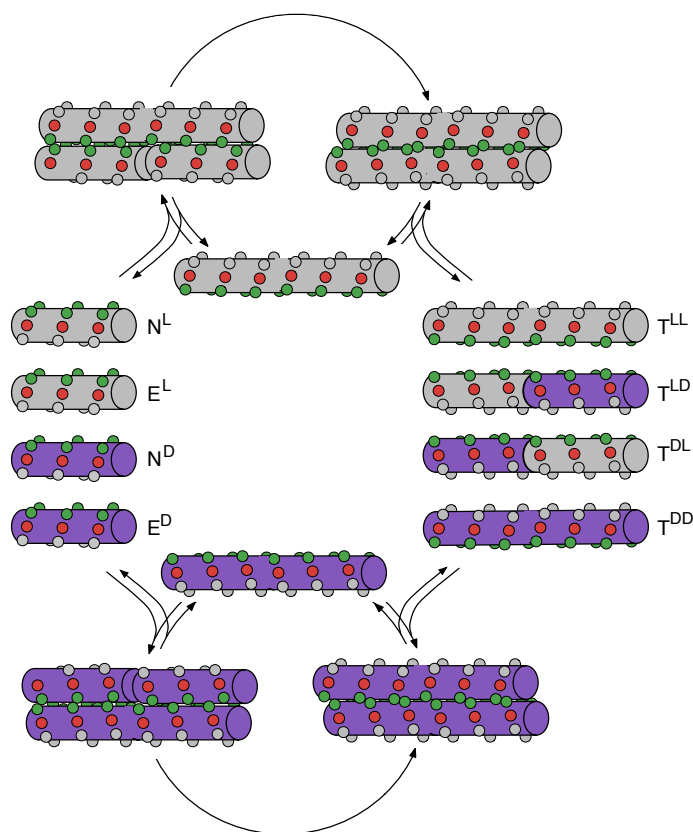
**Scheme 1.25:** The reaction of each nucleophilic fragment **N<sup>1</sup>** or **N<sup>2</sup>** with the electrophilic fragment **E** produces their respective templates **T<sup>1</sup>** or **T<sup>2</sup>** autocatalytically. The thicker line represents more efficient crosscatalysis.



**Figure 1.52:** Representation of the symbiotic system as a graph architecture. Each node represents a unique template with the thin curved arrow around each node denoting autocatalysis. The directed edges (thick black arrows) represent the more efficient crosscatalysis.

Thus when seeding the three component reaction with preformed **T<sup>1</sup>**, a greater increase in the production of **T<sup>2</sup>** than **T<sup>1</sup>** was observed. Likewise, seeding the reaction with **T<sup>2</sup>** produced a greater increase in the rate of production of **T<sup>1</sup>** than **T<sup>2</sup>**.

In a system which could have implications in the search for a mechanism for chiral symmetry breaking and the emergence of homochirality,<sup>150</sup> Ghadiri used enantiomeric substrates to demonstrate a pair of homochiral replicators<sup>151</sup> (Figure 1.53).

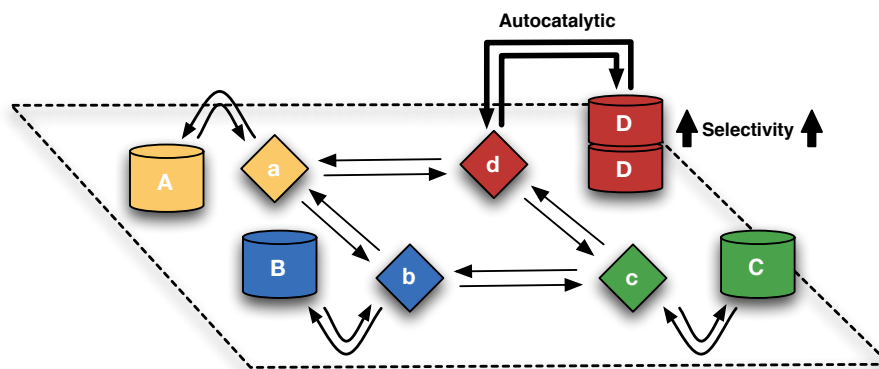


**Figure 1.53:** Chiroselective peptide replicator reported by Ghadiri. Only the reactions to form homochiral templates proceed autocatalytically. Heterochiral templates are only produced by template independent reaction. Figure adapted from reference 151.

The authors used D or L amino acids to produce an enantiomeric pair of peptide fragments for their replicator,  $E^L$ ,  $E^D$ ,  $N^L$  and  $N^D$ . When all four were mixed together in equal concentrations to form a ‘racemic pool’ of reagents, the homochiral templates ( $T^{LL}$  and  $T^{DD}$ ) were the preferentially produced products. In isolated experiments to explain the observed behaviour it became apparent that  $T^{LL}$  and  $T^{DD}$  were autocatalysts for their own formation however the heterochiral template ( $T^{DL}$  and  $T^{LD}$ ) were produced in a template independent manner and were catalytically inert.

### 1.6.2 Combining Replication Strategies With Dynamic Chemistry

A potential method of integrating replication with a DCL<sup>152</sup> is illustrated in Figure 1.54. In this scenario, a dynamic library of four building blocks, **a** through **d** are able to freely exchange with each other and four further products **A** through **D**. One of the products, **D**, is a self-complementary template whose reversible formation from **d** is accelerated by autocatalysis allowing for the formation of a stable product duplex [**D**•**D**]. A critical requirement for the implementation of the protocol outlined in Figure

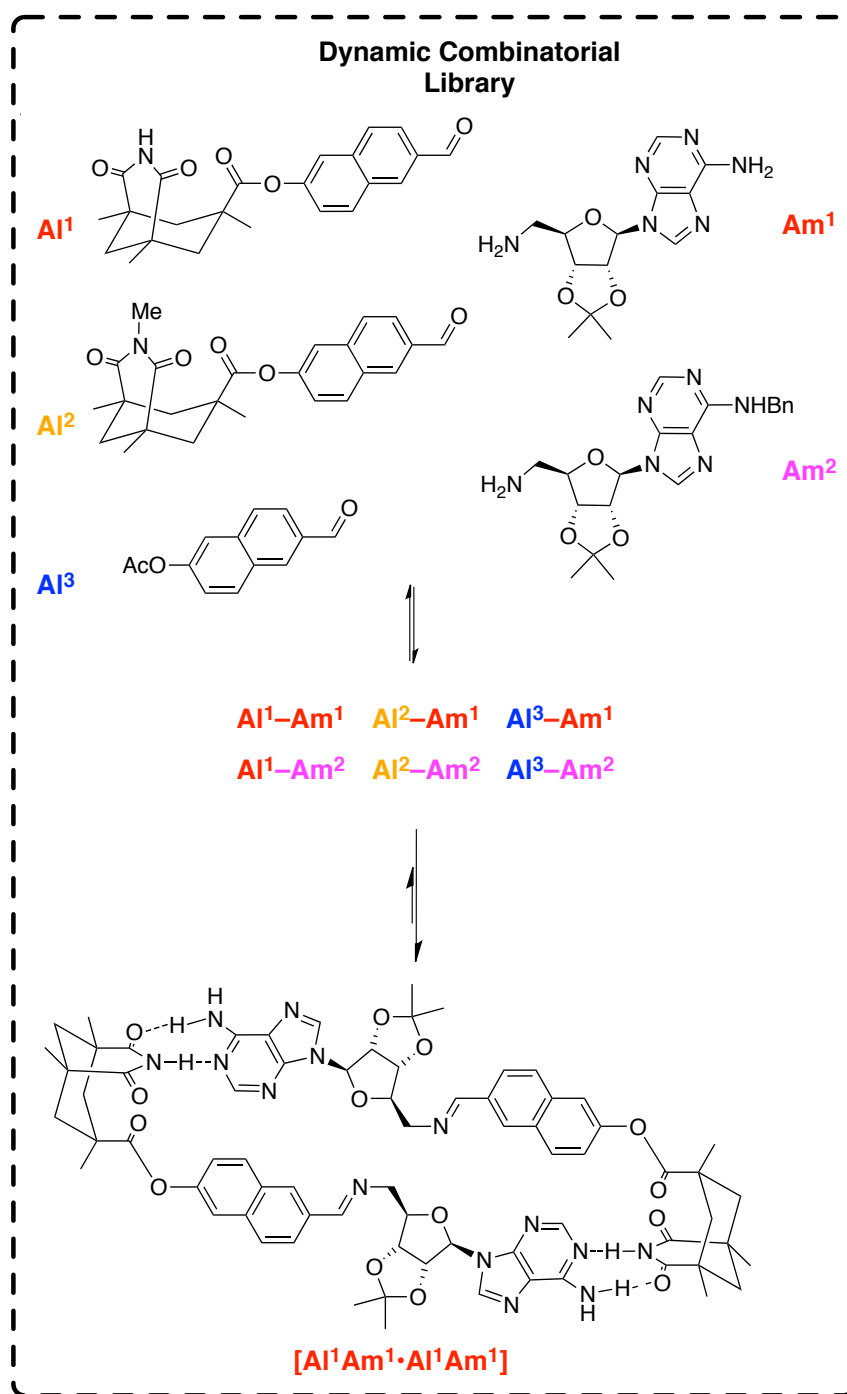


**Figure 1.54:** A model for coupling replication to a dynamic library. Reagents **a**, **b**, **c** and **d** can all exchange with each other. These reagents may be reversibly converted to four products, **A**, **B**, **C** and **D**. Product **D** is able to form a self-complementary duplex and is able to catalyse its own formation therefore should be formed preferentially. The compounds enclosed within the black dashed line are part of a dynamic library. Figure adapted from reference 152.

1.54 is the identification of a suitable reversible covalent bond forming event to furnish the template.

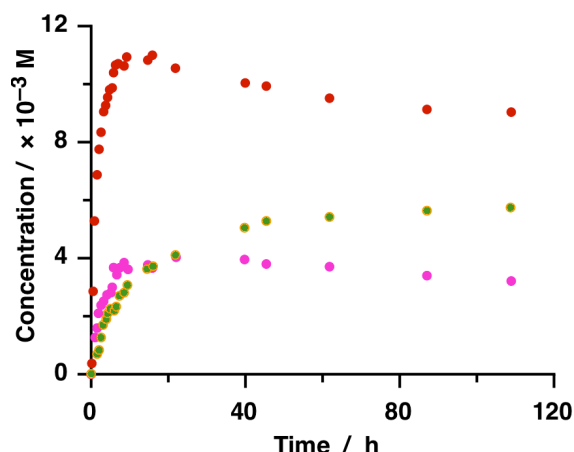
In the first attempt to control a replicating system from such a dynamic pool of reagents, Giuseppone<sup>153</sup> exploited the reversibility of imine bond formation in a system whose design took inspiration from the work of Rebek (Scheme 1.14). Giuseppone's system of five building blocks contained three aldehydes and two amines which upon mixing can produce five imines in a dynamic combinatorial library (Scheme 1.26). Of the five reagents, two are decorated with non-covalent recognition sites, **Al<sup>1</sup>** and **Am<sup>1</sup>**, and upon their condensation reaction form **Al<sup>1</sup>-Am<sup>1</sup>** a self-complementary template which forms the duplex [**Al<sup>1</sup>-Am<sup>1</sup>·Al<sup>1</sup>-Am<sup>1</sup>**] thus stabilising the structure by means of an 'internal' template effect over the other possible products which may arise from the library. The expression of duplex in the system was enhanced 83% in comparison to a recognition disabled control system.

The authors claimed the selection proceeded *via* a self-replicating mechanism by monitoring the evolution of product by <sup>1</sup>H NMR spectroscopy. Examination of the rate of formation of the products in the library was undertaken by monitoring five component reaction. The concentration vs. time profile for the reaction (Figure 1.55) revealed a biphasic reaction characteristic. In the initial stages of the reaction, the kinetic acceleration of the formation of **Al<sup>1</sup>-Am<sup>1</sup>** (red circles) by autocatalysis is the dominant behaviour as an overshoot in the production of **Al<sup>1</sup>-Am<sup>1</sup>** is seen. As the time course of the reaction proceeds, the thermodynamics begin to dominate the



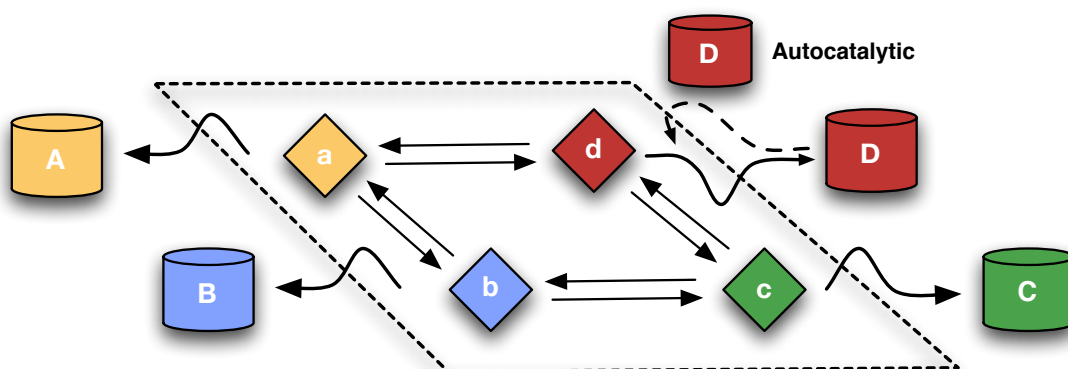
**Scheme 1.26:** Combination of the five reagents A<sup>1</sup>, A<sup>2</sup>, A<sup>3</sup>, Am<sup>1</sup> and Am<sup>2</sup> leads to the formation of an 11 component DCL (5 reagents + 6 combinations). From the DCL, one product is stabilised by hydrogen bonding formed by creating a template duplex. The compounds enclosed within the black dashed line are part of a dynamic library.

overall system behaviour which corrects the overshoot in  $\text{Al}^1\text{-Am}^1$  as well as  $\text{Al}^1\text{-Am}^2$  (purple circles) and their concentration decreases with concurrent increase in condensation products from  $\text{Al}^2$  and  $\text{Al}^3$  with  $\text{Am}^2$  (green and yellow circles).



**Figure 1.55:** Concentration vs. time profile for the five component library with  $Al^1-Am^1$  shown as red circles,  $Al^1-Am^2$  as purple circles and the concentration of imines  $Al^2-Am^2$  and  $Al^3-Am^2$  as yellow and green circles. Data taken from reference 153.

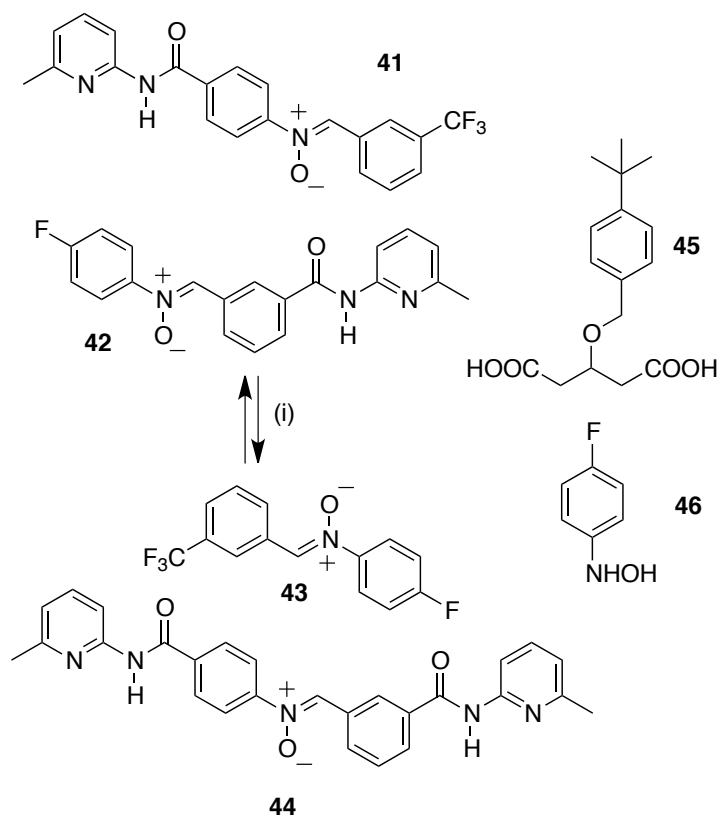
These results served to demonstrate that within a fully dynamic system, consideration of all dynamic processes must be taken into account. In this case, despite the careful design of the recognition mediated processes to form a stable duplex, the covalent bond forming reaction is also an important consideration. It was therefore apparent that for systems based on the model in Figure 1.54 the equilibrium position is purely determined by thermodynamics, that is the relative stability of the product, independent of the pathway to reach that point. The kinetic parameters in the system only govern the rate at which the system will reach equilibrium however will have no effect on that point. If we wish to drive the system away from its predetermined equilibrium position then an irreversible process must be introduced to the system. This realisation leads to the construction of an alternative method described in Figure 1.56.



**Figure 1.56:** A model for coupling irreversible replication to a dynamic library. Reagents **a**, **b**, **c** and **d** can all exchange with each other. These reagents may be irreversibly converted to four products, **A**, **B**, **C** and **D**. Product **D** is able to form a self-complementary duplex and is able to catalyse its own formation therefore should be formed preferentially. The compounds enclosed within the black dashed line are part of a dynamic library. Figure adapted from reference 152.



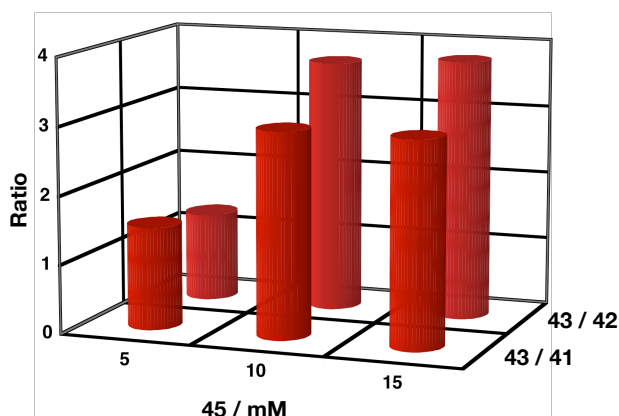
In this new model, reagents **a**, **b**, **c** and **d** are still free to equilibrate between each other in a DCL however in this case the transformation of reagents into products **A**, **B**, **C** and **D** are now irreversible. Once again **D** is a replicator and will therefore consume **d** in an autocatalytic cycle reproducing itself. The autocatalytic formation of **D** will remove **d** from the DCL and as such perturb the equilibrium away from the steady state position of the four building blocks **a**, **b**, **c** and **d**. As such we would expect the system to attempt to reestablish the equilibrium position by generation of more **d** by exchange from **a**, **b** and **c**. The autocatalytic template **D** therefore sets up a feedback loop which drives the DCL to satisfy its own requirements by the generation of more **d**. In order to implement this new model as described by Figure 1.56, it was first necessary to identify a suitable class of reagent capable of participating in an exchanging pool as well as capable of irreversible reaction under the same conditions.



**Scheme 1.27:** Reagents and conditions (i)  $\text{CDCl}_3$ ,  $0^\circ\text{C}$ , 48 h,  $[\mathbf{41}] = [\mathbf{42}] = 10\text{ mM}$ ,  $[\mathbf{46}] = 0.5\text{ mM}$ ,  $[\mathbf{45}] = 5, 10\text{ or }15\text{ mM}$ .

Recently, our laboratory reported<sup>154</sup> that nitrones are able to undergo exchange under mild conditions in non-polar solvents. The dynamic process was demonstrated in the

context of a receptor assisted selection of a host for a dicarboxylic acid guest. When two nitrones, **41** and **42** were mixed together at 25 °C in CDCl<sub>3</sub> with 5 mol% hydroxylamine **46** to catalyse the exchange, a four nitrone dynamic library was established (Scheme 1.27). Nitrones, **41** and **42** bear a single amidopyridine recognition, nitrone **43**, is incapable of recognition and nitrone **44** is decorated with two recognition sites. The target, **45**, a glutaric acid derivative, is complementary to the bisamidopyridine bearing nitrone **44** and as such when **45** is added to the library, the concentration of **44** and consequently **43** in the system should be amplified and **41** and **42** inhibited.



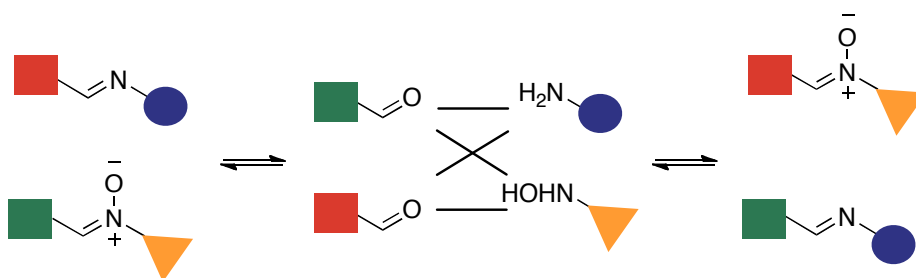
**Figure 1.57:** Ratio of nitrones **43/41** and **43/42** as a function of increasing diacid **45** concentration. Data taken from reference 154.

After 48 h, the equilibrium could be driven to the bisamidopyridine side over the mono amidopyridine nitrone reaching up to a 4:1 ratio (Figure 1.57). These results demonstrated that a dynamic library can be established from interconverting nitrones and that the equilibrium may be perturbed by addition of a template.

Nitrones are also capable of further reaction, participating in an irreversible 1,3 dipolar cycloaddition reaction with a maleimide and their implementation in an autocatalytic self-replicating system was described previously (Scheme 1.21).

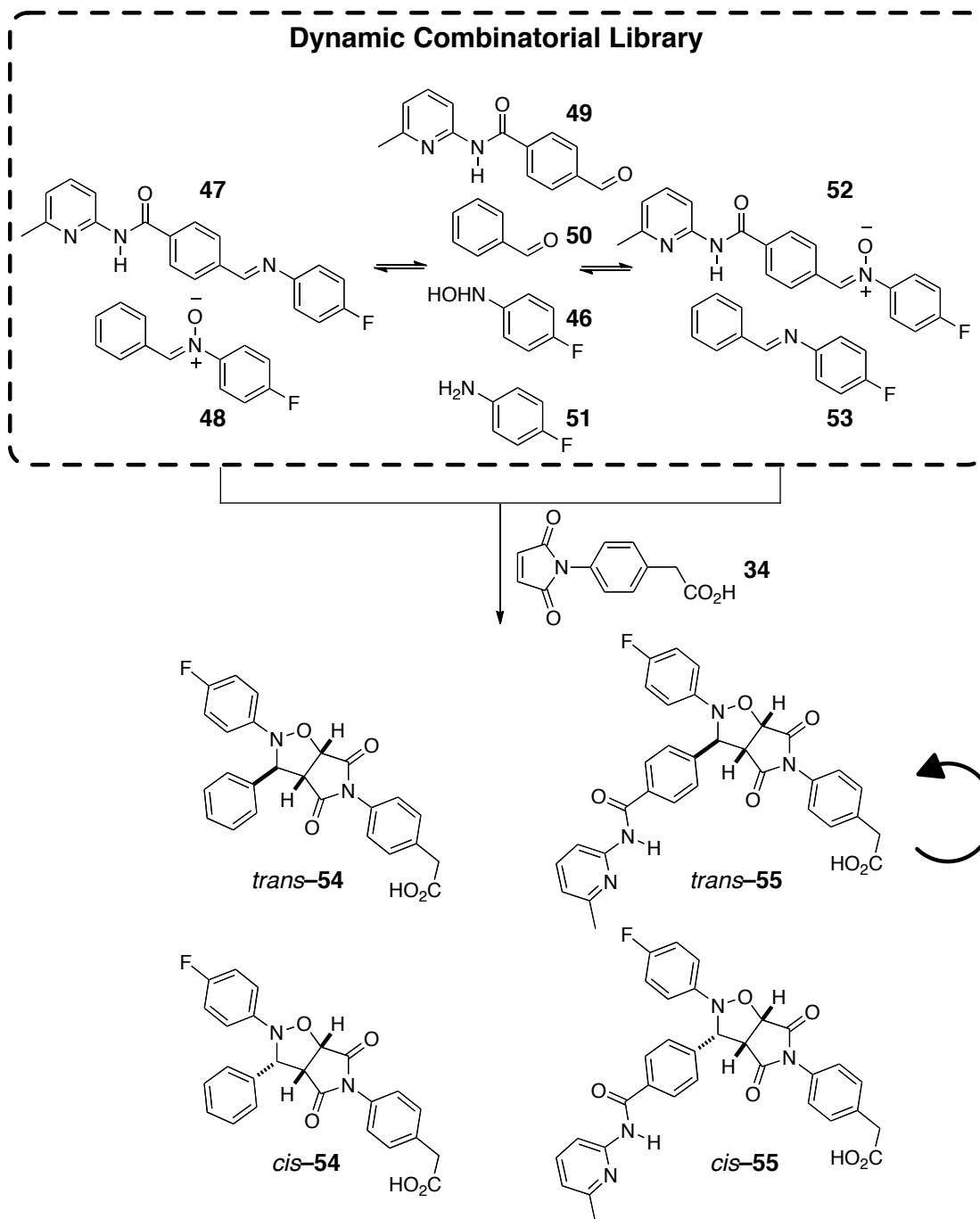
By utilising the dynamic nitrone process coupled to an irreversible cycloaddition reaction a more appropriate example of self-replication from a dynamic library of reagents, as described by the model in Figure 1.56, was produced from our laboratory.<sup>155</sup> The dynamic component of the system was built upon a dynamic

nitron–imine exchange reaction. By understanding that nitrones are able to participate in dynamic processes, we expected that the intermediates in the dynamic process are capable of exchanging with the functionally similar imines. We observed (Scheme 1.28) that upon mixing an equimolar quantity of a nitron with an imine, in  $\text{CD}_2\text{Cl}_2$  saturated with *para*-toluenesulfonic acid, there was breakdown of each of the products to their simpler components. The nitron produced the hydroxylamine and aldehyde and the imine produced the amine and aldehyde. With all four components now available, the hydroxylamine was able to recombine with either of the two aldehydes to form two nitrones and likewise the amine was free to react with either aldehyde to produce two imines. The system eventually was able to reach equilibrium with roughly a 1:1:1:1 ratio of the four products. The net effect is that the functional group located on the aldehyde side of the imine can be transferred to a nitron.



**Scheme 1.28:** Starting on the left hand side, the imine and nitron, which are under dynamic control, can decompose to their building blocks; two aldehydes, a hydroxylamine and an amine. The four simple components can then recombine in any available fashion as described by the black connection lines. These combinations lead back to the original imine and nitron but also to produce a new imine and nitron. The net effect is that a functional group present on an imine, the red square, can be transferred to a nitron.

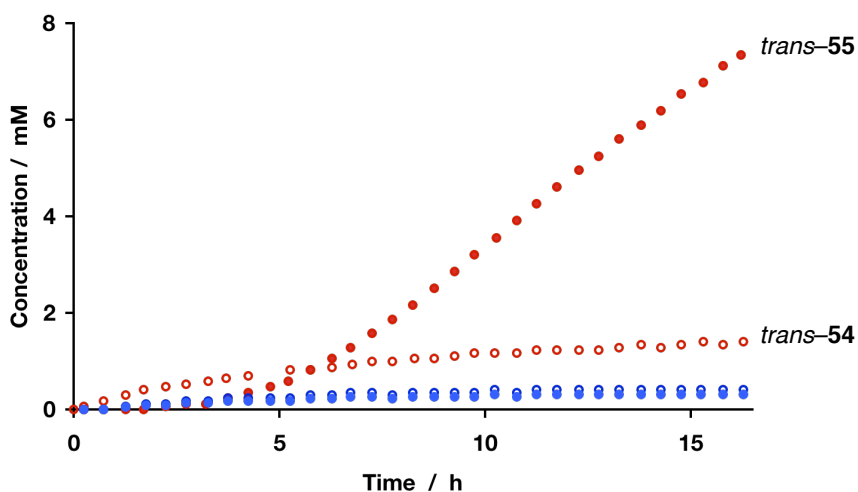
If the rate of consumption of both nitrones by cycloaddition was identical, the library would not be perturbed as both are being removed from the library at equal rates. However if the rate of one reaction was significantly different to the other so that one nitron was consumed significantly faster than the other, the dynamic equilibrium of the system would be disturbed. Le Châtelier's principle states that when a chemical system at equilibrium is disturbed, the equilibrium will shift to counteract the imposed change and a new equilibrium will be established. In our case, the nitron which is consumed will be replenished. Autocatalytic self-replication was utilised as the driving force to consume one nitron from a DCL as shown in Scheme 1.29.



**Scheme 1.29:** Mixing nitronium **48** with imine **47** leads to the formation of an 8 component DCL. The compounds enclosed within the black dashed line are part of a dynamic library. The 1,3 dipolar cycloaddition then consumes nitronium from the library to form the four templates, one of which is an autocatalytic replicator described previously.

The reaction was initiated by mixing three components, **47**, **48** and **34** in equimolar concentrations of 20 mM in  $CD_2Cl_2/p$ -TSA (sat) at 0°C. In mixing the recognition bearing imine **47** with an undecorated nitronium **48**, an eight component DCL was produced, consisting of the added reagents, their simple breakdown components, **49**,

**40**, **46** and **51**, and the new recognition bearing nitrone **52** and undecorated imine **53** which arise from recombination. The presence of maleimide **34** coupled the DCL to an irreversible cycloaddition reaction which furnished four products. From the undecorated nitrone **48**, the undecorated *trans*-**54** and *cis*-**54** templates were produced. From the reaction of nitrone **52**, produced through the DCL, with maleimide **34**, the recognition decorated *trans*-**55** and *cis*-**55** templates are produced. *Trans*-**55** was earlier demonstrated to be an efficient self-replicating template and thus its production will accelerate with time and consume nitrone **52** more rapidly than the undecorated nitrone **48**. There is however, no nitrone **52** in the system upon initiation of the reaction and nitrone **52** must first be provided by the DCL before efficient autocatalysis can proceed. The progress of the reaction was monitored by <sup>1</sup>H NMR spectroscopy by observing a proton spectrum every 30 minutes for 16 h.



**Figure 1.58:** Concentration vs. time profile for the reaction of nitrone **48**, imine **47** and maleimide **34**. Recognition decorated *trans*-**55** represented by filled red circles (●) and *cis*-**55** as filled blue circles (●). Undecorated *trans*-**54** empty red circles (○) and *cis*-**54** empty blue circles (○). Axis not representative of total concentration. Reaction concentration was 20 mM. Data taken from reference 155.

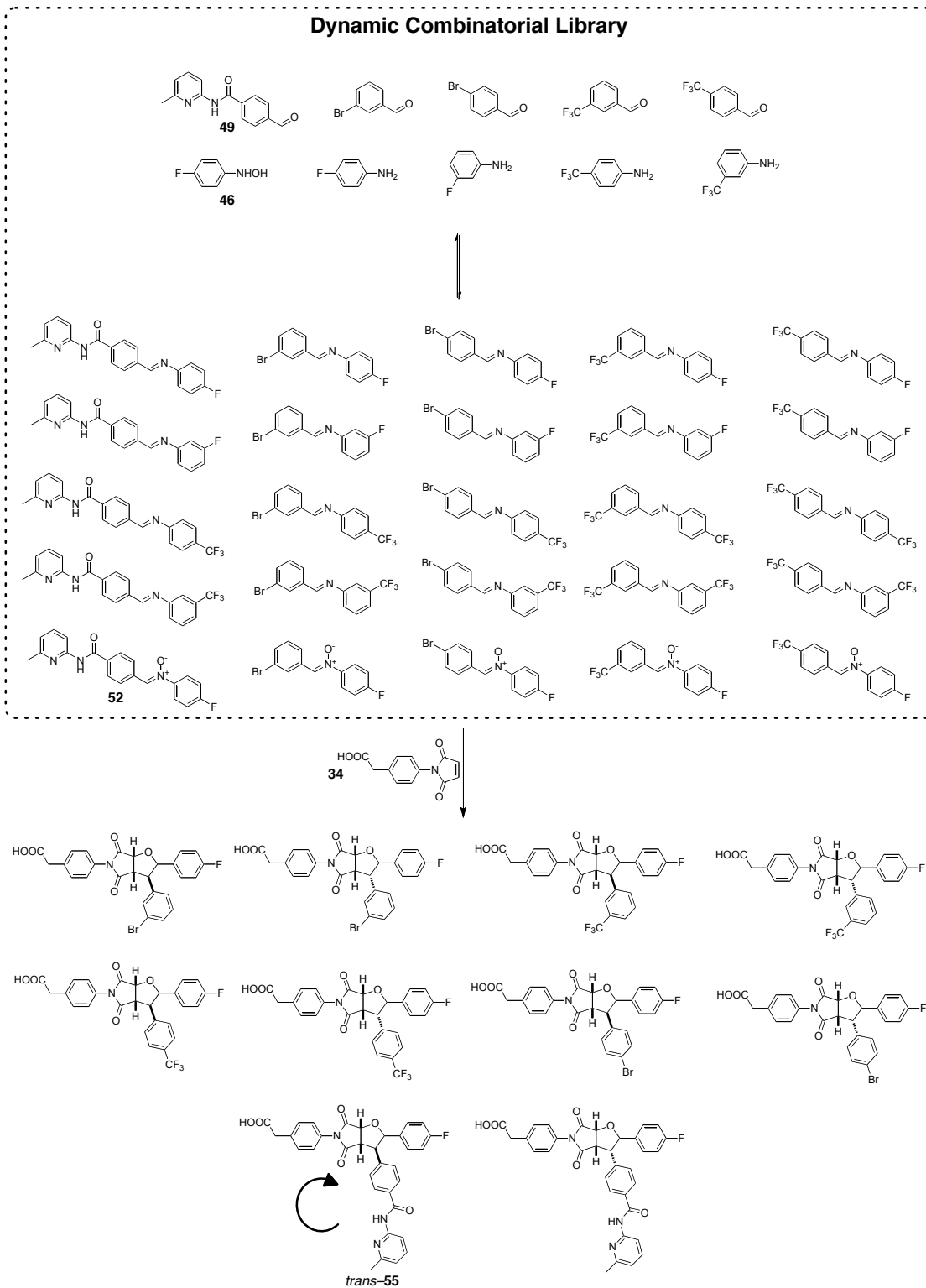
The concentration vs. time profile for the reaction of imine **47**, nitrone **48** and maleimide **34** (Figure 1.58) demonstrated that in the initial stages of the reaction, the production of undecorated *trans*-**54** (empty red circles) was higher than the autocatalytic template *trans*-**55** (filled red circles). This observation is a consequence of two properties in the system, firstly a higher concentration of undecorated nitrone **48** is available for reaction upon initiation and secondly, the concentration of recognition bearing nitrone **52** must reach a concentration above the dissociation

constant ( $K_d = 1$  mM) before efficient recognition mediated process can occur. As time progresses, the DCL provides the nitrone **52** and autocatalytic reaction with maleimide **34** is able to proceed as is seen by the increase in rate of production of *trans*-**55**. As the autocatalytic cycle consumes recognition bearing nitrone **52** from the DCL, the library is producing more nitrone **52** to counteract the change.

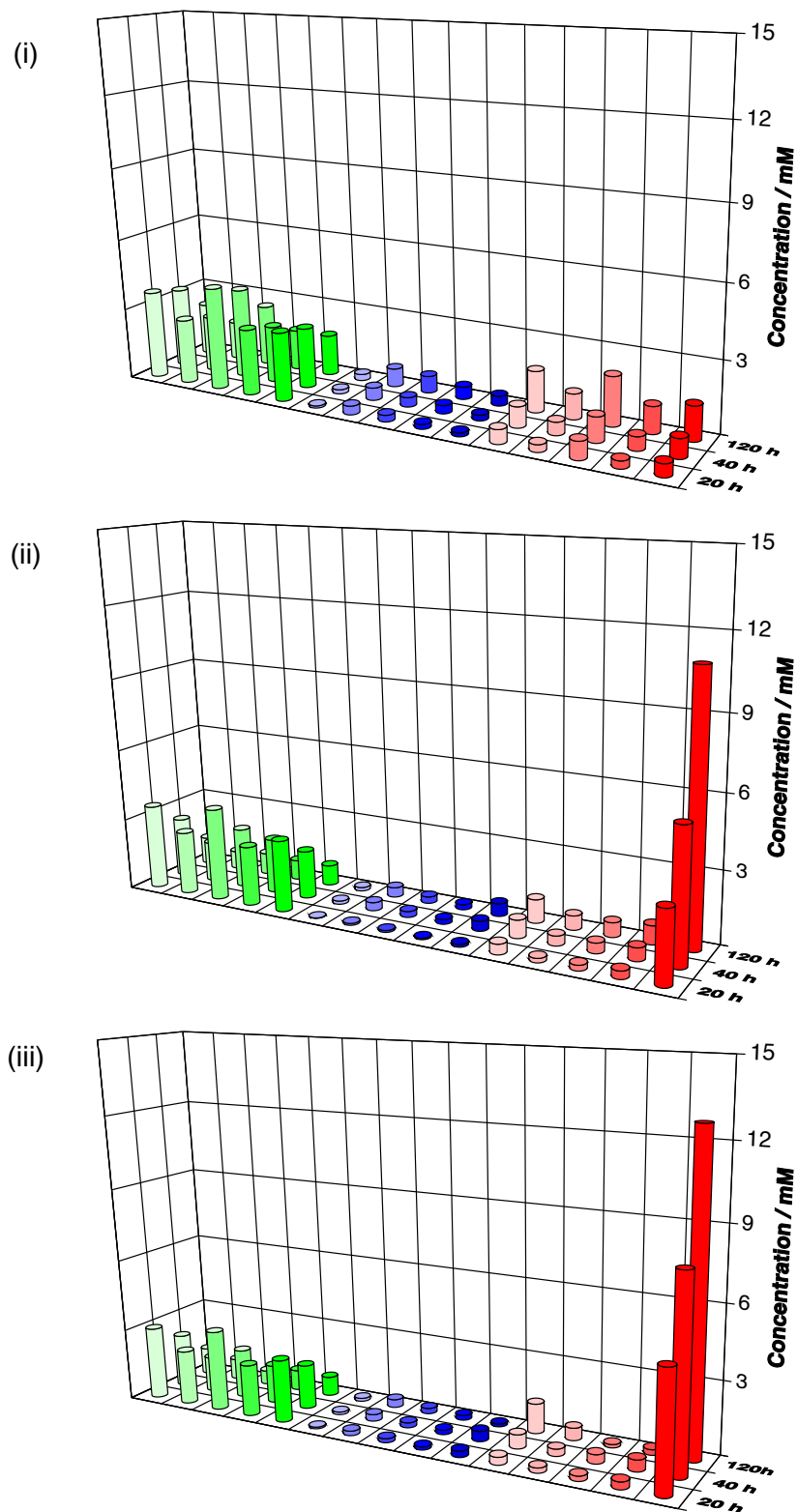
The ability of the DCL to respond to a more severe perturbation was next demonstrated by seeding the reaction with preformed autocatalytic *trans*-**55**. The expectation was that the presence of autocatalytic template upon initiation would serve to remove nitrone **52** at an increased rate and thus drive the system away from equilibrium more rapidly. As a result, the DCL should respond and equilibrate to provide more nitrone **52** more rapidly in comparison to the unseeded system. This expectation came to fruition as the effect was to increase the rate of production of *trans*-**55** so that 11.3 mM was present after 16 h in comparison to 7.7 mM produced in the unseeded reaction after the same time period representing an enhancement of 47%.

Subsequent experiments since then have demonstrated<sup>156</sup> that as the pool of reagents is under dynamic control, it does not matter whether the system is initiated by adding the reagents shown on the left hand side of the DCL (imine **47** and nitrone **48**), the right hand side (imine **53** and nitrone **52**) or in fact the simple components themselves (aldehydes **49** and **50**, hydroxylamine **46** and amine **51**). The same DCL is established and the autocatalytic template *trans*-**55** is always the major product.

To demonstrate that the replicator may still selectively produce itself from a much larger dynamic library, an extended DCL was designed (Scheme 1.30). Based on the same replicating template *trans*-**55**, five aldehydes were selected (one recognition bearing aldehyde **49** and four simple aldehydes), four amines, one hydroxylamine **46** and maleimide **34** for cycloaddition. Upon mixing the eleven components in an equimolar concentration of 20 mM in  $CD_2Cl_2/p$ -TSA (sat) at 0°C, a DCL of five nitrones and twenty imines (thirty five compounds in total) is produced. The DCL is once again connected to the irreversible cycloaddition reaction which produces ten cycloadducts, one of which is the familiar autocatalytic template *trans*-**55**.



**Scheme 1.30:** Extended network produced upon mixing the five aldehydes, four amines and a hydroxylamine. A DCL of 35 components is initially produced before consumption of nitrones by irreversible cycloaddition produces 10 cycloadducts. The compounds enclosed within the black dashed line are part of a dynamic library.



**Figure 1.59:** Results of extended network system showing the concentration of nitrones (green), *cis* (blue) and *trans* (red) cycloadduct products at three time points, 20 h, 40 h and 120 h for (i) the recognition disabled system. It can be seen that there is no selectivity for any product. (ii) The recognition enabled system. The emergence of autocatalytic template *trans*-55 can clearly be observed as the most abundant product of the system. (iii) Template doped recognition enabled system. The observed effect of seeding the reaction is an increase in the initial rate of production of new *trans*-55 which is present in higher concentrations at earlier time points.



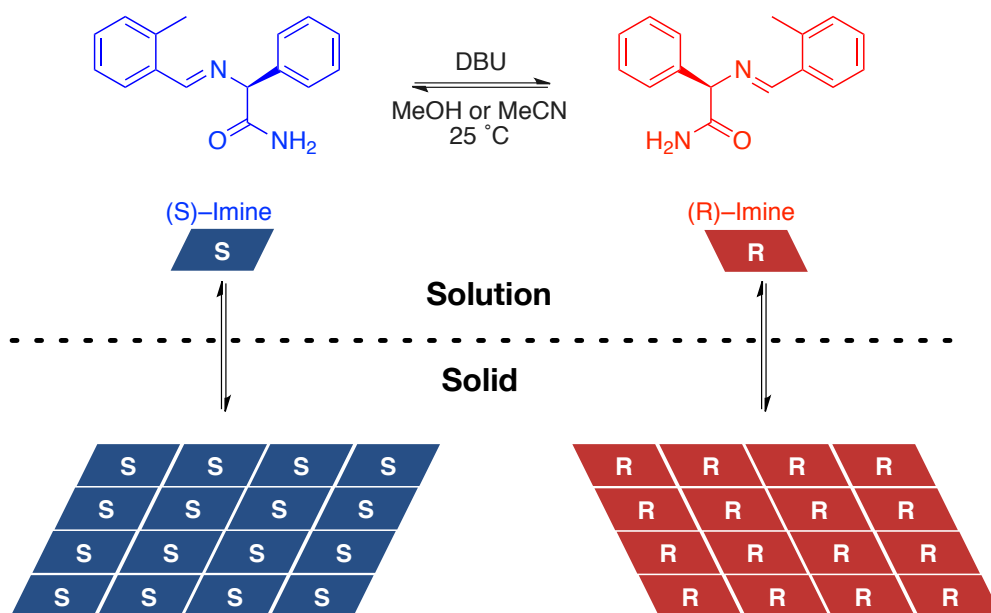
In the absence of recognition effects (achieved by substituting the maleimide acid **34** for a methyl ester), the reaction was slow and proceeded to an overall conversion of ~50% after 120 h with no selectivity and near equal concentrations of the five *trans* cycloadducts and the five *cis* cycloadducts (Figure 1.59 (i)). With the recognition mediated system, the cycloaddition proceeded more rapidly in the production of the major product which was the autocatalytic template *trans*-**55**. The overall conversion after 120 h was now 77% and *trans*-**55** accounted 70% of the product pool reaching a concentration of 10.9 mM, over 7× greater than the next most abundant cycloadduct in solution (Figure 1.59 (ii)).

When the recognition mediated reaction was seeded with preformed *trans*-**55** the observed effect (Figure 1.59 (iii)) was an increase in the initial production of new *trans*-**55** as observed by a higher concentration of new *trans*-**55** at earlier time points (62% enhancement after 20 h in comparison to unseeded reaction). The overall conversion had increased to 79% with *trans*-**55** accounting for 80% of the total cycloadduct pool reaching a concentration of 12.7 mM after 120 h, more than 9× the concentration of the next most abundant cycloadduct in solution.

The systems demonstrated that a replicating template is capable of exploiting and dominating an exchanging pool of reagents in order to amplify its own formation. More significantly, the response of the system to an instruction in the form of added template is a welcome observation as it demonstrates that instructions are able to be processed by the system to produce a response. In this case, the input of replicating template lead to a response which was an enhancement in its own production.

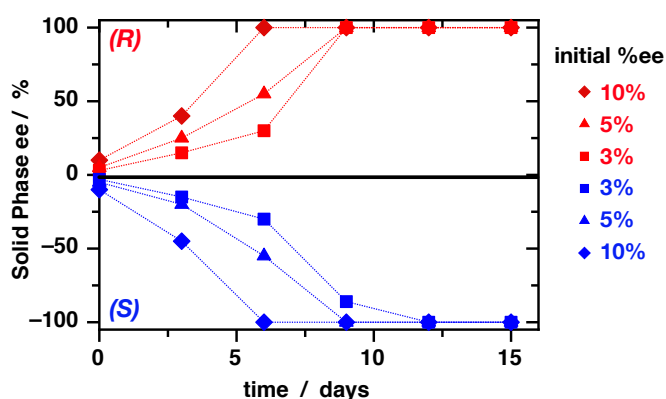
### 1.7 Crystallisation as a Method of Selectivity From a Dynamic Library

The emergence of homochiral biological molecules, *L*-amino acids and *D*-sugars, have been of interest since Pasteur first separated mirror image crystals of a tartrate salt.<sup>157</sup> Blackmond recently explored<sup>158</sup> a system in which two enantiomers form separate *R* and *S* solid phase crystals (known as a conglomerate) and can be induced to undergo rapid racemisation in basic solution (Figure 1.60).



**Figure 1.60:** Racemisation of the imine in basic methanol or acetonitrile provides two enantiomers which may crystallise to form separate homochiral solid crystals. Figure adapted from reference 158.

Solution–solid mixtures of varying ee were magnetically stirred in the absence and in the presence of glass beads to establish a solution–solid equilibrium before initiating the racemisation by adding DBU. Samples of the solid were collected over time and the enantiomeric purity analysed by HPLC. In the presence of glass beads, which cause continuous attrition of the solid crystals by imparting mechanical energy, the authors observed that the ee of the solid rose over time, evolving into a single chiral state from an initial ee imbalance of added solution–solid mixture as low as 2–3% (Figure 1.61). In the absence of glass beads, the solid phase ee remained unaltered from its initial imbalance.



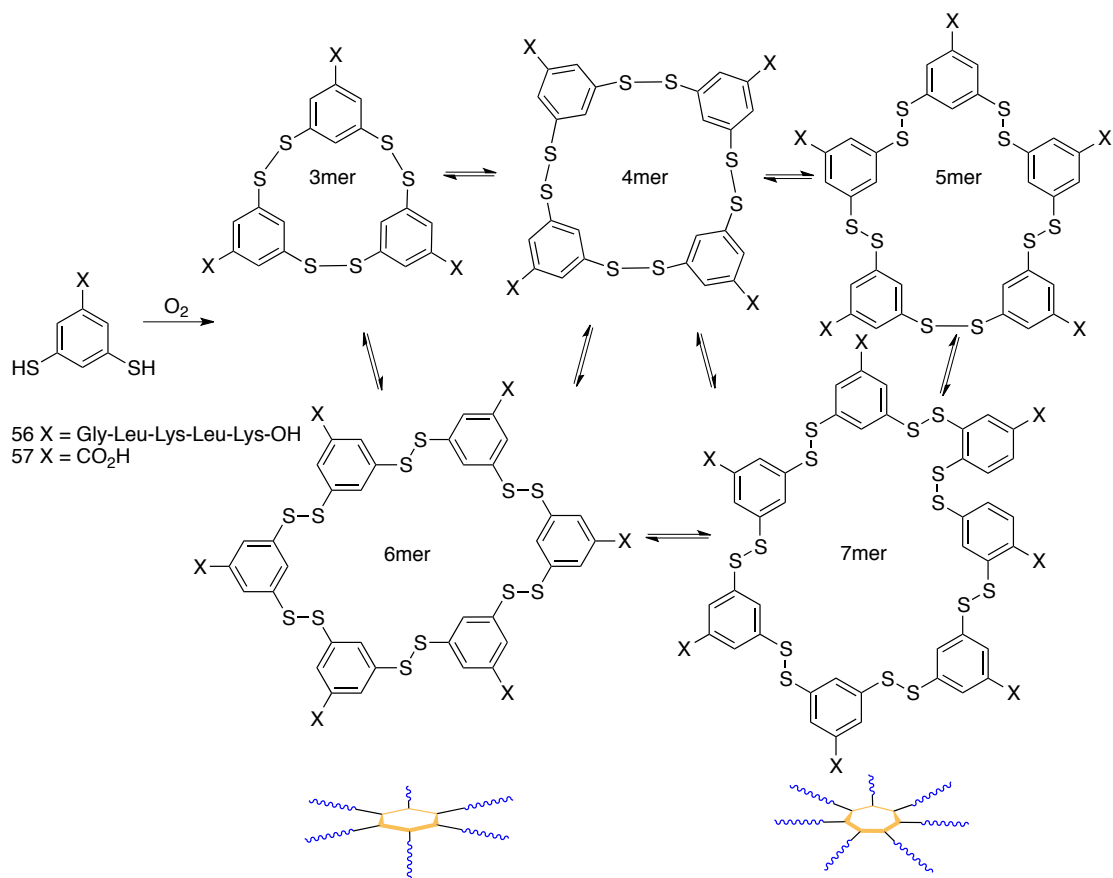
**Figure 1.61:** Attrition–enhanced enrichment of solid phase ee in acetonitrile over time starting from initial ee values as shown. Lines are shown to guide the eye. Data taken from reference 158.

The result demonstrated that the slight enantioimbalance was responsible for directing the formation of a single chiral state. The results were rationalised by considering the dynamic process of crystal dissolution and growth enhanced by attrition along with surface area as a driving force for crystal growth. Smaller crystals, produced by attrition, are able to dissolve more readily than larger crystals. In a saturated solution, this process leads to Ostwald ripening<sup>159</sup> in which large crystals are able to grow at the expense of small crystals regardless of their handedness. A slight imbalance in handedness of large crystals compared to small ones may occur as a result of the minor initial imbalance in ee and along with the continued fragmentation of crystals by mechanical forces leads to a relative increase in surface area of whichever hand is in excess. The authors concluded that although the rate of crystal growth and dissolution have approximately equal rates under near equilibrium conditions, the attrition enriched asymmetry coupled with solution phase racemisation is able to produce a net flow of material towards a single chiral state. The results were the first to demonstrate the process of attrition-enhanced solid phase enantioenrichment of small biologically relevant molecules.

### 1.7.1 Mechanosensitive Self-Replication

In an analogous manner, an alternative approach to self-replication from a dynamic library was recently reported by Otto *et al.*<sup>160</sup> The authors utilised dithiol building blocks which were able to undergo reversible oxidative disulfide formation to form a DCL of macrocycles (Scheme 1.31). The dithiol building blocks were also decorated with a peptide chain with a sequence of alternating hydrophobic (leucine) and hydrophilic (lysine)  $\alpha$ -amino acids which gave the macrocycles the propensity to form fibres held together by  $\beta$ -sheets. The relative stability of each of these macrocycles was expected to be determined by the extent of intermolecular peptide-peptide interaction facilitated by fibre formation.

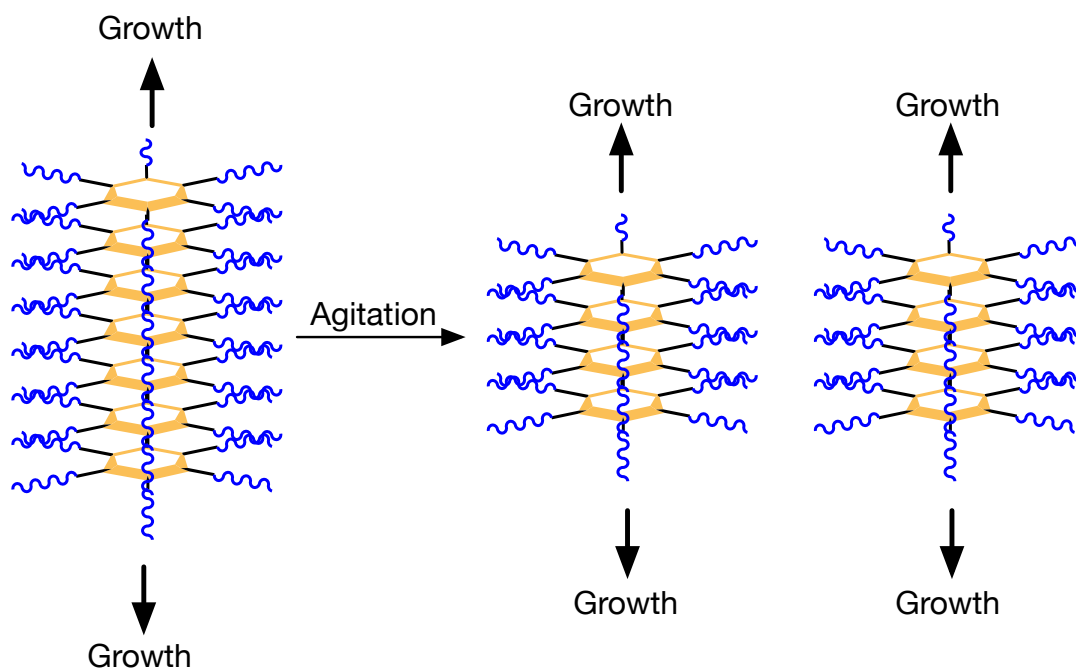
Using a dithiol building block, **57**, devoid of a peptide chain, and thus incapable of fibre formation, a control reaction of stirred **57** in borate buffer at pH 8.0 demonstrated that without intermolecular interactions, the DCL favoured the formation of mostly 3mer and 4mer macrocycles which remained unchanged over 15 days. Repeating the same experiment, instead using the peptide decorated building



**Scheme 1.31:** Dissolving **56** in borate buffer at pH 8.0 in the presence of oxygen from the air allows reversible oxidative disulfide formation to occur forming a library of different sized macrocycles. The 6mer and 7mer are portrayed as cartoons below their structures with a golden ring (dithiols) and blue side chain (amino acid chain).

block, **56**, showed very similar behaviour for the first four days before suddenly changing composition to form 7mer, consuming the 3mer and 4mer in doing so. Interestingly, repeating the experiment with peptide decorated **56** with no agitation provided a DCL with a composition similar to the peptide devoid control **57** DCL of mostly 3mers and 4mers. Only upon agitation did the DCL which consisted mostly of 3mers and 4mers change composition. The authors demonstrated a sigmoidal growth profile for the formation 7mer with stirring and 6mer with shaking. Additionally, when a fresh DCL was seeded with 7mer and stirred or 6mer and shaken, the lag period of the sigmoidal profile was absent. These results indicated that in the absence of intermolecular stabilisation, the thermodynamic equilibrium of cyclic oligomers favours the formation of 3mer and 4mers. However in the presence of intermolecular interactions, 6mer and 7mer were able to form fibres held together by  $\beta$ -sheets allowing stabilisation of these macrocycles.

The mechanism of sigmoidal production of 6mer or 7mer is illustrated in Figure 1.62. The growth of long fibres may proceed on the top or bottom by the macrocycle stabilising an addition macrocycle on its vacant surface. With a single fibre, only two directions of growth are possible. However, agitation of the solution was shown to decrease the average length of fibres which would present more vacant surfaces for growth to occur.



**Figure 1.62:** Dissolving **56** in borate buffer at pH 8.0 in the presence of oxygen from the air allows reversible oxidative disulfide formation to occur forming a library of different sized macrocycles.

However further investigation was needed to explain the difference in selectivity depending on the agitation method. In order to explain this observation the authors described that agitation by shaking was more efficient at breaking long fibres, allowing growth, then stirring was. Also, the authors determined that the formation of fibres was a kinetic process, solutions of only 7mer or 6mer persisted for 3 months without agitation.

To explain the selectivity, the authors proposed that fibres formed from 7mer were able to be broken by the mechanically less stressful stirring and thus growth (more 7mer production) was enhanced by presenting more growth sites. Whereas the mechanical force of stirring was not sufficient to break 6mer fibres therefore less growth sites were present and as such less 6mer was produced. In order to explain the opposite selectivity when shaken, the authors determined that shaking provided

sufficient mechanical force to break 6mer fibres allowing a greater number of growth sites and hence greater 6mer growth. The growth of 6mers was determined to be kinetically faster than 7mer hence with a great enough force to disrupt fibre growth of both 6mer and 7mer by shaking, the more rapid kinetic formation of 6mer allowed a higher concentration of 6mer to be produced.

The approach, coined mechanosensitive self-replication serves to provide an example of mechanical forces acting as a selection pressure to determine the outcome of competition between replicators.

### **1.8 Extended Systems and Molecular Logic**

In each of the above systems, one or more products have utilised their ability to direct their own synthesis to selectively amplify themselves from a pool of reagents. When combining more than one replicating template into a system however, two or more templates are each programmed to produce themselves and thus will compete for common material. The instruction to produce themselves is realised internally when each template is produced.

This instruction could however be issued externally as has been demonstrated in many cases above. In the initial stages of an autocatalytic self-replicating reaction, the rate of reaction is slow as no template is available to direct the reaction. If preformed template is added to the reaction at the start, there exists template to direct the reaction upon initiation and thus the reaction proceeds at its maximum rate upon initiation. In essence we issued an instruction by inputting the template and that instruction is to make that template which is added. This instruction is rather simple however, if there is only one autocatalytic template in the system, only one template instruction can be input, and only one template output can be received. The story becomes a more interesting when more than one autocatalytic reaction is present. If two or more templates produced by a system are self-replicators then two or more instructions may be issued by added either of these templates to a system upon initiation.

This behaviour of issuing instructions and observing an output or response is analogous to binary logical operations which are the cornerstone of modern day

electronics. Binary logic is used to represent inputs and outputs as simple ones or zeros. In an electrical circuit, an input is an electrical signal transmitted through the circuit and is processed by a gate or switch. If input criteria match the required criteria as dictated by Boolean functions, then an electrical signal output is produced. The truth tables for the most common logical operations used by modern day electronics are demonstrated on Figure 1.63.

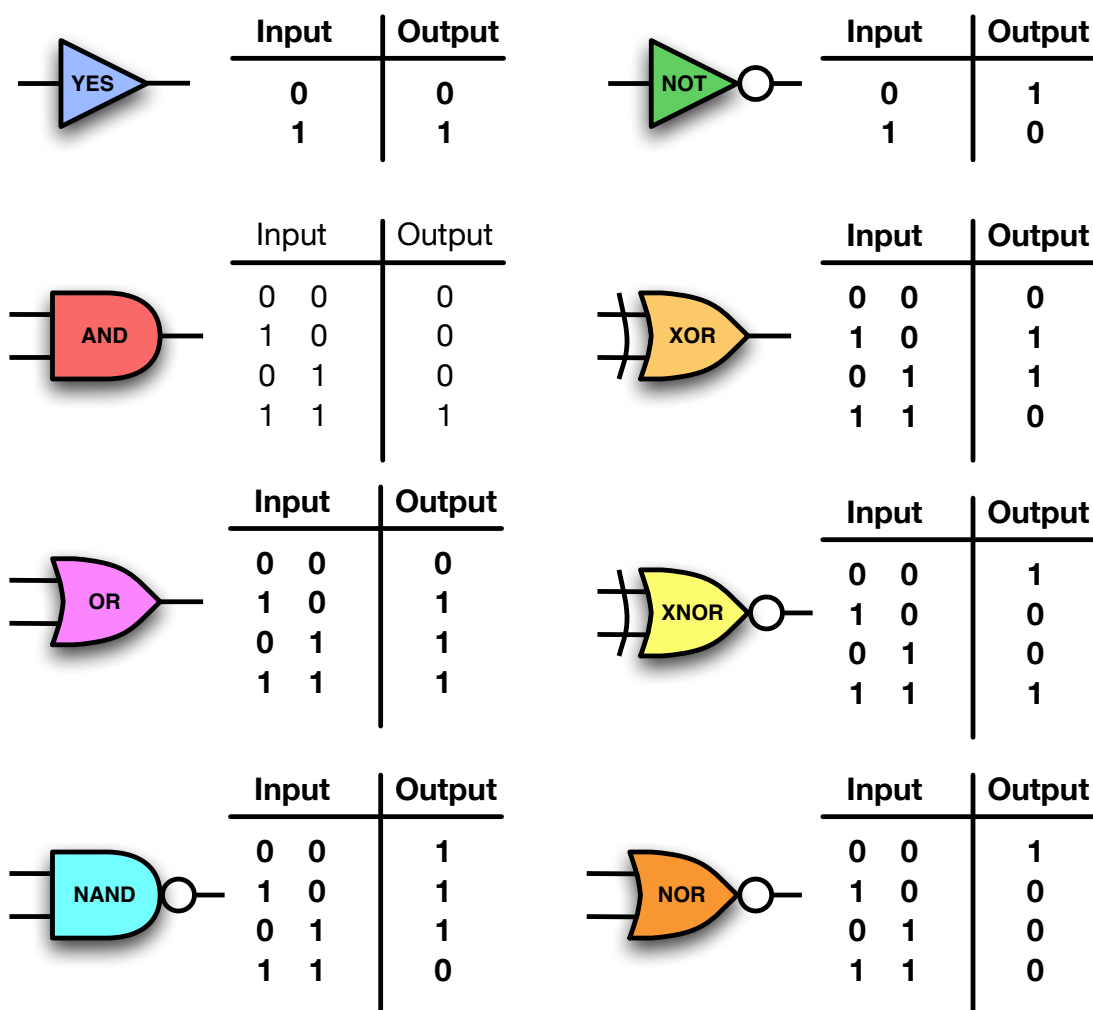


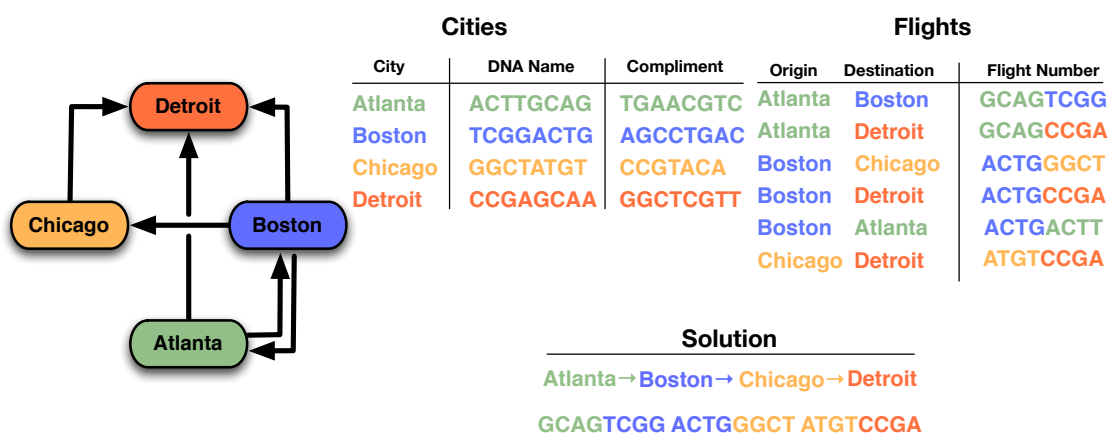
Figure 1.63: One and two input Boolean logic gates with their circuit symbols and truth tables.

A molecular system can also be described in binary logic as either ones or zeros and molecular logic has generated significant interest.<sup>161-164</sup> For an input, a one represents an instruction given to the system by changing its properties which for example could be the addition of a compound. A zero indicates no instruction. For the outputs, a one describes that there is some observed effect of the instruction(s) as there is a change in the properties of the system. A zero describes that no change or output has been observed.

Metastable mechanically interlocked molecules have been demonstrated<sup>165,166</sup> to be able to perform switching operations through the translational motion between two or more states. One of the most prominent examples was demonstrated by Heath and Stoddart<sup>167,168</sup> and made use of a metastable [2] rotaxane in which the macrocyclic component could move between two sites and represents the binary 1, or 0. The [2] rotaxane was fabricated on to a nanowire device and functioned as a molecular memory unit with exceptionally high density. This system is discussed further in Section 5.1.

### 1.8.1 DNA Logic

The concept of molecular computing was first realised by Adleman<sup>169</sup> in 1994 who produced a DNA based computer to solve a Hamiltonian Path problem. A Hamiltonian Path is a route through a directed network which passes through each node only once and the problem is how this can be achieved. In order to demonstrate the technique, a problem is demonstrated by considering four cities (nodes) connected by certain flights (directed edge)<sup>170</sup>. The solution to the problem is a route which passes through all cities only once (Figure 1.64)

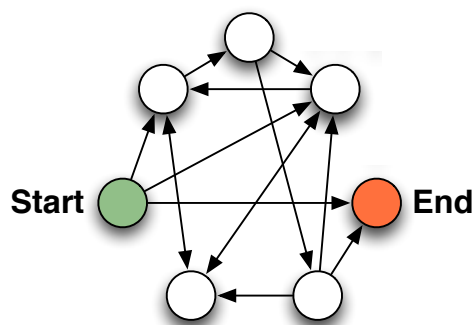


**Figure 1.64.** The four node, six directed edge graph of the Hamiltonian Path problem. Each city (node) is given a DNA name represented by its complementary strand. Each flight between cities (directed edge) is given a flight number which is the second name of each origin city and the first name of the destination city. The solution to the problem is shown below with the route Atlanta→Boston→Chicago→Detroit and represented by the combination of flight numbers. Figure adapted from reference 170.

Using DNA to solve the problem, Adleman assigned each city a unique DNA sequence of 8 bases, to be thought of as a first name of four bases and a second name of the remaining four bases. The complementary city names are the physically



synthesised fragments which would represent the cities in the system. Each flight is given a flight number which is the second name of the city of origin followed by the first name of the destination city. The solution to the Hamilton Path is the route which passes through all cities and only once along the way. The solution is the DNA sequence representing the flights from Atlanta→Detroit which pass through each city only once. When the 'flight' strands hybridise with their complementary cities, it represents a joining of the cities and these joinings would be cemented by a DNA ligase in the reaction mixture. In order to separate the correct pathway from the rest of the solution, two further techniques were used. PCR was used to amplify the correct pathway using two primers working in concert, one which represented the start city and one to represent the final destination. The sequence of correct length was then identified using gel electrophoresis. The actual experiment was carried out with a more complicated example of seven cities with 14 flights (Figure 1.65).

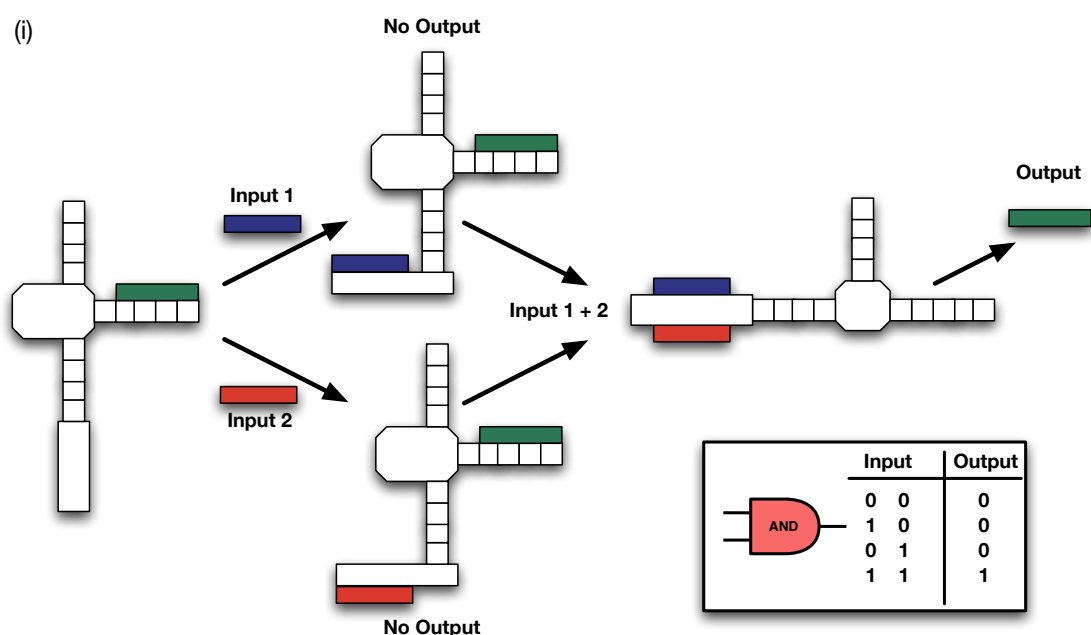


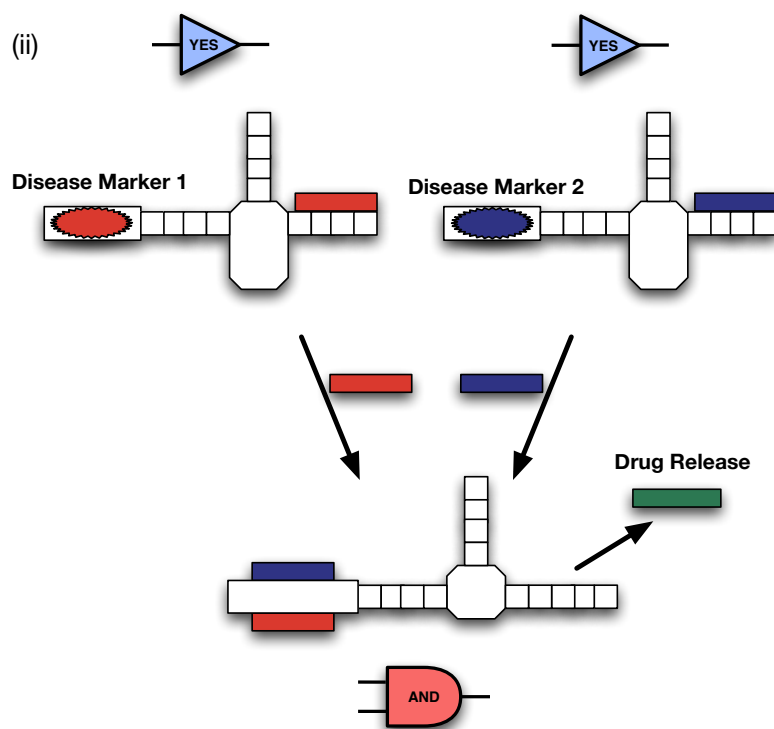
**Figure 1.65:** Directed graph for the 7 city, 14 flight Hamiltonian Path problem solved by Adleman's DNA computer.

The advantages of a DNA computer is that every pathway is probed at the same time in parallel whereas traditional silicone based computers perform tasks in a linear manner. DNA computers therefore have the possibility to outperform traditional computers to simulate the tasks of a large networks and work is on going in this field with Shapiro<sup>171</sup> demonstrating DNA computers successfully solving more complex calculations which reach beyond simple human comprehension, (*i.e.* would require a computer to perform).

The DNA computer demonstrates the promising avenues of research which can be explored by combining the properties of computing with biological processes. DNA serves as the perfect platform for such explorations as it contains information in the sequence of its base pairs and which gives it the potential to be recognised by a

complementary sequence if it possess the correct information. This association may then trigger a downstream process. Stojanovic and co workers have focused their research on molecular logic in DNA based systems.<sup>172,173</sup> Their long term goal is to construct a macromolecular system able to recognise and enter specific cell types. When inside, the system would then be able to sense the molecular markers of diseases and release an appropriate drug accordingly.<sup>172</sup> Accordingly they have developed several ribozyme based logic gates whose self-cleavage activity is regulated by the presence of oligonucleotide strands as inputs based on the work originally demonstrated<sup>174,175</sup> by Breaker who developed allosteric ribozymes using the hammerhead ribozyme secondary structural motif. When an oligonucleotide strand complementary to a region of the ribozyme is present, the association of the strand triggers a conformational change in the RNA structure which leads to a change (activation or deactivation) in the catalytic properties of the ribozyme. The outputs of these systems arise from self-cleavage of the ribozyme and can be oligonucleotides, peptides,<sup>176</sup> small molecules<sup>177</sup> or a fluorescent signal as was demonstrated in a molecular automata of RNA gates which was programmed to be unbeatable in a game of tic-tac-toe.<sup>178,179</sup> Stojanovic successfully demonstrated that ribozyme gates could be created to perform many logical operations such as AND<sup>173</sup> (Figure 1.66 (i)), NOT,<sup>173</sup> XOR<sup>173</sup> and more complicated operations<sup>180,181</sup> as well as chain logic<sup>182</sup> using oligonucleotide outputs from one gate as the input for a subsequent gate.





**Figure 1.66** Schematic representation of Stojanovic's ribozyme based molecular logic gates (i) The AND gate is able to recognise two oligonucleotide inputs 1 (blue) and 2 (red). When only input 1 is bound, there is no catalytic activity of the gate. When only input 2 is bound, likewise there is no catalytic activity of the gate. When both 1 and 2 are bound, the RNA gate is catalytically active and releases the output oligonucleotide. (ii) Schematic representation of the focus to produce an automated therapeutic system using chain logic. Each YES gate recognises a molecular component (protein marker, oligonucleotide etc) of a disease and releases the oligonucleotide products accordingly. When both disease markers are present, both outputs are released and serve as inputs for a downstream AND gate described in Figure 1.63. The overall effect is an output of an appropriate drug when both disease markers are detected. Figures adapted from reference 172.

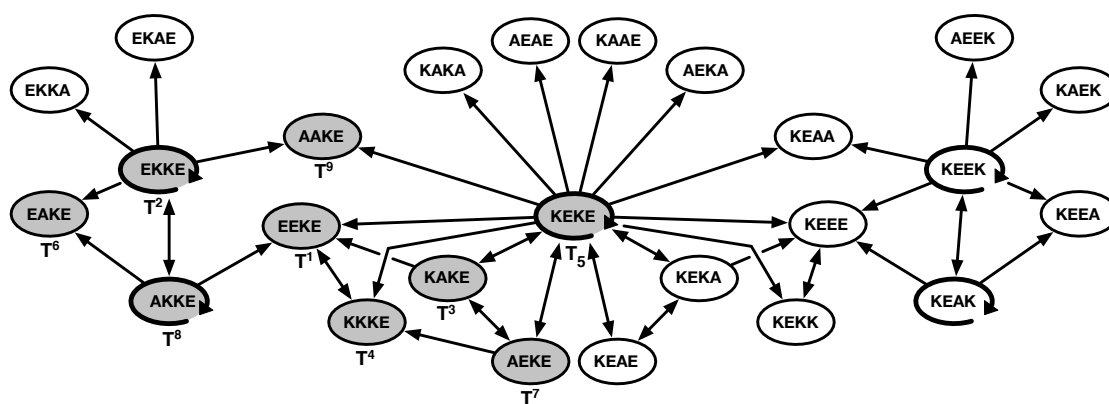
The long term goal of the research is the creation of a therapeutic system such as is demonstrated in Figure 1.66 (ii). The system, comprised of multiple gates, would use disease biomarkers as inputs for a sensor (YES gate) which would release an oligonucleotide output which would serve as an input for a downstream AND gate. Only when both biomarkers are present are both sensors activated and the two outputs released. These two outputs then serve as inputs for the AND gate which releases a small molecule drug in response.

### 1.8.2 Extended Systems and Molecular Logic with Self-Replicators

Self-replication has also been applied to the creation of molecular networks. The inherent complementarity of templates either to themselves as self-replicators or to complementary products as reciprocal replicators or cross-catalysts makes their use

to create highly connected networks an obvious possibility. These networks then have the ability to express forms of logic if inputs are provided.

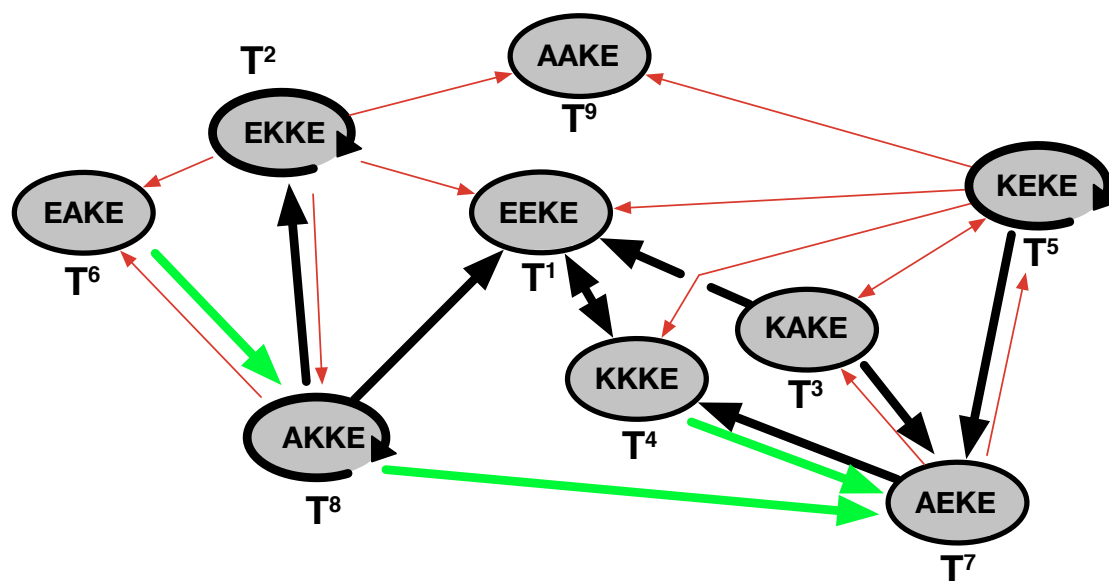
One most extended network of replicators reported comes from the labs of Ghadiri based on self-replicating peptides once again.<sup>183</sup> Starting with 81 sequence similar 32 residue coiled-coil peptide templates, the relative stability of all plausible [substrate•template] assemblies leading to autocatalysis or cross-catalysis were calculated. Those assemblies whose free energies fell within a selection criteria determined from previous experience<sup>184</sup> were selected as viable template directed pathways within the network. From the results, a calculated graphical architecture of the network containing 25 nodes and 53 directed edges was produced (Figure 1.67). Each node of the network represented a unique template sequence with the nomenclature denoting the identity of amino acid residue at the 8,13,15 and 20 position respectively. Each arrow (or directed edge) represented a template directed pathway with the arrow pointing from the template to the product (template→product). Autocatalytic templates are represented by curved arrows.



**Figure 1.67.** Calculated network architecture comprising of 25 nodes joined by 53 vector edges. Each node represents a template or product with the sequence of the varied positions named in the node. Each arrow (directed edge) represents cross-catalytic template assisted pathways with the arrow pointing from the template to the product. Nodes with curved arrows represent autocatalytic products. Nodes in grey and given identity T<sub>1</sub>–T<sub>9</sub> represent the sequences evaluated by experiment. Figure adapted from reference 183.

To validate the network produced by rational computation design, a subunit of the network (highlighted by grey nodes in Figure 1.67) was selected, its fragments and templates synthesised and the behaviour investigated experimentally. The experimental results, graphically represented in Figure 1.68, found that 8 nodes existed in the network with 14 directed edges including the three autocatalytic nodes

(T<sup>9</sup> shown in figure was not found to be part of the network as highlighted by the red arrows). Of the nine possible products, the major products of the system were T<sup>1</sup>, T<sup>2</sup>, T<sup>4</sup>, T<sup>7</sup> and T<sup>8</sup>. Not all of the computationally derived pathways were observed experimentally and this observation was explained to be as a result of the strong autocatalytic cycles of some nodes which inhibited template from being available for crosscatalytic pathways.



**Figure 1.68.** Experimentally derived network architecture comprising of 8 nodes (T<sup>9</sup> not included) joined by 14 vector edges. Each node represents a template or product with the sequence of the varied positions named in the node. Each black and green arrow (directed edge) represents cross template assisted pathways with the arrow pointing from the template to the product. Nodes with curved arrows represent autocatalytic products. Directed edges coloured black were computationally envisaged and experimentally realised. Directed edges coloured red were computationally derived connections which were not realised. Directed edges coloured green are connections not considered by computational studies but found to exist experimentally. Figure adapted from reference 183.

The most interesting experimental results from the network came when attempting to determine the extent of network connectivity. The network reaction was run nine further times, each time in the presence of each of the nine templates T<sup>1</sup> – T<sup>9</sup>. Monitoring the reaction progress over time by HPLC analysis showed that, in the initial stages of the reaction, each added template increases the rate of formation of not only itself, but certain adjacent templates as a result of auto and crosscatalysis. Further to this, these initial rate enhancements lead to an overall increase in yields of all products in the mixture at later time points demonstrating the operation of auto and extended cross-catalytic connections in the network.

The results of the experimental analysis indicated that the theoretical approach to peptide network design can successfully offer a guide for designing relatively complex networks. This work highlights that in a highly connected system, if single events are enhanced, the whole system can benefit.

Ghadiri identified that subunits of the experimental network were able to display logical operations akin to that of computational electronics thus allowing the network to act as a rudimentary molecular computer.<sup>169-171</sup> Further investigation of the experimental system revealed that a fragment of 5 nodes with well established connections was capable of performing logical operations. The authors successfully demonstrated<sup>185</sup> that using preformed templates (nodes) as inputs and observing enhancements of template production as outputs, the logical OR, NOR and NOTIF could be produced.



# 2

## Directing Systems of Multiple Replicators

### 2.1 Complex Systems

The thought of intentionally creating mixtures of compounds tends to make synthetic organic chemists uncomfortable. Traditional organic synthesis is performed in a linear fashion: one set of reagents are chosen to perform one transformation to create a single desired product. This product is then usually purified before subsequent transformation into another compound and so this repeats until ultimately, the target compound is reached. Each step is carefully thought out and conditions optimised each time to achieve the maximum yield of the one desired compound of each transformation. Whilst synthetic chemists do their best to keep things simple, complexity exists all around us in many forms.

The term *complexity* does not refer to something which may be difficult to understand but refers to a system which possess many interconnected processes which can lead to the system expressing some sort of behaviour that is more than the sum of its parts, so called *emergent properties*.

Complex systems can be found all around us. Weather systems are an example of a complex system linked together in the atmosphere. Whilst our knowledge of the fundamental processes responsible for weather pattern formation allows short term weather forecasting, the intricate complexity of the system means we have no real chance of predicting the weather of a particular date say ten years from now. *The butterfly effect*, that a butterfly flapping its wings could cause a hurricane on the other side of the world is a commonly used metaphor which can describe the connectivity of weather patterns. Stock markets around the world are connected with the unfortunate effect that in the quest for rewards, the excessive risk taken by banks in America backfired and upset the financial stability of stock markets around the world

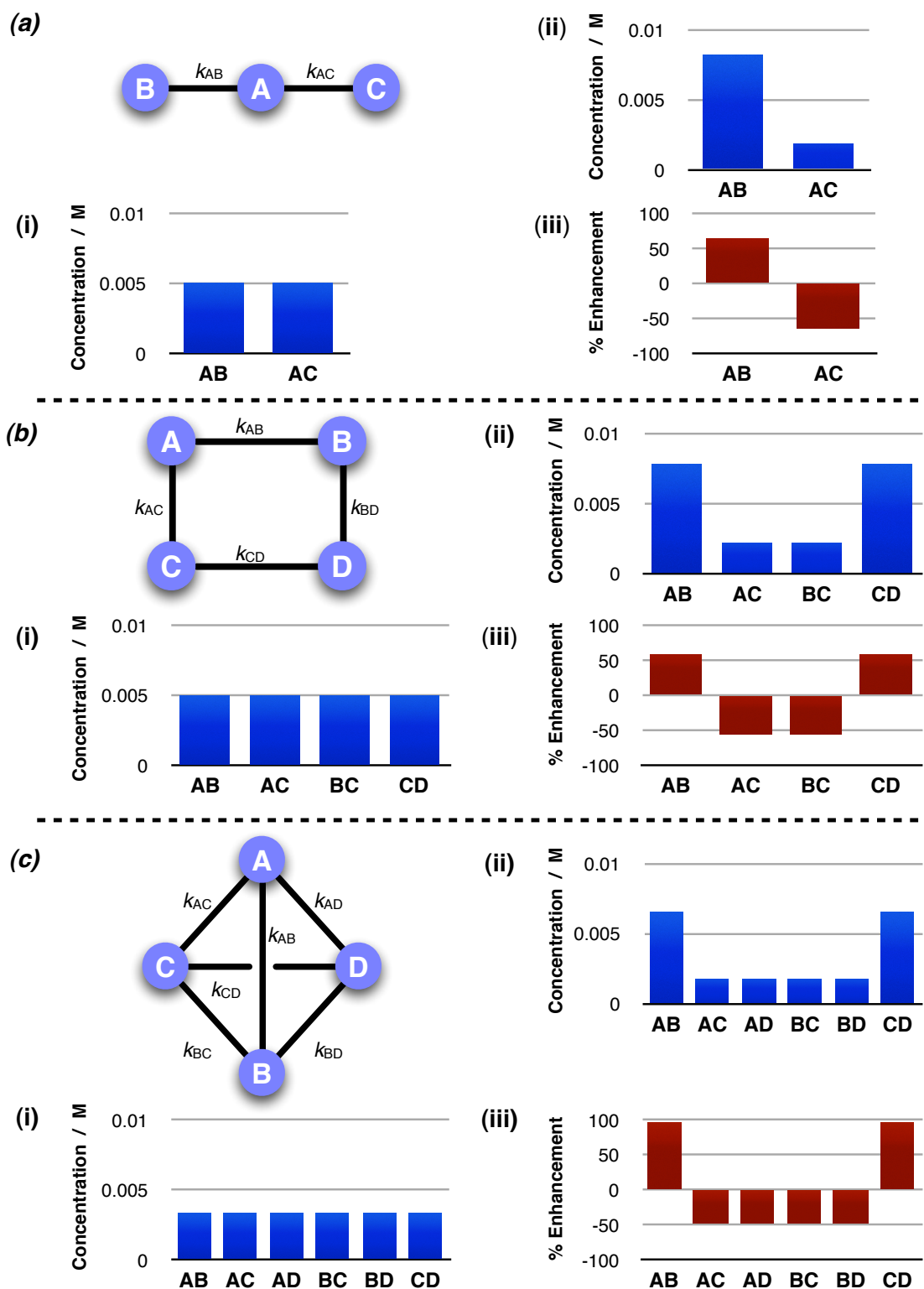


which led to an era of austerity. Life as we know it has evolved to embrace complexity. One of the basic building blocks in our body, the cell, is a complex system filled with many forms of genetic material which interact with and rely on each other to maintain cellular activities. Whilst it is possible to analyse the individual components that occupy the cell, it is impossible to gain an understanding of the entire cell by these simple parts alone. The key to the success of complex systems is the interconnectivity of processes. In chemical terms this complexity can be encompassed by reactions sharing common reagents or the presence of feedback loops influencing reaction rates.

### 2.1.1 Computational Network Experiment

A simple computational simulation, run using SimFit<sup>186</sup> (see Chapter 7.7, Appendix A3 and A4) can illustrate how increasing the number of interconnections within a chemical system can lead to system-level behaviour and enhanced selectivity. In a small system of three equimolar reagents, **A**, **B** and **C**, at 10 mM, reagent **A** is able to react with either **B** or **C** to form **AB** or **AC** respectively (Figure 2.1 (a)). The reaction rate constants for each reaction,  $k_{AB}$  and  $k_{BC}$ , were set to be equal at  $0.03 \text{ M}^{-1}\text{s}^{-1}$ . Upon completion of the reaction, unsurprisingly **AB** and **BC** were observed in the product pool at an equal concentration (5 mM) (Figure 2.1 (a) (i)). If a form of catalysis is simulated thereby increasing the rate of formation of **AB** by a factor of 10 ( $k_{AB} = 0.3 \text{ M}^{-1}\text{s}^{-1}$ ) and the calculation repeated, then at the end of the reaction, unsurprisingly, there is an enhancement in the concentration of **AB** over **BC** (Figure 2.1 (a) (ii) and (iii)). This, rather obvious, result serves to demonstrate that in a connected system, changes in one area of the system can have an effect on connected components. This might be thought of as a traditional chemical reaction in which the conditions are optimised to produce one preferred product over another product.

The number of connections within the system can be increased by expanding the number of components from three to four equimolar reagents. Once again **A** can react with **B** or **C** to form **AB** or **AC** respectively but now the fourth reagent, **D**, can also react with **B** or **C** to form **BD** or **CD** respectively. The network can be represented in the form of a square with each reagent a corner and each side representing a reaction (Figure 2.1(b)). If the four reaction rate constants,  $k_{AB}$   $k_{AC}$   $k_{BD}$



**Figure 2.1:** Three examples of reaction systems. Reagents are represented as blue circles with a black line connecting reagents representing a reaction channel. **(a)**: Three reagents, two reaction channel system, **(b)**: Four reagents, four reaction channel system, **(c)**: Four reagents, six reaction channel system. All calculations performed with starting concentration of 0.010 M of each reagent. Concentration of products after 40 h reaction time shown in graphs when: (i) all reactions rates are equal at  $0.03 \text{ M}^{-1}\text{s}^{-1}$  (ii) the rate of one reaction, to form **AB** is enhanced 10-fold,  $k_{AB} = 0.3 \text{ M}^{-1}\text{s}^{-1}$ . (iii) Graph of the enhancement in concentrations when comparing the results of (ii) to (i). Percentage enhancement calculated from  $(([XY]_{k_{AB} = 0.3 \text{ M}^{-1}\text{s}^{-1}}) - ([XY]_{k_{AB} = 0.03 \text{ M}^{-1}\text{s}^{-1}})) / ([XY]_{k_{AB} = 0.03 \text{ M}^{-1}\text{s}^{-1}}) \times 100\%$ .

and  $k_{CD}$ , are set to be equal at  $0.03 \text{ M}^{-1}\text{s}^{-1}$  then, as expected, each of the four products are formed in an equal concentration (Figure 2.1 (b) (i)). If one reaction in the system, that of **A** and **B** forming **AB**, is accelerated 10-fold whilst holding all other reaction constants as before then the outcome of the system changes to reflect the new conditions (Figure 2.1 (b) (ii)). The product distribution now shows that the concentration of **AB** has increased which would be expected from the increased rate. The concentration of **AB** has been enhanced by 53% in comparison to the uncatalysed system (Figure 2.1 (b) (iii)). This increased reactivity to form **AB** swiftly consumes all of **A** and **B** leaving **C** and **D** to react together to form **CD**. Although the formation of **CD** proceeds at a slower rate, it is able to reach the same concentration as **AB**. Even a small system of only four components is able to display a primitive form of system level behaviour, that an uncatalysed reaction is enhanced not by direct catalysis but through connectivity to a catalysed process.

Next, the effect of increasing the connectivity of the system was investigated. A system of four equimolar reagents was used again with the same starting concentration of each reagent but now two further reaction pathways were added. Each reagent is now allowed to react with all of the other reagents leading to six possible products (Figure 2.1 (c)). The network can now be represented as a tetrahedron with each corner a reagent and each edge a reactive pathway. As before, when the rate constants of each process are identical ( $0.03 \text{ M}^{-1}\text{s}^{-1}$ ), then the concentration of each product reaches the same level at completion (Figure 2.1 (c) (i)). If a form of catalysis is again simulated by increasing the reaction rate to form **AB** 10-fold ( $0.3 \text{ M}^{-1}\text{s}^{-1}$ ) whilst keeping the other rates the same, then at first glance the product distribution is similar with the concentration of **AB** and **CD** up-regulated (Figure 2.1 (c) (ii)). Closer inspection of the results reveals that the level of enhancement of **AB** and **CD** are significantly higher in the tetrahedral system in comparison to the square system (Figure 2.1 (c) (iii)). There is now a 100% enhancement of **AB** and **CD** in comparison to the uncatalysed tetrahedron system, far greater than the square system in which a 53% enhancement of **AB** and **CD** is seen in comparison to its uncatalysed system.

This simulation of a simple chemical system highlights that even a small system of four reagents is able to show signs of system level behaviour, namely the enhancement of products not directly catalysed but as a result of connectivity to a

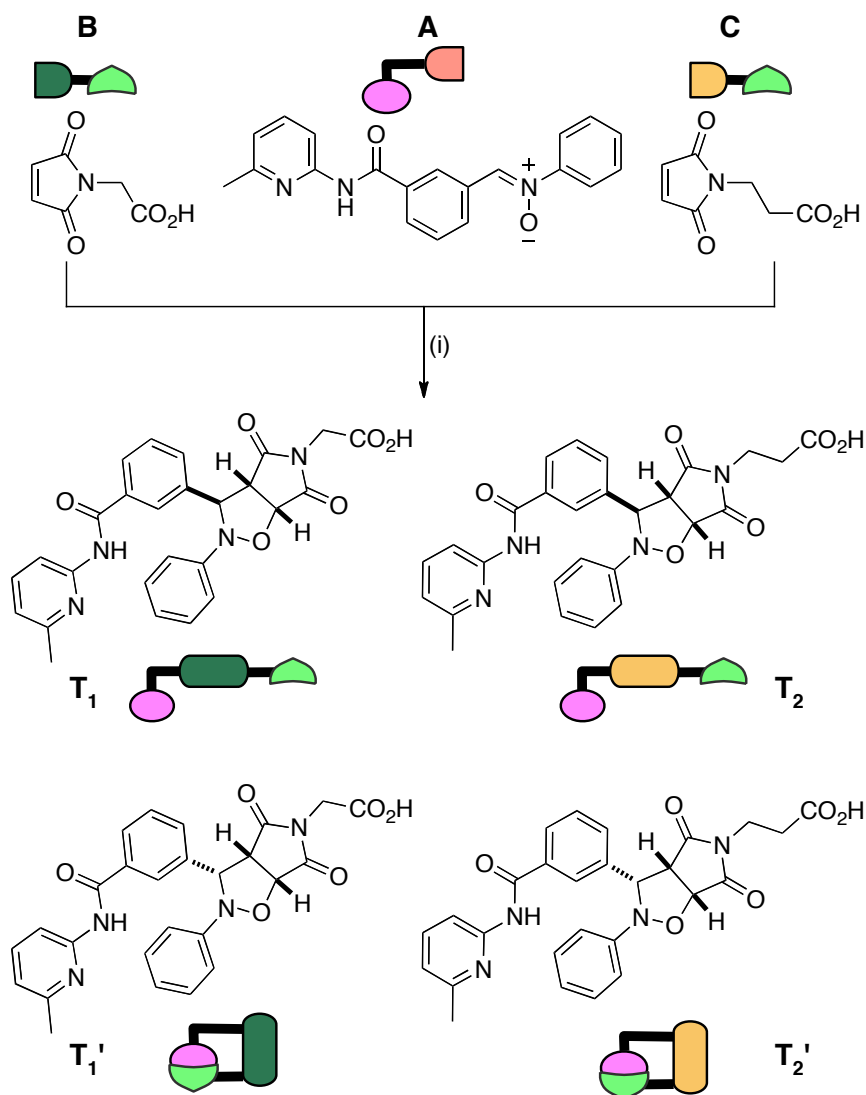
catalysed process. The effect of increasing the complexity of the system, by increasing the connectivity of reagents through reaction pathways, was to increase the enhancement of products in the system.

### 2.1.2 Two Replicators In One System – OR Gate

The template directed nature of autocatalytic self-replicating reactions makes their use ideal for the creation of instructable synthetic networks, that is a network of interconnected reactions whose outcome can be directed to form a particular product as a result of a particular instruction. In the initial stages of reaction of an autocatalytic self-replicator, as there is no template present to catalyse the reaction, the product will only be formed by the bimolecular pathway. This means that in the initial stages the reaction is slow. As time elapses and once sufficient template has been formed the autocatalytic cycle initiates and accelerates the reaction rate. The reaction therefore has two phases, the slow initial phase and the faster autocatalytic phase. A way therefore to 'turn on' the autocatalytic cycle earlier is to dope preformed template into the reaction at the start. This will enhance the product formed through this autocatalytic cycle above any others present in the system.

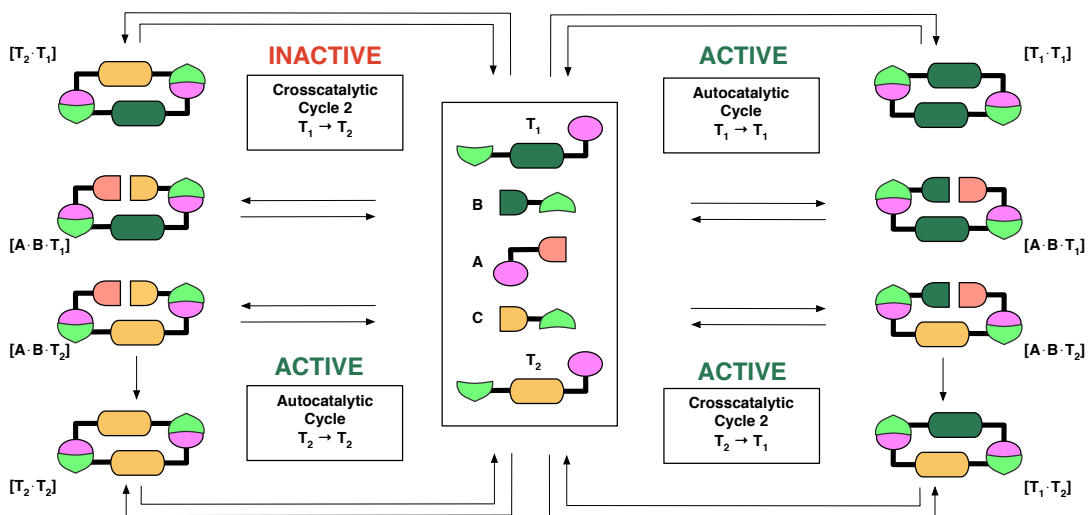
Recently we reported<sup>187</sup> a small system of two interdependent structurally similar replicators (**T**<sub>1</sub> and **T**<sub>2</sub> Scheme 2.1) which when combined together lead to a system which displayed defined system level behaviour (Figure 2.2). The two reported self-replicating templates differed by only one CH<sub>2</sub> group which lead us to expect that they would interact with each other with one acting as a template for the other and *vice versa*. In isolation, we demonstrated that **T**<sub>2</sub> acted as a cross-catalyst for the formation of **T**<sub>1</sub>. However, as a consequence of the shorter length of the template, **T**<sub>1</sub> acted as an inhibitor for the formation of **T**<sub>2</sub>.

We therefore considered that in an equimolar mixture of **A**, **B** and **C**, three of the four possible catalytic cycles would be active (Figure 2.2). The two autocatalytic cycles in which each template produces itself and one cross-catalytic cycle in which **T**<sub>2</sub> catalyses the formation of **T**<sub>1</sub>. This results in the network topology revealing that there exist more catalytic cycles producing **T**<sub>1</sub> than **T**<sub>2</sub>.



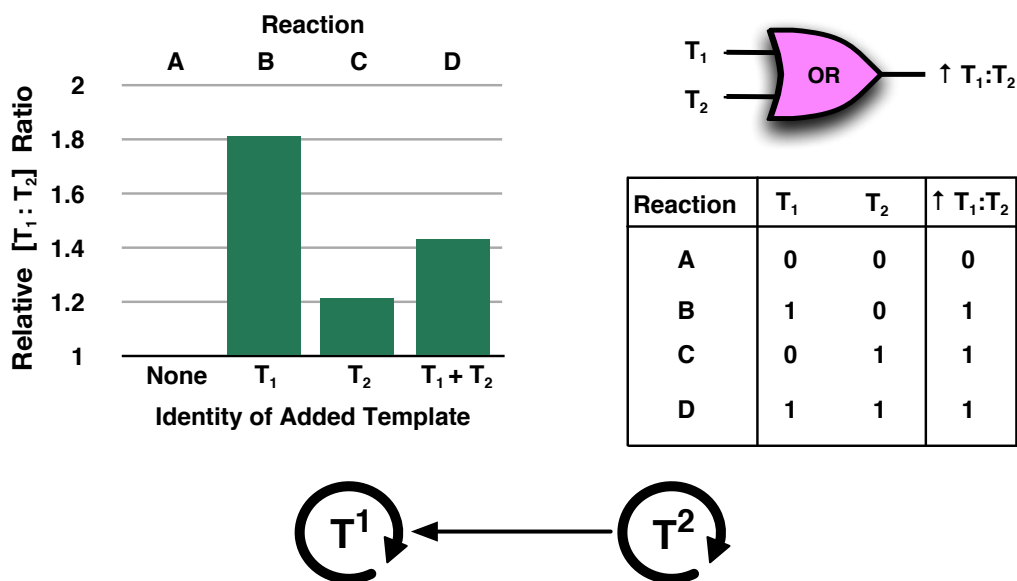
**Scheme 2.1:** Reagents and conditions (i) CDCl<sub>3</sub>, 10 mM, 10 °C, 16 h. Reaction of the three compounds in one pot leads to four possible cycloadduct products, two *trans* and two *cis*. Cartoon representation of each component with unique colour given to aid visualisation of network in following figure.

In the absence of any instruction to an equimolar mixture of **A**, **B**, and **C**, the two replicators coexisted as a result of their comparable efficiency (Figure 2.3 reaction A). When **T<sub>1</sub>** was added to a fresh set of reagents for the reaction, the concentration of newly formed **T<sub>1</sub>** was amplified and the concentration of **T<sub>2</sub>** suppressed as **T<sub>1</sub>** is a catalyst for its own formation and an inhibitor for **T<sub>2</sub>** formation leading to a greater [T<sub>1</sub>]:[T<sub>2</sub>] ratio (Figure 2.3 reaction B). When **T<sub>2</sub>** was added to a fresh set of reagents, **T<sub>1</sub>** was amplified through the cross-catalytic pathway and the more **T<sub>1</sub>** was formed, the more its own autocatalytic pathway dominated the system and inhibited the formation of **T<sub>2</sub>**, once again leading to an increased [T<sub>1</sub>]:[T<sub>2</sub>] ratio (Figure 2.3 reaction



**Figure 2.2:** With structurally similar templates, four possible autocatalytic or cross-catalytic pathways could exist. Three of which were determined to be active and one inactive.

C). When both templates were added,  $T_1$  was amplified as a result of specific autocatalysis and inhibition of  $T_2$  formation again increasing the  $[T_1]:[T_2]$  ratio (Figure 2.3 reaction D).



**Figure 2.3:** Results of doping the system with template(s) relative to the native untemplated results. The ratio of  $[T_1]:[T_2]$  is increased when the identity of the input template is  $T_1$ ,  $T_2$  or Both representing the truth table for the boolean OR operation. Below is the graph architecture of the network. Data taken from reference 187.

If system inputs are the identities that of the added templates and the system output is an increase in  $[T_1]:[T_2]$  ratio (Figure 2.3) then the results can be displayed as a truth table with a 1 representing the presence of an input or output and a 0 the absence of an input or output. The system therefore displayed a system level behaviour analogous to the Boolean Logic OR operation.

## 2.2 Aims of Project

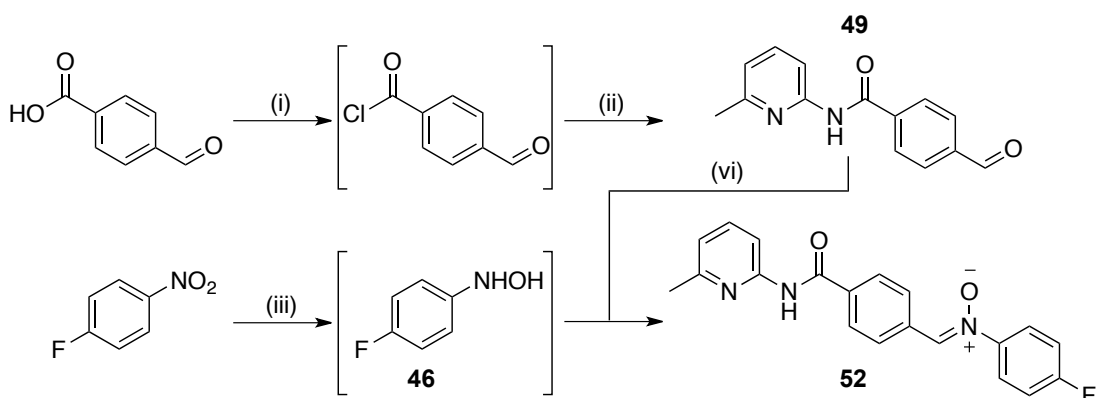
The OR gate system represented a successful demonstration of a small network of multiple organic replicators which coexist and could be instructed by adding a preformed template. In this case however, the level of control in the system is limited. Two different instructions can be given in the form of the two templates however the output is always the same, namely the enhancement of the one template.

Nature on the other hand has evolved to embrace complexity. The cell is host to an array of machinery which operates in harmony through signalling pathways and feedback loops. The result is that the cell is able to recognise a diverse series of environmental changes, such as ion concentration, pH or partial pressure of oxygen, which act as instructions to the cell's machinery which then act accordingly in response. In line with our long term goals, we wish to be able to reproduce the cell's ability to interpret different instructions leading to different responses.

If we are to take inspiration from nature, we must look to create a system of replicators with a greater level of control than our OR gate provided. In order to achieve this goal, we wish to be able to create a network of multiple coexisting replicators which can be addressed by the addition of different preformed templates as before, only this time we would like the system to produce a different output which is dependent on the nature of the input template.

For our OR gate system, the family of replicators used did not allow the degree of control we wished to achieve. In order for our new system to be successful, we need to identify a different family of interdependent replicators. With this in mind, our attention turned to altering the components of the system and searches began for a pair of replicators which shared a common component. Investigation of the reactivity of the C1 maleimide **29** with a plethora of recognition bearing nitrones identified two systems previously reported<sup>136</sup> which behaved as autocatalytic self-replicators of similar efficiencies.

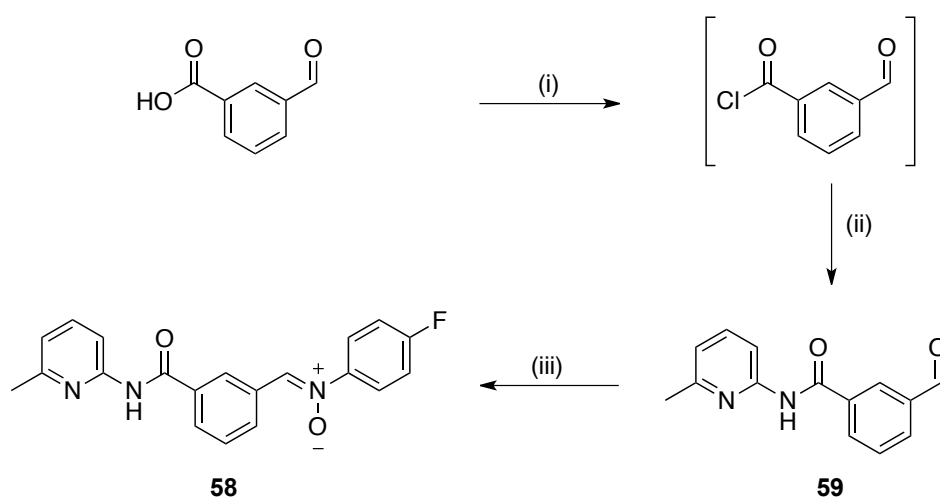
### 2.2.1 Synthesis of *para*-Nitrone 52



**Scheme 2.2:** Reagents and conditions (i) thionyl chloride, toluene, 110 °C, 16 h, quantitative (ii) 2-amino-6-methylpyridine, NEt<sub>3</sub>, CH<sub>2</sub>Cl<sub>2</sub>, 0 °C → RT, 16 h, 72% (iii) Rh/C, NH<sub>2</sub>NH<sub>2</sub>·H<sub>2</sub>O, THF, RT, 0.5 h, quantitative (iv) EtOH, 16 h, 64%.

*para*-Nitron 52 was synthesised in a convergent manner starting with the activation of 4-formylbenzoic acid as an acid chloride before amide coupling with 2-amino-6-methylpyridine to form aldehyde 49. 4-Fluoronitrobenzene was reduced to the hydroxylamine 46 by transfer hydrogenation<sup>188</sup> using 5% rhodium on carbon with hydrazine hydrate and the hydroxylamine condensed with aldehyde 49 to afford the target nitron 52 in satisfactory yields.

### 2.2.2 Synthesis of *meta*-Nitron 58

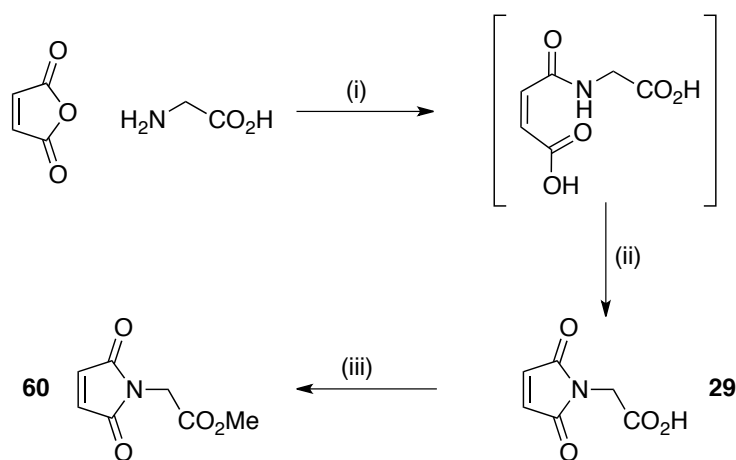


**Scheme 2.3:** Reagents and conditions (i) thionyl chloride, toluene, 110 °C, 16 h, quantitative (ii) 2-amino-6-methylpyridine, NEt<sub>3</sub>, CH<sub>2</sub>Cl<sub>2</sub>, 0 °C → RT, 16 h, 67% (iii) hydroxylamine 46, EtOH, 16 h, 35%.



3-carboxybenzaldehyde was activated by treating with thionyl chloride in toluene to form the acid chloride which was used immediately in an amide coupling reaction with 2-amino-6-methylpyridine in triethylamine and dichloromethane to form aldehyde **59**. With the recognition site installed on the aldehyde, condensation with *N*-(4-fluorophenyl) hydroxylamine **46** in ethanol furnished *meta*-nitron **58** in acceptable yields.

### 2.2.3 Synthesis of Maleimides **29** and **60**



**Scheme 2.4:** Reagents and conditions (i) THF, RT, 16 h, quantitative (ii) HMDS,  $\text{ZnBr}_2$ , MeCN,  $\Delta$ , 2 h 95% (iii) *N,N'*-Dicyclohexylcarbodiimide, MeOH, THF,  $0^\circ\text{C} \rightarrow \text{RT}$ , 40 h, 80%.

Maleic anhydride was first ring opened by glycine before cyclisation with HMDS and zinc bromide furnished the C1 Maleimide **29** in good yields. The recognition disabled control maleimide **60** was obtained in good yields after activating the acid group to ester formation with carbodiimide and then reacting with methanol.

## 2.3 Characterisation of Self-Replicating Systems

The addition of the *para*-fluorine to the self-replicating structures may potentially alter the electronic properties of the nitrones and consequently their reactivity towards C1 maleimide **29**. We also wished to change the solvent from  $\text{CDCl}_3$  to  $\text{CD}_2\text{Cl}_2$  saturated with *para*-toluenesulfonic acid in order to control the water and acid content in the system. In order to confirm each structure still behaved as self-replicators under these conditions, each were characterised again using the standard protocol.

### 2.3.1 Analysis of the reaction of Maleimide 29 with *meta*-Nitron 58

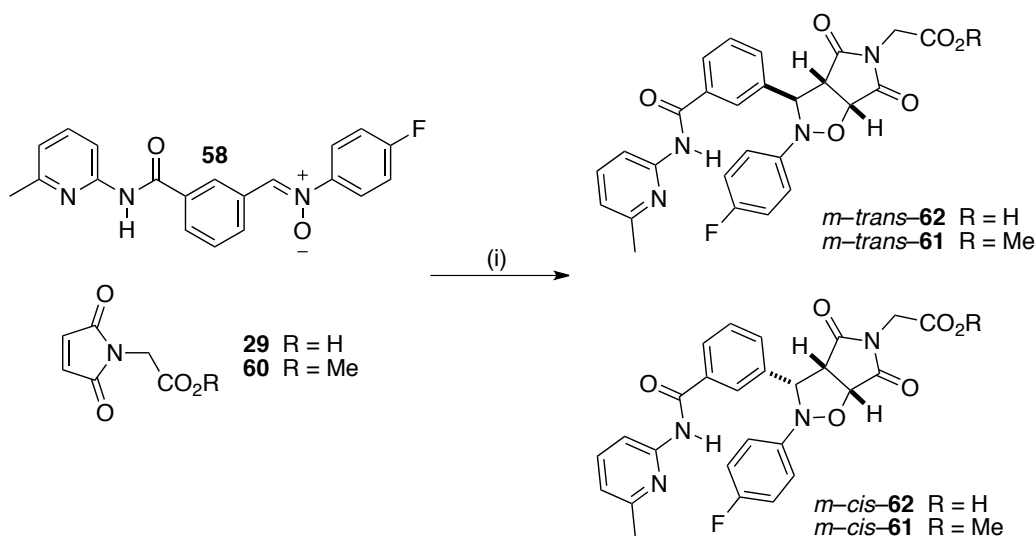
Initially, to establish a background rate of bimolecular reaction in the recognition disabled reaction, maleimide **60**, with the acid group inhibited by a methyl ester, was reacted with *meta*-nitron **58** at 20 mM in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat) at 10 °C and the reaction monitored by <sup>19</sup>F NMR spectroscopy for 16 h, Section 7.6. The concentration vs. time profile for the reaction (Figure 2.4 (i)) shows the linear growth of the two diastereoisomeric products *trans*-**61** and *cis*-**61**. The overall conversion to the major *trans*-**61** cycloadduct was low reaching 14% after 16 h and with a *trans* : *cis* ratio of 2.5 : 1. The reaction profile was subjected to simulation and fitting to a model for a bimolecular reaction using SimFit, see appendix A1. The reaction constants to form *trans*-**61** and *cis*-**61** were found to be  $k_{\text{trans}} = 1.94 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  and  $k_{\text{cis}} = 0.85 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  respectively.

The effect of recognition processes on the reaction was then established by repeating the experiment with the recognition enabled maleimide **29** under the same conditions. The concentration vs. time profile (Figure 2.4 (ii)) now displays that the production of *trans*-**62** is enhanced, reaching 28% conversion after 16 h, double that of the recognition inhibited reaction, and displaying a sigmoidal growth profile. This increase in reaction rate with time is characteristic for an autocatalytic self replicating reaction. The *trans* : *cis* ratio had also increased to 3.1 : 1 highlighting the increase in selectivity for *trans*-**62**, the replicating template. Interestingly the concentration of *cis*-**62** after 16 h was also enhanced, reaching 1.9 mM, in comparison to the bimolecular reaction which reached 1.2 mM. This enhancement can be attributed to the creation an [A•B] complex which promotes the formation of *cis*-**62**, the efficiency of this complex must not be very high as it produces the minor product and is surpassed by the autocatalytic cycle. The results were subjected to kinetic simulation and fitting to a model for a self-replicating reaction using SimFit, see appendix A2, to extract rate information. The rate of reaction within the ternary complex to form *trans*-**62** was determined to be  $1.5 \times 10^{-4} \text{ s}^{-1}$  which reveals the *EM* of the system to be 0.8 M. The rate of reaction in the [A•B] complex to form *cis*-**62** was calculated to be  $6.67 \times 10^{-6} \text{ s}^{-1}$ , significantly slower than the ternary reaction which confirms why the [A•B] complex channel does not dominate the reaction. The *EM* of the [A•B] process was calculated to be 0.008 M, significantly lower than that of the ternary complex.

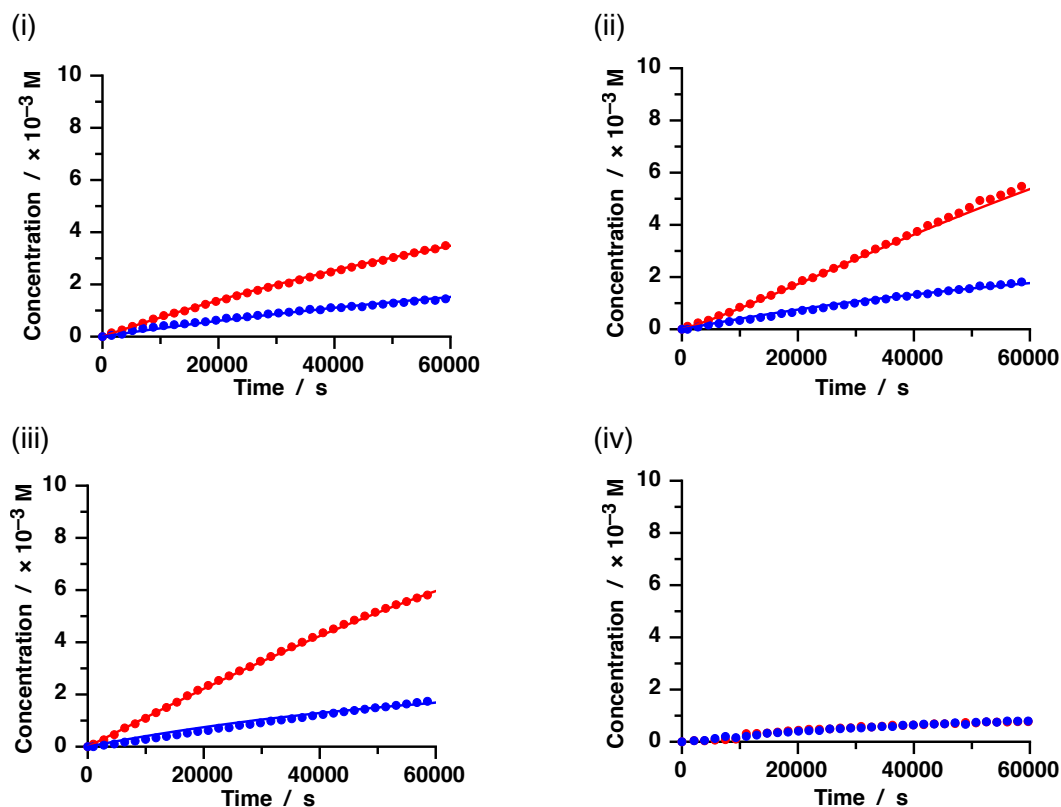
The template directed manner of the reaction was demonstrated by addition of preformed *trans*-**62** to a fresh set of reagents (Figure 2.4 (iii)). With preformed template present upon initiation of the reaction, there will be no lag period and instead the reaction will proceed at the maximum autocatalytic rate from the start. The concentration vs. time profile for this reaction confirmed our expectation and showed that the lag period observed in the native reaction was absent. As the reaction to produce *trans*-**62** proceeds more efficiently from the start, we also observed inhibition of the *cis*-**62** formation with the *trans* : *cis* ratio now at 4:1.

The final reaction to demonstrate the importance of recognition in the system was performed by introducing a competitive inhibitor, benzoic acid (Figure 2.4 (iv)). In this system, however, the increased acid and water content led to degradation of the *meta*-nitron **58** and subsequently the concentration vs. time profile for this reaction showed that production of both *trans*-**62** and *cis*-**62** has been inhibited as a result.

The characteristics of the reaction between maleimide **29** and *meta*-nitron **58** highlight that the system still exhibits self-replicating behaviour after addition of a *para*-fluoride and changing the solvent from CDCl<sub>3</sub> to CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat) although the efficiency of the replicating process has been decreased as the *EM* in this system is now 0.8, by comparison<sup>136</sup> the *EM* of the unfluorinated system in CDCl<sub>3</sub> was determined to be 2.5. This decrease in *EM* could be as a result of a change in reactivity of the fluorinated nitron or increase in acid content in the solvent reducing the efficiency of recognition mediated processes.



**Scheme 2.5:** Reagents and conditions (i) CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat), 10 °C, 20 mM, 16 h



**Figure 2.4:** Concentration vs. time profiles for: (i) Control reaction of nitrene **58** with maleimide ester **60** (ii) Recognition mediated reaction of nitrene **58** with maleimide acid **29** (iii) Template doped reaction of nitrene **58** with maleimide **29** with 10 mol% *trans*-**62** (iv) Competitive inhibitor reaction of nitrene **58** with maleimide **29** and 5 eq. benzoic acid. Experimental data is represented as circles, *trans* cycloadducts (●), *cis* cycloadducts (●). The results of kinetic simulation and fitting using a model for a self-replicating system is shown as lines *trans* cycloadducts (—), *cis* cycloadducts (—). Reaction conditions  $\text{CD}_2\text{Cl}_2/p\text{-TSA}$  20 mM, 10 °C, 16 h. Note axis are expanded to 10 mM for clarity.

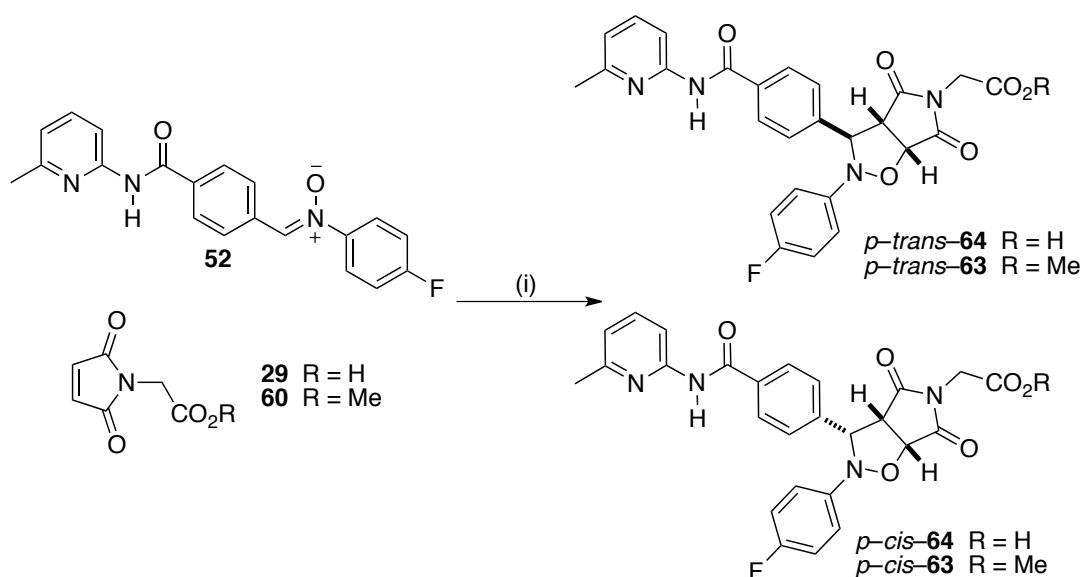
**Table 2.1:** Rate constants determined by kinetic simulation and fitting of experimental data using SimFit.

	$k_{\text{trans}}$	$k_{\text{cis}}$
Bimolecular Rate Constant / $\times 10^{-4} \text{ M}^{-1} \text{ s}^{-1}$	1.94	0.85
$k_{\text{ternary}} / \times 10^{-4} \text{ s}^{-1}$	1.5	—
$EM$ Ternary / M	0.8	—
$K_{\text{Duplex}} / \times 10^5 \text{ M}^{-1}$	7.05	—
Connection $EM / M$ ( $\Delta G^\ddagger / \text{kJmol}^{-1}$ )	1.5 (0.93)	—

### 2.3.2 Analysis of the reaction of Maleimide **29** with *para*-Nitrene **52**

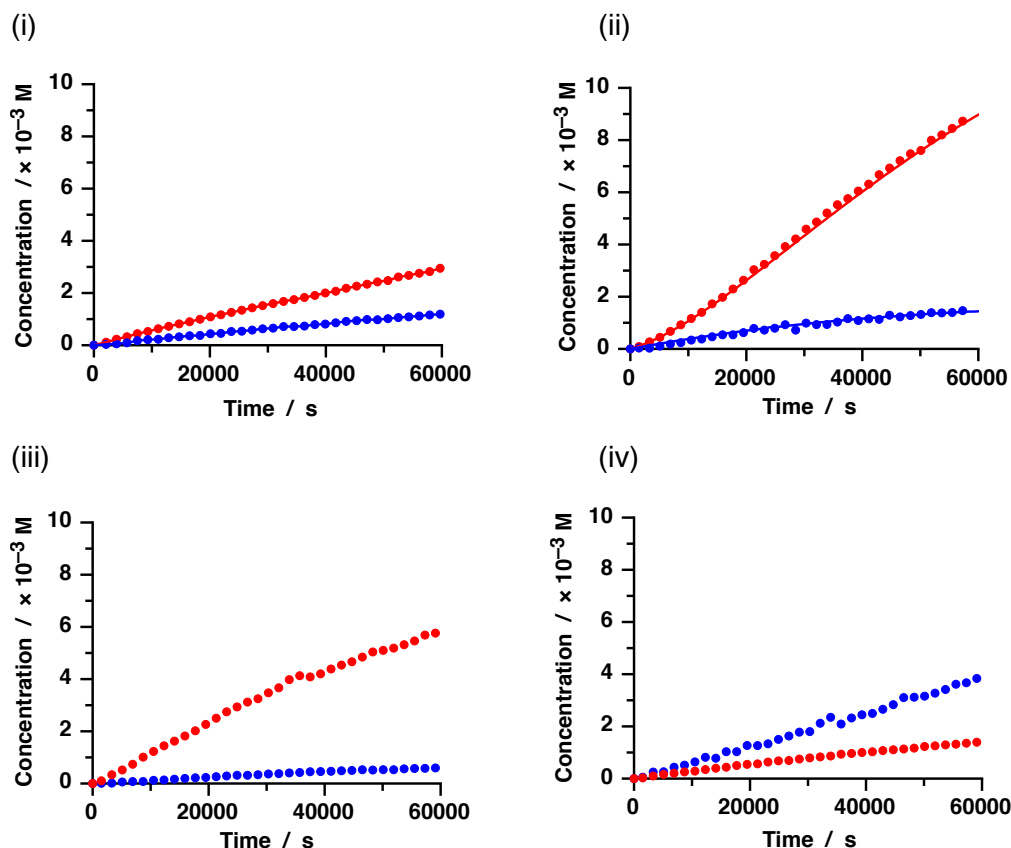
The protocol for characterising and quantifying an autocatalytic self-replicator used in the previous section was also applied to the reaction between *para*-nitrene **52** and maleimide **29**.

The concentration vs. time profile of the bimolecular reaction of recognition disabled maleimide **60** and *para*-nitrone **52**, (Figure 2.5, (i)) showed linear growth of the two diastereoisomeric products *trans*-**63** and *cis*-**63**. The overall conversion to the major *trans*-**63** cycloadduct was once again low reaching 14% after 16 h with a *trans* : *cis* ratio of 2.5 : 1. The reaction profile was subjected to simulation and fitting to a model for a bimolecular reaction using SimFit, appendix A1, and the reaction constants to form *trans*-**63** and *cis*-**63** were determined to be  $k_{\text{trans}} = 1.49 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  and  $k_{\text{cis}} = 0.61 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  respectively.



**Scheme 2.6:** Reagents and conditions (i)  $\text{CD}_2\text{Cl}_2/p\text{-TSA}$  (sat), 20 mM, 10 °C, 16 h.

The recognition mediated reaction was then performed by repeating the experiment with the recognition enabled maleimide **29** under the same conditions. The concentration vs. time profile (Figure 2.5, (ii)) for this reaction now showed non-linear growth with a sigmoidal profile for the production of *trans*-**64**, characteristic for an autocatalytic self replicating reaction. The overall conversion to *trans*-**64** reached 44%, a significant increase in comparison to the recognition disabled system. The *trans* : *cis* ratio was now 6:1 highlighting the selectivity for *trans*-**64**. The production of *cis*-**64** was also enhanced slightly reaching 1.5 mM, an increase from 1.2 mM in the recognition disabled control. This highlights that once again recognition in the system can lead to the creation of an  $[\mathbf{A}\cdot\mathbf{B}]$  complex which promotes the formation of *cis*-**64**. Reaction rate constants were once again obtained from kinetic simulation and fitting to a model for a self-replicating reaction using SimFit, appendix A2. The rate of reaction within the ternary complex to form *trans*-**64** was calculated to be  $3.88 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  which reveals the *EM* of the system to be 2.6 M.



**Figure 2.5:** Concentration vs. time profiles for: (i) Control reaction of nitrene **52** with maleimide ester **60** (ii) Recognition mediated reaction of nitrene **52** with maleimide acid **29** (iii) Template doped reaction of nitrene **52** with maleimide **29** with 10 mol% *trans*-**64** (iv) Competitive inhibitor reaction of nitrene **52** with maleimide **29** and 5 eq. benzoic acid. Experimental data is represented as circles, *trans* cycloadducts (●), *cis* cycloadducts (●). The results of kinetic simulation and fitting using a model for a self-replicating system is shown as lines *trans* cycloadducts (—), *cis* cycloadducts (—). Reaction conditions  $\text{CD}_2\text{Cl}_2/p\text{-TSA}$  20 mM, 10 °C, 16 h. Note axis expanded to 10 mM for clarity.

**Table 2.2:** Rate constants determined by kinetic simulation and fitting of experimental data using SimFit.

	$k_{\text{trans}}$	$k_{\text{cis}}$
Bimolecular Rate Constant / $\times 10^{-4} \text{ M}^{-1} \text{ s}^{-1}$	1.49	0.61
$k_{\text{ternary}} / \times 10^{-4} \text{ s}^{-1}$	3.88	—
$EM_{\text{Ternary}} / \text{M}$	2.6	—
$K_{\text{Duplex}} / \times 10^5 \text{ M}^{-1}$	6.00	—
Connection $EM / \text{M}$ ( $\Delta G^{\ddagger} / \text{kJmol}^{-1}$ )	1.29 (0.6)	—

The template directed manner of the reaction was demonstrated by addition of preformed *trans*-**64** to a fresh set of reagents. The concentration vs. time profile for this reaction (Figure 2.5 (iii)) showed that the lag period observed in the native reaction was absent and the reaction is able to proceed at the maximum autocatalytic

rate from the start of the reaction. After 16 h the *trans* : *cis* ratio had increased to 13:1 highlighting the selectivity for *trans*-**64**.

Finally the competitive inhibitor benzoic acid was added to a fresh set of reagents to demonstrate the importance of recognition. The concentration vs. time profile (Figure 2.5, (iv)) for this reaction showed that production of *trans*-**64** has been inhibited with the rate of reaction proceeding close to the recognition inhibited control owing to benzoic acid inhibiting recognition events within the system.

The characteristics of the reaction between maleimide **29** and *p*-nitro **52** highlight that the system is behaving as an autocatalytic self-replicator selective for the formation of *trans*-**64**. The *EM* of the fluorinated system was calculated to be 2.6 M which is once again a decrease by comparison to the previously reported unfluorinated system which had an *EM* of 3.5 M. This result shows that these changes to the system have also had the effect of slightly decreasing the reactivity of the ternary complex.

### 2.3.3 Isolated Replicator Conclusions

Comparison of the reactivity of *p*-nitro **52** and *m*-nitro **58** towards C1 maleimide **29** shows that both behave as autocatalytic self-replicators in isolation. As a result of both reactions proceeding under identical conditions, they are ideal candidates for combining into one system in which they will compete for common material. As the *p*-nitro system is more efficient in isolation, with a maximum autocatalytic rate 2.5 times faster than the *m*-nitro system, we might expect to see this difference in efficiency transposed into the product pool of a two replicator system with the more efficient replicator the most abundant product in the pool.

### 2.3.4 Cross-Catalysis Investigations

Key to the success of a fully controllable system is the principle that each template will be a specific autocatalyst, that is it will template its own formation and not participate in cross-catalysis. Strong cross-catalysis could produce a system topology similar to that of the OR gate system which would lead to limited control.

With this in mind we performed two further reactions designed to investigate for cross-catalysis in a manner similar to the demonstration of template effects.

Firstly, 20 mol% *m-trans*-**62** was doped into a fresh set of reagents for the *p-trans*-**64** system, (**52** + **29**) in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat) at 20 mM and 10 °C and the reaction monitored by <sup>19</sup>F NMR spectroscopy. The results were extracted as described previously and compared to the native reaction between *p*-nitron **52** and maleimide **29** under the same conditions. In this case, there was no enhancement seen by doping the *p-trans*-**64** system with preformed *m-trans*-**62** in fact there was a small degree of inhibition. This inhibition is a result of the template holding the reactants in an unreactive conformation.

Investigation of the reverse situation, when *p-trans*-**64** was doped into a fresh set of reagents for the *m-trans*-**62** system (**58** + **29**) and compared to an un-doped reaction under the same conditions, highlighted that *p-trans*-**64** was able to enhance the production of *m-trans*-**62** however the enhancement was not as strong as when the same concentration of *m-trans*-**62** was added.

From this result we could conclude that *m-trans*-**62** was a selfish autocatalyst and *p-trans*-**64** was both an autocatalyst and a cross-catalyst however was more efficient as an autocatalyst. These results encouraged us to continue and investigate the properties of combining the two into one system.

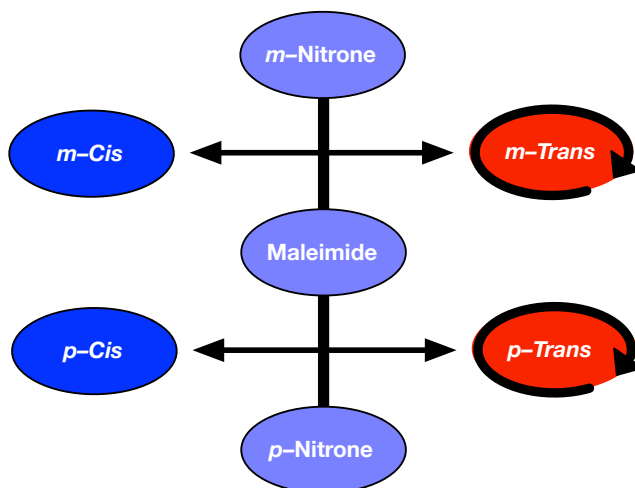
## 2.4 Competitive Two Replicator System

With two independent self-replicators which utilise a common reagent identified, we now possessed the requirements for the creation of a competitive system of two interlinked replicators and set about to demonstrate its operation. In a system consisting of an equimolar concentration of *m*-nitron **58**, *p*-nitron **52** and maleimide **29**, each of the two nitrones will be in competition to consume the maleimide and produce their self-replicating templates (Figure 2.6).

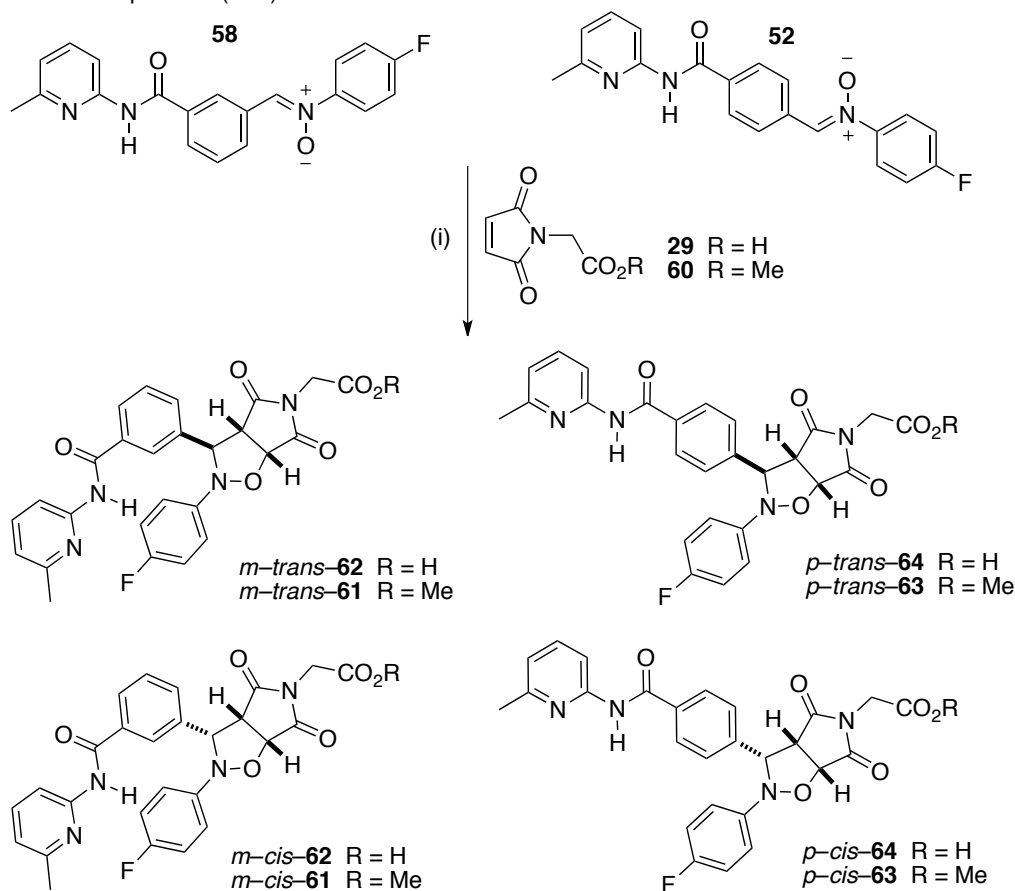
The system can be visualised in a graphical manner (as in Figure 2.6) with the reagents represented as green circles, connections as black lines, reactions as arrows, *trans* products as red circles and *cis* products as blue circles. A common



metabolite (maleimide **29**) may react with (black lines) consumed simultaneously by two pathways (*m*-nitron **58** or *p*-nitron **52**) to form the four cycloadducts. Each of the *trans*-cycloadducts are autocatalytic self-replicators and will catalyse their own formation.



**Figure 2.6:** Graphical representation of the two replicator system. Each of the nitrones can react with a common maleimide to form two autocatalytic products (red) and two inactive waste products (blue).

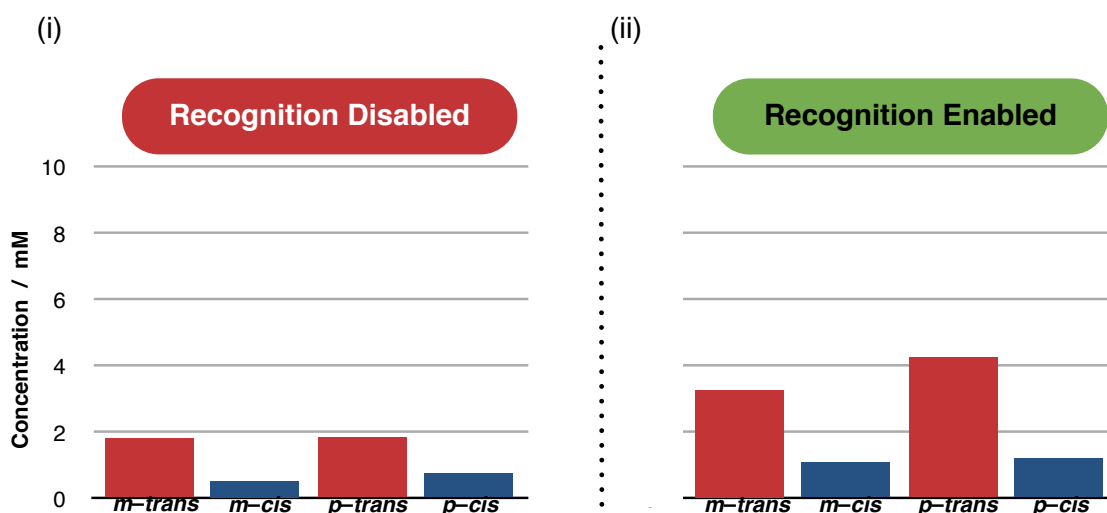


**Scheme 2.7:** The combination of the two nitrones in one system with a maleimide leads to the formation of four products, the two *trans* cycloadducts and two *cis* cycloadducts. Reagents and conditions (i)  $\text{CD}_2\text{Cl}_2/p\text{-TSA}$  (sat), 10 mM, 25 °C, 16 h.

Initially, to demonstrate that recognition mediated processes are still prevalent in the two replicator system, a comparison between a system with recognition mediated reactions and a system in which recognition is inhibited was performed.

The reaction of the recognition inhibited system was investigated by monitoring the reaction of an equimolar mixture of *p*-nitroene **52**, *m*-nitroene **58** with maleimide methyl ester **60** at a concentration of 10 mM in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat) at 25 °C for 16 h, see Chapter 7, Section 7.2 for experimental details (Scheme 2.7). After 16 h, the system was assayed by 376.5 MHz <sup>19</sup>F NMR spectroscopy and the relative concentration of products calculated by deconvolution of the appropriate resonances and revealed that 49% of the maleimide had been consumed by cycloaddition reactions. The two *trans* cycloadducts were produced with near equal concentrations with 1.81 mM *m-trans* **61** and 1.84 mM *p-trans* **63** produced which was expected with the similar rates of bimolecular reactions (Figure 2.7 (i)). Each of their *cis* diastereoisomers were also produced in lower quantities of 0.52 mM and 0.74 mM respectively representing a *trans* : *cis* ratio of 3.4 : 1 for the *meta* cycloadducts and a 2.6 : 1 for the *para* cycloadducts.

We next performed the recognition enabled reaction, achieved by repeating the protocol for the recognition disabled system except substituting the recognition inhibiting ester maleimide **60** and replacing it with the recognition enabled acid maleimide **29**.



**Figure 2.7:** Concentrations of the four cycloadducts in (i) the recognition disabled system with maleimide ester **60** and (ii) the recognition enabled system with acid maleimide **29** after 16 h in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat) at 25 °C and determined by deconvolution of 376.5 MHz <sup>19</sup>F NMR spectrum.

Assaying the system after 16 h in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat) at 25 °C with 376.5 MHz <sup>19</sup>F NMR spectroscopy revealed that the maleimide had been entirely consumed (Figure 2.7 (ii)). Therefore the first observation from the recognition mediated system is the higher reactivity as the recognition disabled system reached only 49% conversion. The two *trans*-cycloadducts are once again the dominant products in the system with closer inspection revealing that there was a greater concentration of *p-trans* **64**, which reached a concentration of 4.24 mM, in the product pool than *m-trans* **62** which reached a concentration of 3.27 mM. This observation was interesting as it highlights that the more efficient replicator in isolation still behaves more efficiently in the system, although, more importantly, the two replicators are able to co-exist in the system as one does not completely dominate the product pool.

Curiously the concentration of each of the two *cis*-cycloadducts are greater in the recognition enabled reaction, where 1.09 mM of *m-cis* **62** and 1.19 mM of *p-cis* **64** were produced, in comparison to the recognition disabled control where 0.52 mM *m-cis* **62** and 0.75 mM *p-cis* **64** were produced. This observation revealed that recognition mediated processes are enhancing the formation of the *cis* cycloadducts. When the replicators were investigated in isolation, in both cases there was an enhancement in the production in the *cis* cycloadduct which was produced through the formation of a reactive [A•B] complex. Obviously these complexes are still able to form in the two replicator system and as a result enhance the formation of the *cis* cycloadducts. Further, the increased number of components in the system will also serve to reduce the efficiency of the replicators. When ternary complexes are formed from the cross-combinations of *m*-nitronone **58** with the *p-trans* **64** template or the *p*-nitronone **52** with the *m-trans* **62** template, these ternary complexes will be inactive or less reactive and thus act as inhibitors in the system. This will lower the efficiency of ternary complexes in the system and, as a result, the reaction through binary [A•B] complexes, which are not affected by the increased complexity, will become more prevalent leading to the observed higher *cis*-cycloadduct concentrations.

#### 2.4.1 Template Instruction of System

As the goal of this project is the creation of a controllable system, the ability of the two replicator network to respond to an instruction was investigated. As there are two

replicating templates in the system, each of these can be used to address the system.

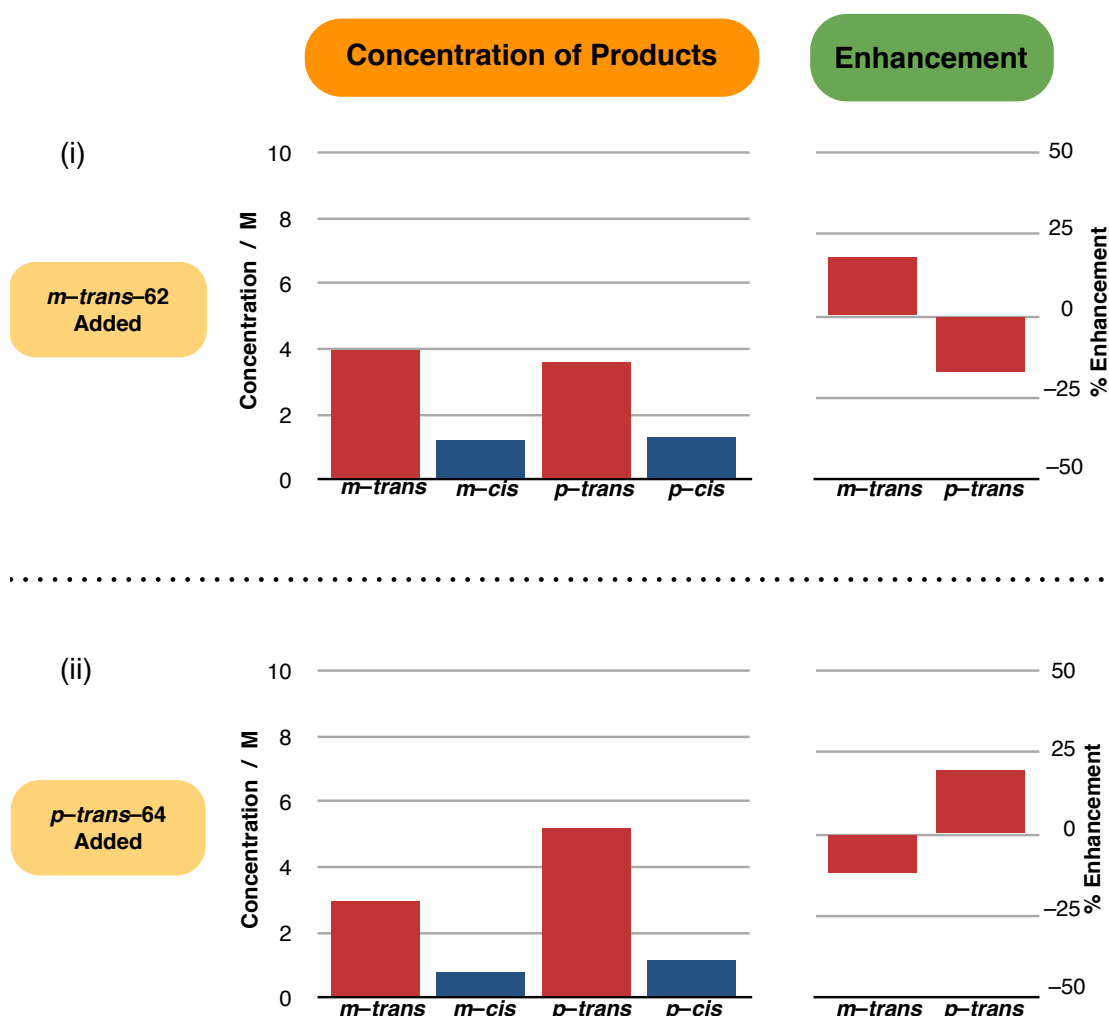
When a fresh set of reagents for the two replicator system is doped with a preformed sample of one of the two replicating *trans* templates. These templates have the ability to direct the formation of themselves through the autocatalytic channel upon initiation of the reaction. The effect of adding preformed template will therefore be that the added template will promote the formation of itself, consuming the nitron it requires along with the common metabolite (maleimide **29**) at an increased rate relative to an undoped system. As the maleimide is consumed more rapidly, less will be available to react with the second nitron in solution to form the other replicating template. The result of the doping will be that the concentration of the template which was added at the start of the reaction will be enhanced (after subtracting added template so that only newly formed template is compared) relative to the undoped system and the other replicating template will be inhibited as a consequence of the consumption of maleimide more swiftly through the other pathway. As the *cis* cycloadducts are not formed by template directed synthesis, these could not be enhanced selectively by any instruction.

In order to test this expectation, we first wished to direct the system to enhance the production of *m-trans* **62**. The instruction for the system to do this was issued by doping 20 mol% of preformed *m-trans* **62** into a fresh set of reagents for the two replicator system. For experimental details see Chapter 7, Section 7.2. The reaction was then repeated under identical conditions as with the undoped system and after 16 h, the system was assayed by 376.5 MHz <sup>19</sup>F NMR spectroscopy and the relative concentration of each product determined by deconvolution. From these results, the percentage enhancements of the *trans* cycloadducts in comparison to the uninstructed system can be calculated according to Equation 2.1 and the result is presented to the right of the concentration of products graph.

$$\% \text{ Enhancement} = \frac{[\text{instructed result} - \text{uninstructed result}]}{[\text{uninstructed result}]} \times 100\%$$

**Equation 2.1:** The percentage enhancement of each of the products in the instructed system relative to the uninstructed system was calculated by subtracting the concentration of product formed in the uninstructed system from the concentration of product formed in the instructed system, dividing the difference by the concentration of the product in the uninstructed system and multiplying by 100%.

After 16 h, the maleimide has once again been entirely consumed and the distribution of products (Figure 2.8 (i)) has altered reflecting the instruction given to the system. The major product is now *m-trans*-**62** of which 3.94 mM had been produced. The second replicating template, *p-trans*-**64**, was the second most abundant product in the solution with 3.58 mM produced. The two *cis* cycloadducts had formed in similar concentrations to the uninstructed systems. From the results we can determine the percentage enhancements in comparison to the uninstructed system to give a clearer indication of the effect of instructing the system. Displayed on the right of the product concentration profile, the enhancement results show that the production of *m-trans*-**62** has been enhanced by 18% at the expense of *p-trans*-**64** which has been inhibited by 17%.



**Figure 2.8:** Concentration of newly formed cycloadducts after 16 h after instruction with: (i) 20 mol% *m-trans* **62** and (ii) 20 mol% *p-trans* **64**. Concentrations of products were determined by deconvolution of  $^{19}\text{F}$  NMR spectra and subtraction of pre-added template. Enhancements were calculated by  $([\text{Instructed}] - [\text{Native}] / [\text{Native}] \times 100\%$ .

This result clearly follows our expectation that addition of the *m-trans-62* template would alter the reactivity in the system, catalysing the formation of itself and inhibiting the production of the other replicating system demonstrating that the system has successfully responded to the instruction of doping with preformed *m-trans-62*.

We now wished to investigate if by changing the nature of the instruction, by adding preformed *p-trans-64*, that we can successfully direct the system to enhance the production of the other replicating template and thus, demonstrate a fully controllable system of two replicators. The instruction was once again issued by addition of preformed template, in this case 20 mol% *p-trans-64*, to a fresh set of reagents. This mixture was incubated and assayed under identical conditions as before and the concentration profile of the results displayed in Figure 2.8 (ii). The results from this instruction firstly show that the distribution of products has once again changed as a result of addition of preformed *p-trans-64*. The major product is now newly formed *p-trans-64* which is now the dominant product in the mixture with 5.17 mM produced, almost double the concentration of the other replicating template *m-trans-62* of which 2.94 mM had been produced. Comparison of these results to the uninstructed system revealed a 19% enhancement of in the production of *p-trans-64* with a concurrent 12% inhibition of the production of *m-trans-62*. Clearly adding preformed *p-trans-64* has altered the outcome of the product distribution in the system, satisfying our desire to enhance *p-trans-64*, in comparison to the undoped system. Most significantly, the results demonstrate that the outcome is different to the results when doped with *m-trans-62*.

It is clear from these results that this network of two replicators is able to be instructed in two different ways, by addition of either of the two replicating templates, and gives rise to two different outcomes with the nature of the outcome dependent on the instruction given. The ability to select between two different outcomes using two different instructions means we have a greater degree of control in comparison to the OR gate system described earlier in which the two instructions gave the same output.

The reactions performed in the controllable two replicator system do however have an advantage over processes in nature. Upon initiation of the reaction all three reagents for the system are present and when the instruction is issued the one step cycloaddition reaction proceeds immediately. In essence there is a kinetic race

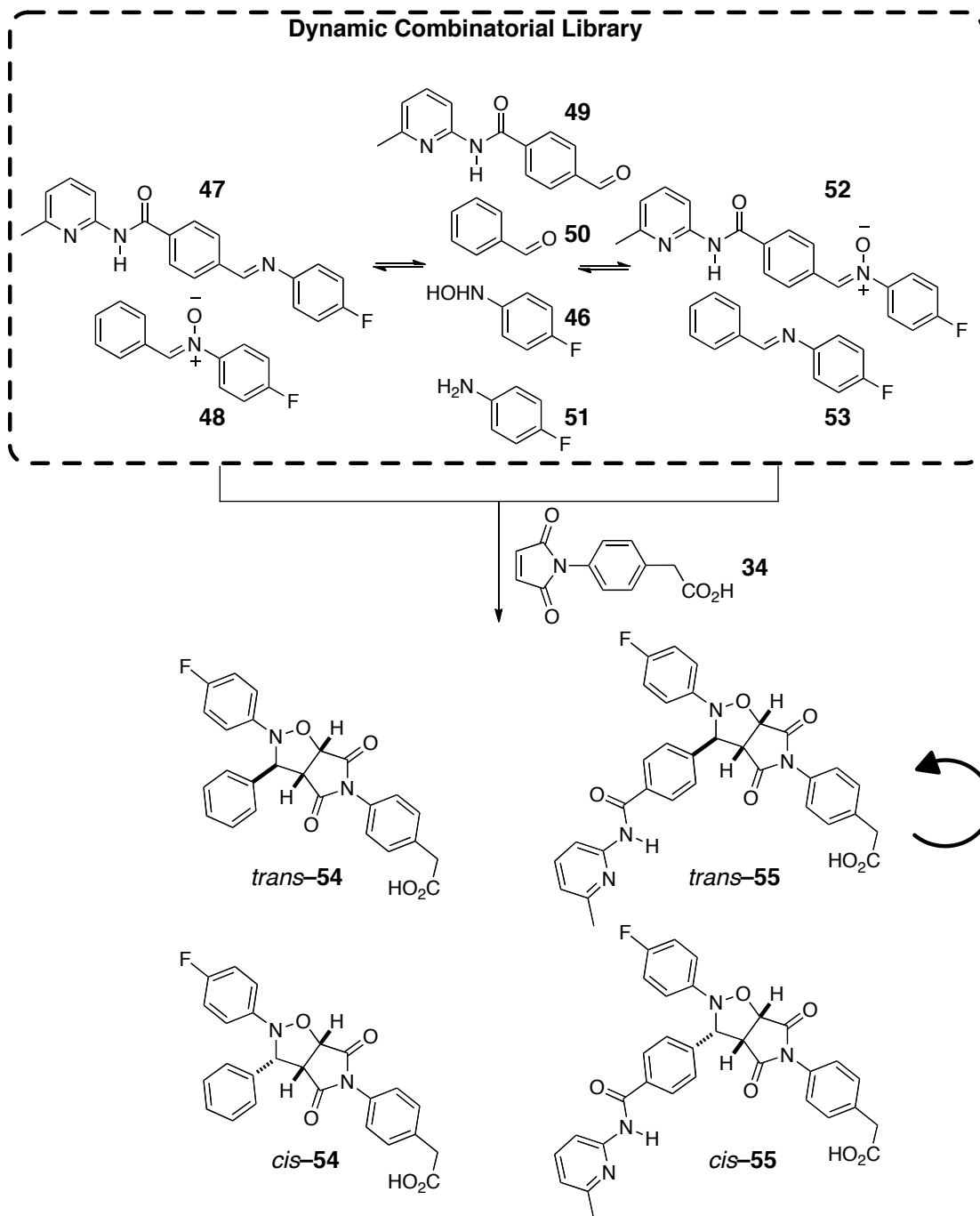
undertaken in the system as the reagents for the reaction of both nitrones with the maleimide are present and the fastest reaction is the winner. In contrast, biological systems do not carry every required reagent for every possible reaction at all times as this would be unfeasible as well as could lead to wastage through undesired reactions. Instead reagents for a specific reaction are produced from simpler, common components connected through a system of complex signalling and metabolic pathways and then reagents required are produced as and when needed in response to stimuli.

## 2.5 Coupling the Two Replicator System to a Dynamic Exchanging Process

We have previously demonstrated,<sup>154,155</sup> Section 1.6.2, that under the reaction conditions for the kinetic 1,3-dipolar cycloaddition, it is possible to subject the reagents to a dynamic imine–nitron exchange reaction.

In our previous example (Scheme 2.8) three components are mixed together to initiate the reaction, simple nitron **48**, recognition bearing imine **47** and maleimide **34**. The recognition bearing nitron (**52**) required to form the replicator is not present at the start of reaction but must be formed through the dynamic nitron–imine exchange which is established upon mixing. As the replicating template, *trans*-**55** consumes the recognition bearing nitron **52** at an increased rate, the dynamic equilibrium will be perturbed and therefore will constantly work to reestablish the equilibrium, replenishing the concentration of nitron **52**. By doing so, more reagents for the autocatalytic reaction are produced which facilitates the formation of more replicating template. As the replicating template influences the redistribution of the system to provide more reagents for its own synthesis, it can be thought of as a primitive feedback loop.

We now wished to use the controllable two replicator system and couple the production of the two recognition bearing nitrones to a dynamic process. In this new system, a dynamic pool of reagents would possess the building blocks for two different replicating templates with the outcome of the system able to be controlled by addition of either of the two templates.

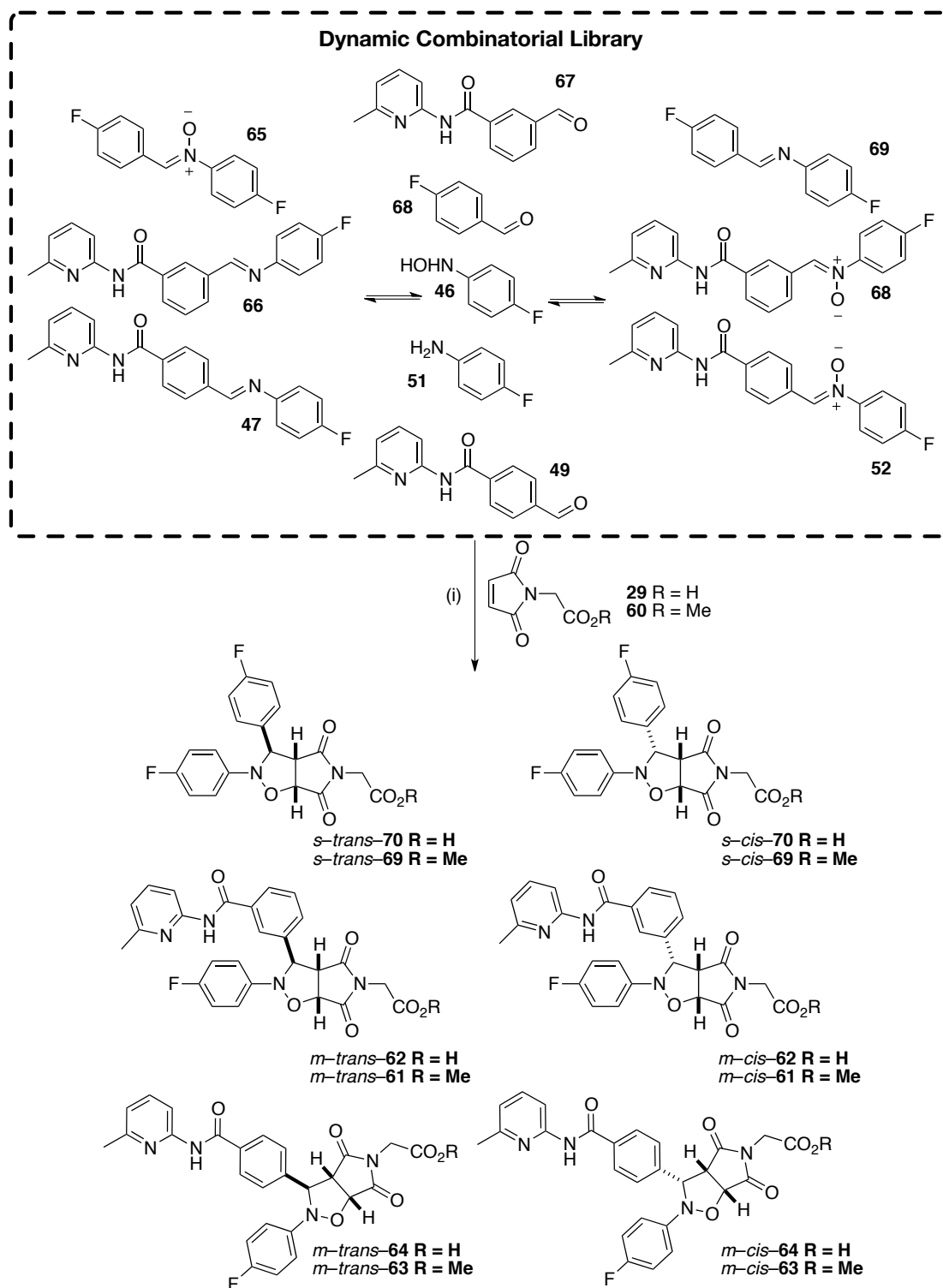


**Scheme 2.8:** The self-replicating template *trans*-55 is able to selectively produce itself from a dynamic reagent pool. The compounds enclosed within the black dashed line are part of a dynamic library. Reagents and conditions (i)  $\text{CD}_2\text{Cl}_2$ / *p*-TSA (sat), 0 °C, 16 h.

A system was designed as outlined in Scheme 2.9 which consisted of four equimolar reagents, *simple*<sup>†</sup>-nitronium 65, *m*-imine 66, *p*-imine 52 and maleimide acid 29. Upon mixing, the only kinetic reaction able to proceed is the cycloaddition between *s*-

<sup>†</sup> The term simple describes that the component does not bear a recognition site.

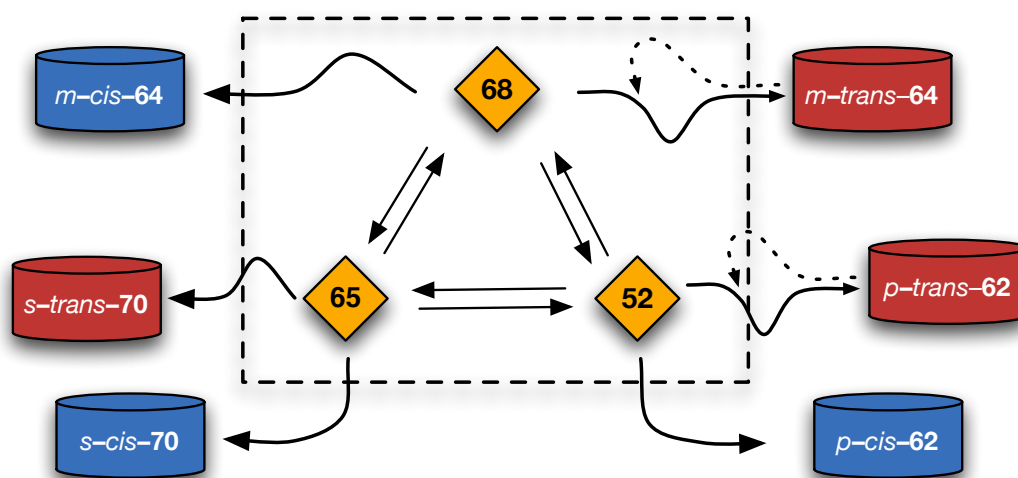




**Scheme 2.9:** Upon mixing the four components for the reaction, imines **66** and **47** and nitrone **65**, will produce an eleven component dynamic library which consists of three nitrone **65**, **68** & **52**, three imines, **66**, **47** & **80** and their component aldehydes, **49**, **67** & **68**, aniline **51** and hydroxylamine **46**. In the presence of maleimide **29**, each of the three nitrone can react in a 1,3-dipolar cycloaddition reaction to furnish six cycloadducts, three *trans* and three *cis*. Of these two cycloadducts, *m-trans-62* and *p-trans-64* are autocatalytic self-replicating templates. The compounds enclosed within the black dashed line are part of a dynamic library. Reagents and conditions (i)  $\text{CD}_2\text{Cl}_2/p\text{-TSA}$  (sat), 10 mM, 25 °C, 40 h.

nitronone **55** and maleimide **29** to create two diastereoisomeric products, *s-trans*-**70** and *s-cis*-**70** by bimolecular processes. Under the conditions of the reaction we have shown that nitrones and imine are able to hydrolyse to their hydroxylamine or amine and aldehyde components. In this case, the hydroxylamine **46** would then be able to recombine with any of the aldehydes to form *m*-nitronone **68** and *p*-nitronone **52**. The amine **51** would also be able to recombine with the simple aldehyde **68** to form the non recognition bearing imine **80**. Without any interference from a kinetic reaction, the exchange would reach an equilibrium with the ratio of nitrones and imines present reflected in the thermodynamic stability of each of the compounds. As the system is coupled to a kinetic reaction, however, this equilibrium is constantly disturbed as nitrones will be consumed at a different rates with *m*-nitronone **68** and *p*-nitronone **52** consumed faster than *s*-nitronone **65** as a consequence of autocatalysis.

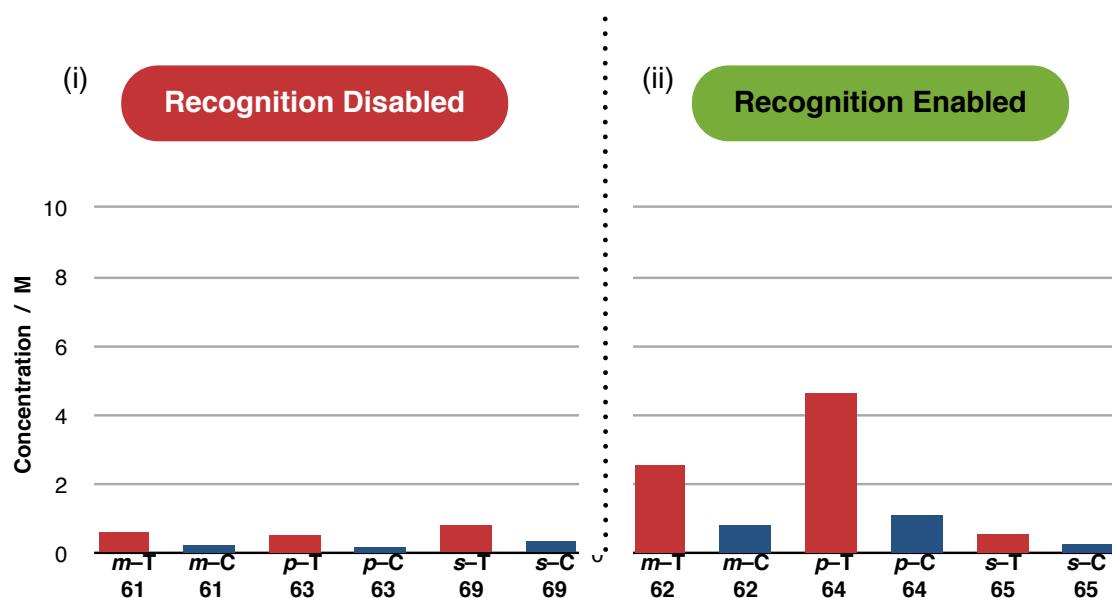
The dynamic system therefore will work to re-establish the original equilibrium according to Le Châteliers principle and in doing so provide more *m*-nitronone **68** and *p*-nitronone **52**. The system may be visualised graphically such as in Figure 2.9 using the same style used in the introduction, Section 1.6.2.



**Figure 2.9:** Representation of the system in which the three reactive reagents within the black dashed line exchange under thermodynamic control and will each react irreversibly to produce their two diastereoisomeric products. Two of these product are produced autocatalytically as represented by the second arrow looping back into the dynamic box. The other components of the dynamic library are omitted for clarity.

We first set about to demonstrate that once again, recognition mediated processes are prevalent in the system after coupling to a dynamic process by comparison of a recognition disabled control system to a recognition enabled system. Initially, for the recognition disabled control reaction, an equimolar mixture of the four components,

*s*-nitronone **65**, *m*-imine **66**, *p*-imine **47** and maleimide ester **60** at 10 mM in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat), were incubated at 25 °C for 40 h after which point the system was assayed by 376.5 MHz <sup>19</sup>F NMR spectroscopy and the relative concentration of each product determined by deconvolution.



**Figure 2.10:** Concentration of products after 40 h incubation at 25 °C for: (i) bimolecular control reaction with maleimide ester **73**. (ii) recognition mediated system with maleimide **72**. Abbreviations used: *m* – meta, *p* – para, *s* – simple, T – trans and C – cis.

After 40 h, the recognition disabled system reached 28% overall conversion with similar concentrations of the three *trans* cycloadducts (0.61 mM *m*-*trans*-**61**, 0.53 mM *p*-*trans*-**63** and 0.83 mM *s*-*trans*-**69**) and the three *cis* cycloadducts (0.24 mM *m*-*cis* **61**, 0.20 mM *p*-*cis*-**63** and 0.35 mM *s*-*cis*-**69**) (Figure 2.10 (i)). The slightly higher concentrations from the simple nitronone are a result of the higher abundance of the simple nitronone upon initiation of the reaction.

The recognition enabled system was then set up using the free acid maleimide **29** in the four component system instead of the ester. The system of four equimolar components, *s*-nitronone **65**, *m*-imine **66**, *p*-imine **47** and maleimide **29** at 10 mM in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat), were subjected to the same reaction conditions (25 °C, 40 h) and assayed by 376.5 MHz <sup>19</sup>F NMR spectroscopy and the relative concentration of each product determined by deconvolution and displayed in (Figure 2.10 (ii)).

In the case of the recognition mediated system, there was complete consumption of maleimide after the same time. The two replicating templates *m*-*trans*-**62** and *p*-

*trans*-**64** are now the most abundant products in the system reaching a concentration 2.57 mM and 4.64 mM respectively. These two products now account for 72% of the product pool. The cycloadducts arising from the simple nitron, *s-trans*-**70** and *s-cis*-**70**, are the two lowest produced products in the pool reaching a concentration of 0.56 mM and 0.27 mM respectively. These values are lower than the simple cycloadducts in the control system which highlights the strength of the autocatalytic processes at redistributing the equilibrium of nitron reagents in the system, producing more recognition bearing nitron at the expense of the simple nitron and leaving less simple nitron to react with the maleimide.

Interestingly, by slowly feeding the nitron reagents for the two autocatalytic cycles by the dynamic exchange reaction rather than providing both reagents at the start of the reaction, as in the kinetic system (Section 2.4), the ratio of the two replicating templates has changed. The more efficient of the two replicators in isolation, *p-trans*-**64**, is produced more selectively in the dynamic system than in the kinetic system with the ratio of *p-trans*-**64** to *m-trans*-**62** increased from 1.3:1 in the kinetic system to 1.8:1 in the dynamic system.

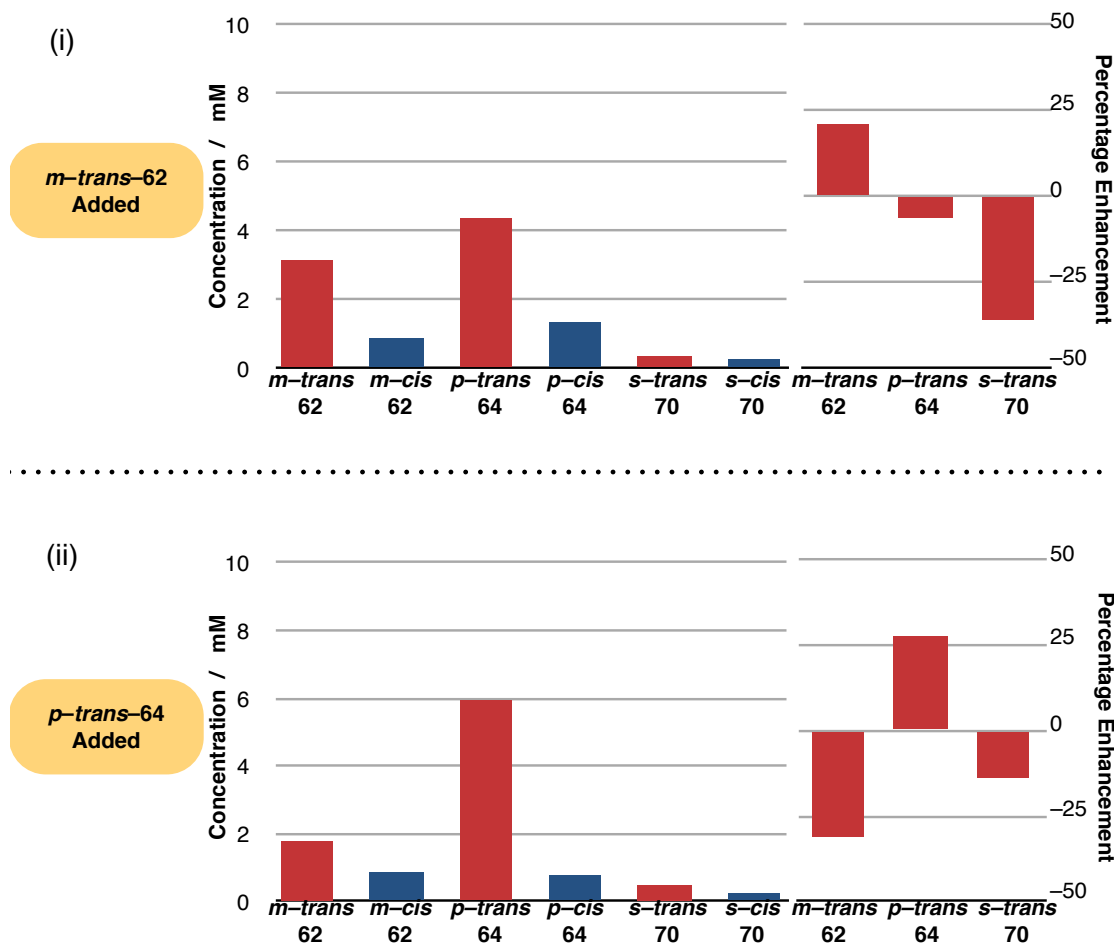
### 2.5.1 Template Direction of a Dynamic Reagent Pool

With the successful creation of a system of two coexisting replicators selectively forming themselves from a dynamic reagent pool we now wished to attempt to control the outcome of system by instructing it to produce a specific template using template doping as instructions as was successfully described in the kinetic system (Section 2.4.1).

The instruction was issued in the same manner, by doping a fresh set of reagents with 20 mol% of one of the two replicating templates. The mixture was then incubated under the same conditions and assayed by 376.5 MHz <sup>19</sup>F NMR spectroscopy to extract the relative concentrations of products in the system by deconvolution. For experimental details, see Chapter 7, Section 7.3.

Upon issuing the instruction for the system to produce *m-trans*-**62**, once again the system reached quantitative conversion however the distribution of products in the system had changed (Figure 2.11 (i)), from the non-instructed system. The

concentration of newly formed *m-trans*-62 reached a concentration of 3.12 mM with the second replicating template *p-trans*-64 reaching a concentration of 4.35 mM. These two templates now account for 75% of the product pool with the production of *m-trans*-62 enhanced by 21% whilst the production of *p-trans*-64 was inhibited by 6%. Clearly the system is still able to accept and respond to an instruction by adding preformed template however, the instruction has not been able to enhance the final concentration of *m-trans*-62 enough for it to become the dominant product in the pool.



**Figure 2.11:** Concentration of newly formed cycloadducts after 40 h as determined by  $^{19}\text{F}$  NMR spectroscopy after instruction with: (i) 20 mol% *m-trans*-62 and (bottom) 20 mol% *p-trans*-64.

Upon issuing the second instruction, to produce *p-trans*-64, by doping the mixture of reagents with 20 mol% *p-trans*-64, the system once again reached quantitative conversion and displayed a different distribution of products (Figure 2.11 (ii)). There was a clear enhancement in the production of *p-trans*-64 which reached a concentration of 5.93 mM, by far the most abundant product in the system

accounting for 59% of the total product population and greater than double the concentration of the next most abundant cycloadduct. The second replicating template *m-trans*-**62** was the second most abundant product in the mixture reaching a concentration of 1.79 mM. Clearly the production of *p-trans*-**64** was enhanced in comparison to the uninstructed system, with an enhancement of 28% whereas the production of the second replicating template was inhibited by 31%.

The results demonstrate that even when the pool of reagents for the replicators are under dynamic control, the outcome of the system is still able to be selectively controlled by instructing with either of the two preformed templates. The level of enhancement when the mixture is doped with *p-trans*-**64** is higher than when the mixture is doped with *m-trans*-**62**. As the *p-trans*-**64** template is a more efficient replicator, when it is doped into the mixture, is able to consume its reagent, *p*-nitron **52**, from the dynamic pool at the greatest rate (and greater than the rate at which *m*-nitron **68** is consumed when *m-trans*-**62** is doped into the mixture). The equilibrium of the dynamic process will be perturbed and will work to re-establish the equilibrium by providing more *p*-nitron **52**. This feedback loop will work to increase the selectivity of the more efficient autocatalytic cycle in the system.

## 2.6 Summary of Two Replicator Systems

We have successfully demonstrated that a system of two replicators which share a common reagent can be created and that the two are able to coexist in solution. The more efficient replicator in isolation maintains its efficiency in the system existing as the dominant product in the kinetic system. When the reagents for the autocatalytic cycles are slowly fed into the system by coupling to a dynamic process, the selectivity for the more efficient autocatalytic cycle increases. Each system was able to be instructed to produce one or the other of the two replicating templates by doping with prefabricated template and the magnitude of the enhancements were greater in the dynamic system than in the kinetic system.

For sufficient reactivity, the systems needed to be incubated at 25 °C to increase the rate of cycloadditions but by running the reactions at this temperature, the strength of recognition events in the systems are lowered relative to performing the reaction at 10 °C and the rate of bimolecular reactions are increased which overall has the effect

of lowering the efficiency of recognition mediated processes. Ideally, we would prefer to be able to work at a lower temperature to take advantage of higher association constants and slower bimolecular reactions which should have the effect of increasing recognition mediated processes. For the replicating templates used in these examples, lowering the temperatures would require extended reaction times as a consequence of their relatively low efficiency. Ideally we would like to use replicators with higher efficiency to show greater enhancements in comparison to bimolecular processes.

## 2.7 Change of Maleimide

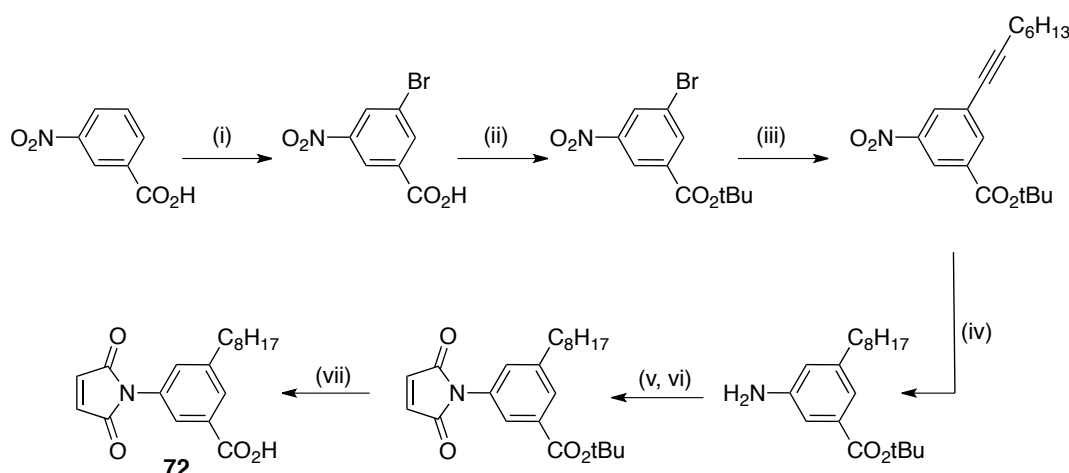
When looking to redesign the two replicators to increase their efficiency, it was logical to look first of all to alter the maleimide that they share rather than redesign both recognition bearing nitrones. Therefore we now attempted to identify an alternative maleimide which would react with the two recognition bearing nitrones and each behave as a minimal replicator.

Our attention turned to a benzoic acid maleimide **71**, (Figure 2.12) used previously in the group with limited success. The introduction of a benzoic acid group as a recognition site as opposed to an acetic acid has the effect of increasing the association constant.



**Figure 2.12:** The original benzoic acid maleimide **71** which was poorly soluble in non-polar solvents and the newly designed maleimide **72** decorated with an alkyl chain to increase solubility.

The reaction between maleimide **71** and *p*-nitron **52** has been shown<sup>156</sup> to behave as a minimal self-replicator however the template produced has low solubility in the non-polar solvents used. It was therefore desirable to modify the maleimide to increase its solubility in non-polar solvents without adding any groups that would interfere with the reaction or recognition site. The modification chosen was to add a lipophilic alkyl chain as shown in Figure 2.12.



**Scheme 2.10:** Reagents and conditions (i) NBS, H<sub>2</sub>SO<sub>4</sub>, Δ, 16 h, 100% (ii) BOC<sub>2</sub>O, DMAP, THF, RT, 60 h, 81% (iii) Octyne, Pd(P(Ph)<sub>3</sub>)<sub>2</sub>Cl<sub>2</sub>, CuI, PPh<sub>3</sub>, NEt<sub>3</sub>, Δ, 16 h, 88% (iv) Pd/C, MeOH, THF, H-Cube, 50 °C, 50 bar H<sub>2</sub>, 99% (v) Maleic anhydride, THF (vi) ZnBr<sub>2</sub>, HMDS, MeCN, Δ, 1.5 h, 84% over two steps (vii) TFA, CH<sub>2</sub>Cl<sub>2</sub>, RT, 1 h, 96%.

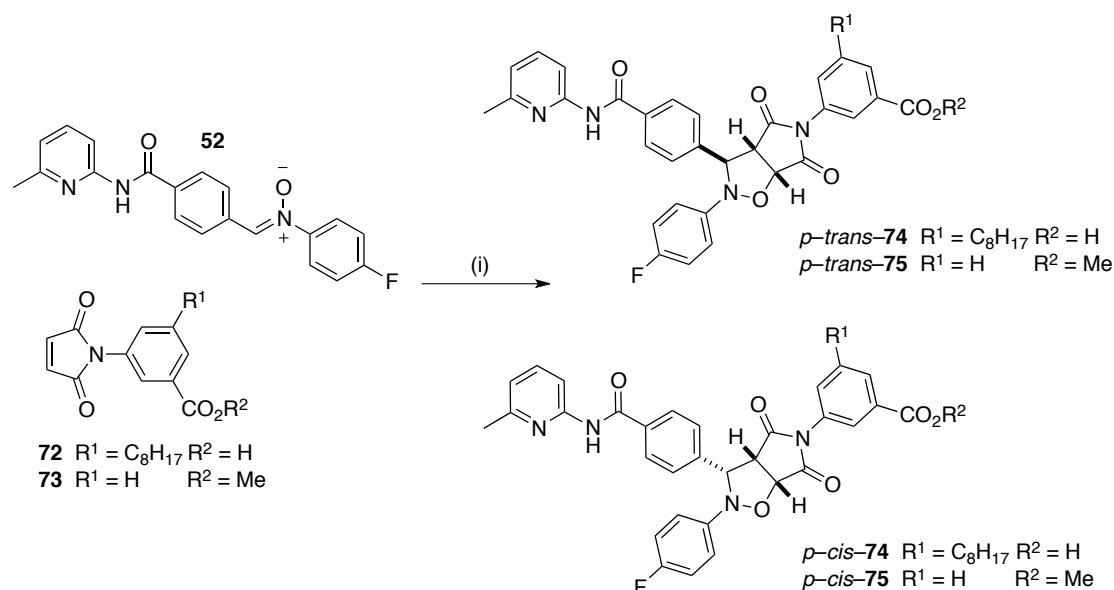
Our experience within the group with Sonogashira reactions made this reaction an ideal choice to employ to install an alkyl chain to the maleimide in a synthetic scheme as shown in Scheme 2.10. In order to prepare for the Sonogashira reaction, a halide had to be added to the ring which was achieved by brominating 3-nitrobenzoic acid with *N*-bromosuccinimide in sulfuric acid.<sup>189</sup> The acid group next had to be shielded with a protecting group stable to the basic Sonogashira reaction conditions. The acid was thus protected as a *tert*-butyl ester obtained after reacting the acid with BOC anhydride in the presence of catalytic DMAP in THF. Following a successful Sonogashira reaction to install octyne, the next step was to hydrogenate the nitro group to an amine to make it available for subsequent maleimide formation. Initial attempts at hydrogenation under standard conditions using a H<sub>2</sub> balloon and 10% Pd/C in a methanol/THF mix showed that the nitro group was readily reduced to the amine however after complete nitro reduction there was also a mix of partially reduced alkyne compounds. Extended reaction times of several days, refreshing catalyst and H<sub>2</sub> balloons at best yielded 50% of the desired product after column purification. An alternative procedure using the H-Cube®, a continuous flow hydrogenation reactor, was utilised with the optimum conditions found after performing small scale test reactions. The nitro and alkyne were successfully reduced using 50 bar of H<sub>2</sub> and at 50 °C over a 10% Pd/C catalyst to give the saturated amine in a quantitative yield. The maleimide formation was carried out in two steps. Firstly, reaction of the amine with maleic anhydride in THF at room temperature provides the ring opened maleimide which was immediately cyclised



using ZnBr<sub>2</sub> and HMDS in MeCN affording the *t*-butyl ester protected maleimide in an 84% yield over both steps. Finally the acid was deprotected by stirring the maleimide ester in a TFA : CH<sub>2</sub>Cl<sub>2</sub> mix (1:10) for one hour to yield the target free acid maleimide **72** which satisfyingly was substantially more soluble in CDCl<sub>3</sub> than the undecorated original maleimide **71**.

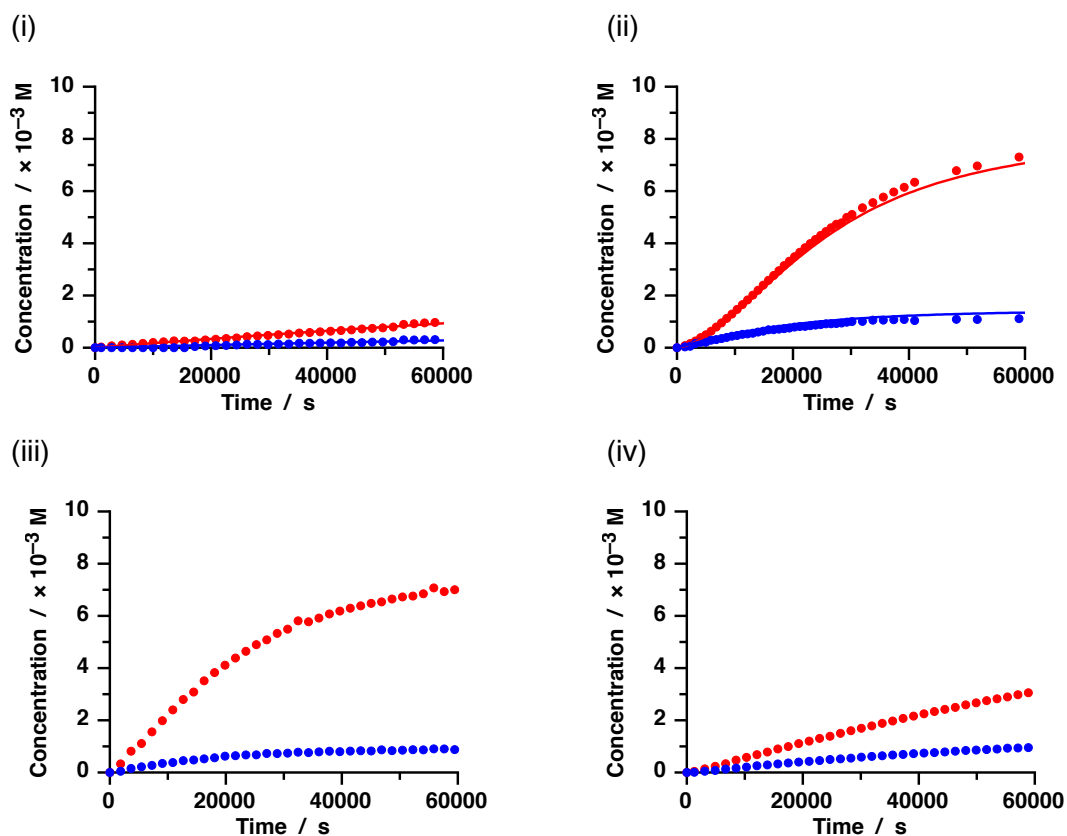
### 2.7.1 Analysis of the reaction of maleimide **72** with *p*-nitrone **52**

With the new maleimide in hand, we set about to ensure addition of the alkyl group does not disrupt the behaviour of the self-replicator. We therefore set about to demonstrate both of the replicators in the standard fashion as demonstrated previously (Section 2.3) starting with the reaction of maleimide **72** with *p*-nitrone **52**. Initially, the rate of bimolecular reaction in the system was investigated by monitoring the recognition disabled control reaction of *p*-nitrone **52** with maleimide ester **73**, at 10 mM in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat) at 10 °C for 16 h, by 470.3 MHz <sup>19</sup>F NMR spectroscopy using the standard kinetic experimental procedure, Chapter 7, Section 7.6. The relative concentrations of species in the reaction were determined by deconvolution of appropriate resonances and a concentration vs. time profile (Figure 2.13 (i)) was constructed. The results show that the recognition disabled reaction proceeded slowly reaching 13% conversion overall with the major product *p*-*trans*-**75** reaching 1.0 mM with its diastereoisomer *p*-*cis*-**75** reaching 0.3 mM representing a *trans* : *cis*



**Scheme 2.10:** Reagents and conditions (i) CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat), 10 mM, 10 °C, 16 h.

ratio of 3.1:1. The results were subject to kinetic simulation and fitting to a model for a bimolecular reaction using SimFit to extract rate information. The reaction rate constants for  $k_{p\text{-trans-75}}$  and  $k_{p\text{-cis-75}}$  were determined to be  $1.77 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  and  $0.53 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  respectively.



**Figure 2.13:** Concentration vs. time profiles for: (i) Control reaction of nitrene **52** with maleimide ester **73** (ii) Recognition mediated reaction of nitrene **52** with maleimide acid **72** (iii) Template doped reaction of nitrene **52** with maleimide **72** with 10 mol% *trans*-**74** (iv) Competitive inhibitor reaction of nitrene **52** with maleimide **72** and 5 eq. benzoic acid. Experimental data is represented as circles, *trans* cycloadducts (●), *cis* cycloadducts (●). The results of kinetic simulation and fitting using a model for a self-replicating system is shown as lines *trans* cycloadducts (—), *cis* cycloadducts (—). Reaction conditions  $\text{CD}_2\text{Cl}_2/p\text{-TSA}$  10 mM,  $10^\circ\text{C}$ , 16 h.

**Table 2.3:** Rate constants determined by kinetic simulation and fitting of experimental data using SimFit.

	$k_{\text{trans}}$	$k_{\text{cis}}$
Bimolecular Rate Constant / $\times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$	1.77	0.53
$k_{\text{ternary}} / \times 10^{-4} \text{ s}^{-1}$	71.6	—
EM Ternary / M	40	—
$K_{\text{Duplex}} / \times 10^8 \text{ M}^{-1}$	3.52	—
Connection EM / M ( $\Delta G^\ddagger / \text{kJMol}^{-1}$ )	40.5 (8.71)	—

The recognition mediated reaction was then performed by monitoring the reaction of recognition enabled maleimide **72** with *p*-nitrone **52** under the same conditions. After 16 h, the reaction had reached 84% conversion with *p-trans-74* the major product reaching 7.3 mM and *p-cis-74* reaching 1.1 mM (Figure 2.13 (ii)). The *trans* : *cis* ratio therefore was 6.6:1 indicating an increase in selectivity for the *p-trans-74*. The concentration vs. time profile for this reaction now showed non-linear growth with a sigmoidal profile for the production of *p-trans-74*, characteristic for an autocatalytic self replicating reaction. The production of *cis-74* was also enhanced (1.1 mM) in comparison to the recognition disabled control (0.3 mM) indicating that recognition mediated processes are responsible for increasing the production of the *cis* cycloadduct indicating the presence of a recognition mediated [A•B] complex channel. The results were once again subjected to kinetic simulation and fitting to a model for a self-replicating reaction using SimFit to extract reaction rate information. The rate of reaction within the ternary complex,  $k_{\text{ternary}}$ , was calculated to be  $7.16 \times 10^{-3} \text{ s}^{-1}$  which reveals the effective molarity of the system to be 40 M. The reaction rate of the binary [A•B] complex reaction,  $k_{[\text{A}\cdot\text{B}]}$ , was calculated to be  $6.7 \times 10^{-3} \text{ M}^{-1}\text{s}^{-1}$  from which we could calculate the effective molarity of the binary complex channel to be 7.8 M. These results highlighted that the whilst the binary complex is an efficient process within the system it is surpassed by the efficiency of the autocatalytic cycle.

The template directed nature of the reaction was demonstrated by addition of 10 mol % preformed *p-trans-74* to a fresh set of reagents for the reaction. The mixture was then subjected to the same conditions and monitored as before by 470.3 MHz  $^{19}\text{F}$  NMR spectroscopy. From the results, a concentration vs. time profile was constructed (Figure 2.13 (iii)), which clearly shows that the lag period observed in the native reaction is absent and the reaction proceeds at the maximum autocatalytic rate upon initiation. After 16 h the *trans* : *cis* ratio had increased from 6.6:1 in the undoped system to 13:1 when doped with preformed template. This result highlights that the template directed reaction is selective for the production of *p-trans-74* and not involved in any cross-catalytic pathway to enhance *p-cis-74*.

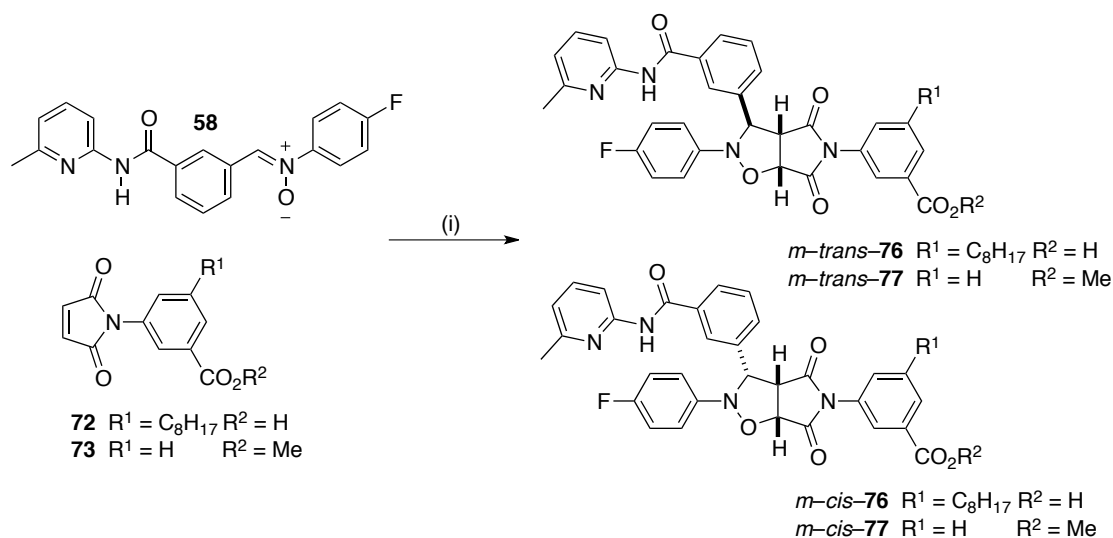
Finally, five equivalents of the competitive inhibitor benzoic acid was added to a fresh set of reagents for the replicating system to demonstrate the importance of recognition mediated events. The reaction was performed under the same conditions

and monitored as before by 470.3 MHz  $^{19}\text{F}$  NMR spectroscopy for 16 h. From the results, a concentration vs. time profile (Figure 2.13 (iv)), was once again constructed and in this case, the production of *p-trans*-**74** has been inhibited in comparison to the native system (Figure 2.13 (ii)). This decrease in autocatalytic efficiency is a consequence of benzoic acid disrupting the formation of the ternary complex by competing for binding to the amindopyridine binding site. The rate of formation of *p-cis*-**74** proceeded at close to the native rate, further evidence for the presence of an **[A•B]** complex channel as an increase in concentration of competitive inhibitor will statistically have less of an effect on a binary complex as with a ternary complex.

From the results of the standard protocol of autocatalytic self-replicator characterisation, it is clear that the dominant reaction of *p*-nitrone **52** with maleimide **72** is the autocatalytic formation of *p-trans*-**74**. There is however a binary complex channel present which is able to enhance the formation of *p-cis*-**74** but not significantly enough to become the major pathway.

### 2.7.2 Analysis of the reaction of maleimide **72** with *m*-nitrone **68**

Our attention now turned to characterising the behaviour of our second self-replicator, the reaction of *m*-nitrone **68** with maleimide **72** (Scheme 2.11). Accordingly, the standard protocol for identifying and characterising autocatalytic self-replicators, as described previously, was carried out to investigate the reaction between *m*-nitrone **68** and maleimide **72** under the same conditions as used for the *p*-nitrone **52** replicating system.

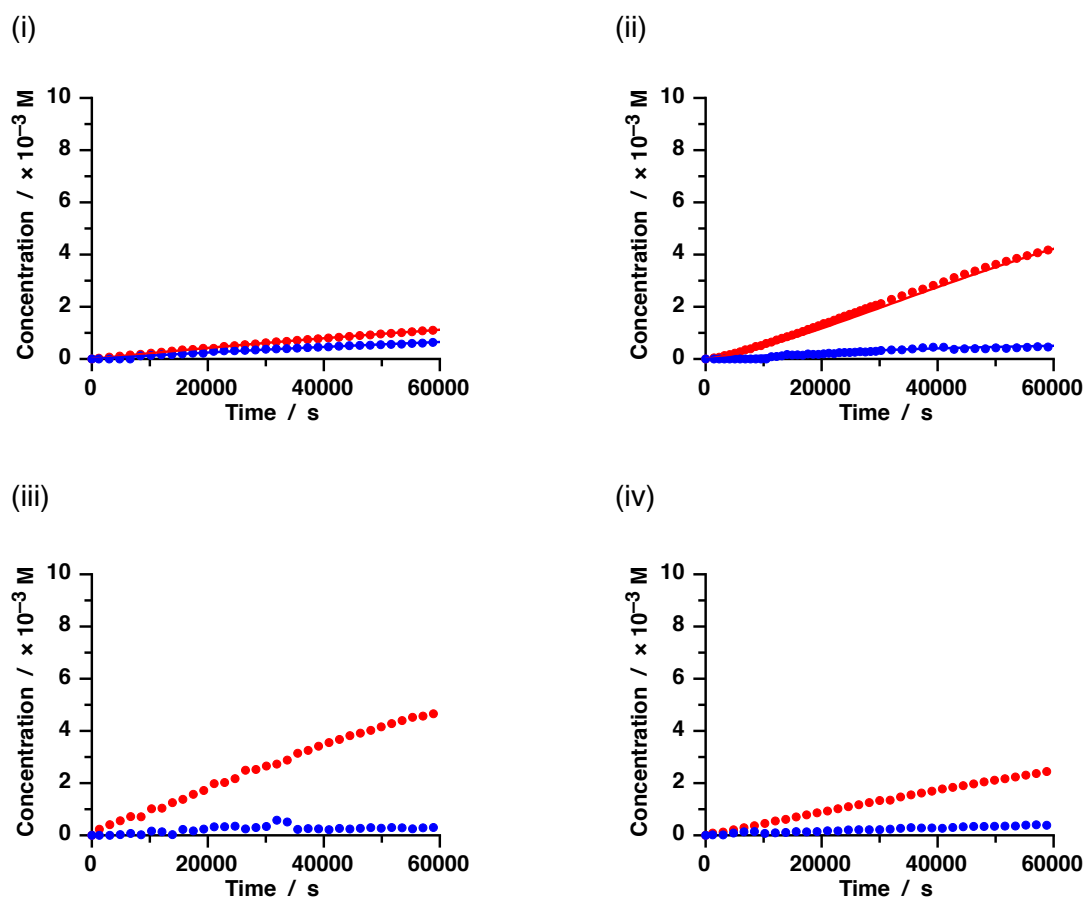


**Scheme 2.11:** Reagents and conditions (i)  $CDCl_3$ , 10 mM, 10 °C, 16 h.

Initially, to characterise the rate of bimolecular reaction in the system, the reaction of *m*-nitron **68** with the recognition disabled maleimide ester **73** was performed at 10 mM in  $CD_2Cl_2/p$ -TSA (sat) at 10 °C and monitored by 470.3 MHz  $^{19}F$  NMR spectroscopy for 16 h in the standard way as described in the experimental section, Chapter 7, Section 7.6. From the results, a concentration vs. time profile was produced (Figure 2.14 (i)) which revealed that the recognition disabled reaction proceeded slowly reaching 17% overall conversion, with the major product *m-trans-77* reaching 1.1 mM with its diastereoisomer *p-cis-77* reaching 0.6 mM representing a *trans* : *cis* ratio of 1.7:1. The results were once again subjected to kinetic simulation and fitting to a model for a bimolecular reaction using SimFit to extract rate information. The reaction rate constants for  $k_{m-trans-77}$  and  $k_{m-cis-77}$  were determined to be  $2.27 \times 10^{-4} M^{-1}s^{-1}$  and  $1.32 \times 10^{-4} M^{-1}s^{-1}$  respectively.

The recognition mediated reaction was next performed by reacting recognition enabled maleimide **72** with *m*-nitron **68** under the same conditions and monitored as before by 470.3 MHz  $^{19}F$  NMR spectroscopy for 16 h. From the results, a concentration vs. time profile was constructed (Figure 2.14 (ii)) which revealed a dramatically different reaction profile. After 16 h, the reaction had reached 47% overall conversion with *m-trans-76* the major product which reached a concentration of 4.2 mM with its diastereoisomeric partner *m-cis-76* reaching 0.5 mM. The *trans* : *cis* ratio therefore was 8:1 indicating an increase in selectivity for the *m-trans-76*. The concentration vs. time profile now showed a sigmoidal profile for the production

of *m-trans*-**76**, characteristic for an autocatalytic self replicating reaction. The production of *m-cis*-**76** was also inhibited reaching (0.5 mM) in comparison to the recognition disabled control (0.6 mM) indicating that the binary complex channel reaction to form *m-cis*-**76**, as was present in the *p*-nitro system, was inefficient if present at all in this case. Reaction rate constants were once again obtained from kinetic simulation and fitting to a model for a self-replicating reaction using SimFit. The rate of reaction within the ternary complex,  $k_{\text{ternary}}$ , was calculated to be  $3.8 \times 10^{-3} \text{ s}^{-1}$  from which the effective molarity of the replicator was calculated to be 17 M.



**Figure 2.14:** Concentration vs. time profiles for: (i) Control reaction of nitrone **68** with maleimide ester **73** (ii) Recognition mediated reaction of nitrone **68** with maleimide acid **72** (iii) Template doped reaction of nitrone **68** with maleimide **72** with 10 mol% *trans*-**76** (iv) Competitive inhibitor reaction of nitrone **68** with maleimide **72** and 5 eq. benzoic acid. Experimental data is represented as circles, *trans* cycloadducts (●), *cis* cycloadducts (●). The results of kinetic simulation and fitting using a model for a self-replicating system is shown as lines *trans* cycloadducts (—), *cis* cycloadducts (—). Reaction conditions  $\text{CD}_2\text{Cl}_2/p\text{-TSA}$  10 mM, 10 °C, 16 h.

The template directed manner of the reaction was demonstrated by addition of 10 mol% preformed *m-trans*-**76** to a fresh set of reagents for the recognition mediated reaction which was then performed under the same conditions and

**Table 2.4:** Rate constants determined by kinetic simulation and fitting of experimental data using SimFit.

	$k_{\text{trans}}$	$k_{\text{cis}}$
Bimolecular Rate Constant/ $\times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$	2.27	1.32
$k_{\text{ternary}} / \times 10^{-4} \text{ s}^{-1}$	38	—
$EM \text{ Ternary} / \text{ M}$	17	—
$K_{\text{Duplex}} / \times 10^8 \text{ M}^{-1}$	10.1	—
Connection $EM / \text{ M}$ ( $\Delta G^{\circ} / \text{ kJmol}^{-1}$ )	116 (11.2)	—

monitored the reaction as before by 470.3 MHz  $^{19}\text{F}$  NMR spectroscopy. After deconvolution of the results, a concentration vs. time profile was created (Figure 2.14 (iii)) which revealed that the lag period observed in the native reaction was now absent and the reaction proceeded at the maximum autocatalytic rate from the start of the reaction. The overall conversion had increased to 50% with the concentration of *m-trans-76* increased to 4.7 mM with the production of *m-cis-76* inhibited with only 0.3 mM produced. The *trans* : *cis* ratio had therefore increased to 15:1 highlighting that the autocatalytic cycle is selective for *m-trans-76* and not involved in any cross-catalytic pathway to enhance *m-cis-76*.

Finally five equivalents of the competitive inhibitor benzoic acid was added to a fresh set of reagents for the recognition mediated reaction to demonstrate the importance of recognition events in the system. The reaction was performed under the same conditions and monitored as before. From the results, a concentration vs. time profile was constructed (Figure 2.14, (iv)) which showed that production of *m-trans-76* has been inhibited owing to benzoic acid disrupting the formation of the ternary complex. The rate of formation of *m-cis-76* remained unaltered indicating its formation proceeded through a bimolecular reaction only.

From the results of following the standard protocol of autocatalytic self-replicator characterisation, it is clear that the dominant reaction of *m-nitrone 68* with maleimide **72** is the autocatalytic formation of *m-trans-76* which has an  $EM$  of 17 M.

Comparison of the two replicating templates reveals that the autocatalytic reaction of maleimide **72** with *p-nitrone 52* is more efficient than the reaction *m-nitrone 68*. The

*EM* of the *p*-nitronone **52** system is 40 M in comparison to 17 M for the *m*-nitronone **68** system and also the rate of reaction in the ternary complex is 1.9 times faster in the *p*-nitronone **52** system than the *m*-nitronone **68** system. The differences in efficiencies of the two replicators is greater when utilising the new benzoic acid maleimide **72** than when the glycine derived maleimide **29** was used as described previously in this chapter. Therefore we might expect to observe a hierarchy of reactivity in a three component system of the two nitrones, **52** & **68**, mixed with maleimide **72** with the more efficient replicator produced most abundantly in the system.

### 2.7.3 Investigation for Cross-Catalysis

As a consequence of the structural similarity of the two replicating template, we investigated the potential for cross-catalysis as demonstrated before in Section 2.3.4. The reactions were performed using the standard kinetic experimental procedure as used above and described in Chapter 7, Section 7.6

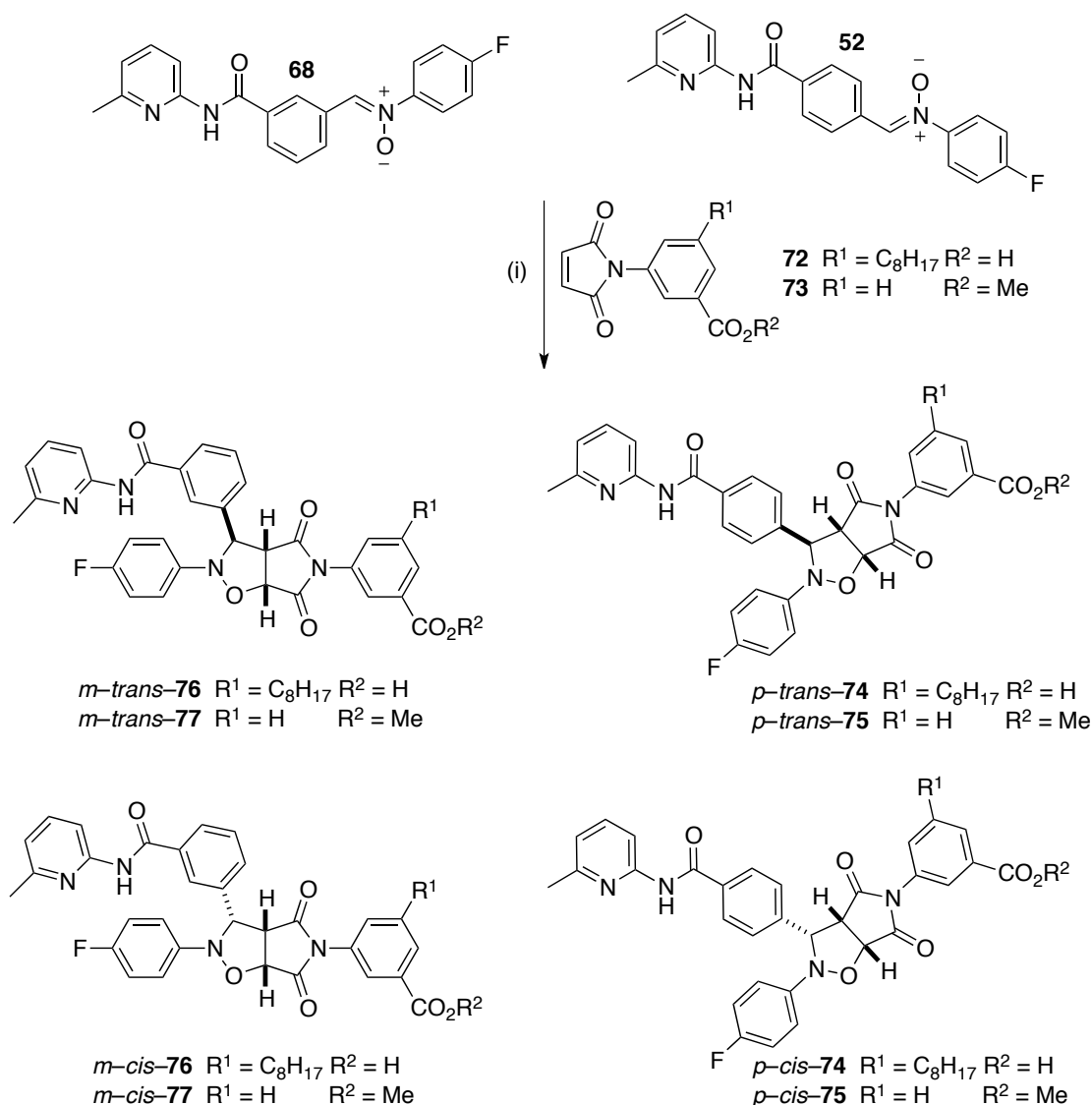
Initially, 20 mol% *m-trans*-**76** was doped into the reaction of *p*-nitronone **52** with maleimide **72** and the reaction performed under the same conditions as before, 10 mM in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat) at 10 °C, and the reaction monitored by 470.3 MHz <sup>19</sup>F NMR spectroscopy. The results were extracted as described previously and compared to the native reaction between *p*-nitronone **52** and maleimide **72** under the same conditions. In this case, no enhancement in the production of *p-trans*-**74** was observed but instead a small degree of inhibition was evident. These results demonstrated that *m-trans*-**76** could not act as a template for the formation of *p-trans*-**74**.

Next, *p-trans*-**74** was doped into the reaction of *m*-nitronone **68** with maleimide **72** system and the reaction performed under the same conditions as before, 10 mM in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat) at 10 °C, and the reaction monitored by 470.3 MHz <sup>19</sup>F NMR spectroscopy. The results were compared to the un-doped reaction of maleimide **72** with *m*-nitronone **68** under the same conditions and revealed a slight enhancement in the production of *m-trans*-**76** however not as strong as when *m-trans*-**76** itself is doped into its own system.



From these results we could conclude that *m-trans*-**76** is a selfish autocatalyst and *p-trans*-**74** is a strong autocatalyst and a weak cross-catalyst for the production of *m-trans*-**76**. We concluded that the efficiency of the two autocatalytic cycles are higher than the cross-catalytic cycle and as a result we were satisfied with the choice of two replicators to combine into one system.

## 2.8 Two Replicator Kinetic System



**Scheme 2.12:** Reagents and conditions (i) CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat) 10 mM, 10 °C, 16 h.

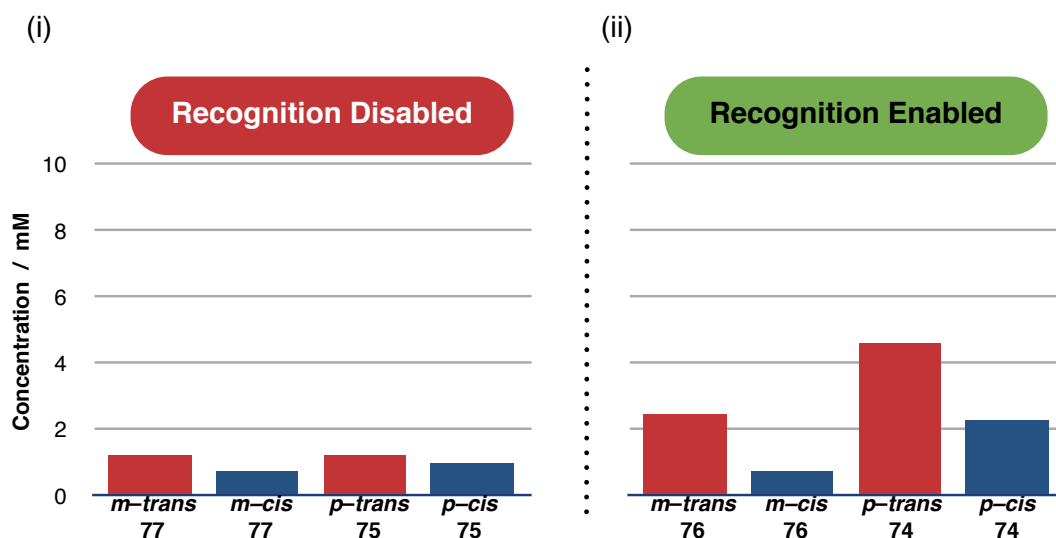
We have successfully demonstrated two new autocatalytic self-replicating systems which use the common maleimide **72**. We now wished to create a three component system of *m*-nitrone **68**, *p*-nitrone **52** and maleimide **72** in which the reactions to

form the two replicating templates compete for the common component (Scheme 2.12). In isolation, *p-trans*-**74** was a more efficient replicator possessing a higher *EM* and a reaction rate in the ternary complex 1.9 times faster than the *m-trans*-**76** system. As a consequence of the difference in efficiency of each of the replicators in isolation, we would expect to see a hierarchy formed with the more efficient replicator most abundant in the product pool, however this may be disrupted by the presence of one cross-catalytic pathway in the system. The system may be visualised graphically to show the same topology as in Figure 2.6.

Initially, we wished to investigate the rate of bimolecular reaction in the three component system. In order to achieve this, an equimolar mixture of the recognition disabled maleimide **73**, *m*-nitroene **68** and *p*-nitroene **52** at 10 mM in CD<sub>2</sub>Cl<sub>2</sub>/ *p*-TSA (sat) and incubated at 10 °C for 40 h using the standard experimental procedure as used before and described in Chapter 7, Section 7.2. After this time, the system was assayed by 376.5 MHz <sup>19</sup>F NMR spectroscopy and the relative concentration of each product determined by deconvolution and displayed in Figure 2.15 (i).

After 40 h, 41% of maleimide **73** had been consumed to form cycloadducts with the two *trans* cycloadducts *p-trans*-**75** and *m-trans*-**77** produced in near equal concentrations of 1.20 mM and 1.21 mM respectively. Their diastereoisomeric partner cycloadducts *p-cis*-**75** and *m-cis*-**77** were produced in lower concentrations of 0.94 mM and 0.73 mM respectively. These results demonstrated that in the absence of recognition processes there is low conversion and selectivity produced in the system.

The recognition enabled system was next investigated by creating an equimolar mixture of recognition enabled maleimide **72**, *p*-nitroene **52** and *m*-nitroene **68** and performing the reaction under the same conditions as before. After 40 h the reaction was assayed by the same method with the relative concentration of products determined by deconvolution and the results displayed in Figure 2.15 (ii). The results revealed that the system now reached quantitative conversion with the major product *p-trans*-**74** which reached a concentration of 4.6 mM, nearly half of the total product pool. The second replicating template, *m-trans*-**76**, is the second most abundant product in solution reaching 2.4 mM however the unexpected result is that *p-cis*-**74** reaches a modest concentration of 2.3 mM. The fourth cycloadduct, *m-cis*-**76** was produced in a low concentration of 0.73 mM.

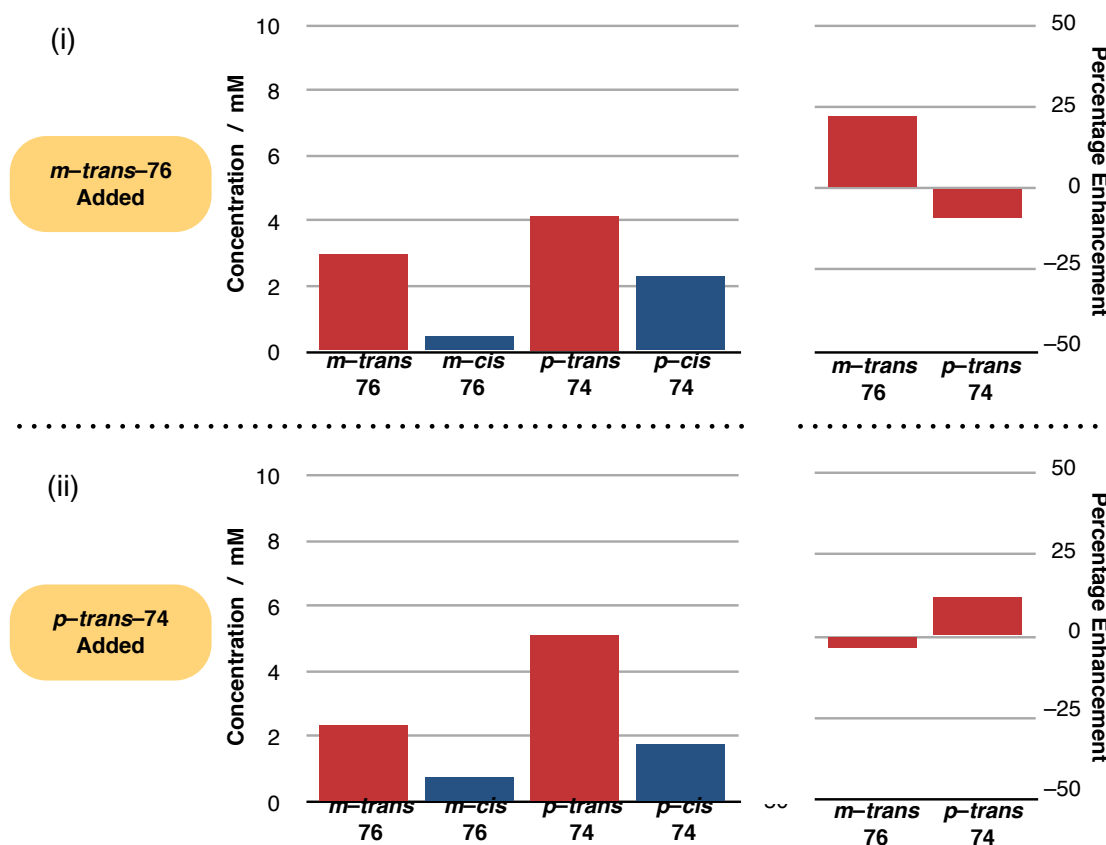


**Figure 2.15:** Concentration profile of the four cycloadducts after 40 h as determined by deconvolution of 376.5 MHz  $^{19}\text{F}$  NMR spectroscopy. (i) Results of the recognition disabled control reaction with maleimide **73**. (ii) Results of the recognition mediated reaction with maleimide **72**.

These results demonstrate that recognition mediated processes are responsible for increasing the production of three of the four cycloadducts. Two of these, *p-trans-74* and *m-trans-76*, are the replicating templates which are autocatalytically produced and selectively produce themselves from the mixture. The concentration of *p-cis-74* is enhanced as a consequence of *p*-nitrone **52** and maleimide **72** being able to form a reactive  $[\mathbf{A}\cdot\mathbf{B}]$  complex which promotes the formation of *p-cis-74* as described in Section 2.7.1. In isolation when *p*-nitrone **52** reacts with maleimide **72**, the autocatalytic cycle to form *p-trans-74* outperforms the  $[\mathbf{A}\cdot\mathbf{B}]$  complex channel which forms *p-cis-74*. In the two replicator system however, the increased concentration of recognition bearing components leads to a higher degree of competitive inhibition through the formation of unproductive ternary complexes. The formation of the binary complex does not rely on the formation of ternary complexes and is therefore not affected by the increased number of recognition bearing substrates. As a result, the formation of the product through the  $[\mathbf{A}\cdot\mathbf{B}]$  complex channel is up-regulated which is why there is an increase in the production of *p-cis-74*. As this  $[\mathbf{A}\cdot\mathbf{B}]$  complex channel is not template directed, we are unable to control the outcome of reaction down this pathway directly.

## 2.8.1 Template Instruction of Two Replicator System

We once again wished to take advantage of the template directed nature of autocatalytic self-replicators to instruct the system to produce one or the other template as desired by providing a doping of one of the two templates to direct the formation of itself. In order to achieve this, a fresh set of reagents for the two replicator system was doped with 20 mol% of one or the other pre-fabricated template. The reactions were then subjected to the same reaction conditions as before and assayed as before by  $^{19}\text{F}$  NMR spectroscopy.



**Figure 2.16:** Concentration of newly formed cycloadducts after 40 h after instruction with: (i) 20 mol% *m-trans-76* and (ii) 20 mol% *p-trans-74*. Concentrations were determined by deconvolution of  $^{19}\text{F}$  NMR spectra and subtraction of pre-added template.

The first instruction issued to the system was to make *m-trans-76*. The instruction was issued by doping a fresh set of reagents for the three component recognition mediated system with 20 mol% of preformed *m-trans-76* using the standard experimental procedure as described in Chapter 7, Section 7.2. The reaction was then incubated under the same conditions as before, assayed after 40 h by  $^{19}\text{F}$  NMR spectroscopy.

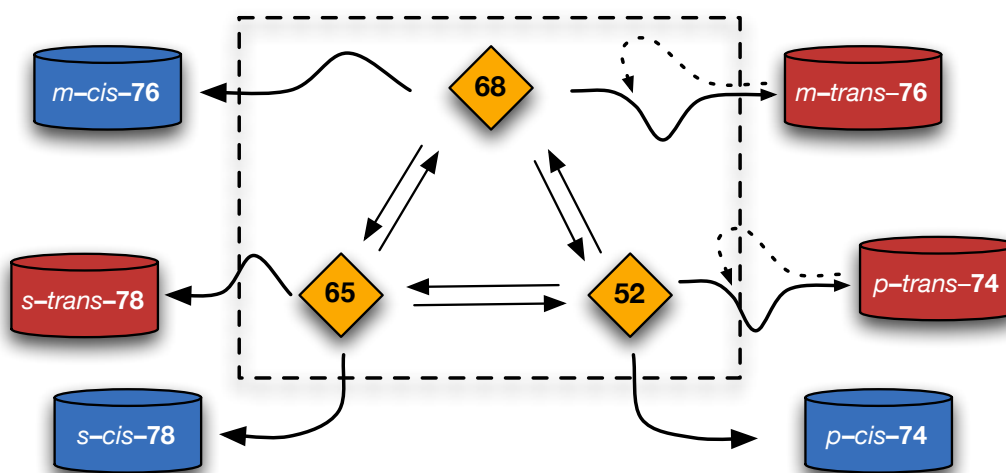
MHz  $^{19}\text{F}$  NMR spectroscopy and the relative concentration of each product calculated from deconvolution. The results (Figure 2.16 (i)) revealed the two self-replicating *trans* cycloadducts were once again the two major products with 2.98 mM of *m-trans-76* and 4.16 mM of *p-trans-74* produced. Closer inspection revealed that in comparison to the uninstructed system, there was a 21% enhancement in the production of *m-trans-76* and a 9% inhibition in the production of *p-trans-74*. These results followed our expectation that providing preformed *m-trans-76* would enhance the formation of itself at the expense of the other replicating template. The level of enhancement was not significant enough to make *m-trans-76* the dominant product in solution owing to the more efficient *p-trans-74* replicator. The concentrations of the two *cis* cycloadducts are also inhibited owing to the more efficient consumption of the maleimide to form *m-trans-76*.

To demonstrate that the system is fully controllable, we next wished to instruct the system to enhance the production of *p-trans-74*. The instruction was once again issued by doping a fresh set of reagents for the three component recognition mediated system with 20 mol% of preformed *p-trans-74* using the same experimental procedure as described in Chapter 7, Section 7.2. The relative concentration of each product was determined by deconvolution of the appropriate resonances arising from each species and the concentration of products displayed in Figure 2.16 (ii). In this case, the major product of the system was *p-trans-74* which had been produced at a concentration of 5.12 mM. The second most abundant product was the second replicating template, *m-trans-76*, which reached a concentration of 2.34 mM. Closer inspection revealed there was the desired enhancement in the final concentration of *p-trans-74* with a decrease in the concentration of *m-trans-76*. In comparison to the uninstructed system, there was a 12% enhancement in the production of *p-trans-74* and a 3.5% inhibition in the production of *m-trans-76*. The level of enhancement after this instruction was not as large as the *m-trans-76* instruction, possibly as a result of the cross-catalytic pathway in the system, however, the desired behaviour of a controllable two replicator system has been demonstrated.

## 2.9 Coupling Two Replicator System to a Dynamic Process

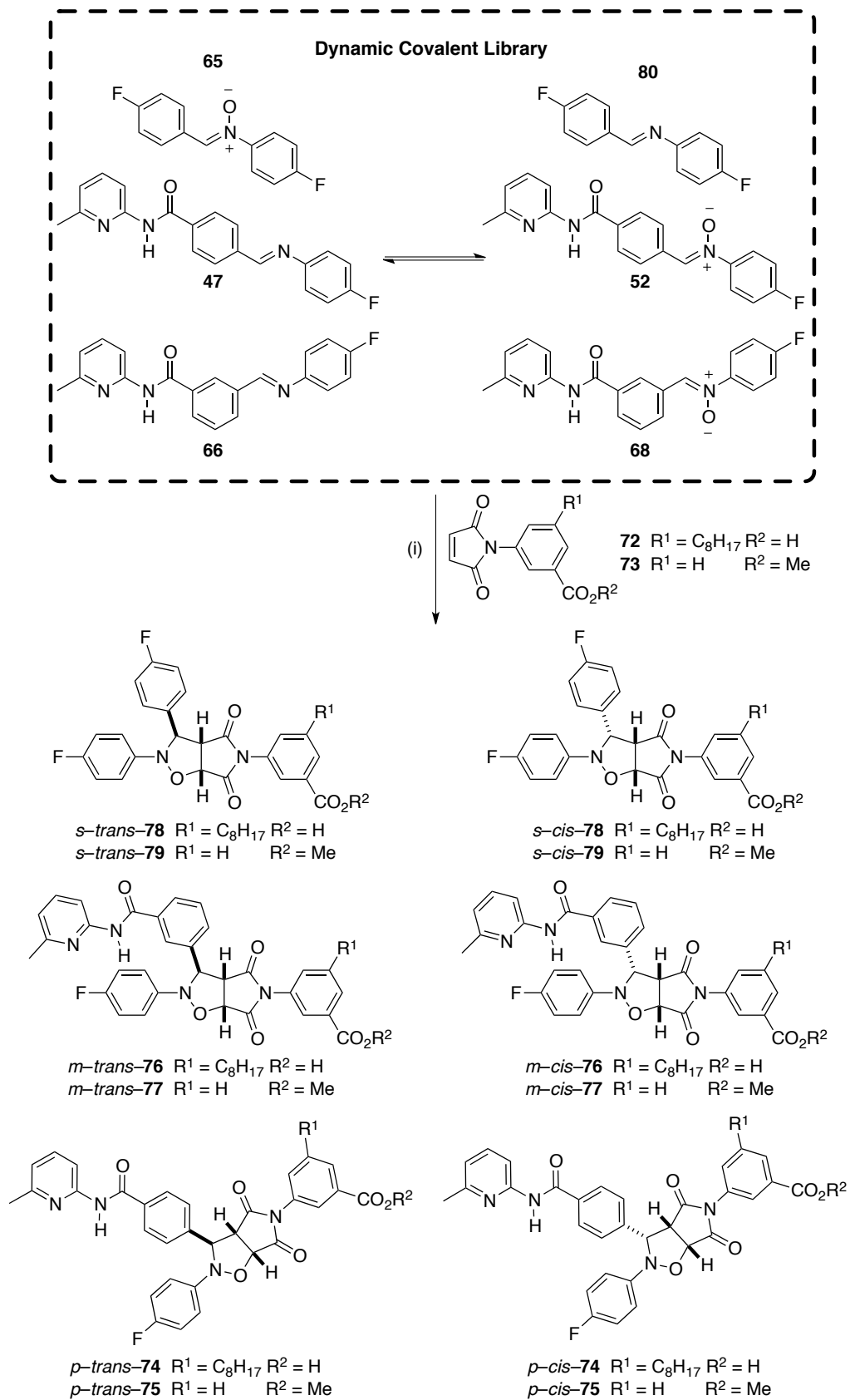
As described previously, Section 2.5, the two replicator system can be coupled to a dynamic process which has the effect of slowly feeding nitron reagents into the system. In our previous example with the glycine derived maleimide, coupling the system to a dynamic process lead to a system in which the two replicators selectively produced themselves from the mixture. The dynamic reagent pool system was able to be directed as desired by instruction leading to greater enhancements in the dynamic system than in the kinetic system when the two nitrones were provided at the start. We would therefore like to apply this protocol to this new system with the new maleimide.

The new system as demonstrated in Scheme 2.13 was conceived and initially the native reactivity was investigated by comparison of a recognition disabled system using the maleimide ester with a recognition enabled system with the free acid maleimide. The system may be visualised graphically as in Figure 2.17.



**Figure 2.17:** Representation of the system in which the three reactive reagents within the black dashed line exchange under thermodynamic control and will each react irreversibly to produce their two diastereoisomeric products. Two of these product are produced autocatalytically as represented by the second arrow looping back into the dynamic box. The other components of the dynamic library are omitted for clarity.

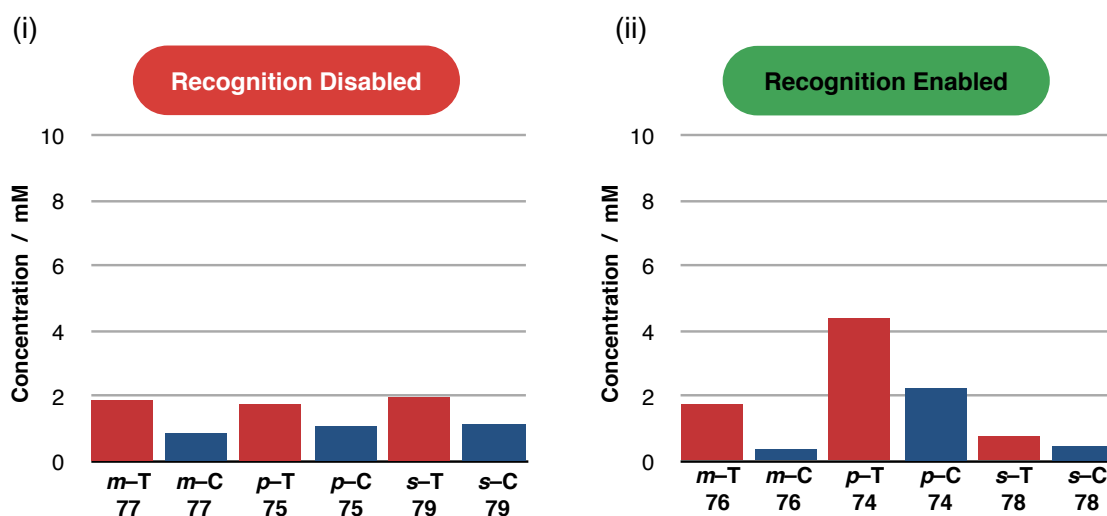
The recognition disabled reaction was performed by mixing four components, *p*-imine **47**, *m*-imine **66**, *s*-nitron **65** and maleimide ester **73** at 10 mM, in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat) and incubating at 10°C for 5 days (120 h), see Chapter 7, Section 7.2 for experimental procedure. The time span for the reaction was chosen to ensure



**Scheme 2.13:** Reagents and conditions (i)  $\text{CD}_2\text{Cl}_2/p\text{-TSA}$  (sat) 10 mM, 10 °C, 16 h. The compounds enclosed within the black dashed line are part of a dynamic library.

complete conversion of maleimide to product in the recognition enabled systems to follow. After 5 days, the system was assayed by 376.5 MHz  $^{19}\text{F}$  NMR spectroscopy and the relative concentrations of products determined by deconvolution. The results, displayed in Figure 2.18 (i), revealed that after 5 days, there was 87% conversion of maleimide to cycloadducts. There was a near equal concentration of each of the three *trans* cycloadducts as well as slightly lower similar concentrations of each of the *cis* cycloadducts. This result was expected as with no recognition mediated reactions prevalent in the system each of the three cycloadditions should proceed at its bimolecular rate which is similar in each case.

The recognition mediated system was next investigated by mixing four components, *p*-imine **47**, *m*-imine **66**, *s*-nitron **65** and free acid maleimide **72** at 10 mM in  $\text{CD}_2\text{Cl}_2/p\text{-TSA}$  (sat) and incubated at 10°C for 5 days (120 h) as before, see Chapter 7, Section 7.3 for experimental procedure.



**Figure 2.18:** Concentration of cycloadducts after 120 h in (i) recognition disabled system using maleimide ester **73** and (ii) recognition enabled system using free acid maleimide **72**.

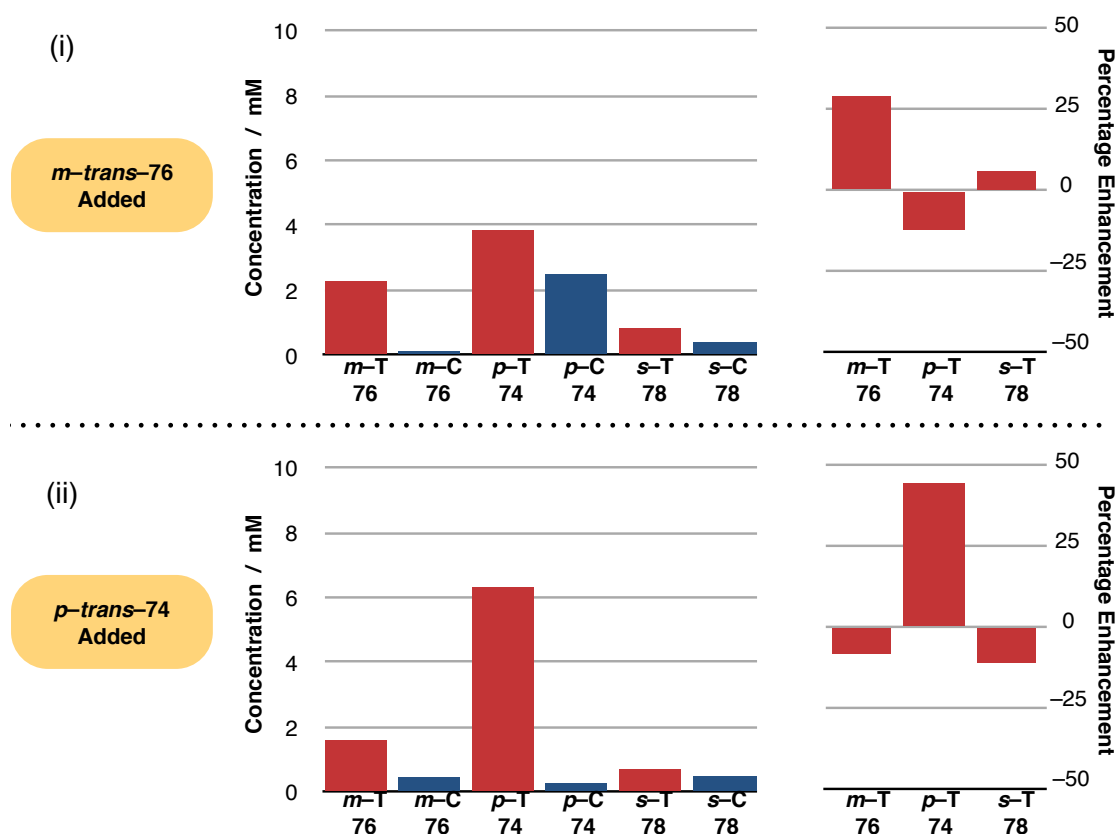
Upon completion, the system was assayed by 376.5 MHz  $^{19}\text{F}$  NMR spectroscopy and the concentration of products calculated after deconvolution (Figure 2.18 (ii)). The results revealed that the major product in the system was the more efficient replicator *p*-*trans*-**74** of which 4.38 mM was produced. The next most abundant product was now the product of the **[A•B]** complex channel product *p*-*cis*-**74** which reached a concentration of 2.25 mM. The next most abundant product was the second replicating template *m*-*trans*-**76** of which 1.75 mM was produced. The remaining



products, all of which produced though bimolecular processes only were produced in lower concentrations. The results demonstrate that in this case, the effect of slowly feeding in nitrene through a dynamic process has been to enhance the prevalence of the [A•B] complex channel as it now becomes the second most abundant product in solution. The most efficient replicator, *p-trans*-74, was still the major product from the mixture.

### 2.9.1 Template Direction of Dynamic Two Replicator System

With the native distribution of products established in the recognition enabled system, we now wished to demonstrate control of the system by providing preformed template to direct the formation of one or the other product.



**Figure 2.19:** Concentration of newly formed cycloadducts after 120 h after instruction (i) instruction with 20 mol% *m-trans*-76 and (ii) instruction with 20 mol% *p-trans*-74. Concentrations were determined by deconvolution of  $^{19}\text{F}$  NMR spectra and subtraction of pre-added template.

The first instruction issued to the system was to make *m-trans*-76 and the instruction was issued by adding 20 mol% of preformed *m-trans*-76 to a fresh set of the four

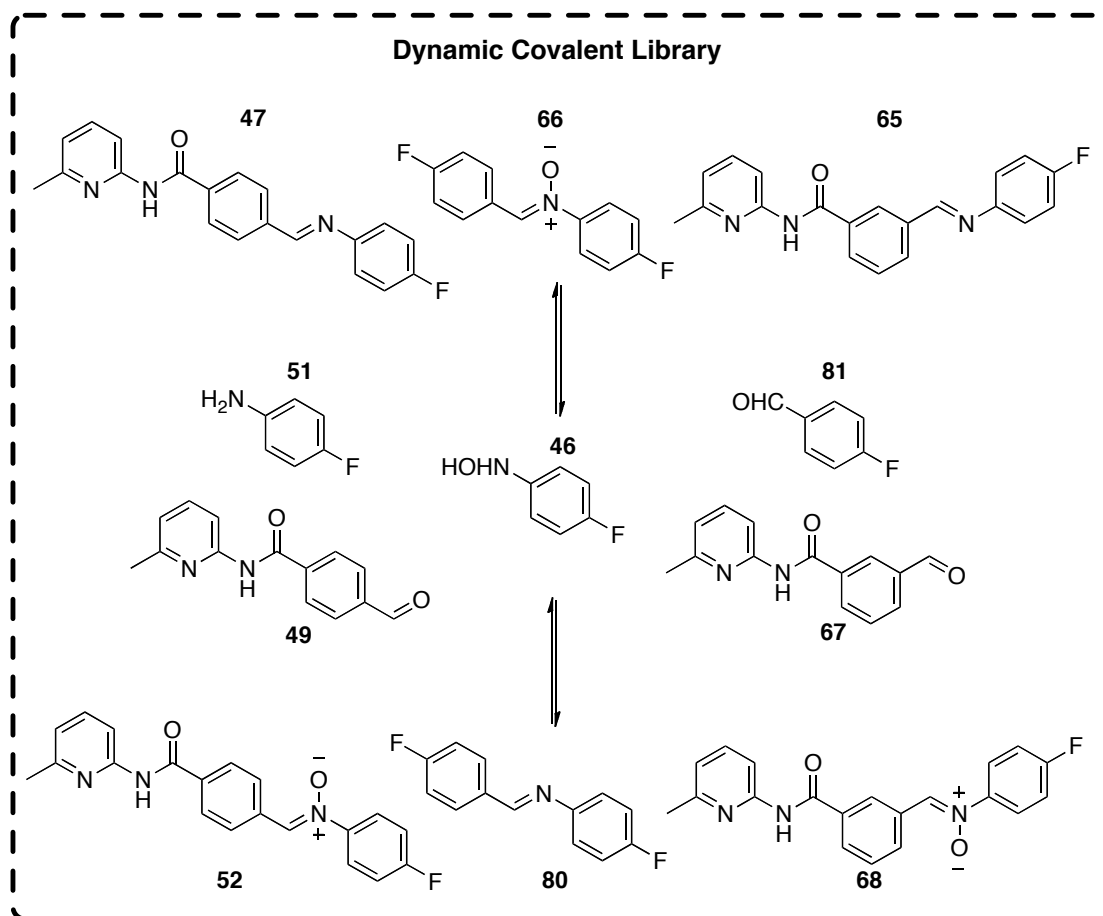
reagents, *p*-imine **47**, *m*-imine **66**, *s*-nitron **65** and maleimide **72** at 10 mM in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat) and incubated at 10°C for 5 days (120 h) and then assayed before, see Chapter 7, Section 7.3 for experimental procedure. The results (Figure 2.19 (i)) revealed a different distribution of products from the uninstructed system. The concentration of newly formed *m-trans*-**76** had risen to 2.26 mM, with the concentration of *p-trans*-**74** falling to 3.84 mM representing a 29% increase in the production of *m-trans*-**76** with 12% inhibition in the production of *p-trans*-**74**.

Next, to demonstrate that the network is able to respond to a different instruction, 20 mol% of preformed *p-trans*-**74** was added to a fresh set of the four reagents and the mixture incubated and assayed as before. The results (Figure 2.19 (ii)) once again revealed a different distribution of products from the uninstructed system. The major product of the system was newly formed *p-trans*-**74** which had been produced at a concentration of 6.33 mM, with the concentration of *m-trans*-**76** falling to 1.61 mM. Comparison of these results to the uninstructed system revealed a 44% increase in the production of *p-trans*-**74** with 11% inhibition in the production of *m-trans*-**76**. The doping also had the effect of down-regulating the prevalence of the [A•B] complex channel reaction of *p*-nitron **52** with maleimide **72** assembling to form [52•72] before reaction to furnish *p-cis*-**74**. The concentration of *p-cis*-**74** had fallen to 0.28 mM in comparison to the uninstructed system in which 2.25 mM was produced representing an 88% inhibition.

From these results, it is clear that the system created of two interlinked replicators coupled to a dynamic process was once again able to, first of all, coexist with both replicating templates produced in greater quantities over simple bimolecular processes, although the existence of a [A•B] complex channel served to upregulate the formation of *p-cis*-**74**. The system was then able to be successfully instructed to upregulate the formation of whichever of the two replicating templates was desired using a doping of preformed template. The higher efficiency of the autocatalytic cycle to produce *p-trans*-**74** in comparison to *m-trans*-**76** meant that in all cases, *p-trans*-**74** was the major product of the system and when *p-trans*-**74** was used as an instruction the level of enhancement was very strong. Once again in comparison to the kinetic system (Section 2.8 + 2.8.1) the level of enhancement of each template when instructed is greater than when the system is coupled to a dynamic process.

## 2.10 Direction of a Mixture of Simple Components

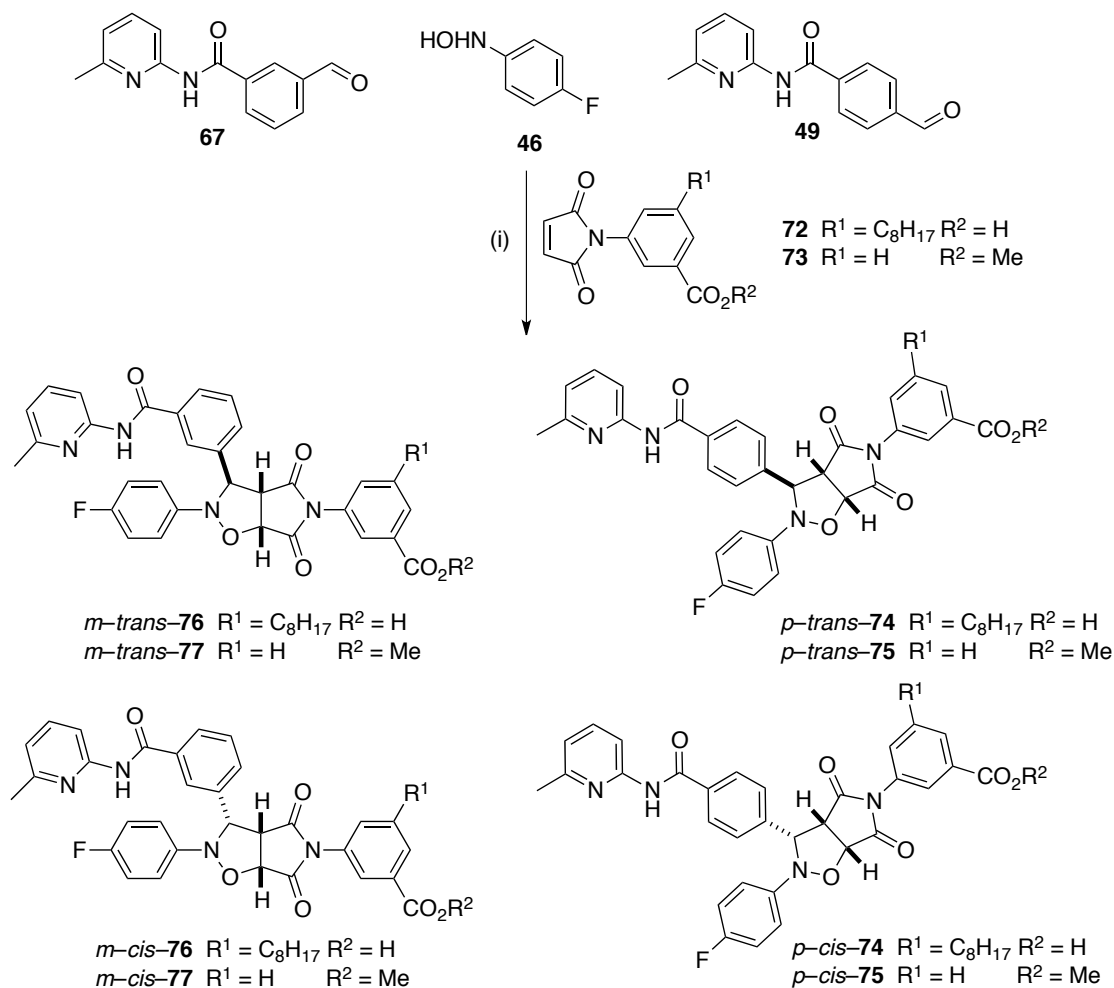
During the operation of the two replicator system coupled to a dynamic process, the success of the system relies on the ability of the nitrones and imines to be able to exchange their functionalities through the dynamic process. The nitron and imines present at the start of the reaction are able to hydrolyse to their hydroxylamine, amine and aldehyde components aided by the water and acid in the solvent (Scheme 2.14). Recombination of these hydrolysed components or direct attack of hydroxyamine on the imines will lead to the formation of the two new nitrones and simple imine setting up a dynamic library of 11 components.



**Scheme 2.14:** Under the conditions of the dynamic system ( $\text{CD}_2\text{Cl}_2/p\text{-TSA (sat)}$ ) imines are able to hydrolyse to their building blocks before recombination leads to newly formed imine and nitron. The compounds enclosed within the black dashed line are part of a dynamic library.

We have demonstrated upon mixing two recognition bearing imines, *p*-imine **47** and *m*-imine **66** and *s*-nitron **65** along with maleimide **72**, the major products formed are those from recognition mediated processes from nitrones bearing recognition

sites which have been formed through dynamic exchange. This results indicates that the rate of exchange must exceed the rate of cycloaddition in the bimolecular processes.

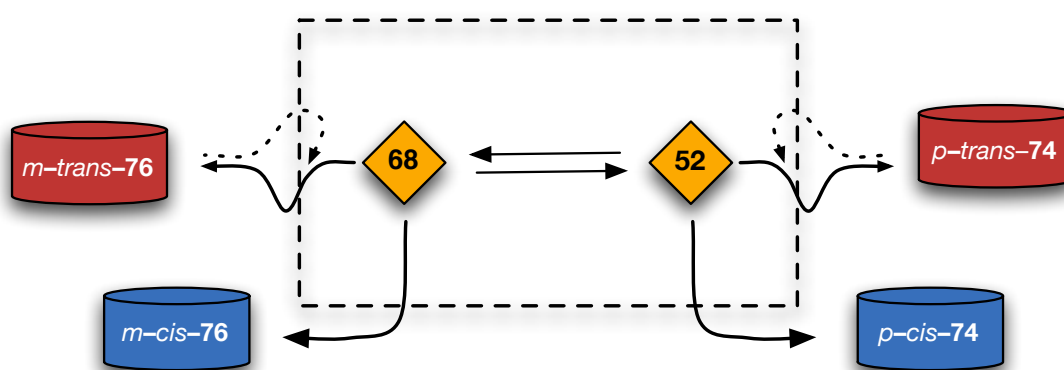


**Scheme 2.15:** Reagents and conditions (i)  $CD_2Cl_2/p$ -TSA (sat), 10 mM, 10 °C, 40 h.

We therefore envisaged that the two replicator system could be designed to initiate from a mixture of only its subcomponents required for the formation of nitrones, that is upon mixing two aldehydes **49** and **67** with a hydroxylamine **46**. From this mixture, the two recognition bearing nitrones, *p*-nitron **52** and *m*-nitron **68**, will be formed and as only a single equivalent of hydroxylamine is used, an equilibrium of the two nitrones will be formed with the ratio of their concentrations dependent of their relative stability. The nitrones can subsequently be consumed by cycloaddition reaction with maleimide **72**. As the rates of reaction of the two nitrones with maleimide **72** are not equal, the original equilibrium of the two nitrones will be

perturbed and the system will then work to re-establish the original equilibrium in accordance with Le Châtelier's principle. As the rate of cycloaddition reaction to form the two replicating templates can be influenced by doping either of the preformed templates into the reaction upon initiation, we expect that the outcome of the system should be controllable once again.

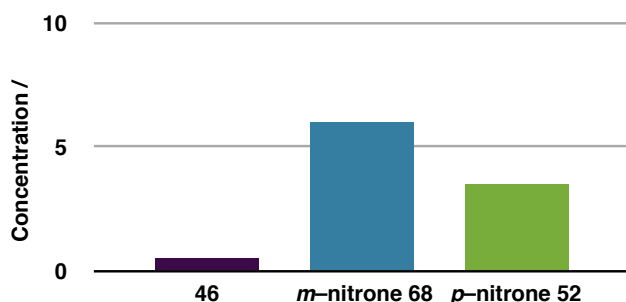
We can again visualise this system in a graphical manner as in Figure 2.20. The visualisation reveals that there are now only two nitron nodes which can each produce a *trans* template autocatalytically and a *cis* template by bimolecular reaction which means there are less kinetic traps for the reagents in this system in comparison to Figure 2.17.



**Figure 2.20:** Representation of the system in which the two reactive reagents within the black dashed line exchange under thermodynamic control and will each react irreversibly to produce their two diastereoisomeric products. Two of these product are produced autocatalytically as represented by the second arrow looping back into the dynamic box.

Initially we wished to characterise the outcome of simply mixing the two aldehydes **67** and **49** with hydroxylamine **46** to observe the relative concentrations of each of the two nitrones establishing a baseline behaviour for the dynamic reaction. In order to achieve this, an equimolar solution of aldehydes **67** and **49** and hydroxylamine **46** at 10 mM in  $\text{CD}_2\text{Cl}_2/p\text{-TSA}$  (sat) was created. The mixture was incubated at 10 °C for 24 h and after this time was assayed by 376.5 MHz  $^{19}\text{F}$  NMR spectroscopy. The relative concentration of the hydroxylamine and two nitrones was then determined (Figure 2.21) by deconvolution of the appropriate resonance arising from each of the species and revealed the equilibrium point of the mixture (incubation for 48 h and 72 h showed identical results). At this point, 95% of the hydroxylamine had been incorporated into a nitron. The major product in the library was *m*-nitron **68** which existed at a concentration of 6 mM with *p*-nitron **52** accounting for 3.5 mM of the

mixture with 0.5 mM of hydroxylamine **46** remaining. These results were interesting as the major product, *m*-nitrone **68**, is the starting material for the production of *m*-*trans*-**76**, the less efficient of the two replicators.

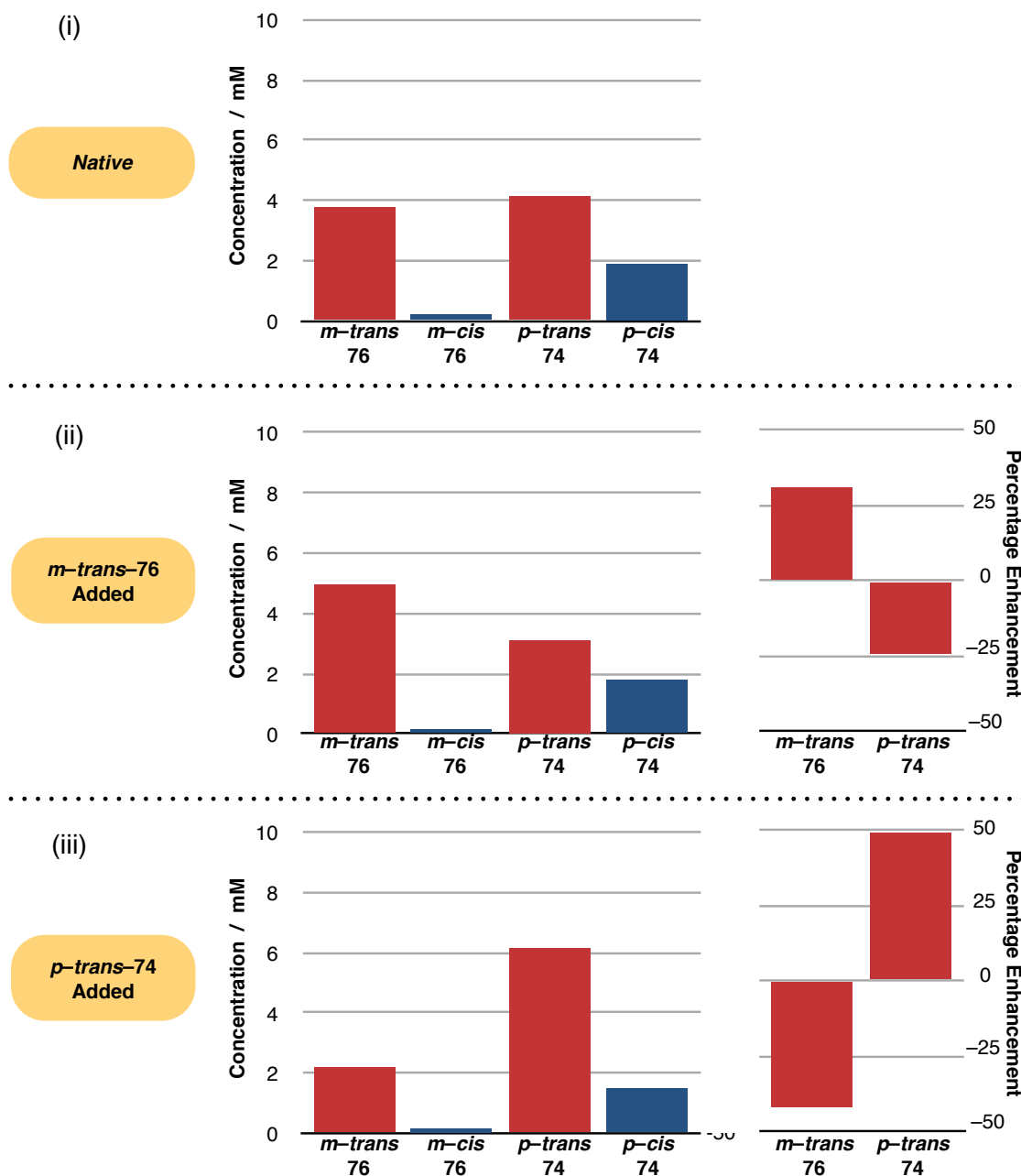


**Figure 2.21:** Relative concentration of hydroxylamine **46**, *m*-nitron **68** and *p*-nitron **52** after 24 h in an equimolar mixture of the three components at 10 mM in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat) at 10 °C.

In order to see what effect this equilibrium would have on the outcome of the relative concentration of the two replicating templates, we next coupled this mixture to the 1,3 dipolar cycloaddition reaction by introducing maleimide **72** to a fresh mixture of aldehydes **67** and **49** and hydroxylamine **46**, for experimental details, see Chapter 7, Section 7.4. An equimolar mixture of aldehydes **67** and **49**, hydroxylamine **46** and maleimide **72** at 10 mM in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat) was incubated at 10 °C for 40 h after which time the system was assayed by 376.5 MHz <sup>19</sup>F NMR spectroscopy and the relative concentration of products determined by deconvolution of the appropriate product resonance signals.

The results for the native<sup>†</sup> system (Figure 2.22 (i)), revealed the two major products of the reaction were the two replicating templates, *p*-*trans*-**74** and *m*-*trans*-**76** which are now produced in similar concentrations of 4.12 mM and 3.77 mM respectively. In this case, for the first time, the production of *p*-*trans*-**74** does not dominate the product pool. This result of the two replicators reaching similar concentrations, even though the production of *p*-*trans*-**74** is ~1.9× more efficient, is a result of the imbalance in the concentrations of the two reagent nitrones in the system as described above (Figure 2.21) with there being an abundance of reagent to produce the *m*-*trans*-**76** template, i.e. the result was largely a consequence of the ratio of the equilibria between nitrones  $K_{68} > K_{52}$  as opposed to the difference in rates  $k_{m-trans-76} < k_{p-trans-74}$  seen thus far.

<sup>†</sup> The term Native, used throughout this thesis, is used to describe a recognition mediated reaction which is performed with no added template.



**Figure 2.22:** Concentration of products after 40 h incubation at 10 °C for: (i) native reaction (ii) 20 mol% *m-trans*-76 instruction (iii) 20 mol% *p-trans*-74 instruction.

With the native distribution of products in the system established, the ability of the system to respond to instruction was next investigated in a manner similar to before. Firstly, we wished to issue the instruction to the system to make *m-trans*-76 with this command achieved by doping a fresh set of reagents, aldehydes **67** and **49**, hydroxylamine **46** and maleimide **72** with 20 mol% *m-trans*-76. The mixture was then incubated under the same conditions for 40 h and assayed by 376.5 MHz  $^{19}\text{F}$  NMR spectroscopy as before. The results (Figure 2.22 (ii)) revealed quantitative

conversion once again and the major product of the mixture was newly formed *m-trans-76* with 4.94 mM produced. The next most abundant product was expectedly the second replicating template *p-trans-74* which had been produced at a concentration of 3.11 mM. Comparison of these results to the native system revealed a 31% enhancement in the production of *m-trans-76* with concurrent 24% inhibition in the production of *p-trans-74*. Clearly the system has responded to the instruction to make *m-trans-76* with our doping and produced the desired result.

We next wished to demonstrate that the system was still able to respond to the second instruction of upregulating *p-trans-74*, which was issued by doping 20 mol% of preformed *p-trans-74* to a fresh set of the four reagents and incubated under identical conditions as before. After 40 h, the mixture was assayed as before to produce a concentration profile of the products (Figure 2.22 (iii)). At first glance, it is clear to see that *p-trans-74* has been enhanced as it dominates the product pool reaching a concentration of 6.1 mM. The second replicating template, *m-trans-76* is the next most abundant product and reaches a concentration of 2.19 mM. Comparison of these results to the native system revealed a significantly high enhancement in the production of *p-trans-74* of 49% with a 42% inhibition in the production of *m-trans-76*.

Clearly these results demonstrate that a system of two replicators is able to be produced by mixing their simple aldehydes, hydroxylamine and maleimide components and allowing the dynamic formation of the nitron to occur *in situ*. Furthermore, these results demonstrated that a greater enhancement in products in response to instructions can be reached using simpler components. The results also reveal that the outcome of the ratios of the two replicating templates, which in all cases thus far were dominated by *p-trans-74*, could be influenced by the distribution of reagents in the nitron dynamic pool. In this case, the nitron required for the less efficient replicator was the more stable product in the dynamic pool and thus existed at a higher concentration. This property of the dynamic library was translated into the final kinetic product pool as the less efficient replicator is able to be produced at higher concentrations as a result of it having a greater concentration of its resources.

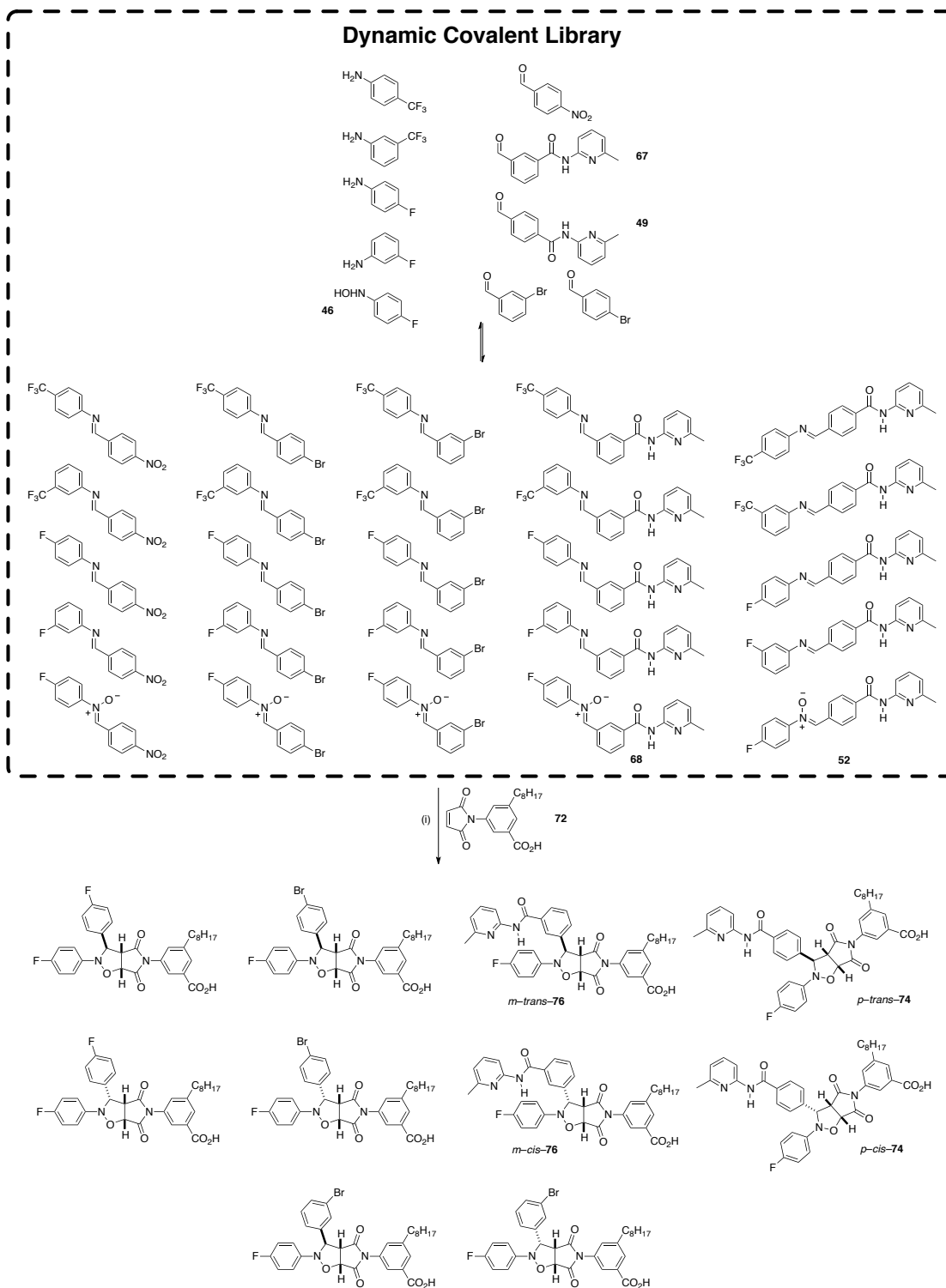


## 2.11 Extended System of Dynamic Reagents

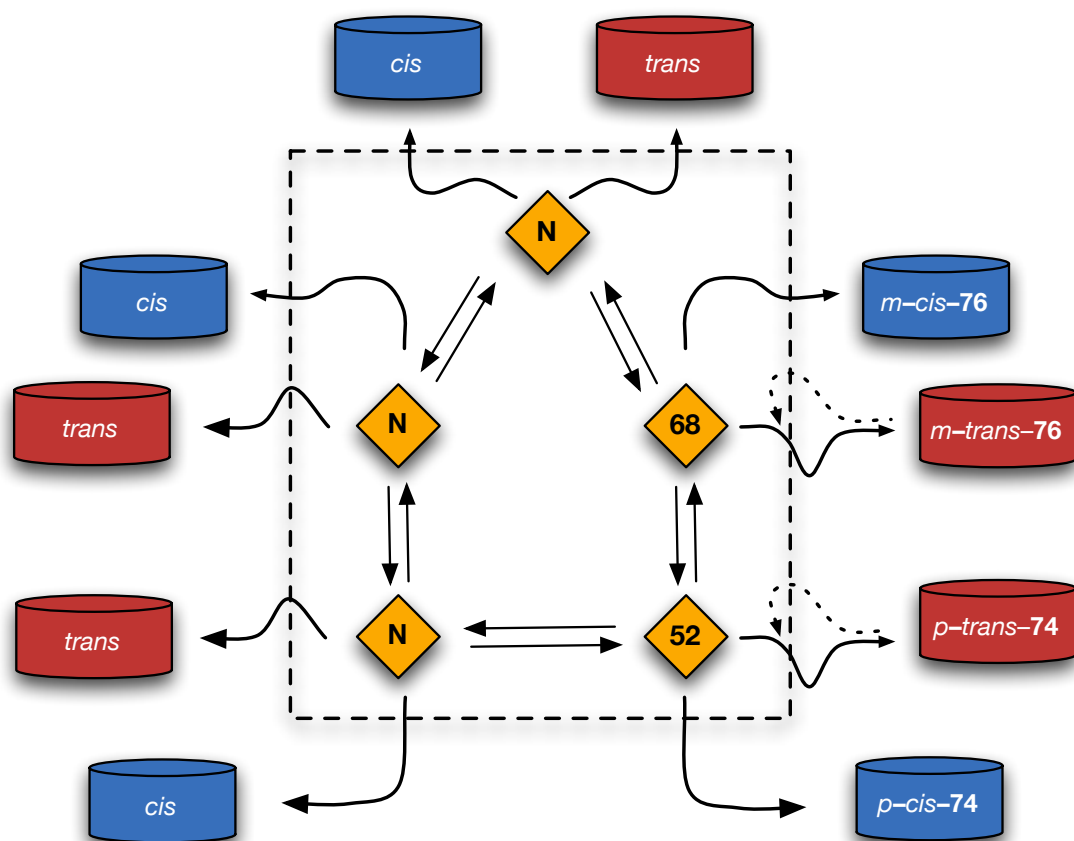
We have so far demonstrated that a network of two replicators can be created (Section 2.8) that are able to coexist and respond to an instruction (Section 2.8.1). We have subsequently demonstrated that increasing the complexity of the system by introducing dynamic processes acting on the reagents (Section 2.9 and 2.10) is also able to be utilised without disrupting the ability of the two replicators to efficiently produce themselves and in fact the presence of these dynamic processes can lead to systems which give greater enhancements than the native kinetic system of two nitrones and a maleimide (Section 2.8). We now wished to extend the mixture of reagents used to increase the complexity of the network further and observe if the two replicator system is robust enough to exist and be instructable in the face of a more complex system.

To achieve this goal, the dynamic pool of reagents was extended by increasing the number of components. Mixing five aldehydes, four anilines, one hydroxylamine and maleimide under the same conditions as the above systems, (10 mM in  $\text{CD}_2\text{Cl}_2$ ,/*p*-TSA(sat)), leads to the formation of a dynamic library consisting of 35 components (Scheme 2.16). Upon mixing the eleven reagents in the system, 20 imines are produced along with 5 nitrones which are capable of further reaction with the maleimide to form 10 cycloadducts, 5 *trans* and 5 *cis*. In the dynamic library, two of the nitrones are the recognition bearing *p*-nitrone **52** and *m*-nitrone **68**, the building blocks for the two replicating templates *p-trans*-**74** and *m-trans*-**76**. We also recognised that the formation of one of the *cis* cycloadducts, *p-cis*-**74**, is also promoted through a recognition mediated [A•B] complex channel and with the increased complexity and competition for material, it may be expected that this binary process will become more prevalent with the decrease in efficiency in ternary complexes as seen in previous systems. The remaining *trans* and *cis* cycloadducts will be formed through the bimolecular reaction and as a result could not be controlled by any template instruction.

The system may be visualised as in Figure 2.23 which shows that in this case, there are a significant number of kinetic traps in which the reagent pool may be irreversibly consumed through. Only two pathways lead to self-replicating templates and thus there is a much greater degree of competition in the system.



**Scheme 2.16:** Upon mixing the 5 aldehydes with 4 anilines and hydroxylamine, a dynamic library of 35 components is produced. Each of the nitrones formed is able to react with maleimide **72** to form ten products, 5 *trans* and 5 *cis* with 2 of these *trans* templates the replicating templates *p-trans-74* and *m-trans-76*. The compounds enclosed within the black dashed line are part of a dynamic library. Reagents and conditions (i)  $\text{CD}_2\text{Cl}_2/\text{p-TSA}$  (sat),  $10^\circ\text{C}$ , 10 mM, 7 days.



**Figure 2.23:** Representation of the system in which the five reactive reagents within the black dashed line exchange under thermodynamic control and will each react irreversibly to produce their two diastereoisomeric products. Two of these products are produced autocatalytically as represented by the second arrow looping back into the dynamic box. The two previously described nitrones and cycloadducts are explicitly labelled whilst the others are referred to as **N**. Likewise only the two reported replicating templates are explicitly labelled. Whilst equilibrium arrows are shown between nitrones, there is no explicit pathway within the system and all nitrones will be interconverting with each of the others.

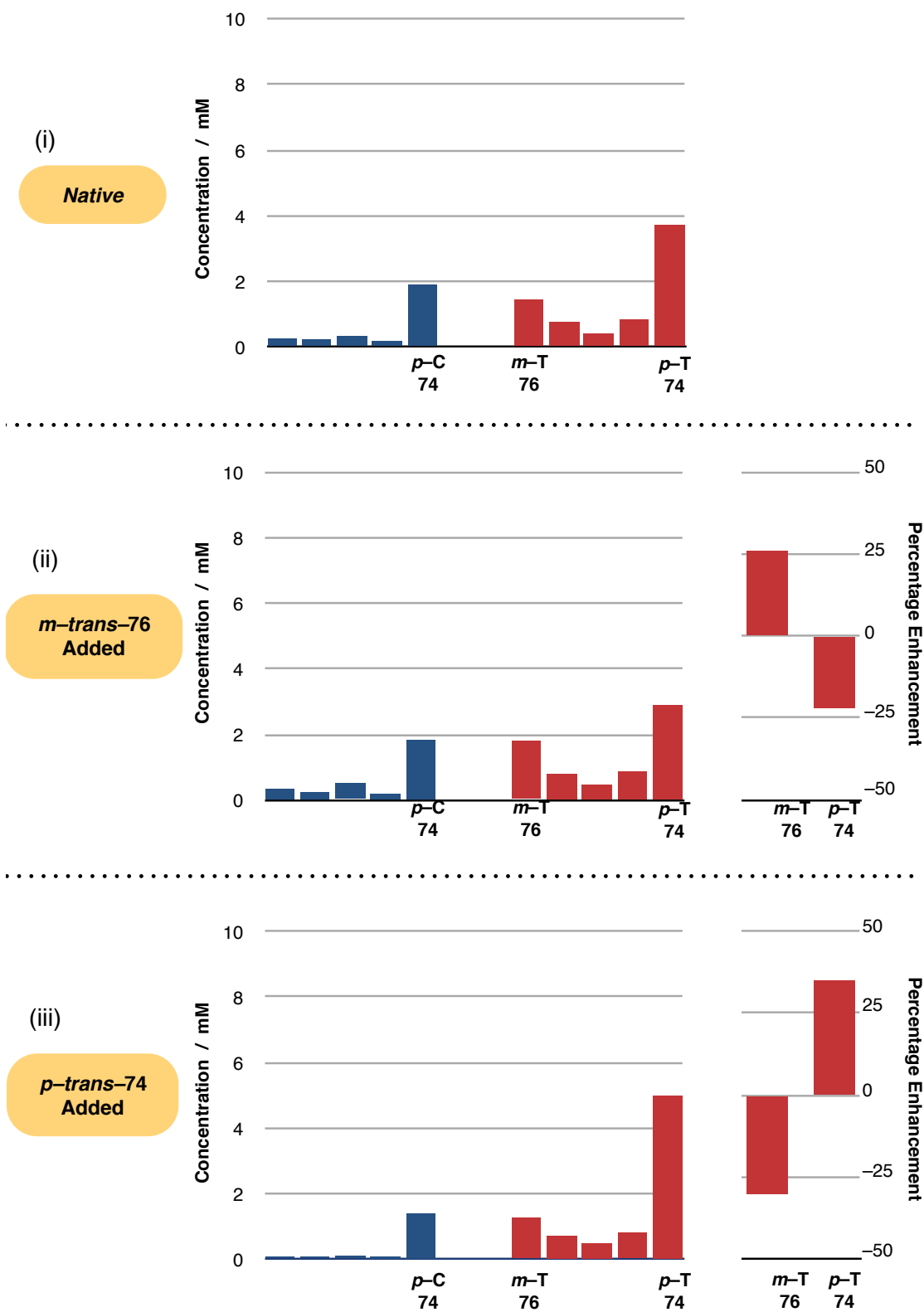
Upon creation of the system as above, see Chapter 7, Section 7.4 for full experimental details, the mixture would then be incubated at 10 °C for 7 days to ensure complete consumption of the hydroxylamine to cycloadduct after which time the mixture would be analysed by 376.5 MHz  $^{19}\text{F}$  NMR spectroscopy and the relative concentration of each product determined by deconvolution.

The results of devolution (Figure 2.24 (i)) revealed quantitative conversion with the major product from the system being *p-trans-74*, the more efficient replicator, which was produced to a concentration of 3.72 mM. The next most abundant product was *p-cis-74*, the product of the **[A•B]** binary complex channel which reached a concentration of 1.9 mM. In a more complex system, the **[A•B]** complex channel was

expected to thrive as seen previously in Section 2.9. The second replicating template, *m-trans-74*, had reached a concentration of 1.43 mM which was still almost double the concentration of the closest other *trans* cycloadduct (0.80 mM). By increasing the complexity of the system, we have reduced the relative efficiency of the recognition mediated processes proceeding through a ternary complex (replicators) in comparison to those proceeding through a binary complex (**[A•B]** complex channel). These results confirm that the two replicating templates are still able to be selectively produced from the extended mixture of reagents however the binary complex channel has become substantially more prevalent as a result of the increase in complexity.

We next attempted to investigate whether the system was still addressable by issuing the instruction to produce *m-trans-76* which was achieved by adding 20 mol% preformed *m-trans-76* to a fresh mixture of the 11 reagents and then incubating the mixture under identical conditions as before. Once again the mixture was assayed after 7 days by 376.5 MHz <sup>19</sup>F NMR spectroscopy and the relative concentration of products was determined by deconvolution and presented in Figure 2.24 (ii). The results revealed that the major product in the mixture was still *p-trans-74* however its production had been inhibited and only reached a concentration of 2.89 mM. The production of *m-trans-76* on the other hand had been enhanced as desired up to a concentration of 1.80 mM. Comparison of these results to the native system reveals a 26% enhancement in the production of *m-trans-76* with a 22% inhibition of the production of *p-trans-74*. The concentration of the binary complex product was not affected a great deal with the concentration reaching 1.84 mM which represents an inhibition of 3.5%.

Clearly the results demonstrated that the addition of preformed *m-trans-76* was responsible for directing the formation of itself at the expense of the other replicating template and thus demonstrating that the system was still addressable.



**Figure 2.24:** Distribution of *cis* (blue) and *trans* (red) cycloadduct products after 7 days as determined by deconvolution of the reactions when (i) no instruction given, results are of native recognition mediated system (ii) system doped with 20 mol% *m-trans-76* and (iii) system doped with 20 mol% *p-trans-76*.

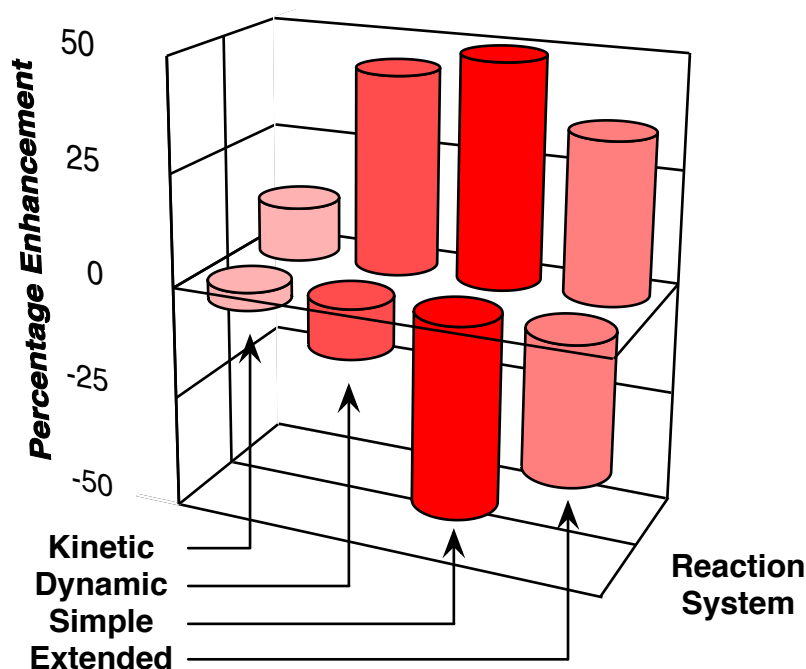
We now wished to demonstrate that the system would respond to a different instruction by doping with *p-trans-74* and observe if there was enhancement of itself. The instruction was once again issued by doping 20 mol% preformed *p-trans-74* to a fresh set of the 11 reagents and incubating the system under the same conditions for 7 days using the standard experimental protocol as described in Chapter 7, Section 7.4. After 7 days, the mixture was assayed exactly as before by 376.5 MHz  $^{19}\text{F}$  NMR spectroscopy and the relative concentrations of products determined by deconvolution. The results (Figure 2.24 (iii)) revealed that the major product was once again newly formed *p-trans-74* which now has increased in concentration to 5.00 mM dominating the product pool. The production of the second replicating template, *m-trans-76*, has been inhibited reaching a concentration of 1.26 mM. The production of the [A•B] binary reaction channel, *p-cis-74*, has also been inhibited reaching a concentration of 1.4 mM. Comparison of these results to the native system reveals the production of *p-trans-74* has been enhanced by 35% while the production of *m-trans-76* was inhibited by 30%. The [A•B] binary complex channel was also inhibited by 36%. These results demonstrated that the system was able to be directed to selectively form the second template by providing the necessary instruction.

Clearly, the series of results demonstrated that the two replicators within the extended library were able to selectively produce themselves from the mixture above the bimolecular products. The outcome of the distribution of products from the mixture was once again able to be directed towards either of the two replicating templates by addition of the appropriate preformed template to the reaction upon initiation.

## 2.12 Summary of Two Replicator Systems

We have successfully demonstrated that when mixing the reagents for two replicators (nitrones and maleimide) so that the two have to compete for a common component, the two replicators are able to coexist. This system was then able to be directed to selectively enhance the production of one or the other of the two replicating templates by addition of the appropriate preformed template. Following the lesson learned by the computational network simulation which showed that increasing the number of connections and overall complexity increased the relative

enhancement in selectivity, we increased the complexity of the system by introducing a dynamic processes (imine – nitron exchange) which acted on the reagents for the two replicators. The result of the dynamic reagent pool system was the creation of a network of two replicators which once again coexisted in the absence of any instruction with the product distribution able to be direct towards one or the other replicating template by the addition of the appropriate preformed template as an instruction and the level of enhancements in this dynamic reagent pool system were greater than in the simple kinetic system.



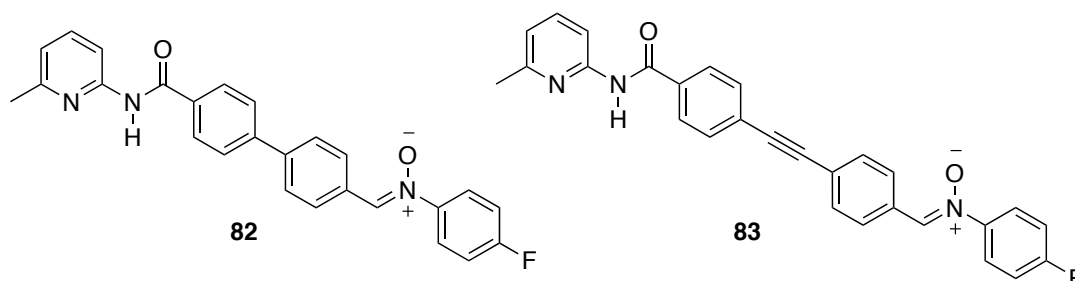
**Figure 2.25:** Comparison of enhancements for the results of the *p-trans-74* instructed systems. Kinetic represents the results of Section 2.8 when the two nitrones are mixed with a maleimide. Dynamic represents the results of Section 2.9 (mixture of imines, nitron and maleimide). Simple represents the results of Section 2.10. Extended represents the results of Section 2.11.

When the pool of reagents were connected to a dynamic reaction, a third reactive nitron was introduced to the system which acted as a further kinetic sink potentially lowering the possible enhancement which could be achieved. In order to avoid this kinetic sink yet still possess a dynamic pool of reagents to increase connectivity in the system, we next demonstrated that the simple components required to form the two nitrones could be used as reagents. These will then dynamically form the two nitrones *in situ* which themselves are precursors for the formation of the two self-replicating templates. In this case, the production of the two replicating templates

became almost equal despite their differences in efficiency. This observation was a consequence of the inherent reactivity mixture of simple components of the two aldehydes with the hydroxylamine. This system could also be directed by instruction with preformed template and lead the greatest observed enhancements in product distribution. Finally, we subjected the two replicators to an extended dynamic pool of reagents, forcing the replicating templates to find their reagents from a dynamic pool of 35 components. The two templates were able to selectively produce themselves above the concentration of products of bimolecular reactions demonstrating that they were indeed selectively producing themselves from this mixture however the enhancements upon instruction were not as great. This result is likely to be a consequence of the increased number of kinetic traps, that is five nitrones, each of which can irreversibly consume the maleimide. Addition of preformed templates to direct the formation of one or the other replicating template was carried out and was still able to selectively enhance the production of the desired template.

### 2.13 Design and Characterisation of a Third Self-Replicating System

Thus far, we have described the design and characterisation of two minimal self-replicating systems based on the reaction of a nitron with the common maleimide **72**. We now wished to increase the number of replicators in the system to three which would increase the complexity of the network allowing three possible instructions to be given in the form of each of their templates and desirably the network should be controllable and lead to three different outcomes in response to the identity of the added template.



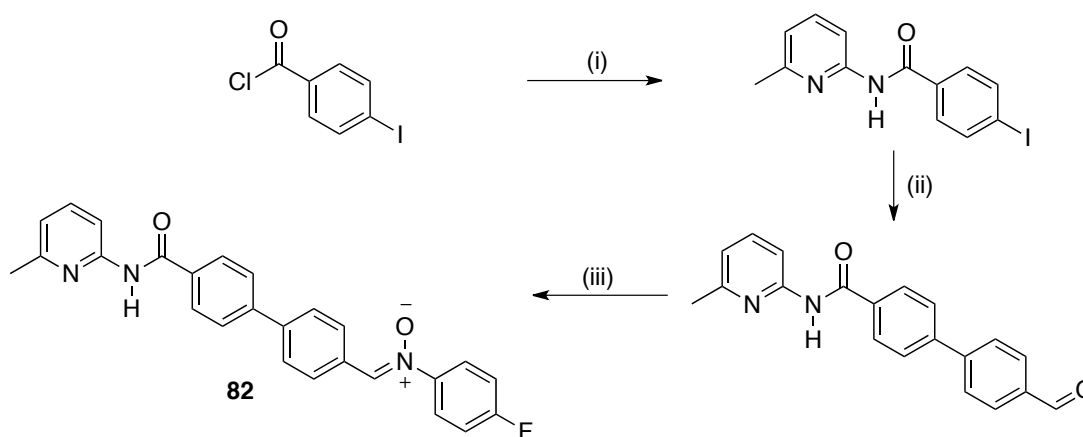
**Figure 2.26:** Nitron **82** based on a biphenyl group to increase the distance between the reaction centre and the recognition site and nitron **83** with an extended alkyne group installed.



In order to fulfil this goal, we require a third replicating template which is produced by the reaction of a new nitron with the common maleimide. Using experience obtained from earlier studies in the group of length segregated replicators,<sup>156</sup> we identified two nitrones, Figure 2.26, which were targeted for synthesis and subsequent investigation of their reaction with maleimide **72** to investigate for self-replicating behaviour. Both of the two new nitrones were designed to include a group which increased the distance between the amino pyridine recognition side and the nitron reaction. Subsequent reaction of the nitrones with maleimide **72** will produce templates which are longer and thus should not interact in a cross-catalytic manner with the reaction of the shorter nitrones **52** and **68** with maleimide **72**.

### 2.13.1 Synthesis of Biphenyl Nitron **82**

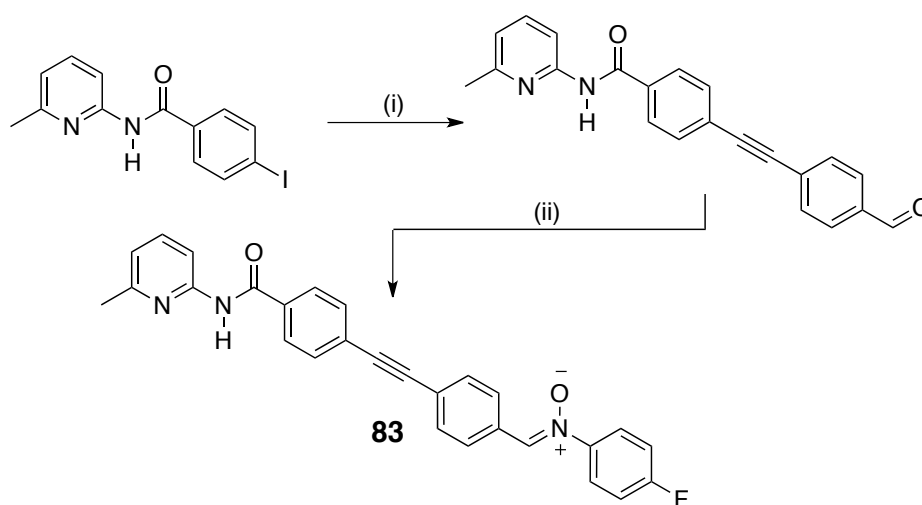
The synthesis of biphenyl nitron **82** began from an amide coupling of 4-iodobenzoyl chloride with 2-amino-6-methylpyridine under standard conditions with a little THF added to aid solubility of the acid chloride to provide the amide in a quantitative yield. The biphenyl aldehyde was produced in good yields after a Suzuki reaction of the aryl iodide with 4-formylphenylboronic acid. Subsequent condensation of the aldehyde with *N*-4-fluorophenylhydroxylamine **46** furnished nitron **82** which, alas, was near completely insoluble in non-polar solvents CD<sub>2</sub>Cl<sub>2</sub> and CDCl<sub>3</sub> which meant that it could not be used for investigation of the creation of a self-replicating system.



**Scheme 2.17:** Reagents and conditions (i) 2-amino-6-methylpyridine, NEt<sub>3</sub>, CH<sub>2</sub>Cl<sub>2</sub> / THF (4:1), 25 °C, 16 h, 99%, (ii) 4-formylphenylboronic acid, Pd(P(Ph)<sub>3</sub>)<sub>4</sub>, K<sub>2</sub>CO<sub>3</sub>, THF, 40 °C, 16 h, 75% (iii) **46**, EtOH, 25 °C, 16 h, 80%.

### 2.13.2 Synthesis of Long Nitron 83

Nitron **83**, lengthened using an acetylene spacer was synthesised starting from an iodoamide which was coupled to 4-ethynylbenzaldehyde by means of a Sonogashira reaction to produce the recognition bearing aldehyde in a good yield. Subsequent condensation of the aldehyde with hydroxylamine **46** in ethanol furnished nitron **83**, known as *l*-nitron **83** (*l* = long), in acceptable yields.



**Scheme 2.18:** Reagents and conditions (i) 4-formylphenylacetylene, Pd(P(Ph)<sub>3</sub>)<sub>2</sub>Cl<sub>2</sub>, CuI, PPh<sub>3</sub>, NEt<sub>3</sub>, 45 °C, 16 h, 75% (ii) **46**, EtOH, 25 °C, 16 h, 45%.

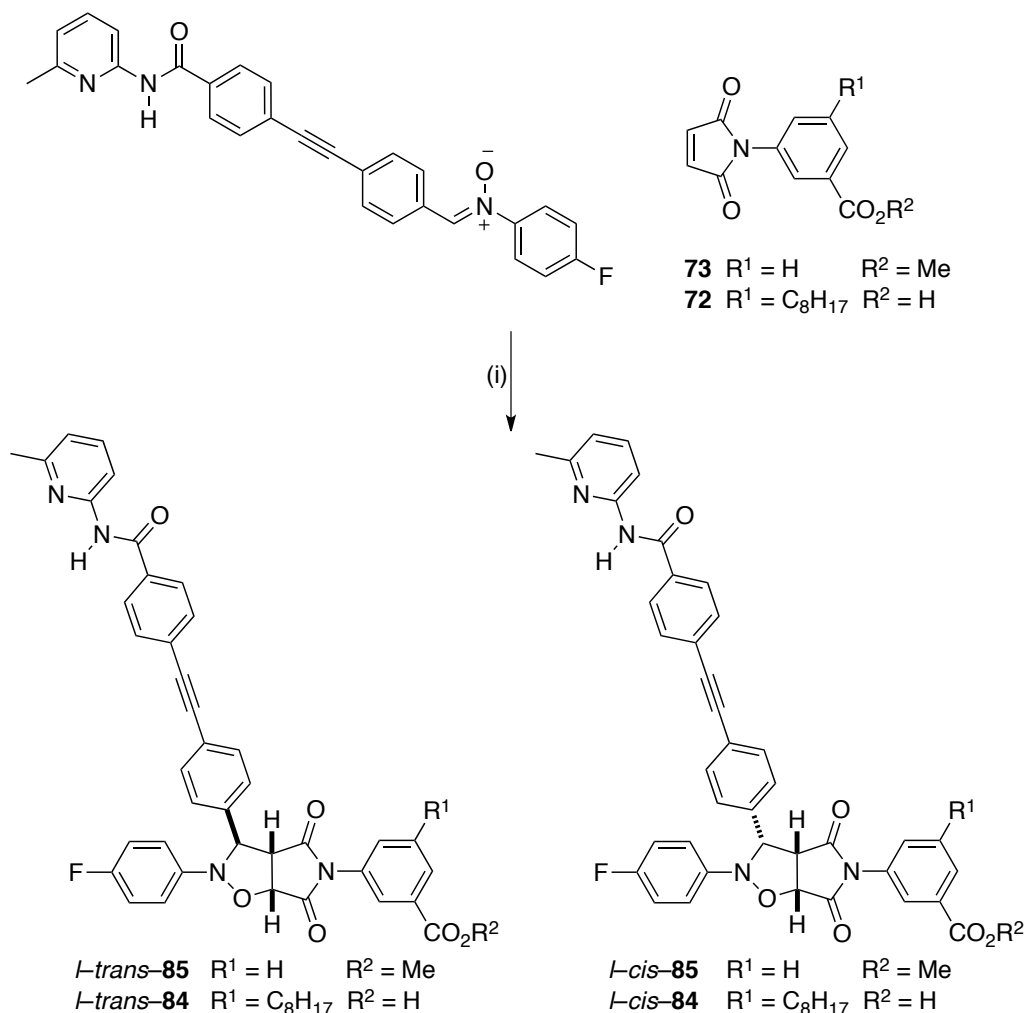
Pleasingly, *l*-nitron **83** was soluble in CDCl<sub>3</sub> and CD<sub>2</sub>Cl<sub>2</sub> at concentrations high enough for investigation of its reaction with maleimide **72**, that is, concentrations greater than 20 mM.

### 2.13.3 Analysis of the Reaction of *l*-Nitron 83 with Maleimide 72

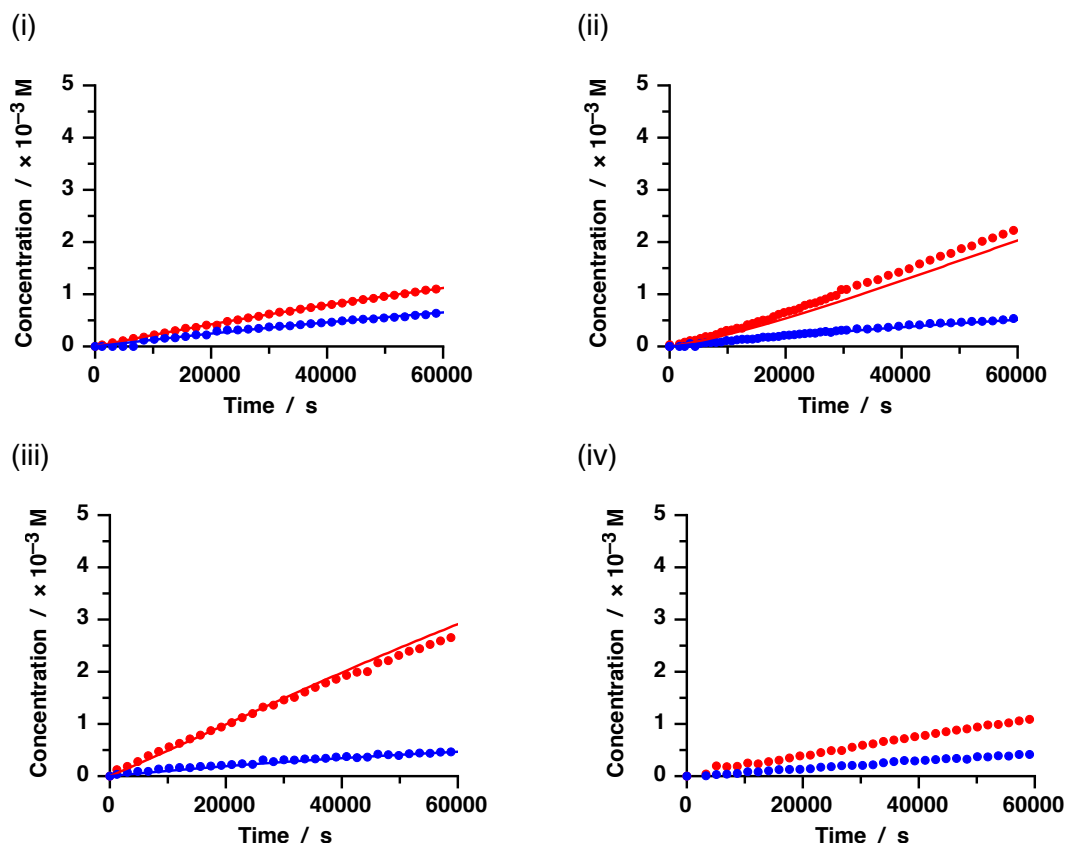
With a third nitron in hand, we set about to investigate whether its reaction with maleimide **72** proceeded through autocatalytic self-replication in the standard manner as utilised previously in this Chapter and under the same reaction conditions (10 mM, CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat), 10 °C, 16 h) as was used to investigate the previous reactions, see Chapter 7, Section 7.6 for experimental details.

Initially, the rate of bimolecular reaction in the system was investigated by monitoring the recognition disabled control reaction of *l*-nitron **83** with maleimide ester **73** using

470.3 MHz  $^{19}\text{F}$  NMR spectroscopy using the standard kinetic experimental procedure, Chapter 7, Section 7.6. The relative concentrations of species in the reaction were determined by deconvolution of appropriate resonances and a concentration vs. time profile, Figure 2.27 (i), was constructed. The results show that the recognition disabled reaction proceeded slowly reaching 17% conversion overall with the major product *l-trans*-**85** reaching 1.1 mM with its diastereoisomer *l-cis*-**85** reaching 0.6 mM representing a *trans* : *cis* ratio of 1.8:1. The results were subjected to kinetic simulation and fitting to a model for a bimolecular reaction using SimFit to extract rate information. The reaction rate constants for  $k_{l\text{-trans-85}}$  and  $k_{l\text{-cis-85}}$  were determined to be  $1.39 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  and  $0.44 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  respectively.



**Scheme 2.19:** Reagents and conditions (i)  $\text{CD}_2\text{Cl}_2/p\text{-TSA}$  (sat), 10 mM, 10 °C, 16 h



**Figure 2.27:** Concentration vs. time profiles for: (i) Control reaction of nitrone **83** with maleimide ester **73** (ii) Recognition mediated reaction of nitrone **83** with maleimide acid **72** (iii) Template doped reaction of nitrone **83** with maleimide **72** with 10 mol% *trans*-**84** (iv) Competitive inhibitor reaction of nitrone **83** with maleimide **72** and 5 eq. benzoic acid. Experimental data is represented as circles, *trans* cycloadducts (●), *cis* cycloadducts (●). The results of kinetic simulation and fitting using a model for a self-replicating system is shown as lines *trans* cycloadducts (—), *cis* cycloadducts (—). Reaction conditions CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA 10 mM, 10 °C, 16 h. Note axis are expanded to 5 mM to observe behaviour more easily.

**Table 2.5:** Rate constants determined by kinetic simulation and fitting of experimental data using SimFit.

	$k_{\text{trans}}$	$k_{\text{cis}}$
Bimolecular Rate Constant / $\times 10^{-4} \text{ M}^{-1} \text{ s}^{-1}$	1.39	0.44
$k_{\text{ternary}} / \times 10^{-4} \text{ s}^{-1}$	15.0	—
$EM_{\text{ternary}} / \text{M}$	10.8	—
$K_{\text{duplex}} / \times 10^8 \text{ M}^{-1}$	6.37	—
Connection $EM / \text{M}$ ( $\Delta G^{\circ} / \text{kJmol}^{-1}$ )	4.29 (1.01)	—

The recognition mediated reaction was then performed by monitoring the reaction of recognition enabled maleimide **72** with *l*-nitrone **83** under the same conditions and deconvoluting the spectra in the same method to create a concentration vs. time profile (Figure 2.27 (ii)). After 16 h, the reaction conversion had increased reaching

27% with *l-trans*-**84** the major product reaching a concentration of 2.2 mM with the production of the diastereoisomer *l-cis*-**84** reaching 0.5 mM. The *trans* : *cis* ratio therefore was 4.4:1 indicating an increase in selectivity for the *l-trans*-**84** in comparison to the recognition disabled system. The results were once again subjected to kinetic simulation and fitting to a self-replicating reaction using SimFit to extract reaction rate information. The rate of reaction within the ternary complex,  $k_{\text{ternary}}$ , was calculated to be  $1.50 \times 10^{-3} \text{ s}^{-1}$  which by comparison to the rate of the control reaction, reveals the effective molarity of the template directed reaction to be 10.8 M.

The template directed nature of the reaction was demonstrated by addition of 10 mol % preformed *l-trans*-**84** to a fresh set of reagents for the reaction. The mixture was then subjected to the same conditions and monitored as before by 470.3 MHz  $^{19}\text{F}$  NMR spectroscopy. From the results, a concentration vs. time profile was constructed (Figure 2.27 (iii)) which showed that the lag period observed in the native reaction is absent and the reaction proceeds at the maximum autocatalytic rate upon initiation. After 16 h the overall conversion had increased to 32% with the production of the major product, *l-trans*-**84** increased to 2.72 mM with the concentration of *l-cis*-**84** reaching 0.47 mM. The *trans* : *cis* ratio therefore had increased from 4.4:1 in the undoped system to 5.8:1 when doped with preformed template. This result highlights that the template directed reaction is selective for the production of *l-trans*-**84** and not involved in any cross-catalytic pathway to enhance *l-cis*-**84**.

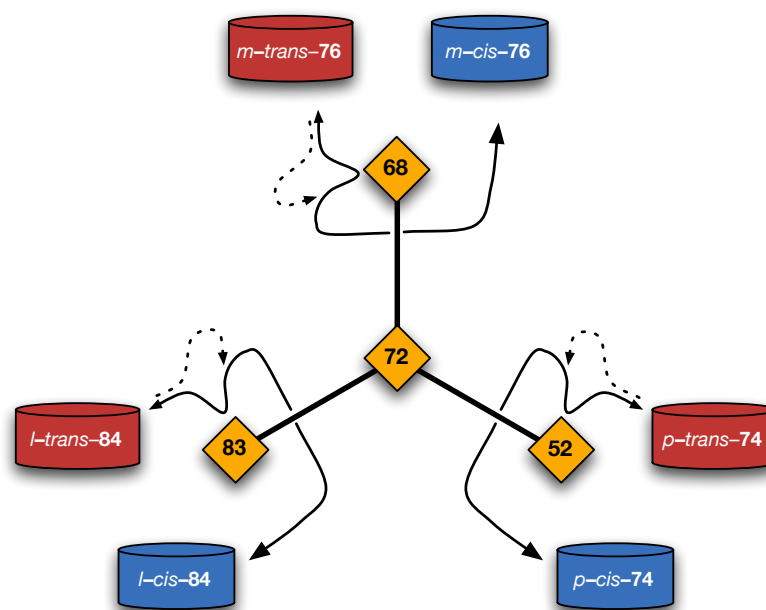
Finally, five equivalents of the competitive inhibitor benzoic acid was added to a fresh set of reagents for the replicating system to demonstrate the importance of recognition mediated events. The reaction was performed under the same conditions and monitored as before by 470.3 MHz  $^{19}\text{F}$  NMR spectroscopy for 16 h. From the results, a concentration vs. time profile (Figure 2.27 (iv)) was once again constructed and in this case, the production of *l-trans*-**84** has been inhibited in comparison to the native system (Figure 2.27 (ii)). This expected decrease in autocatalytic efficiency is once again as a consequence of benzoic acid disrupting the formation of the ternary complexes.

From these results of analysis of the reaction of maleimide **72** with *l*-nitron **83** following the standard protocol of autocatalytic self-replicator characterisation, it is

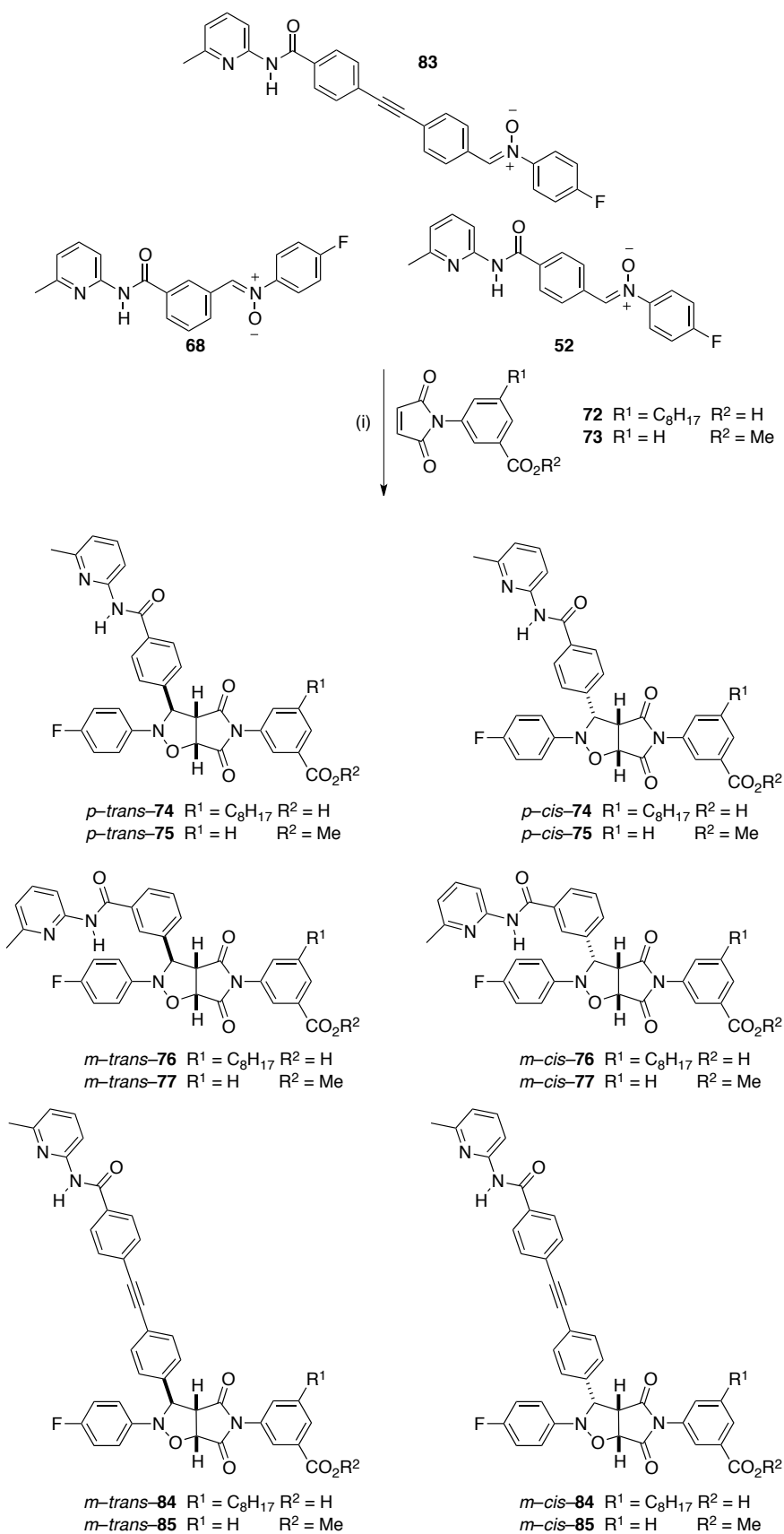
clear that the dominant reaction between *l*-nitroene **83** and maleimide **72** is the autocatalytic formation of *l*-*trans*-**84**.

## 2.14 Expanding the Replicator Network – Three Replicator System

With a third self-replicating reaction using maleimide **72** discovered, we set about to integrate it into a system in which it is combined with the two previously described self-replicating reactions to create a network of three interconnected replicators. The first network we wished to create would be the combination of the three recognition bearing nitroene reagents, **52**, **68** & **83**, with the common maleimide **72** (Scheme 2.20). In a network of three replicators, there are three template directed reactions occurring in the reaction mixture. As such three different templates can be added to the mixture upon initiation and thus there is the possibility that the outcome of the system could be directed in three ways towards the promotion of whichever template is added upon initiation. The system may be visualised as in Figure 2.28.



**Figure 2.28:** Representation of the system in which the common maleimide **72** may react with either nitrones **52**, **68** or **83**. Each reaction irreversibly consumes the common maleimide to produce their two diastereoisomeric products. Three of these product are produced autocatalytically as represented by the second arrow looping back.

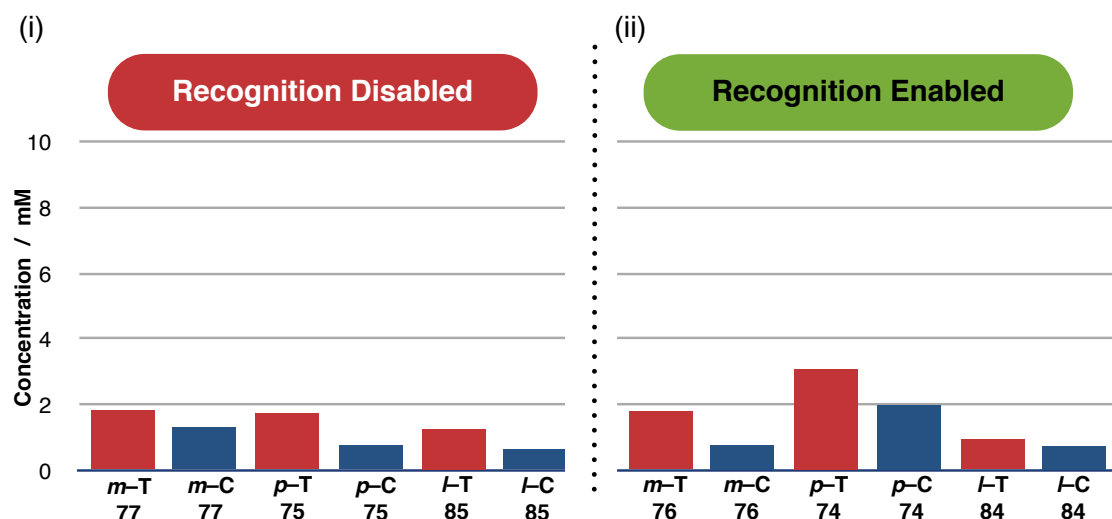


**Scheme 2.20:** Upon mixing the three nitrones with a maleimide, six possible products are formed. Reagents and conditions (i)  $CD_2Cl_2/p$ -TSA (sat),  $10^\circ C$ , 10 mM, 16 h.

Initially, we wished to demonstrate that recognition mediated processes are still responsible for the direction of reactions in the mixture of the three replicators and this was achieved by comparison of a recognition disabled system with a recognition enabled system. In the recognition disabled system of three nitrones with a maleimide, there would be no preference of any reaction other than the other and thus we would expect that the three *trans* cycloadducts would be produced in near equal concentrations with the three *cis* cycloadducts produced at equal concentration to each other but all lower than that of the *trans* products. In the recognition enabled system, we would expect that the three more efficient template directed self-replicating reactions would facilitate the fabrication of themselves above simple bimolecular reactions as seen before in the two replicator system. We would also expect to observe a distribution of products which reflect the different rates of reactions in each of the self-replicating reactions. Finally, we expected that the previously observed [A•B] complex channel product *p-cis-74* will likely be present and efficient as a result of the increased concentration of recognition bearing compounds reducing efficient ternary complex formation.

The recognition disabled system was produced by mixing an equimolar amount of *l*-nitrone **83**, *m*-nitrone **68** and *p*-nitrone **52** with recognition disabled maleimide ester **73** at a concentration of 10 mM in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat) in a manner the same as with the two replicator system in Section 2.8. The mixture was then incubated at 10 °C for 40 h before assaying with 376.5 MHz <sup>19</sup>F NMR spectroscopy and the relative concentration of each product determined by deconvolution of the appropriate resonance signals, see Chapter 7, Section 7.6 for experimental details and example spectrum. The results of the system (Figure 2.29 (i)) revealed that 75% of the maleimide had been consumed with similar concentrations of the three *trans* cycloadducts produced of 1.82 mM, 1.73 mM and 1.26 mM. The three *cis* cycloadducts were also produced in lower concentrations of 1.31 mM, 0.75 mM and 0.63 mM. This result is expected as there would be no selectivity without the presence of recognition mediated template direction of the reaction.





**Figure 2.29:** The concentration of cycloadduct products after 40 h for (i) the recognition disabled system in which the three nitrones are mixed with maleimide ester **73** and (ii) the recognition enabled system in which maleimide acid **72** is mixed with the three nitrones. Abbreviations used *m-T* for *meta-trans*, *m-C* for *meta-cis*, *p-T* for *para-trans*, *p-C* for *para-cis*, *l-T* for *long-trans* and *l-C* for *long-cis*.

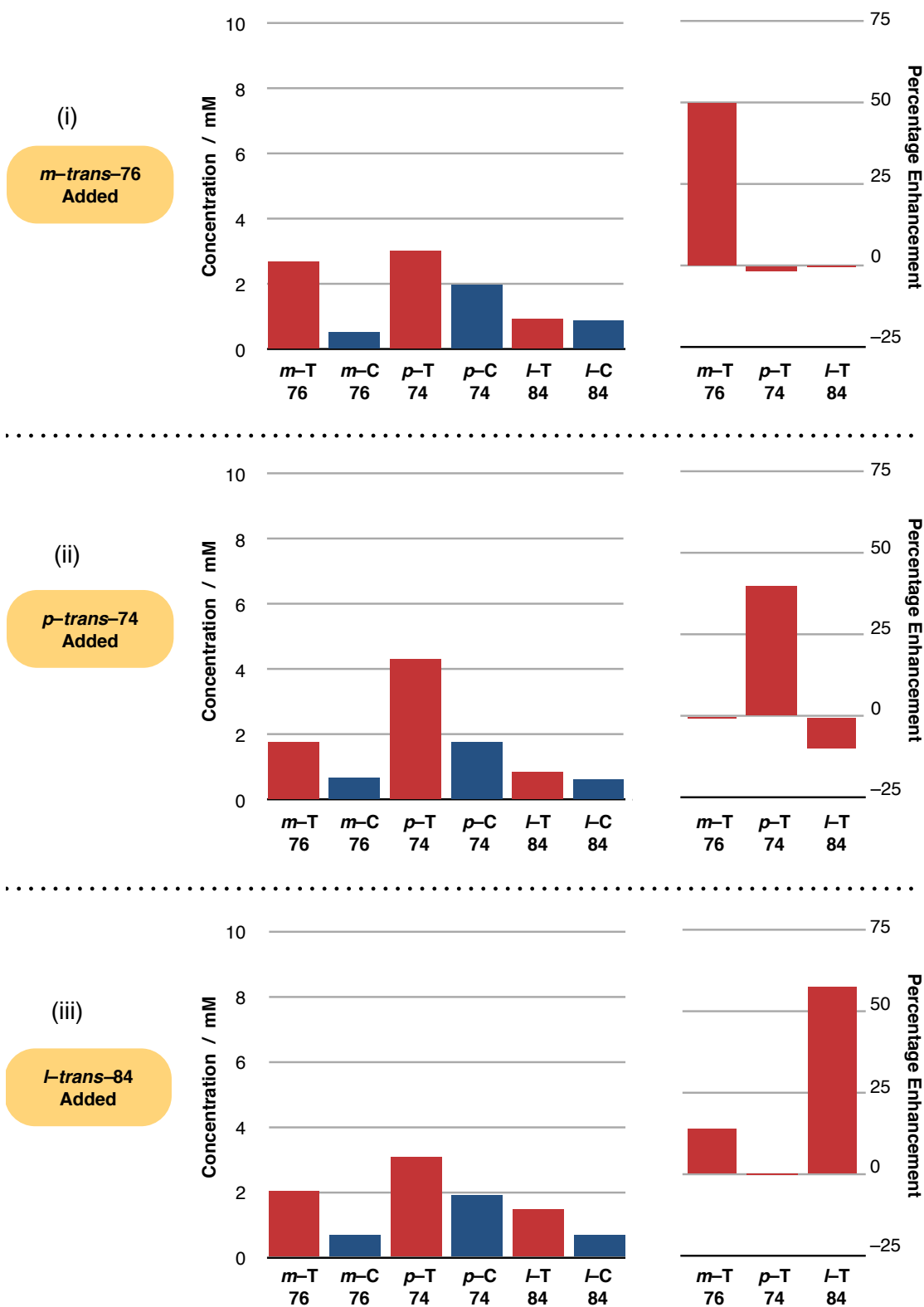
Next, the reaction was repeated under the same conditions, this time using the recognition enabled maleimide **72** to reveal the effects of allowing recognition mediated processes to occur in the system. The results (Figure 2.29 (ii)) revealed that after 40 h of reaction with the recognition mediated maleimide **72**, 92% of the maleimide had been converted into cycloadduct. The major product was *p-trans-74*, the most efficient replicating template, of which 3.08 mM had been produced. The next most abundant product was *p-cis-74*, the product of the [A•B] binary complex channel, which once again is produced in a significantly large quantity of 1.99 mM. The remaining two replicating templates *m-trans-76* and *l-trans-84* are produced in lower concentrations of 1.72 mM and 0.94 mM respectively, both greater than the remaining *cis* products formed through only bimolecular processes. It is interesting to note that the concentration of *m-trans-76* and *l-trans-84* produced in the recognition mediated system is lower than in the control system. This is as a consequence of the more efficient recognition mediated processes of the replicating template *p-trans-74* and [A•B] complex channel consuming the maleimide at an increases rate leaving less to react with *m-nitron* **68** and *l-nitron* **83** therefore producing less *m-trans-76* and *l-trans-84* as a result.

### 2.14.1 Template Direction of Three Replicator System

We now wished to direct the outcome of the system in the same manner as before by adding preformed template for one of the self-replicating reactions in order to direct the system to produce the added template. With three template directed self-replicating reactions present in the system, three possible instructions can be issued.

Firstly we wished to direct the system to produce *m-trans*-**76** which can be achieved by adding preformed *m-trans*-**76** to a mixture of the three nitrones and maleimide **72** upon initiation. The mixture was created in the same manner as before with the two replicator systems. An equimolar mixture of *m*-nitron **68**, *p*-nitron **52**, *l*-nitron **83** and maleimide **72** at 10 mM in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat) along with 20 mol% *m-trans*-**76** was subjected to the same reaction conditions as the uninstructed system (10 °C, 40 h) and then assayed by 376.5 MHz <sup>19</sup>F NMR spectroscopy. The results of the experiment are shown in Figure 2.30 (i) which revealed a different distribution of products in comparison to the uninstructed system. The overall conversion was now quantitative with the major product still *p-trans*-**74** of which 3.03 mM had been produced representing a slight inhibition of 1.6%. The concentration of *m-trans*-**76** produced had now increased to 2.69 mM representing an enhancement of 57% in comparison to the uninstructed system. Clearly the addition of preformed template *m-trans*-**76** has had the effect of increasing the production of itself leading to inhibition in the production of other templates in the system with *l-trans*-**84** also inhibited slightly with 0.94 mM produced (0.3% inhibition).

Next, a fresh mixture of the three nitrones and maleimide was doped with the second of the three templates, *p-trans*-**74**, by adding 20 mol% preformed template using the same experimental procedure as utilised above with the expectation that adding this template will enhance its own production within the system. The mixture of reagents along with the template doping was incubated under the same conditions for the same time before assaying by 376.5 MHz <sup>19</sup>F NMR spectroscopy and calculating the relative concentration of products by deconvolution of the appropriate resonance signals in the spectrum.



**Figure 2.30:** The concentration of cycloadduct products with percentage enhancements relative the the native system (Figure 2.26 (ii)) of each system after 40 h reaction time after doping with (i) 20 mol% *m-trans-76* (ii) 20 mol% *p-trans-74* and (iii) 20 mol% *l-trans-84*. Abbreviations used *m-T* for *meta-trans*, *m-C* for *meta-cis*, *p-T* for *para-trans*, *p-C* for *para-cis*, *l-T* for *long-trans* and *l-C* for *long-cis*.

The results of the instruction (Figure 2.30 (ii)) revealed a different distribution of products once again with respect to the uninstructed system. There was now quantitative conversion of the maleimide to cycloadduct. The major product of the system was once again *p-trans*-**74**, the most efficient replicator, of which 4.33 mM had been produced. In comparison to the uninstructed system, which produced 3.08 mM *p-trans*-**74**, the effect of adding the instruction has been a 40% enhancement in the production of *p-trans*-**74**. The other two replicating templates produced in the system are both inhibited with the concentration of *m-trans*-**76** reaching 1.78 mM (inhibited 0.6% from 1.79 mM in the uninstructed system) and the concentration of *l-trans*-**84** reaching 0.85 mM (inhibited 9.6% from 0.94 mM in the uninstructed system). These results followed our expectation that enhancing the rate of production of the *p-trans*-**74** template will consume the maleimide faster leaving less to react with other nitrones in the system with the overall result an enhancement of *p-trans*-**74** with concurrent inhibition of all other templates in the system.

Finally, a fresh mixture of the three nitrones and maleimide was instructed with the third self-replicating template, *l-trans*-**84**, in the same manner as before. The results (Figure 2.30 (iii)) revealed once again an increase in conversion to quantitative conversion. Comparison of the results to the uninstructed system revealed that the concentration of newly formed *l-trans*-**84** had increased from 0.94 mM to 1.48 mM, a 57% enhancement. Clearly adding preformed *l-trans*-**84** has had the effect of enhancing its own formation within the system. Curiously there was also an enhancement of *m-trans*-**76** of which 2.04 mM had been produced, an increase of 14% from 1.79 mM in the uninstructed system. This result clearly was not expected nor could be explained from physical template effects as doping preformed *l-trans*-**84** into the reaction of *m-nitron*-**68** with maleimide **72** had no effect in the production of *m-trans*-**76**. The result therefore could not be explained by consideration of physical template effects and perhaps its origins lay in experimental error in construction of the three replicator system or arises from the breakdown of material in the system complicating the overall conversion. None the less, the major enhancement of the system was the desired production of *l-trans*-**84** in response to the doping of preformed template.

From these results it is encouraging to observe that the more complicated three replicator system may still be directed to enhance the production of whichever self-

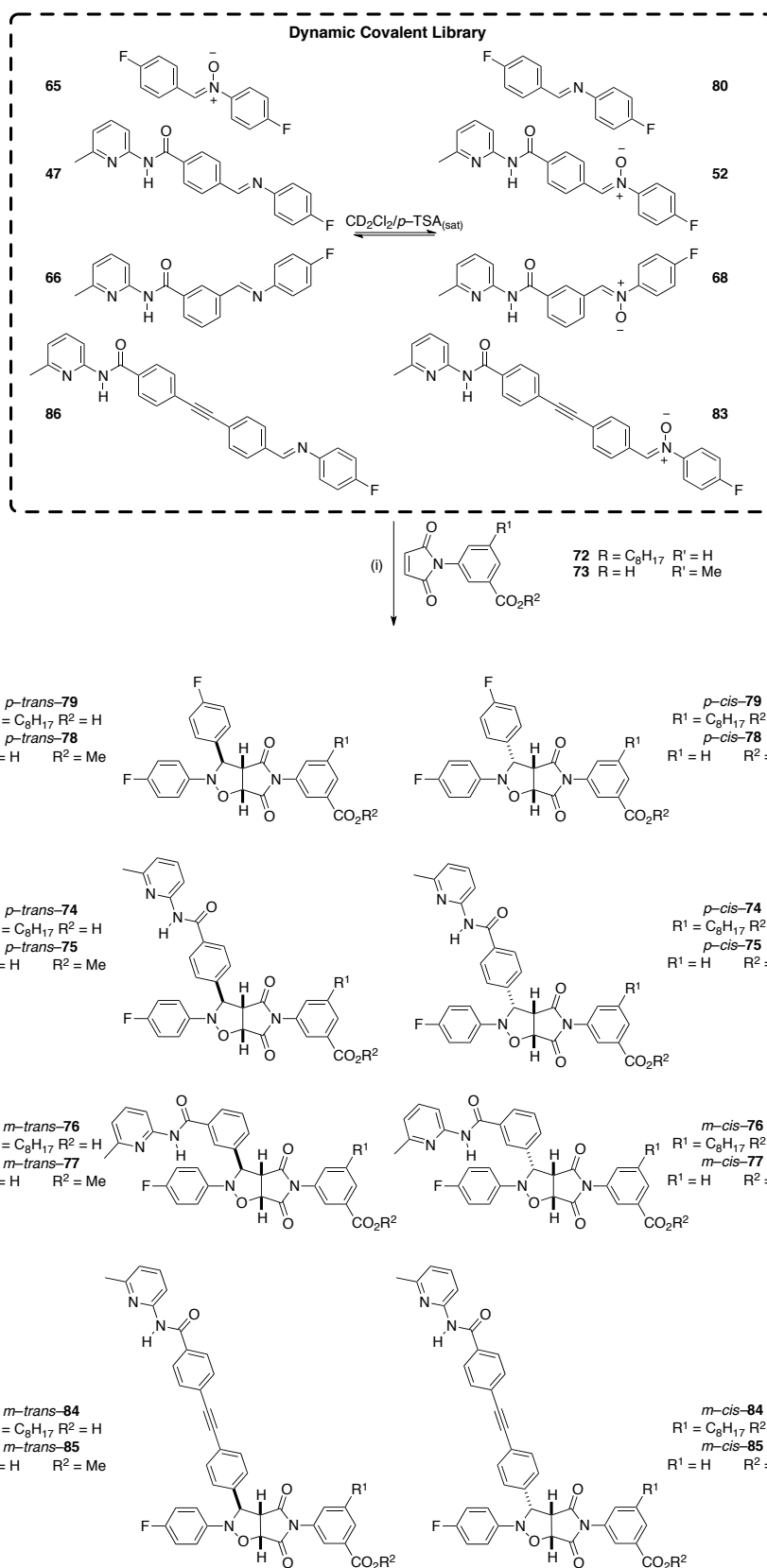
replicating template is desired by doping preformed template as an instruction. The levels of percentage enhancements in the three replicator system are also greater in comparison to the two replicator system, Section 2.8.1. The origins of the greater enhancements lay in a system level effect that in a three replicator system, there is lower production of each of the templates as a result of a greater number of reactions consuming reagent in the system. Upon instruction, the template directed reaction therefore has a relatively greater effect than in a two replicator system.

### 2.15 Coupling Three Replicator System to a Dynamic Pool of Reagents

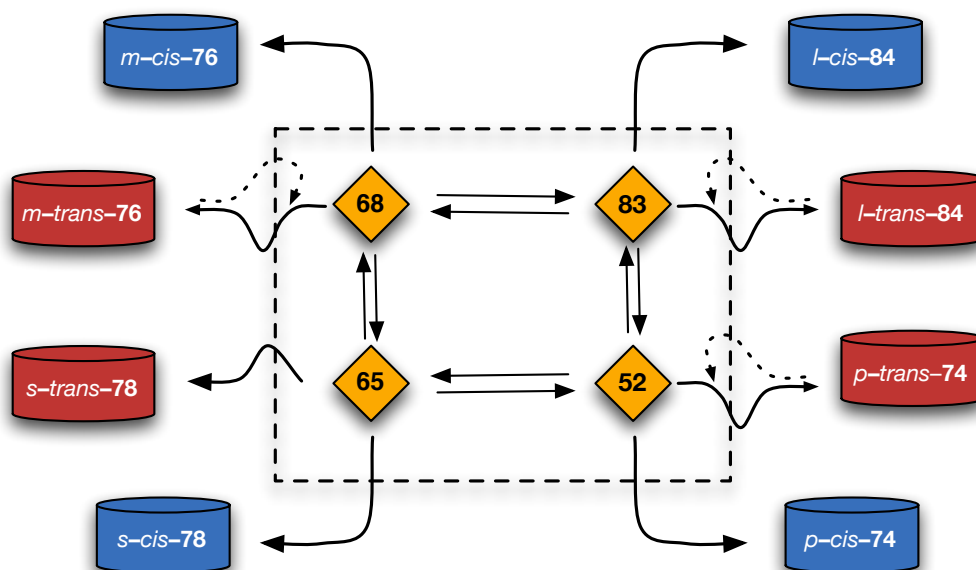
After successfully demonstrating that the product distribution of the three replicator network may successfully be directed to increase the production of whichever template is desired, we now set about to couple the network to a dynamic process in the same manner as previously described in Section 2.5.

In a mixture of *s*-nitronone **65** with three recognition bearing imines, *m*-imine **66**, *p*-imine **47** and *l*-imine **86** in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat) the known dynamic process of the breakdown and recombination of nitrones and imines will occur (Scheme 2.25). From this process, each of the three recognition bearing nitrones will be formed. As there are now in total four nitrones in the dynamic pool, the concentration of each of these nitrones will be lower in the three replicator system in comparison to the two replicator dynamic system in Section 2.9. In the two replicator system, the 10 mM of nitronone added upon initiation would be spread between the two recognition bearing nitrones and the simple nitronone. In the three replicator system, the same 10 mM concentration of nitronone added upon initiation will now be spread over three recognition bearing nitrones and the simple nitronone.

Irreversible cycloaddition reaction of each of these four nitrones will lead to the formation of eight cycloadduct templates. Three of these templates will be produced autocatalytically whilst the remaining five do not participate in template directed synthesis and serve as kinetic traps in the system. The system may be visualised as in Figure 2.31 with the four reagents under dynamic control with irreversible reaction to produce the eight templates.



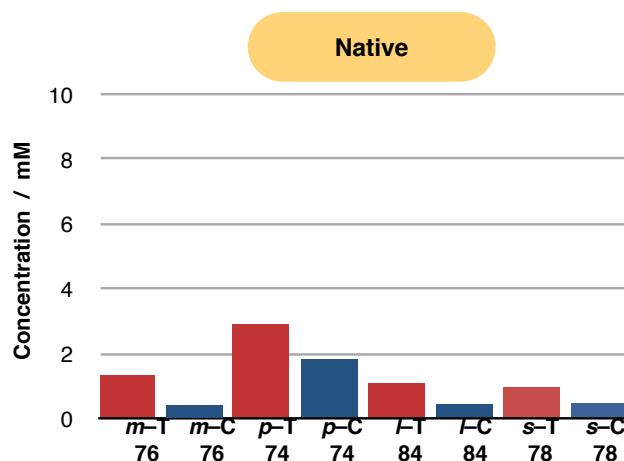
**Scheme 2.25:** Upon mixing the four components, a dynamic library of reagents is produced (enclosed by the black dashed line). Irreversible reaction of either of the nitronium ions with maleimide **72** leads to a product pool of eight unique templates, three of which are produced autocatalytically. Reagents and conditions (i) CD<sub>2</sub>Cl<sub>2</sub>/p-TSA (sat), 10 °C, 120 h.



**Figure 2.31:** Representation of the system in which the four reactive reagents within the black dashed line exchange under thermodynamic control and will each react irreversibly to produce their two diastereoisomeric products. Three of these product are produced autocatalytically as represented by the second arrow looping back into the dynamic box. Whilst equilibrium arrows are shown between nitrones, there is no explicate pathway within the system and all nitrones will be interconverting with each of the others. Not all reagents within the dynamic pool are shown for clarity.

Initially we set about to establish the native distribution of cycloadducts in a system in which the *s*-nitronone **65** is incubated with the three recognition bearing imines, **47**, **66** & **86**, in the presence of maleimide **72**. In order to achieve this, an equimolar solution of *s*-nitronone **65**, *m*-imine **66**, *p*-imine **47**, *l*-imine **86** and maleimide **72** were mixed together at 10 mM in  $\text{CD}_2\text{Cl}_2/p\text{-TSA}$  (sat) and incubated at 10 °C for 120 h. For experimental details, see Chapter 7, Section 7.3. After this time, the mixture was assayed by 376.5 MHz  $^{19}\text{F}$  NMR spectroscopy and the relative concentration of products calculated by deconvolution with the results presented in Figure 2.32. The results of the native experiment revealed that the major product within the mixture was once again *p-trans-74*, the most efficient of the three replicators, of which 2.90 mM had been produced. The next most abundant product was *p-cis-74*, the product of the  $[\mathbf{A}\cdot\mathbf{B}]$  complex channel reaction, of which 1.80 mM had been produced. Once again, the increased complexity of the mixture has reduced the efficiency of ternary complexes in the reaction and allowed the binary complex channel to increase in prevalence. The remain two replicating templates, *m-trans-76* and *l-trans-84*, were the two next abundant products with a concentration of 1.31 mM and 1.06 mM produced respectively. Both of these are still produced above the concentration of the simple bimolecular reaction to form *s-trans-78* which reached 0.94 mM and above

the three *cis* cycloadducts each of which are formed only by the bimolecular reaction at similar concentrations of 0.41 mM, 0.43 mM and 0.46 mM of *m-cis-76*, *l-cis-84* and *s-cis-78* respectively. It is interesting to observe that the hierarchy of replicator efficiency, i.e. *p-trans-74* > *m-trans-76* > *l-trans-84* is faithfully reproduced in the outcome of the three replicator system with the concentration of *p-trans-74* (2.90 mM) > *m-trans-76* (1.31 mM) > *l-trans-84* (1.06 mM).



**Figure 2.32:** The concentration of cycloadduct products after 120 h at 10 °C. Abbreviations used *m-T* for *meta-trans*, *m-C* for *meta-cis*, *p-T* for *para-trans*, *p-C* for *para-cis*, *l-T* for *long-trans*, *l-C* for *long-cis*, *s-T* for *simple-trans* and *s-C* for *simple-cis*.

### 2.15.1 Template Direction of Three Replicator System with Dynamic Reagent Pool

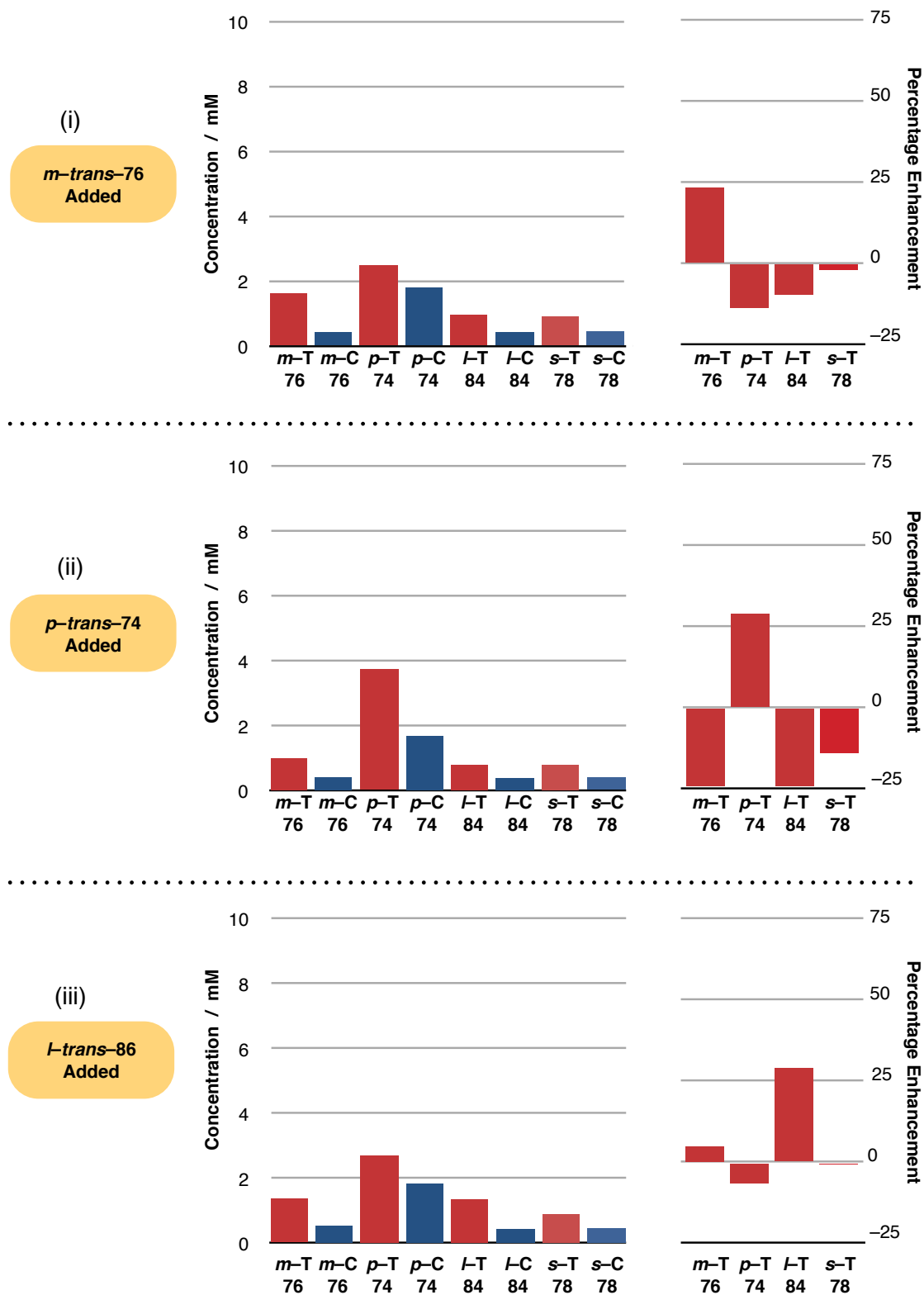
Next, we set about to change the distribution of products in the system by instructing a fresh mixture of reagents with a template. The first template used as an instruction was *m-trans-76* of which a 20 mol% doping was used. This doping was added to a fresh set of the reagents for the system, *s*-nitroene **65**, the three imines **47**, **66** and **86** and maleimide **72** at 10 mM in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat). For experimental details, see Chapter 7, Section 7.3. The mixture was then incubated under the same conditions as the native system (10 °C, 120 h) and then assayed by 376.5 MHz <sup>19</sup>F NMR spectroscopy. The relative concentration of each product was then determined by deconvolution of appropriate resonance signals and the results of the mixture are displayed in Figure 2.33 (i). The distribution of products in the system in this case was altered in comparison to the uninstructed system. Whilst the major product was still *p-trans-74*, the concentration produced was lower now reaching only 2.50 mM, a decrease of 14%. In contrast, the concentration of *m-trans-76* had now increased to



1.62 mM which is an increase of 23%. The concentration of the remaining two *trans* cycloadducts produced has also decreased with 10% less *l-trans-84* and 2% less *s-trans-78* as a consequence of the increased rate of formation of *m-trans-76* consuming the maleimide. The concentration of each of the *cis* cycloadducts produced was similar within experimental error to the native system. Clearly the addition of *m-trans-76* template has selectively directed the formation of *m-trans-76* within the mixture at the expense of the remaining *trans* cycloadducts as the concentration of each has been reduced.

In order to direct the outcome of the system to a different composition, a fresh mixture of the reagents for the system was doped with 20 mol% *p-trans-74* using the same experimental procedure as for the *m-trans-76* instructed system above. The results (Figure 2.33 (ii)) revealed that upon instruction of the mixture with *p-trans-74*, there is a different distribution of products once more. The major product, as expected, is *p-trans-74* of which 3.74 mM had been produced, an enhancement of 29% in comparison to the native system. The concentration of the two replicating templates, *m-trans-76* and *l-trans-84*, have decreased 24% to 1.00 mM and 14% to 0.80 mM respectively. The concentrations of each of the *cis* cycloadducts have also all either decreased or remained at a similar level. In this case, the outcome of the system has been altered by the instruction with *p-trans-74* to produce more of itself and as a result of the increased rate of consumption of maleimide for this process, the remaining reactions in the system have been inhibited resulting in an increase in selectivity for *p-trans-74*.

Finally, a fresh mixture of reagents was doped with *l-trans-84* with the expectation that this template would instruct the system to increase the rate of production of *l-trans-74* at the expense of the other templates in solution. The distribution of products in the system (Figure 2.33 (iii)) revealed that once again there was a different distribution of products caused by the addition of template. In this case, the major product was still *p-trans-74* however the concentration produced had been inhibited by 6% relative to the native system, decreasing to 2.72 mM. The concentration of newly formed *l-trans-84*, the template desired to be enhanced by instruction, had reached 1.36 mM, an increase relative to the native system of 29%.



**Figure 2.33:** The concentration of cycloadduct products after 120 h and percentage enhancement relative to the native product distribution after instruction with (i) 20 mol% *m-trans-76* (ii) 20 mol% *p-trans-74* and (iii) 20 mol% *l-trans-84*. Abbreviations used *m-T* for *meta-trans*, *m-C* for *meta-cis*, *p-T* for *para-trans*, *p-C* for *para-cis*, *l-T* for *long-trans*, *l-C* for *long-cis*, *s-T* for *simple-trans* and *s-C* for *simple-cis*.

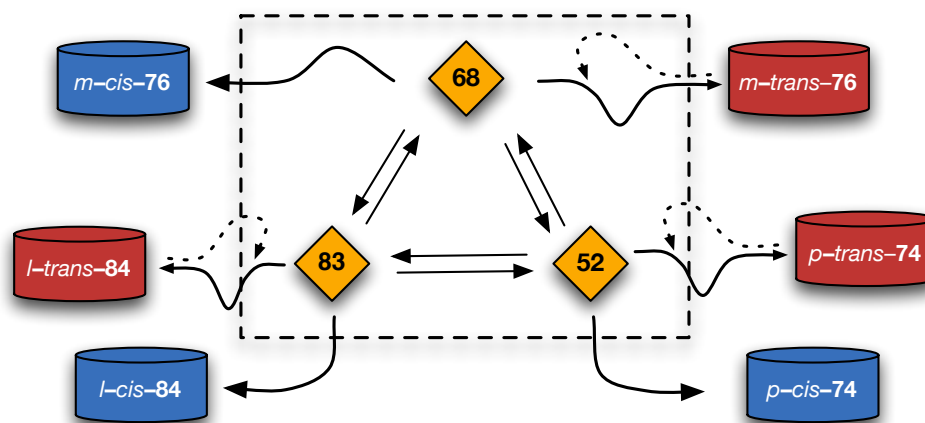
Curiously, the concentration of *m-trans*-**76** produced within the mixture was 1.38 mM, an increase of 5% relative to the native system. As discussed previously, *l-trans*-**84** cannot act as a template to catalyse the cycloaddition reaction of *m*-nitron **68** with maleimide **72** to form *m-trans*-**76**. An explanation could be that in the native reaction, only 92% overall conversion had been reached whereas in the template doped reactions, 100% conversion was observed thus obscuring the ability to make a more accurate comparison. Alternatively degradation of the reagents over the reaction time course, more specifically of the hydroxylamine component of the nitrones, may have in this case complicated analysis. Otherwise there is a more complicated interplay of dynamic equilibria vs. kinetic reactions which would warrant further investigation in the future.

These results demonstrate the successful creation of a three replicator system in which one of the reagents for the reaction is provided by a dynamic pool of reagents. The outcome of the system, that is the distribution of products, can be manipulated by adding one of the preformed templates to amplify the desired product. Whilst the magnitude of enhancements offered by the instruction are not as great as observed previously in the two replicator system, the three replicator system here is still able to respond to a given instruction in a manner that would be expected. One may assume that with the increased number of recognition processes occurring as a result of the addition of the third replicating reaction, the overall efficiency of each individual process is lowered as a result which could explain the lower magnitude of enhancements.

## 2.16 Three Replicator System From Simple Components

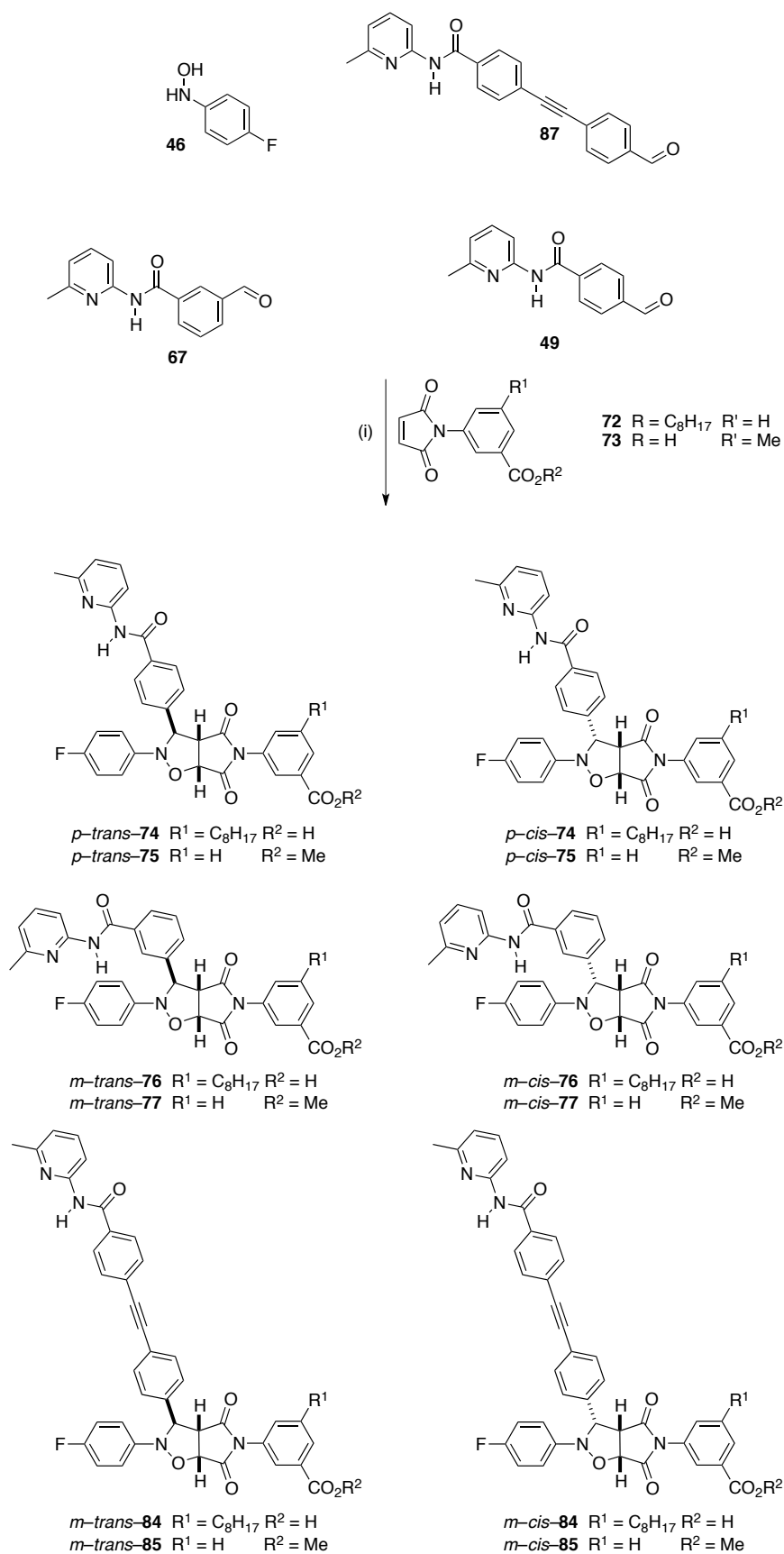
With the successful design of a three replicator system which responds to instruction in a simple kinetic system, as in Section 2.14, and a dynamic reagent pool, Section 2.15, we finally looked to establish whether a smaller dynamic pool of subcomponents could be employed to control the production of the nitrones as described previously for the two replicator system in Section 2.10 and if such a system would produce greater enhancements as observed before. Such a system may also be visualised as in Figure 2.34 which displays that there is a smaller reagent pool in comparison to the earlier Section 2.15. There is also a greater proportion of templates produced by the system which are done so autocatalytically,

three of the six. In such a system there are fewer kinetic traps and as a result we would expect to see a greater degree of control of the system .

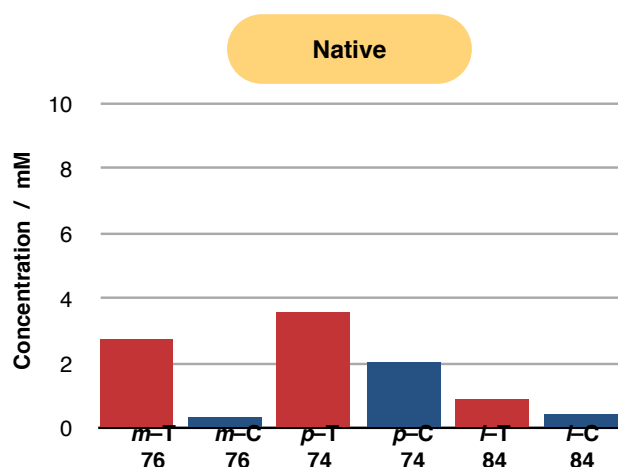


**Figure 2.34:** Representation of the system in which the three reactive reagents within the black dashed line exchange under thermodynamic control and will each react irreversibly to produce their two diastereoisomeric products. Three of these product are produced autocatalytically as represented by the second arrow looping back into the dynamic box.

To produce such a system, three recognition bearing aldehydes **49**, **67** and **87** were mixed with hydroxylamine **46** in equimolar amounts at 10 mM in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat), as displayed in Scheme 2.27. These aldehydes are the precursors for the three recognition bearing nitrones, *m*-nitrone **68**, *p*-nitrone **52** and *l*-nitrone **83**, which are formed upon reaction of the appropriate aldehyde with hydroxylamine **46**. Also present in the system is maleimide **72** which can undergo cycloaddition with the nitrones once formed. Initially to assess the distribution of products without any instruction by preformed template, an equimolar mixture of aldehydes **49**, **67** and **87**, hydroxylamine **46** and maleimide **72** were dissolved together at 10 mM in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat) and incubated at 10 °C for 120 h. For experimental details see Chapter 7, Section 7.4. After this time, the mixture was assayed by 376.5 MHz <sup>19</sup>F NMR spectroscopy and the relative concentration of products determined by deconvolution of the appropriate product resonance signals.



**Scheme 2.27:** Reagents and conditions (i) CD<sub>2</sub>Cl<sub>2</sub>/p-TSA (sat), 10 mM, 10 °C, 120 h.



**Figure 2.35:** The concentration of cycloadduct products after 120 h at 10 °C. Abbreviations used *m*-T for *meta*-*trans*, *m*-C for *meta*-*cis*, *p*-T for *para*-*trans*, *p*-C for *para*-*cis*, *l*-T for *long*-*trans*, *l*-C for *long*-*cis*, *s*-T for *simple*-*trans* and *s*-C for *simple*-*cis*.

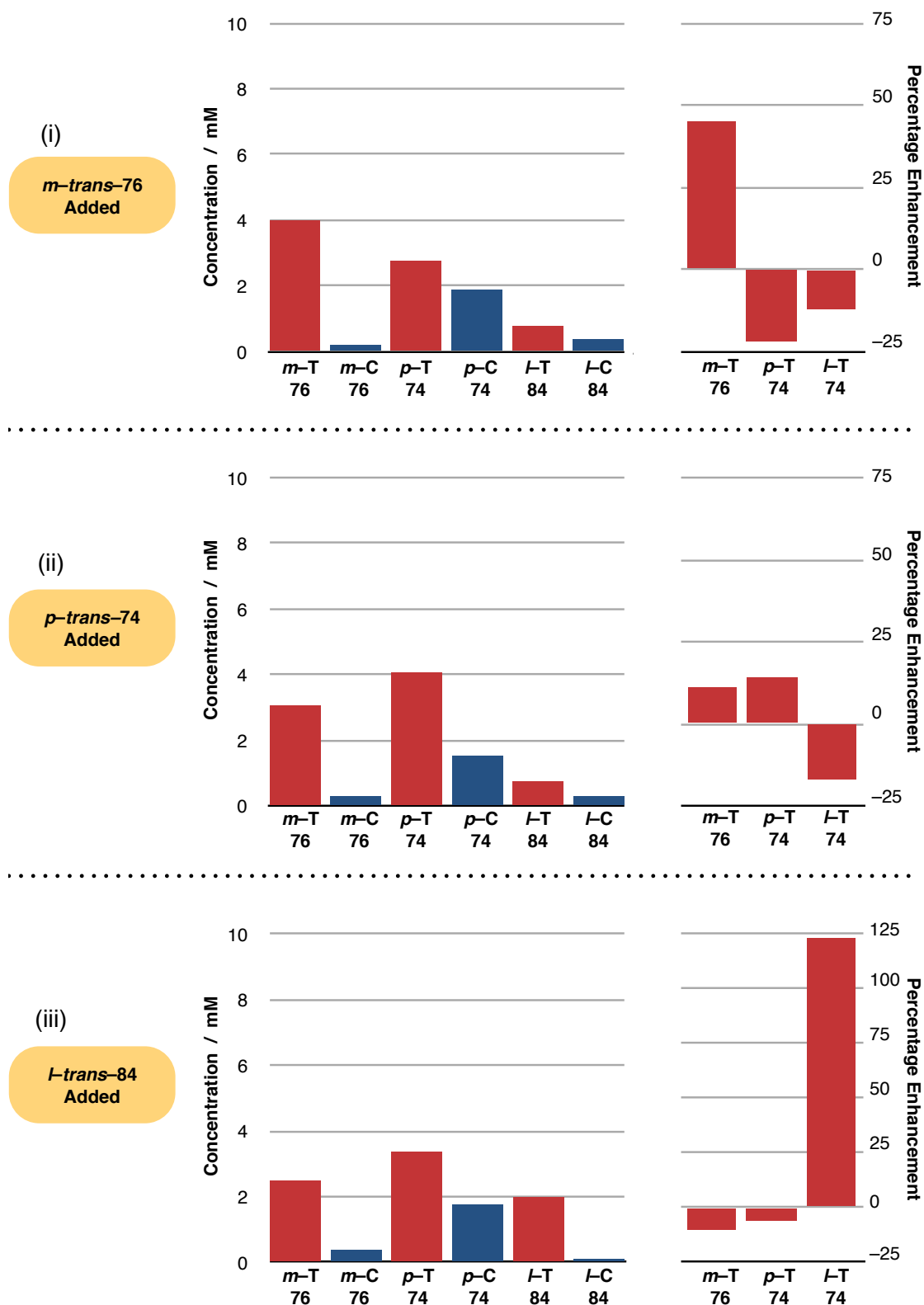
The product distribution (Figure 2.35) revealed that the major product of the mixture was once again the more efficient replicating template *p*-*trans*-**74** of which 3.57 mM had been produced. The second most abundant product in this case was now the second most efficient template *m*-*trans*-**76** of which 2.75 mM had been produced. This result is similar to the two replicator system described earlier in which the concentration of the *m*-*trans*-**76** was closer to the *p*-*trans*-**74** as a result of the preference of the formation of *m*-nitron **68** over *p*-nitron **52** in the dynamic reagent pool. We may expect that the same process is active in this system and is responsible for the result. The product of the [A•B] complex channel reaction, *p*-*cis*-**74** is the third most abundant product in the mixture with 2.04 mM produced. 0.89 mM of the final replicating template *l*-*trans*-**84** was produced along with 0.41 mM of its diastereoisomeric partner *l*-*cis*-**84** and finally 0.34 mM of *m*-*cis*-**76** formed also through the bimolecular reaction only.

### 2.16.1 Template Direction of Simple Components

To a fresh set of reagents, a doping of preformed *m*-*trans*-**76** was issued to direct the system to alter the outcome of the distribution of products with the desire to produce more *m*-*trans*-**76**. The doping was achieved by adding 20 mol% *m*-*trans*-**76** to a fresh mixture of aldehydes **49**, **67** and **87**, hydroxylamine **46** and maleimide **46** and incubated under the same conditions for 120 h. See Chapter 7, Section 7.4 for experimental detail. The result of the reaction (Figure 2.36 (i)) revealed a different distribution of products within the mixture, the concentration of newly formed *m*-

*trans*-**76** had increased to 4.00 mM, an enhancement of 45%, which now meant it was the most abundant product in the mixture. The concentration of the other two replicating templates produced in the system, *p-trans*-**74** and *l-trans*-**84** had decreased 23% to 2.76 mM and 12% to 0.38 mM respectively indicating that the upregulation of the reaction to form *m-trans*-**76** was responsible for the swift consumption of reagent and leaving less to react to form other templates. The production of the three *cis* cycloadducts had also decreased in all cases.

Next, the system was instructed with a different instruction by doping a fresh set of reagents with *p-trans*-**74** in the same manner as above. The expectation was that the outcome of the system would now change and enhance the production of *p-trans*-**74**. The distribution of products in this case (Figure 2.36 (ii)) revealed that the addition of template had caused the mixture to produce a different distribution of products relative to the native reaction. The major product was *p-trans*-**74** of which 4.07 mM had been produced, an enhancement of 14% relative to the native system. The next most abundant product was then *m-trans*-**76** of which 3.06 mM had been produced which represents an enhancement of 11%. This result was unexpected as although *p-trans*-**74** had been demonstrated to template the formation of *m-trans*-**76** from *m-nitrone* **68** and maleimide **72**, an enhancement in the formation of *m-trans*-**76** in a system which had been instructed by *p-trans*-**74** had not been observed in any previous example. In this case, the interplay of dynamic equilibria and kinetic rates may be producing an unexpected composition of reactants in the mixture. We already know that *m-nitrone*-**68** is produced more favourably over *p-nitrone*-**52** from a smaller reagent pool, Section 2.10. Now that a further aldehyde is present in the mixture, the hydroxylamine component will be consumed to produce a further nitron, which if is thermodynamically more favoured than *p-nitrone*-**52** will lead to a higher dilution of the nitron and thus make its own recognition mediated events less efficient at the expense of others, namely *m-nitrone*-**68**, in solution. Such behaviour would require further investigation in the future. The concentration of *l-trans*-**84** produced had been inhibited as expected by 17% as only 0.74 mM had been produced. The concentration of the remaining *cis* cycloadducts had all been inhibited as expected from increased consumption of reagent to form *p-trans*-**74**.



**Figure 2.36:** Concentration of each of the cycloadduct products after 120 h reaction time and percentage enhancement relative to the native product distribution after instruction of a fresh set of reagents with (i) 20 mol% *m-trans-76* (ii) 20 mol% *p-trans-74* and (iii) 20 mol% *l-trans-84*. Abbreviations used *m-T* for *meta-trans*, *m-C* for *meta-cis*, *p-T* for *para-trans*, *p-C* for *para-cis*, *l-T* for *long-trans*, *l-C* for *long-cis*, *s-T* for *simple-trans* and *s-C* for *simple-cis*.



The final objective was to instruct the system with *l-trans*-**84**, the third replicating template, in the same manner as above with the expectation that the outcome of the system would now reflect the instruction and enhance the production of *l-trans*-**84**. The distribution of products in the mixture (Figure 2.36 (iii)) reveal that the addition of template had once again altered the distribution of products relative to the native reaction. The major product once again was still *p-trans*-**74** of which 3.35 mM had been produced, an inhibition of 6% relative to the native system. The next most abundant product was then *m-trans*-**76** of which 2.49 mM had been produced, once again inhibited relative to the native reaction this time by 9%. The concentration of newly formed *l-trans*-**84** had been affected significantly by the doping of template with 1.98 mM now produced which, relative to the native reaction, is an enhancement of 123%, the greatest value of enhancement seen thus far. The concentration of the three remaining *cis* cycloadducts had all been inhibited as expected from increased consumption of reagent to form *l-trans*-**84**. This result was interesting as the weakest replicating template in isolation had the greatest template effect when instructed in the system.

The results demonstrated that a network of three interconnected replicators was successfully created from a mixture of their simple aldehyde, hydroxylamine and maleimide reagents and allowing the dynamic formation of the nitron to occur *in situ*. The relative distribution of products of the three replicator system faithfully reproduced the relative reactivities of the isolated systems, that is the most efficient replicator in isolation, *p-trans*-**74**, was the most abundant replicator produced in the mixture followed by the next most efficient, *m-trans*-**76**, and finally *l-trans*-**84**. The distribution of products in the system was once again able to be tailored by instruction using preformed template to direct the formation of one particular reaction, enhancing the production of that template relative to the native system whilst also inhibiting, in all but one case, the production of the remaining templates within the mixture.

## 2.17 Overall Conclusions

The work presented in this Chapter was designed to investigate the relationship between two or more self-replicating templates when they are formed concurrently in

the same solution. The survival of the fittest theory might suggest that the more efficient replicator would dominate the product pool however the results established that the network of two replicators was able to co-exist. The concentration of each of the two replicators produced was not equal but weighted towards the replicating template which was found to be more efficient in isolation unless the reagents were thermodynamically controlled in a dynamic reagent pool. The distribution of products in the system was then able to be directed towards the formation of whichever template was desired by issuing an instruction to the mixture upon initiation in the form of preformed template. Using this method, the rate of the self-replicating reaction provided with preformed template was accelerated upon initiation consuming the finite concentration of maleimide more rapidly and leaving less to react to form other products whose rate of formation was not accelerated. The overall effect being the enhancement of the concentration of instructed product and inhibition of non instructed product in the system, in short, the outcome of the system could be controlled.

The complexity of the system was then increased by introducing dynamic processes into the system, firstly in the form of a dynamic nitron / imine exchange which has the effect of splitting the concentration of nitron reactive site across three or four reagents and then through dynamic exchange replenishing the concentration of nitron which is being removed most rapidly by reaction with the maleimide. In each case, the levels of enhancements were greater in systems coupled to a dynamic process than the simple kinetic race.

A redesign of the maleimide reagent from a glycine derived maleimide to a benzoic acid derived maleimide **72** led to the formation of three new self-replicating templates whose behaviour was characterised using the established protocol. The networks created using the redesigned maleimide suffered from the presence of an [A•B] complex channel which reduced the conversion of material to form replicating templates however the outcome of these systems were still able to be controlled even when the dynamic pool of reagents was extended to include three further nitrones which could only react in the bimolecular reaction.

Increasing the number of self-replicating reactions present within the system led to the creation of a three replicator network which was still controllable even though

there existed a hierarchy of different replicator efficiencies. Finally the reagent pool of the three replicator system was coupled to a dynamic process led to the creation networks whose outcomes were still able to be directed by doping with preformed template once again.

These results demonstrate that networks of multiple small molecule self-replicators can be created and their native outcome can be altered as desired by instruction. When coupling the reagents to dynamic processes, the template directed reactions are responsible for disturbing the equilibrium of the dynamic pool which in turn causes a redistribution of the reagent pool replenishing reagents which are consumed at an increased rate providing more for reaction and thus demonstrating the operation of a feedback loop.

In order to obtain a high degree of selectivity from a small system of chemical reagents we have investigated the properties of a fully kinetic system, and systems which are under varying degrees of dynamic control. Systems fully under thermodynamic control, such as the examples of Guiseppone and Philp (Section 1.6.2), will always reach an equilibrium point and whilst the addition of instructions to these systems allow them to reach this point quicker, the point will always be the same. In contrast systems under kinetic control were able to respond to instruction which lead to alternative outputs depending on the nature of the input. It is therefore essential to include an irreversible process into the design of controllable systems although it would be favourable to subject the reagent pool to a dynamic process in order to create feedback loops leading to greater selectivity.

# 3

## An Orthogonal Reciprocal Replicating System

### 3.1 Replication and the Origin of Life

The exact mechanism from which life evolved on earth has long fascinated scientists and to this day continues to elude us. Modern scientific theories on the origin of life predominantly fall into two rival categories; replicator–first or metabolism–first.

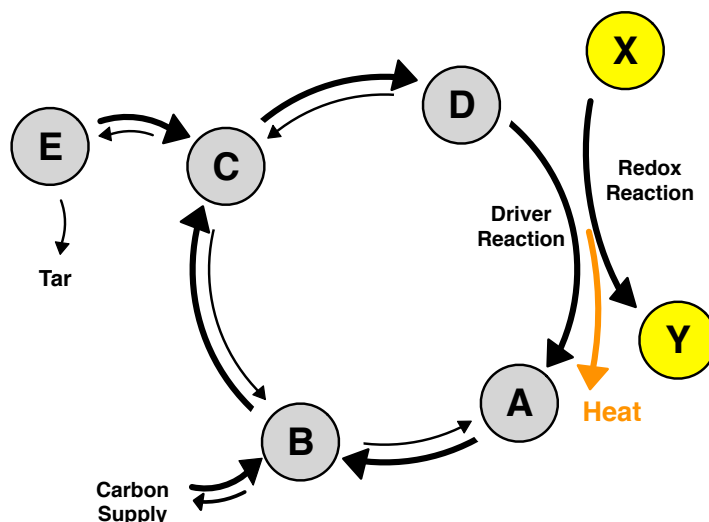
Both models utilise a pool of molecules formed on early earth by non–biological processes<sup>†</sup> or provided by extraterrestrial sources however it is what happens next to these molecules in which the theories differ.

The metabolism–first theory requires the spontaneous formation of compartments which contain mixtures of the molecules.<sup>190,191</sup> These molecules may then undergo cycles of reactions linking common components, for example the **ABCD** cycle in Figure 3.1, and thus forming a chemical network. In order to overcome the negative change in entropy that compartmentalisation brings about, the network must have a driver reaction, one which releases heat into the outside environment thus increasing the entropy of the outside environment compensating for the decrease in the chemical network. The irreversible driver reaction consumes the chemical energy and serves as the engine driving the organisation of the network. Branch reactions will occur, for example moving material between **D** and **E** outside of the cycle, however the driving force of the **ABCD** cycle will favour material moving back into the network thus maximising the release of energy through the cycle. For the network to grow, material must be gained faster through a carbon supply than it is consumed by precipitating or ‘taring’. For the system to survive and grow, a method of reproduction

---

<sup>†</sup> The methods of ‘non–biological processes’ and molecules which are produced by these methods is widely known as pre–biotic synthesis, an area of equal merit in the origin of life studies as the identity of available prebiotic material may determine the appropriate pathway for life to form.

must develop. For example if the network is trapped in the pores of a rock, when the network has grown enough, the material may spill over into an adjacent pore. The newly filled pore may then continue driving the same chemical network as the parent pore as all the material required has spilled into the new pore. The term ‘compensational genome’ is used to describe the heredity information stored by small molecules in a network rather than in a polymeric list in DNA or RNA.



**Figure 3.1:** Each grey node is a small organic molecule which may react between two forms as indicated by reaction arrows. Coupling the driver reaction  $D \rightarrow A$  to an energy source such as a redox reaction on a mineral surface (yellow) provides energy to drive the system whilst releasing heat to the surrounding environment to compensate for negative entropy. An adjacent process will provide material to the system as a carbon supply whilst branching reactions will be present which removes material from the system eventually leading to loss of material by precipitation as tar. Figure adapted from reference 190.

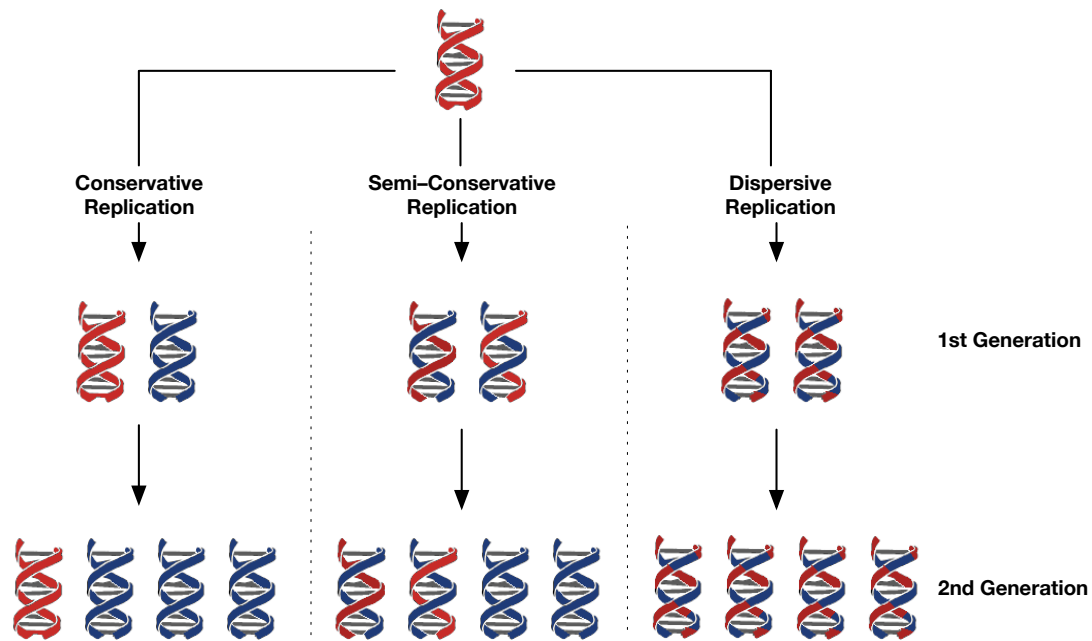
The replication–first hypothesis on the other hand assumes a different fate for the small prebiotic molecules. In this hypothesis, the small molecules may react together, joining to form a chain capable of copying itself. This chain may then reproduce through replication, consuming more of the required small molecules. This progenitor RNA may then undergo mutation and evolution to select replicators which are more efficient at reproducing and storing information than its original chain and these new chains will be selected and survive.

Life today however, is not composed of single selfish molecules replicating only themselves but an interplay of replication and metabolism systems in cells. The blueprints for creating all of the machinery required to produce and maintain these systems is stored in our genes, long chains of DNA which code for a particular

protein *via* RNA chains. As cells divide allowing organisms to grow or replace dead cells, the DNA in its genome must be replicated so that both daughter cells have the same genetic information as their parent.

### 3.1.1 Reciprocal Replication

DNA comprises of two polymeric chains associated together by hydrogen bonding between nucleotide bases in a process called base pairing allowing each chain to wrap around each other to form the iconic double helix structure.<sup>66</sup> Upon the discovery of the double helix structure, Watson and Crick commented that each single strand would serve as a template for the synthesis of its complementary partner, a process known as semi-conservative replication.<sup>67</sup> To confirm this hypothesis, Meselson and Stahl<sup>192</sup> designed a reaction to monitor the replication process by isotope labelling (Figure 3.2).  $^{14}\text{N}$  is by far the most abundant isotope of nitrogen in nature however  $^{15}\text{N}$  may be easily obtained. The authors grew *E. coli* cultures in a medium enriched with  $^{15}\text{N}$  for several generations and investigated their DNA by centrifugation which found that the DNA of cells grown in  $^{15}\text{N}$  medium had a higher density than that of DNA grown in a  $^{14}\text{N}$  medium. Next, *E. coli* cells with  $^{15}\text{N}$  enriched DNA were incubated in a  $^{14}\text{N}$  medium and allowed to divide. After a single round of replication, DNA was extracted from the *E. coli* to be investigated by centrifugation and was compared to samples of pure  $^{14}\text{N}$  DNA and  $^{15}\text{N}$  DNA. The first generation was found to have an intermediate density between pure  $^{14}\text{N}$  DNA and  $^{15}\text{N}$  DNA. This result dismissed the conservative replication mechanism (Figure 3.2 left) which should result in equal amounts of both  $^{14}\text{N}$  DNA and  $^{15}\text{N}$  DNA however the result was still consistent with expected results for a dispersive replication mechanism (Figure 3.2 right). The *E. coli* cultures were allowed to divide once more in  $^{14}\text{N}$  medium and the DNA extracted from the second generation and analysed. The second generation DNA was found to be composed of two equal amounts of DNA with different densities, one corresponding to pure  $^{14}\text{N}$  DNA and the other corresponding to the intermediate density observed after a single division. These results were inconsistent with the dispersive mechanism of replication which would require a single density of DNA, whilst satisfying the expected results of the semi-conservative mechanism.



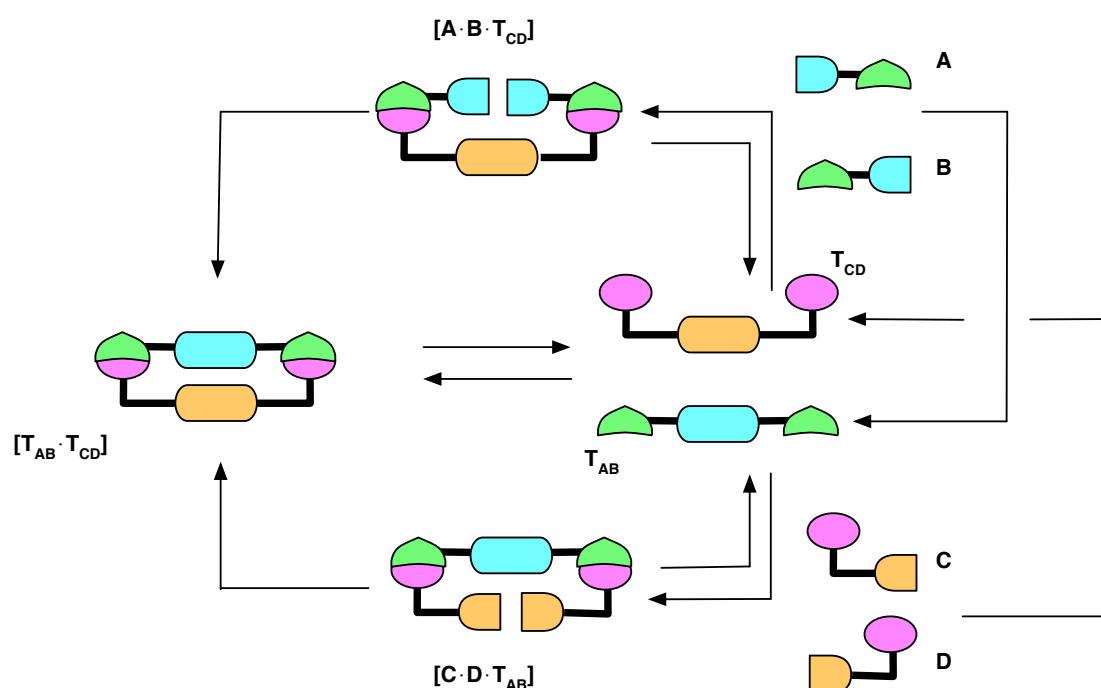
**Figure 3.2:** The three proposed mechanisms of DNA replication, The conservative, semi-conservative and dispersive mechanism. Starting from  $^{15}\text{N}$  enriched *E. coli* DNA (red), culturing in  $^{14}\text{N}$  medium and allowing replication, the conservative replication mechanism will produce an equal amount of pure  $^{14}\text{N}$  DNA and  $^{15}\text{N}$  DNA, the semiconservative mechanism will produce mixed DNA of one  $^{14}\text{N}$  strand and one  $^{15}\text{N}$  strand and the dispersive mechanism will produce two mixed strands of DNA. After the second generation of replication, the conservative replication mechanism will produce 75% of pure  $^{14}\text{N}$  DNA and 25% of pure  $^{15}\text{N}$  DNA, the semiconservative mechanism will consist of 50% of the mixed DNA with one  $^{14}\text{N}$  strand and one  $^{15}\text{N}$  strand and 50% of pure  $^{14}\text{N}$  DNA and the dispersive mechanism will produce equal amounts of mixed DNA.

This semiconservative replication mechanism by which new DNA is produced is a form of template directed synthesis, albeit aided by enzymes. Firstly, the double helix is unwound to expose the two strands. Each strand is now able to act as a template through base pairing to arrange complementary bases on to its surface allowing the enzyme-catalysed synthesis of a new complementary strand. The overall result is that two new double-stranded DNA chains are produced and can be passed on to the daughter cells upon division.

DNA specifically may be thought of as being produced by a reciprocal template mechanism in which, very simply, three primary reagents are present. One template and two complementary building blocks. The template pre-organises the building blocks on its surface and promotes their reaction leading to an increase in rate of formation and selectivity of the complementary partner. The presence of the template instruction at the start of the replication process is crucial to the success of DNA replication. However, when life came to evolve there would have been no instruction to follow. As mentioned previously, the replication-first theory attributed the origins of

life to a single replicating molecule. Another possibility is that instead of a single molecule replicating from a pool, two complementary molecules whose synthesis proceeds in a reciprocal replicating fashion, similar to DNA synthesis, may have been the initial spark of life. Certainly today, reciprocal replication is crucial to life.

In a reciprocal replicating system, both of the complementary templates are produced *in situ* from four building blocks (Figure 3.3).



**Figure 3.3:** Four component reciprocal replicating system. Complementary recognition sites are coloured purple and green. Reactive sites are sites are orange for T<sub>CD</sub> and blue for T<sub>AB</sub>.

At the start of a reciprocal replication cycle, only the four reagents **A**, **B**, **C** and **D** are present, there are no templates available to catalyse any reaction. Therefore, the only way to initiate the cycle is through a bimolecular reaction. For example the reaction of **A** and **B** through a bimolecular mechanism will furnish T<sub>AB</sub>. The recognition sites on T<sub>AB</sub> are able to associate their complementary reagents, in this case **C** and **D**, to arrange them on to the template forming the ternary [C·D·T<sub>AB</sub>] complex. This complex will promote the reaction between **C** and **D** to form the duplex [T<sub>AB</sub>·T<sub>CD</sub>]. Dissociation of the duplex now releases two templates, T<sub>AB</sub> and T<sub>CD</sub>, into solution. T<sub>AB</sub> can continue to catalyse the reaction of **C** and **D** as described, however, T<sub>CD</sub> will be able to associate with its complementary reagents **A** and **B** to form the ternary [A·B·T<sub>CD</sub>] complex which will promote the reaction between **A** and **B** in a



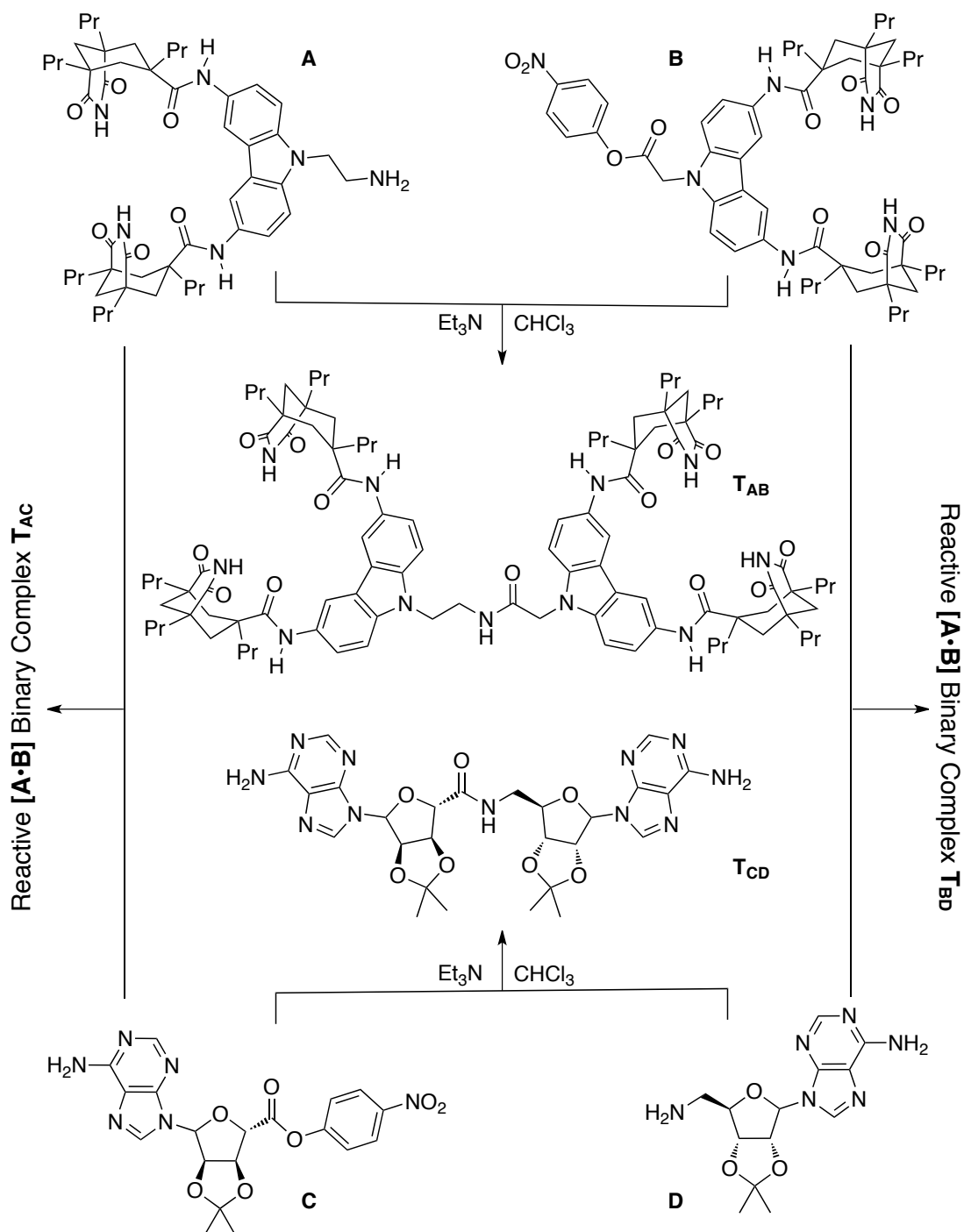
specific manner to produce the  $[T_{AB} \cdot T_{CD}]$  duplex which can once again dissociate and begin another cycle. The formation of each template is mutually beneficial as  $T_{AB}$  will produce  $T_{CD}$  which will in turn produce  $T_{AB}$ , a form of positive feedback. Without the molecular recognition facilitating the specific arrangement of reagents on to the template, only slow bimolecular reactions subject to errors will be present. The reciprocal replication cycle is able to subdue these errors by increasing the rates of the production of the two complementary templates. In a reciprocally replicating system, two different templates complementary to each other are formed. Reciprocal replication, the mechanism by which DNA stores and transmits genetic information, may be seen as a more efficient method of storing information. The reciprocal relationship between the two templates means that each template has the information required to produce the other which in turn may reproduce the other. Thus the information stored in one template is sufficient to reproduce both.

### 3.1.2 Synthetic Reciprocal Template Effects

Previous attempts to produce a synthetic reciprocal replicating system have emerged from experience in minimal self-replication though have not been able to successfully produce an efficient reciprocal replicating system. Rebek *et al.* were able to show<sup>193,194</sup> reciprocal template effects in a synthetic system (Scheme 3.1).

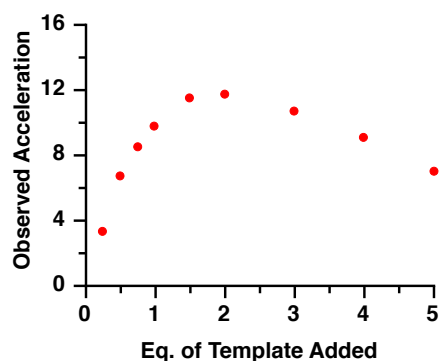
The amide bond formation between the adenosine decorated molecules **A** and **B** to give template  $T_{AB}$  was found to be accelerated twelve-fold, relative to the bimolecular rate, in the presence of two equivalents of  $T_{CD}$ . The *di*-imide cleft on **A** and **B** strongly binds the adenosine groups on  $T_{CD}$  ( $K_a = 10^4$ – $10^5$  M<sup>-1</sup>) which leads to the formation of the reactive ternary  $[A \cdot B \cdot T_{CD}]$  complex (as well as unreactive  $[A \cdot A \cdot T_{CD}]$  and  $[B \cdot B \cdot T_{CD}]$ ). The reactive ternary complex will orientate the amine in **A** into close proximity to the activated *p*-nitrophenyl ester in **B** and render the reaction *pseudo* intramolecular thus accelerating the rate.

Increasing the concentration of added template was seen to increase the initial rate of reaction up to the addition of two equivalents (Figure 3.4). Above this concentration, adding more template was seen to facilitate less of a rate acceleration as a result of the two reactive compounds being increasingly separated as complexes on different templates.



**Scheme 3.1:** Reciprocal template system proposed by Rebek *et al.*

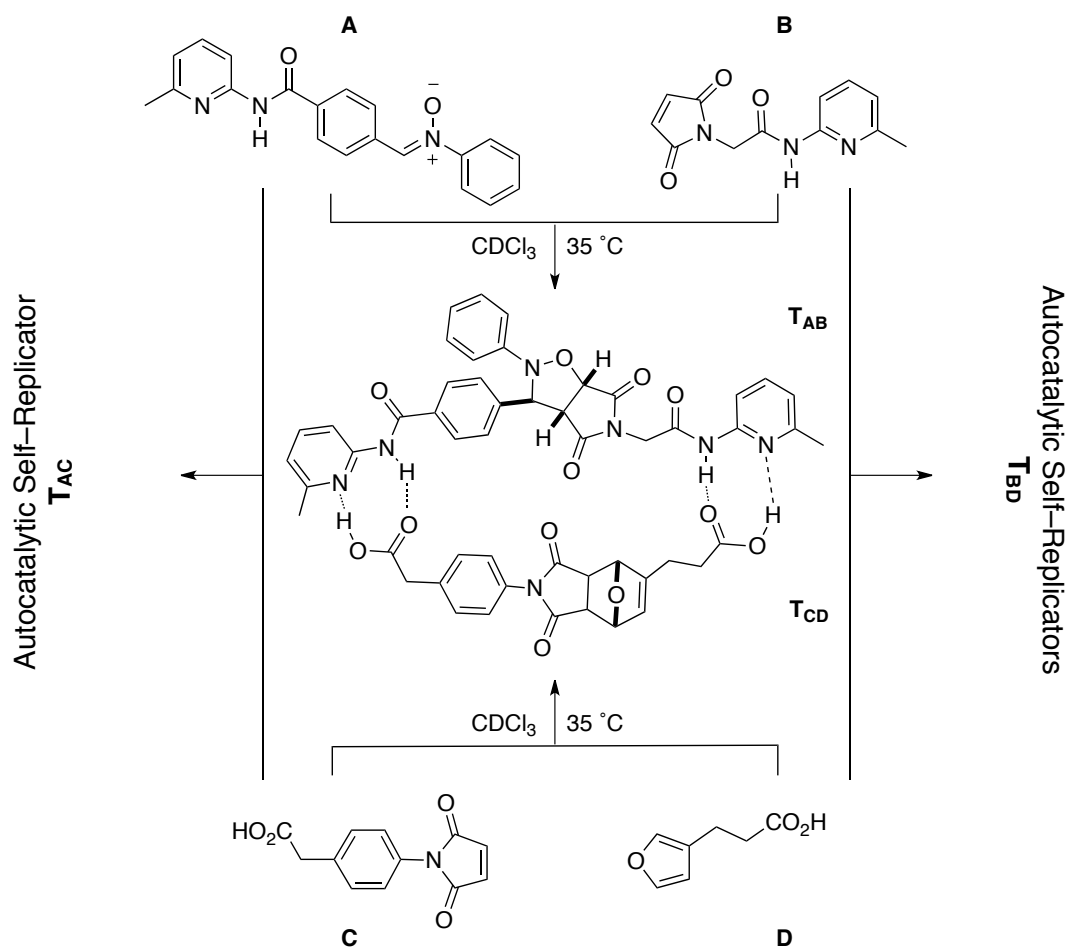
The reciprocal behaviour of the system was revealed when it was found that the amide bond formation between **C** and **D** to give template **T<sub>CD</sub>** was found to be accelerated five-fold, relative to the bimolecular rate, in the presence of two equivalents of preformed **T<sub>AB</sub>**.



**Figure 3.4:** Observed acceleration of the formation of  $T_{AB}$  in the presence of varying quantities of  $T_{CD}$  shown as red points. Data adapted from reference 194.

These results demonstrated a reciprocal replication cycle with each template shown to accelerate the formation of its reciprocal partner in isolation, however when all four components (**A**, **B**, **C** and **D**) were added together at the start of the reaction, reciprocal replication *i.e.* the formation of  $T_{AB}$  and  $T_{CD}$  was not observed. In this case the formation of reactive  $[A \cdot B]$  binary complexes with complementary reagents,  $[A \cdot C]$  and  $[B \cdot D]$ , was observed leading to the production of  $T_{AC}$  and  $T_{BD}$ . The rate of reaction to form  $T_{AC}$  and  $T_{BD}$  was found to be greatly accelerated in comparison to the native bimolecular reactions as a result of molecular recognition forming the reactive binary complexes. The rate of reaction to form  $T_{AC}$  was found to be accelerated 3500-fold in comparison to the recognition disabled control reaction and likewise the rate of reaction to form  $T_{BD}$  was found to be 500-fold greater than its corresponding recognition disabled control. These accelerations in rate as a result of the *pseudo*-intramolecular reaction through a reactive  $[A \cdot B]$  complex were significantly greater than the reciprocal template effects and as such the reciprocal replication mechanism desired in a four component system could not be observed.

More recently, Philp *et al.* were also able to display<sup>195</sup> reciprocal template effects in another four component system. In this system (Scheme 3.2) the reaction of **A** with **B** gives rise to two diastereomeric products *trans*  $T_{AB}$  and *cis*  $T_{AB}$  through a 1,3-dipolar cycloaddition. Likewise, reaction of **C** with **D** gives rise to a pair of diastereoisomeric products *exo*  $T_{CD}$  and *endo*  $T_{CD}$  via a Diels–Alder [4+2] cycloaddition. In isolation, the diastereoselectivity of the reaction of **A** with **B** is relatively poor producing a 3:1 ratio of *trans* : *cis*. However, in the presence of *exo*  $T_{CD}$  the reaction to form *trans*  $T_{AB}$  was accelerated 3.6× with a concurrent inhibition of *cis*  $T_{AB}$  leading to an increase in diastereoselectivity of 14:1 *trans* : *cis*. The increase in rate and selectivity was



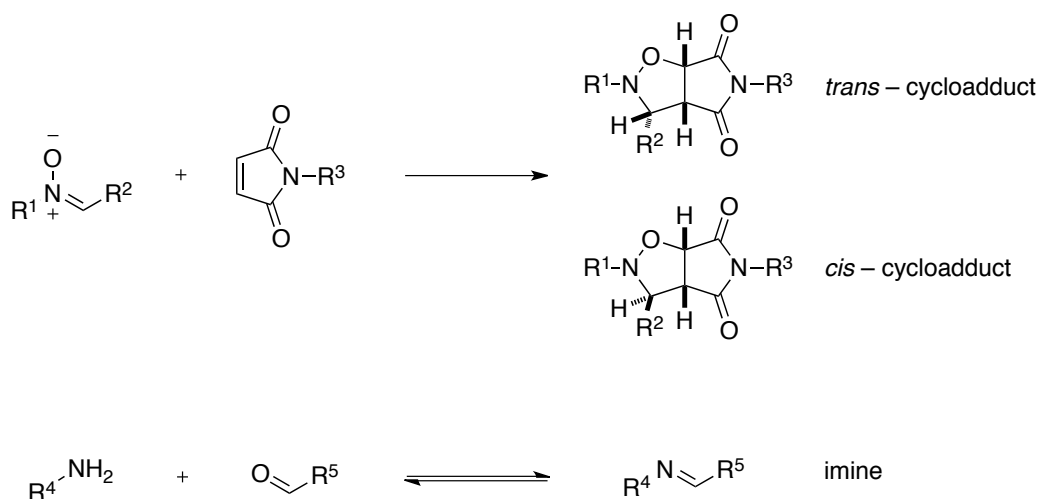
**Scheme 3.2:** Reciprocal template system proposed by Philp *et al.*

achieved by the association of the carboxylic acid of **A** and **B** with their complementary amidopyridine groups on *exo* **T<sub>CD</sub>**. This orientated **A** and **B** into a reactive conformation whilst holding them in a specific orientation promoting the formation of one diastereoisomer. The reciprocal behaviour of the system was confirmed when *trans* **T<sub>AB</sub>** was likewise added to a solution of **C** and **D**. In isolation the reaction of **C** and **D** was slow and poorly diastereoselective producing a 1:1.3 ratio of *exo* **T<sub>CD</sub>** : *endo* **T<sub>CD</sub>**. In the presence of *trans* **T<sub>AB</sub>** the reaction to form *exo* **T<sub>CD</sub>** was accelerated 2.3× leading to an increase in diastereoselectivity of 2:1 *exo* **T<sub>CD</sub>** : *endo* **T<sub>CD</sub>**. Once again, however, when all four components (**A**, **B**, **C** and **D**) were added together at the start of the reaction, reciprocal replication was not observed as the major product. In this case, **A** and **C** react to form *trans* **T<sub>AC</sub>**, a very efficient minimal self-replicator<sup>139</sup> and **B** and **D** react with each other form a pair of diastereoisomeric products *exo* **T<sub>BD</sub>** and *endo* **T<sub>BD</sub>**, both of which were found to be minimal self-replicators.<sup>138</sup>

In both reported cases, the interaction of complementary reagents leads to a reactive complex which prevents reciprocal replication. In order to create a purely reciprocal replicating system these complementary cross combinations must be eliminated. One way this can be achieved is to alter the bond forming reaction between one pair of reagents to a reaction which is orthogonal to the other. In this way **A** will react with only **B** and **C** will only react with **D**. This will eliminate the irreversible reactions consuming complementary reagents.

### 3.2 Design of an Orthogonal Reciprocal Replicator

In order to apply an orthogonal reactivity strategy and create a successful reciprocal replicator, careful consideration of the covalent bond forming reactions must be taken. The 1,3-dipolar cycloaddition between a maleimide and a nitron has been well utilised previously in the group and we have a sound understanding of its reactivity. The cycloaddition forms two diastereoisomeric products, the ratio of which can be used as indicators for the reactivity and selectivity of the system. This reaction was chosen to be one of the two within the new reciprocal system. The second, orthogonal, reaction would have to be able to proceed under the same conditions as the cycloaddition but not possess reagents or products which would react with the nitron, maleimide or cycloadduct.



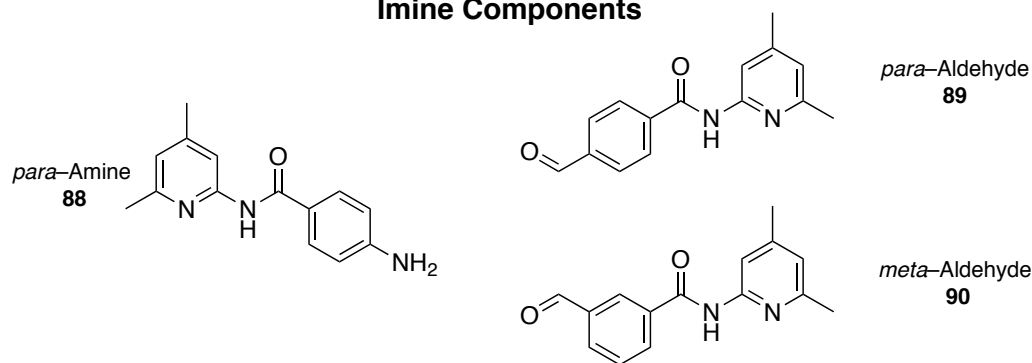
**Scheme 3.3:** Top, the irreversible 1,3-dipolar cycloaddition between a nitron and a maleimide produces two diastereoisomeric products, *trans* and *cis*, and bottom the reversible imine condensation of an amine with an aldehyde.

We have previously shown<sup>140</sup> that imine formation between an aniline and an aldehyde is able to proceed with the aid of recognition mediated processes in a non-polar solvent under similar conditions to the 1,3-dipolar cycloaddition. Thus, imine formation was chosen for the second reaction in the system. As imine formation is known to be reversible<sup>196</sup> this choice added a dynamic component to the reaction which could be interesting to exploit later as reversibility could be useful as an error checking mechanism.

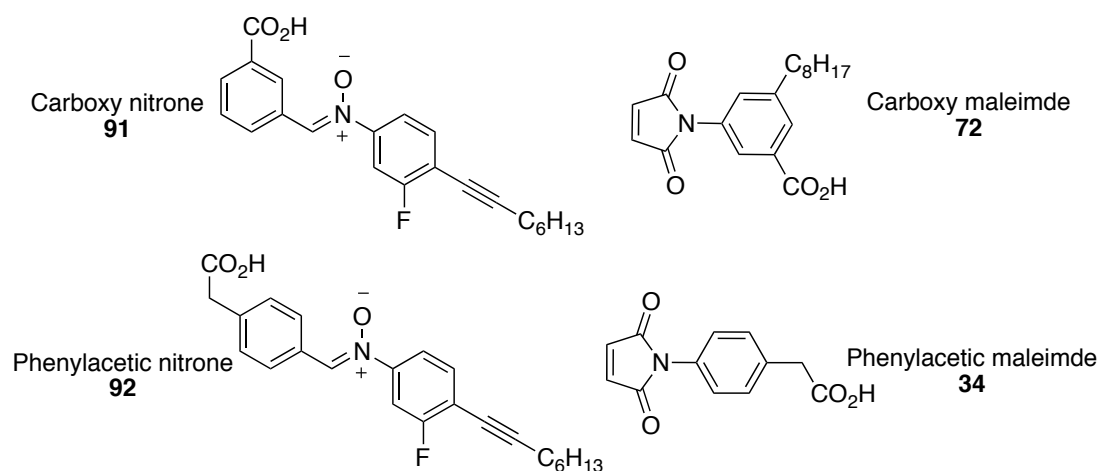
The molecular recognition would again be facilitated by amidopyridine carboxylic acid interactions which have been successfully utilised thus far. For a reciprocal replicating system, each individual template must not be self-complementary in nature but instead complementary to the other template. As such the components for the 1,3-dipolar cycloaddition would be decorated with the carboxylic acid recognition sites with their complementary partners the amidopyridine sites placed on the orthogonal imine reaction components (Figure 3.5).

With the essential recognition and reactive components selected, a series of compounds to bear these features were proposed which would lead to a diversity of structures to be investigated for the reciprocal system. The series comprised of one aniline which could be reacted with either of two aldehydes to give two different imine templates, Table 3.1. The cycloaddition reagents comprised of two nitrones (**91** and **92**) and two maleimides (**72** and **34**). There are four possible combinations of each nitron with each maleimide which gives rise to eight possible templates; four *trans* cycloadducts and four of the diastereoisomeric *cis* cycloadducts. Experience and molecular modelling has shown that *trans* cycloadducts form with the recognition sites open to solution, as required by replication, however, *cis* cycloadducts will form with the recognition sites closed potentially associated together and unavailable for further effective recognition processes.

### Imine Components



### Cycloaddition Components



**Figure 3.5:** (Top) Substrates for the imine condensation bearing the amidopyridine recognition site. (Bottom) Family of substrates designed for the 1,3-dipolar cycloaddition bearing the acid recognition site.

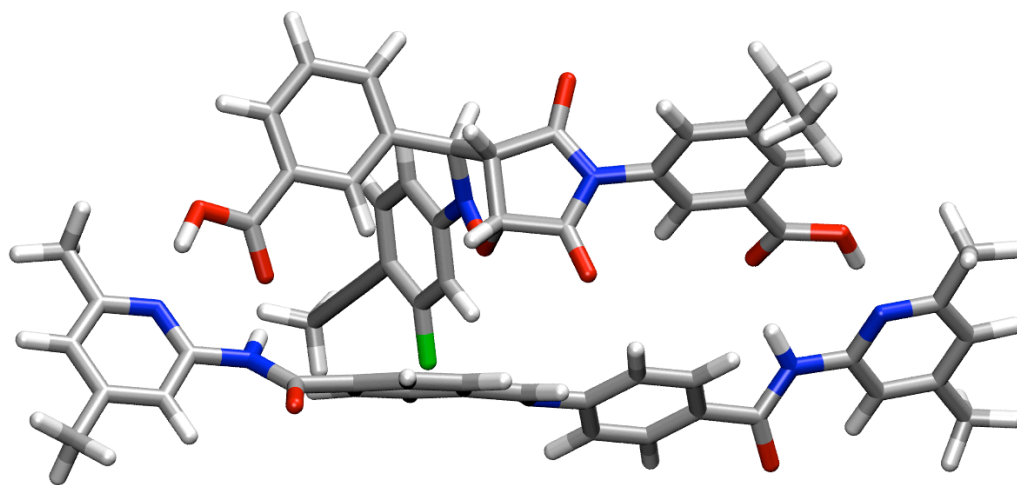
**Table 3.1:** Four unique combinations of each Nitron with each Maleimide leads to the formation of four unique pairs of diastereoisomeric products. Likewise the individual combination each of the two aldehydes with amine **88** leads to two imine products.

Nitron	Maleimide	Product Templates
carboxy Nitron <b>91</b>	carboxy Maleimide <b>72</b>	<i>trans</i> - <b>93</b> + <i>cis</i> - <b>93</b>
carboxy Nitron <b>91</b>	phenylacetic Maleimide <b>34</b>	<i>trans</i> - <b>94</b> + <i>cis</i> - <b>94</b>
phenylacetic Nitron <b>92</b>	carboxy Maleimide <b>72</b>	<i>trans</i> - <b>95</b> + <i>cis</i> - <b>95</b>
phenylacetic Nitron <b>92</b>	phenylacetic Maleimide <b>34</b>	<i>trans</i> - <b>96</b> + <i>cis</i> - <b>96</b>
Aldehyde	Amine	Product Imine
<i>para</i> -Aldehyde <b>89</b>	Amine <b>88</b>	<i>para</i> -Imine <b>97</b>
<i>meta</i> -Aldehyde <b>90</b>	Amine <b>88</b>	<i>meta</i> -Imine <b>98</b>

### 3.2.1 Molecular Modelling

To assess the viability of any of the templates to form a reciprocal replicating system a molecular model of each of the four *trans* cycloadducts and two imines was created. Once optimised in isolation, each *trans* cycloadduct template was individually docked with each of the imine templates and their duplex subjected to an energy minimisation and conformational search to optimise the duplex structure. Once complete, each duplex was visually inspected to identify stable duplexes by looking for relaxed structures, that is no strained bonds or fused rings, whilst also satisfying the four hydrogen bonding sites in the duplex.

Of the eight duplexes analysed, the most promising result was observed when docking *trans*-**93** (the cycloadduct of *carboxy*-nitrone **91** and *carboxy*-maleimide **72**) with *para*-imine **97** (the product of *para*-aldehyde **89** and amine **88**).



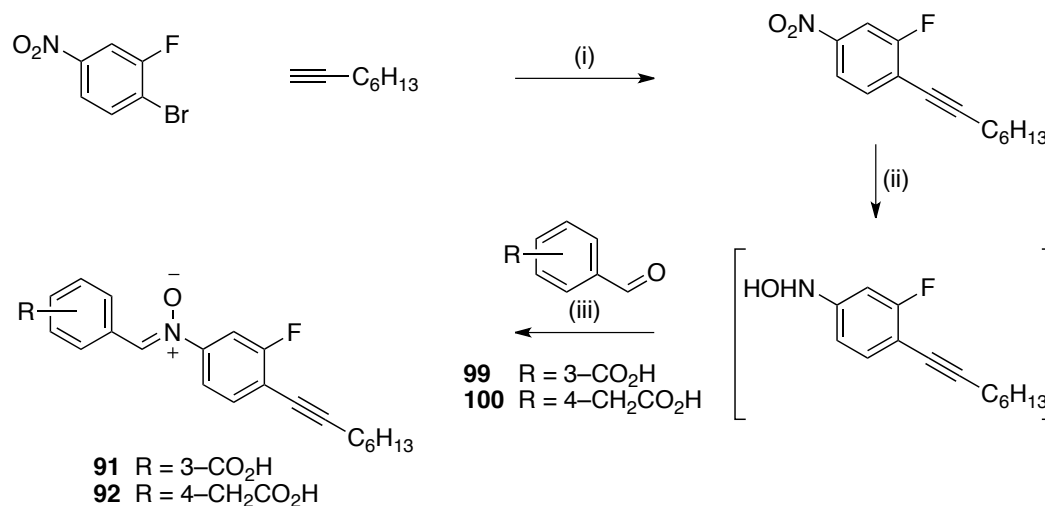
**Figure 3.6:** Molecular mechanics model of the template duplex [*trans*-**93**•*para*-imine **97**]. Carbon atoms are represented in grey, oxygen atoms in red, nitrogen atoms in blue, fluorine atoms in green and hydrogen atoms in light grey. The long alkyl chain on nitrone **91** and maleimide **72** was omitted to reduce computational time whilst retaining the alkyne group to maintain electronic properties.

The model (Figure 3.6) revealed that all four hydrogen bonds were satisfied in the duplex with no straining of bonds or fused rings. Even so, each of the systems were experimentally investigated in order to confirm our hypothesis.



### 3.2.2 Synthesis of Nitrones **91** and **92**

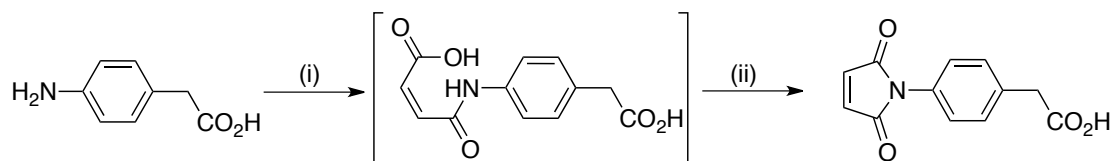
The two target nitrones were designed to include a fluorine tag for monitoring the reactions by  $^{19}\text{F}$  NMR spectroscopy and are decorated with a lipophilic alkyl chain to increase the solubility of the nitrones and subsequent cycloadducts in non-polar solvents.



**Scheme 3.4:** Reagents and conditions (i) Pd(P(Ph)<sub>3</sub>)<sub>2</sub>Cl<sub>2</sub>, CuI, PPh<sub>3</sub>, NEt<sub>3</sub>, Δ, 16 h, 59%, (ii) 5% Rh/C, NH<sub>2</sub>NH<sub>2</sub>·H<sub>2</sub>O, THF, RT, 0.5 h, quantitative, (iii) aldehyde **99** or **100**, EtOH, RT, 16 h, **91** 52% and **92** 43%.

The synthesis of each nitron began from the Sonogashira coupling of 1-bromo-2-fluoro-4-nitrobenzene with 1-octyne which produced the solubilised nitro which was significantly more soluble in CDCl<sub>3</sub> than its bromo precursor. The nitro group was then reduced to the hydroxylamine by a transfer hydrogenation reaction using rhodium on carbon with hydrazine hydrate to yield the hydroxylamine, no reduction of alkyne group was observed. As a consequence of its instability, the hydroxylamine was not isolated and purified but used immediately and reacted with the appropriate aldehyde in ethanol to yield the target nitrones **91** and **92** in acceptable yields. Each of the nitrones was found to be sensitive to hydrolysis when dissolved in CDCl<sub>3</sub> or CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat). We expected that this was as a consequence of the residual acidity of the solvent. In order to overcome this problem, one equivalent of the bulky base collidine was added to NMR spectroscopy samples in either solvent which was found to greatly reduce the rate of hydrolysis and as a result an equivalent of collidine was included in the deuterated solvents when performing all reactions for investigation of the reciprocal replication system.

### 3.2.3 Synthesis of Maleimides **34** and **72**



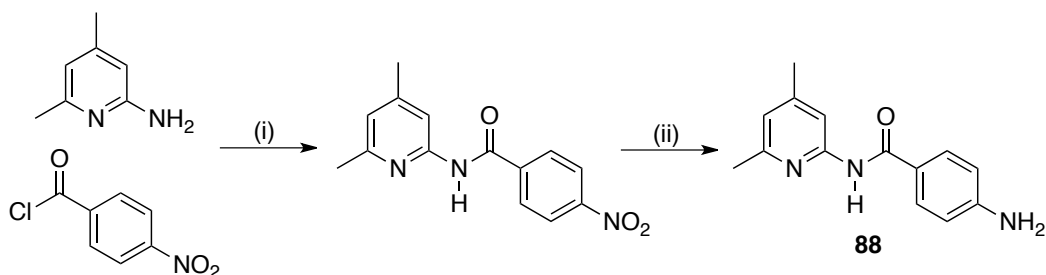
**Scheme 3.5** Reagents and conditions (i) maleic anhydride, AcOH, RT, 16 h, quantitative (ii) AcOH, Δ, 2 h, 33%

The synthesis of phenylacetic maleimide **34** began from 4-aminophenylacetic acid which was reacted with maleic anhydride in acetic acid to furnish the acyclic intermediate which was immediately thermally cyclised to yield the target maleimide.

The synthesis of maleimide **72** was covered previously in Chapter 2 and thus shall not be discussed again here.

### 3.2.4 Synthesis of Aniline **88**

The recognition bearing aniline **88** was readily synthesised in two steps.

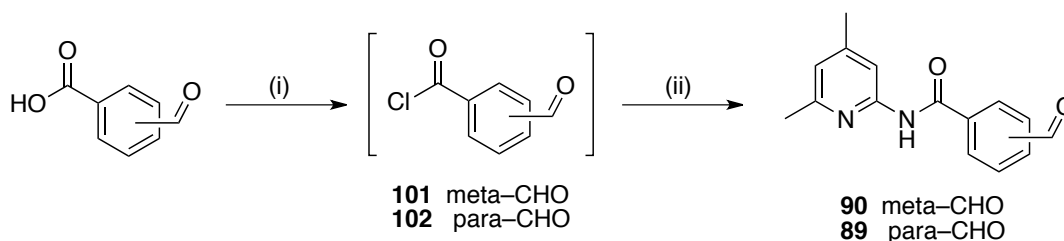


**Scheme 3.6:** Reagents and conditions (i) NEt<sub>3</sub>, CH<sub>2</sub>Cl<sub>2</sub>, 0°C->RT, 16 h, 90% (ii) 10% Pd/C, H<sub>2</sub>, MeOH, THF, RT 99%.

Amide coupling of 2-amino-4,6-dimethylpyridine with 4-nitrobenzoyl chloride in CH<sub>2</sub>Cl<sub>2</sub> and triethylamine yielded the amide in good yields before standard hydrogenation with 10% Pd/C and H<sub>2</sub> in MeOH and THF yielded the target amine **88** in a quantitative yield.

### 3.2.5 Synthesis of Aldehydes **89** and **90**

Aldehydes **89** and **90** were readily synthesised in two steps.

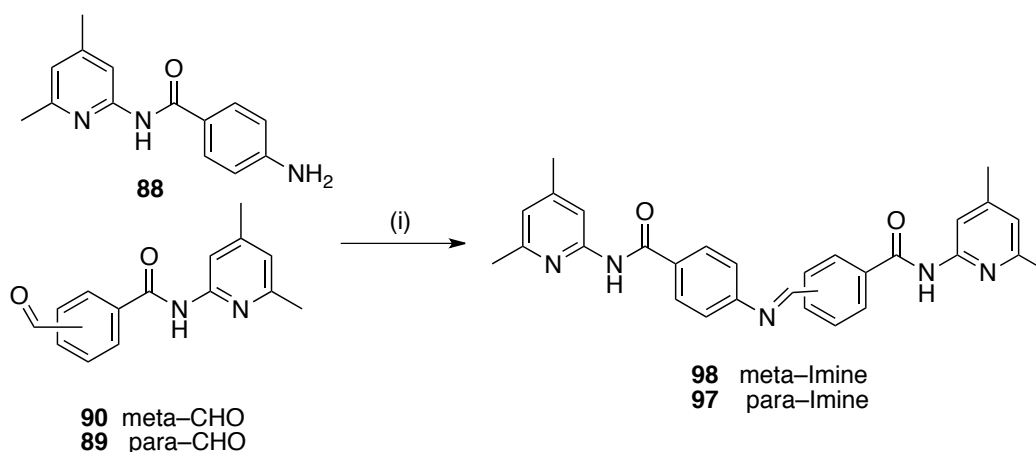


**Scheme 3.7:** Reagents and conditions (i)  $\text{SOCl}_2$ , Toluene, **101** and **102** quantitative (ii) 2,6-dimethylaminopyridine,  $\text{NEt}_3$ ,  $\text{CH}_2\text{Cl}_2$ ,  $0\text{ }^\circ\text{C} \rightarrow \text{RT}$ , 16 h, **89** 82% and **90** 75%

In order to prepare for an amide couple, each carboxybenzaldehyde was activated as an acid chloride by refluxing the acid with thionyl chloride in toluene. The acid chlorides were used immediately to avoid hydrolysis, each reacted with 2-amino-4,6-dimethylpyridine in  $\text{CH}_2\text{Cl}_2$  and triethylamine to furnish the target aldehydes **89** and **90** in good yields.

### 3.2.6 Synthesis of Imines **97** and **98**

To assess the ability of each imine to act as a template for the cycloaddition of each of the four combinations of nitrones and maleimides, the two imine templates must be pre-synthesised.

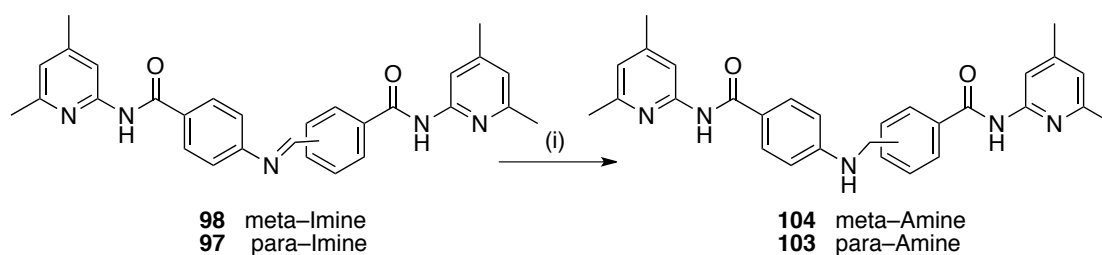


**Scheme 3.8:** Reagents and conditions (i) EtOH, RT, 40 h, **98** 34% and **97** 86%

Condensation of amine **88** with aldehyde **89** or **90** in ethanol furnished imine **97** or **98** respectively in acceptable yields. The imines were able to be stored in the freezer ( $-20\text{ }^{\circ}\text{C}$ ) until required.

### 3.2.7 Amine Synthesis

As a result of the formation of the imines being reversible, we chose to fix the structure by reducing the imine group to the corresponding amine. This reduction had the advantage of kinetically trapping the structure as the amine would not be able to undergo the reverse reaction. The amine would however also possess an extra degree of conformational flexibility around the amine bond as opposed to the rigid imine centre.



**Scheme 3.9:** Reagents and conditions (i) NaBH<sub>4</sub>, THF, RT, 16 h, **104** 54% and **103** quantitative.

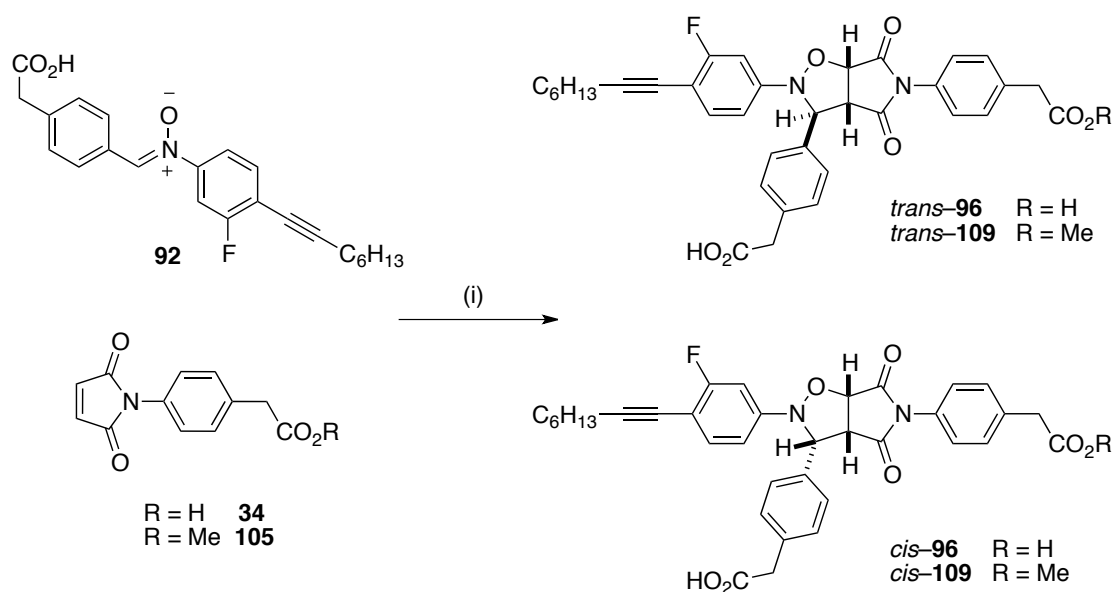
Each imine was reduced using sodium borohydride in THF to furnish amine **103** and **104** in reasonable yields after purification by recrystallisation.

### 3.3 Template Effects in a Reciprocal System

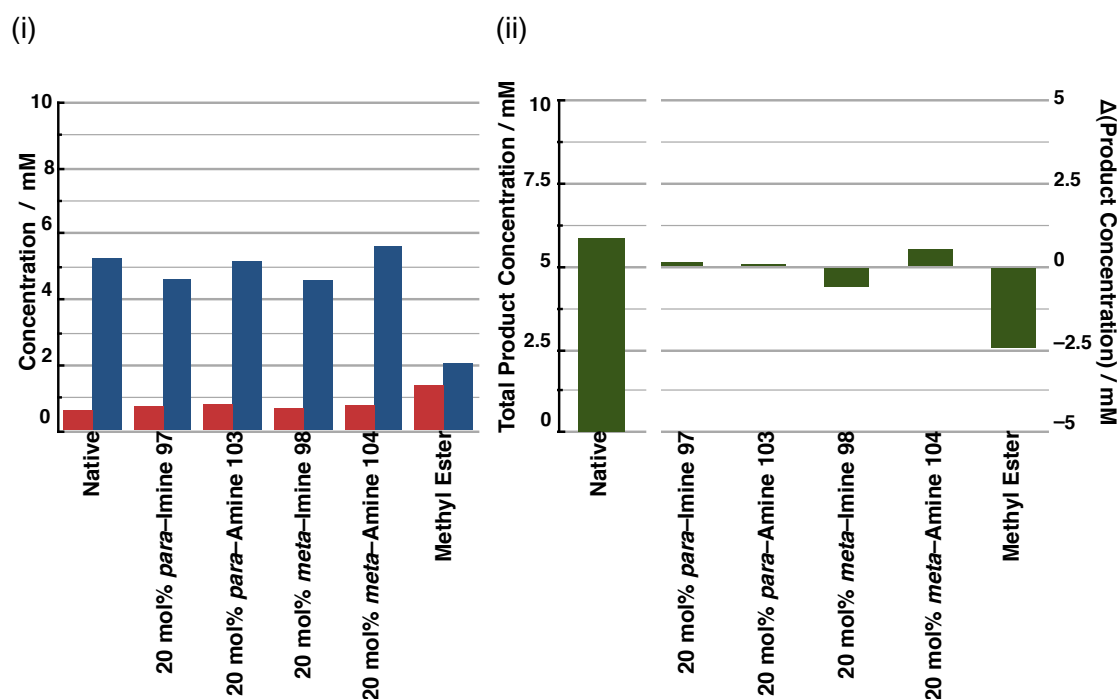
With all the reagents at hand, a series of experiments designed to screen for reciprocal template effects were carried out. Each unique combination of nitron and maleimide was incubated together at 10 mM in CD<sub>2</sub>Cl<sub>2</sub> / *p*-TSA (sat), collidine (10 mM). This reaction was incubated at 10 °C for 16 h and then assayed by <sup>19</sup>F NMR spectroscopy. The relative concentration of each of the species in the mixture was then determined by deconvolution. These results constituted the native behaviour of each reaction.

Concurrently, each reaction was also performed in the presence of 20 mol% of a template; imine **97** or **98** or amine **103** or **104**. These reactions were performed under identical conditions and were assayed in the same manner. From these results, a direct comparison could then be made between the product distribution of the native reactions to the product distribution of the reactions doped by each of the four templates. The expectation from doping the cycloaddition reaction with an imine or amine template would be that the template would associate the maleimide and nitronone reagents on its recognition sites which would promote the formation of the *trans* cycloadduct as demonstrated by molecular modelling. As the reaction is rendered pseudo intramolecular, the rate of cycloaddition reaction to produce the *trans* cycloadduct should occur at an increased rate relative to its untemplated reaction. Therefore when an imine or amine template is successful in associating the two substrates and promoting their reaction, we would expect to see an increase in the production of the *trans* cycloadduct as well as an increase in the overall conversion with respect to the native systems. During the course of the screening, recognition mediated behaviour was observed during the native reactions between maleimide and nitronone. We expected this could be as a result of acid dimerisation and as a result we also performed recognition inhibited reactions with one acid group blocked as an ester.

### 3.3.1 Phenylacetic Nitron 92 and Phenylacetic Maleimide 34

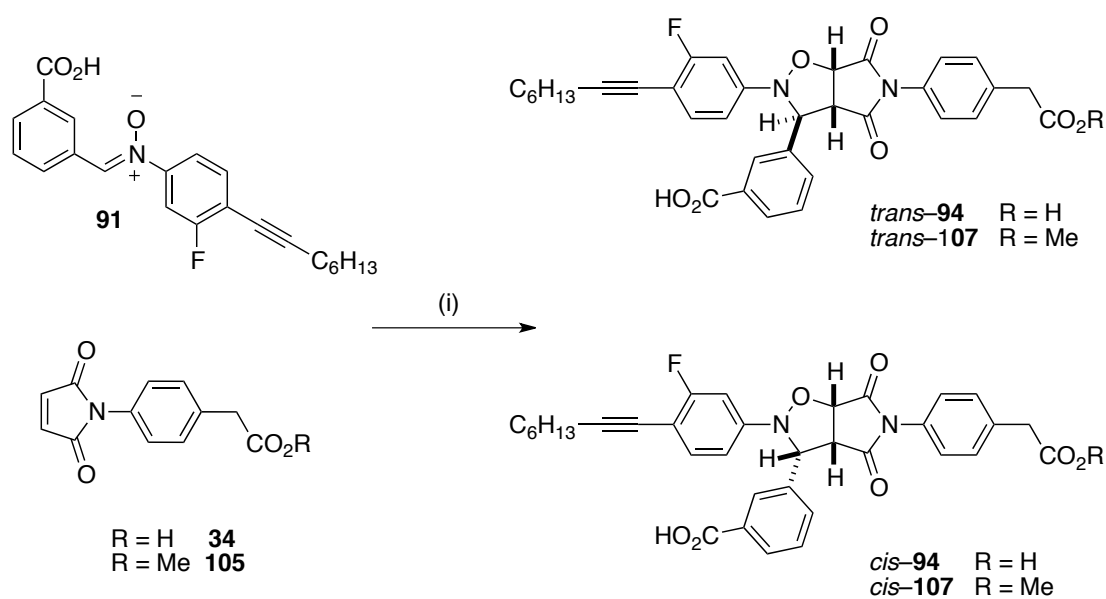


**Scheme 3.10:** Reagents and conditions (i) 0% or 20 mol% imine **97** or **98** or amine **103** or **104**  $\text{CD}_2\text{Cl}_2$  /  $p$ -TSA (sat), collidine, 10 mM, 10 °C, 16 h.

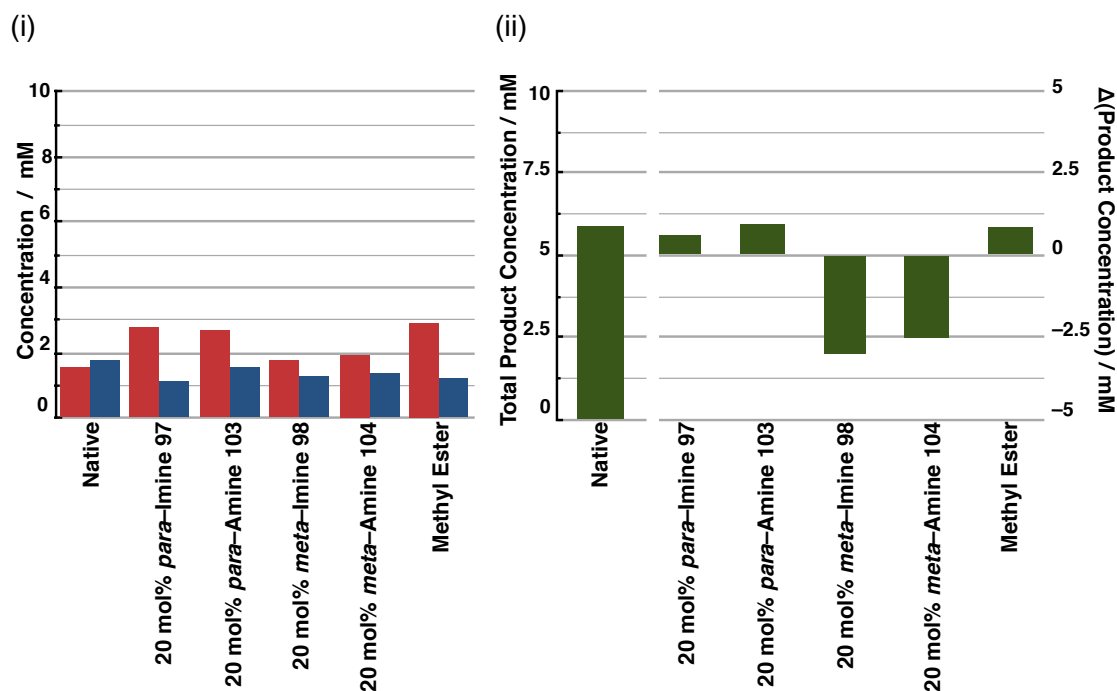


**Figure 3.7:** (i) Concentration of *trans* (red) and *cis* (blue) cycloadducts for each of the reactions after 16 h as determined by deconvolution of 376.5 MHz  $^{19}\text{F}$  NMR spectroscopy with the identity of added template displayed below each result. (ii) The first result is the total concentration of the native reaction (no template added) with subsequent results the change in total product concentration of each template doped reaction.

### 3.3.2 Carboxy Nitron 91 and Phenylacetic Maleimide 34

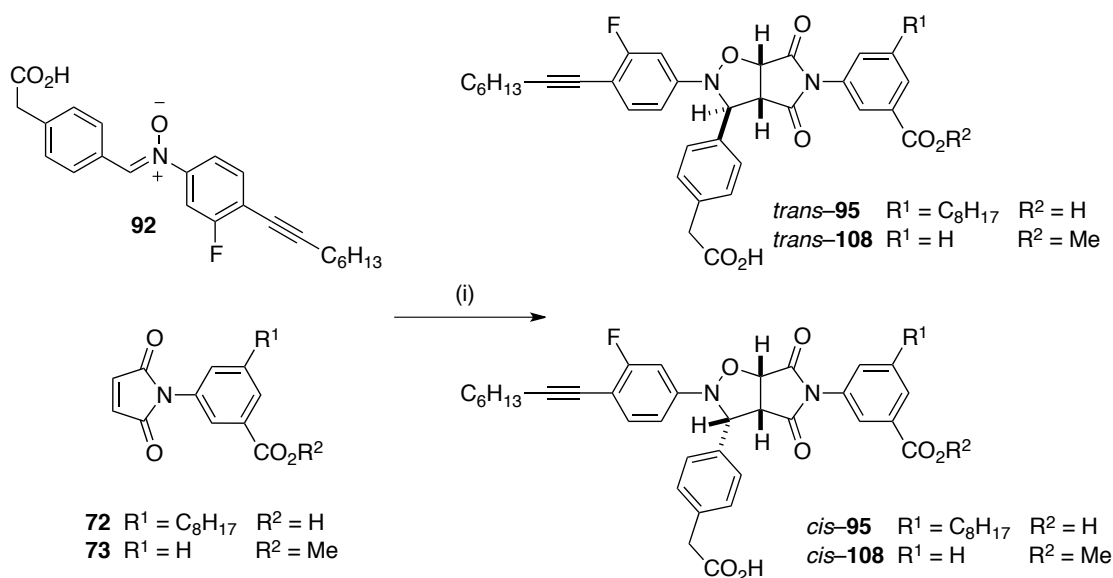


**Scheme 3.11:** Reagents and conditions (i) 0% or 20 mol% imine **97** or **98** or amine **103** or **104** CD<sub>2</sub>Cl<sub>2</sub> / *p*-TSA (sat), collidine, 10 mM, 10 °C, 16 h.

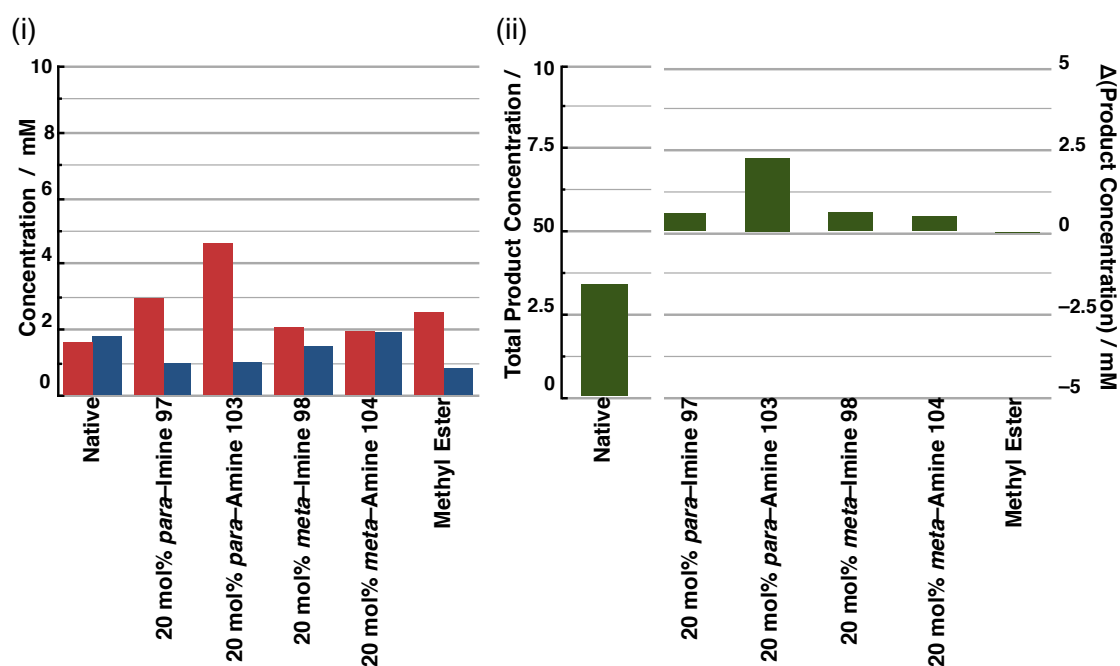


**Figure 3.8:** (i) Concentration of *trans* (red) and *cis* (blue) cycloadducts for each of the reactions after 16 h as determined by deconvolution of 376.5 MHz <sup>19</sup>F NMR spectroscopy with the identity of added template displayed below each result. (ii) The first result is the total concentration of the native reaction (no template added) with subsequent results the change in total product concentration of each template doped reaction.

### 3.3.3 Phenylacetic Nitron 92 and Carboxy Maleimide 72



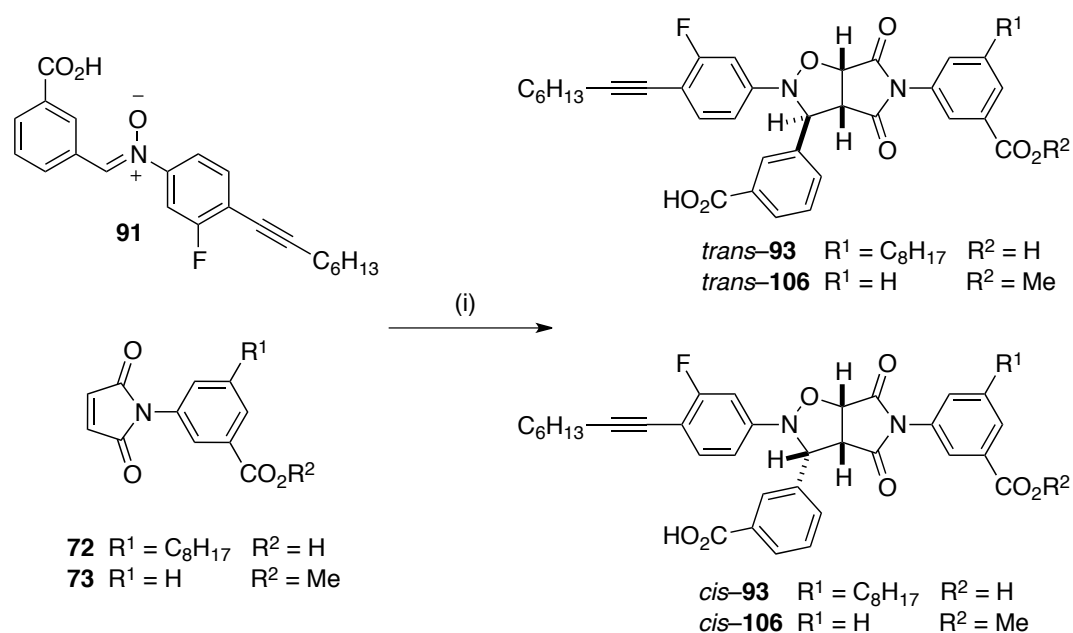
**Scheme 3.12:** Reagents and conditions (i) 0% or 20 mol% imine **97** or **98** or amine **103** or **104**  $\text{CD}_2\text{Cl}_2$  / *p*-TSA (sat), collidine, 10 mM, 10 °C, 16 h.



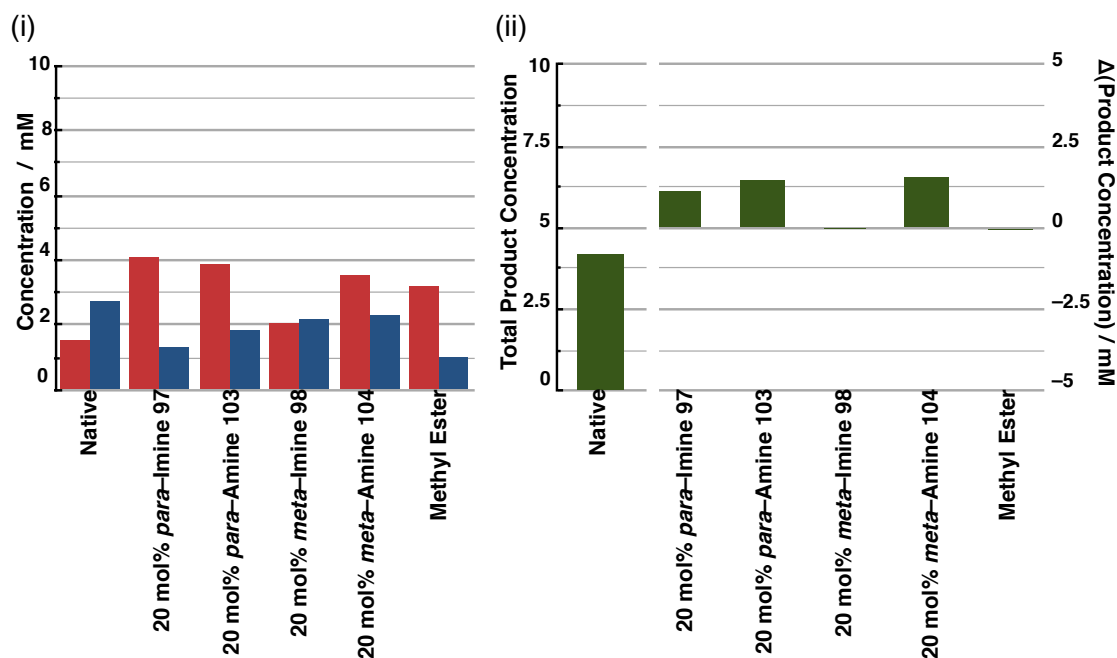
**Figure 3.9:** (i) Concentration of *trans* (red) and *cis* (blue) cycloadducts for each of the reactions after 16 h as determined by deconvolution of 376.5 MHz  $^{19}\text{F}$  NMR spectroscopy with the identity of added template displayed below each result. (ii) The first result is the total concentration of the native reaction (no template added) with subsequent results the change in total product concentration of each template doped reaction.



### 3.3.4 Carboxy Nitron 91 and Carboxy Maleimide 72



**Scheme 3.13:** Reagents and conditions (i) 0% or 20 mol% imine **97** or **98** or amine **103** or **104**  $\text{CD}_2\text{Cl}_2$  /  $p$ -TSA (sat), collidine, 10 mM, 10 °C, 16 h.



**Figure 3.10:** (i) Concentration of *trans* (red) and *cis* (blue) cycloadducts for each of the reactions after 16 h as determined by deconvolution of 376.5 MHz  $^{19}\text{F}$  NMR spectroscopy with the identity of added template displayed below each result. (ii) The first result is the total concentration of the native reaction (no template added) with subsequent results the change in total product concentration of each template doped reaction.

In each of the native reactions, there was an unexpected and interesting result. In the absence of recognition mediated processes, the *trans* : *cis* ratio for a 1,3 dipolar cycloaddition should be close to 3:1. In each aforementioned case however, the ratio was reversed favouring the production of the *cis* cycloadduct. The highest selectivity for the *cis* cycloadduct in the native reaction was seen in the reaction between phenylacetic nitron **92** and phenylacetic maleimide **34** (Scheme 3.10 and Figure 3.7). Addition of imine or amine templates to this reaction had little effect in altering the *trans* : *cis* ratio or the overall conversion and these interesting results warranted further investigation which will be discussed shortly, Section 3.4.

The clearest evidence for a successful template directed reaction was observed in the reaction of carboxy nitron **91** and carboxy maleimide **72** (Section 3.3.4, Scheme 3.13). The results of the native reaction (Figure 3.10) showed the reaction proceeded with low selectivity producing 1.5 mM *trans*-**93** and 2.7 mM *cis*-**93** representing a *trans* : *cis* ratio of 1:1.8 and overall conversion of 42%. Addition of 20 mol% *para*-imine **97** had the effect of reversing the selectivity now producing predominantly *trans*-**93**, 4.1 mM with 1.3 mM *cis*-**93** representing a new *trans* : *cis* ratio of 3.2:1 and an overall conversion which has now increased to 54%. Addition of 20 mol% of the conformationally more flexible *para*-amine **103** template also displayed the same relative enhancement of the *trans* cycloadduct over the *cis* cycloadduct however less efficiently with 3.9 mM *trans*-**93** and 1.9 mM *cis*-**93** produced representing a *trans* : *cis* ratio of 2.1 : 1 and an overall conversion of 58%. In both cases however we observed the desired increase in the production of the *trans* cycloadduct with simultaneous inhibition of the *cis* cycloadduct.

Similar enhancements of *trans* cycloadducts were observed in the remaining systems, Section 3.3.2 and Section 3.3.3, however the most promising result was observed using the *para*-imine / amine templates with carboxy maleimide **72** and carboxy nitron **91**.

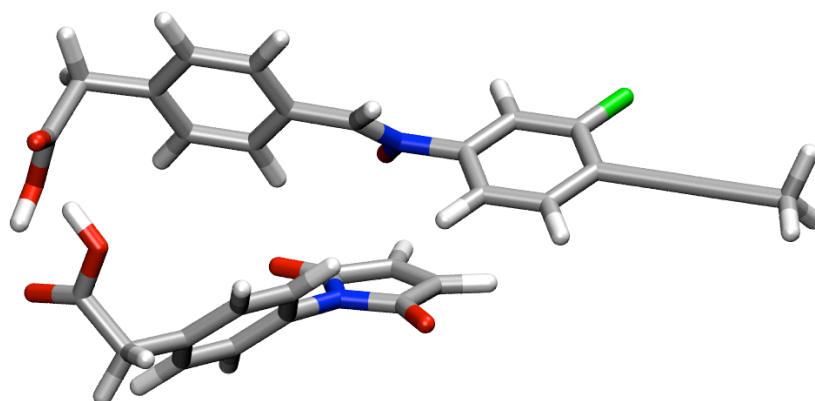
From the whole series, a pattern can be inferred as to the efficiency of the templates. In the majority of cases, addition of the amine template was less effective than its imine counterpart. This can be attributed to the fact that although the amine is kinetically stable, it bears an extra degree of flexibility which arises from the saturated sp<sup>3</sup> nitrogen of the amine. The unsaturated imine bears a conformationally

restricted  $sp^2$  nitrogen and as such the imine template will have less degrees of freedom which will make it more efficient at holding the complementary reagents in a reactive conformation.

### 3.4 Investigation of the Origin of *Cis* Selectivity

The unexpected result arising from the reaction between phenylacetic nitron **92** with phenylacetic maleimide **34** warranted further investigation. Conventionally, a 1,3-dipolar cycloaddition between a maleimide and a nitron will produce a 3:1 ratio of *trans* : *cis* cycloadducts in the absence of any recognition mediated effects. In this case, however, the diastereoselectivity has been reversed as the major product is the *cis* cycloadduct with 0.6 mM *trans*-**96** and 5.3 mM *cis*-**96** produced representing a *trans* : *cis* ratio of 1:8.8 which suggests that some form of recognition mediated process is occurring. The addition of any template has little effect in promoting the *trans* cycloadduct and subsequent overall conversion. Our hypothesis was that dimerisation of the acid groups would potentially be responsible for the creation of an efficient [**A**•**B**] complex which orientates the reactive sites into close proximity for the production of the *cis* cycloadduct. The first evidence to support this theory was observed in the reaction in which one of the acid groups was blocked as a methyl ester (Figure 3.7). The result of this effect was the dramatic decrease in the selectivity for the *cis* cycloadduct, with 1.4 mM *trans*-**109** and 2.1 mM *cis*-**109** produced representing a *trans* : *cis* ratio of 1:1.4 and a lower overall conversion of 35%. We next turned to molecular modelling to assess the possibility of an acid dimerisation giving rise to a reactive [**A**•**B**] complex.

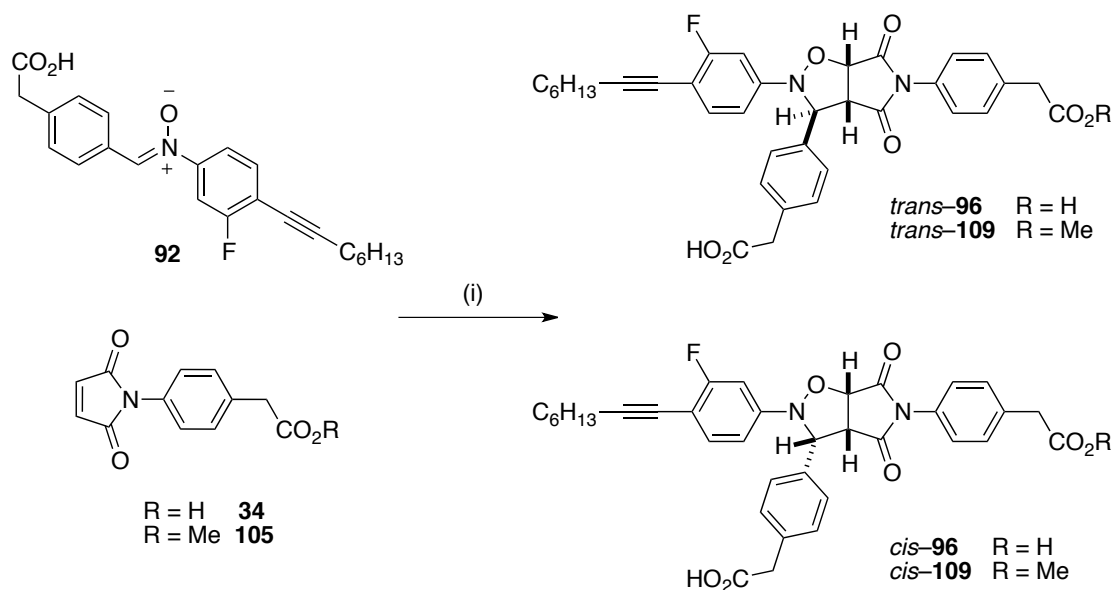
A molecular model of the cycloadduct product *cis*-**96** was created and subjected to an energy minimisation and conformational search to identify the lowest energy conformer. From this, the transition state leading to this product could then be identified (Figure 3.11).



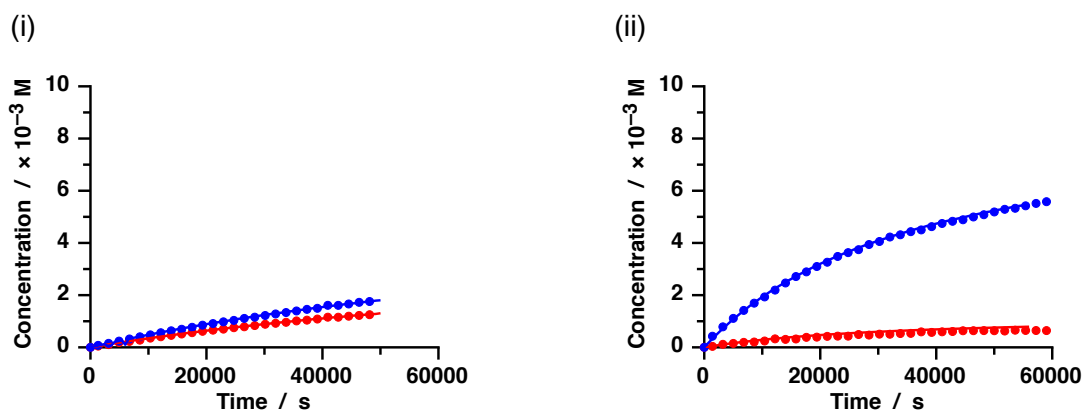
**Figure 3.11:** Molecular model of the transition state of the reaction of phenylacetic maleimide **34** and phenylacetic nitron **92** leading to the formation of *cis*-**96**. Carbon atoms are represented in grey, oxygen atoms in red, nitrogen atoms in blue, fluorine atoms in green and hydrogen atoms in light grey. The alkyl chain on nitron **92** and maleimide **34** was omitted to reduce computational time whilst retaining the alkyne group to maintain electronic properties.

The molecular model revealed that the formation of the acid dimer with two hydrogen bonds supported the transition state. The ability of the acid duplex to create a reactive [A•B] complex appears to be the clear source for the *cis* selectivity in the system. In order to confirm this observation, we next performed a kinetic analysis of the reaction to confirm this behaviour and measure the efficiency of the [A•B] process.

Initially to establish the native reactivity of the 1,3 dipolar cycloaddition a recognition disabled control reaction was carried out between nitron **92** and maleimide ester **105** in CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat), collidine at 10 °C and monitored for 14 h by <sup>19</sup>F NMR spectroscopy. A concentration vs. time profile was constructed after deconvolution of the <sup>19</sup>F NMR spectra and revealed (Figure 3.12 (i)) a slow rate of formation of each of the two cycloadducts. The kinetic profile was fitted to a model for a bimolecular reaction using SimFit and revealed the rate of reactions,  $k_{\text{trans}}$  and  $k_{\text{cis}}$  to be  $3.79 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  and  $5.28 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  respectively.



**Scheme 3.14:** Reagents and conditions (i)  $\text{CD}_2\text{Cl}_2$  / *p*-TSA (sat), Collidine, 10 mM, 10 °C, 16 h.



**Figure 3.12:** Concentration vs. time profiles displaying the results of kinetic monitoring of (i) the recognition disabled control reaction of nitrone **92** and with maleimide ester **73** and (ii) recognition enabled reaction of nitrone **92** with maleimide acid **34** (ii). *Cis*-cycloadducts (●), *trans*-cycloadducts (●). The results of fitting the experimental data to a kinetic model for an  $[\text{A}\cdot\text{B}]$  complex reaction model shown as lines, *cis*-cycloadducts (—) *trans*-cycloadducts (—). Reaction conditions  $\text{CD}_2\text{Cl}_2$  / *p*-TSA (sat), Collidine, 10 mM, 10 °C, 16 h.

**Table 3.2:** Results of kinetic simulation and fitting of the kinetic results from the native reaction and the control reactions.

	$k_{\text{trans}}$	$k_{\text{cis}}$
Bimolecular Rate / $\times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$	3.79	5.28
Recognition mediated / $\times 10^{-4} \text{ s}^{-1}$	—	6.08
$EM / M$	—	1.2

The influence of recognition in the system was next investigated by monitoring the reaction between nitron **92** and maleimide acid **34** performed and assayed under identical conditions. The concentration vs. time profile (Figure 3.12 (ii)) constructed for the recognition mediated reaction shows an increase in the rate of production of *cis*-**96** and inhibition of the production of *trans*-**96**. The maximum rate of production of *cis*-**96** was observed at the start of the reaction when the concentration of reagents is greatest, consistent with a reaction promoted *via* an **[A•B]** binary complex. The reaction was again subjected to kinetic simulation and fit to a model for an **[A•B]** binary complex reaction using SimFit to extract rate information, Table 3.2, see appendix. From the results we ascertained the dimerisation constant of the acid groups to be 41 M<sup>-1</sup> and the rate of reaction in the binary complex,  $k_{AB}$ , to be 6.08 × 10<sup>-4</sup> s<sup>-1</sup> which reveals the *EM* of the reaction through the binary complex to be 1.2 M. The results of the kinetic simulation and fitting are consistent with the expected low dimerisation constant of an acid and also highlight the efficiency of the **[A•B]** complex in promoting the reaction. It is clear that the presence of two acid groups in the system allows recognition between the nitron and maleimide through acid dimerisation bringing the reactive sites of the nitron and maleimide into close proximity, promoting their reaction to form the *cis* cycloadduct. Key to the success of the **[A•B]** channel in promoting the reaction between these reagents was likely to be the extra conformational flexibility afforded by the CH<sub>2</sub> group. This will allow the substrates to arrange themselves into a reactive conformation after association of the acid groups.

In contrast, the screening results showed the lowest enhancement of *cis* cycloaddition is produced when carboxy nitron **91** and carboxy maleimide **72** are used, Section 3.3.4. Dimerisation of the two acid groups in this case will be the most conformationally restricted with no CH<sub>2</sub> linker to allow flexibility. In this case it is unlikely that they will be able to produce a reactive conformation.

As our ultimate goal is the creation of a reciprocal system in which the rate of cycloaddition is promoted by a complementary template, the selfish recognition mediated behaviour between phenylacetic nitron **92** and maleimide **34** is not advantageous and although interesting is not a constructive property for a reciprocal system. Instead our attention focused on the systems which displayed low *cis*

selectivity in their native reactions thus avoiding [A•B] binary complexes and an increase in *trans* selectivity for their template doped reactions.

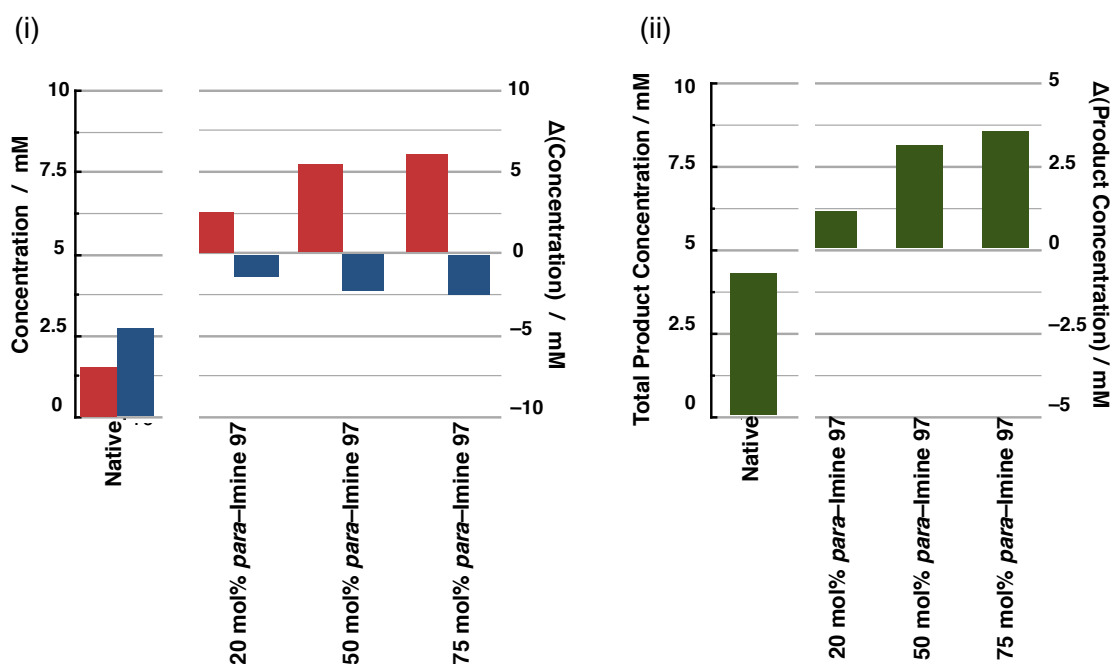
### 3.5 Acceleration of the Reaction of *m*-Nitron 91 with *m*-Maleimide 72 by an Imine Template

The most promising result came from the group of reactions between carboxy nitron **91** and carboxyl maleimide **72**. The native reaction once again displayed deviation from the customary 3:1 *trans* : *cis* ratio, in this case the *cis* cycloadduct was the major product reaching 2.7 mM and 1.5 mM of *trans*-**93** representing a *trans* : *cis* ratio of 1:1.8 and an overall conversion of 4.2 mM. The enhancement of the *cis* cycloadduct is not as great as in the aforementioned phenylacetic nitron **92** and maleimide **34** system, Section 3.3.1, which highlights that the lack of conformational freedom of the benzoic acid recognition site restricts the formation of a reactive [A•B] binary complex. When the reaction between nitron **91** and maleimide **72** was doped with 20 mol% *para*-imine **97** there was a clear enhancement in the formation of the *trans*-**93** of which 4.1 mM was produced with concurrent inhibition of *cis*-**93** of which 1.3 mM was produced representing a *trans* : *cis* ratio of 3.1:1 and an increased overall conversion of 54%. It is clear that the effect of adding the *para*-imine **97** template was to enhance the formation of the *trans*-**93** cycloadduct as expected from the molecular modelling. When 20 mol% of the structurally similar *para*-amine **103** was doped into the cycloaddition reaction, a similar enhancement of *trans*-**93** was observed. In this case, 3.9 mM of *trans*-**93** and 1.8 mM *cis*-**93** were produced representing a *trans* : *cis* ratio of 2.1:1 and an overall conversion of 57%. Comparison of the doping results from the amine and the imine highlighted that the more conformationally rigid imine was the more efficient template engineering a higher concentration of *trans*-**93** and higher *trans* : *cis* ratio. The increased conformational flexibility of the amine template means it is less efficient at acting as a template for the cycloaddition reaction to produce *trans*-**93**.

Doping the cycloaddition reaction with *meta*-imine **98** provided only a minor enhancement of *trans*-**93** which reached 2.1 mM however *cis*-**93** was still the dominant product reacting 2.2 mM. The different geometry of the recognition sites and structural rigidity of the *meta*-imine template is in this case not producing an efficient ternary [*meta*-imine **98**•nitron **91**•maleimide **72**] complex. *meta*-Amine **104**

which has some extra conformational freedom is able to better accommodate the arrangement of nitrene **91** and maleimide **72** into a reactive conformation which leads to the production of 3.5 mM *trans*-**93** and 2.3 mM *cis*-**93**, an increase in *trans* selectivity in comparison to the *meta*-imine however less efficient than the *para* templates.

Our attention therefore focused on the *para*-imine template and a further series of experiments were performed to demonstrate the ability of the *para* templates to increase the production of *trans*-**93**. The cycloaddition reactions were performed in the same manner as before and under the same conditions this time in the presence of increasing concentrations of *para*-imine **97** at 50 mol% and 75 mol%. We would expect that increasing the concentration of the template added should increase the selectivity for the production of *trans*-**93** and further inhibit *cis*-**93** formation.



**Figure 3.13:** (i) Concentration of *trans* (red) and *cis* (blue) cycloadducts for the native reaction with no added template after 16 h and then the change in concentration of each product with each subsequent template doped reaction. (ii) The total concentration for the native reaction followed by the change in total concentration of each template doped reaction.

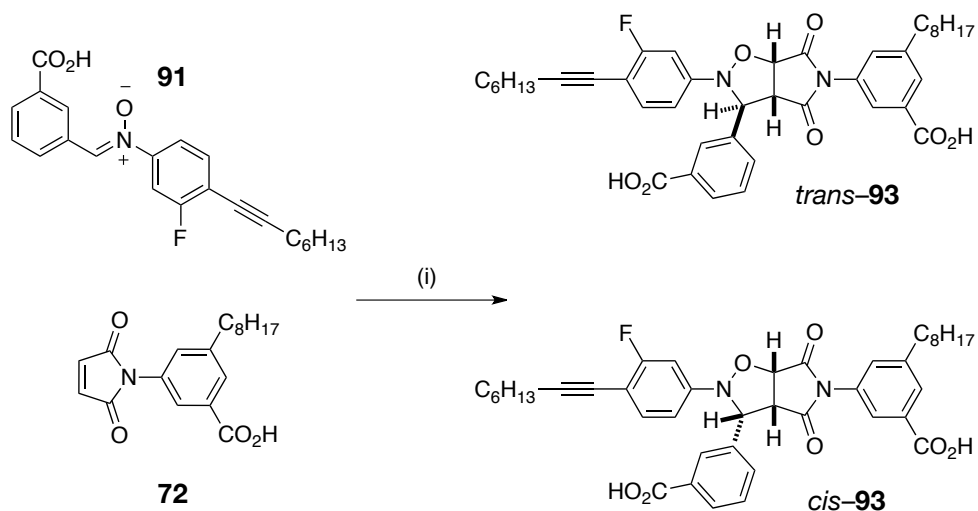
Increasing the concentration of *para*-imine **97** present upon initiation of the reaction (Figure 3.13 (i)) produces a significant increase in the selectivity for *trans*-**93** further evidence of the template effect. When 50 mol% of *para*-imine **97** was added to the



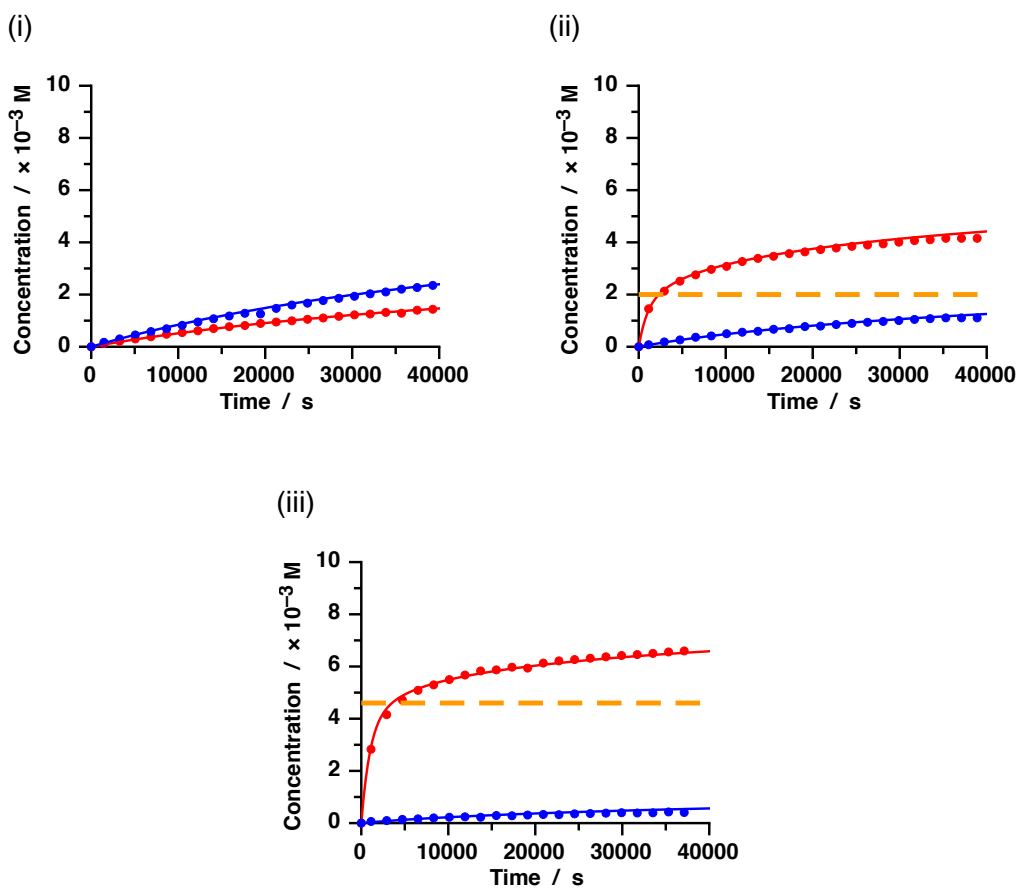
cycloaddition, the concentration of *trans*-**93** produced was increased to 7.0 mM with the formation of *cis*-**93** inhibited forming 0.5 mM representing a 14:1 *trans* : *cis* ratio and 75% overall conversion. Further increasing the concentration of *para*-imine **97** added to the cycloaddition to 75 mol% engineered a further increase in selectivity for *trans*-**93**. In this case 7.6 mM *trans*-**93** and only 0.2 mM *cis*-**93** was produced representing a *trans* : *cis* ratio of 38:1 and an overall conversion of 78% (Figure 3.13 (ii)).

From the trend of increasing the production of *trans*-**93** with increasing concentration of *para*-imine **97** we could draw two conclusions. The first was that the imine template was successfully arranging the nitron and maleimide into a reactive conformation which promoted the formation of the *trans* cycloadduct as anticipated from our molecular modelling. Secondly, with the observation that increasing the concentration of template doping, we see a further increase in selectivity and overall conversion, we expect that there is a degree of product inhibition present. That is after a molecule of *para*-imine **97** has successfully promoted the formation of the *trans* cycloadduct, the template duplex [*trans*-**93**•*para*-imine **97**] is so strongly associated together that there is a very low amount of dissociation and subsequent turnover. Strong product inhibition is a known phenomena which has previously been observed<sup>135,197</sup> within systems investigated by the group. Nevertheless, these results bode well for the for the formation of a reciprocal system and warranted further investigation by kinetic analysis.

Initially we wished to establish the native reactivity of the 1,3 dipolar cycloaddition in this system. Accordingly we performed the native reaction between nitron **91** and maleimide **72** at 10 mM in CD<sub>2</sub>Cl<sub>2</sub> / *p*-TSA (sat), collidine at 10 °C and monitored the reaction for 12 h by <sup>19</sup>F NMR spectroscopy. The relative concentration of each product was calculated by deconvolution to produce a concentration vs. time profile (Figure 3.14 (i)) which revealed a low overall conversion in the system reaching 38% after 12 h. The major product of the system was *cis*-**93** which reached a concentration of 2.4 mM with 1.4 mM of *trans*-**93** produced representing a *trans* : *cis* ratio of 1:1.7.



**Scheme 3.15:** Reagents and conditions (i) *para*-Imine **97** 0, 20 or 50 mol%, CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat)/collidine, 10 mM, 10 °C, 12 h



**Figure 3.14:** Concentration vs. time profiles of the reaction of nitron **91** with maleimide **72** in the presence of various concentrations of *para*-imine **97** (---) doping to form *trans*-**93** (●) and *cis*-**93** (●). (i) 0 mol% *para*-Imine **97** added. (ii) 20 mol% *para*-Imine **97** added. (iii) 46 mol% *para*-Imine **97** added. The results of kinetic simulation and fitting of the experimental data (points) to a template directed reaction model shown as lines, *cis*-cycloadducts (—), *trans*-cycloadducts (—). Reaction conditions: CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat)/collidine, 10 mM, 10 °C, 12 h.

**Table 3.3:** Results of kinetic simulation and fitting of the kinetic results from the native reaction and the imine doped reactions.

	$k_{\text{trans}}$	$k_{\text{cis}}$
Control / $\times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$	6.04	9.86
Template Directed Reaction / $\times 10^{-4} \text{ s}^{-1}$	28.1	—
$EM / M$	4.65	—
Connection $EM / M$ ( $\Delta G / \text{kJmol}^{-1}$ )	4.5 (3.5)	— —

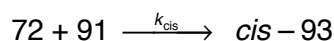
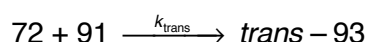
When preformed *para*-imine **97** was added to a fresh set of reagents and the reaction performed and assayed under the same conditions, a contrasting concentration vs. time profile showing a biphasic reaction was observed (Figure 3.14 (ii)). Addition of 20 mol% *para*-imine **97** facilitated a rapid initial rate of formation of *trans*-**93** and simultaneous inhibition of *cis*-**93**. Once the concentration of *trans*-**93** reached approximately 2 mM the rate of formation of *trans*-**93** slows down and the reaction then proceeds at approximately the rate observed in the native reaction which indicates that the reaction is no longer proceeding *via* template directed synthesis but rather through a bimolecular reaction. As 2 mM of *para*-imine **97** was added, it was clear that the imine was capable of strongly accelerating the rate of cycloaddition to form *trans*-**93**. However, as the rate enhancement ceases after 2 mM of cycloadduct was produced it was clear that the product was inhibiting the template as it was unable to turnover and participate in further catalysis.

Further evidence to support this observation was obtained when doping a fresh set of reagents with a higher concentration of *para*-imine **97** (Figure 3.14 (iii)). Addition of 46 mol% *para*-imine **97** at the start of the reaction once again increased the initial rate of reaction to form *trans*-**93** further suppressing the formation of *cis*-**93**. The enhanced reaction rate proceeded until the concentration of *trans*-**93** reaches close to 5 mM, a concentration approximately equal to added *para*-imine **97**. After this point the reaction rate once again proceeded at close to the native reaction rate indicating that no further template directed reaction was proceeding. Once again, the reaction to form *trans*-**93** was strongly accelerated by addition of the template imine.

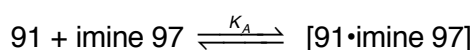
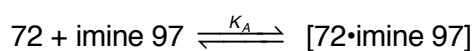
The lack of turnover of the catalyst can be attributed to a very strong product duplex [*para*-imine **97**•*trans*-**93**]. In order to obtain further insight into the behaviour of the system, we turned to kinetic simulation and fitting using SimFit (Table 3.3). Fitting the

experimental results of the native reaction to a bimolecular model revealed the reaction rate constants  $k_{\text{trans}}$  and  $k_{\text{cis}}$  to be  $6.04 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  and  $9.86 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  respectively. These rates were used in the construction of a kinetic model to describe the template directed formation of *trans*-93 on the imine template (Figure 3.15).

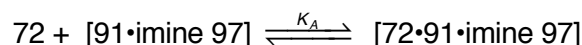
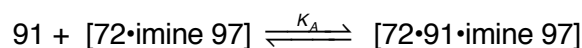
*Bimolecular Cycloadditions Reactions:*



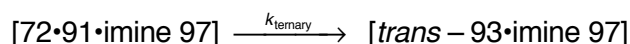
*Binary Complex Formation:*



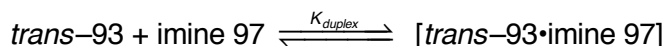
*Constructive Ternary Complex Formation:*



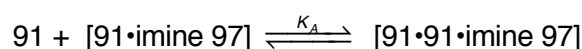
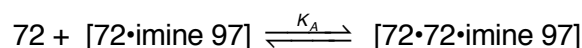
*Reaction Promoted Through Ternary Complex:*



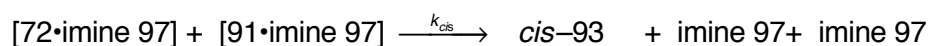
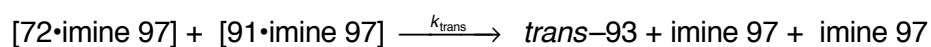
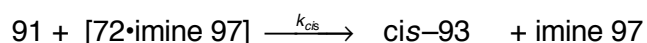
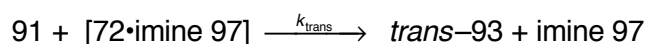
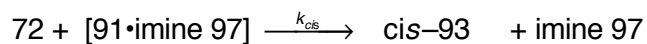
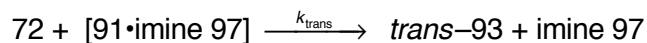
*Association of Ternary Complex:*



*Unconstructive Ternary Complex Formation:*



*Bimolecular Reactions of Binary Complexes:*



**Figure 3.15:** Kinetic model constructed to represent the template directed reaction of nitron 91 and maleimide 72.

The results of the simulation and fitting from this model revealed the rate of reaction to form *trans*-**93** on *para*-imine **97** to be  $28.1 \times 10^{-4} \text{ s}^{-1}$ . This result revealed the *EM* of the reaction on the template to be 4.7 M highlighting the efficiency of the template at promoting the cycloaddition. The results of the simulation and fitting also revealed the binding energy of the [*para*-imine **97**•*trans*-**93**] complex which was calculated to be  $-41.9 \text{ kJmol}^{-1}$ . Considering each recognition event between the acid and amidopyridine recognition sites were  $3470 \text{ mol}^{-1}$  we were able to calculate the connection free energy of binding to be  $3.5 \text{ kJmol}^{-1}$ .

*Free energy of association of a single amidopyridine with benzoic acid:*

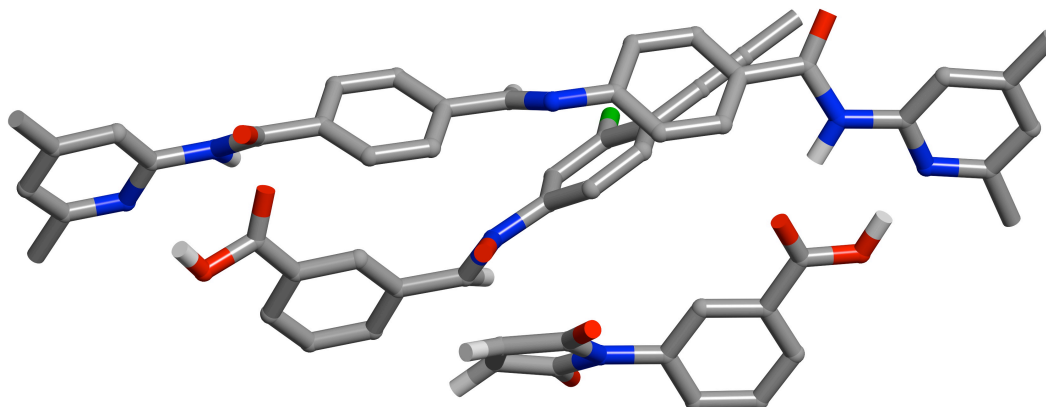
$$\begin{aligned}\Delta G &= -RT \ln K_a \\ &= -8.314 \times 283 \times \ln 3470 \\ &= -19.2 \text{ kJmol}^{-1}\end{aligned}$$

*Calculation of connection free energy:*

$$\begin{aligned}\Delta G_{\text{duplex}} &= \Delta G_A + \Delta G_B - \Delta G_s \\ -41.9 &= (-19.2) + (-19.2) - \Delta G_s \\ \Delta G_s &= 3.5 \text{ kJmol}^{-1}\end{aligned}$$

**Equation 3.3:** Gibbs free energy calculations for (top) the binding energy of a single association and (bottom) the connection free energy.

With a favourable free energy of connectivity, the imine template is able to bind the two acid groups of *trans*-**93** stronger than if it were simply binding the two free reagents. This positive cooperativity highlights the [*para*-imine **97**•*trans*-**93**] template duplex is a strong and stable complex consistent with our experimental evidence for strong product inhibition.



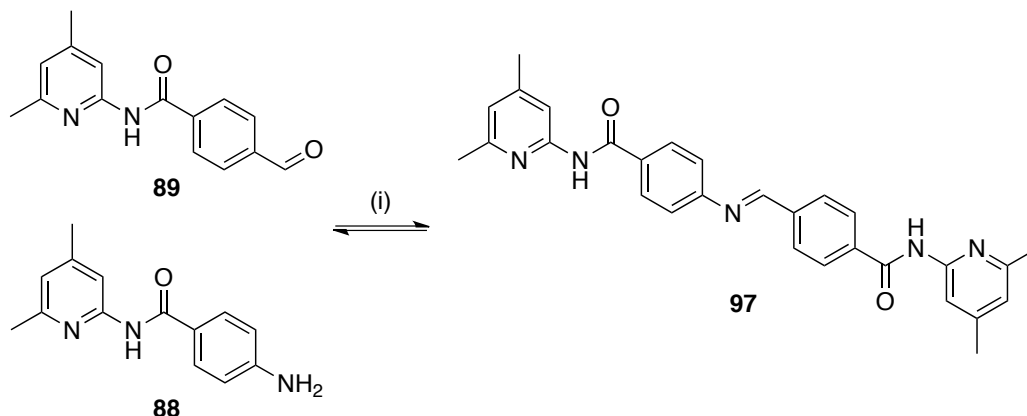
**Figure 3.16:** Molecular model of the transition state of the ternary complex [Nitrone **91**•maleimide **72**•*para*-imine **97**] calculated at the DFT level of theory. Carbon atoms are represented in grey, oxygen atoms in red, nitrogen atoms in blue, fluorine atoms in green and selected hydrogen atoms in light grey. Most hydrogen atoms were omitted for clarity. The alkyl chain on nitrone **91** and maleimide **72** was omitted to reduce computational time whilst retaining the alkyne group to maintain electronic properties.

A molecular model of the transition state within the ternary complex [Nitrone **91**•maleimide **72**•*para*-imine **97**] was constructed to demonstrate that *para*-imine **97** was able to direct the formation of *trans*-**93**. The model (Figure 3.16) clearly shows that the amidopyridine recognition sites on *para*-imine **97** are associated with the acid groups of maleimide **72** and nitrone **91** by all four hydrogen bonds and the reactive sites have been brought into an orientation which bonds formed by the reaction have started to form. The orientation of the transition state also reveals that the protons on the maleimide double bond and the nitrone reactive site are pointed in opposite directions which leads to the formation of the *trans* cycloadduct thus confirming that the *para*-imine **97** is able to support the transition state for the formation of *trans*-**93**.

### 3.5.1 Reciprocal Template Effect

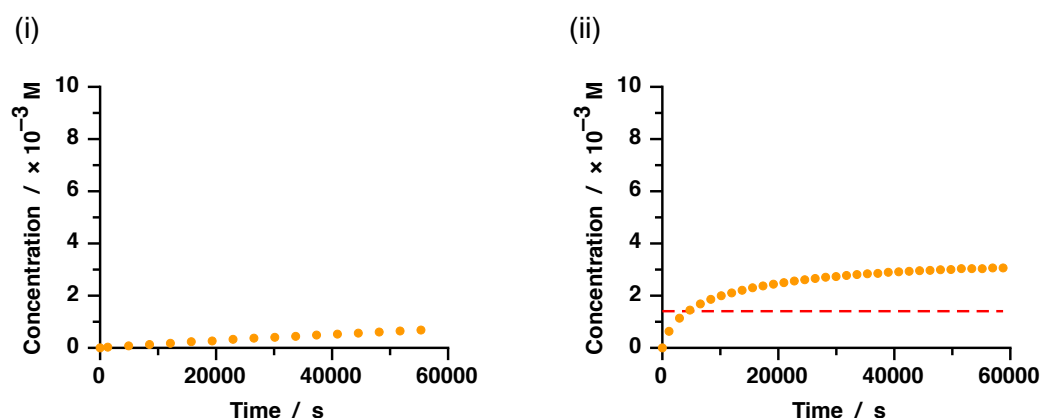
In order to successfully create a fully reciprocal system, we must be able to prove that each product is a catalyst for the formation of their complementary partner. We

have so far proved that *para*-imine **97** is a template for the cycloaddition to selectively form *trans*-**93**. For a fully reciprocal system, we must also be able to show that *trans*-**93** is a template for the formation of *para*-imine **97** from its components.



**Scheme 3.16:** Reagents and conditions (i) 40 mol% benzoic acid or 14 mol% *trans*-**93** (added as a 3:1 *trans* : *cis* mixture), CD<sub>2</sub>Cl<sub>2</sub>, *p*-TSA (sat), collidine, 10 mM, 10 °C, 16 h

The native reaction of amine **88** with aldehyde **89** was performed under the same conditions as the cycloaddition reactions, 10 mM in CD<sub>2</sub>Cl<sub>2</sub>, *p*-TSA (sat), collidine at 10 °C, and was monitored by 500.1 MHz <sup>1</sup>H NMR spectroscopy for 16 h. The native reaction without template was performed in the presence of 40 mol% benzoic acid to simulate the acid concentration of *trans*-**93** and ensure any effect is not as a consequence of simple acid catalysis.



**Figure 3.17:** Concentration vs. time profile for the formation of imine **97** (●) from amine **88** and aldehyde **89** (i) with no added template and (ii) with 14 mol% *trans*-**93**. The concentration of added *trans*-**93** is shown as a dashed red line for reference. Reaction conditions: CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat), Collidine, 2 eq. benzoic acid, 10 mM, 10 °C, 16 h.

In the absence of any added template, the concentration vs. time profile (Figure 3.17 (i)) showed the reaction to form *para*-imine **97** was slow reaching 8% conversion after 16 h. When the reaction was repeated under the same conditions with the

addition of 14 mol% preformed *trans*-**93**, the concentration vs. time profile (Figure 3.17 (ii)) clearly shows an increase in the rate of formation of *para*-imine **97**. The rate of formation of *para*-imine **97** in the presence of template initially proceeds at an increased rate until the concentration of product is approximately equal to the concentration of added template. After the reaction has reached this point, the reaction rate decreases and proceeds at a constant rate close to that of the native rate of reaction with no added template. This observed behaviour is a classic example of a template effect in a dynamic system as the added *trans*-**93** template stabilises the formation of *para*-imine **97**. The ability of *trans*-**93** to template the formation of *para*-imine **97** and *vice versa* demonstrates the reciprocal nature of the reactions within this system.

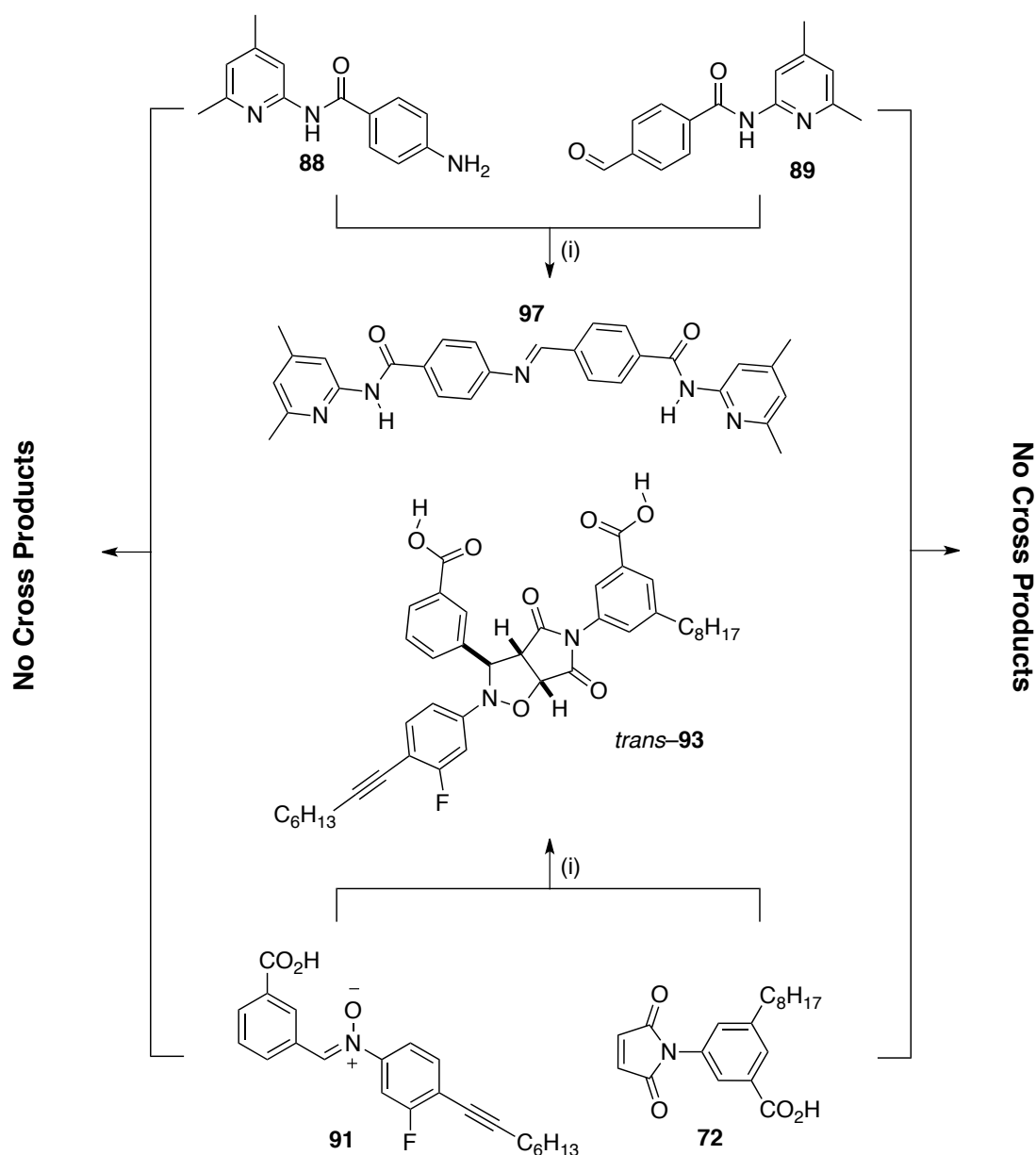
### 3.6 Four Component Reciprocal Replicator

For the system to behave as a reciprocal replicator each template must be able to promote the reaction of its complementary partner *in situ* from a system in which only the four reagents are present at the start of reaction.

This would lead to a sigmoidal concentration vs. time profile for the formation of each of *trans*-**93** and *para*-imine **97** as at the start of the reaction, no template for either reaction would be present and as such only the native bimolecular route is available for the formation of each template. Once some of either template is formed, it will then be able to promote the formation of its reciprocal partner. The duplex should then (ideally) dissociate both templates back into the reaction so that they may continue to catalyse both of the two reciprocal reactions.

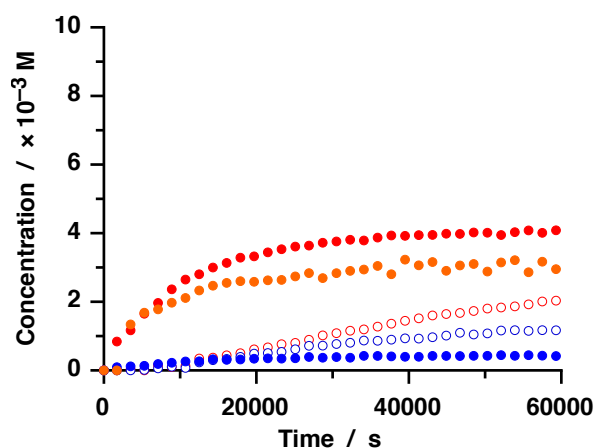
In order to investigate this behaviour, a four component system (Scheme 3.17) was created in which an equimolar concentration of nitrone **91**, maleimide **72**, amine **88** and aldehyde **89** were dissolved together and the reaction performed under identical conditions as before, 10 mM in CD<sub>2</sub>Cl<sub>2</sub>, *p*-TSA (sat), collidine, 10 °C, and the reaction progress monitored by 499.9 MHz <sup>1</sup>H and 376.5 MHz <sup>19</sup>F NMR spectroscopy every 30 minutes for 16 h. Upon completion the relative concentration of each species was determined by deconvolution of the appropriate resonances.





**Scheme 3.17:** Reagents and conditions (i)  $\text{CD}_2\text{Cl}_2$ , *p*-TSA (sat), collidine, 10 mM, 10 °C, 16 h

The concentration vs. time profile of the four component system (Figure 3.18) showed that the two most abundant products of the system after 16 h were the pair of reciprocal templates, *para*-imine **97** and *trans*-**93**. The reciprocal replicating system is behaving as more than the sum of its parts as each of the reciprocal templates are formed more efficiently within the four component system than in their native two component reactions. The concentration of *trans*-**93** in the four component system reaches 4.1 mM after 16 h with a *trans* : *cis* ratio of 9.5:1, an increase when compared to the native system in which only 1.8 mM of *trans*-**93** was

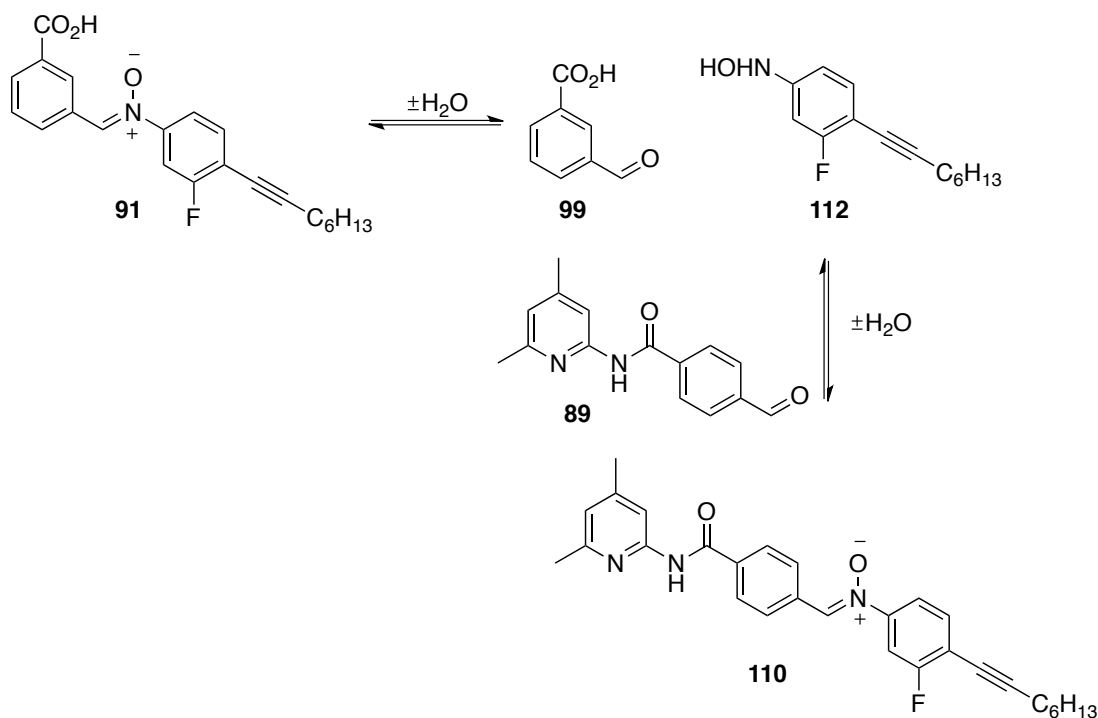


**Figure 3.18:** Concentration vs. time profile of the native four component system displaying the production of *trans*-**93** (●), *cis*-**93** (●), imine **97** (●), and the cycloadducts arising from nitrone **110** *trans*-**111** (○) and *cis*-**111** (○). Reaction conditions: CD<sub>2</sub>Cl<sub>2</sub> / *p*-TSA (sat), Collidine, 10 mM, 10 °C, 16 h.

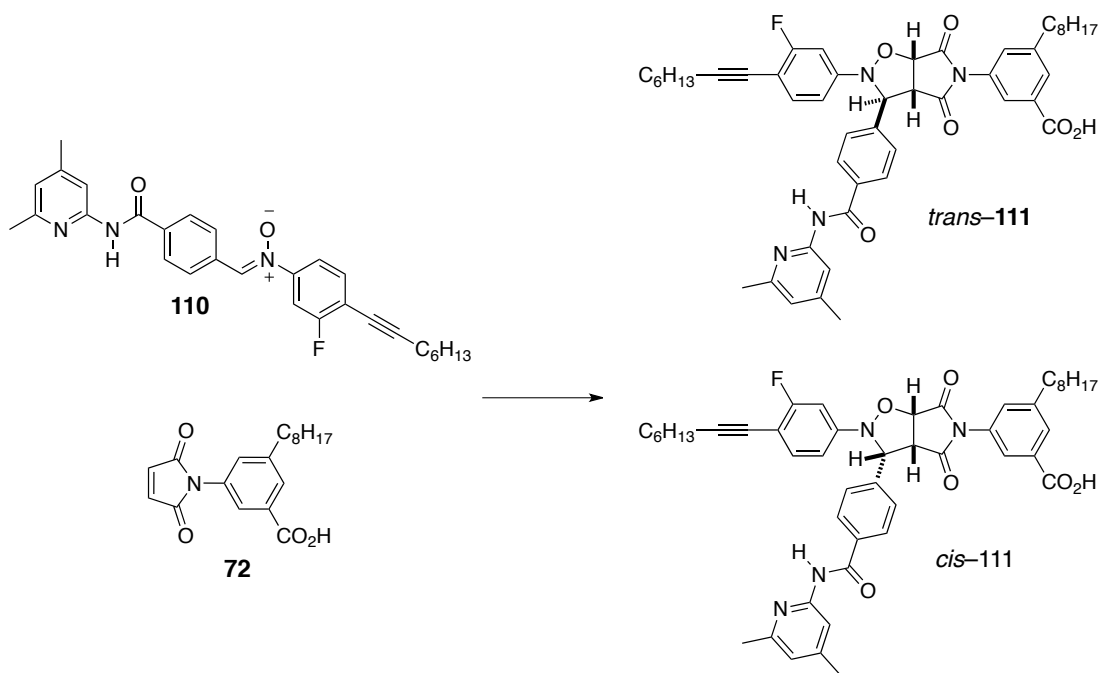
produced after the same time with a reversed *trans* : *cis* ratio of 1:1.7. The formation of *para*-imine **97** was also enhanced reaching 3 mM in comparison to its native reaction<sup>†</sup> which reached 0.8 mM after the same time. At the start of the reaction, the rate of formation of the two reciprocal templates were closely linked with their concentrations equal for the first 10,000 s. After this point, the concentration of each continued to rise at slightly different rates before levelling off at around 30,000 s. The slowing of the formation of the *trans*-**93** was as a result of hydrolysis of nitrone **91** which had been completely consumed by 30,000 s and subsequently a new nitrone was formed *in situ* after recombination of the hydroxylamine **112** (from breakdown of nitrone **91**) with aldehyde **89** present in greater concentration in solution (Scheme 3.18). This recombination nitrone **110** is similar to a previously described nitrone (**52**) which is able to react with maleimide **72** as an autocatalytic minimal replicator, Section 2.7.1, with a rate of reaction in the ternary complex of  $71.6 \times 10^{-4} \text{ s}^{-1}$  and an *EM* of 40 (Scheme 3.19). This autocatalytic reaction rate is 2.5 times greater than the template directed reaction determined for the reciprocal template directed reaction (Table 3.3) and as such the autocatalytic template is able to be produced in a significant quantity in the system.

The reaction to form the autocatalytic template *trans*-**111** and its diastereoisomer *cis*-**111** serve to consume components from the four component system, namely

<sup>†</sup> The native reaction was also performed in the presence of 200 mol% benzoic acid to simulate the acid concentration of the four component reaction. This showed no difference in results to the reaction with 60 mol% added.



**Scheme 3.18:** Nitrone **91** is able to be hydrolysed to hydroxylamine **112** and aldehyde **99** in the presence of water in the system under the conditions of the reaction. Hydroxylamine **112** is then able to recombine with aldehyde **89**, present in the four component system, to produce nitrone **110**.



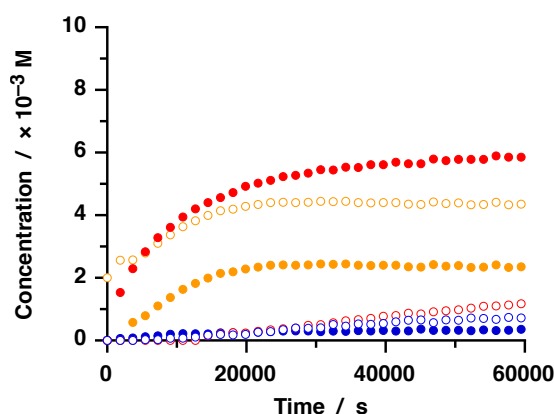
**Scheme 3.19:** Reaction of nitrone **110**, formed in situ within the four component system, with maleimide **72** is able to form two cycloadducts: *trans*-**111** by a minimal self-replication mechanism and *cis*-**111** via an [A·B] binary complex channel.

maleimide **72**, aldehyde **89** and nitrone **91** (as a consequence of hydrolysis) which disrupts the efficiency of the reciprocal replicating system. Although the byproducts

*trans*-111 and *cis*-111 are present within the product pool in the four component system, the reciprocal templates are produced most efficiently and remain the dominant pathway within the system.

### 3.6.1 Template Instruction of Reciprocal System

The importance of reciprocal template effects in the system can be demonstrated by addition of one of the reciprocal templates to a fresh set of reagents. Addition of *para*-imine 97 to the four components will serve as a template to accelerate the formation of *trans*-93 and inhibit *cis*-93 formation from the very start of reaction and should also consume nitron 91 swiftly leaving it less susceptible to hydrolysis and therefore subdue the formation of the autocatalytic byproduct *trans*-111 and its diastereoisomer *cis*-111.



**Figure 3.19:** Concentration vs. time profile for the four component reaction system doped with with 30 mol% *para*-imine 97 showing the evolution of *trans*-93 (●), *cis*-93 (●), newly formed *para*-imine 97 (●) and total *para*-imine 97 (○) as well as the products of the cycloaddition of the recombinant nitron 110 with maleimide 72 to form *trans*-111 (○) and *cis*-111 (○). Reaction conditions: CD<sub>2</sub>Cl<sub>2</sub> / *p*-TSA (sat), Collidine, 10 mM, 10 °C, 16 h.

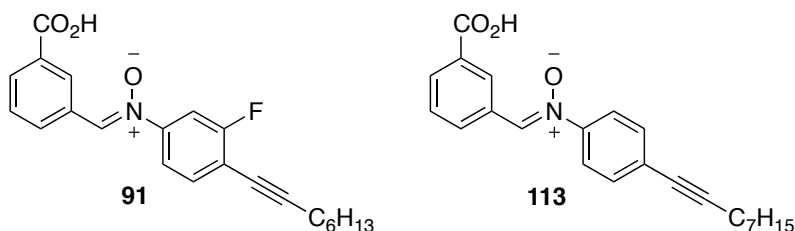
The *para*-imine 97 doped reaction was performed by addition of 20 mol% imine to the four components and performing and assaying the reaction under the same conditions as previously described. The concentration vs. time profile for the imine doped reaction (Figure 3.19) most significantly shows an enhancement in the formation of *trans*-93 which has reached 5.9 mM, an increase from 4.1 mM reached in the native four component system. The increase in selectivity is also reflected in the *trans* : *cis* ratio which was now 17.5 : 1 in comparison to the native reaction which was 9.5 : 1. Addition of reciprocal template also has the effect of reducing the production of the two byproducts *trans*-111 and *cis*-111 as a consequence of the

increased rate of consumption of nitron **91**. Interestingly however, addition of *para*-imine **97** to the four component reaction also serves to inhibit the formation of new *para*-imine **97**, of which 2.35 mM was produced, in comparison to the native four component reaction in which 2.95 mM of *para*-imine **97**. This inhibition is not a physical template effect but instead a consequence of the dynamic nature of imine formation. As *para*-imine **97** is stabilised by *trans*-**93**, the concentration of *trans*-**93** in the system must first satisfy the concentration of pre-added imine before any additional imine is able to be formed from amine and aldehyde. As there is now a lower concentration of cycloaddition reagents (10 mM) in the system compared to imine reagents (10 mM + added *para*-imine **97**), the cycloaddition reagents will become exhausted before the imine reagents. Therefore less of the fresh reagents are consumed to produce new *para*-imine **97**. The total concentration of *para*-imine **97** within the instructed four component system is greater (4.35 mM) than the native system (2.95 mM).

It was clear from these results that there is an inherent problem with the stability of nitron **91** as cycloadducts *trans*-**111** and *cis*-**111** were still present and this problem was now addressed.

### 3.7 Evolution of the Design of the Reciprocal System

The stability of nitron **91** to hydrolysis has proven to be problematic during the investigation thus far. Other research within the group was carried out using a structurally similar nitron which was not decorated with a fluorine atom adjacent to the reactive site which proved to be more stable to hydrolysis (Figure 3.20).



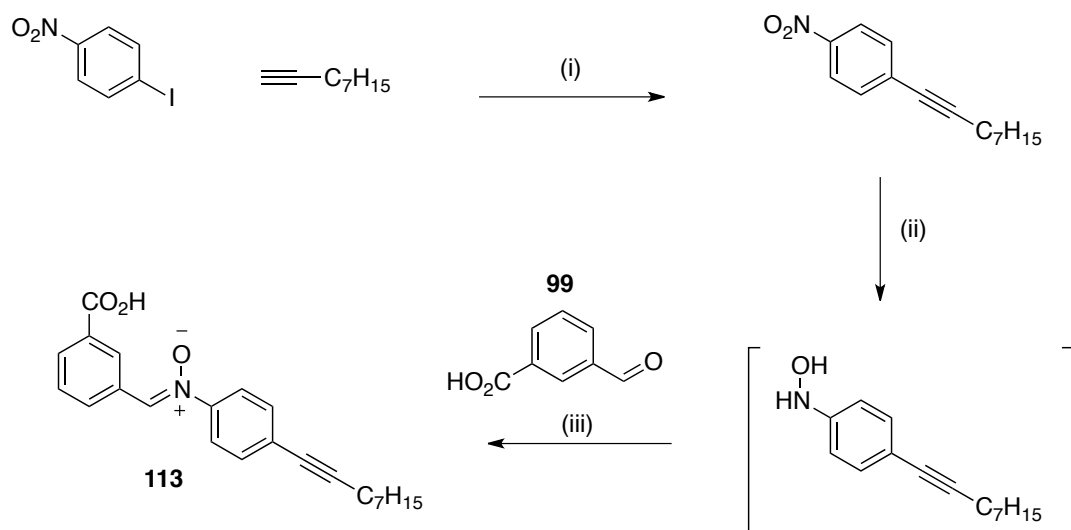
**Figure 3.20:** Fluorine bearing nitron **91** used thus far and non-fluorinated nitron **113** identified by the group.

The increased relative stability of nitron **113** to hydrolysis can be rationalised by considering the effect of the extra fluorine on the nitron. With fluorine present, as in nitron **91**, the electronegative effect of the atom will withdraw electron density from the nitron reactive site, destabilising the positive charge on the nitrogen atom, resulting in an increase in the electrophilicity of the reactive site and increasing the rate of hydrolysis with water. In contrast without fluorine present, as with nitron **113**, the reactive site has a lower electrophilicity and thus less susceptible to hydrolysis.

The fluorine atom adjacent to the nitron reactive site was included simply to aid the monitoring of cycloadditions during the initial screening reactions by  $^{19}\text{F}$  NMR spectroscopy and as the four component reaction is monitored by  $^1\text{H}$  NMR spectroscopy it has become a burden serving only to destabilise nitron **91**. The increased stability of nitron **113** makes it more desirable to use within the four component system and thus it was used to replace the fluorine tagged nitron **91**.

### 3.7.1 Synthesis of Nitron **113**

Nitron **113** was prepared from a similar route to the previously described fluorine decorated nitron **91**.



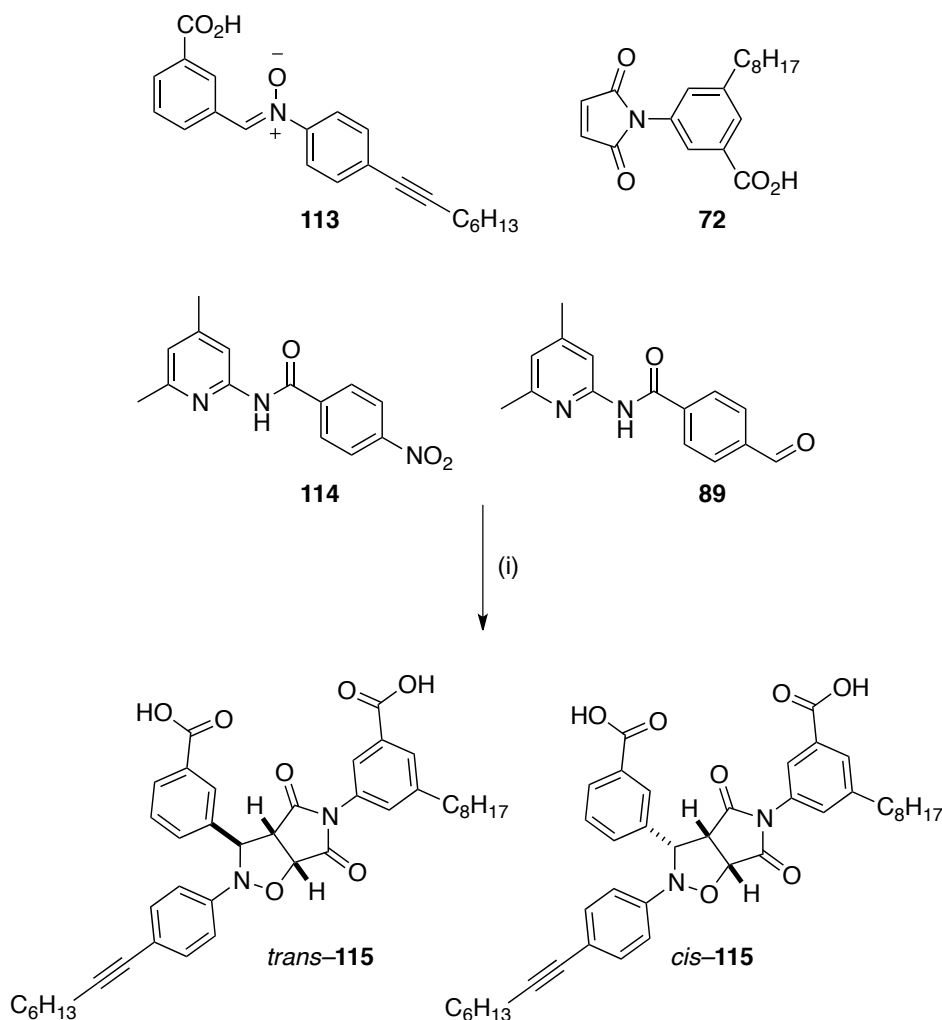
**Scheme 3.20:** Reagents and conditions (i)  $\text{Pd}(\text{P}(\text{Ph})_3)_2\text{Cl}_2$ ,  $\text{CuI}$ ,  $\text{PPh}_3$ ,  $\text{NEt}_3$   $50^\circ\text{C}$ , 16 h, 75% (ii)  $\text{NH}_2\text{NH}_2\cdot\text{H}_2\text{O}$ ,  $\text{Rh/C}$ ,  $\text{THF}$ ,  $\text{RT}$ , 0.5 h, quantitative (iii)  $\text{EtOH}$ ,  $\text{RT}$ , 16 h, 52%.

For the purposes of increasing solubility of the nitron, an alkyl group was once again desired to be included on the structure. A Sonogashira reaction was employed to

successfully add a nonyl group to 4-iodo-nitrobenzene before the nitro group was reduced to the hydroxylamine by transfer hydrogenation using Rh/C and hydrazine. The hydroxylamine was immediately condensed with 3-formylbenzoic acid to yield the target nitron **113** a reasonable yield.

### 3.7.2 Reactivity of Nitron **113** towards Maleimide **72**

The increased stability of nitron **113** in solution meant that the additives included in previous systems could be removed and the reaction able to be carried out in fresh  $\text{CDCl}_3$  which had simply been dried using molecular sieves (the water content was determined to be 9 ppm by Karl-Fischer titration).



**Scheme 3.21:** Reagents and conditions (i)  $\text{CDCl}_3$ ,  $10^\circ\text{C}$ , 10 mM, 40 h.

For an accurate evaluation of the new four component system, a reaction was designed to investigate the native reactivity of the cycloaddition reaction under the conditions within the four component system to ensure simple removal of the fluoride did not have any significant effect on the selectivity of the system. In order to achieve this, the four component reaction, as described in Scheme 3.21, was performed under the conditions to be used for the new four component system, 10 mM in CDCl<sub>3</sub> at 10 °C. In this four component system, only the cycloaddition reaction can take place as the reactive amine **88** has been replaced by the unreactive nitro **114** which is unable to condense with aldehyde **89**. The design of the reaction also means that two equivalents of amidopyridine recognition units are present during the cycloaddition ensuring a more accurate understanding of the cycloaddition reaction within the four component reaction.

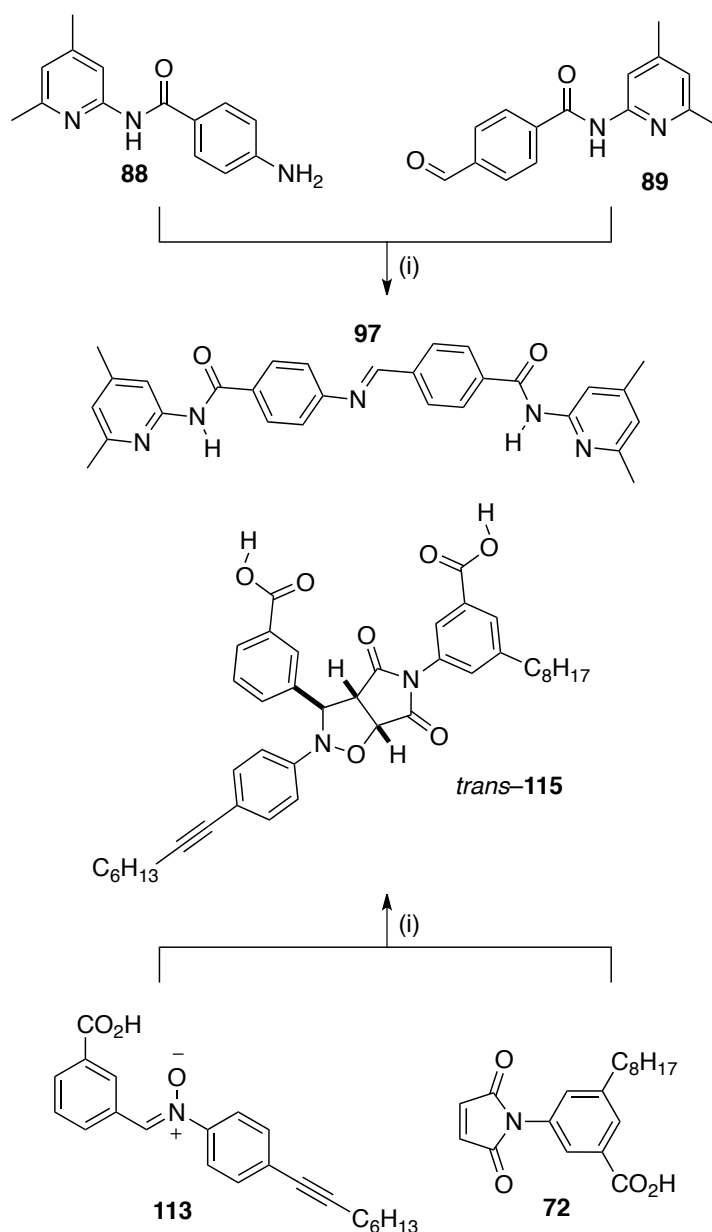
After 16 h, the native reaction produced 1.8 mM of *trans*-**115** and 0.6 mM *cis*-**115** representing a *trans* : *cis* ratio of 3:1 and an overall conversion of 24%. After 40 h, 3.3 mM of *trans*-**115** and 1.1 mM of *cis*-**115** had been produced, maintaining the *trans* : *cis* ratio of 3:1 and now with an overall conversion of 44%. These results highlight that in the four component system, the native cycloaddition reaction is slow and with low *trans* : *cis* selectivity. This result reveals a reversal of *trans* : *cis* selectivity from the previous control (Figure 3.13 (i)) which was 1:1.7. It is likely that in the previous control reaction there could have been recognition mediated behaviour of a single hydrogen bond (as the second was blocked as the methyl ester) facilitating the formation of the *cis* cycloadduct. In this new control reaction the addition of amidopyridine recognition sites to the reaction mixture would disrupt the formation of such a single hydrogen bond complex as stronger hydrogen bonds would be formed with the amidopyridine units.

### 3.8 Reciprocal Replicating System

The four component reaction was repeated with the new nitro **113** (replacing the F-tagged nitro **91**). An equimolar reaction of nitro **113**, maleimide **72**, aldehyde **89** and amine **88** (Scheme 3.20) at 10 mM in CDCl<sub>3</sub> was incubated at 10°C and assayed every 30 minutes for 40 h by 500.1 MHz <sup>1</sup>H NMR spectroscopy. Upon completion, the relative concentration of each species was determined by

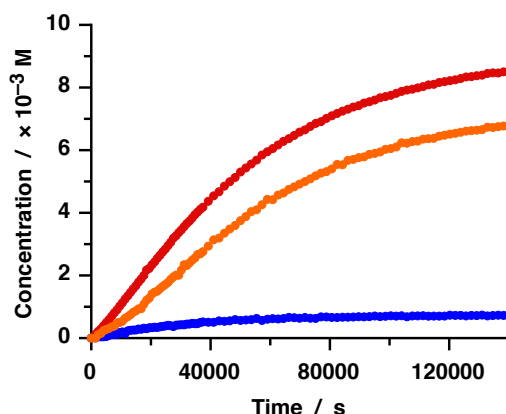


deconvolution of the appropriate resonances and a concentration vs. time profile created (Figure 3.21).



**Scheme 3.22:** Reagents and conditions (i)  $CDCl_3$ ,  $10^\circ C$ , 10 mM, 40 h.

The concentration vs. time profile for the four component system clearly shows that once again the two dominant products are the reciprocal templates, *trans*-**115** and *para*-imine **97**. The expected diastereomeric cycloadduct *cis*-**115** was present in the system however the production of the two cycloaddition by-products as seen in the previous system was strongly inhibited with the autocatalytic *trans* cycloadduct barely detectable after 40 h with less than 2% conversion (not shown on graph). The



**Figure 3.21:** Concentration vs. time profile for the four component system. *Trans*-115 (●), *cis*-115 (●), *para*-imine 97 (●). Reaction conditions: CDCl<sub>3</sub>, 10 mM, 10 °C, 40 h.

profile of the formation of both of the reciprocal templates clearly shows a sigmoidal profile consistent with a reciprocal replicator mechanism. The concentration of *trans*-115 reaches 8.6 mM after 40 h whilst its diastereoisomeric partner reaches 0.7 mM representing a *trans* : *cis* ratio of 12:1 and an overall conversion of 93%.

The production of the reciprocal *para*-imine 97 template was once again closely linked to the concentration of *trans*-115. The concentration of *para*-imine 97 reached 6.9 mM after 40 h however throughout the course of the reaction, the concentration of *para*-imine 97 was consistently lower than that of *trans*-115. This is likely a result of the dynamic nature of imine formation.

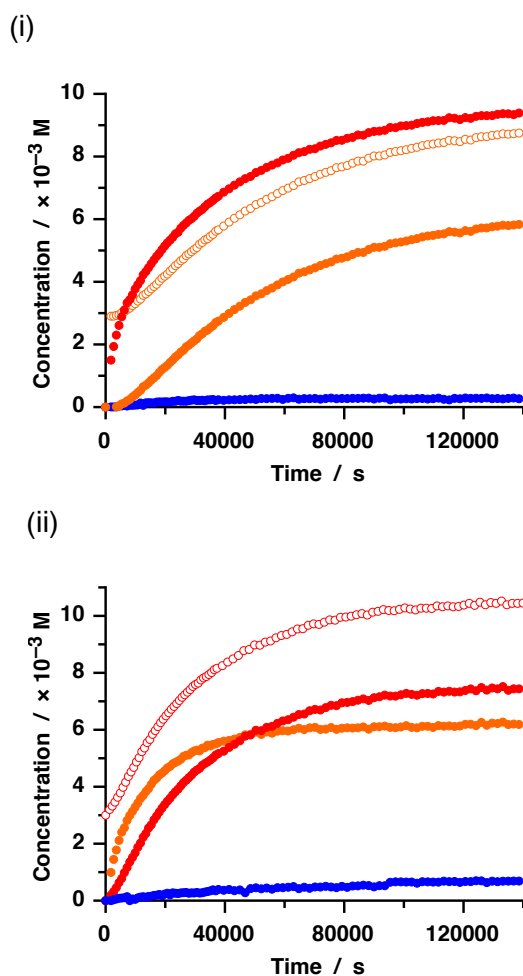
In isolation, imine condensation will, after time, reach an equilibrium point where the rate of forward and reverse reactions are equal. In order to find this point, a reaction containing amine 88 and aldehyde 89 was incubated at 10 °C in the presence of 200 mol% benzoic acid to simulate the acid concentration of the four component system and was assayed by 400.1 MHz <sup>1</sup>H NMR spectroscopy on a regular basis. After 14 days, the condensation had halted after reaching 20% conversion to imine 97.

In the four component system, this equilibrium is disturbed as *trans*-115 will act as a template to stabilise the formation of *para*-imine 97. As the concentration of *trans*-115 increases with time, it will eventually pass the equilibrium concentration of *para*-imine 97. As the concentration of *trans*-115 continues to increase it will continue to stabilise the formation of *para*-imine 97 at concentrations higher than its equilibrium point. Above this point, any free imine not stabilised by *trans*-115 will revert to its

starting materials as the system tries to recover to its original equilibrium point. This ensures that the concentration of *para*-imine **97** will never be able to surpass that of *trans*-**115** in the system.

### 3.8.1 Template Instruction of Reciprocal Replicating System

The importance of reciprocal template effects in the system was once again demonstrated by addition of the reciprocal templates individually to a fresh set of reagents.



**Figure 3.22:** Concentration vs. time profile of (i) four component system with 30 mol% *para*-imine **97** doped (ii) four component system with 30 mol% *trans*-**115** doped. Newly formed *trans*-**115** (●), newly formed *cis*-**115** (●), newly formed *para*-imine **97** (●), total *trans*-**115** (○) and total *para*-imine **97** (○). As preformed imine is added in (i) and preformed *trans*-**115** in (ii) there will be an initial concentration of these compounds in the mixtures as seen by the offset in concentration at t=0. Reaction conditions: 10 mM, CDCl<sub>3</sub>, 10 °C, 40 h

The four component system was first doped with preformed *para*-imine **97**. This was achieved by addition of 30 mol% *para*-imine **97** to an equimolar system of the four

reagents under the same conditions as before and the reaction performed and assayed in an identical manner to construct a concentration vs. time profile of each of the products (Figure 3.22 (i)). The concentration vs. time profile for the imine doped reaction clearly showed a rapid increase of initial rate of formation of *trans*-**115**. Formation of new imine was once again dependant on the concentration of *trans*-**115** as it can be seen that the production of new imine **97** displays a lag period and does not begin to increase until 30 mol% of *trans*-**115** has been produced, thus satisfying the pre-added template. Once again the formation of new imine **97** is inhibited as the total concentration of imine is dependant on the concentration of *trans*-**115** which is able to stabilise the dynamic imine formation.

Addition of preformed *trans*-**115** to the four component system was also performed by addition of 30 mol% *trans*-**115** to a fresh set of reagents and the reaction performed under the same conditions and assayed as before to produce a concentration vs. time profile (Figure 3.22 (ii)). The results showed a rapid initial rate of formation of the reciprocal imine template before the reaction rate slowed down. The formation of *trans*-**115** displayed a sigmoidal profile consistent with the expected behaviour that firstly the pre-added *trans* template must be satisfied before formation of new *trans* cycloadduct is formed.

### 3.9 Environmental Dependence of Replication Mechanism

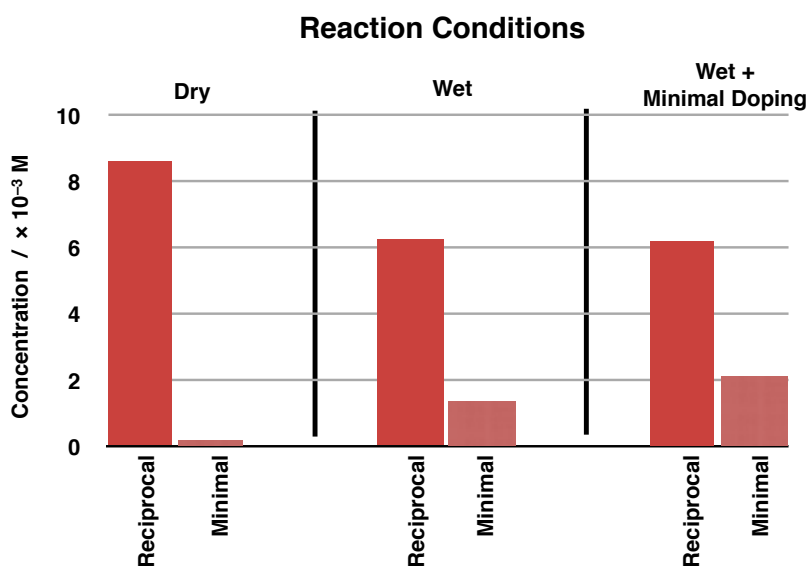
The susceptibility of the nitron used in the four component system to hydrolysis and its subsequent recombination into an alternative nitron, capable of autocatalytic cycloaddition with maleimide **72**, gives rise to an interesting scenario in which the environment in which the four component reaction is performed will promote either reciprocal or minimal replication.

In the examples reported thus far, we have attempted to create a system in which only reciprocal replication was present and subdue the minimal replication pathway arising from the breakdown of the acid bearing nitron and subsequent recombination to form an amidopyridine bearing nitron. We successfully demonstrated that by removing the fluoride group, we could produce a nitron (**113**) with increased stability under the conditions of the reaction and thus avoid hydrolysis.

If the rate of hydrolysis of the nitron could be enhanced in the four component reaction, we could expect to see a greater formation of the amidopyridine bearing nitron and subsequent template formed from autocatalytic minimal replication.

In an attempt to create conditions which would promote hydrolysis, we decided to saturate  $\text{CDCl}_3$  with water before undertaking the reaction.  $\text{CDCl}_3$  was stirred for 20 mins with  $\text{D}_2\text{O}$  before allowing the two layers to settle and separating the  $\text{CDCl}_3$  from  $\text{D}_2\text{O}$ . The water content in the  $\text{CDCl}_3$  was determined by Karl–Fisher titration to be >1500 ppm and this was used as the solvent to run the four component reaction under ‘wet’ conditions.

The non fluorinated four component reaction was then performed, as described above in Section 3.8, in ‘wet’  $\text{CDCl}_3$  and incubated at 10 °C for 40 h before assaying with 400.1 MHz  $^1\text{H}$  NMR spectroscopy.



**Figure 3.23:** Concentrations of *trans* cycloadducts from the reciprocal replication cycle (*trans*–115) and minimal replication cycle (*trans*–116) from the four component system under; (left) ‘dry’ reaction conditions (middle) under the ‘wet’ conditions and (right) wet conditions plus doping with 20 mol% minimal template (*trans*–116).

Under the new ‘wet’ conditions, after 40 h the ratio of reciprocal template, *trans*–115, to minimal template, *trans*–116, had dramatically changed (Figure 3.23). In the dry conditions, 8.6 mM *trans*–115 was produced in comparison to just 0.19 mM *trans*–116 representing a reciprocal:minimal replication template ratio of 49:1. Under the new ‘wet’ conditions, the concentration of *trans*–115 produced had decreased to 6.2 mM whilst the concentration of *trans*–116 had increased to 1.3 mM representing a reciprocal:minimal replication template ratio of 5:1. These results are explained as a

consequence of the increased rate of hydrolysis producing a greater concentration of amidopyridine bearing nitron which means more will be available for autocatalytic cycloaddition.

The minimal pathway can be promoted further by repeating the wet reaction under the same conditions but now in the presence of 20 mol% of the minimal template *trans*-**116**. Addition of the minimal template will serve to increase the rate of formation of product through the minimal replication pathway. The results of this reaction showed (Figure 3.23 (right)), a further increase in the production of the minimal template. The concentration of *trans*-**115** was similar at 6.2 mM whilst the concentration of *trans*-**116** had now increased 2.1 mM representing a reciprocal : minimal replication template ratio of 3:1.

Whilst we were able to increase the formation of minimal replicator template by changing the environment of the four component reaction, the relatively strong efficiency of the reciprocal replicator proved to be greater than the minimal replicator. One factor which contributed to the relative stability was the use of the non-fluorinated nitron which we have found to be more resistant to hydrolysis. In order to attempt to increase the prevalence of the minimal replication pathway, the four component reaction should be repeated with the fluorinated nitron **91**. This nitron was found to be more susceptible to hydrolysis and therefore under wet conditions, the rapid hydrolysis should allow the formation of amidopyridine bearing nitron **110** more rapidly and thus we should be able to further encourage minimal replication.

### 3.10 Conclusions

The formation of a synthetic four component reciprocal replicating system has long been plagued by the misfortunate combination of complementary reagents to form either reactive [A•B] complexes such as in the system described by Rebek *et al* or efficient minimal replicators as in the system described by Philp *et al*. One way to inhibit the formation of these reactive complexes is to orthogonalise the reactive sites so that even if the complementary complexes do form they will not be able to react as they do not have complementary reactive sites. This approach has been successfully applied here to this system of four components incorporating a 1,3-dipolar cycloaddition as one reaction and imine condensation as the second.

We have successfully identified a pair of reciprocal templates which are formed by orthogonal reactions and demonstrated that each of the templates was a successful catalyst for the formation of its reciprocal partner. When the four components for the two reciprocal templates were combined into one system, the results of the first system incorporating a fluorine tagged nitron showed that the reciprocal templates were the major products formed from the reaction and the significance of each template was demonstrated by doping the four component reaction. At this stage the efficiency of the system was reduced by the instability of the nitron which through hydrolysis into its aldehyde and hydroxylamine and subsequent recombination with the added aldehyde lead to the production of an autocatalytic minimal replicator which consumed reagents required for the reciprocal replicator and lead to an overall reduction in efficiency.

In order to overcome the stability problem in the pursuit of a more efficient reciprocal replicator, the fluorine tag was removed to produce a nitron with greatly increased stability to hydrolysis. Utilisation of this new nitron in the four component reaction lead to the creation of a highly efficient reciprocal replicating system. These results demonstrate the first successful implementation of a reciprocal replicating system created from small, synthetic organic molecules and vindicate the design method of employing orthogonal reactive sites to achieve a system which suffers less from by-products.

The behaviour of the system in the presence of water was very interesting. When the reagents were subjected to 'wet' conditions, the breakdown of the nitron and subsequent recombination with a second aldehyde in solution lead to the formation of a nitron which facilitated a second mechanism of replication, minimal replication. Whist the minimal replicator was not the major product of the system, the its prevalence could be enhanced further by doping with the minimal template. These results give a demonstration of a system in which there is environmental dependence of the mechanism of replication.

# 4

## Rational Design of a Bifunctional Replicator

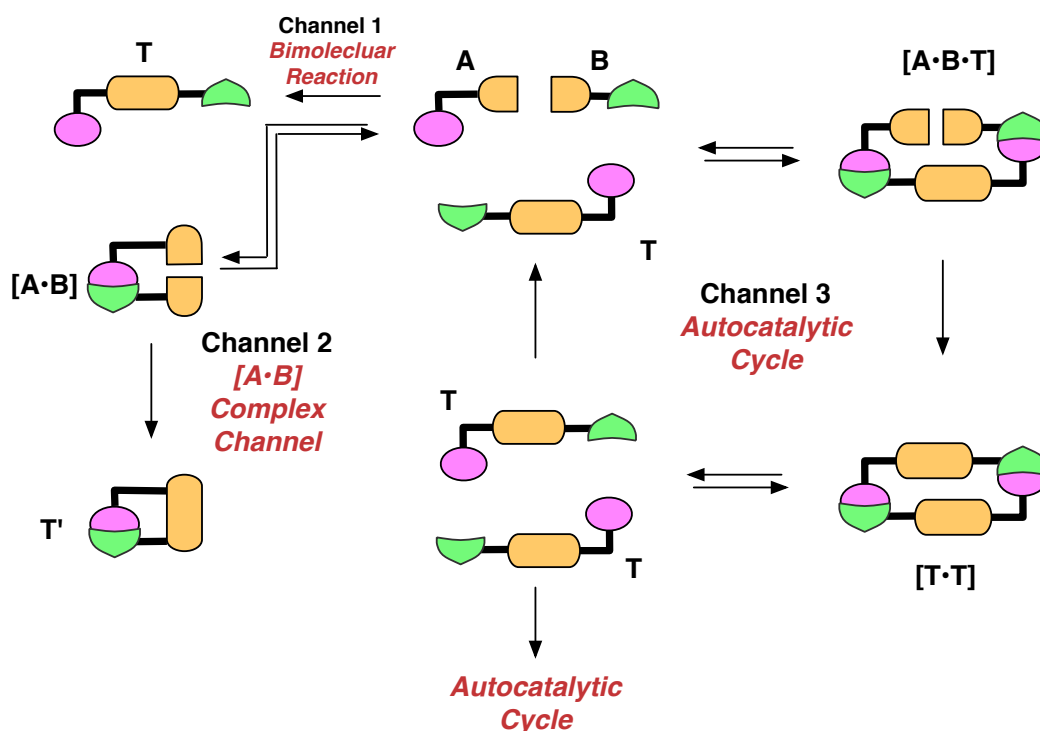
### 4.1 Introduction

Thus far, the work presented in this thesis has centred on the design, characterisation and application of templates which serve a single purpose, namely catalysing the formation of themselves (self-replication) or a complementary partner (reciprocal replication). In order to implement replication strategies into other areas of supramolecular chemistry, these molecules must be able to perform more than a single purpose. Therefore the latter half of this thesis will focus on the investigation of bifunctional templates, self-replicating templates that are able to perform an additional behaviour.

For the synthetic chemist, self-replication may be viewed as a powerful synthetic tool as the product selectively catalyses its own formation at the expense of other side reactions leading to greater yields and purity. The minimal model of self-replication (Figure 4.1) demonstrates that initiation of the autocatalytic cycle in a self-replicating reaction occurs when preformed template **T** associates its building blocks, **A** and **B**, which organises them into a reactive ternary complex [**A**•**B**•**T**]. In order for **T** to successfully associate **A** and **B**, the recognition sites of **T** need to be open to solution. In the minimal model, preformed open template **T** has to be formed by a bimolecular reaction, Channel 1, before the autocatalytic cycle can proceed. This bimolecular reaction is not directed by recognition processes but instead driven by random collisions which makes the reaction slow and unselective. Only when enough of **T** is formed in this manner can the autocatalytic cycle begin to produce more **T** through the more efficient autocatalytic cycle in Channel 3. Therefore it is during the early stage of a self-replicating reaction that the rate is slowest and the system is error-prone which leads to unwanted by products and lower selectivity.

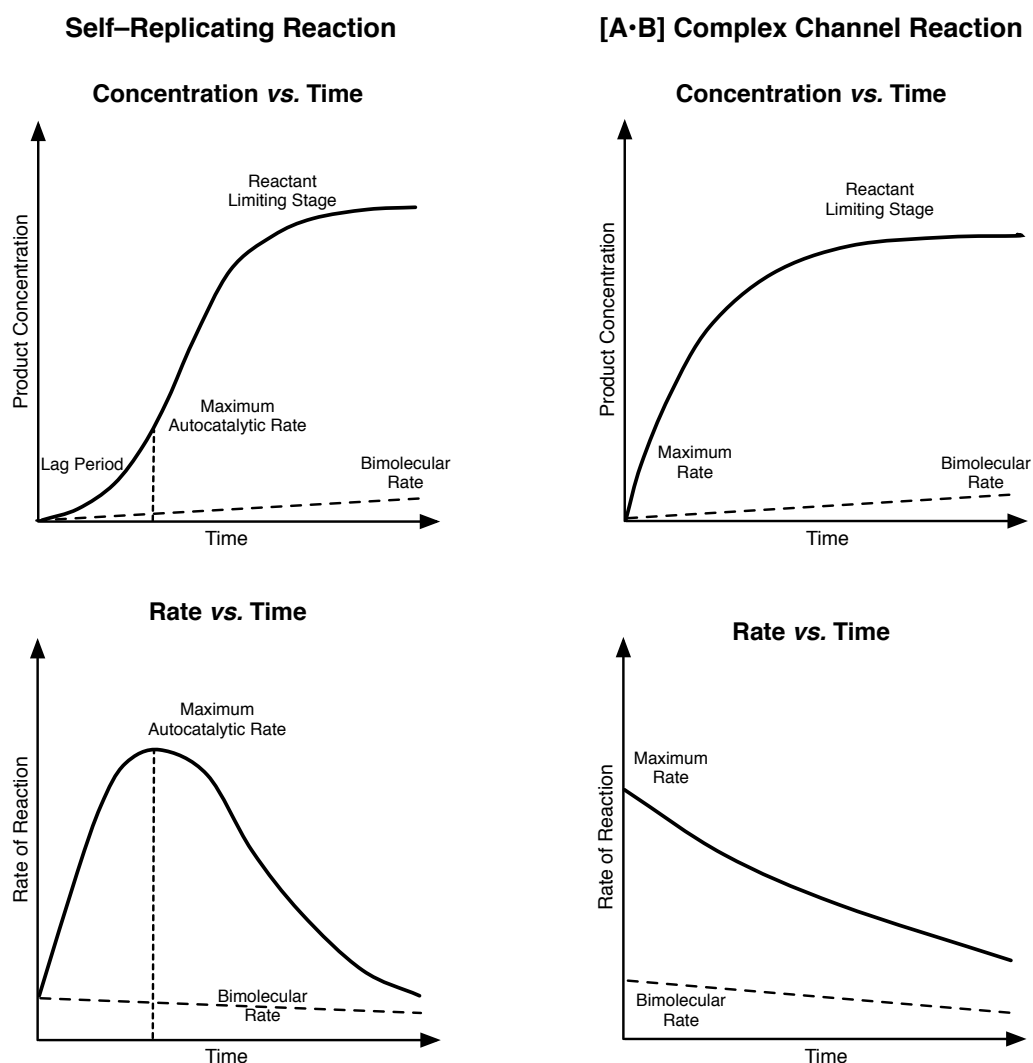


The minimal model also describes a further recognition mediated pathway, the  $[A \cdot B]$  complex channel, Channel 2. In contrast to the kinetic behaviour of the autocatalytic cycle channel, the  $[A \cdot B]$  complex channel is most efficient in the initial stages of reaction. The maximum rate of reaction through this channel occurs when the concentration of  $[A \cdot B]$  is highest. As the concentration of  $[A \cdot B]$  is dependent on the association constant of the recognition sites ( $K_a$ ), which is constant, and the concentration of **A** and **B**, the reaction will progress at the maximum rate when the concentration of **A** and **B** is highest, *i.e.* upon initiation of the reaction. The resultant *pseudo*-intramolecular reaction will then produce template **T'** at an increased rate compared to the bimolecular reaction. The drawback of reaction through Channel 2 is that the template formed, **T'**, is catalytically inert as the recognition sites used to assemble the product remain associated together which inhibits the template from taking part in any further recognition mediated processes.



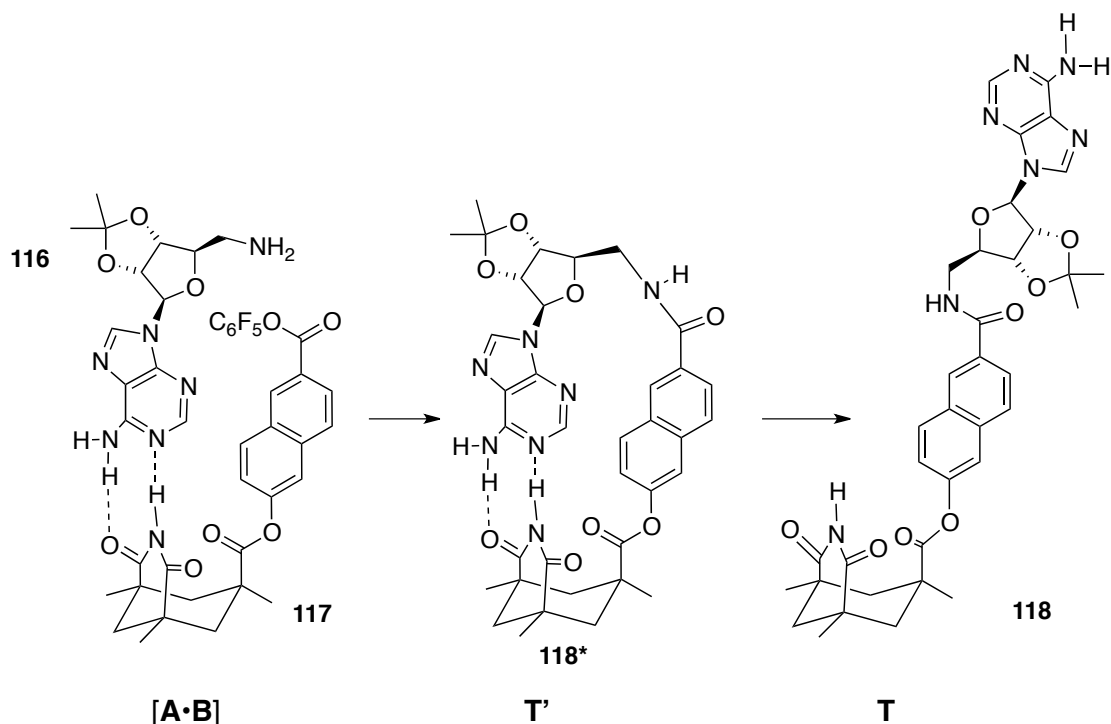
**Figure 4.1:** Minimal model of replication. Reactive sites are coloured orange with complementary recognition sites green and purple.

Ultimately, we would like to design a system in which there is a fast initial rate of reaction which is selective for only one product, which could be achieved by exploiting an  $[A \cdot B]$  complex pathway, however it would be desirable that this product is still catalytically active in other recognition mediated reactions such as producing itself (autocatalytic self-replicating) or other templates (cross-catalysis).



**Figure 4.2:** Top: Concentration vs. time profiles and bottom for (i) an autocatalytic self-replicating reaction showing a characteristic sigmoidal shape and (ii) an  $[A \cdot B]$  complex channel reaction showing the maximum rate to occur upon initiation of the reaction.

In 1990, Rebek *et al* reported<sup>114</sup> the design of a synthetic self-replicating template (Scheme 4.1) **118** which was produced by amide bond formation of amine **116** with activated ester **117** as discussed in Chapter 1, Section 1.5.5. For this reaction, the concentration vs. time profile did not display a sigmoidal profile however autocatalysis was demonstrated by doping preformed template into a fresh mixture of reagents which had the effect of increasing the initial rate of reaction therefore demonstrating that the template was directing the reaction. The absence of a lag period suggested that self-replication through a ternary complex was not the only reaction pathway present.



**Scheme 4.1:** Association of reagents into an **[A•B]** complex promotes the amide coupling reaction to form closed template **T'** with *cis*-amide. Isomerisation of the amide bond from *cis* to *trans* leads to the formation of the open template **T**.

The authors went on to explain that an alternative reaction mechanism though an **[A•B]** binary channel was present. In this pathway, upon association of the reagents amine **116** and activated ester **117** into an **[A•B]** binary complex **[116•117]** (Scheme 4.1) the reactive sites were brought into close proximity which rendered the reaction *pseudo*-intramolecular and thus proceeded at an increased rate relative to the bimolecular reaction leading to the formation of a closed template **T'** which was the *cis*-amide **118\***. This template was then able to undergo isomerisation of the amide bond from *cis* to the more stable *trans* isomer which furnished the open template **T**, *trans*-amide **118**. Using kinetic simulation and fitting, the authors were able to determine that the **[A•B]** binary complex reaction proceeded with an *EM* of 0.16 M whilst the ternary complex reaction had an *EM* of 3.0 M. Although the *EM* of the ternary complex reaction is appreciably greater than the **[A•B]** binary complex reaction a significant fraction of the product *trans*-amide **118** was determined to be formed through the **[A•B]** binary pathway. In order to account for this discrepancy, the authors considered the states in which the population of molecules will exist under the reaction conditions. At a reaction concentration of 10 mM in the presence of 3 mM of preformed template, the authors determined that 67% of the reagents will exist in their free states, 27% will exist in an **[A•B]** binary complex whilst just 1.5% will exist

in a ternary complex. The relative contributions were estimated to be 1, 4.3 and 4.6 for the background bimolecular reaction,  $[A \cdot B]$  binary complex reaction and ternary complex reaction respectively. Thus with a greater population of material in an  $[A \cdot B]$  binary complex, this pathway will significantly contribute to the total product conversion and as such obscure the appearance of a sigmoidal concentration vs. time profile.

Whilst the ability of the undesired  $[A \cdot B]$  complex product  $T'$  (**118\***) to isomerise into the desired open template  $T$  (**118**) was unplanned, Rebek's system showed that it is possible for a product formed *via* the  $[A \cdot B]$  complex channel to open its recognition sites to solution to promote further catalysis.

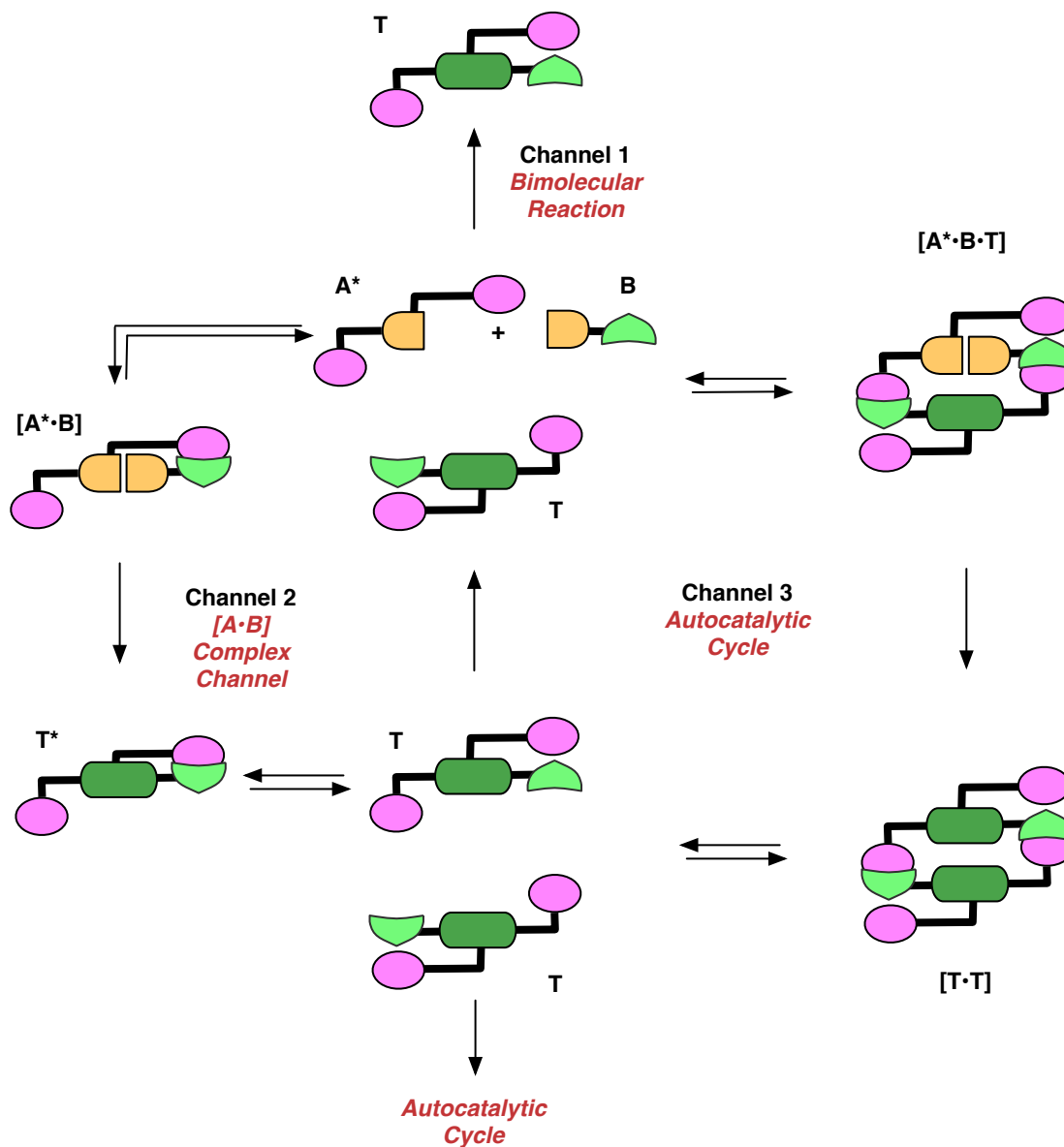
#### **4.2 Combining $[A \cdot B]$ Complex Channel and Autocatalysis. A Fresh Perspective**

The observed behaviour of the Rebek system lay fortuitously in the ability of the molecule to isomerise from a *cis* to a *trans* amide which was capable of further catalysis. Ultimately, the long term goal of our research is the integration of products from recognition mediated processes into higher ordered supramolecular superstructures. In order to achieve this auspicious goal, it is critical to have a fundamental understanding of the structural effects which control behaviour which arises from recognition mediated process and in particular how to design for a particular desired behaviour.

Our attention therefore turned to the design of a system which would deliberately combine the advantageous behaviour of the  $[A \cdot B]$  complex channel and autocatalytic self-replication, namely a fast initial rate of reaction with control over product formation and autocatalytic behaviour.

In order to achieve such behaviour, the redesign of one of the original reagents into a bifunctional substrate was anticipated (Figure 4.3). The bifunctional substrate,  $A^*$  would be decorated with two recognition groups each complementary to the recognition group on  $B$ . One recognition site would be designed to allow for  $[A \cdot B]$  complex channel behaviour, facilitating the formation of a reactive binary complex  $[A^* \cdot B]$  which would lead to the formation of  $T^*$  at a rate greater than the simple bimolecular reaction. The second recognition site on substrate  $A^*$  would be designed

in such a way as to facilitate autocatalytic channel behaviour as seen with previous minimal replicators, facilitating the formation of a ternary  $[A^* \cdot B \cdot T]$  complex.



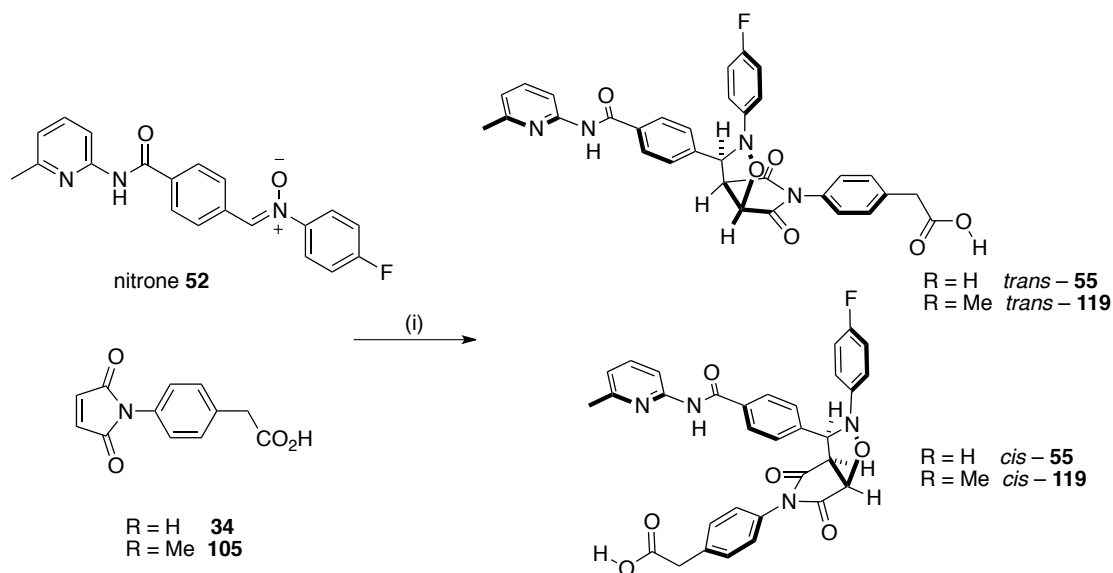
**Figure 4.3:** Representation of desired model of replication. Three reaction channels exist as with the minimal model except there is now a direct connection between channel two and channel three as the product of the  $[A \cdot B]$  complex channel is able to open its recognition sites to solution and act as a template for its own formation *via* the creation of a ternary complex. Reactive sites are coloured orange with complementary recognition sites green and purple.

The anticipated behaviour of the system would be that upon initialisation of the reaction **A\*** and **B** would associate to form  $[A^* \cdot B]$  and react to form **T\*** at an increased rate relative to the bimolecular reaction. The recognition sites associated in **T\*** should then be able to dissociate to form **T**. With all recognition sites now open to solution, **T** should then be able to act as a template for the formation of more **T**

through an autocatalytic channel. The effect of coupling the  $[A \cdot B]$  complex channel to the autocatalytic channel has been to create a recognition mediated pathway which forms the autocatalytic template **T** faster than the bimolecular rate which would otherwise be required and thereby satisfying our desired behaviour of a system which exploits both an  $[A \cdot B]$  complex channel, for fast initial rate of reaction, and an autocatalytic channel for further catalysis.

### 4.3 Design of the Bifunctional System

Previous work in our laboratory has identified<sup>139</sup> the self-replicating system in which nitron **52** and maleimide **34** react together to selectively produce *trans*-**55** via autocatalytic self-replication and *cis*-**55** as a non-catalytic byproduct.



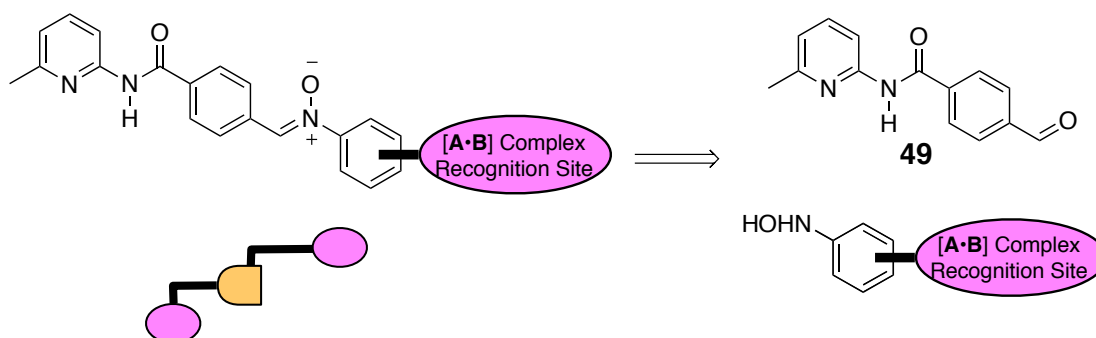
**Scheme 4.2:** Previously reported autocatalytic self-replicating system. Reaction of nitron **52** with maleimide **34** leads to the formation of two cycloadducts, *trans*-**55** which is an autocatalytic self-replicating template and *cis*-**55** a byproduct. Reagents and conditions (i)  $CDCl_3$ , various temperatures.

The concentration vs. time profile of this reaction at low temperature displays a sigmoidal shape, characteristic of an autocatalytic reaction, for the formation of *trans*-**55**. In this system, association of the building blocks into an  $[A \cdot B]$  complex does not orientate the reactive sites into a conformation which enhances the rate of reaction for the formation of either product. This system was chosen as a starting point to which addition of a second recognition facilitating  $[A \cdot B]$  complex channel behaviour site could be investigated.

#### 4.4 Designing an [A•B] Complex Pathway

The first step when designing a system with a fast initial rate of reaction is to design substrates which could associate to form an [A•B] complex which promotes reaction between the substrates. This will increase the rate of reaction relative to the bimolecular rate and also control stereochemistry of the product as the recognition sites hold the reactive sites in a fixed conformation for reaction. With a successful replicating template *trans*-55 already designed, any change to the structure of the existing recognition sites would very likely reduce the efficiency of the template to behave as a self-replicator. The focus of the design to add a second recognition site was to identify an area to add the second recognition site which would not alter the geometry of the existing recognition sites on either **A** or **B**.

With this in mind a site on the nitron was identified which would be synthetically accessible to change without disturbing any existing recognition sites (Scheme 4.3).

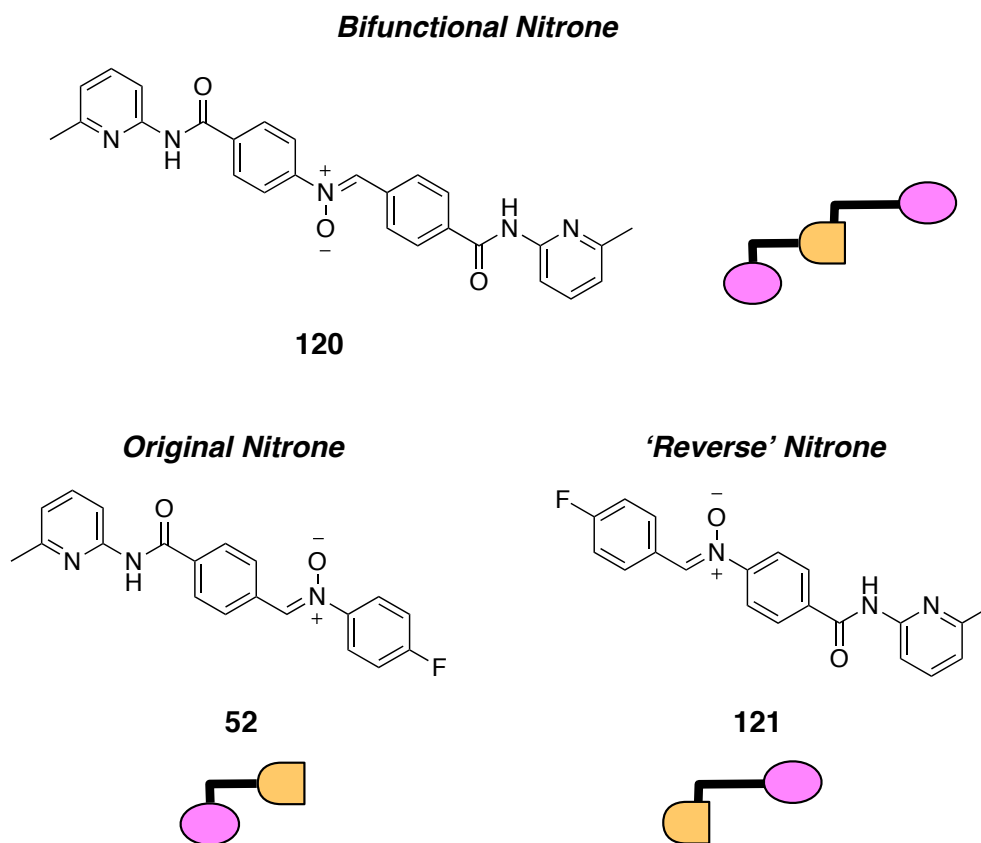


**Scheme 4.3:** Design of the bifunctional nitron with the existing recognition site known to facilitate autocatalytic behaviour with the potential location of a recognition site which facilitates [A•B] complex channel behaviour. Retrosynthesis of the nitron shows that it can be produced from the existing recognition bearing aldehyde **49** with an appropriate recognition bearing hydroxylamine.

Locating the second recognition site on the currently unfunctionalised phenyl ring adjacent to the nitron reactive site would add the second recognition site far from the original recognition site and so not cause any steric problems by adding another group. Retrosynthetic analysis of the new nitron bearing two recognition sites showed that the nitron could be synthesised by condensation of the appropriate recognition bearing hydroxylamine with existing recognition bearing aldehyde **49**, (Scheme 4.3).

The new nitron thus carries two recognition sites, one on either side of the nitron reactive site. On the carbon side of the reactive site lays the original recognition site to allow behaviour as a self-replicator and the nitrogen side of the reactive site lays a recognition site designed to promote [A•B] complex formation.

As the second recognition site on the nitron was required to be complementary to the phenyl acetic acid group on the maleimide, the amidopyridine recognition was chosen for the second recognition site. Addition of the amidopyridine recognition site *para* to the nitron reactive site (Figure 4.4 (Top)) leads to a new nitron, the bifunctional nitron **120**.



**Figure 4.4:** (Top) Nitron **120** bearing two recognition sites with (left), the mono-functional original nitron **52** showing the amidopyridine recognition site on the carbon atom side of the nitron reactive site with added F to simulate electron with drawing amide and (right) mono-functional design of an amidopyridine recognition site on to the other (nitrogen) side of the nitron reactive site leads to the 'reverse' nitron **121**.

Before the behaviour of a bifunctional nitron bearing two recognition groups can be investigated, the kinetic behaviour arising from each of the two recognition sites on



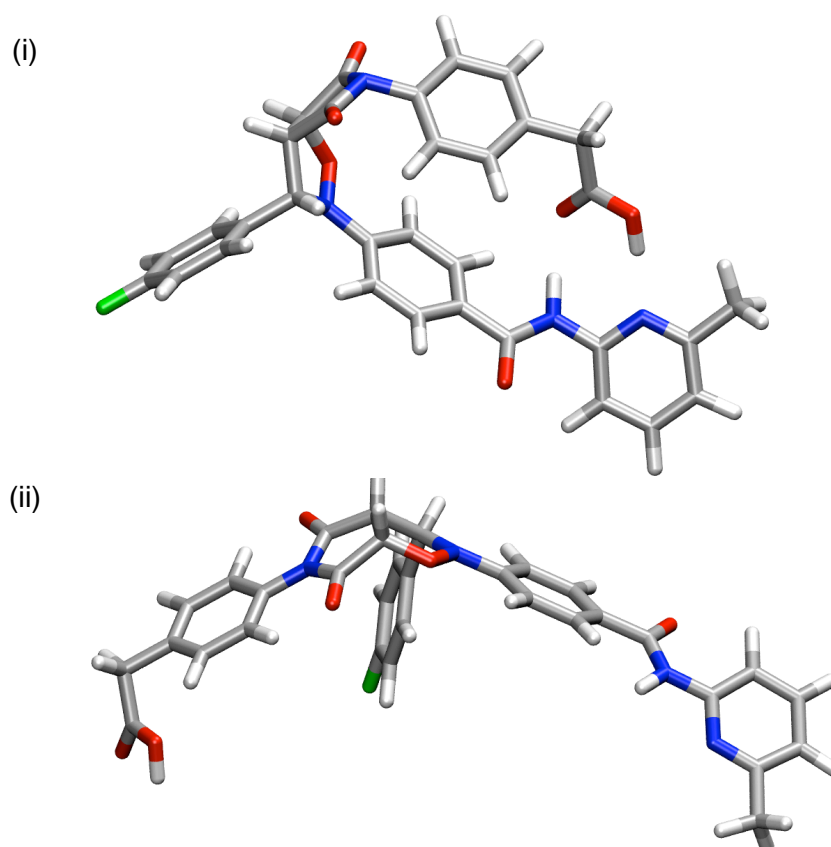
the nitrone must first be characterised individually. This will ensure that a successful [A•B] complex channel reaction is designed in isolation and then added to the self-replicating system. Thus the bifunctional nitrone can be split in half at the nitrone reactive site to create two mono-functional nitrones (Figure 4.4 left and right) each with one recognition site to facilitate the desired recognition mediated behaviour.

We are already familiar with the behaviour of the original nitrone **52** (Figure 4.4 left) which reacts autocatalytically with maleimide **34** in a self-replicating reaction selective for the *trans* cycloadduct **55**. The structure of the original nitrone has been modified slightly to include a *para* fluoride group to simulate the electronic properties of the electron withdrawing amide group, present in the bifunctional nitrone, and aid monitoring of the reaction progress by <sup>19</sup>F NMR spectroscopy. Any effects these changes have on the self-replicating behaviour will be investigated later.

The second nitrone **121** designed (Figure 4.4 right) has an amidopyridine group on the other side of the nitrone reactive site and is dubbed the '*Reverse*' nitrone as a result of the reversal of the reactive site with respect to the recognition site. A *para* fluoride group has also been positioned to simulate the electronic properties of the electron withdrawing *para* amide, as will be present in the bifunctional nitrone **120**, and also to facilitate facile kinetic monitoring of the reaction progress by <sup>19</sup>F NMR spectroscopy.

In order to obtain an insight into the expected behaviour of the *reverse* nitrone upon reaction with maleimide **34**, we turned to molecular modelling. Models of the two diastereoisomeric cycloadducts, *trans* and *cis*, were constructed and subjected to energy minimisation and conformational search calculations. The lowest energy conformers for the formation of the *trans* and *cis* cycloadducts were selected and are displayed in Figure 4.5.

The model of the *trans* cycloadduct reveals the structure to have a folded shape in which the recognition sites are in close proximity after the cycloaddition inferring that they will have participated in the construction of a transition state. The molecular model obtained of the *cis* cycloadduct reveals the recognition sites to be far apart and open inferring they will have played no role in forming the transition state. The results of the molecular modelling indicate that the *trans* cycloadduct was likely to be

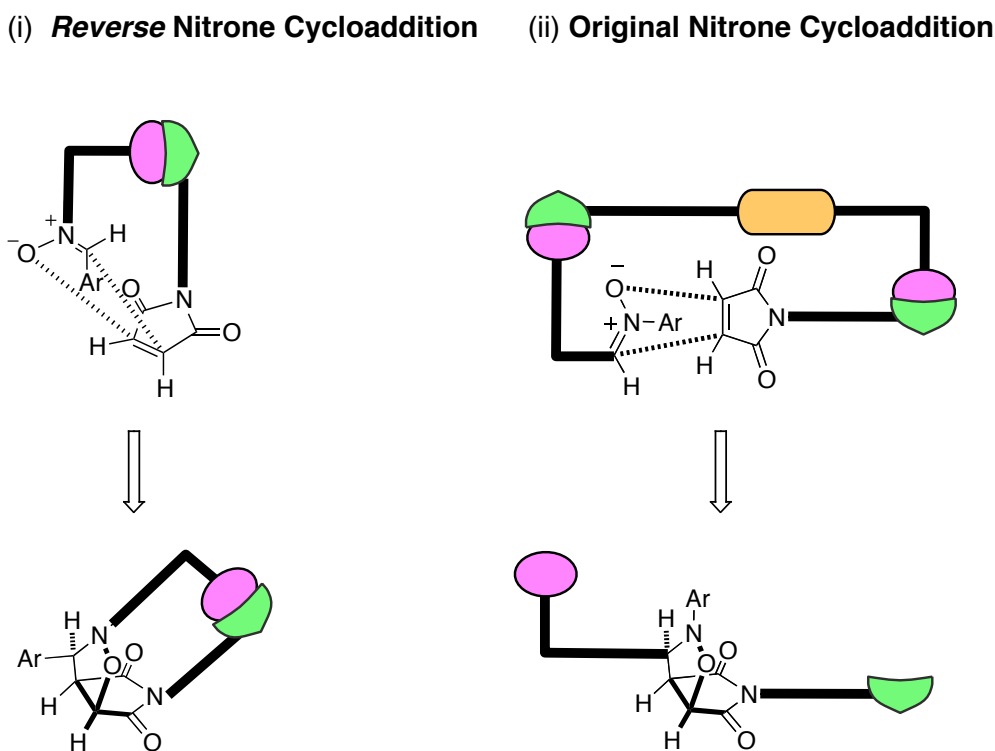


**Figure 4.5:** Molecular model of (i) the *trans*-122 cycloadduct showing the two recognition sites in close proximity after cycloaddition inferring their participating in construction (ii) the *cis*-122 cycloadduct showing an open structure with recognition sites far apart therefore unlikely to have been formed through an [A•B] complex channel. Molecular models constructed using Maestro and calculation carried out using AMBER\* force field with CHCl<sub>3</sub> solvation model with MacroModel. Representation of atoms: carbon atoms shown in grey, oxygen atoms in red, nitrogen atoms in blue, fluorine in green and hydrogen in light grey.

constructed by the desired [A•B] complex channel behaviour. One might anticipate that with recognition sites open, as in provided by the *cis* cycloadduct, the template may be capable of template directed self-replication. We expected that the orientation of the *para*-fluorophenyl ring, pointing between the recognition sites, would prevent the formation of a ternary complex as it may be sterically unfavourable and thus inhibit self-replication.

By considering the transition states leading to each of the two cycloadducts with respect to the recognition sites, it is clear why the *trans* cycloadduct is selected from the formation of an [A•B] complex (Figure 4.6 (i)). With the recognition sites satisfied, the transition state which arises from an [A•B] complex shows that the orientation of the nitro and maleimide reactive sites in this model have the proton on the nitro and the protons on the maleimide orientated in opposite directions leading to the

*trans* cycloadduct. This selectivity serendipitously leads to the formation of a cycloadduct with the same diastereoisomeric core to that of the self-replicator which is also selective for the *trans* cycloadduct as shown in Figure 4.6 (ii). This result indicates that both recognition sites on the bifunctional nitron, although through a different recognition mediated behaviour, will be selective for the *trans* cycloadduct.



**Figure 4.6:** Representations of the recognition mediated cycloaddition transition state of (i) the reverse nitron to form the *trans* cycloadduct via an [A•B] complex and (ii) the original nitron to form the *trans* cycloadduct via the autocatalytic channel on preformed template.

#### 4.4.1 Opening Recognition Sites on AB Template

Key to the success of the bifunctional nitron system would be the ability of the recognition sites used to promote template formation through an [A•B] complex to dissociate allowing the sites to be open to solution after reaction. In order to evaluate whether this behaviour could occur, we performed molecular dynamic calculations on the *trans* template formed from an [A•B] complex channel.

Initially, a molecular model of *trans*-122 was created using Macromodel with the calculation carried out with the AMBER\* force field and constraining the distance

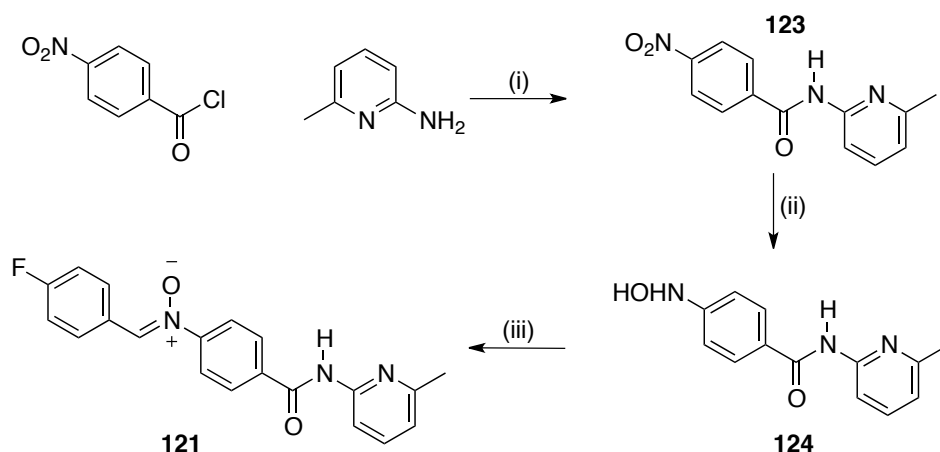
between the intramolecular recognition sites to  $3 \text{ \AA} \pm 2 \text{ \AA}$  to represent the position of the recognition sites after reaction to form the *trans* cycloadduct. This model demonstrated that after cycloaddition, the association of the recognition sites caused a slight strain of the fused rings at the cycloaddition site.

Next, a molecular dynamics simulation was performed using Macromodel subjecting the model of *trans*-**122** to thermal stimulation at  $0 \text{ }^\circ\text{C}$ . The result of the calculation was that the recognition sites were released from association and opened to solution, relieving the strain at the fused rings.

The desired behaviour of the cycloadduct opening the recognition sites to solution was therefore expected and we proceeded to investigate the properties of the system experimentally.

#### 4.4.2 Synthesis of Nitron 121

With molecular modelling proving reassuring insights into the expected reactivity, the reagents were prepared with standard synthetic chemistry starting with the target nitron **121**.



**Scheme 4.4:** Reagents and conditions (i)  $\text{CH}_2\text{Cl}_2$ ,  $\text{NEt}_3$ ,  $0 \text{ }^\circ\text{C} \rightarrow \text{RT}$ , 16 h, 97% (ii) 5% Rh/C,  $\text{NH}_2\text{NH}_2 \cdot \text{H}_2\text{O}$ , THF, RT, 1 h, 84% (iii) 4-fluorobenzaldehyde, EtOH, RT, 16 h, 57%

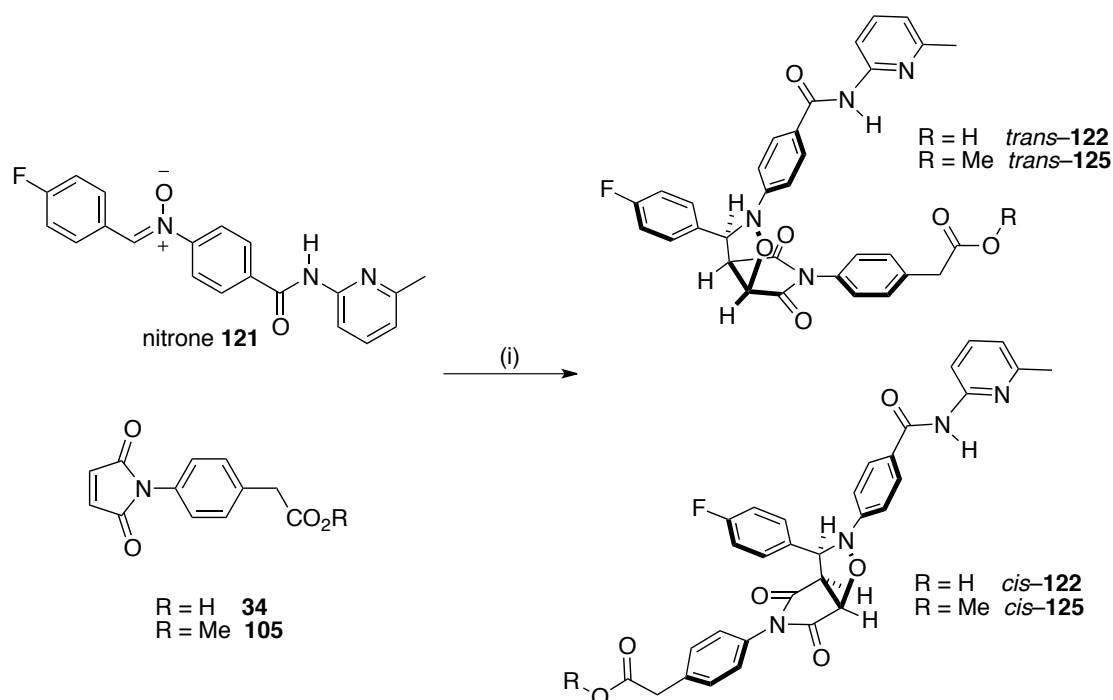
Nitron **121** was synthesised in three steps. Firstly the nitro amide **123** was prepared from an amide coupling of 4-nitrobenzoyl chloride and 2-amino-6-methylpyridine in dichloromethane and triethylamine. The nitro group was then reduced to the

hydroxylamine **124** using transfer hydrogenation with 5% Rh/C and hydrazine hydrate in THF before a condensation of the hydroxylamine with 4-fluorobenzaldehyde yielded the target nitron **121**.

#### 4.4.3 Synthesis of Nitron 52

The overall goal of this project is to create a system taking advantage of [A•B] complex channel behaviour as well as autocatalytic self-replicating behaviour. Therefore, for comparison, the isolated self-replicating system must be created and characterised. A slight modification is made to the nitron from the reported replicating system which is the inclusion of a *p*-fluoride substituent on the phenyl ring which serves to allow facile monitoring of the system by <sup>19</sup>F NMR spectroscopy as well as mimicking the electronic properties for the *p*-amide. The synthetic route to produce nitron **52** was previously discussed in Section 2.2.1.

#### 4.4.4 Kinetic Behaviour of [A•B] System



**Scheme 4.5:** Reaction of the maleimide can lead to the formation of the *trans* or *cis* diastereoisomeric products. Reagents and conditions (i) CDCl<sub>3</sub>, 0 °C, 16 h.

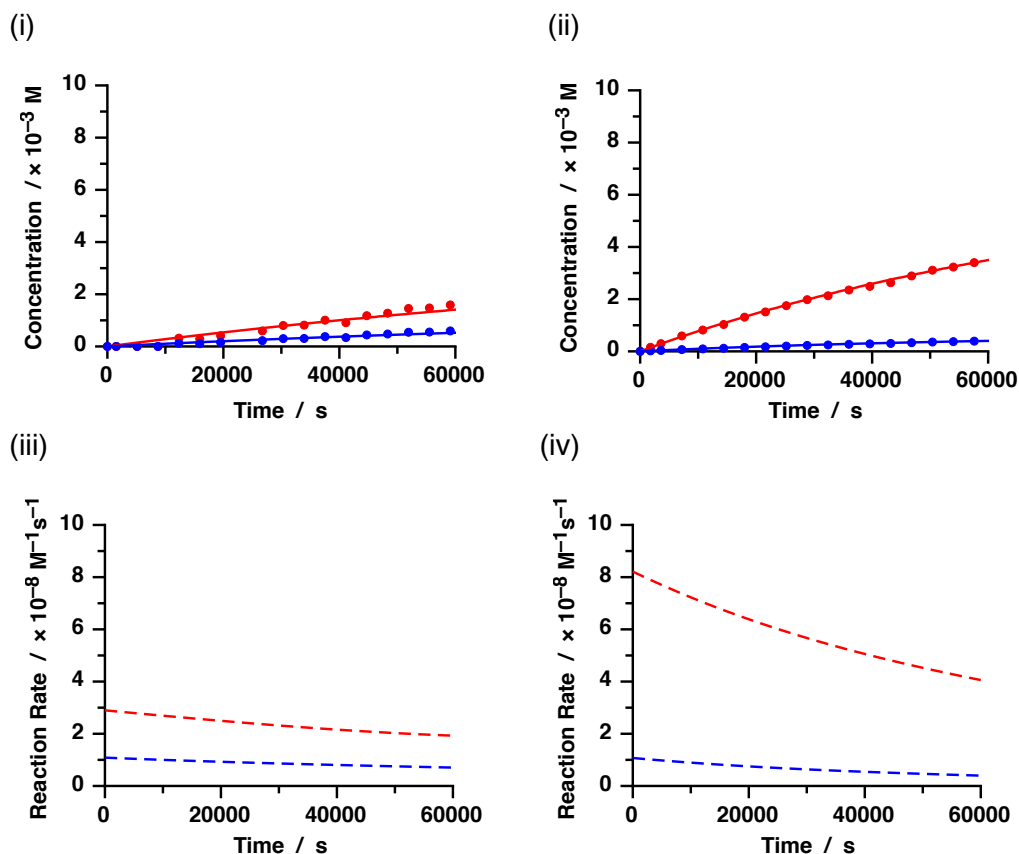
Before the recognition mediated pathway is investigated, it is important to characterise the native bimolecular reaction between the nitron and maleimide. The product conversion and *trans:cis* ratio of this control will serve as the baseline

behaviour for the background reaction and with this known, any change caused by recognition mediated process can easily be identified.

The recognition inhibited control reaction was carried out by mixing nitrone **121** with the maleimide ester **105** at a concentration of 10 mM in CDCl<sub>3</sub> and incubated at 0 °C. The progress of the reaction was monitored by 470.3 MHz <sup>19</sup>F NMR spectroscopy every 30 minutes for 16 h and the relative concentrations of species in the reaction at each time point calculated by deconvolution of the appropriate resonances from the spectra to create a concentrations vs. time profile.

The concentration vs. time profile for the control reaction (Figure 4.7 (i)) revealed the reaction progresses slowly under these conditions producing 1.4 mM *trans*-**125** and 0.52 mM *cis*-**125** representing an overall conversion of 19% and a *trans* : *cis* ratio of 2.7 : 1. Kinetic simulation and fitting was carried out using the SimFit program and revealed the rate constants  $k_{trans}$  and  $k_{cis}$  to be  $2.92 \times 10^{-4} \text{ Ms}^{-1}$  and  $1.08 \times 10^{-4} \text{ Ms}^{-1}$  respectively. By taking the first derivative of a curve fitted to the experimental data in the concentration vs. time profile, the equation can be plotted and describes the rate of reaction at each time point. The rate of reaction to form *trans*-**125** and *cis*-**125** is shown (Figure 4.7 (iii)) below the concentration vs. time profile and shows the expected behaviour for a simple bimolecular reaction.

The recognition mediated reaction was next performed by reacting maleimide **34** with nitrone **121** under the same conditions. The reaction was monitored as before and a concentration vs. time profile constructed after deconvolution of the spectra. The concentration vs. time profile for the recognition mediated reaction (Figure 4.7 (ii)) clearly demonstrated an enhancement in the production of *trans*-**122** of which 3.4 mM was produced after 16 h at the expense of *cis*-**122** of which 0.39 mM was produced after 16 h representing an overall conversion of 38% and a *trans* : *cis* ratio of 9:1. The rate vs. time profile (Figure 4.7 (iv)) produced in the same manner as before showed that the rate of production of *trans*-**122** was increased in comparison to the control reaction. The maximum rate of production of *trans*-**122** occurs at the start of the reaction, consistent with the behaviour of a reaction proceeding through an [A•B] complex channel. The effect of the recognition in the system has been the selective enhancement in the formation of the *trans* cycloadduct as expected from our molecular modelling predictions.

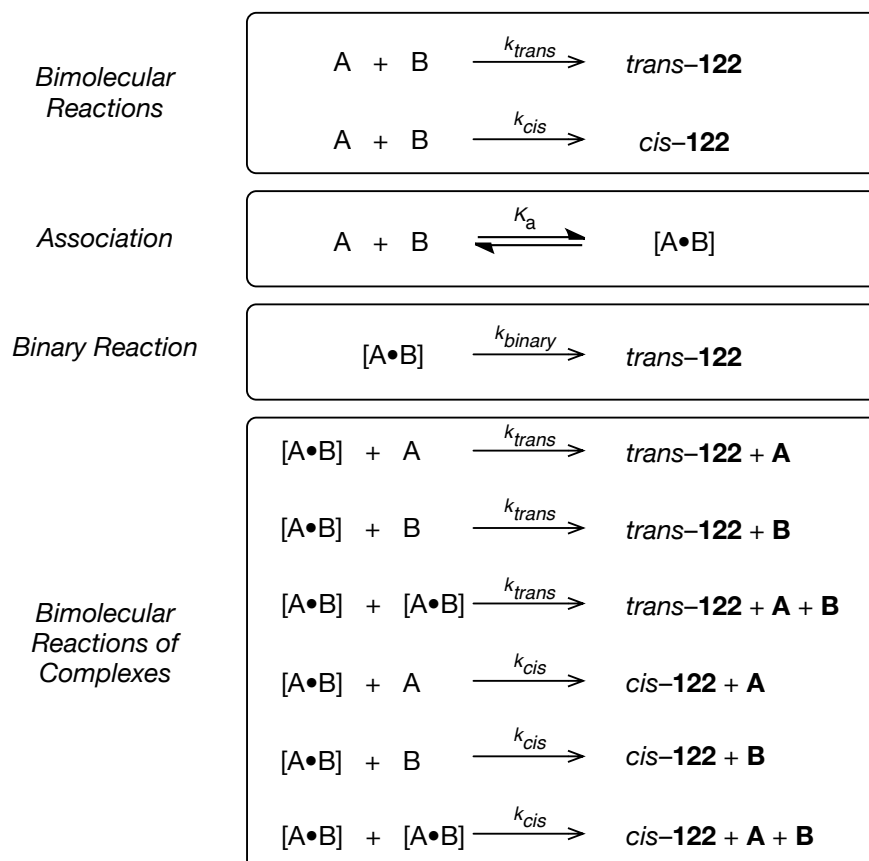


**Figure 4.7:** (i) Concentration vs. time profile for the recognition inhibited control reaction of nitrone **121** with maleimide **105** (ii) Concentration vs. time profile for the recognition mediated reaction of nitrone **121** with maleimide **34**. Experimental data shown as points with *trans* products represented as (●) and *cis* products represented as (●). The best fit of the data to a kinetic model for an [A•B] complex reaction is represented as lines with *trans* products represented by (—) and *cis* products by (—). (iii) Rate vs. time profile for the recognition inhibited control reaction of nitrone **121** with maleimide **105**. (iv) Rate vs. time profile for the recognition mediated reaction of nitrone **121** with maleimide **34**. The rate of production of the *trans* cycloadduct shown as a red dashed line and the *cis* cycloadduct as a blue dashed line. Reaction conditions: 10 mM,  $CDCl_3$ , 0 °C, 16 h.

**Table 4.1:** Reaction rate constant data calculated from kinetic simulation and fitting using SimFit.

	$k_{trans}$	$k_{cis}$
Bimolecular Rate Constant / $\times 10^{-4} M^{-1}s^{-1}$	2.92	1.08
$k_{binary} / \times 10^{-4} s^{-1}$	0.0737	—
EM Binary Complex / M	0.025	—

Kinetic simulation and fitting of the system was carried out using a simple kinetic model (Scheme 4.6) using the values of  $k_{trans}$  and  $k_{cis}$  found previously from the control reaction and the  $K_a$  reported<sup>139</sup> previously as 1020  $M^{-1}$  in  $CDCl_3$  at 0 °C. The model also assumes that the reactants associated together are able to react in a bimolecular reaction with the same rate constants as when in isolation.



**Scheme 4.6:** Simple kinetic model to used to represent the process in the recognition mediated system.  $k_{trans}$  and  $k_{cis}$  are the reaction rate constants as determined from the recognition inhibited control.  $K_a$  the association constant between an amidopyridine and a phenylacetic acid at 0 °C in  $CDCl_3$  and  $k_{binary}$  the rate of reaction in the binary complex. Nitron **121** is represented as **A** and maleimide **34** as **B** for simplicity.

Kinetic simulation and fitting revealed the reaction rate constant of cycloaddition in the  $[A \cdot B]$  binary complex  $k_{binary}$  to be  $7.37 \times 10^{-6} \text{ s}^{-1}$ . The *EM* of the recognition mediated reaction was therefore determined to be 0.025 M, a modest increase compared to the actual reaction concentration of 0.010 M indicating that the reaction is behaving as if the concentration was 2.5 × greater than it actually is.

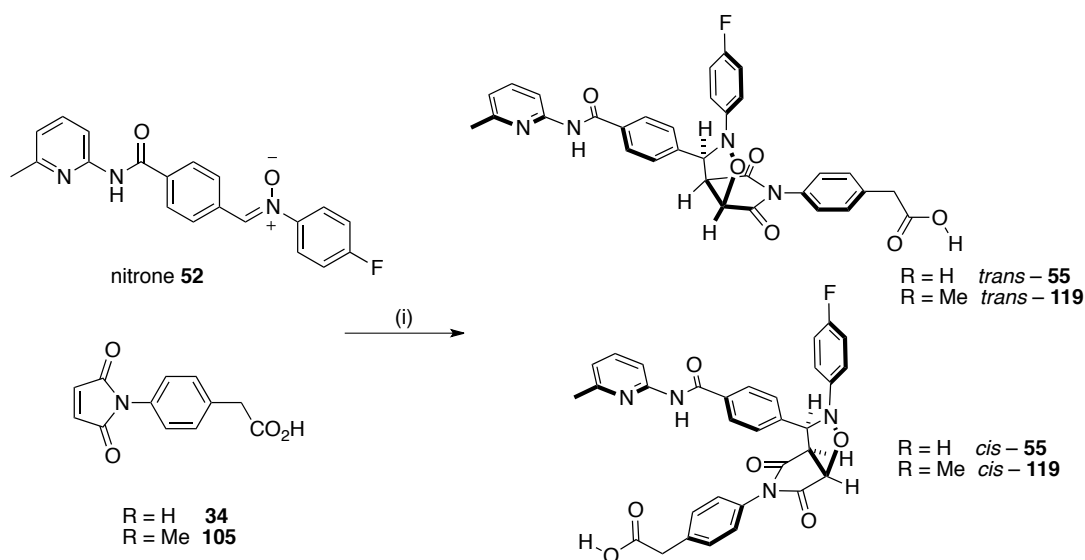
A further control to prove the reaction is not autocatalytic, *i.e.* template directed, was carried out by adding preformed *trans*-**122** to a fresh reaction of nitron **121** and maleimide **34** and monitoring the reaction under identical conditions as before. No change in the rate of formation of *trans*-**122** or *cis*-**122** was identified and no change of *trans* : *cis* ratio was seen indicating that there was no template directing effect on the system.



From the results, it was clear that the recognition site designed on to the nitrogen side of the nitron reactive site is able to associate the maleimide to form an  $[A \cdot B]$  binary complex and selectively promote the cycloaddition to form the *trans* cycloadduct

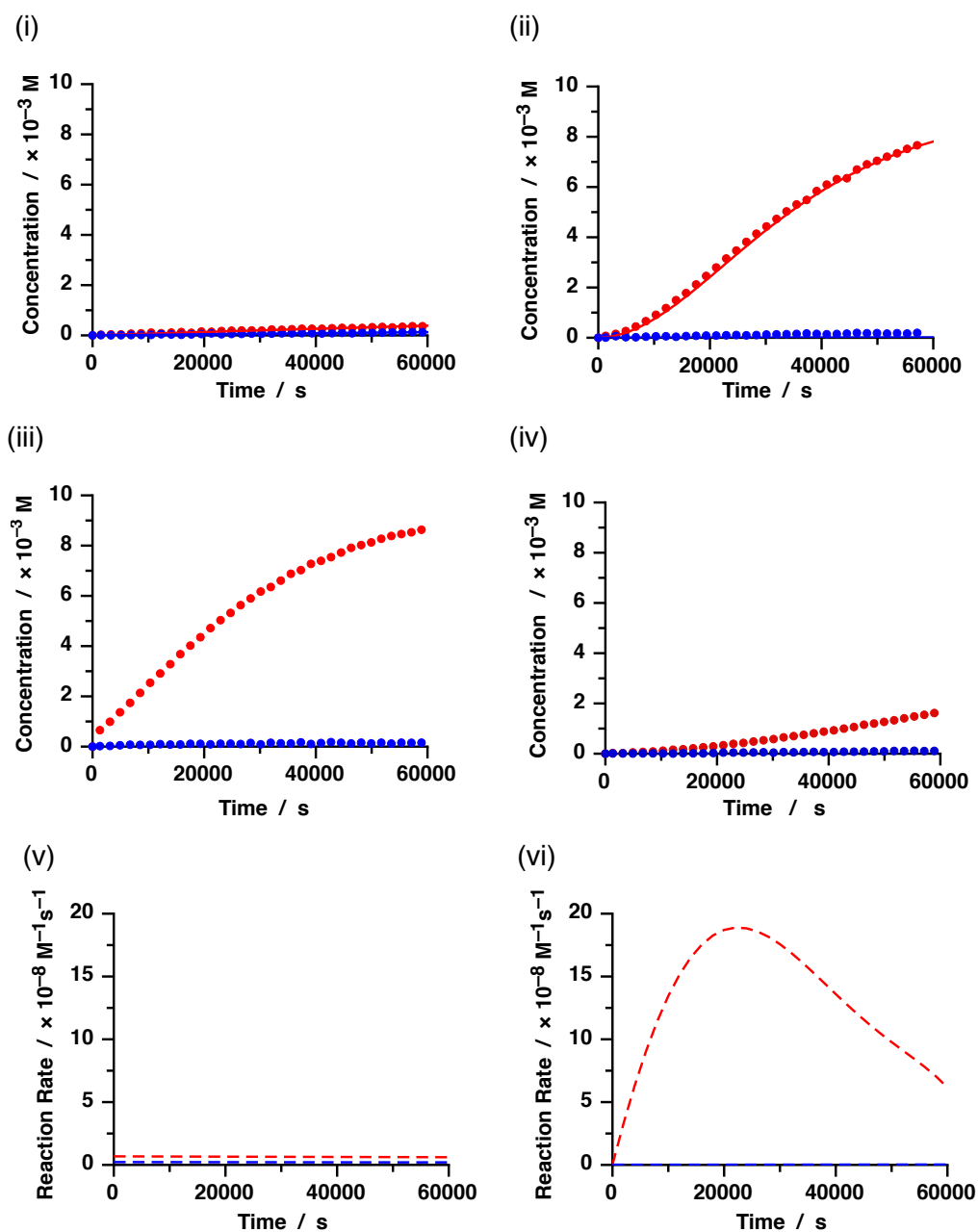
#### 4.5 Kinetic Behaviour of the Self-Replicating System

Whilst the behaviour of the self-replicating system is well understood, it was important to investigate whether or not altering the electronic properties of the nitron by introducing an electron withdrawing group *para* to the nitron reactive site would have any effect on its reactivity.



**Scheme 4.7:** Reaction of the maleimide with nitron **52** can lead to the formation of the *trans* or *cis* diastereoisomeric products. Reagents and conditions (i)  $CDCl_3$ ,  $0\text{ }^\circ\text{C}$ , 16 h.

In order to determine the reactivity of the 1,3-dipolar cycloaddition in the absence of recognition effects, the recognition disabled control reaction was performed by reacting nitron **52** with maleimide ester **105** under the same conditions as before, (10 mM, in  $CDCl_3$ ,  $0\text{ }^\circ\text{C}$ ) and assaying in the same manner. The concentration vs. time profile (Figure 4.8 (i)) showed the reaction to be slow under these conditions producing 0.36 mM *trans*-**119** and 0.13 mM *cis*-**119** representing an overall conversion of 5% after 16 h and a *trans* : *cis* ratio of 2.8:1. Subjecting the results to kinetic simulation and fitting allowed us to obtain the reaction rate constants of  $k_{trans}$  and  $k_{cis}$  to be  $6.81 \times 10^{-5}\text{ Ms}^{-1}$  and  $2.25 \times 10^{-5}\text{ Ms}^{-1}$  respectively. The rate vs. time profile (Figure 4.8 (v)) for the control reaction once again displayed the expected profile for a simple bimolecular reaction.



**Figure 4.8:** Concentration vs. time profiles of (i) Recognition disabled control reaction, (ii) Recognition mediated reaction, (iii) template doped recognition mediated reaction (iv) recognition mediated reaction with competitive inhibitor benzoic acid (5 eq). Experimental data shown as points with *trans* products represented as (●) and *cis* products represented as (●). The best fit of the data to a kinetic model for a self-replicating reaction is represented as lines with *trans* products represented by (—) and *cis* products by (—). (v) Rate vs. time profile for the recognition inhibited control reaction of nitron **52** with maleimide **105**. (vi) Rate vs. time profile for the recognition mediated reaction of nitron **52** with maleimide **34**. The rate of production of the *trans* cycloadduct shown as a red dashed line and the *cis* cycloadduct as a blue dashed line. Reaction conditions: 10 mM, CDCl<sub>3</sub>, 0 °C, 16 h.

The reaction mediated behaviour was next observed by performing the reaction of maleimide acid **34** with nitron **52** under the same conditions and monitored as before. The concentration vs. time profile (Figure 4.8 (ii)) now shows a characteristic

**Table 4.2:** Reaction rate constant data calculated from kinetic simulation and fitting using SimFit.

	$k_{trans}$	$k_{cis}$
Bimolecular Rate Constant / $\times 10^{-5} \text{ M}^{-1}\text{s}^{-1}$	6.81	2.25
$k_{ternary} / \times 10^{-5} \text{ s}^{-1}$	186	—
$EM$ Ternary Complex / M	27	—
$K_{Duplex} / \times 10^6 \text{ M}^{-1}$	3.01	—
Connection $EM$ / M ( $\Delta G^{\circ} / \text{kJmol}^{-1}$ )	2.9 (2.4)	—

sigmoidal shape for the production of *trans*-55 of which 7.66 mM had been produced after 16 h along with 0.19 mM of *cis*-55 representing an overall conversion of 79% with a *trans* : *cis* ratio of 40:1. The sigmoidal profile is a classic characteristic of an autocatalytic self-replicating reaction. The results of the reaction were subjected to kinetic simulation and fitting to a self-replicating reaction model using SimFit. The results revealed the rate of reaction within the ternary complex to be  $186 \times 10^{-5} \text{ s}^{-1}$  from which the  $EM$  of the reaction was determined to be 27 M. A rate vs. time profile was once again produced (Figure 4.8 (vi)) and revealed that the rate is slow in the initial stages before increasing as the reaction progresses and finally slowing as the reactants are consumed. This rate vs. time profile shows the characteristic non-linear behaviour for a self-replicating reaction with the maximum rate reached after ~22 000 s.

In order to demonstrate that the reaction is template directed, preformed *trans*-55 was added to a fresh set of reagents for the recognition mediated reaction and the reaction performed and monitored as before. The concentration vs. time profile for this reaction (Figure 4.8 (iii)) show the lag period is absent and the reaction proceeds at the maximum autocatalytic rate from the start which is consistent with the template acting as a catalyst for its own formation.

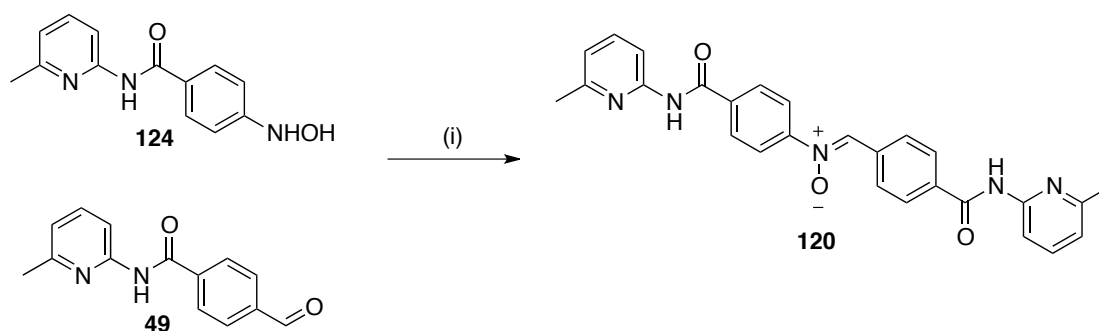
A further control reaction was carried out to demonstrate the crucial effect of recognition within the system. An excess of a competitive inhibitor, benzoic acid, was added to a fresh set of reagents for the recognition mediated system and the reaction performed under the same conditions and monitored as before. The concentration vs. time profile for this reaction (Figure 4.8 (iv)) demonstrates that the effect this had

was to strongly inhibit the autocatalytic reaction, reducing the production of *trans*-**55** to 1.7 mM and *cis*-**55** to 0.11 mM representing an overall conversion of 18% with the *trans* : *cis* ratio falling to 17:1. This is a consequence of the benzoic acid interfering with the recognition processes by associating with the amidopyridine therefore inhibiting the formation of the the [A•B•T] ternary complex and thus reducing the efficiency of the replicator.

From the results, it is clear that the reaction of maleimide **34** with nitrone **52** is autocatalytic and selective for the production of *trans*-**55**. There has been no detrimental effect from adding the fluoride group *para* to the nitrone reactive site.

#### 4.6 Synthesis of Bifunctional Nitrone **120**

With the desired recognition mediated successfully designed by placing a recognition site on either side of the nitrone reactive site and evaluating in mono functional nitrones, the bifunctional nitrone can now be produced to evaluate the effect of combining the two recognition sites on to one molecule.



**Scheme 4.8:** Reagents and conditions (i) EtOH, RT, 16 h, 54%

Bifunctional nitrone **120**, decorated with two recognition sites, was synthesised from two previously prepared compounds. Condensation of aldehyde **49** with hydroxylamine **124** furnished the target nitrone in acceptable yields.

##### 4.6.1 Kinetic Analysis of the Reaction of Nitrone **120** with Maleimide **34**

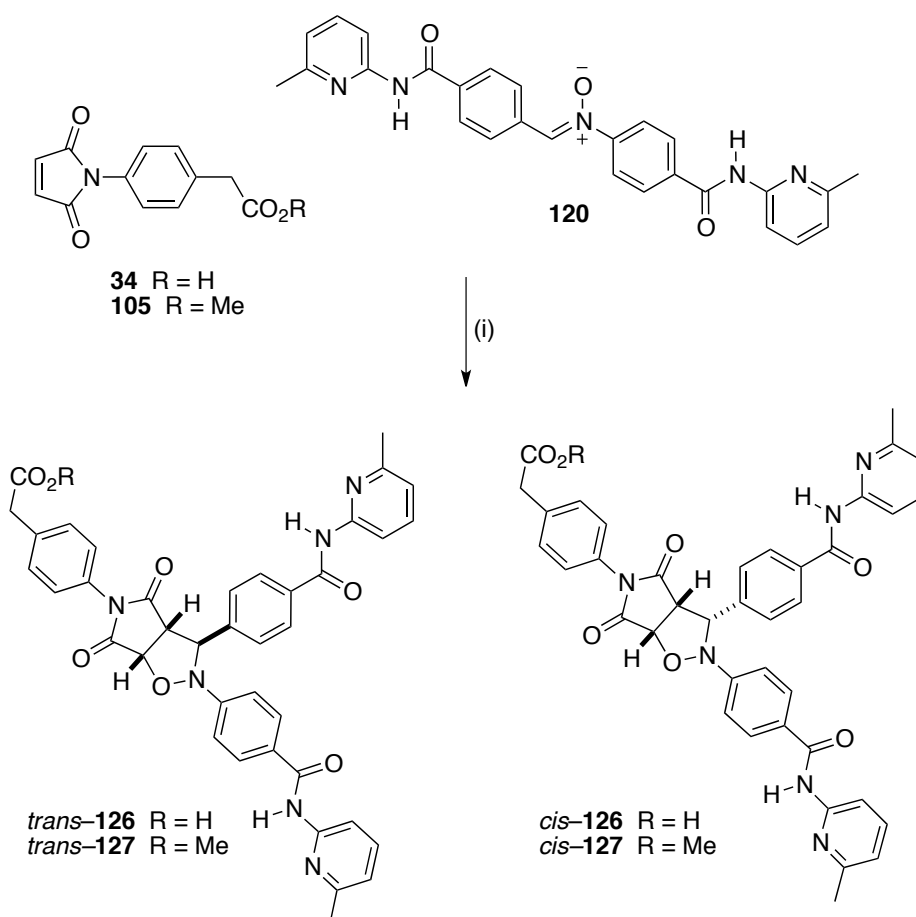
As a consequence of the poor solubility of bifunctional nitrone **120** in chloroform, a lower concentration of 5 mM had to be employed to perform the reactions. Whilst this

was not ideal to change to the condition of the reactions, the concentration is still above the dissociation constant ( $K_d$ ) of the amidopyridine with the phenylacetic acid and therefore high enough to permit efficient recognition mediated behaviour.

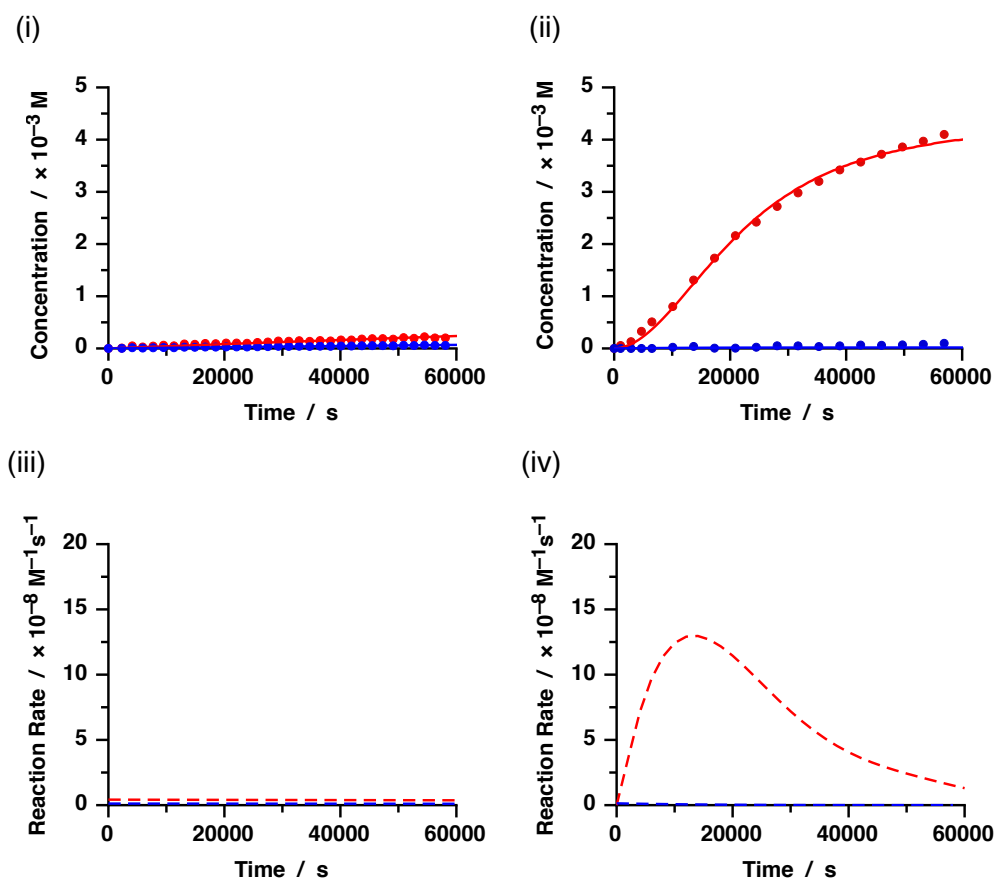
$$K_a = \frac{1}{K_d} \quad \therefore \quad K_d = \frac{1}{K_a}$$

$$K_d = \frac{1}{1020} \cdot \frac{1}{M^{-1}} = 0.99 \text{ mM}$$

**Equation 4.1:** The dissociation constant  $K_d$  is inversely proportional to the experimentally determined association constant  $K_a$ . When the concentration of reagents equals the dissociation constant, there will be a 50:50 ratio of associated to unassociated material. At concentrations below the  $K_d$  isolated reagents will be in excess. At concentrations above  $K_d$  associated reagents will be in excess.



**Scheme 4.9:** Reagents and conditions (i) 5 mM, CDCl<sub>3</sub>, 0 °C, 16 h



**Figure 4.9:** Concentration vs. time profiles of (i) Recognition disabled control reaction, (ii) Recognition mediated reaction. Experimental data shown as points with *trans* products represented as (●) and *cis* products represented as (●). The best fit of the data to a kinetic model for an including both an [A•B] complex reaction and self-replicating reaction is represented as lines with *trans* products represented by (—) and *cis* products by (—). (iii) Rate vs. time profile for the recognition inhibited control reaction of nitrone **120** with maleimide **105**. (iv) Rate vs. time profile for the recognition mediated reaction of nitrone **120** with maleimide **34**. The rate of production of the *trans* cycloadduct shown as a red dashed line and the *cis* cycloadduct as a blue dashed line. Reaction conditions 5 mM, CDCl<sub>3</sub>, 0 °C, 16 h.

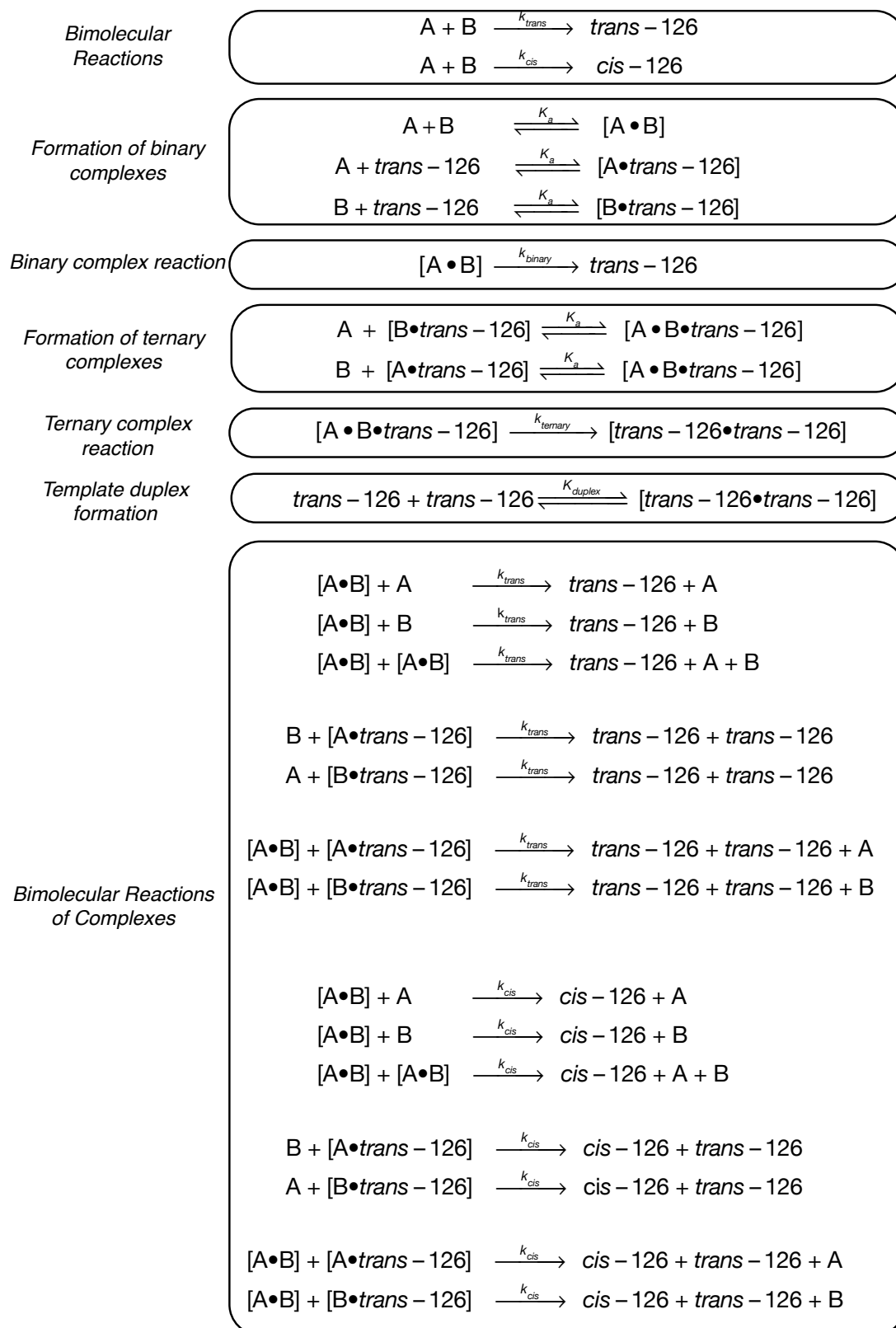
**Table 4.3:** Reaction rate constant data calculated from kinetic simulation and fitting using SimFit.

	$k_{\text{trans}}$	$k_{\text{cis}}$
Bimolecular Rate Constant / $\times 10^{-4} \text{ M}^{-1} \text{ s}^{-1}$	1.69	0.48
$k_{\text{ternary}} / \times 10^{-4} \text{ s}^{-1}$	36.4	—
EM Ternary Complex / M	21.5	—
$K_{\text{Duplex}} / \times 10^6 \text{ M}^{-1}$	1.92	—
Connection EM / M ( $\Delta G^\ddagger / \text{kJmol}^{-1}$ )	1.88 (1.4)	—

In order to characterise the behaviour of the cycloaddition in the absence of recognition effects, the recognition disabled control reaction of nitrone **120** with

maleimide ester **105** was performed under the same conditions described earlier (5 mM, CDCl<sub>3</sub>, 0 °C) and monitored by 500.1 MHz <sup>1</sup>H NMR spectroscopy every 30 minutes for 16 h. After deconvolution of each spectrum a concentration vs. time profile was constructed (Figure 4.9 (i)) which demonstrated that the reaction was slow under these conditions with 0.2 mM *trans*-**127** and 0.06 mM *cis*-**127** produced representing an overall conversion of 5% and a *trans* : *cis* ratio of 3.3:1. Kinetic simulation and fitting of the experimental data using SimFit revealed the reaction rate constants of  $k_{\text{trans}}$  and  $k_{\text{cis}}$  to be  $1.69 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  and  $0.48 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  respectively.

The effect of combining two recognition sites on one nitrone was next investigated by monitoring the recognition mediated reaction between nitrone **120** with the acid maleimide **34** under the same conditions and monitored once again by 500.1 MHz <sup>1</sup>H NMR spectroscopy. Deconvolution of the spectra once again was undertaken to produce a concentration vs. time profile for the reaction (Figure 4.9 (ii)) which showed an increase in production of *trans*-**126** with 4.1 mM produced after 16 h along with 0.1 mM *cis*-**126** representing an overall conversion of 82% and a *trans* : *cis* ratio of 41:1. In order to extract reaction rate information, a kinetic model, shown in Scheme 4.10, was produced which combined the processes of an [A•B] complex model with the processes of a self-replicating reaction model. The model assumed that the rate of reaction in the [A•B] complex,  $k_{\text{binary}}$ , was consistent with that of the mono-functional system and allowed the rate of reaction within the ternary complex,  $k_{\text{ternary}}$ , and template duplex strength,  $K_{\text{duplex}}$ , to vary. The experimental data was then subjected to kinetic simulation and fitting to the model using SimFit the results of which, shown in Table 4.3, revealed that the rate of reaction in the ternary complex was  $36.4 \times 10^{-4} \text{ s}^{-1}$  from which the *EM* of the reaction was determined to be 21.5. It is interesting to note that the results highlight the efficiency of the ternary complex reaction is lower in the bifunctional system than in its mono-functional system. This may be a consequence of an increase in competition for binding sites. The concentration of amidopyridine sites on the nitrone is double what it was in the monofunctional system and each will compete to associate with the complementary phenylacetic acid. As each amidopyridine recognition site has the same  $K_a$  with the phenylacetic acid then on average the maleimide will spend half the time on each site.

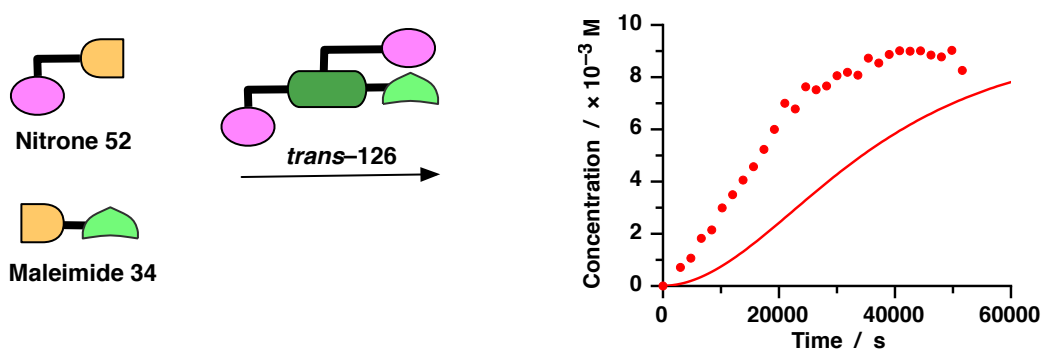


**Scheme 4.10:** Simple kinetic model to used to represent the process in the recognition mediated system.  $k_{trans}$  and  $k_{cis}$  are the reaction rate constants as determined from the recognition inhibited control.  $K_a$  the association constant between an amidopyridine and a phenylacetic acid at 0 °C in  $CDCl_3$ ,  $k_{binary}$  the rate of reaction in the binary complex as determined by the mono-functional system,  $K_{duplex}$  the association constant of the template duplex and  $k_{ternary}$  the rate of reaction in the ternary complex. Nitron **120** is represented as A and maleimide **34** as B for simplicity.

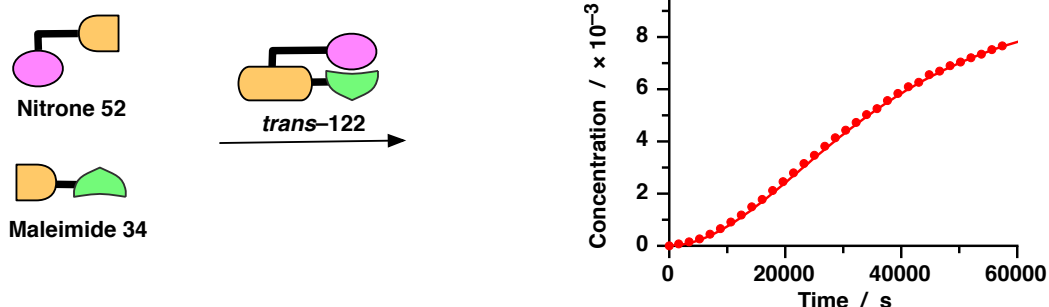


The association of the template duplex was also calculated to be weaker with a  $K_a$  of  $1.96 \times 10^6 \text{ M}^{-1}$ , lower than the mono functional self-replicating template duplex which was  $3.01 \times 10^6 \text{ M}^{-1}$ . This may also be as a result of the increased amidopyridine concentration or as a consequence of the increased steric bulk on the template. A rate vs. time profile was produced (Figure 4.9 (iv)) which revealed that the rate of reaction increases with time consistent with a self-replicating reaction however in this case, the maximum rate is reached after a shorter reaction time of  $\sim 12,000 \text{ s}$  in comparison to the mono-functional replicating system which took  $\sim 22,000 \text{ s}$  to reach the maximum rate. The reaction of maleimide **34** with bifunctional nitrone **120** through the binary reaction channel produces *trans*-**126**, capable of template directed formation of itself, more rapidly than otherwise possible through the simple bimolecular reaction. With the template present earlier in the reaction, the maximum autocatalytic rate is reached at an earlier stage of the reaction.

(i)



(ii)



**Figure 4.10:** (i) Reaction of mono-functional nitrone **52** with maleimide **34** in the presence of 10 mol% bifunctional template *trans*-**126**. Concentration vs. time profile for this reaction shown to the right with filled red circles representing the production of *trans*-**55** in the doped reaction and the red line representing the production of *trans*-**55** in the native reaction of nitrone **52** with maleimide **34** with no doping. (ii) Reaction of mono-functional nitrone **52** with maleimide **34** in the presence of 10 mol% mono-functional template *trans*-**122**. Concentration vs. time profile for this reaction shown on the right with filled circles representing the production of *trans*-**55** in the doped reaction with the red line representing the production of *trans*-**55** from the native reaction of nitrone **52** with maleimide **34** with no doping.

Two further control reactions were carried out to confirm the ability of bifunctional *trans*-**126** to act as a template for its own formation in a self-replicating fashion. Firstly, preformed 10 mol% *trans*-**126** was added to a fresh set of reagents for the mono functional self-replicating system, nitrene **52** and maleimide **34**, Figure 4.10 (i). This reaction was designed to probe whether or not the recognition sites in the bisamidopyridine *trans*-**126** are able to associate fresh reagents. The reaction was performed under the same conditions as before and the concentration vs. time profile for this system showed no lag period for the formation of *trans*-**55** which confirms that the recognition sites on the bifunctional template must be open and able to associate fresh reagents allowing it to catalyse the reaction.

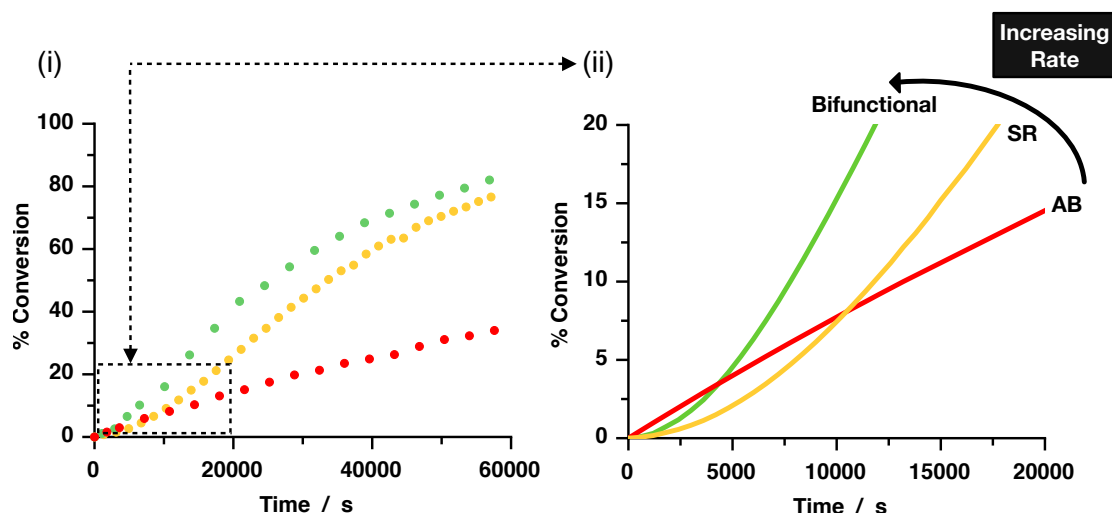
To confirm this observed behaviour arises from the amidopyridine recognition site on self-replicating side, as opposed to the [A•B] side, a further control was designed, Figure 4.10 (ii). 10 mol% preformed mono-functional *trans*-**122**, the template formed from the [A•B] complex channel reaction, was added to a fresh set of reagents for the self-replicating system and the reaction performed under the same conditions as before. The concentration vs. time profile for this system showed no increase in the rate of formation of the replicating template *trans*-**55** showing that the [A•B] complex product is catalytically inert.

These experiments and controls confirm that bifunctional *trans*-**126** is able to be assembled through a binary complex reaction at a rate above the simple bimolecular reaction and act as an autocatalyst for its own formation in a self-replicating manner as its recognition sites are open to solution after construction of the template and not associated together in a closed fashion. Although each of the individual self-replicating and [A•B] complex channel behaviours of the bisamidopyridine nitrene are less efficient than in isolation with *mono*-amidopyridine nitrenes, the combination of the two in one system still leads to a more efficient rate of cycloaddition.

#### 4.7 Relative Efficiencies of the Three Recognition Mediated Systems

An interesting observation is apparent when comparing the reaction profiles of the bifunctional nitrene to that of the mono-functional systems (Figure 4.11 (i)). The reaction which reaches the lowest overall conversion is that of the [A•B] complex

reaction (red) which reaches 34%. The self-replicating reaction (orange) reaches 77% with the bifunctional reaction (green) reaching the greatest conversion of 82%.



**Figure 4.11:** Concentration vs. time profiles of (i) the production of *trans* cycloadducts through the three recognition mediated behaviours (ii) and an enlargement of the initial 6 h with lines to aid visualisation showing that the initial rate of reaction of the bifunctional nitron **120** (●) proceeds similarly to the **[A•B]** complex channel (●) and both are more rapid than the self-replicating (SR) channel (●).

Closer inspection of the initial stages of the three reactions provides further evidence of the desired behaviour, Figure 4.11 (ii). Comparison of the initial rate of reaction in the mono-functional **[A•B]** complex system (red) to the initial rate in the mono-functional self-replicating system (orange) shows the expected behaviour that in the initial stages of the reaction, the **[A•B]** complex channel proceeds at a faster rate to the self-replicating system. However as a consequence of the **[A•B]** complex system being catalytically inert, once enough self-replicating template is produced, the autocatalytic cycle is initiated and the rate of production of the self-replicating template increases with time with the concentration produced eventually overtaking the **[A•B]** complex channel after roughly 10 000 s.

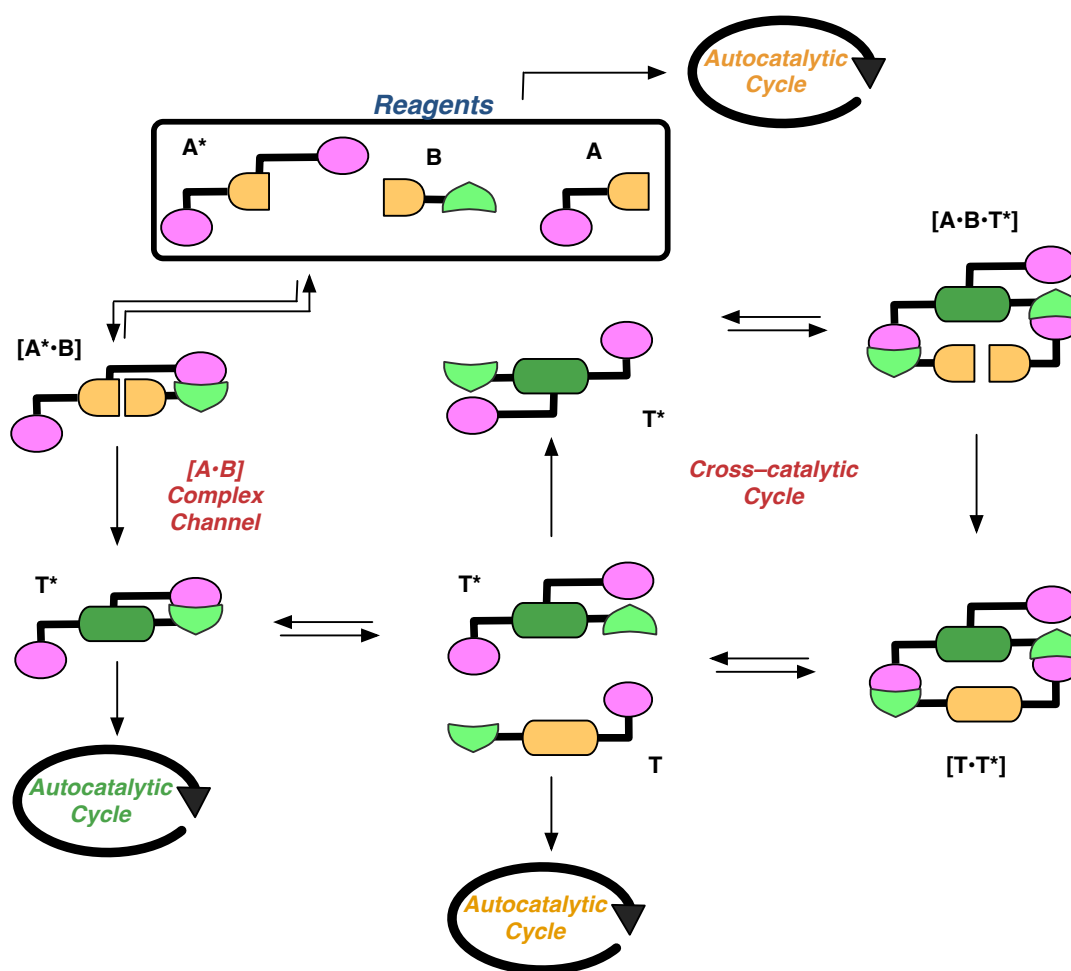
Inspection of the bifunctional system (green) shows that the production of bifunctional product is more rapid than the self-replicating product upon initiation although slightly less efficient than the **[A•B]** complex channel. With template produced more rapidly, the autocatalytic cycle initiated sooner allowing for more rapid production of template with time in comparison to the mono-functional self-replicating system.

This observation demonstrates that the bifunctional system is able to exploit both **[A•B]** complex channel and autocatalytic channel behaviour to facilitate a more rapid

fabrication of the *trans* cycloadduct and thus signifies the successful design of the desired bifunctional system.

#### 4.8 Cross-Catalytic Doping Reactions

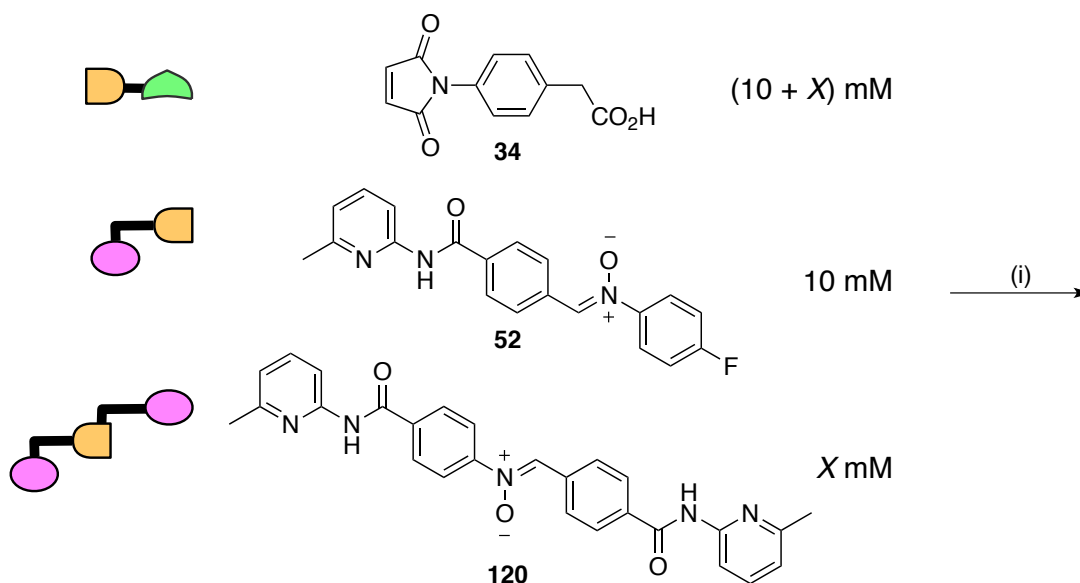
The desired behaviour programmed into the bifunctional nitrone is the ability to react with the acid maleimide to produce a template capable of further autocatalysis more rapidly than otherwise would occur by a simple bimolecular reaction. This autocatalytic template is able to catalyse its own formation and, as a result of the structural similarity, was also shown to accelerate formation of mono-functional self-replicating template from nitrone **52** and maleimide **34**.



**Figure 4.12:** In a system of three components, bifunctional nitrone **A\***, maleimide **B** and mono-functional **A**, in addition to the previously described autocatalytic channels, a cross catalytic channel is present. After reaction of **A\*** with **B** through a binary complex channel, bifunctional template **T\*** may be capable of participating in a cross catalytic reaction. Association of mono-functional nitrone **A** and maleimide **B** on template **T\*** renders the reaction to form **T** *pseudo*-intramolecular. As **T\*** is able to be produced at a rate greater than monofunctional **T**, the overall effect of doping **A\*** into the reaction should be an increased rate of production of **T** relative to an undoped reaction.

As the bifunctional nitron is able to react with maleimide **34** to produce autocatalytic template *trans*-**126** more rapidly than the mono-functional nitron under the same conditions, a combination of the two nitrones into one system should lead to an interesting catalytic effect (Figure 4.12).

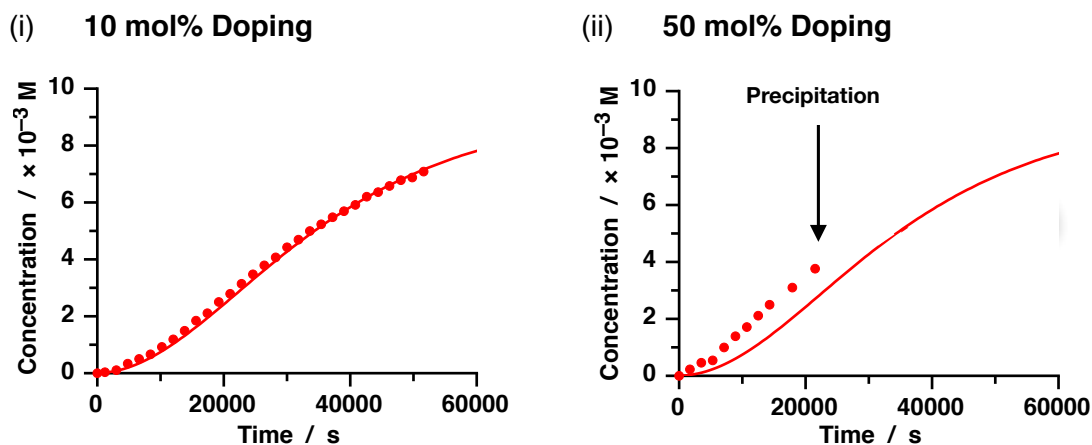
In a system which possessed mono functional nitron **52** (1 eq), bifunctional nitron **120** (*X* eq) and maleimide **34** ((1 + *X*) eq), we would expect the following behaviour. The bifunctional nitron **120** should react rapidly with maleimide **34**, via the combination of the [A•B] and autocatalytic channels available to it, to produce the bifunctional template *trans*-**126**. This process should occur more rapidly than the bimolecular reaction of mono functional nitron **52** with maleimide **34** to form *trans*-**55**. With the bifunctional template, *trans*-**126** available in solution, it will be able to catalyse the reaction of the mono functional nitron **52** with maleimide **34**. The overall affect is that the mono functional template will be produced more rapidly in the presence of bifunctional nitron than in isolation.



**Scheme 4.11:** Reagents and conditions (i) *X* = 1 or 5 mM, CDCl<sub>3</sub>, 0 °C, 16 h

A reaction was designed (Scheme 4.11) in which a substoichiometric amount of bifunctional nitron **120** was added to the reagents for the self-replicating system nitron **52** and maleimide **34** with enough maleimide to react entirely with both nitrones. The reaction would be performed under the same conditions as the isolated examples, at 10 mM of nitron **52** in CDCl<sub>3</sub> at 0 °C and monitored by 500.1 <sup>1</sup>H NMR spectroscopy by recording a spectrum every 30 minutes for 16 h. The relative

concentration of product in the reaction at each time point was calculated by deconvolution of the spectra and a concentration vs. time profile constructed for the reactions.



**Figure 4.13:** Concentration vs. time profiles for the production of mono functional *trans*-55 (●) in the presence of (i) 10 mol% bifunctional nitrone 120 and (ii) 50 mol% bifunctional nitrone 120. The undoped reaction between mono functional nitrone 52 and maleimide 34 is shown for comparison (red line).

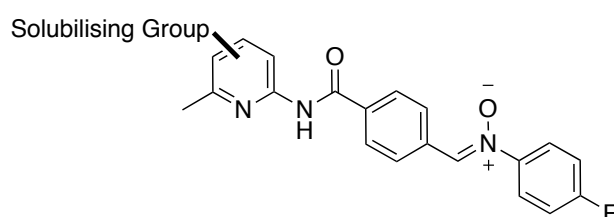
When 10 mol% of the bifunctional nitrone 120 was added to the system, there was no observed change in the rate of production of *trans*-55 (Figure 4.13 (i)). In retrospect, we might have expected this result as addition of 10 mol% nitrone 120 is equal to a concentration of 1 mM which we calculated to be the  $K_d$  of the acid with an amidopyridine, Section 4.6.1. Therefore below this concentration, reagents will exist in isolation rather than associated together. Without efficient recognition processes, there will be no rapid formation of the bifunctional template which would then catalyse the mono functional system. Instead the mono functional system would behave in its native state as was seen with the doped data (Figure 4.13 (i) red circles) being equal to the undoped data, red line.

To overcome this problem, the concentration of doped bifunctional nitrone was increased to 50 mol% which was equal to 5 mM, a concentration above the  $K_d$  and therefore we expected to observe the desired recognition mediated behaviour. The concentration vs. time profile for this reaction (Figure 4.13 (ii)), showed a more encouraging result with an increased rate of production of *trans*-55 in the early stages of reaction. The rate of reaction then decreased after roughly 30,000 s which was as a result of the precipitation of the template from solution.

The poor solubility of the reagents used thus far was a constant struggle within the project. Bifunctional nitron **120** itself was poorly soluble in  $\text{CDCl}_3$  with a concentration of 5 mM the maximum obtainable. The subsequent template formed upon reaction with maleimide **34** also was poorly soluble in  $\text{CDCl}_3$  and would precipitate from 5 mM solutions if left to stand longer than a few hours. Mono functional nitrones were more soluble however the products of their cycloadditions were poorly soluble again precipitating if left to stand for a prolonged period of time. Bifunctional nitron **120** was readily soluble in more polar solvents such as THF and DMSO however using these solvents for the reactions was unfeasible as recognition between the reagents would be disrupted by solvent interactions. Instead the more desirable option was to modify the structures of the reagents to include groups to increase their solubility in non-polar solvents.

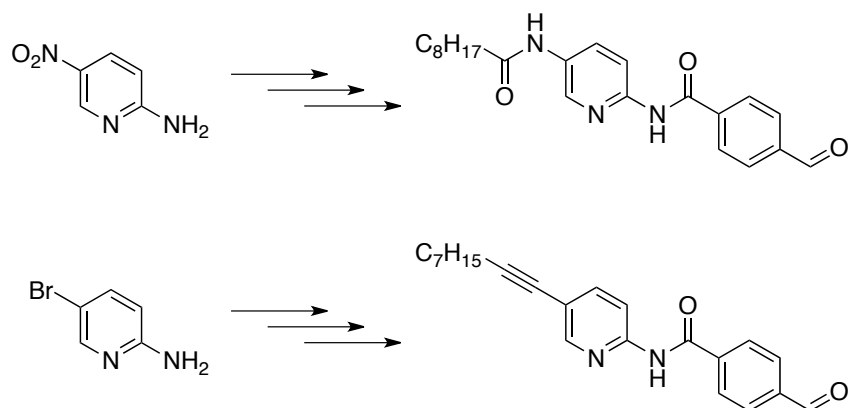
#### 4.9 Increasing the Solubility of Reagents in Non-Polar Solvents

When attempting to improve the solubility of compounds in non-polar solvents it is common to include non-polar alkyl chains to the structure of the compound. Using the mono-functional nitron from the self-replicating system, we looked to identify a position where an alkyl chain could be added through a synthetically facile route and not interfere sterically with the recognition or reactive site (Figure 4.14). The most attractive strategy identified was to add the alkyl chain to the aminopyridine ring.



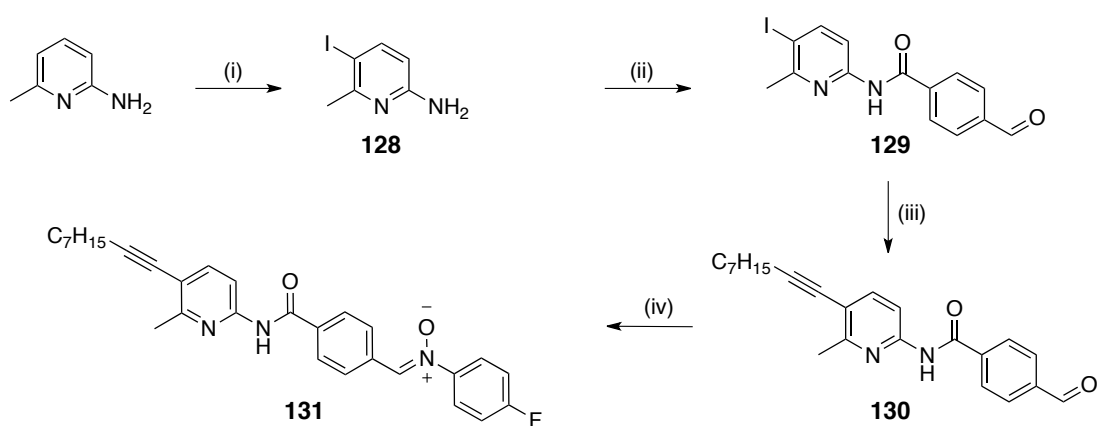
**Figure 4.14:** Functionalisation of the pyridine ring with a group to increase solubility.

The first attempted strategy (Scheme 4.12) was to couple 2-amino-5-nitropyridine with 4-formylbenzoyl chloride to form the aldehyde before reduction of the nitro and subsequent second amide coupling with nanonyl chloride to add a long chain. Synthetically the initial amide coupling was troublesome as the amine was highly deactivated and required sodium hydride deprotonation and elevated reaction temperature to obtain only poor yields and thus this route was rejected.



**Scheme 4.12:** Synthetic routes to the precursor aldehyde for the nitrone with possible groups to increase solubility of the subsequent nitrone and cycloadducts.

Concurrently, a second strategy was investigated which would have the advantage of installing the alkyl chain without adding another polar amide site which could possibly have interfered with recognition processes. The second strategy started with amide coupling of 2-amino-5-bromopyridine to form the aldehyde before a Sonogashira reaction with nonyne to install the alkyl chain. This route was once again unsuccessful producing low yields of products, this time as a consequence of the low reactivity of the Sonogashira reaction. The product itself was encouragingly more soluble in chloroform and as a result the synthetic route was altered to try to increase the reactivity of the Sonogashira reaction. One of the most obvious starting points when attempting to increase the reactivity of a Pd catalysed reaction would be to change the alkyl halide from a bromide to an iodide group to promote oxidative insertion.



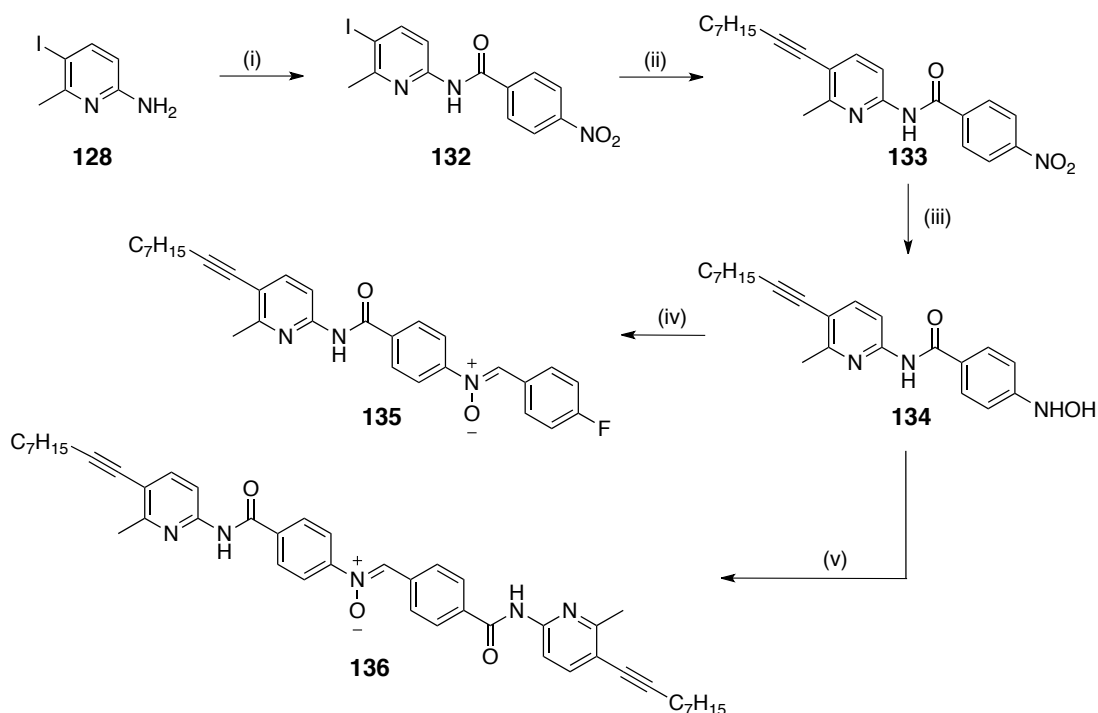
**Scheme 4.13:** Reagents and conditions (i) Periodic acid,  $\text{H}_2\text{SO}_4$ , AcOH,  $60^\circ\text{C}$ , 20 h, 76%, (ii) 4-formylbenzoyl chloride,  $\text{CH}_2\text{Cl}_2$ ,  $\text{NEt}_3$ ,  $0^\circ\text{C} \rightarrow \text{RT}$ , 16 h, 98%, (iii) nonyne,  $\text{Pd}(\text{P}(\text{Ph})_3)_2\text{Cl}_2$ , CuI,  $\text{NEt}_3$ ,  $50^\circ\text{C}$ , 16 h, 90%, (iv) *N*-(4-fluorophenyl)hydroxylamine **46**, EtOH, RT, 16 h, 56%.

Instead of purchasing the relatively expensive 2-amino-5-iodopyridine **128**, a procedure was identified<sup>198</sup> which would iodinate 2-amino-6-methylpyridine and was



amenable to large scale reactions (Scheme 4.13). The iodination was achieved using periodic acid in concentrated sulfuric acid and acetic acid which furnished 2-amino-5-iodo-6-methylpyridine **128** in good yields. An amide coupling performed with the iodinated aminopyridine with 4-formylbenzoyl chloride under standard conditions produced the iodo aldehyde **129** in a quantitative yield before a Sonogashira reaction inserted nonyne under mild conditions and in good yields. The aldehyde was then condensed with hydroxylamine **46** as before to form the target nitron **131** for studying the self-replicating system.

The alkyne functionalised nitron **131** was readily soluble in  $\text{CDCl}_3$  compared to its undecorated predecessor **52** and as a result nonyne was chosen again to be added to **121**, the *reverse* nitron, to ensure the association constants of each recognition site were consistent as the alkyne group would undoubtedly alter their binding ability. This also had the added benefit of decorating the diamidopyridine nitron **120** with two alkyl chains thus having the maximum effect when increasing solubility in  $\text{CDCl}_3$ .



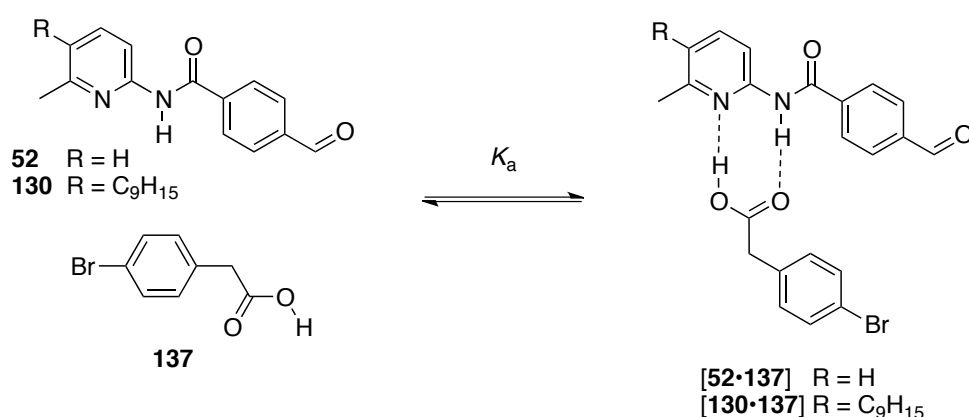
**Scheme 4.14:** Reagents and conditions (i) 4-nitrobenzoyl chloride,  $\text{NEt}_3$ ,  $\text{CH}_2\text{Cl}_2$ ,  $0^\circ\text{C} \rightarrow \text{RT}$ , 16 h, 62% (ii) nonyne,  $\text{Pd}(\text{P}(\text{Ph})_3)_2\text{Cl}_2$ ,  $\text{CuI}$ ,  $\text{NEt}_3$ ,  $50^\circ\text{C}$ , 16 h 66% (iii)  $\text{NH}_2\text{NH}_2 \cdot \text{H}_2\text{O}$ ,  $\text{Rh/C}$ , THF, 0.5 h, RT, 93% (iv) 4-fluorobenzaldehyde, EtOH, RT, 16 h, 80% (v) aldehyde **130**, EtOH, RT, 16 h, 50%.

Nitron **135**, the solubilised target for investigating the  $[\mathbf{A} \cdot \mathbf{B}]$  reaction was synthesised (Scheme 4.14) starting from the previously prepared iodo aminopyridine

**128** which was coupled with 4-nitrobenzoyl chloride to form the iodoamide **132**. The Sonogashira reaction was again employed to insert the nonyne group which greatly improved the solubility of the nitroamide **133**. Reduction using the standard rhodium on carbon with hydrazine hydrate procedure readily formed hydroxylamine **134** which was found to be more stable than previously the prepared hydroxylamines and was able to be stored in the freezer without degradation for several months. Condensation of **134** with 4-fluorobenzaldehyde yielded nitrone **135** for investigating the [A•B] complex channel reaction and condensation of **134** with the recognition bearing aldehyde **130** yielded the bisamidopyridine nitrone **136**. Gratifyingly, nitrone **136** was found to have greatly increased solubility in CDCl<sub>3</sub>.

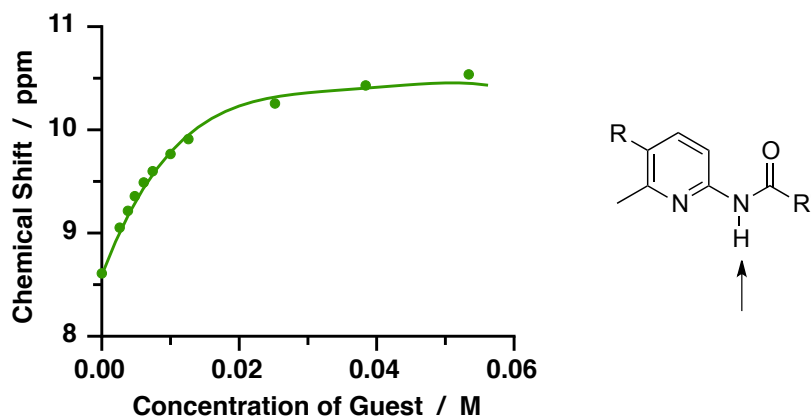
#### 4.10 Determination of Association Constant

By introducing the electron withdrawing alkyne group on to the pyridine ring we expected to observe a change in association constant between the amidopyridine and phenylacetic acid. Accordingly the new association constant was determined by the 300.1 MHz <sup>1</sup>H NMR spectroscopy by titrating model compounds for the phenyl acetic acid into solutions of amidopyridine. The phenyl acetic acid group on the maleimide was represented by 4-bromophenylacetic acid **137** as it was more soluble in CDCl<sub>3</sub> and permitted the high concentrations which were necessary. The amidopyridine recognition bearing aldehyde **130** was used as a model compound for the nitrones as a result of its structural and electronic similarity.



**Scheme 4.15:** Association of model compounds, amidopyridine bearing aldehyde with 4-bromophenylacetic acid **137**, was used to calculate the association constant  $K_a$  in CDCl<sub>3</sub> at various temperatures.

The titration experiments were carried out by preparing ten samples of aldehyde **130** at 10 mM in  $\text{CDCl}_3$  and into each of these, a varying concentration of 4-bromophenylacetic acid **137**, from 1 mM up to 50 mM was added. 300.1 MHz  $^1\text{H}$  NMR spectroscopy was then used to obtain a proton spectrum of each sample. The sample chamber of the NMR spectrometer was cryostatically cooled to the desired temperature, 0 °C or 10 °C, and each sample tube allowed to equilibrate in the sample chamber for 30 minutes before recording a proton spectrum.



**Figure 4.15:** Example of a concentration vs. chemical shift profile for the change in resonance signal for the  $N$ -H proton with concentration of the guest. Experimental data points shown as green circles and the calculated 1:1 binding isotherm shown as a solid green line.

The chemical shift of three signals, arising from the amide  $N$ -H and two pyridine ring protons, were individually plotted against the concentration of guest added, such as demonstrated for the  $N$ -H proton in Figure 4.15. Each profile showed a characteristic change of chemical shift which allowed a 1:1 binding isotherm to be fitted and taking the average value obtained from monitoring the three protons revealed the association constant at various temperatures.

**Table 4.4:** Association constants determined using the model compounds at various temperatures.

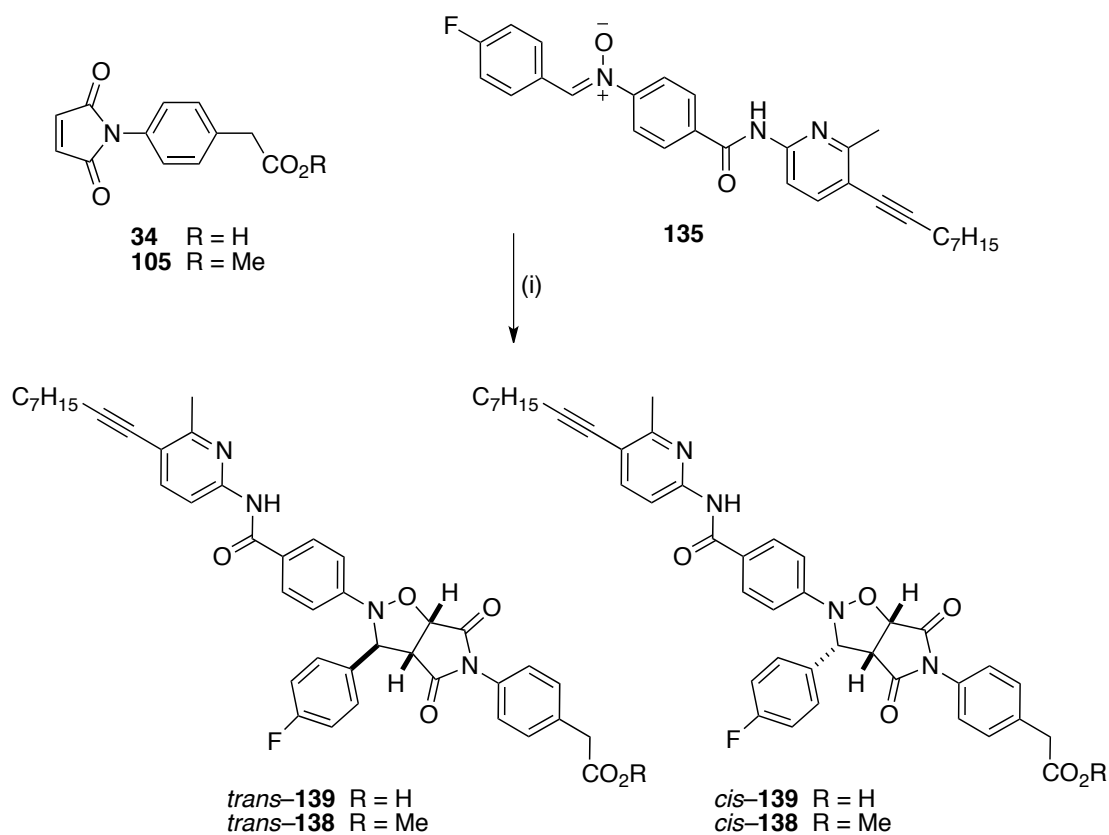
	$K_a / \text{M}^{-1}$ $\text{R} = \text{H}$	$K_a / \text{M}^{-1}$ $\text{R} = \text{C}_9\text{H}_{15}$
0 °C	1020	580
10 °C	680	370

Comparison of the  $K_a$  with and without the nonyl substituent highlights that introducing an alkyne group causes a reduction of the  $K_a$ . This reduction in  $K_a$  was expected as an alkyne group is electron withdrawing which, in this case, would

withdraw electron density away from the pyridine ring leaving it more positive and as a result it will be less able to accept a H-bond from the acid in comparison to when R = H. Crucially the  $K_a$  does not drop enough to cause any concern that association would not take place at concentrations worked with thus far. With a  $K_a$  of  $580 \text{ Mol}^{-1}$ , the  $K_d$  ( $1/K_a$ ) will be  $1.7 \text{ mM}$ , therefore at concentrations below this value the reagents would exist mostly in their isolated form.

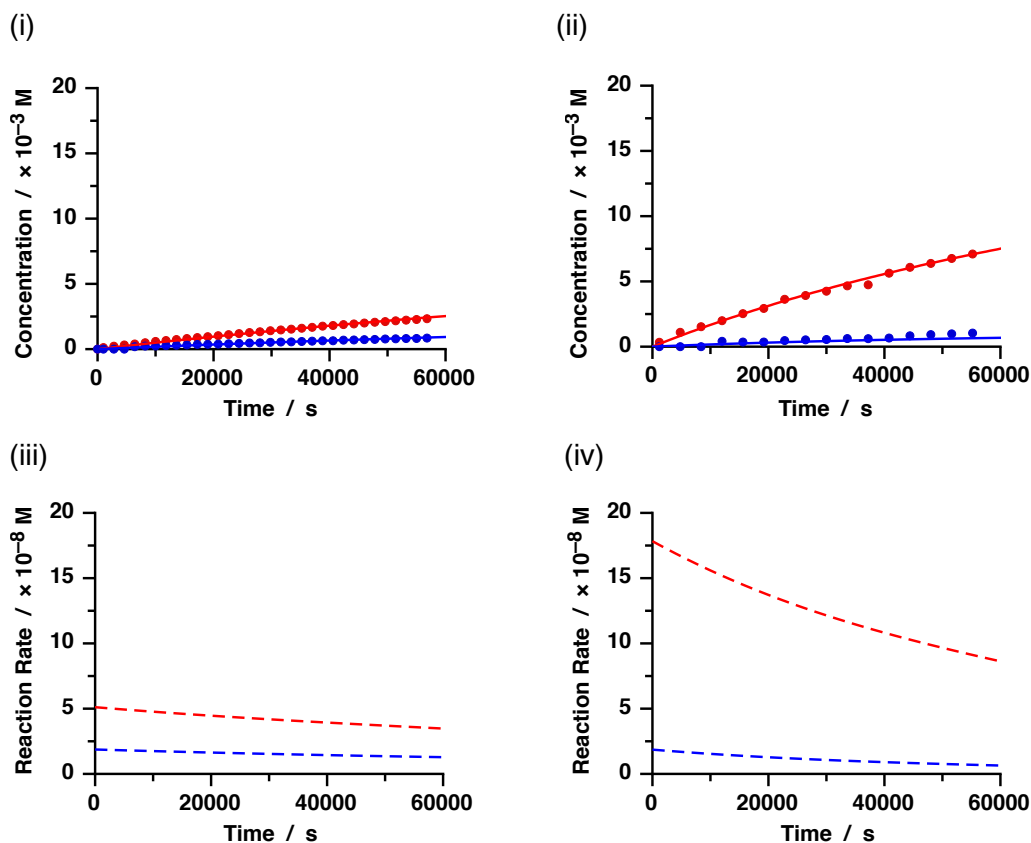
#### 4.11 Kinetic behaviour of [A•B] system

With a new family of nitrones synthesised, the kinetic behaviour of each was re-evaluated to investigate for any change in reactivity or selectivity arising from the introduction of a nonyl group and subsequent lowering of the  $K_a$  and any associated steric effects. With the advantage of increased solubility of the reagents, all of the reactions were now carried out at the increased concentration of  $20 \text{ mM}$  to promote recognition mediated process.



**Scheme 4.16:** Reagents and conditions (i)  $20 \text{ mM}$ ,  $\text{CDCl}_3$ ,  $0^\circ\text{C}$ ,  $16 \text{ h}$ .

The kinetic behaviour of the **[A•B]** system was investigated as before with the bimolecular reaction of the recognition disabled system characterised initially by monitoring the reaction of nitrone **135** with maleimide ester **105** at 20 mM in CDCl<sub>3</sub> at 0 °C. The reaction was monitored by 500.1 MHz <sup>1</sup>H NMR spectroscopy with a proton spectrum recorded every 30 minutes for 16 h. The relative concentration of each species in solution could be determined at each time point by deconvolution of the appropriate resonances and a concentration vs. time profile constructed.



**Figure 4.16:** Concentration vs. time profile for (i) the recognition inhibited control reaction of nitrone **135** with maleimide **105** and (ii) the recognition mediated reaction of nitrone **135** with maleimide **34**. Experimental data shown as points with *trans* products represented as (●) and *cis* products represented as (●). The best fit of the data to a kinetic model for an **[A•B]** complex reaction is represented as lines with *trans* products represented by (—) and *cis* products by (—). (iii) Rate vs. time profile for the recognition inhibited control reaction of nitrone **135** with maleimide **34**. (iv) Rate vs. time profile for the recognition mediated reaction of nitrone **135** with maleimide **34**. The rate of production of the *trans* cycloadduct shown as a red dashed line and the *cis* cycloadduct as a blue dashed line. Reaction conditions: 20 mM, CDCl<sub>3</sub>, 0 °C, 16 h.

**Table 4.5:** Reaction rate constant data calculated from kinetic simulation and fitting using SimFit.

	$k_{\text{trans}}$	$k_{\text{cis}}$
Bimolecular Rate Constant / $\times 10^{-4} \text{ M}^{-1} \text{ s}^{-1}$	1.28	0.47
$k_{\text{binary}} / \times 10^{-4} \text{ s}^{-1}$	0.086	—
<i>EM</i> Binary Complex / M	0.067	—

The control reaction (Figure 4.16 (i)) proceeds slowly as before with 2.3 mM *trans*-**138** and 0.5 mM *cis*-**138** representing an overall conversion of 15% after 16 h with a *trans* : *cis* ratio of 2.8:1. Kinetic simulation and fitting was performed using SimFit to extract reaction rate information and the reaction rate constants  $k_{\text{trans}}$  and  $k_{\text{cis}}$  were determined to be  $1.28 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  and  $0.47 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  respectively. A rate vs. time profile was produced for the reaction (Figure 4.16 (iii)) which showed the expected profile for a bimolecular reaction.

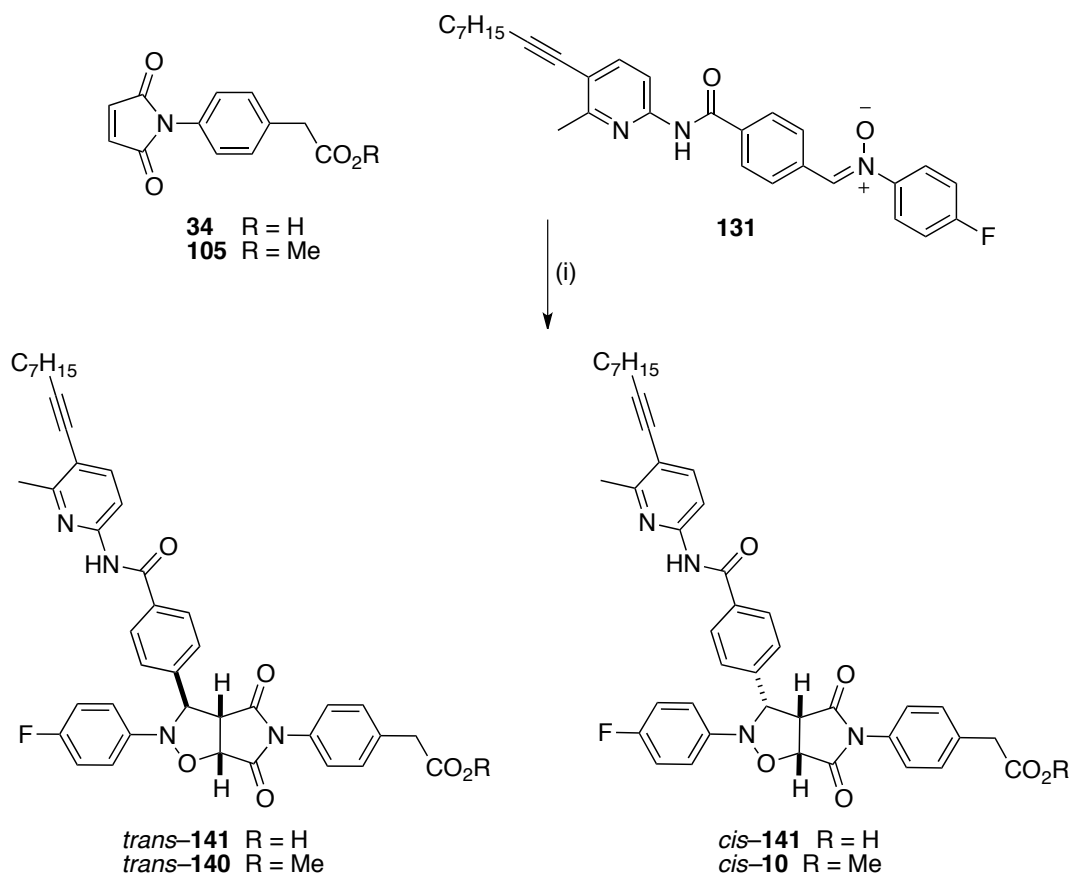
The recognition mediated behaviour was then observed by monitoring the reaction between nitrone **138** and maleimide **34** using the same methods under the same conditions. The concentration vs. time profile (Figure 4.16 (ii)) clearly shows an increase in production of the *trans* cycloadduct with 7.1 mM of *trans*-**136** produced along with 1.04 mM of *cis*-**136** representing an overall conversion of 41% and a *trans* : *cis* ratio of 6.8:1. Kinetic simulation and fitting was performed once again using SimFit and revealed the reaction rate constant within the binary complex  $k_{\text{binary}}$  to be  $0.086 \times 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  from which the effective molarity for the reaction to form the *trans* cycloadduct in the binary complex to be 0.086 M once again a modest increase compared to the actual reaction concentration of 0.020 M indicating that the effect of the recognition is to increase the effective concentration of reagents  $3.3 \times$  greater than they actually are.

It is once again apparent that the presence of the recognition site serves to arrange the reagents into a reactive  $[\mathbf{A}\cdot\mathbf{B}]$  complex holding their orientation of approach to promote the formation of the *trans* cycloadduct engineering a modest increase in rate of reaction.

#### 4.12 Kinetic behaviour of self-replicating system

Kinetic analysis of the mono functional self-replicating system was performed to investigate any effect as a consequence of addition of the nonyl group on the amidopyridine binding site (Scheme 4.17).

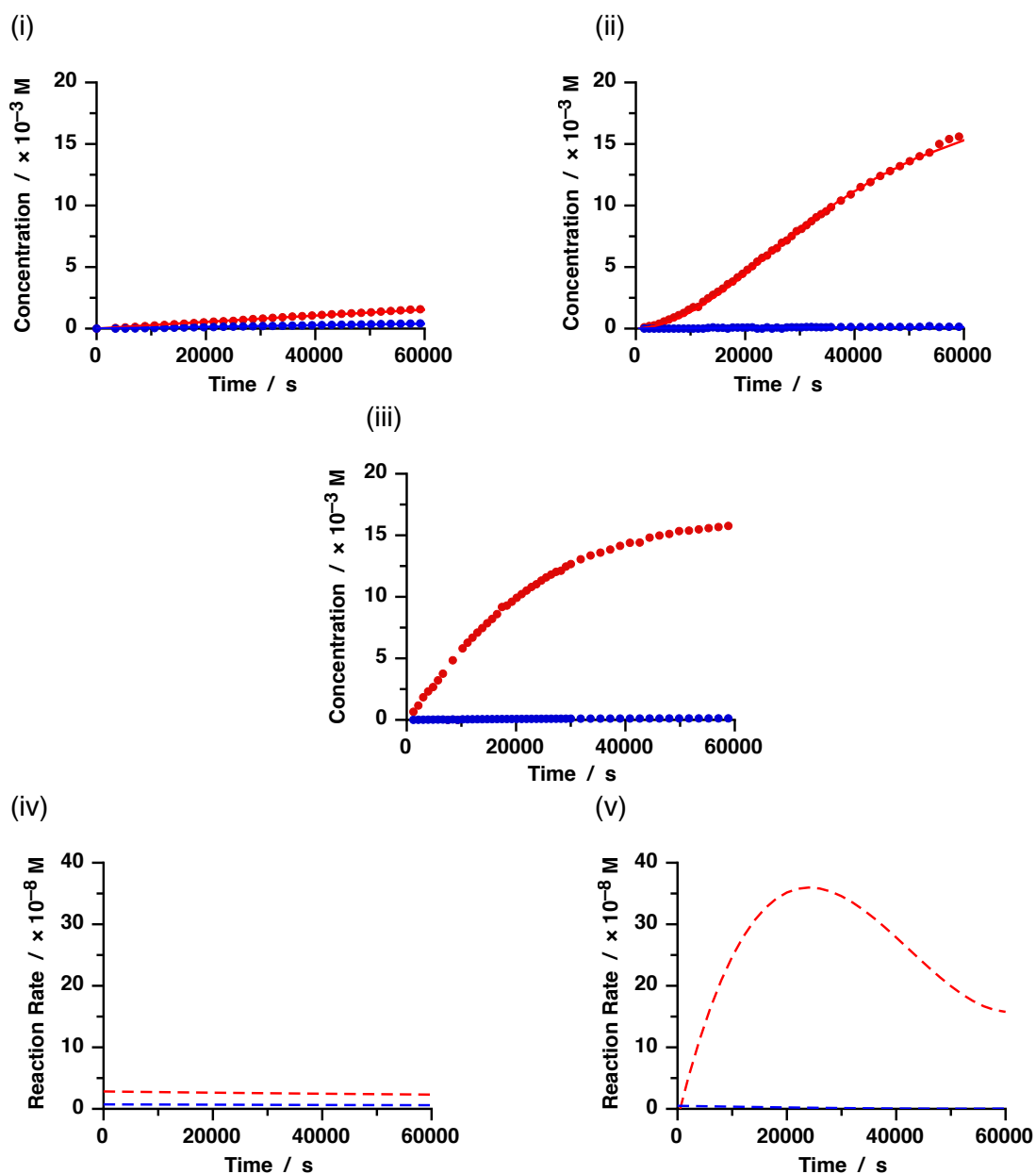
In order to characterise the background bimolecular reaction, the recognition inhibited control reaction was performed under the same reaction conditions (20 mM,  $\text{CDCl}_3$ ,  $0^\circ\text{C}$ , 16 h) using the maleimide ester **105** and nitrone **131** and monitoring by



**Scheme 4.17:** Reagents and conditions (i) 20 mM, CDCl<sub>3</sub>, 0 °C, 16 h.

500.1 MHz <sup>1</sup>H NMR spectroscopy for 16 h. The relative concentration of products were calculated by deconvolution and a concentration vs. time profile created (Figure 4.17 (i)) which showed the reaction had proceeded slowly with 1.57 mM *trans*-**140** and 0.42 m *cis*-**140** produced representing an overall conversion of 10% and a *trans* : *cis* ratio of 3.7:1. Kinetic fitting and simulation using SimFit revealed the rate constants  $k_{trans}$  and  $k_{cis}$  to be  $7.07 \times 10^{-5} \text{ M}^{-1}\text{s}^{-1}$  and  $1.81 \times 10^{-5} \text{ M}^{-1}\text{s}^{-1}$  respectfully. A rate vs. time profile was produced for the reaction (Figure 4.17 (iv)) which once again showed the expected profile for a simple bimolecular reaction.

The recognition mediated behaviour was next observed by performing and monitoring the reaction of nitron **131** with maleimide **34** under the same conditions. The concentration vs. time profile (Figure 4.17 (ii)) clearly showed an increase in production of the *trans*-**141** which reached 15.8 mM along with 0.93 mM *cis*-**141** representing an overall conversion of 84% and a *trans* : *cis* ratio of 17:1. The concentration vs. time profile revealed a sigmoidal shape characteristic of an autocatalytic self-replicator. A rate vs. time profile was also produced for the reaction



**Figure 4.17:** Concentration vs. time profiles of (i) Recognition disabled control reaction, (ii) Recognition mediated reaction, (iii) 10 mol% template doped recognition mediated reaction Experimental data shown as points with *trans* products represented as (●) and *cis* products represented as (●). The best fit of the data to a kinetic model for a self replicating reaction is represented as lines with *trans* products represented by (—) and *cis* products by (—). (iv) Rate vs. time profile for the recognition inhibited control reaction of nitrone **131** with maleimide **105**. (v) Rate vs. time profile for the recognition mediated reaction of nitrone **131** with maleimide **34**. The rate of production of the *trans* cycloadduct shown as a red dashed line and the *cis* cycloadduct as a blue dashed line. Reaction conditions: 20 mM,  $\text{CDCl}_3$ , 0 °C, 16 h.

(Figure 4.17 (v)) which revealed that the rate of reaction increases as the reaction time progresses characteristic of an autocatalytic reaction. Kinetic simulation and fitting using SimFit calculated the reaction rate on the ternary complex  $k_{\text{ternary}}$  to be  $79 \times 10^{-5} \text{ M}^{-1}\text{s}^{-1}$  revealing the *EM* of the ternary reaction to be 11.2 M. The template



duplex was reduced to  $3.86 \times 10^5 \text{ M}^{-1}$  reflecting the weaker individual recognition interactions. The connection *EM* was determined to be 1.1 M with the free energy of connectivity calculated to be  $0.3 \text{ kJmol}^{-1}$  demonstrating a positively cooperative mode of binding the second recognition unit.

**Table 4.6:** Reaction rate constant data calculated from kinetic simulation and fitting using SimFit.

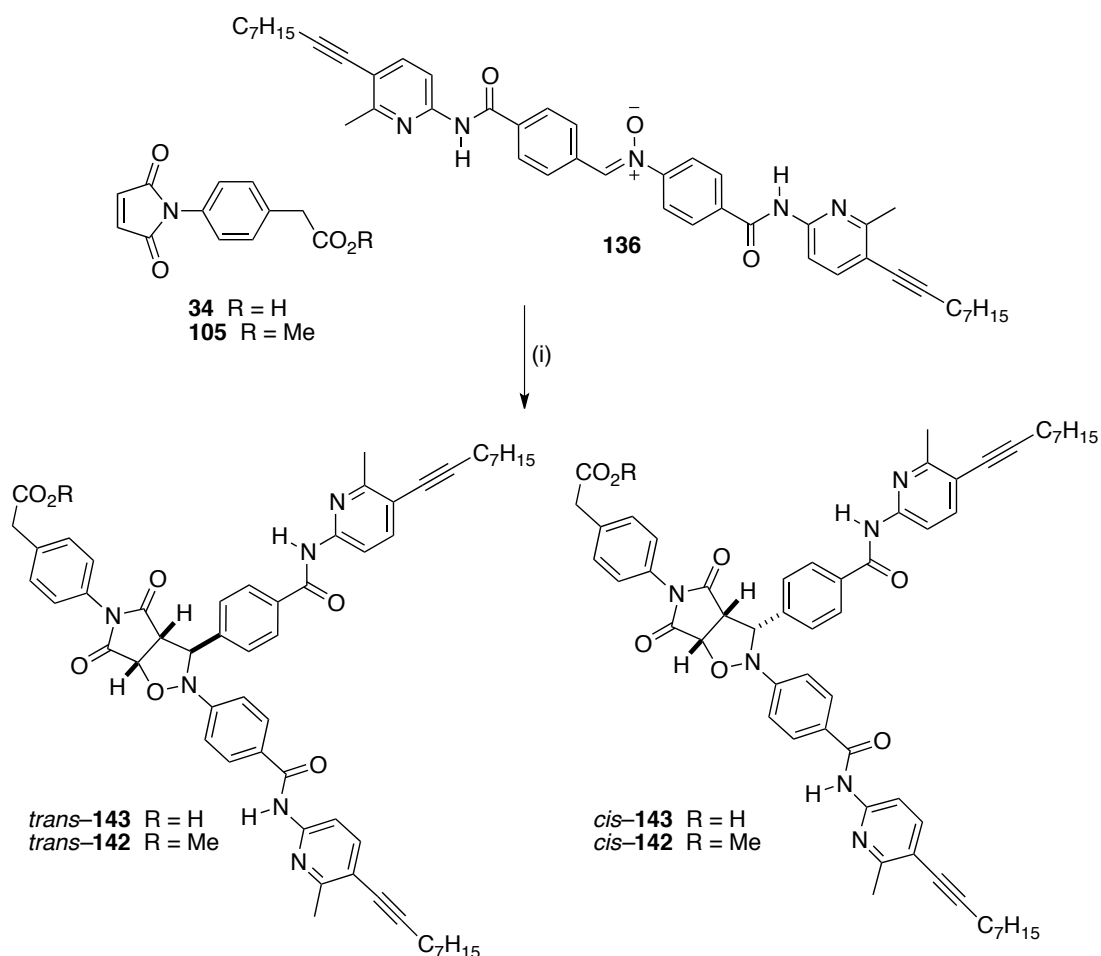
	$k_{\text{trans}}$	$k_{\text{cis}}$
Bimolecular Rate Constant / $\times 10^{-5} \text{ M}^{-1} \text{ s}^{-1}$	7.07	1.81
$k_{\text{ternary}} / \times 10^{-5} \text{ s}^{-1}$	79	—
<i>EM</i> Ternary Complex / M	11.2	—
$K_{\text{Duplex}} / \times 10^5 \text{ M}^{-1}$	3.86	—
Connection <i>EM</i> / M ( $\Delta G^\ddagger / \text{kJmol}^{-1}$ )	1.1 (0.3)	—

The template directed nature of the reaction was confirmed by demonstrating that preformed *trans*-**141** could catalyse the reaction. A template doped reaction was performed by adding 10 mol% *trans*-**141** to a fresh set of reagents for the recognition mediated reaction which was then performed and monitored under the same conditions. The concentration vs. time profile (Figure 4.15 (iii)) showed that the reaction proceeded at the maximum rate upon initiation, with no lag period, demonstrating that *trans*-**141** was acting as a catalyst for its own formation.

The results from the characterisation of the self-replicating system demonstrated that the introduction of the nonyl group on the amidopyridine binding site have not fundamentally altered the ability of the reaction to perform in an autocatalytic manner. There was however subtle changes as the rate within the ternary complex and the resultant *EM* for the system was determined to be lower than the original system. The template duplex was also determined to be lower on account of the weaker association events used in its dimerisation.

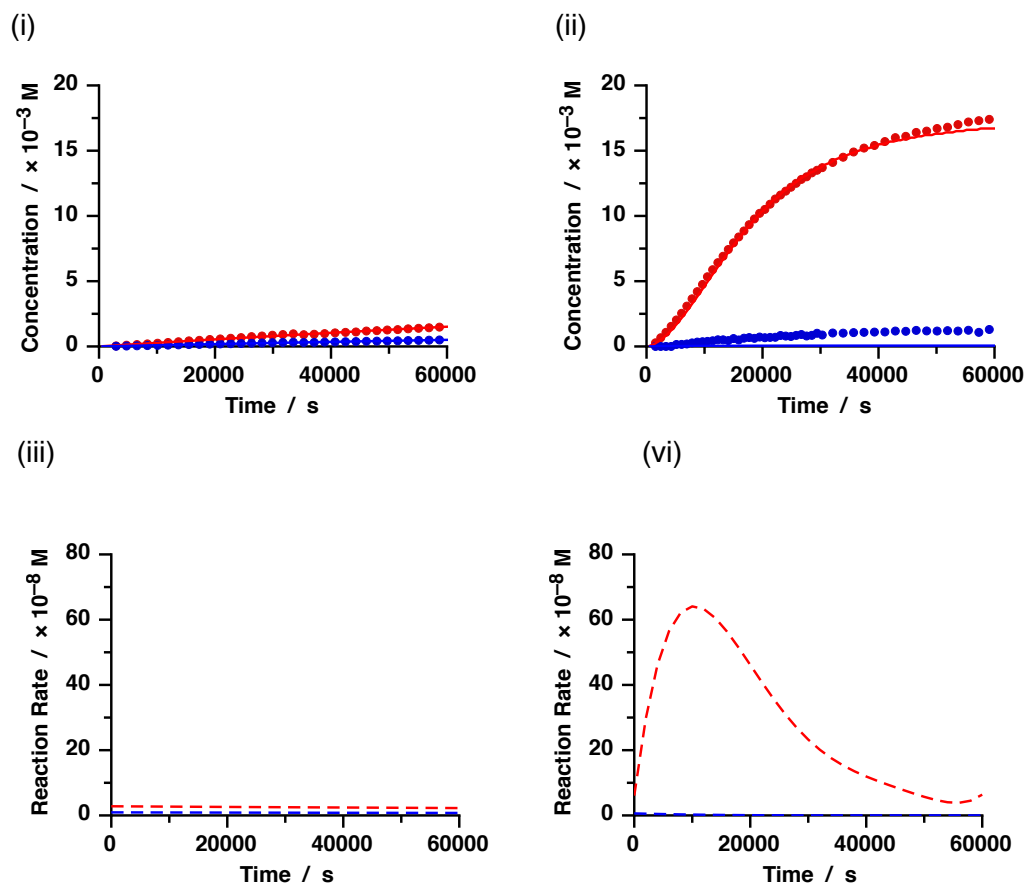
#### 4.13 Kinetic Behaviour of the Bifunctional System

With the kinetic behaviour of each of the two amidopyridine recognition sites characterised in isolation on two mono functional nitrones, we once again looked to investigate the behaviour of a system which combines both into one bifunctional nitron.



**Scheme 4.18:** Reagents and conditions (i) 20 mM, CDCl<sub>3</sub>, 0 °C, 16 h.

Initially, the bimolecular reaction within the system was investigated by performing the recognition inhibited control reaction of nitron **136** with maleimide **105** under the same reaction conditions as before (20 mM, CDCl<sub>3</sub>, 0 °C, 16 h) and monitoring by 500.1 MHz <sup>1</sup>H NMR spectroscopy for 16 h. Once complete, the relative concentration of products were calculated by deconvolution and a concentration vs. time profile created (Figure 4.18 (i)) which showed the reaction had proceeded slowly with 1.48 mM *trans*-**142** and 0.50 mM *cis*-**142** produced representing an overall conversion of 10% and a *trans* : *cis* ratio of 3.0:1. Kinetic fitting and simulation for a bimolecular reaction was performed using SimFit which revealed the rate constants  $k_{trans}$  and  $k_{cis}$  to be  $7.01 \times 10^{-5} \text{ M}^{-1}\text{s}^{-1}$  and  $2.35 \times 10^{-5} \text{ M}^{-1}\text{s}^{-1}$  respectively. A rate vs. time profile was produced for the reaction (Figure 4.18 (iv)) which once again showed the expected profile for a simple bimolecular reaction.



**Figure 4.18:** Concentration vs. time profiles of (i) Recognition disabled control reaction, (ii) Recognition mediated reaction. Experimental data shown as points with *trans* products represented as (●) and *cis* products represented as (●). The best fit of the data to a kinetic model of an [A•B] complex reaction and self-replicating reaction is represented as lines with *trans* products represented by (—) and *cis* products by (—). (iii) Rate vs. time profile for the recognition inhibited control reaction of nitron **136** with maleimide **105**. (iv) Rate vs. time profile for the recognition mediated reaction of nitron **136** with maleimide **34**. The rate of production of the *trans* cycloadduct shown as a red dashed line and the *cis* cycloadduct as a blue dashed line. Reaction conditions: 20 mM, CDCl<sub>3</sub>, 0 °C, 16 h.

**Table 4.6:** Reaction rate constant data calculated from kinetic simulation and fitting using SimFit.

	$k_{trans}$	$k_{cis}$
Bimolecular Rate Constant / $\times 10^{-5} \text{ M}^{-1}\text{s}^{-1}$	7.01	2.35
$k_{ternary} / \times 10^{-5} \text{ s}^{-1}$	612	—
EM Ternary Complex / M	87	—
$K_{Duplex} / \times 10^6 \text{ M}^{-1}$	7.12	—
Connection EM / M ( $\Delta G^{\circ} / \text{kJmol}^{-1}$ )	20.9 (6.9)	—

The effect of recognition in the system was next investigated by performing the recognition mediated reaction between maleimide **34** and nitron **136** under the

same conditions and monitoring as before. In this case, the concentration vs. time profile (Figure 4.18 (ii)) for the reaction revealed the selective formation of the *trans* cycloadduct with 17.4 mM *trans*-**143** produced along with 1.3 mM *cis*-**143** after 16 h. The concentration vs. time profile for the reaction display a sigmoidal shape but with a significantly shorter lag period in comparison to the mono-functional self-replicating system. A rate vs. time graph was once again produced which revealed that the maximum rate of reaction was reached after ~11,000 s, which is significantly earlier than the mono-functional replicator which took ~22,000 s to reach the maximum rate.

Kinetic simulation and fitting of the experimental results to a kinetic model combining the properties of an [A•B] complex channel reaction with a self-replicating reaction as before (Scheme 4.10) was undertaken using SimFit. The rate constant for reaction in the ternary complex,  $k_{\text{ternary}}$ , was determined to be  $612 \times 10^{-5} \text{ s}^{-1}$  and the template duplex strength,  $K_{\text{Duplex}}$ , was found to be  $7.12 \times 10^6 \text{ M}^{-1}$ . From these results, the *EM* of the reaction was determined to be 87 M representing that recognition processes were significantly enhancing the reaction. The connection *EM* was also determined to be 20.9 M from which the connection free energy was determined to be 6.9 kJmol<sup>-1</sup> revealing a positively cooperative binding mechanism.

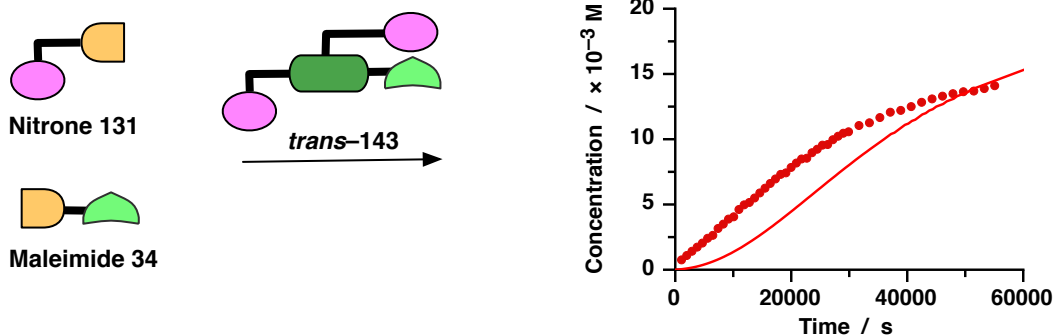
#### 4.13.1 Further Investigations and Characterisation of Origins of Bifunctional Behaviour

To demonstrate that the bifunctional system is able to take advantage of the autocatalytic cycle, preformed bifunctional *trans*-**143** was added to a fresh set of reagents for the recognition mediated mono-functional self-replicating system. 10 mol% *trans*-**143** was added to a reaction of nitron **131** with maleimide **34** under the same conditions as above and the reaction monitored once more by 500.1 MHz <sup>1</sup>H NMR spectroscopy for 16 h after which the spectra were deconvoluted to produce a concentration vs. time profile for the reaction (Figure 4.19 (i)).

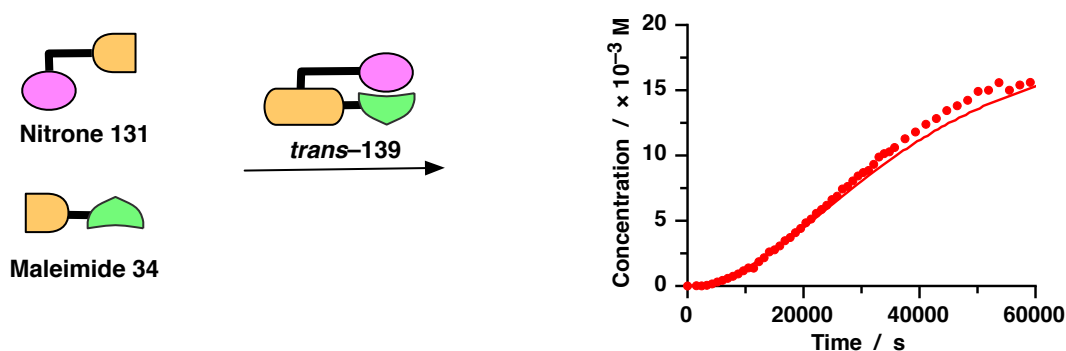
The concentration vs. time profile for the reaction revealed that the presence of *trans*-**143** in the reaction had the effect of removing the lag period in the native reaction and allowing the self-replicating reaction of nitron **131** with maleimide **34** to proceed at the maximum autocatalytic rate upon initiation. This result demonstrates

that the mono-functional reagents are able to be organised on the bifunctional template into a reactive conformation which enhances the production of the mono-functional *trans*-141 cycloadduct confirming that the bifunctional system is able to participate in the autocatalytic channel.

(i)



(ii)



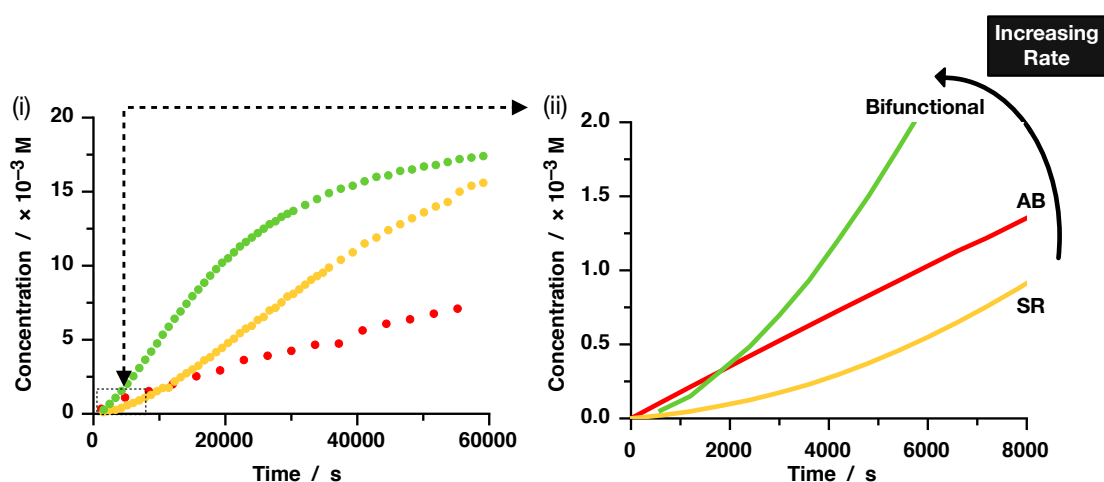
**Figure 4.19:** (i) Reaction of mono-functional nitron **131** with maleimide **34** in the presence of 10 mol% bifunctional template *trans*-143. Concentration vs. time profile for this reaction shown to the right with filled red circles representing the production of *trans*-141 in the doped reaction and the red line representing the production of *trans*-141 in the native reaction of nitron **131** with maleimide **34** with no doping. (ii) Reaction of mono-functional nitron **131** with maleimide **34** in the presence of 10 mol% mono-functional template *trans*-139. Concentration vs. time profile for this reaction shown on the right with filled circles representing the production of *trans*-141 in the doped reaction with the red line representing the production of *trans*-141 from the native reaction of nitron **131** with maleimide **34** with no doping.

To demonstrate that this catalytic effect was a result of the recognition site from the self-replicating side of the bifunctional template and the not [A•B] side, a further control reaction was performed (Figure 4.19 (ii)). Pre-synthesised mono-functional *trans*-139, from the [A•B] system, was added to the reagents for the mono functional self-replicating system. We would expect that this template would have no catalytic effect and no change in the production of mono functional *trans*-141 was anticipated.

The reaction was performed by addition of 10 mol% *trans*-**139** to nitrone **131** and maleimide **34**, the reaction performed under the same conditions and monitored by the same method as before to construct a concentration vs. time profile (Figure 4.19 (ii)). The results demonstrated that there was no noticeable difference in the rate of production of the mono-functional *trans*-**141** with or without *trans*-**141** present. These results demonstrated the product of the mono functional [A•B] complex channel was not able to catalyse the reaction of the mono functional self-replicating channel.

#### 4.14 Comparison of Efficiencies of Functionalities

Direct comparison of the mono-functional systems with the bifunctional system can now be made as the reactions were all performed at the same concentration. Comparison of the rate of production of the *trans* cycloadducts (Figure 4.18 (i)) demonstrated that the bifunctional system (green) is the most efficient producer of *trans* product followed by the mono-functional self-replicating system (orange) and then the mono functional [A•B] complex system (red). Upon closer inspection of the initial 2 h (Figure 4.18 (ii)) in the initial stages of the reaction, the mono-functional [A•B] complex channel was the most efficient at producing *trans* cycloadduct with the mono-functional self-replicating the least efficient as expected. The initial rate of reaction of the bifunctional system was similar to the [A•B] complex reaction rate which allowed the production of the bifunctional template to occur more rapidly than



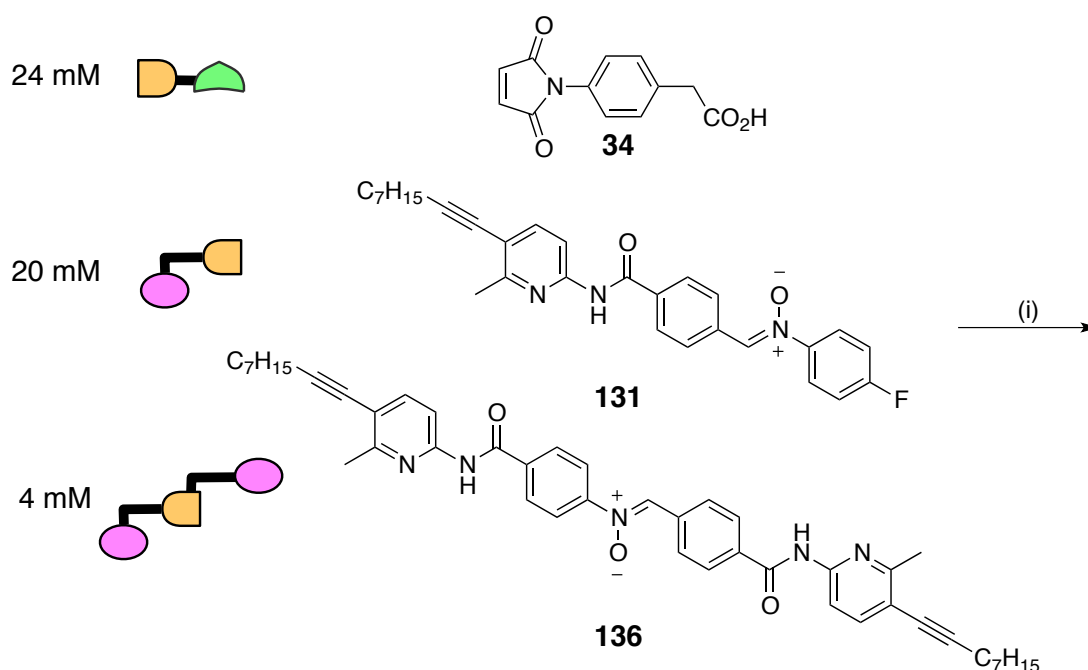
**Figure 4.20:** Concentration vs. time profiles of (i) the production of *trans* cycloadducts through the three recognition mediated behaviours and (ii) an enlargement of the initial 2 h showing that the initial rate of reaction of the bifunctional nitrone (●) proceeds similarly to the [A•B] complex channel (●) and both are more rapid than the self-replicating channel (●).

the mono-functional self-replicating reaction. The rate of the bifunctional system then increases with time significantly earlier than the mono-functional self-replicating system as the autocatalytic cycle is initiated earlier in the bifunctional system by template which has been produced taking advantage of the **[A•B]** complex channel.

This observation demonstrates that the bifunctional system is once again able to exploit both **[A•B]** complex channel and autocatalytic channel behaviour to facilitate a more rapid fabrication of the *trans* cycloadduct in comparison to both mono functional systems and thus signifies the successful design of the desired bifunctional system.

#### 4.15 Cross-Catalytic doping Reactions

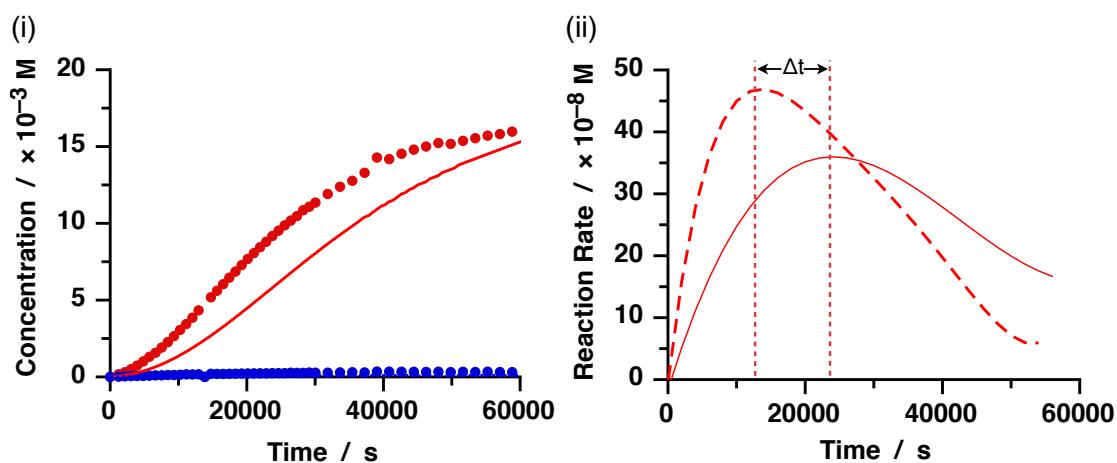
The increased efficiency of the reaction to form the bifunctional template and its cross-catalytic behaviour with the mono-functional system, Section 4.13.1, led us to design a system in which a substoichiometric quantity of bifunctional nitrene **136** could be added to the reagents for the mono functional self-replicating system as previously explored, Section 4.8, Figure 4.11. Bifunctional nitrene **136** will be able to react with maleimide **34** to produce template *trans*-**143** more rapidly than the bimolecular reaction of mono-functional nitrene **131** with maleimide **34** would produce *trans*-**141** through a bimolecular reaction. With bifunctional template present



**Scheme 4.19:** Reagents and conditions (i)  $\text{CDCl}_3$ ,  $0^\circ\text{C}$ , 16 h.

in solution, it shall be able to direct the reaction of mono-functional nitrene **131** with maleimide **34** to produce *trans*-**141**. Since bifunctional template *trans*-**143** is available earlier in the reaction, the production of *trans*-**141** should proceed at an increased rate relative to the native undoped reaction.

The reaction was performed by addition of 20 mol% (4 mM) of bifunctional nitrene **136** to a CDCl<sub>3</sub> solution of mono functional nitrene **131** (20 mM) and enough maleimide **34** to satisfy all reactive sites (24 mM). The reaction as then performed under the same conditions (0 °C, CDCl<sub>3</sub>, 16 h) and monitored by the same method as before.



**Figure 4.21:** (i) Concentration vs. time profile for the production of mono-functional *trans*-**141** (●) and *cis*-**141** (●) from nitrene **131** and maleimide **34** in the presence of bifunctional nitrene **136**. The red line represents the production of *trans*-**141** from nitrene **131** and maleimide **34** undoped at a concentration of 24 mM. (ii) Rate vs. time profile for the doped reaction shown as a dashed red line and for the undoped reaction at 24 mM as a filled red line.

The concentration vs. time profile for the doped reaction (Figure 4.21 (i) filled circles) highlights that the production of mono functional *trans*-**141** was accelerated in the presence of bifunctional nitrene **136** in comparison to the undoped reaction performed at 24 mM (Figure 4.21 (i) red line). For a clearer comparison, the rate vs. time profile for each reaction was also plotted (Figure 4.21 (ii)) which revealed that for the doped reaction, dashed line, the maximum rate is greater than the undoped reaction and is reached in roughly half the time. The effect of this is that the autocatalytic cycle in the mono-functional system is initiated sooner leading to a more rapid template directed production of *trans*-**141** earlier in the reaction with the overall effect that not only the conversion is increased, albeit only slightly from 77%



to 81%, but more significantly the ratio of trans: cis products from the reaction has increased from 17:1 in the undoped system to 52:1 now in the doped system.

In order to prove the results of the doping experiment were as a result of the template directed nature of the bifunctional template and not as a result of the increase in reagent concentration, a further control reaction was performed.

To demonstrate that having an excess of maleimide **34** in comparison to the mono-functional nitrone **131** does not increase the relative rate of reaction, the reaction of 20 mM nitrone **131** with 24 mM maleimide **34** in CDCl<sub>3</sub> was performed at the same temperature for 16 h as before and monitored in the same manner. The results from this reaction revealed no discernible difference for the production or selectivity of *trans*-**141** in comparison to the reaction which had a 1:1 ratio of reagents at 20 mM. and demonstrated that the increase in reaction rate was not consequential of a simple increase in concentration of the reagents.

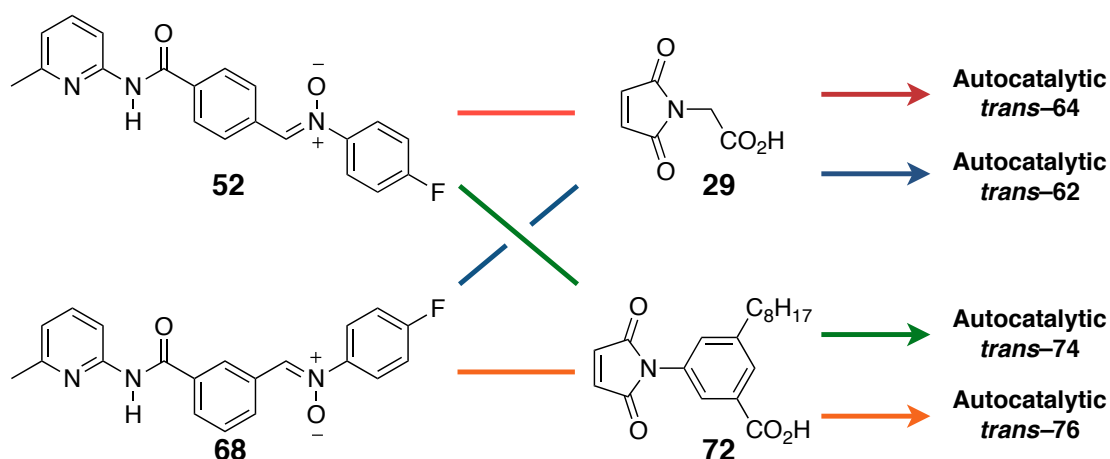
#### 4.16 Conclusions From Bifunctional System

In order to reach our target of creating a rationally designed bifunctional replicating system, capable of incorporating both [A•B] complex behaviour for rapid initial fabrication and autocatalytic self-replicating behaviour for catalytic turn over, we have successfully designed and demonstrated in isolation each of the desired behaviours. First the successful design of an [A•B] complex channel behaviour was demonstrated with the promotion of a 1,3-dipolar cycloaddition reaction between maleimide **34** and nitrone **121**. Autocatalytic self-replicating behaviour had already been successfully established in a previous publication<sup>139</sup> and the substrates for this system were evolved and their behaviour demonstrated once again. Finally a new bifunctional nitrone decorated with two recognition groups, as designed from combination of the above two reaction channels, was produced and demonstrated to utilise both reaction channels to permit rapid fabrication of product which was shown to still be catalytically active. The rapid fabrication of the bifunctional template was then utilised in a cross-catalytic system in which the bifunctional template would be produced *in situ* faster than the mono-functional autocatalytic template and therefore initiate the autocatalysis earlier than would be possible otherwise, therefore giving the replication cycle a 'kickstart'.

#### 4.17 Future Work: Adding Bifunctionality to Existing Replicators.

The successful design of the bifunctional system was achieved by taking an established self-replicating system and decorating the nitronone reagent with a second recognition site which programmed the  $[A \cdot B]$  complex behaviour. Critical to this success was the ability of this  $[A \cdot B]$  complex to promote the formation of the *trans* cycloadduct core, the same stereochemical information at fused core as is produced from the self-replicating autocatalytic cycle. Using this approach, we sought to identify if other self-replicating systems previously identified could be modified to include the bifunctional behaviour.

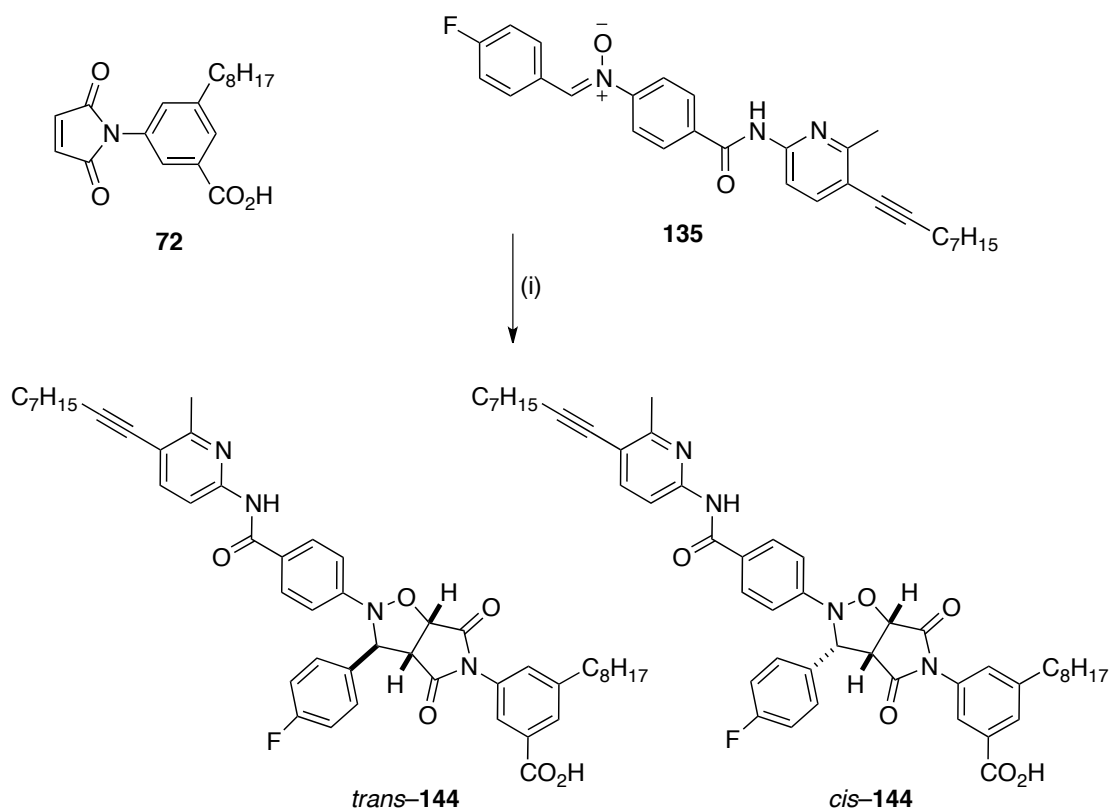
Several autocatalytic self-replicating systems based on the 1,3-dipolar cycloaddition between a nitronone and a maleimide have been discussed previously within this thesis and summarised in Scheme 4.20. Each of the two nitronones in Scheme 4.20 bear the *para* fluoro group on the phenyl ring which is essentially superfluous and has the potential to be replaced with a recognition site which would allow  $[A \cdot B]$  complex behaviour. This approach has already been demonstrated with the reaction of *para* nitronone **52** with maleimide **34** but could equally be applied to the reaction of *para* nitronone **52** with maleimide **29** or **72** or indeed with their reactions with *meta* nitronone **68**.



**Scheme 4.20:** The reaction of each of the nitronones with each of the maleimides represented by a coloured line is autocatalytic for the production of the *trans* cycloadduct.

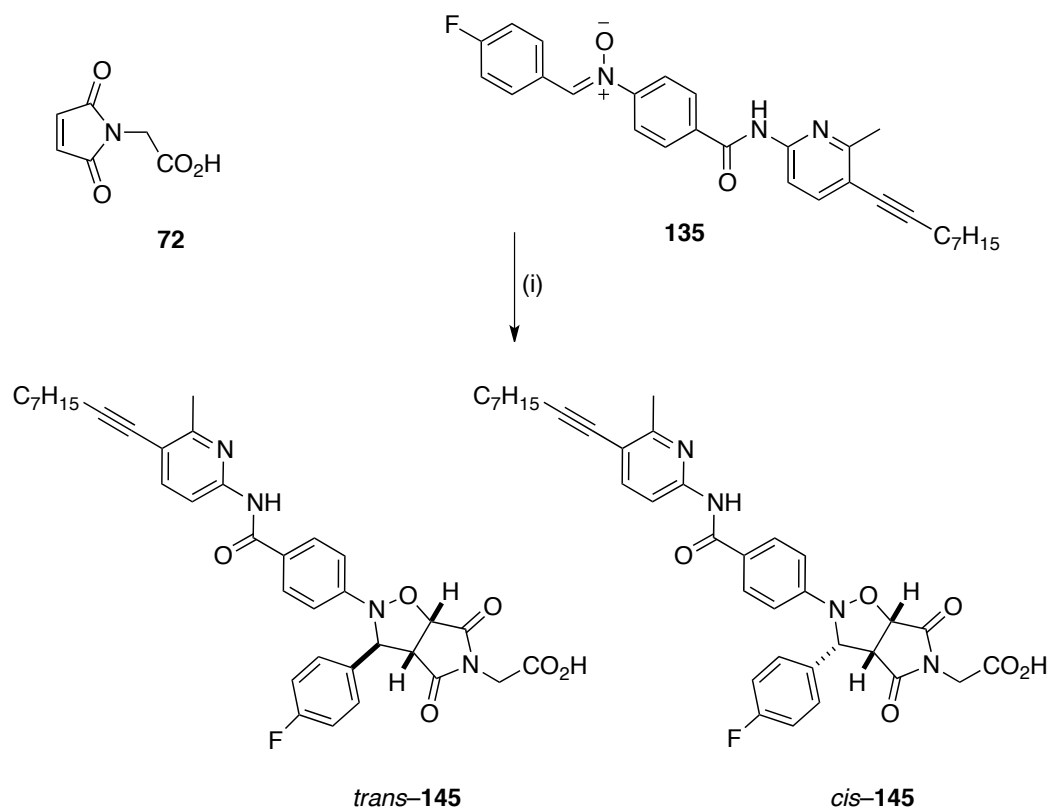
To successfully design bifunctionality into these replicating systems, the additional recognition site must be able to form a reactive  $[A \cdot B]$  complex with the maleimide used which promotes the formation of the *trans* cycloadduct. In order to identify

whether the same *reverse* nitron is able to form a reactive **[A•B]** complex with either of the two maleimides in Scheme 4.20, the reaction of the mono-functional nitron **135** with each of the two maleimide was undertaken to investigate for the desired *trans* selectivity.



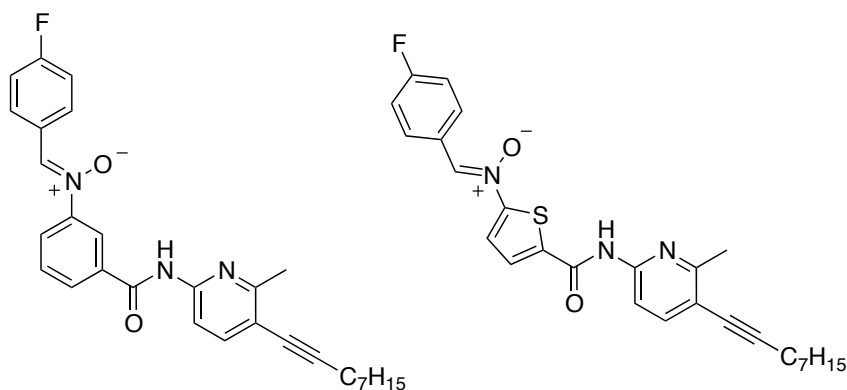
**Scheme 4.21:** Investigation for **[A•B]** complex behaviour of the reverse nitron with a different maleimide. (i) Reagents and conditions 10 °C,  $CDCl_3$ , 16 h

The first reaction investigated was that of nitron **135** with maleimide **72** (Scheme 4.21). A 0.8 mL sample of a 20 mM reaction solution of maleimide **72** with nitron **135** in  $CDCl_3$  was incubated at 10 °C for 16 h before assaying by 376.5 MHz  $^{19}F$  NMR spectroscopy. The relative concentration of each product was then determined by deconvolution of the appropriate resonances arising from the products and starting materials. The results of the reaction was that after 16 h, a 1:1 ratio of *trans* : *cis* cycloadducts were present. In this case, clearly this system did not display the desired *trans* selectivity which would arise from **[A•B]** complex channel behaviour and thus the system was abandoned.



**Scheme 4.22:** Investigation for **[A•B]** complex behaviour of the reverse nitron with a different maleimide. (i) Reagents and conditions 10 °C, CDCl<sub>3</sub>, 16 h

By the same experimental method, the reaction of nitron **135** with maleimide **29** (Scheme 4.22) was performed and assayed after 16 h once again by 376.5 MHz <sup>19</sup>F NMR spectroscopy. The relative concentration of products was determined by deconvolution and in this case revealed that the major product of the reaction was the *cis* cycloadduct. From these results it was clear that in this case that any recognition mediated reactions was not promoting the formation of the *trans* cycloadduct.



**Figure 4.20:** Two possible redesigns of the *reverse nitron* which may be produced in an attempt to realise **[A•B]** complex channel behaviour upon reaction with either maleimide **29** or **72**.

In light of each of these results, the current design of the *reverse* nitron was unable to produce a reactive [A•B] complex which promotes the formation of the *trans* cycloadduct with either of the two maleimides. In order to achieve such behaviour, the position of the recognition site relative to the reaction site would need to be redesigned. In order to produce a different geometry between the recognition site and reactive site, two designs could be envisaged (Figure 4.20). Firstly, a simple alteration of the substitution of the phenyl ring could move the recognition from the *para* to the *meta* substitution which would decrease the distance between the reactive site and recognition site. This approach was briefly attempted however synthetic problems prevented its production and subsequent investigation. Alternatively a small heterocyclic ring such as thiophene could be employed as the spacer between the reactive and recognition sites.

#### 4.18 Overall conclusions

Nanoscale construction strategies using non-covalent interactions to associate small building blocks into supramolecular structures offers an attractive opportunity to use simple building blocks to produce higher ordered nano-structures. These supramolecular assemblies may be capable of macroscopic functions and one day may be capable of fulfilling the expectations which lay with nanotechnology. In order to achieve these expectations, the future of nanoscale construction requires a fundamental understanding of recognition mediated processes at the molecular scale. Such an understanding would permit a two fold degree of control in the assembly behaviour, efficiently directing reactions to selectively form products which may then be implemented as building blocks for supramolecular assemblies.

The work presented within this chapter serves to demonstrate that by careful consideration of recognition mediated events, a reaction may be accelerated by two different recognition mediated pathways. The [A•B] complex pathway is able to accelerate a reaction upon initiation by allowing the reagents to recognise each other, bringing their reactive sites into close proximity and promoting reaction. The product of this reaction is catalytically inert and would normally take no further part in an autocatalytic cycle. The autocatalytic pathway on the other hand, as described earlier, is able to produce a template which is capable of participation in template directed synthesis giving rise to the non linear kinetics inherent with self-replicating

reactions. It is in the early stages of a self-replicating cycle in which the reaction proceeds slowly and is prone to errors leading to lower selectivity. With careful consideration of the design, the desirable properties of an **[A•B]** complex channel and autocatalytic reaction were successfully combined. The resultant system was able to accelerate a reaction upon initiation through a reactive **[A•B]** complex to produce a product capable of autocatalysis. The potential to increase the efficiency of autocatalytic self-replicators is an attractive prospect as it will allow full recognition mediated control of a reaction at all stages. This higher degree of control may help to all but eliminate byproduct formation.

It would be desirable in the near future to implement this technique on to the less efficient replicators, which show only modest selectivities as there is the potential that, with a greater degree of control, far superior enhancements in selectivities could be achieved.



# 5

## Autocatalytic Supramolecular Polymer Synthesis

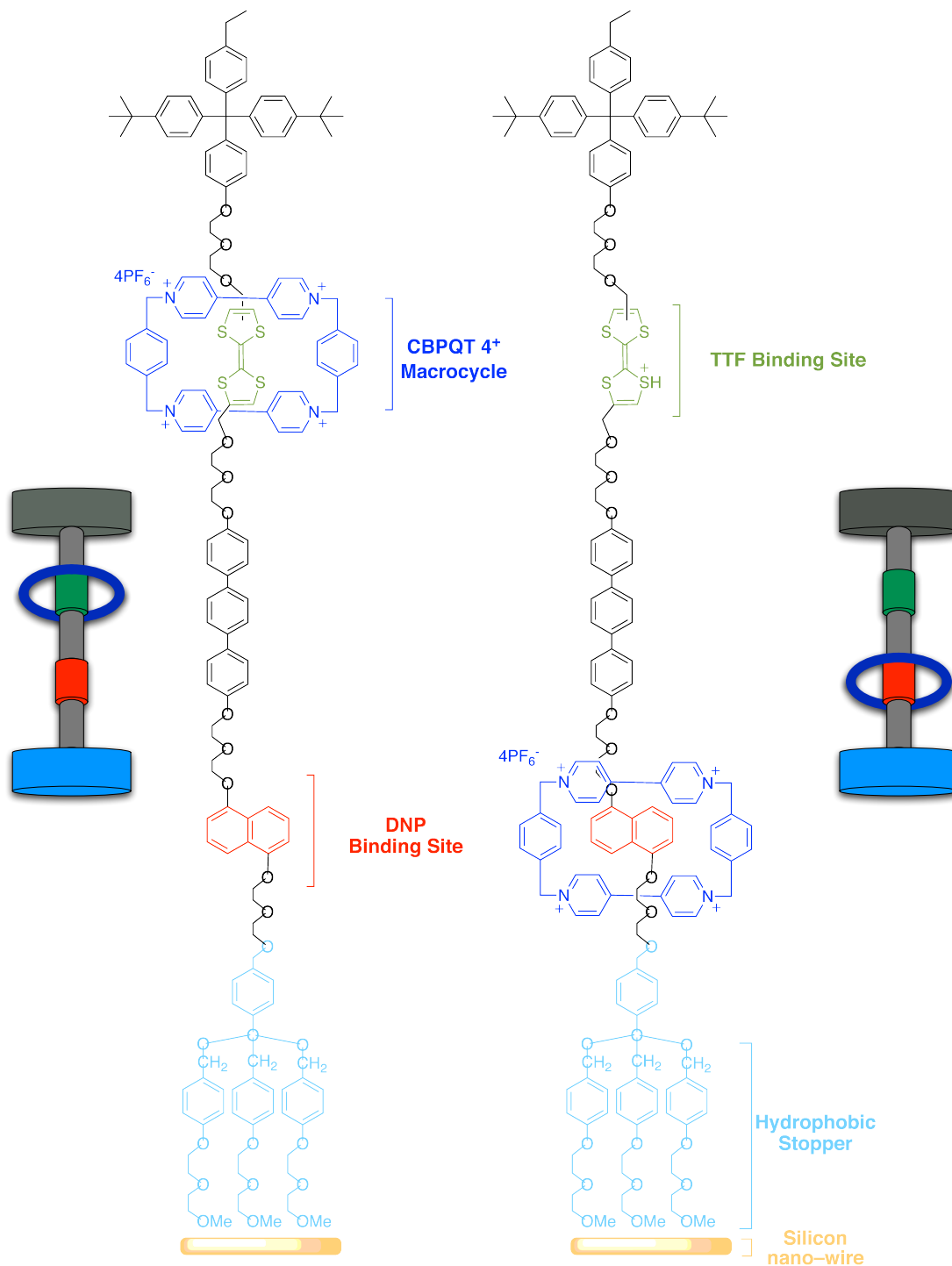
### 5.1 Nanoscale Construction – Engineering with Replication

The current trend of engineering small components is primarily concerned with the ‘top down’ design approach, that is taking an existing product and making it smaller thereby squeezing more into the same space. For example, this trend of miniaturisation is paramount in the integrated circuit manufacturing industry with the design of computer chips and memory devices. With each new generation of microprocessor or flash memory drive released, the size of components etched on to the integrated circuit decreases allowing for greater densities to be achieved. In 1965 Moore noted<sup>199</sup> an observation that the number of transistors placed on an integrated circuit roughly doubled every two years. Whilst he then only predicted that this trend would continue for a further 10 years, his prediction, later coined ‘Moore’s Law’ has roughly held true to this day.

For Moore’s law to continue, the manufacturing industry needs to continue to scale down the size of devices on integrated circuits using increasingly sophisticated methods of optical lithography and new techniques to meet the urgent challenges expected to arise by 2017.<sup>200</sup> Eventually these methods of scaling down technology will reach a theoretical limit when they enter the dimensions of the molecular level.

Information is stored digitally by modern computers using the binary (or base–2) system. All forms of information, be it text, audio, pictures etc can be converted into a series of binary digits each of which can be represented by just two values, 1 or 0. Recently, Stoddart *et al* have demonstrated<sup>167</sup> a ‘bottom up’ engineering technique ahead of its time in which a molecular electronic memory device capable of switching between two stages, representing a 1 or 0, allowing information to be stored. The device, a wholly synthetic molecule, was so small that when it was assembled on to a





**Figure 5.1:** Molecular structure and cartoon representation of the bistable [2] rotaxane used by Stoddart *et al.* The hydrophobic stopper (light blue) was adsorbed on to a silicon nano-wire surface (yellow). The ground state conformation, left, with the CBPQT<sup>4+</sup> ring (dark blue) located on the TTF site corresponds to the low conductance or 0 state. Oxidation of the TTF ring to the TTF<sup>+</sup> or TTF<sup>2+</sup> relocates the CBPQT<sup>4+</sup> ring to the dioxynaphthalene site (red) with the TTF<sup>+</sup> site subsequently reduced back to TTF to produce the metastable high conductance or 1 state, right.

surface, the density of devices on the surface corresponded to the projection of device density Moore's law expected to be reached in 2020.

The device was a bistable [2] rotaxane<sup>168</sup> (Figure 5.1) a mechanically interlocked molecular architecture comprising of two components; a linear dumbbell like thread encircled by a macrocycle, the second component. The thread possessed two  $\pi$ -electron rich sites, a tetrathiafulvalene (TTF) and a 1,5-dioxynaphthalene (DNP), each of which is capable of associating with the  $\pi$ -electron deficient cyclobis(paraquat-*p*-phenylene) (CBPQT<sup>4+</sup>) macrocyclic ring which encircles the binding site. The macrocycle is restricted from leaving the thread by two large stoppers on either end of the thread. At equilibrium, the CBPQT<sup>4+</sup> ring resides predominantly on the TTF site (>90%) and this is known as the ground state co-conformation. Oxidation of the TTF site to the TTF<sup>+</sup> radical leads to coulombic repulsion between the TTF<sup>+</sup> site and CBPQT<sup>+</sup> ring which facilitates movement of the macrocycle to the DNP site. Upon reduction of the TTF<sup>+</sup> site back to TTF, the macrocycle remains at the DNP site for a short time in a metastable state before relaxation to the original equilibrium.

In order to create a molecular memory circuit, a monolayer of the bistable [2] rotaxanes was assembled. One hundred were adsorbed by their hydrophobic stopper on to a silicon nanowire surface and sandwiched by a titanium top wire (Figure 5.1). The bistable [2] rotaxane was then able to serve as data storage element with the location of the ring between two binding sites representing the information bit 1 or 0. The ground state conformation (Figure 5.1 left) in which the CBPQT<sup>4+</sup> ring was located on the TTF site, corresponded to a low conductance or 0 state. Switching of the device, by oxidation of the TTF<sup>0</sup> site to TTF<sup>+</sup>, relocated the macrocycle on to the DNP binding site. Reduction of the TTF<sup>+</sup> site back to TTF<sup>0</sup> produced the metastable state (Figure 5.1 right) which corresponds to a high conductance or 1 state. The metastable state can then relax back to the ground state and had a half-life of 0.5 h.

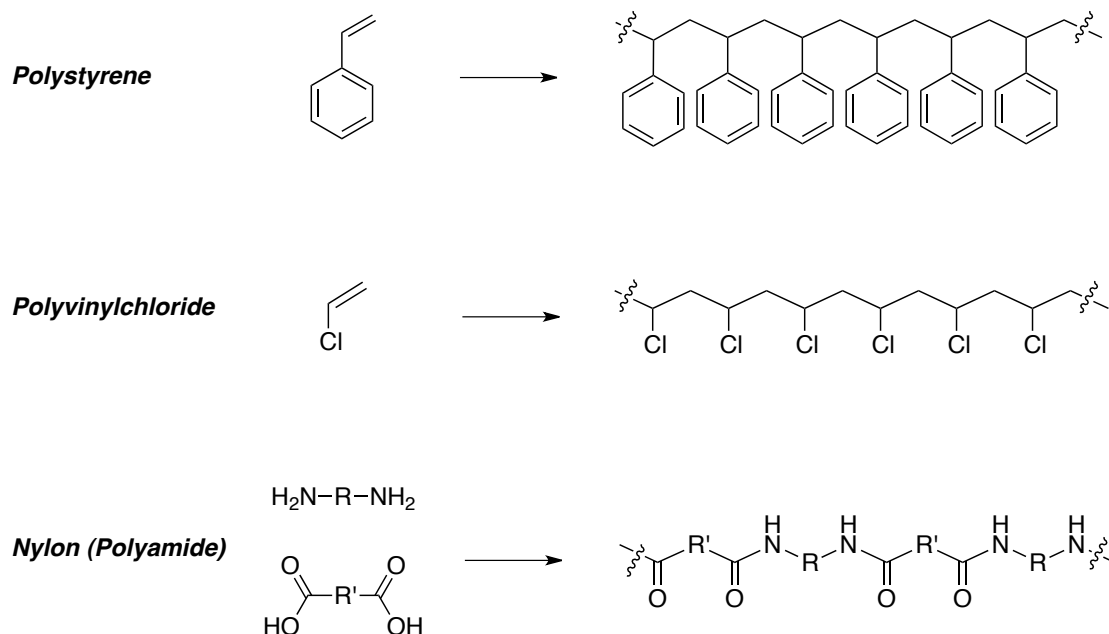
The advantages of engineering on the molecular scale offer an unprecedented opportunity for the design of new electronic devices. As these devices are created at the molecular level there is the possibility to control their behaviour by manipulating events at the atomic level. Whilst the production of molecular electronic devices is an

exciting prospect, the area of research comes with its own challenges. One drawback with molecular engineering could lay with the synthesis of functionally complex molecules required to perform such functions. The synthesis of the Stoddart rotaxane required 11 linear synthetic steps with an overall yield of 0.6%.

As well as the creation of electronic devices at the molecular level, supramolecular chemistry may also be applied to create unique forms infrastructure in which the control of recognition events at the molecular level are critical to the properties of the materials. The creation of polymeric arrays of monomers held together by non-covalent interactions, so called supramolecular polymers, are a prime example of the application of supramolecular chemistry to the creation of materials.

### 5.1.1 Conventional Covalent Polymers

Conventional synthetic polymers are long chain macromolecules made up of short repeating units which are covalently linked together. Well known examples of polymers in everyday use include polystyrene, polyvinyl chloride (PVC) and nylon (Figure 5.2).

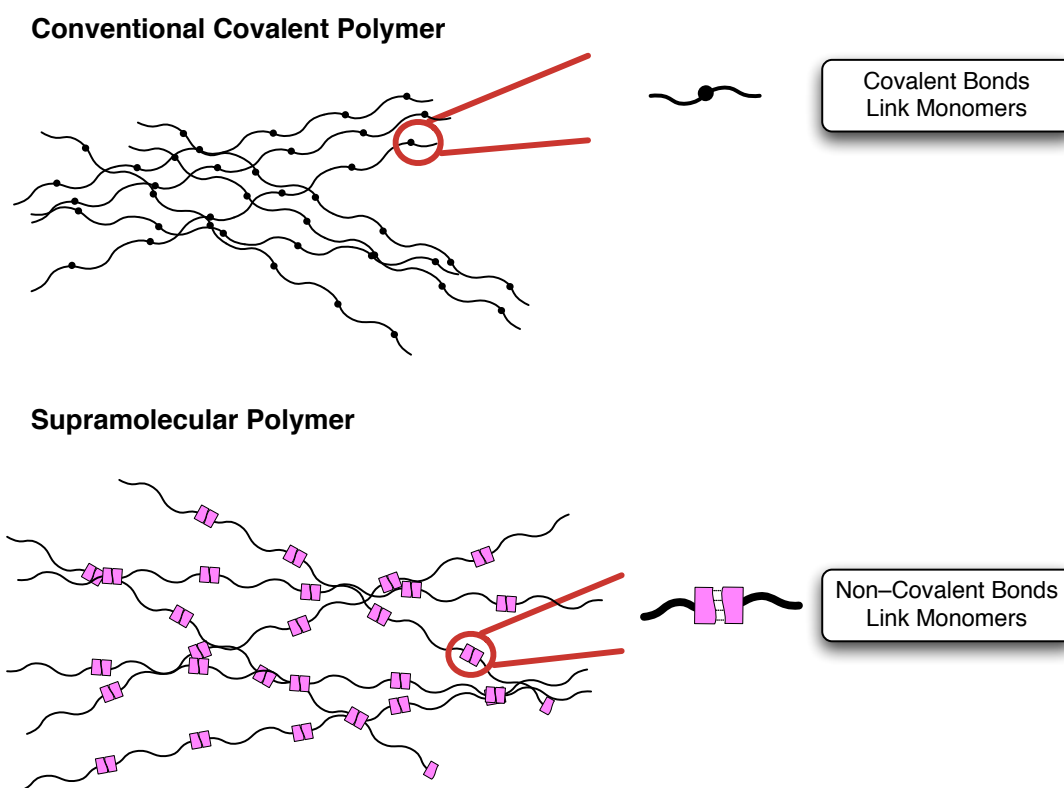


**Figure 5.2:** (top) Polystyrene polymer produced from polymerisation of styrene. (middle) Polyvinyl chloride produced from the polymerisation of vinyl chloride. (bottom) Nylon copolymer produced from the condensation of a diamine with a diacid to form a polyamide.

Synthetic polymers gain their wide range of materials properties from the entanglement of macromolecular chains and in the case of nylon additional secondary non-covalent interactions through co-operative hydrogen bonding interactions are responsible for their properties.

## 5.2 Supramolecular Polymers

Supramolecular polymers,<sup>201,202</sup> on the other hand, are polymeric arrays of a short repeating monomer which are held together by moderately strong, highly directional, non-covalent secondary interactions such as ion pairing, metal ligand bonds or hydrogen bonding (Figure 5.3). The dynamic nature of the interactions which hold the monomers together allows for the formation of responsive materials whose properties respond to a change in environment such as temperature or solvent and unique properties such as the ability to self-heal fractures.<sup>203-205</sup> As small increases in temperature can lead to a large reduction in molecular interaction strength and as a result solution viscosities, supramolecular polymer materials can be easily processed at significantly lower temperatures making their handling more convenient.

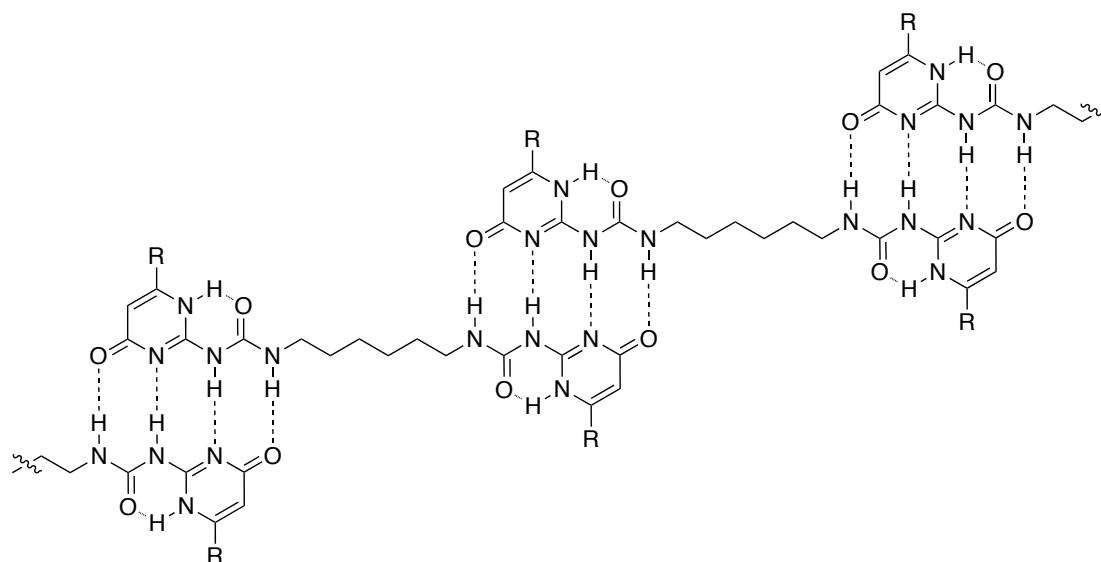


**Figure 5.3:** Top, representation of a covalent polymer made of long macromolecules which themselves are repeating monomeric units covalently linked. Bottom, representation of a supramolecular polymer of long chain macromolecules assembled by highly directional, non-covalent interactions.

The most abundant examples of supramolecular polymers found in the literature exploit H bonding as the non-covalent linker.<sup>30,201,205-209</sup>

For hydrogen bonding to be used as the non-covalent interaction between monomer units, a strong recognition motif is required. The properties describing the design of a strong hydrogen bonding motif have already been discussed in Chapter 1, Sections 1.2.1 → 1.2.5 and thus shall not be reiterated here, instead the application of Meijer's 4-H bond UPy motif to form a supramolecular polymer shall be looked at in more detail.<sup>207</sup>

The successful supramolecular polymer system constructed by Meijer *et al* comprised of two ureidopyrimidinone (UPy) units covalently linked by an aliphatic hexyl chain (Figure 5.4). The UPy groups are homoditopic and produce a very strong self-complementary quadruple hydrogen bonding array able to dimerise with the arrangement of donors (**D**) and acceptors (**A**) in either a **DDAA** structure or **DADA** structure depending on the identity of pyrimidine ring substituent.<sup>17</sup> As a consequence of the higher instance of constructive secondary interaction, the **DDAA** array is able to dimerise with a higher association constant of  $2 \times 10^7 \text{ M}^{-1}$  in chloroform.



**Figure 5.4:** Dimerisation of bifunctional UPy derivatives when  $R = C_{13}H_{27}$  via a DDAA quadruple hydrogen bond array leads to the formation of long chains held together by non-covalent interactions.

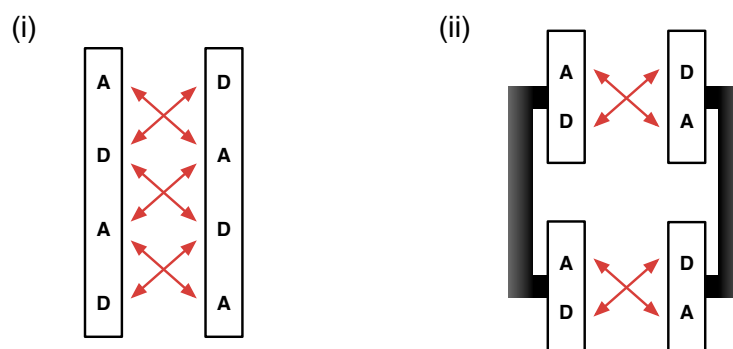
When two UPy units were connected (Figure 5.4) the hydrogen bonding between UPy units allowed the formation of long chains held together by non-covalent interactions. When the bisUPy was dissolved in chloroform to permit the non-covalent interaction, a highly viscous solution in chloroform was observed at concentrations as low as 40 mM as a result of tangling and interweaving of these long chains. Removal of the solvent lead to the formation of flexible plastic materials comprised entirely of the supramolecular polymer.

The future success of supramolecular polymer formation largely depends on the ability to synthesise their highly functional monomers some of which require sophisticated synthetic chemistry to obtain. Therefore applying an autocatalytic reaction to produce monomers for supramolecular polymer formation would be a highly desirable goal. Our research interests in autocatalytic self-replication therefore lead us to investigate the application of the process to the formation of monomeric units for polymers.

### 5.2.1 Applying Self-Replication to Supramolecular Polymer Synthesis

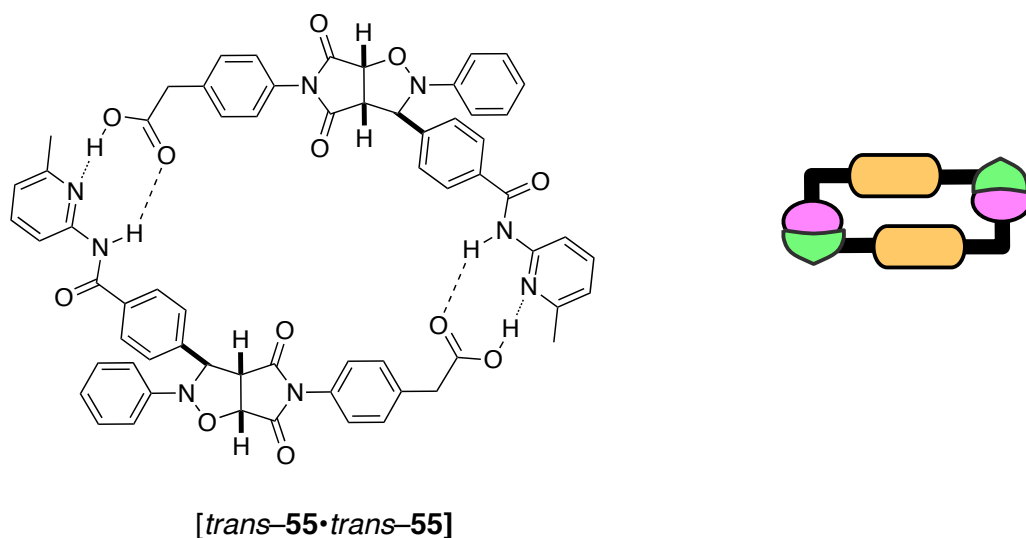
The strong  $K_a$  of UPy units was a result of the preorganisation of the H bond donor and acceptors to form the four hydrogen bond array with the order **AADD**. Of the two possible arrangements of donor and acceptor atoms in a self complementary template, Chapter 1 Section 1.2.5, the **AADD** arrangement allows the greatest number of constructive secondary interactions and theoretically should allow a higher  $K_a$  than an **ADAD** arrangement. The self-replicating templates presented within this thesis thus far are examples of quadruple hydrogen bonding arrays of arrangement **ADAD**, readily produced in an efficient and selective manner. Once produced, these templates exist as dimers in non-polar solvents.

Although this arrangement of donor and acceptor atoms theoretically would carry the greatest number of destructive secondary interactions, the design of the self-replicating template does not align all four H bond sites in close proximity, rather in two pairs of **D** and **A** atoms (Figure 5.5). The result of this design is that the number of destructive secondary interactions is decreased as the distance between the middle pair of **D** and **A** atoms becomes too great for an electrostatic interaction to occur ( $>12 \text{ \AA}$ ).



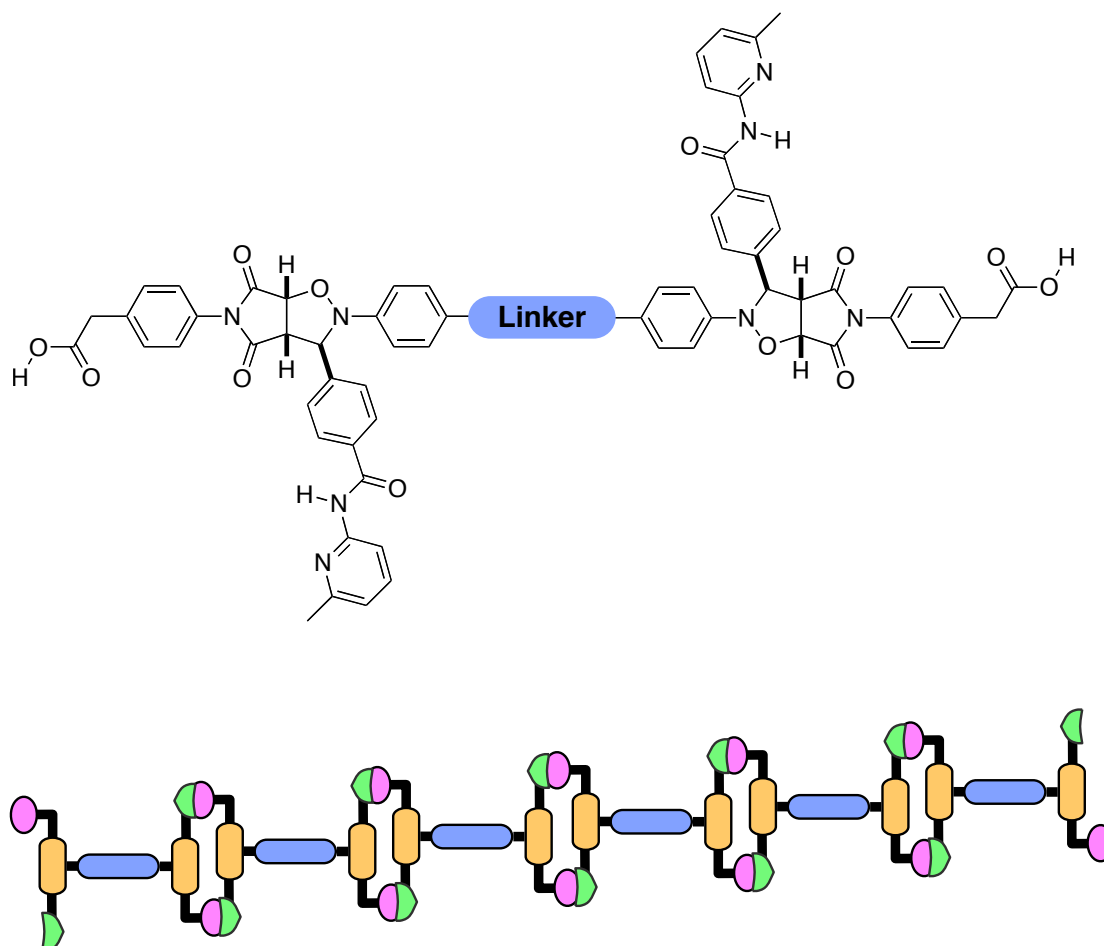
**Figure 5.5:** (i) Arrangement of self-complementary **ADAD** 4 H bond array with the four destructive secondary interactions highlighted as red arrows showing the maximum possible number of destructive secondary interactions. (ii) Arrangement of H bond donor and acceptor atoms in a self-replicating template. The same **ADAD** arrangement is achieved however the acceptor and donor atoms are arranged in two pairs held apart by a spacer unit. As a result, there is too great a distance between the inner middle pair of **A** and **D** atoms to allow secondary interactions to occur.

Earlier in this thesis, Chapter 4 Section 4.5, the self-replicating template *trans*-**55** was introduced (Figure 5.6). This template, formed from the autocatalytic reaction of nitrene **52** with maleimide **34**, is a quadruple hydrogen bond (**DADA**) dimer with an association constant ( $K_a$ ) of  $3 \times 10^6 \text{ M}^{-1}$  in  $\text{CDCl}_3$  at  $0^\circ \text{C}$  (template dimer  $K_a$ s of up to  $10^8 \text{ M}^{-1}$  have been presented in this thesis, Section 2.7.1 and 2.7.2). Whilst in this case the  $K_a$  is lower in magnitude to the UPy polymer system, it is strong enough to be considered appropriate for use as the non-covalent interaction for the creation of a supramolecular polymer.



**Figure 5.6:** Autocatalytic template duplex and cartoon, below linking duplexes together to give rise to a supramolecular polymer

Considering the successful UPy polymer was constructed by covalently linking two UPy units together (Figure 5.4) we envisaged that covalently linking two self-replicating templates would also lead to the formation of a supramolecular polymer (Figure 5.7).



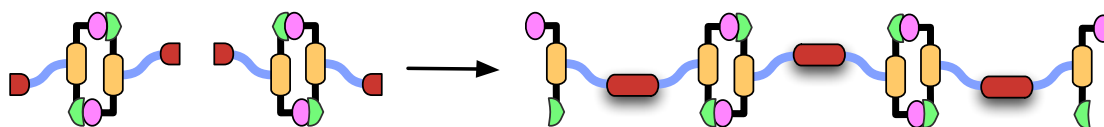
**Figure 5.7:** Linking together two templates gives rise to a monomer which can polymerise as a result of the hydrogen bonding between templates.

### 5.3 First Generation Monomer Design

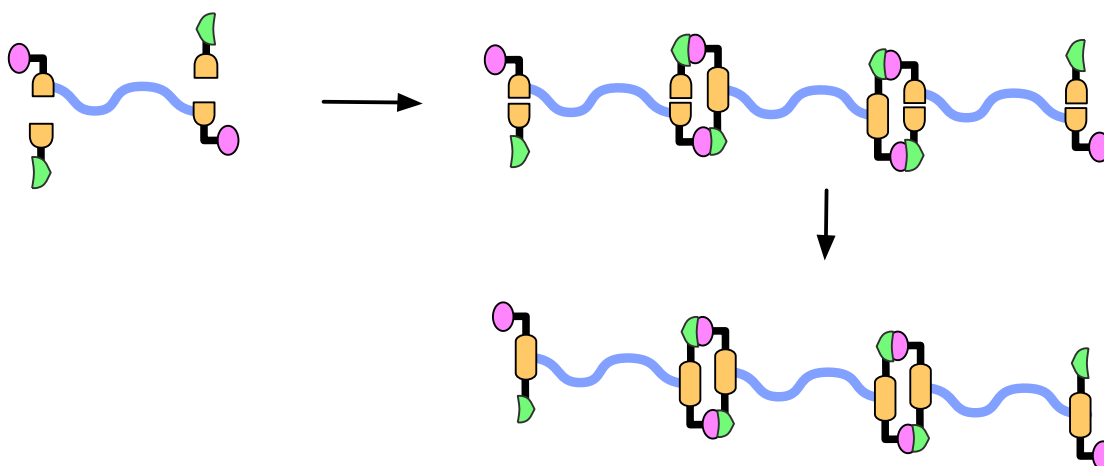
When identifying a potential method to link to templates together, two possible synthetic scenarios were conceived. Firstly, the two templates could be connected covalently after the autocatalytic template formation (Figure 5.8 (i)). This scenario would require an additional reaction to be performed after the autocatalytic template formation. This additional covalent bond forming reaction would have to be carefully designed to be orthogonal to the 1,3 dipolar cycloaddition to minimise interference.



(i) **Linking Template Dimers**



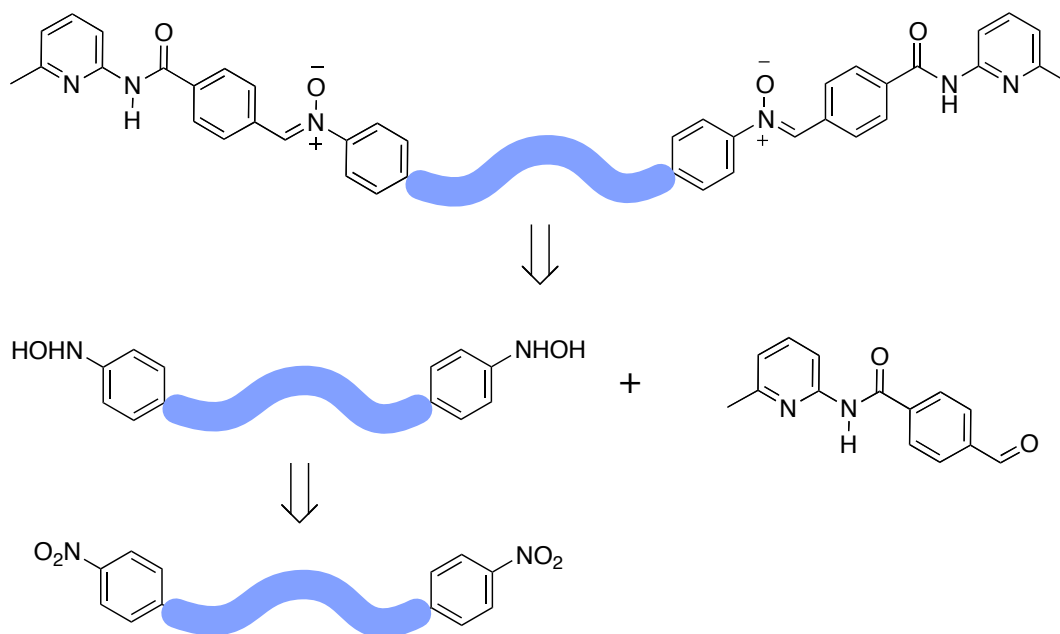
(ii) **Autocatalytic Template Linking**



**Figure 5.8:** Two scenarios for covalently linking two autocatalytic templates. (i) Post autocatalytic reaction linking can be achieved by performing an additional covalent reaction after the formation of each template. (ii) Autocatalytic template linking utilises a disubstrate which employs the autocatalytic reaction to produce the bistemplate and polymer formation

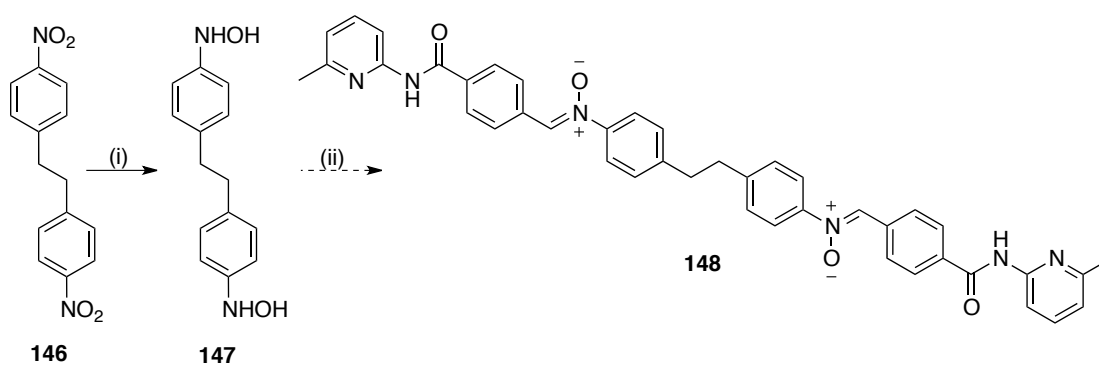
The more favourable second scenario required the use of a bifunctional substrate (Figure 5.8 (ii)). This substrate possesses the covalent linking group which after cycloaddition will produce a bistemplate. The bistemplate should then spontaneously form a supramolecular polymer *in situ*. This method has the advantage of utilising the well established autocatalytic reaction to produce each of the two templates which then act as the non-covalent interactions between monomers. This method can be termed ‘self-replicating supramolecular polymerisation’.

The autocatalytic template linking method required the production of a bifunctional substrate. For our 1,3 dipolar cycloaddition, this would mean that either a bisnitron or a bismaleimide would need to be produced. By considering a retrosynthetic analysis for the creation of a bisnitron, we envisaged that such a product could be achieved without the need for any additional synthetic steps with respect to the previously reported mono nitrones.



**Scheme 5.1:** Retrosynthetic analysis of the bisnitron showing its formation from a recognition bearing aldehyde and dihydroxylamine which in turn can be produced from a dinitro compound.

Our attention therefore focused on designing a bisnitron which consisted of two of the original recognition bearing nitrones covalently linked together. A retrosynthetic analysis revealed that the bisnitron could be obtained from the dicondensation of two equivalents of the recognition bearing aldehyde **49** with a dihydroxylamine which in turn could be obtained after reduction of a dinitro compound. Our initial attempt at the synthesis of a bisnitron therefore started by identifying a suitable dinitro compound.

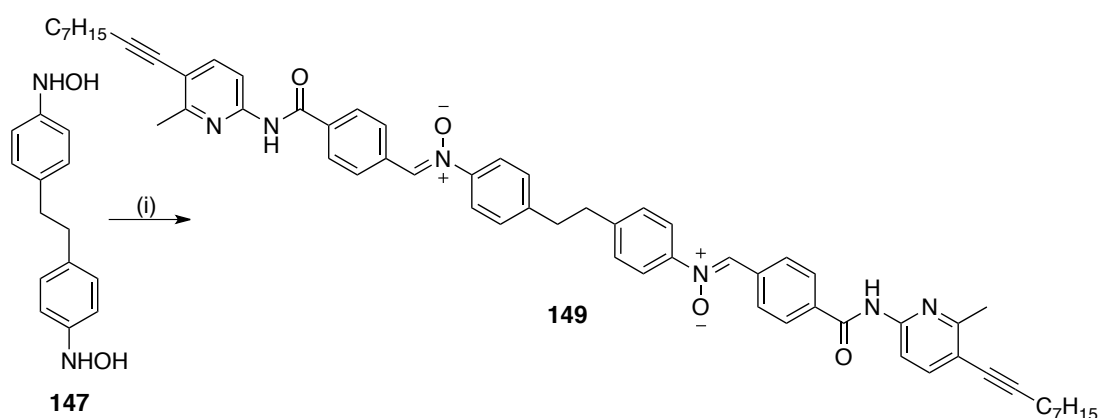


**Scheme 5.2:** Reagents and conditions (i) Rh/C,  $\text{NH}_2\text{NH}_2 \cdot \text{H}_2\text{O}$ , THF, RT, 1 h, quantitative (ii) 2 eq aldehyde **49**, EtOH, RT, 48 h

We identified the commercially available dinitro compound 4,4'-dinitrobibenzyl **146** to use for our initial attempt to create a bisnitron. The use of this dinitro would lead to a bisnitron connected by a CH<sub>2</sub>CH<sub>2</sub> group which will have minimal effect on the electronic properties of the nitron reactive site. The synthesis began by the reduction of 4,4'-dinitrobibenzyl using transfer hydrogenation with Rh/C and hydrazine hydrate to yield dihydroxylamine **147** which was used immediately in a dicondensation reaction with two equivalents of aldehyde **49**. The product of the dicondensation reaction was very poorly soluble in the non-polar solvents which would be required for the cycloaddition reaction and even in DMSO. As a result this compound was not suitable for use in creating the supramolecular polymer and rejected.

#### 5.4 Second Generation Monomer Design

The poor solubility of the initial bisnitron design led us to attempt to increase its solubility. As shown previously, decorating compounds with lipophilic alkyl chains has the effect of increasing the solubility of the nitrones and subsequent cycloadducts in the non-polar solvents used in our investigations. The introduction of a nonyl group on to the recognition bearing aldehyde **130**, Section 4.9, proved successful and the first attempt to increase the solubility of the bisnitron was attempted by using this aldehyde in the nitron synthesis.



**Scheme 5.3:** Reagents and conditions (i) aldehyde **130**, EtOH, RT, 48 h, 89%

The dicondensation reaction was repeated using two equivalents of aldehyde **130** with dihydroxylamine **147** to yield bisnitron **149** which had increased solubility in

$\text{CDCl}_3$  with a maximum concentration of 5 mM able to be reached. Whilst the addition of the nonyl groups most certainly increased the solubility of the bisnitron in chloroform, the resultant templates were poorly soluble and higher concentrations were desired to adequately investigate supramolecular polymer formation.

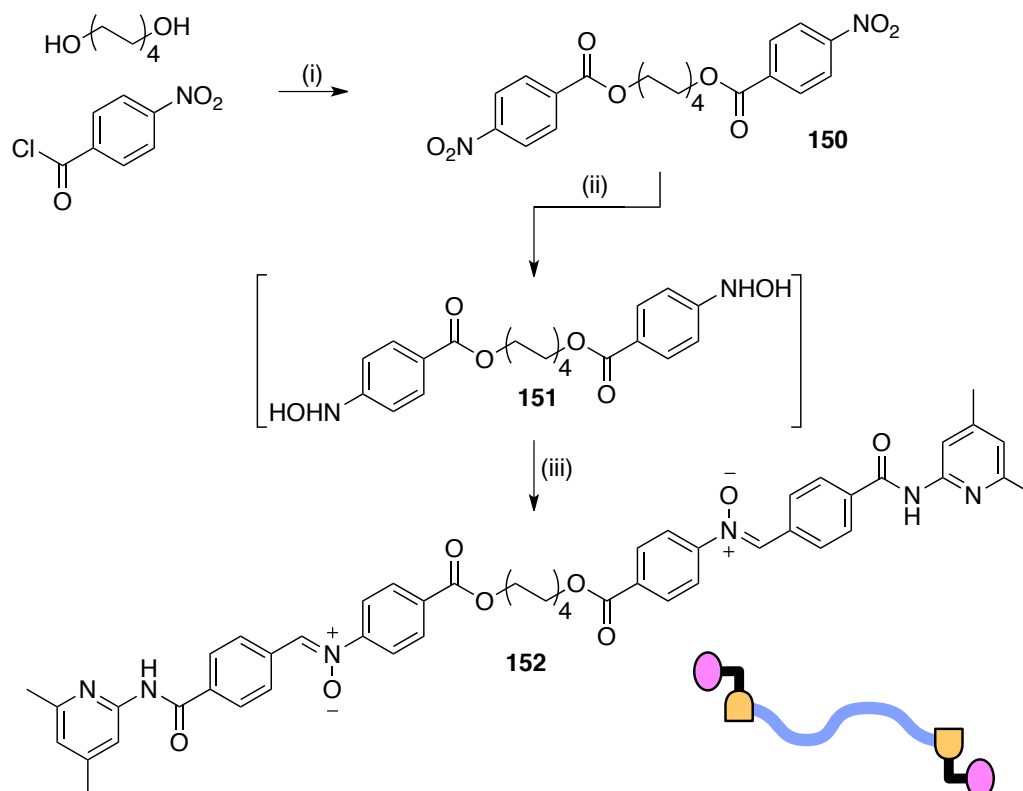
The poor solubility of the bisnitron in non-polar solvents was not entirely as a result of the chemistry of the linking group but certainly was not aided by it. 4,4'-dinitrobibenzyl was itself poorly soluble in non-polar solvents and thus by using this compound as a linking group we cannot expect to produce a very soluble bisnitron.

### 5.5 Third Generation Monomer Design

The use of the nonyl decorated aldehyde had the added drawback of decreasing the  $K_a$  between the amidopyridine and acid. Ideally we would desire the  $K_a$  to be as high as possible to increase the strength of interactions between the template duplex and thus encourage long chain polymer formation. For this reason, the nonyl decorated amidopyridine aldehyde **130** was abandoned and instead we chose to use the dimethylaminopyridine which was shown<sup>210</sup> to have a higher  $K_a$  with phenylacetic acids. The focus of the synthesis therefore changed as new methods of increasing the solubility of the bisnitron without causing any disruption to the amidopyridine or phenylacetic acid groups was desired.

As a consequence of the very poor solubility of the initial dinitro compound used as the linking group, the possibility of using a linker group which was very soluble to aid the solubility of the bisnitron was explored.

We identified the possibility of utilising 4-nitrobenzoyl chloride as a source of an aryl nitro group which could be reacted with a dinucleophile to produce a dinitro compound. This approach had the advantage that the acid chloride could be coupled with a range of different dinucleophiles, in the short term giving rise to a facile route to produce a series of different linker groups and in the long term could be used to fine tune the properties of a supramolecular polymer by altering the chemical composition of the linking group.



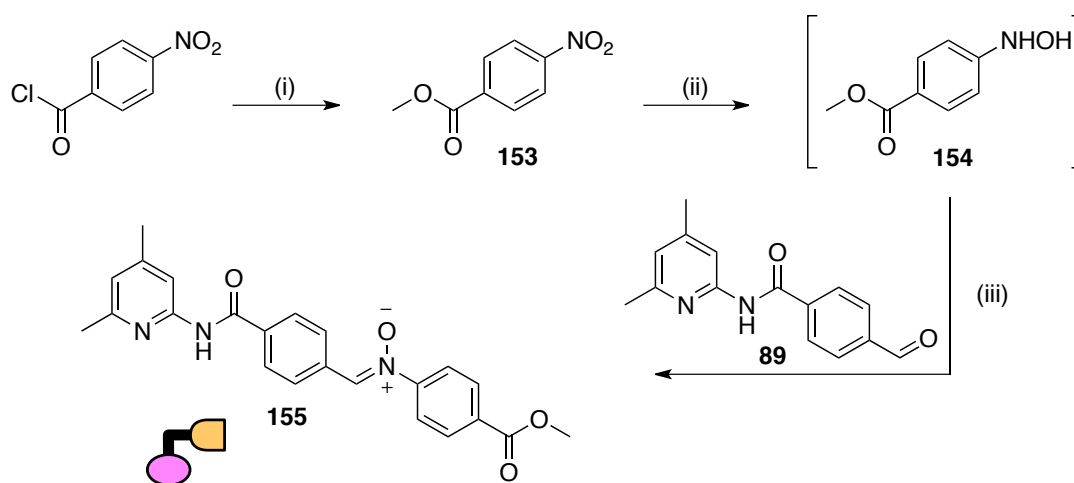
**Scheme 5.4:** Reagents and conditions (i)  $\text{CH}_2\text{Cl}_2$ ,  $\text{NEt}_3$ ,  $0\text{ }^\circ\text{C}\rightarrow\text{RT}$ , 16 h, 52% (ii)  $\text{Rh/C}$   $\text{NH}_2\text{NH}_2\cdot\text{H}_2\text{O}$ , THF, RT, 1 h, quantitative (iii) 3 eq. aldehyde **89**, EtOH, RT, 48 h, 69%

Initially 1,8 octanediol was selected to be used as the linking group as the alkyl chain should increase the solubility of the bisnitrone and subsequent cycloadducts. Dinitro **150** was synthesised by ester formation between 1,8–octanediol with an excess of 4–nitrobenzoyl chloride. At this point, the dinitro compound was particularly soluble in chloroform which was encouraging. Dinitro **150** was reduced to the dihydroxylamine **151** using transfer hydrogenation with rhodium on carbon with hydrazine hydrate in THF and used immediately in a condensation reaction with an excess of aldehyde **89** in ethanol. Normally when attempting the condensation reaction the minimum volume of ethanol required to dissolve the hydroxylamine is used with the aldehyde added to this solution. By using the minimum volume of ethanol required, the product nitrone which is inevitably less soluble is encouraged to precipitate out and can be isolated by filtration. During the first attempt at the diconsensation, the produced isolated by precipitation was found to be a mono–nitrone mono–hydroxylamine, *i.e.* had reacted once and then precipitated. In order to overcome the precipitation of mono reacted compounds, the reaction was repeated under diluted conditions to enable the

dihydroxylamine to react fully before precipitating. The optimum concentration of reagents found to use was 3 equivalents of aldehyde, a 50% excess, with the dihydroxylamine in ethanol at 0.015 M. Under these conditions only bisnitronone **152** was isolated. Encouragingly, bisnitronone **152** proved to be readily soluble in chloroform and was investigated up to a concentration of 200 mM.

### 5.5.1 Design and Synthesis of an Model Mono–Nitronone

In the design of the bisnitronone, we have included the electron withdrawing ester group *para* to the nitronone reactive site. In order to ensure the change in electronic properties of the nitronone were tolerated by the autocatalytic 1,3–dipolar cycloaddition an appropriate mono nitronone was synthesised to act as a control compound so that a kinetic analysis may be performed for the reaction with maleimide **34** to ensure the reaction remained an efficient autocatalytic reaction and therefore the cycloaddition on each site of the bisnitronone may be understood.

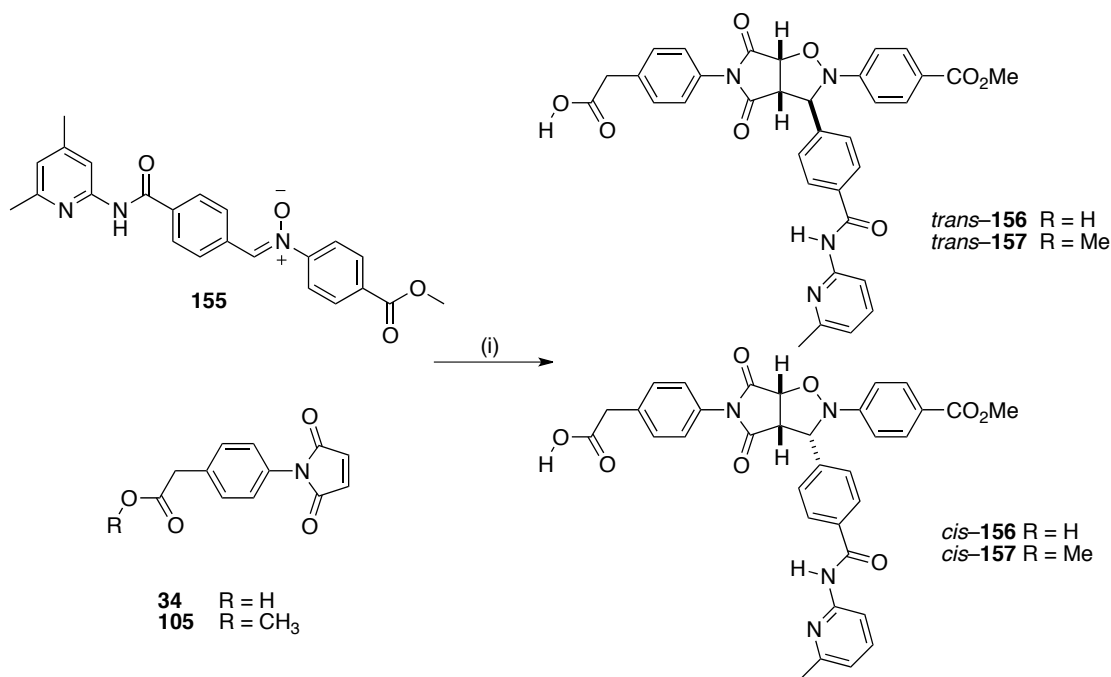


**Scheme 5.5:** Reagents and conditions (i) MeOH, CH<sub>2</sub>Cl<sub>2</sub>, NEt<sub>3</sub>, 0 °C→RT, 16 h, quantitative (ii) Rh/C NH<sub>2</sub>NH<sub>2</sub>·H<sub>2</sub>O, THF, RT, 1 h, quantitative (iii) 3 eq aldehyde **89**, EtOH, RT, 48 h, 39%

The methyl ester was chosen as an appropriate control as it would possess the electronic properties of a *para*–ester group attached to the nitronone reactive site and it also was consistent with the ester having the aryl and alkyl sides.

### 5.5.2 Kinetic Analysis of the Reaction of mono Nitrone **155** with Maleimide **34**

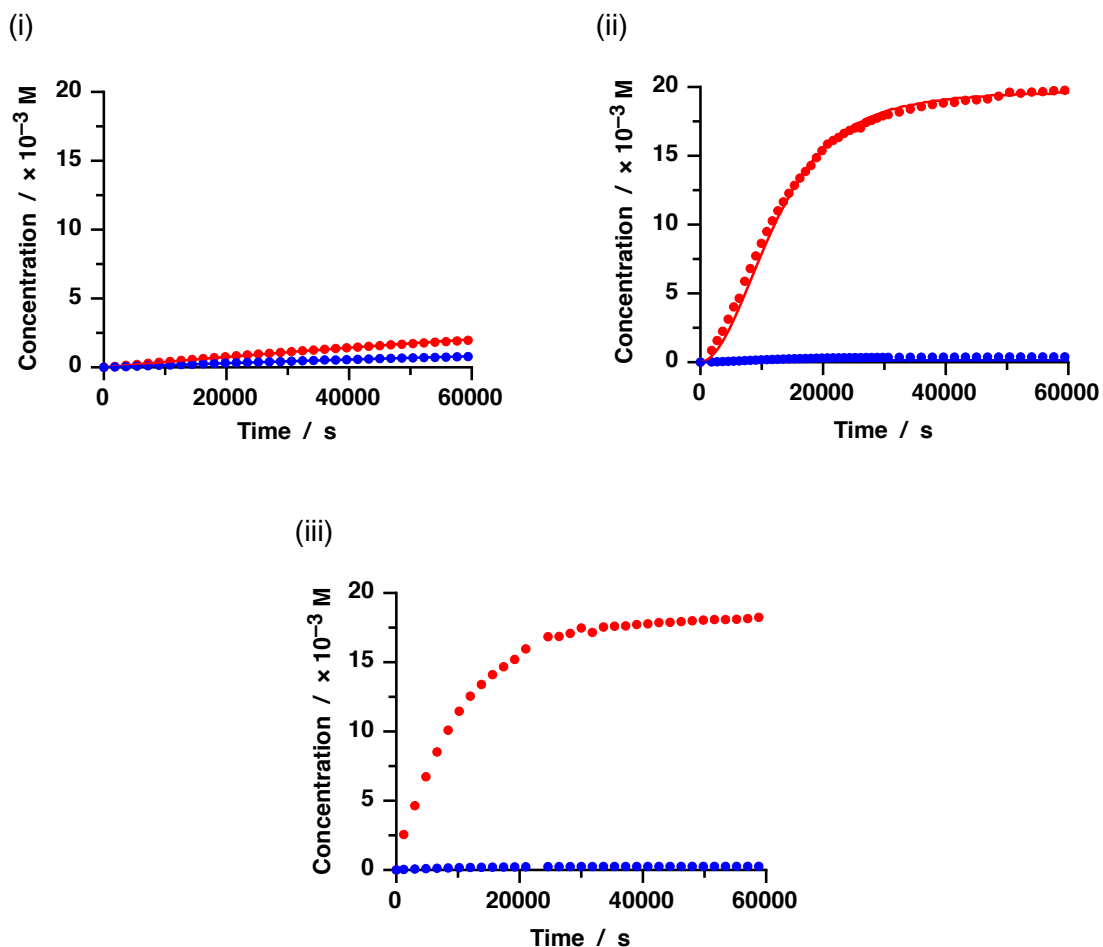
In order to confirm the ester group did not perturb the autocatalytic reaction with maleimide **34** and to gain an insight into the strength of the template duplex with the dimethyl amidopyridine a kinetic analysis of the reaction of nitrone **155** with maleimide **34** was performed. This control reaction is an important test as increasing the strength of the recognition events in the system may have the effect of increasing the strength of the template duplex to the point where a slower dissociation may effect the overall efficiency of the autocatalytic cycle.



**Scheme 5.6:** Reagents and conditions (i) CDCl<sub>3</sub>, 20 mM, 0 °C, 16 h

In order to determine the native reactivity of the cycloaddition in the absence of recognition effects, the recognition inhibited control reaction of nitrone **155** with the maleimide methyl ester **105** was performed at 20 mM in CDCl<sub>3</sub> at 0 °C and monitored by 500.1 MHz <sup>1</sup>H NMR spectroscopy for 16 h. After deconvolution of appropriate resonances a concentration vs. time profile was constructed (Figure 5.9 (i)) which showed that the recognition disabled reaction proceeded slowly under the conditions used with 1.96 mM *trans*-**157** and 0.78 mM of *cis*-**157** produced after 16 h, representing a *trans* : *cis* ratio of 2.5:1 for this reaction and an overall conversion of 14%. In order to obtain rate information, we undertook kinetic simulation and fitting with SimFit using the standard kinetic model for a bimolecular reaction. The results of

the fitting revealed the reaction rate constants  $k_{\text{trans}}$  and  $k_{\text{cis}}$  to be  $9.90 \times 10^{-5} \text{ Ms}^{-1}$  and  $3.92 \times 10^{-5} \text{ Ms}^{-1}$  respectively.



**Figure 5.9:** Concentration vs. time profiles for; (i) the recognition disabled control reaction between nitrone **155** and maleimide ester **105** (ii) the recognition mediated reaction between nitrone **155** and maleimide **34** and (iii) the template doped recognition mediated reaction of nitrone **155** with maleimide **34** in the presence of 20 mol% *trans*-**156**. Experimental data shown as points with *trans* products represented as (●) and *cis* products represented as (●). Simulated data is represented as lines with *trans* products represented by (—) and *cis* products by (—). Reaction conditions: 20 mM,  $\text{CDCl}_3$ ,  $0^\circ\text{C}$ , 16 h.

**Table 5.1:** Rate constant information as calculated from kinetic simulation and fitting using SimFit.

	$k_{\text{trans}}$	$k_{\text{cis}}$
Bimolecular Rate / $\times 10^{-5} \text{ M}^{-1}\text{s}^{-1}$	9.90	3.92
Ternary Complex Reaction / $\times 10^{-5} \text{ s}^{-1}$	516	—
$EM / M$	50.9	—
$K_a$ Template Duplex / $\times 10^7 \text{ M}^{-1}$	2.66	—
Connection $EM / M$ ( $\Delta G^\ddagger / \text{kJmol}^{-1}$ )	8.2 (4.8)	—



Next, the recognition mediated reaction of maleimide **34** with nitron **155** was performed under the same conditions and assayed as before by 500.1 MHz  $^1\text{H}$  NMR spectroscopy. After deconvolution a concentration vs. time profile was constructed which showed (Figure 5.9 (ii)) the selective enhancement in the production of *trans*-**156** and inhibition of *cis*-**156**. The profile of the production of *trans*-**156** displayed a sigmoidal shape with a very short lag period, characteristic of an autocatalytic reaction as seen with previous replicating systems. After 16 h, 19.6 mM of *trans*-**156** had been produced along with 0.4 mM of *cis*-**156** representing a *trans* : *cis* ratio of 49:1 and quantitative overall conversion. In order to obtain a deeper understanding of any subtle differences caused by using the dimethyl amidopyridine recognition site and *para* ester group, we performed kinetic simulation and fit the reaction data to the model for a self-replicating reaction using SimFit. The rate of reaction within the ternary complex was determined to be  $5.16 \times 10^{-3} \text{ s}^{-1}$  which reveals the *EM* of the system to be 50.9 M. The  $K_a$  of the template duplex was calculated to be  $2.66 \times 10^7 \text{ M}^{-1}$ . This association constant represents an increase in strength in comparison to the system reported previously, Section 4.5, in which the mono methyl amidopyridine template had a dimerisation constant of  $3.01 \times 10^6 \text{ M}^{-1}$ . It was reassuring to observe that after the changes made to the system to increase the binding strength between templates, there has been no loss of reactivity of the autocatalytic cycle.

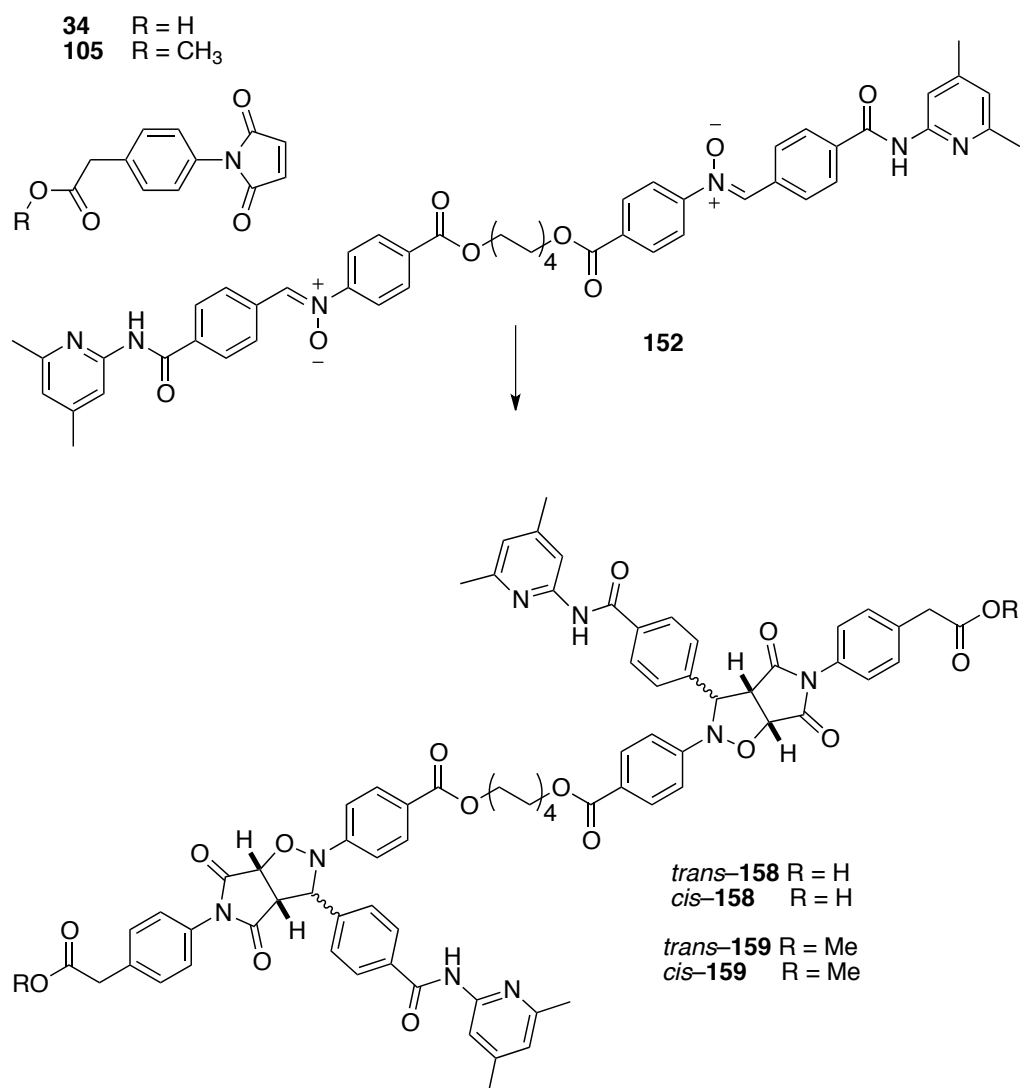
Finally to demonstrate that the template was indeed responsible for directing the outcome of the reaction, the reaction of nitron **155** with maleimide **34** was performed in the presence of 20 mol% preformed *trans*-**156** under the same conditions and assayed as before. The concentration vs. time profile (Figure 5.9 (iii)) revealed that the lag period observed in the undoped reaction was absent and the reaction proceeded at the maximum rate of reaction from the start showing that preformed template added to the reaction was free to act as a catalyst. After 16 h, 18.6 mM of *trans*-**156** had been produced along with 0.26 mM *cis*-**156** representing a *trans* : *cis* ratio of 70 : 1, an increase in comparison to the undoped recognition mediated reaction.

The results from the analysis of the mono nitron display that after modification of our original replicating motif, by strengthening the recognition events and adding an electron withdrawing group, the reaction is still autocatalytic with the product *trans*-**156** acting as a template to promote the formation of itself. Encouragingly, the

increased strength of template duplex,  $2.66 \times 10^7 \text{ M}^{-1}$ , is comparable with the strength of dimerisation of two UPy units which was  $2 \times 10^7 \text{ M}^{-1}$ .

### 5.5.3 Kinetic Analysis of the reaction of Bisnitron 152 with Maleimide 34

With the results of the mono nitron system indicating that the mono template was formed by autocatalytic self-replication, the next investigation undertaken was the reaction of bisnitron **152** with maleimide **34** to produce a bitemplate.

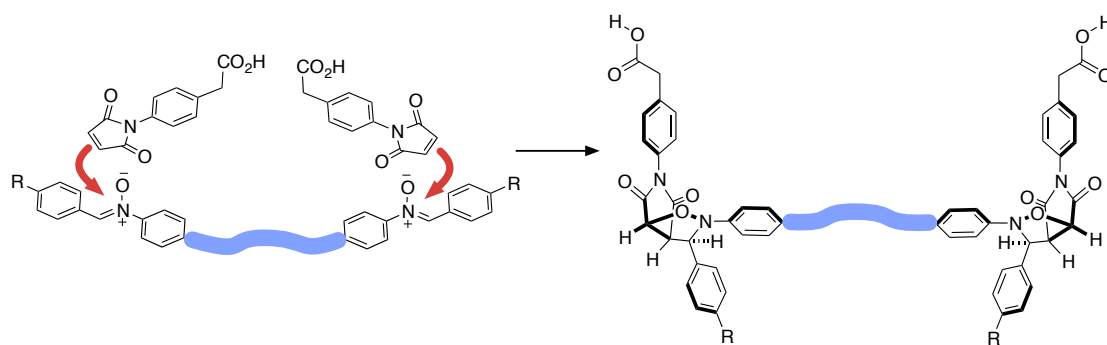


**Scheme 5.7:** Reagents and conditions (i) 10 mM bisnitron, 20 mM maleimide, 6 mM 4-bromo-*tert*-butylbenzene, CDCl<sub>3</sub>, 0 °C, 16 h

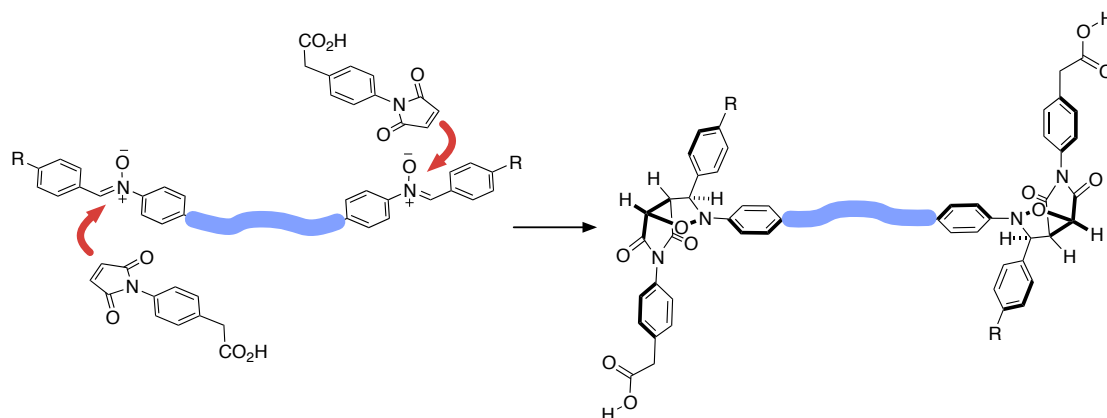
Each reaction was performed with a concentration of 10 mM of bisnitron (equivalent to 20 mM of the mono nitron) and 20 mM of maleimide, one equivalent per nitron reactive site.

Initially, the recognition disabled bimolecular reaction of bisnitron **152** with maleimide **105** was performed and monitored by 500.1 MHz  $^1\text{H}$  NMR spectroscopy for 16 h. From the reaction eight unique products were expected, two singly reacted intermediates products and six fully reacted products, as a result of there being a linker between two reactive sites. The intermediates in which one of the nitron sites has reacted would be; 'nitron-*trans*' describing where the bisnitron has reacted once to form the *trans* cycloadduct and nitron-*cis* where the bisnitron has reacted once to form the *cis* cycloadduct.

(i) '*syn*' Reaction



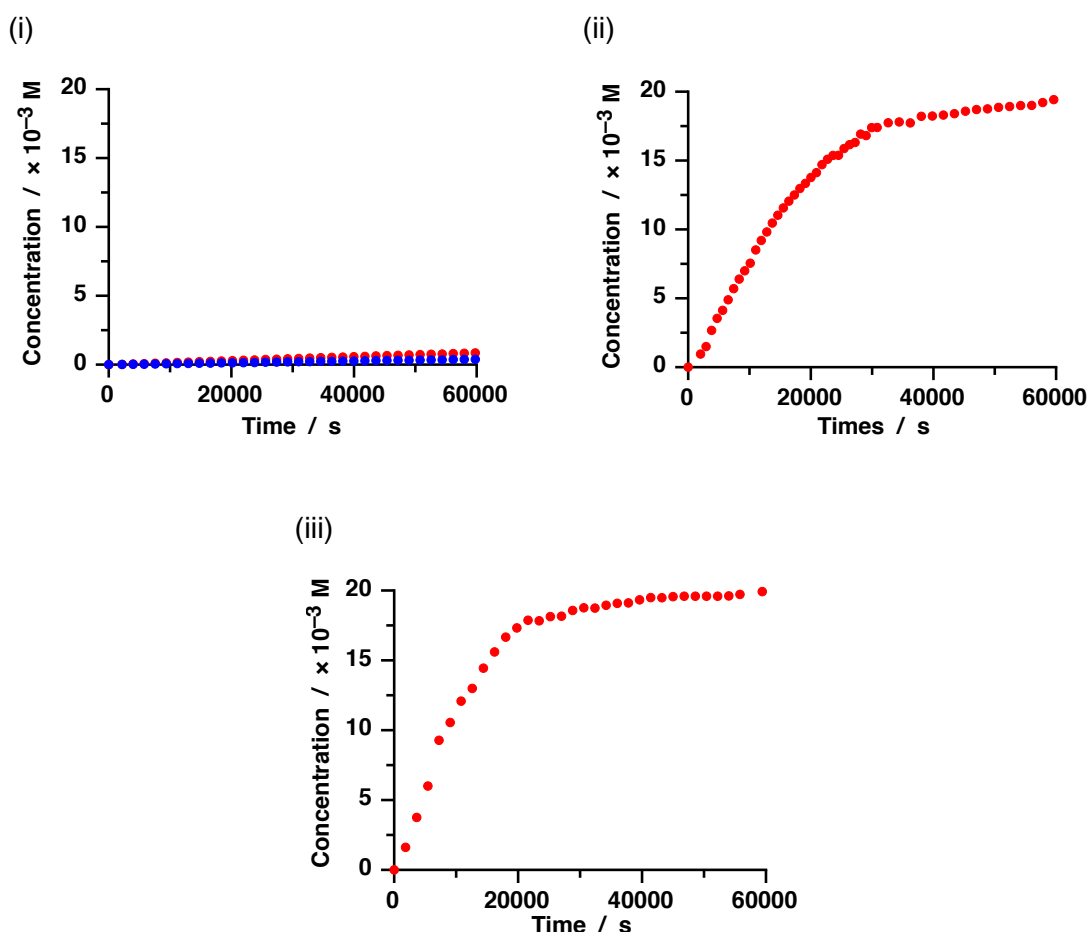
(ii) '*anti*' Reaction



**Figure 5.10:** Diagrams of the two possible manners of attack of the maleimide on the bisnitron (i) the pseudo '*syn*' reaction of two maleimides on the same side of the bisnitron and (ii) the pseudo '*anti*' reaction of two maleimides on opposite sides of the bisnitron.

From these two intermediates, a further cycloaddition reaction may take place at the second nitron reactive site. This cycloaddition will form either the *trans* or *cis*

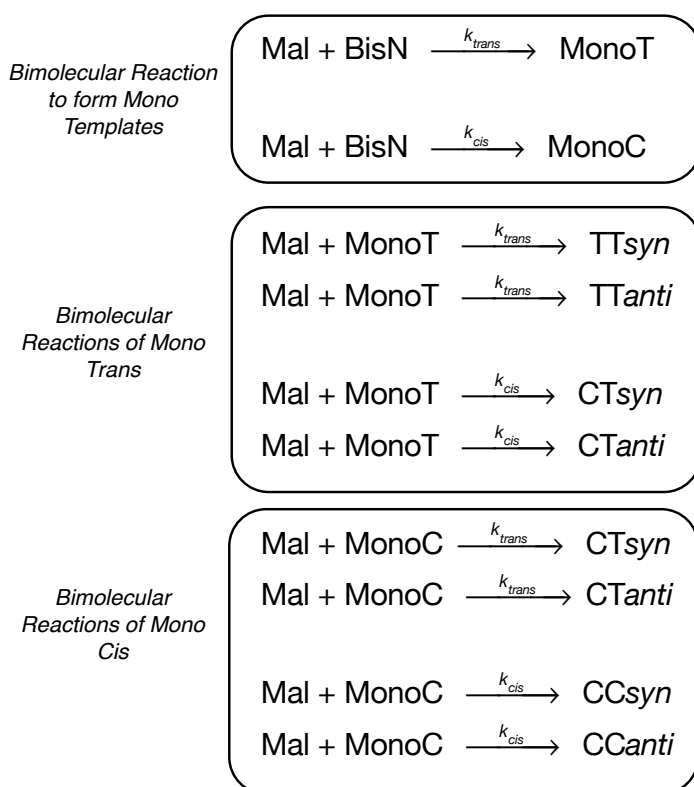
cycloadducts and may also do so in two possible orientations relative to the first cycloadduct, either *syn* or *anti*. If the nitrene is considered to be flat, the maleimide may attack on either the top face or the bottom face. With two nitrene sites linked together, there is the possibility that each maleimide may attack on the same side in a pseudo '*syn*' manner (Figure 5.10 (i)) or may attack on opposite sides in a pseudo '*anti*' manner (Figure 5.10 (ii)). A second cycloaddition reaction will therefore lead to the formation of six fully reacted products; *syn bistrans* – in which two *trans* cycloadducts are linked in the *syn* manner; *anti bistrans* – in which two *trans* cycloadducts are linked in the *anti* manner; *syn biscis* – in which two *cis* cycloadducts are linked in the *syn* manner; *anti biscis* – in which two *cis* cycloadducts are linked in the *anti* manner; *syn tran–cis* – in which a *trans* and *cis* cycloadduct are linked in the *syn* manner and finally; *anti trans–cis* – in which a *trans* and *cis* cycloadduct are linked in the *anti* manner.



**Figure 5.11:** Concentration vs. time profiles for; (i) the recognition disabled control reaction of bisnitrene **152** and maleimide ester **105** (ii) the recognition mediated reaction of bisnitrene **152** and maleimide **34** and (iii) the template doped recognition mediated reaction of bisnitrene **152** with maleimide **34** in the presence of 10 mol% *trans*-**158**. Experimental data shown as points with *trans* products represented as (●) and *cis* products represented as (●). Reaction conditions: 10 mM bisnitrene, 20 mM maleimide,  $\text{CDCl}_3$ , 0 °C, 16 h.

From 500.1 MHz  $^1\text{H}$  NMR spectroscopy, the eight potential products could not be observed as only one signal was observed for the *trans* cycloadduct and one signal for the *cis* cycloadducts. The overlap of signals might be expected as the two sites are far from each other and as a result each should have little chemical effect on the other. As a result of this, during deconvolution of the  $^1\text{H}$  NMR spectra, the total *trans* concentration and total *cis* concentration was reported.

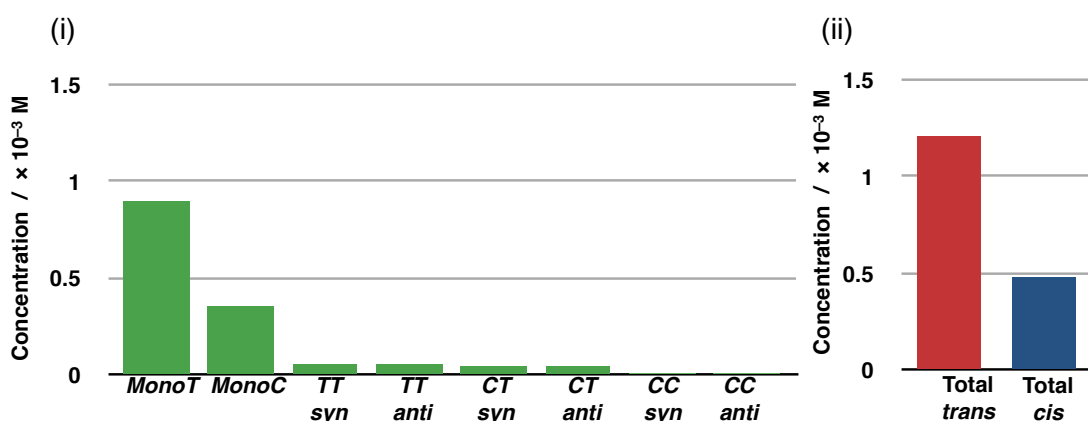
After 16 h, the concentration vs. time profile for the recognition disabled reaction (Figure 5.11 (i)) showed slow conversion with a total of 1.92 mM *trans* cycloadduct present and 0.86 mM *cis* cycloadduct present representing *trans* : *cis* ratio of 2.2:1 and an overall conversion of 14%. In order to understand how the distribution of the five products would constitute the total conversion, a kinetic model was created which described the system.



**Scheme 5.8:** Kinetic model produced to describe the reactions which could occur in a mixture of bisnitron **152** with maleimide **105**. Abbreviations used in model: Mal = maleimide; BisN = bisnitron, monoT = nitron-*trans*: monoC = nitron-*cis*: TTsyn = *syn* bis*trans* template: TTanti = *anti* bis*trans* template: CCsyn = *syn* bis*cis* template: CCanti = *anti* bis*cis* template: CTsyn = *syn* *cis*-*trans* template: CTanti = *anti* *cis*-*trans* template.

The model (Scheme 5.8) describes all the possible reactions of maleimide **105** (Mal) with bisnitron **152** (BisN) and all possible intermediates. Each of the reaction rates

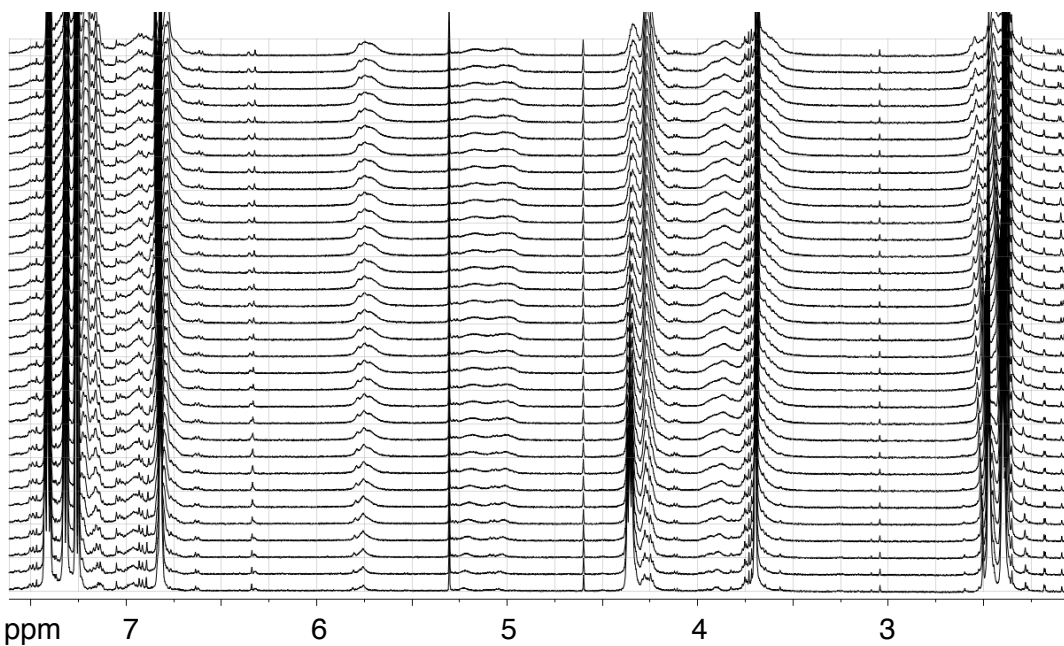
for  $k_{trans}$  and  $k_{cis}$  were assumed to be equal to the rates determined for the reaction of mono-nitrone **155** with maleimide **105** as described above, Section 5.5.2, and SimFit was employed to perform the kinetic simulation. The result of the simulation (Figure 5.12 (i)) showed the major products were the singly reacted intermediate products with a total *trans* concentration of 1.21 mM and a total *cis* concentration of 0.48 mM (Figure 5.12 (ii)) representing an overall conversion of 9% which in comparison to the experimental results of 14% overall conversion was a little lower than was expected indicating that the bisnitrone was slightly more reactive in the 1,3 dipolar cycloaddition.



**Figure 5.12:** Results of kinetic simulation (i) Concentration of all possible unique mono and bis reacted products (ii) Total *trans* and total *cis* concentration of products for comparison to experimental results.

The recognition mediated reaction was next performed by creating a mixture of 10 mM bisnitrone **152** with 20 mM maleimide **34** in CDCl<sub>3</sub> and incubating at 0 °C for 16 h. Our initial attempt at evaluating the results was obstructed by a broadening of product signals, *vide infra* Section 5.6, which made accurate deconvolution not possible.

The broadened product resonances occurred in the region of the spectrum where resonances arising from the *trans* cycloadduct occur so we could infer that the major product of the system was the *trans* cycloadduct, however an exact concentration at each time point could not be accurately determined. In order to overcome this hurdle, the reaction was performed once again under the same conditions this time with an internal standard added. In order to follow the reaction progress in this case, the resonance signal arising from the starting material could be monitored in comparison to the internal standard signal which meant the concentration of starting material



**Figure 5.13:** Stack plot of partial 500.1 MHz  $^1\text{H}$  NMR spectroscopy spectra at 0 °C. The *trans* cycloadduct resonance signals between 5.5 ppm and 6 ppm usually sharp and used for deconvolution are in this case broad and unable to be accurately analysed by deconvolution.

could be accurately determined at each time point. The concentration of starting material at each time point could then be subtracted from the total concentration to find the concentration of total product at each point. 6 mM of 1-bromo-4-*tert*-butylbenzene was chosen as an appropriate internal standard as the *tert*-butyl group had a  $^1\text{H}$  NMR spectroscopy resonance which occurred in an isolated region of the spectrum and also the compound contained no groups which would be able to interact with any of the reagents or products of the cycloaddition reaction.

The concentration vs. time profile for the appearance of *trans*-**158** in the recognition mediated reaction displayed a significant enhancement in product formation (Figure 5.11 (ii)). There was close to quantitative consumption of starting material after 16 h. Encouragingly, the rate of formation of the product in the bisnitron system displayed a sigmoidal shape which closely resembled the rate of production of *trans*-cycloadduct in the mono nitron control system which was demonstrated to be self-replicating. From these results we could deduce that after the first cycloaddition of the bisnitron, there was little or no inhibition of reaction at the second nitron reactive site which could have arisen from the first cycloadduct causing steric inhibition.

To demonstrate that the product was responsible for direction of the reaction, 20 mol % preformed *trans*-**158** was doped into a fresh set of reagents for the recognition mediated reaction and the reaction performed and assayed as before to produce a concentration vs. time profile (Figure 5.11 (iii)). The results from this doping experiment revealed that the reaction proceeded from the maximum autocatalytic rate upon initiation of the reaction, removing the lag period seen in the native recognition mediated reaction and thus the bitemplate must be able to direct the reaction.

The results from kinetic analysis of the bisnitron reactions were encouraging as no problems from linking two nitron reactive sites were encountered and the reactions were still shown to be autocatalytic in nature.

#### 5.5.4 Cross Doping Reactions

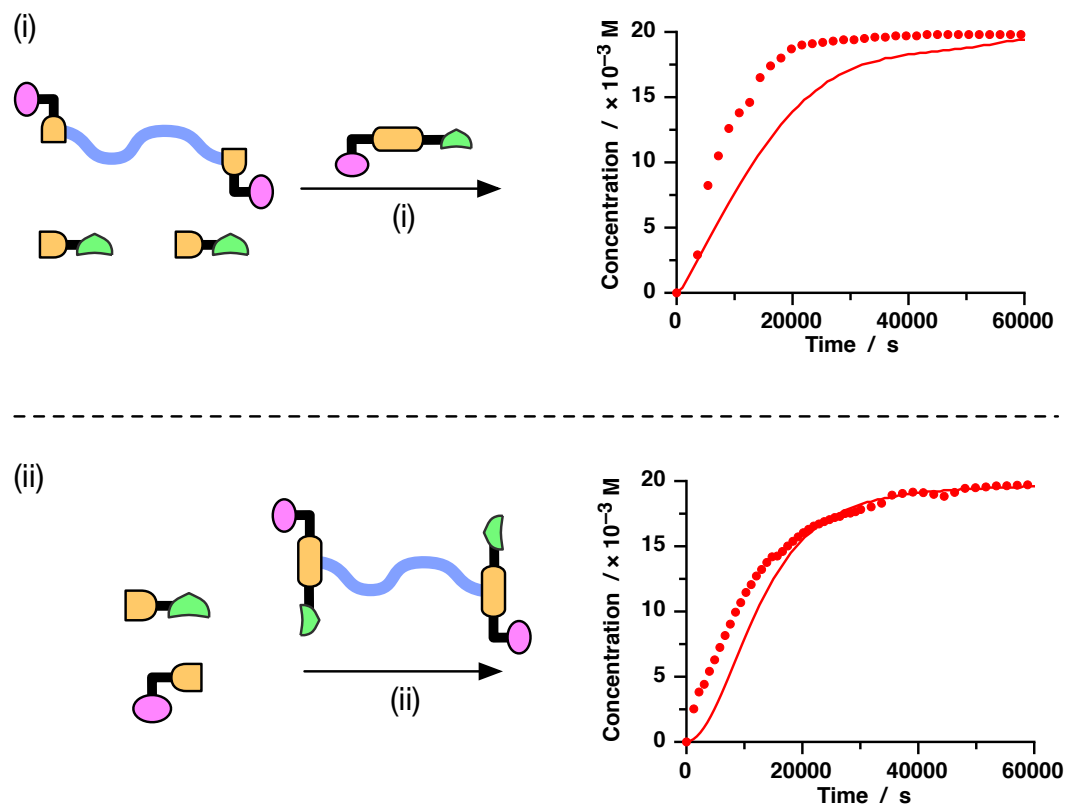
To further demonstrate that the cycloadduct templates were responsible for directing the reaction, two further doping reactions were designed. Firstly, to demonstrate that the a free mono cycloadduct was the responsible template for enhancing the formation of each of the cycloadditions on the bisnitron, preformed mono template *trans*-**156** was added to the reagents for the bisnitron system. In order to perform this reaction, 20 mol% *trans*-**156** was added to the reagents for the recognition mediated bisnitron system and the reaction performed under the same conditions, 10 mM bisnitron **152**, 20 mM maleimide **34** in CDCl<sub>3</sub> at 0 °C and the reaction monitored by <sup>1</sup>H NMR spectroscopy for 16 h.

The concentration vs. time profile constructed after deconvolution (Figure 5.14 (i)) showed that the reaction proceeded at the maximum autocatalytic rate upon initiation demonstrating that the added mono cycloadduct *trans*-**156** had served as a template at the start of the reaction.

Similarly, to demonstrate that the templates on the biscycloadduct were free to direct the cycloaddition of free reagents, 10 mol% of biscycloadduct *trans*-**158** was doped into a fresh set of reagents for the recognition mediated mono nitron system and the reaction performed and monitored under the same conditions. After deconvolution of the 500.1 <sup>1</sup>H NMR spectra, a concentration vs. time profile was constructed (Figure



5.14 (ii)) which showed that addition of the bitemplate has had the effect of increasing the initial rate of reaction, removing the short lag period seen in the native recognition mediated system.



**Figure 5.14:** Concentration vs. time profile for (i) recognition mediated reaction of bisnitron **152** with maleimide **34** doped with 20 mol% mono template *trans*-**156** (ii) recognition mediated reaction of mono nitron **155** with maleimide **34** doped with 10 mol% bitemplate *trans*-**158**. Trans cycloadducts are represented by (●) with the native recognition mediated reaction (—) displayed for reference. Reagents and conditions for (i) CDCl<sub>3</sub>, 0 °C, 10 mM nitron, 20 mM maleimide, 16 h, 20 mol% template doping (ii) CDCl<sub>3</sub>, 0 °C, 20 mM, 16 h, 10 mol% template doping.

From these results, it is clear that the templates formed from 1,3-dipolar cycloaddition were responsible for promoting their own formation. In the case of the bitemplate construction, the close proximity of a template which was covalently linked did not have any effect to inhibit the cycloaddition on the second reaction.

Additionally, the dynamic nature of template duplex formation in the bitemplate system was demonstrated. When preformed bitemplate was added to the reagents for either the mono nitron or bisnitron system it engineered an initial increase in the rate of reaction verifying that the templates are not irreversibly associated together but instead the duplex is able to dissociate and then each template may associate with reagents and promote the reaction.

### 5.5.5 Conclusions From Kinetic Analysis

The results from the kinetic analysis of the mono nitron system confirm that the alterations made to the nitron have not inhibited the autocatalytic nature of its reaction with maleimide **34**. By comparing to systems previously reported, Section 4.5, the mono system described in this section represents a more efficient replicating system as the reaction has a higher *EM* as well as a stronger template duplex.

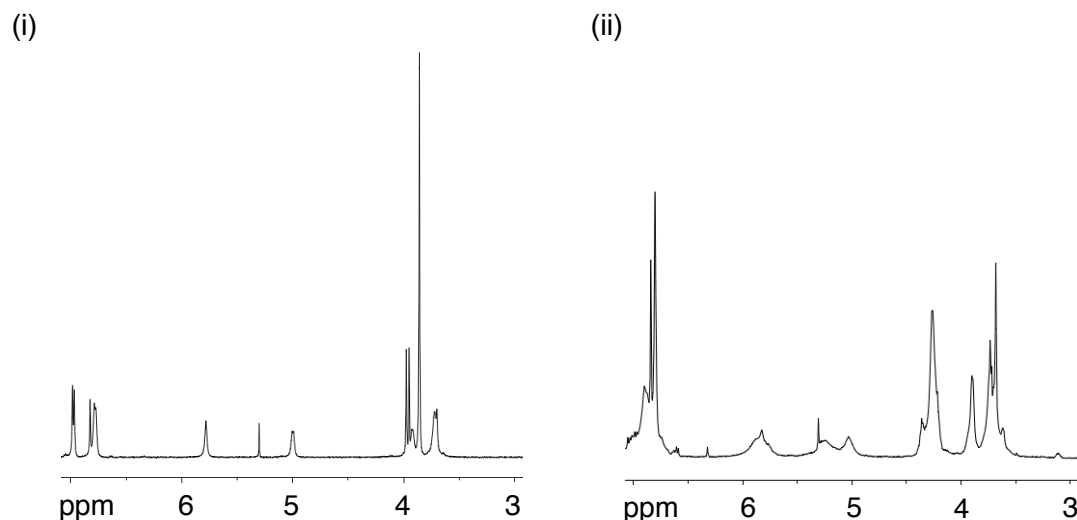
There was originally concern for the reactivity of the bisnitron. Whilst one cycloaddition should proceed readily without any steric interference, after this first cycloaddition to produce one template, the second nitron reactive site is now connected to a large steric group which could disrupt the efficiency of the subsequent second cyclisation however these concerns were dismissed.

Results of the kinetic analysis of the bisnitron system were obscured by broadening of product signals however sufficient evidence that the formation of product in the bisnitron system closely resembles that of the mono system was established which indicated a similar mechanism of formation and dismissed any concerns of reactivity of the dinitone.

### 5.6 Properties of Bistemplate – Evidence for Supramolecular Polymerisation

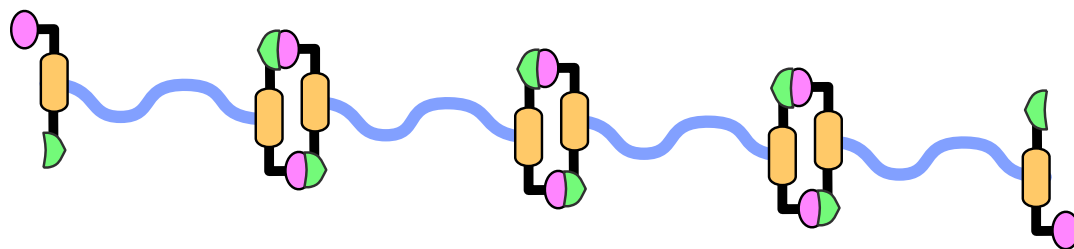
Although the broadening of product resonance signals (Figure 5.15) from <sup>1</sup>H NMR spectroscopy in chloroform obstructed the deconvolution of spectra for kinetic analysis, it was an encouraging observation as evidence for the creation of polymeric products. In chloroform, a non-polar solvent, we would expect that each template will exist as a duplex.

As a consequence of covalently linking two templates, linear polymeric chains of different lengths will be formed (Figure 5.16 (i)). As a result of the slightly different environment of each cycloadduct a broadening of <sup>1</sup>H NMR spectroscopy resonance signals would be observed. In addition to this, on account of the flexible long chain used to link the templates together, there was also the possibility that chains could fold around allowing the ends to create a template duplex and lead to the formation of macrocyclic products (Figure 5.16 (ii)). These would create further unique proton resonance signals further broadening the proton spectrum.

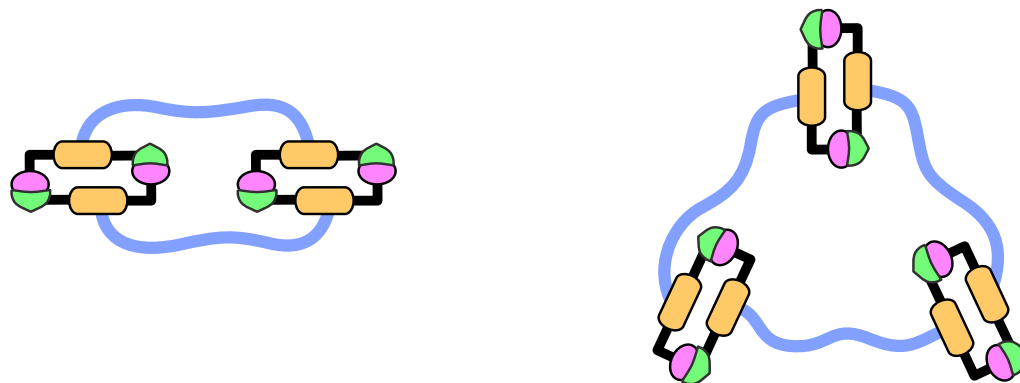


**Figure 5.15:** Partial 500.1 MHz  $^1\text{H}$  NMR spectra at 0 °C of, (i), *mono*-template *trans*-156 showing sharp product resonance signals. (ii), *bitemplate* *trans*-158 showing broadened resonance signals.

**(i) Linear Polymer Formation**

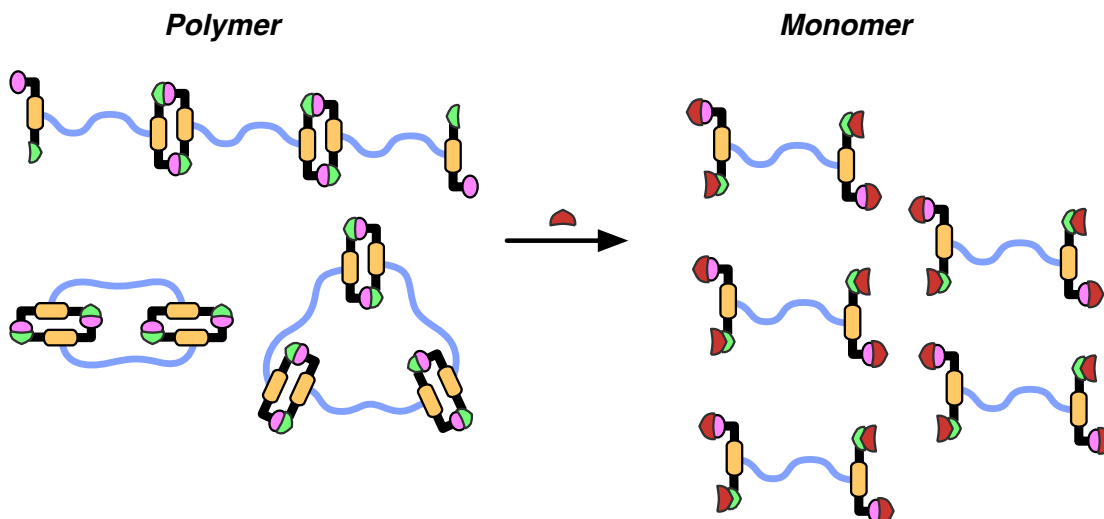


**(ii) Macrocyclic Product Formation**



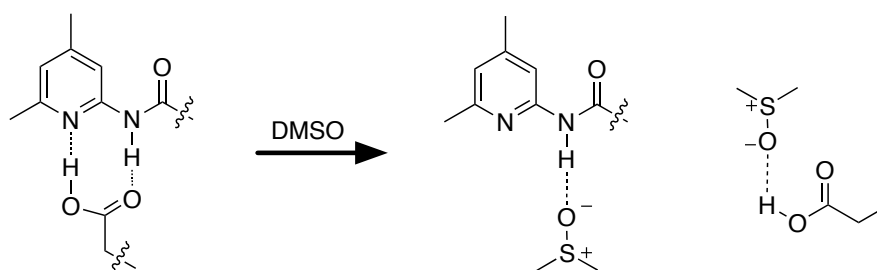
**Figure 5.16:** (i) Representation of linear supramolecular formation with flexible linker represented. (ii) Representation of supramolecular macrocycle formation made possible by flexibility of linker group.

As the interactions responsible for associating monomers together are reversible and non-covalent in nature, these interactions will be sensitive to the environment of the system. By blocking the interactions between monomers, no association will be able to occur and only monomer will exist in solution (Figure 5.17).



**Figure 5.17:** Representation of addition of a competitive binder (DMSO) to the mixture of linear polymers and macrocyclic species. The competitive inhibitor, which is in great excess, disrupts hydrogen bonding between templates by associating with the recognition groups. Therefore only the monomeric species will exist in solution reflected by the appearance of sharp resonance signals in the proton spectrum.

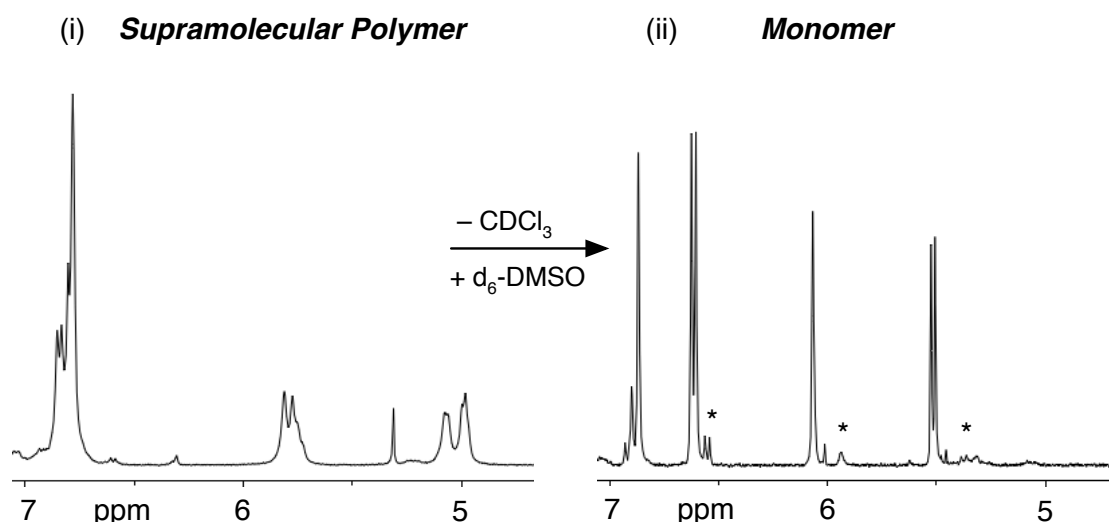
The non-covalent interactions used in this case are hydrogen bonds allowing the templates formed to dimerise in chloroform with a high  $K_a$  of  $2.66 \times 10^7 \text{ Mol}^{-1}$ . Key to the success of this dimerisation is the solvent used. Chloroform is a non-polar solvent which does not disrupt the formation of hydrogen bonds. Changing the solvent the monomers are dissolved in to DMSO, which is more polar and a hydrogen bond acceptor, will create an environment which disrupts the hydrogen bonding between templates (Scheme 5.9). DMSO will be a competitive binder at the recognition sites and as it is present in a very high excess will obstruct template duplex formation. Therefore only monomer will exist in solution.



**Scheme: 5.9 :** Association of the dimethyl amidopyridine group with a phenylacetic acid in non-polar solvent. After addition of DMSO, the interactions are disrupted by competitive binding arising from DMSO molecules.

This environmental effect was experimentally demonstrated by taking freshly produced bitemplate *trans*-**158** which was synthesised in  $\text{CDCl}_3$  at 50 mM and  $0^\circ\text{C}$ . After incubating for two days to ensure quantitative conversion a 400.1 MHz  $^1\text{H}$  NMR

spectrum was recorded (Figure 5.18 (i)) before removing the  $\text{CDCl}_3$  under vacuum and then dissolving the product in  $d_6$ -DMSO at the same concentration. A 400.1 MHz  $^1\text{H}$  NMR spectrum was once again recorded (Figure 5.18 (ii)) which in this case revealed that the broadened signals were not present but instead sharp signals, reminiscent of mono templates, were observed. These results clearly confirmed that the *trans*-**158** was indeed the major product.



**Figure 5.18:** (i) Partial 500.1 MHz  $^1\text{H}$  spectrum of bistemplate at 50 mM in  $\text{CDCl}_3$  at RT showing broadened product resonance signals. (ii) Proton spectrum of bistemplate at 50 mM in  $d_6$ -DMSO showing sharp resonance signals as a result of the disruption of hydrogen bonding between template duplexes enabling a  $^1\text{H}$  NMR spectrum of the monomeric unit to be observed. The \* are shown to draw attention to the second series of *trans* cycloadducts described in the main body.

This proton spectroscopy evidence supports the expected behaviour that in non-polar solvents which allowed the formation of template duplexes, oligomeric chain and macrocyclic products were present in solution whereas changing the environment to a solvent which inhibited template duplex formation lead to the observation of only monomer in solution. One interesting observation is also apparent from the  $^1\text{H}$  NMR spectrum of the bistemplate **158** in  $d_6$ -DMSO and that is of the presence of a series of resonance signals arising from a *trans* cycloadduct which are different to the major product, highlighted in Figure 5.18 (ii) by small ★s.

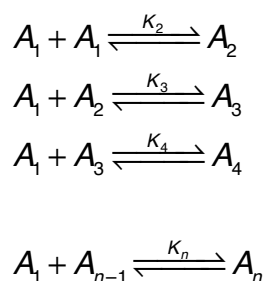
This observation was rationalised by the aforementioned *syn* or *anti* product formation discussed previously (Figure 5.10). Whilst the stereochemistry at each cycloadduct core will be '*trans*' the *syn* and *anti* products will have absolute configurations different to each other and thus a pair of diastereoisomers will have

been produced. These diastereoisomers will show slightly different properties with respect to  $^1\text{H}$  NMR spectroscopy resonance signals and thus may explain the presence of two sets of *trans* cycloadducts. Interestingly the ratio of these signals is unequal with the ratio of 9:1 observed indicating that there is a preference for one of the two configurations. We might expect that the pseudo '*anti*' configuration would encounter the least steric inhibition for reaction and be the preferred product.

## 5.7 Concentration Dependence on Polymer Length

Supramolecular polymers as well as covalent polymers gain their material properties as a consequence of the entanglement of long chains. The extent of supramolecular polymerisation in solution can be described by the degree of polymerisation<sup>201,202</sup> which is the average length, in monomer units, of polymer chain length. For desirable materials properties, this value is preferred to be in the hundreds or thousands.

For supramolecular polymerisation, a theoretical model for the polymerisation of a self-complementary monomer  $A_1$  into linear chains of length  $A_n$  can be created (Figure 5.19).



**Figure 5.19:** Theoretical association model for a self-complementary monomer  $A_1$

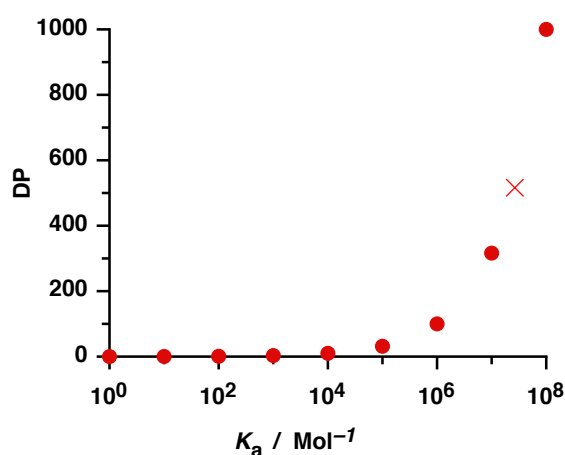
Although the addition of monomer to growing chains is considered sequentially, association may occur through random combination of oligomers. For example, a tetramer ( $A_4$ ) is shown arise from association of a monomer ( $A_1$ ) with a tetramer ( $A_3$ ) although in addition it may arise from a combination of two dimers ( $A_2 + A_2$ ). Since the model considers the thermodynamic equilibrium and not the kinetically most favoured pathway, the selection of expressions does not affect the outcome so long

as all possible species are included. As the model includes an infinite number of unknown parameters ( $K_n$ ) some simplifying assumptions are necessary.

The simplest model of supramolecular polymer growth, the Isodesmic model,<sup>201</sup> assumes that each association between monomers is equal and independent of chain-length ergo  $K_n = K_2$  for  $n > 1$ . Using this assumption, the DP can now be calculated from the dimerisation constant and the concentration of monomer in solution using Equation 5.1 which directly relates DP to the association constant between monomers  $K$  and the concentration of monomers  $C_0$ .<sup>207</sup>

$$DP = \frac{1 + \sqrt{1 + 4KC_0}}{2} \approx \sqrt{KC_0}$$

**Equation 5.1:** The degree of polymerisation (DP) can be calculated knowing the association constant between monomers ( $K_a$ ) and the concentration of monomer ( $C_0$ ) in solution.



**Figure 5.20:** Demonstration of the relationship between the association constant  $K_a$  and the degree of polymerisation DP at a concentration ( $C_0$ ) of 10 mM. The red X displays the calculated DP from the bitemplate duplex association constant.

The relationship between the  $K_a$  and the DP at a concentration of 10 mM is demonstrated in Figure 5.20. At low  $K_a$ s, of  $10^0 \text{ Mol}^{-1}$  up to roughly  $10^5 \text{ Mol}^{-1}$ , the DP is low reaching ~30, too short to form chains with macroscopic properties. With a  $K_a$  above  $10^6$  the value of DP increases dramatically and longer chains may form. From Equation 5.1, the DP of the bitemplate at 10 mM in  $\text{CDCl}_3$  at  $0^\circ \text{C}$  (conditions of kinetic analysis) was calculated to be 516 from the  $K_a$  value of  $2.66 \times 10^7 \text{ Mol}^{-1}$  calculated from the mono nitron system. This DP value falls within of the DP range

desired for supramolecular polymer formation. In comparison, using this equation, the DP of the UPy system was calculated to be 938 (In practise the DP was experimentally determined to be 700 using viscosity measurements. The deviation from the ideal DP length was later attributed to the existence of macrocyclic products).

As the DP is directly proportional to the concentration of the monomer, a successfully created supramolecular polymer should increase in length and molecular size with an increase in monomer concentration. In order to investigate this property we turned to diffusion ordered spectroscopy (DOSY).

### 5.7.1 Diffusion Ordered Spectroscopy

The translational movement of molecules in solution is known as Brownian motion or simply diffusion. The rate of movement in solution depends on the physical properties for example the shape or size of the molecule, temperature and viscosity. When the molecule is assumed to be spherical in shape, the diffusion coefficient  $D$  of a molecule in a solvent can be described by the Stokes–Einstein equation, Equation 5.1.<sup>211,212</sup>

$$D = \frac{kT}{6\pi\eta r_s}$$

**Equation 5.2:** The Stokes–Einstein equation where  $D$  is the diffusion coefficient,  $k$  the Boltzmann constant,  $T$  the temperature,  $\eta$  the viscosity of the solvent and  $r_s$  the hydrodynamic radius of the molecule.

The Stokes–Einstein equation demonstrates that the diffusion coefficient  $D$  is inversely proportional to the radius of the molecule,  $r_s$ . therefore the larger the molecule, the smaller  $D$  will be.

A spectroscopic technique available for measuring the diffusion coefficient ( $D$ ) is diffusion NMR spectroscopy using pulsed field gradient NMR spectroscopy.<sup>211</sup> Using a gradient, a spatially directed short pulse, molecules can be spatially labelled in the NMR spectroscopy sample tube. If these labelled molecules move during the following diffusion time,  $\Delta$ , the spins will change orientation and, upon magnetisation by a second gradient pulse, will acquire a net change in phase. The change in phase will lead to a partial cancellation of the observed magnetisation and therefore

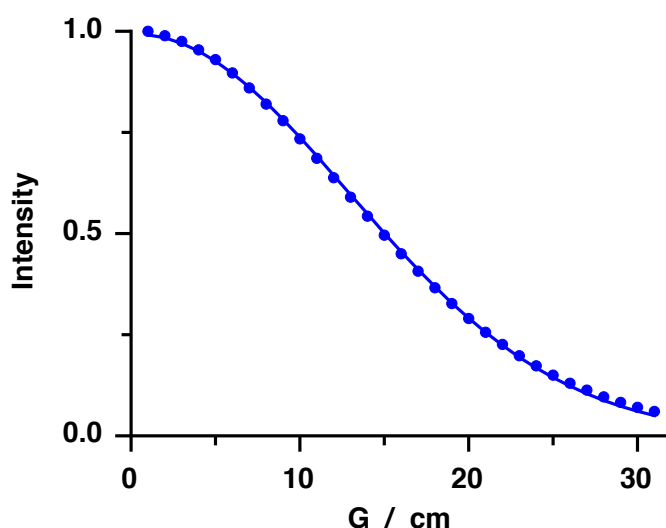


attenuation of the observed signal. The level of attenuation is dependent on the time between gradients ( $\Delta$ ), the strength of the applied gradients ( $g$ ) and the diffusion coefficient of the molecule ( $D$ ) and can be described as in Equation 5.3.

$$I = I_0 e^{-D\gamma^2 g^2 \delta^2 (\Delta - \gamma/3)}$$

**Equation 5.3:**  $I$  is the observed signal intensity,  $I_0$  is the unattenuated signal intensity,  $D$  the diffusion coefficient,  $\gamma$  the gyromagnetic moment of the observed nucleus,  $g$  the gradient strength,  $\delta$  is the length of the gradient and  $\Delta$  is the diffusion time.

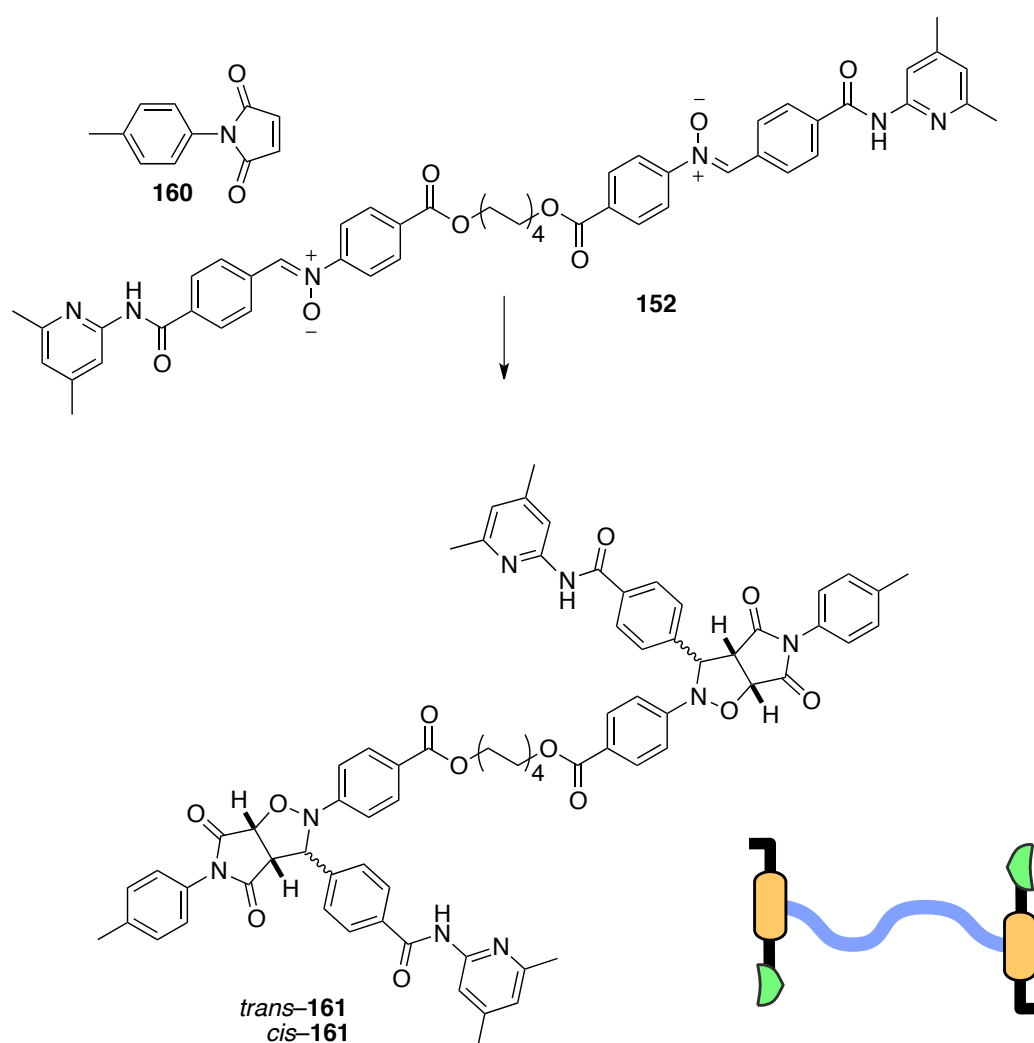
In order to determine  $D$ , a series of  $^1\text{H}$  NMR spectroscopy experiments can be performed which allows the gradient strength to vary from 2% to 95% whilst keeping the length of gradient  $\delta$  and diffusion time  $\Delta$  constant. The results of the experiment show a maximum observed signal intensity at low gradient strength with a decay in intensity as the gradient strength increases leading to a low observed signal intensity at high gradient strength. The results of a typical experiment are displayed in Figure 5.21 (blue points). The attenuation caused by diffusion decay may then be analysed using the SimFit algorithm within Bruker's Topspin software which fits a simulated decay curve (blue line) to the experimental results and from this curve the diffusion coefficient  $D$  is extracted. This technique of analysis is known as diffusion ordered spectroscopy (DOSY).



**Figure 5.21:** Example of Observed Intensity ( $I_0$ ) vs. gradient point graph showing the decay of signal intensity with increasing gradient strength. Experimentally observed results displayed as filled circles and results of fitting using SimFit package within Topspin 2.0 shown as a line.

Larger molecules which diffuse slowly have a smaller diffusion coefficient in comparison to smaller molecules which diffuse faster and therefore have a larger diffusion coefficient. Equation 5.1 illustrates that the concentration  $C$  of the monomer

is directly proportional to the degree of polymerisation DP. Therefore by increasing the concentration of monomer in solution, we expect that the molecular size will increase. As the molecular size increases, the diffusion coefficient should D should decrease. Therefore by increasing the concentration of monomer in solution, if supramolecular polymerisation is occurring, the molecular size should increase leading to slower translational motion and therefore the diffusion coefficient D should decrease. In order to investigate for this change in D with concentration we carried out a series of experiments designed to investigate for a change in molecular size with a change in concentration of polymer.



**Scheme 5.10:** Control bitemplate used for DOSY experiments designed to be the same size as the monomer unit but is unable to dimerise an account of the removal of the acid association site. Reagents and conditions (i)  $\text{CDCl}_3$ , RT, 1 week, quantitative.

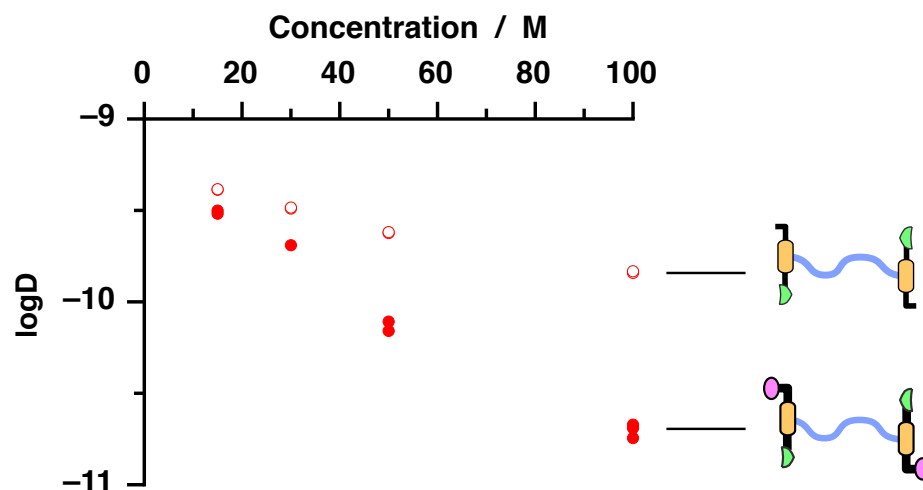
Bisnitronone **152** was dissolved in CDCl<sub>3</sub> at concentrations of 15 mM, 30 mM, 50 mM and 100 mM along with two equivalents maleimide **34** and incubated in NMR spectroscopy tubes at 0 °C, to allow for complete efficient and selective autocatalytic production of *trans*-**158**, and left for 3 days to ensure complete conversion. Once complete, each sample of monomer at a different concentration was analysed by 2D DOSY NMR spectroscopy to determine the diffusion coefficient with the expectation that the diffusion co-efficient should decrease with an increase in concentration.

However, simply increasing the concentration of a substance in solution will have an effect on the viscosity of the solution and hence the rate of molecular diffusion and as a result, in addition to the samples of monomer, an appropriate control series was designed to account for any change in diffusion coefficient with concentration. The control compound would need to be the same size as the monomer, to reproduce their effect on the solution simply by increasing concentration, but not possess the ability to polymerise.

The reaction of bisnitronone **152** with the recognition disabled maleimide **160** will produce a bitemplate (mixture of *trans* + *cis*) which has the same dimensions as the monomer and as such is a good representation for use as a control compound. The recognition disabled bitemplates were produced in a similar manner as before, by mixing a 2:1 ratio of maleimide **160** with bisnitronone **152** at the appropriate concentration (15 mM, 30 mM, 50 mM and 100 mM) in CDCl<sub>3</sub> in NMR spectroscopy tubes and allowed to react for one week at room temperature to ensure quantitative conversion before assaying by 2D DOSY NMR spectroscopy to determine the diffusion co-efficient. The expectation from this control would be that the diffusion co-efficient would be roughly the same value or decrease slightly with increasing concentration as a consequence of amide aggregation at increasing concentrations.

### 5.7.2 Results of DOSY Experiments

The results of the DOSY experiments of the recognition disabled control monomer (Figure 5.22 open circles) demonstrated that the diffusion coefficient decreased slightly with increasing concentration in line with our expectations.



**Figure 5.22:** Concentration vs. logD from the results of DOSY NMR spectroscopy experiments with logD of polymer (●) and control monomer (○)

The results of the DOSY experiments of the recognition enabled monomer (Figure 5.22 filled circles) demonstrate some important results.

Firstly, at all concentrations, the diffusion coefficient of the recognition enabled monomer is lower than the recognition disabled control monomer. This result indicates that the average molecular size of the recognition enabled monomer is larger than the size of the recognition disabled control monomer. In order to achieve this, the monomeric units of the polymer must successfully be associating to form a supramolecular polymer with chains longer than the monomer.

Secondly, the diffusion coefficient of the recognition enabled monomer decreases with increasing concentration indicating the formation of longer chains at higher concentrations fulfilling the expected properties of a supramolecular polymer following Equation 5.1.

Finally, the difference in diffusion coefficient between the recognition enabled monomer and recognition disabled control monomer increases with concentration. For example, at 15 mM, logD of the recognition disabled control monomer and recognition enabled monomer is  $-9.39$  and  $-9.53$  respectively, having a difference of  $0.14$ . At 100 mM however, logD of the recognition disabled control monomer and recognition enabled monomer is  $-9.84$  and  $-10.70$  respectively with a difference now of  $0.86$ . This observation that the diffusion coefficient increases with concentration

relative to the control monomer, demonstrates that the average supramolecular polymer chain length (and hence DP) is increasing with concentration obeying the expected behaviour from Equation 5.1. As the concentration decreases, the diffusion coefficients of the recognition disabled control monomer and recognition enabled monomer converge. This can be expected as at infinite dilution, the dimerisation of recognition enabled monomers will not occur, instead each monomer will exist as an isolated molecule and should show behaviour identical to the recognition disabled control monomer which is unable to dimerise.

The results of the DOSY experiments confirmed our expectation that the recognition enabled monomers formed by autocatalytic self-replication were able to spontaneously assemble into polymeric chains and thus demonstrate the creation of a supramolecular polymer by autocatalytic self-replication.

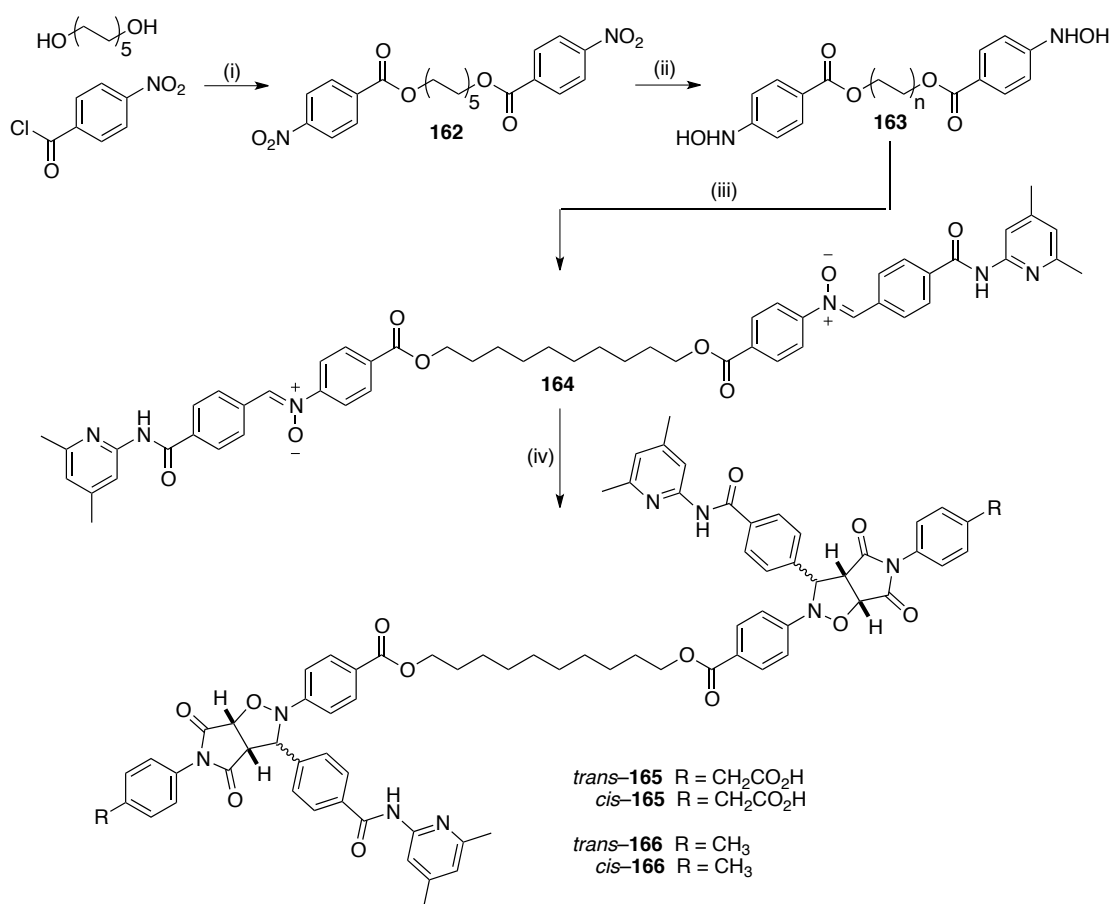
### 5.7.3 Extension of Linker

The C<sub>8</sub>H<sub>16</sub> linker used to tether the two templates together had a positive effect on increasing the solubility of the reagents and products in the non-polar solvent used, chloroform. One drawback which may be envisaged from using such a flexible linking group is that the templates may not form long chains but instead loop round to form rings of varying sizes. These rings will be smaller than the polymeric chains, lowering the average molecular size and increasing the diffusion constant. By increasing the flexibility of the linking group, we allow greater conformational freedom which will allow the formation of rings. The only way to increase the flexibility of the C<sub>8</sub>H<sub>16</sub> chain is to increase the chain length. By increasing the chain from eight to ten carbons (C<sub>10</sub>H<sub>20</sub>), we increase the conformational flexibility of the linking group. If adding two extra carbon groups to the linker does not facilitate the formation of small rings rather than chains then we would expect that the polymer size would increase as a result of a longer linker group. This in turn would lead to a smaller diffusion constant relative to the control monomer.

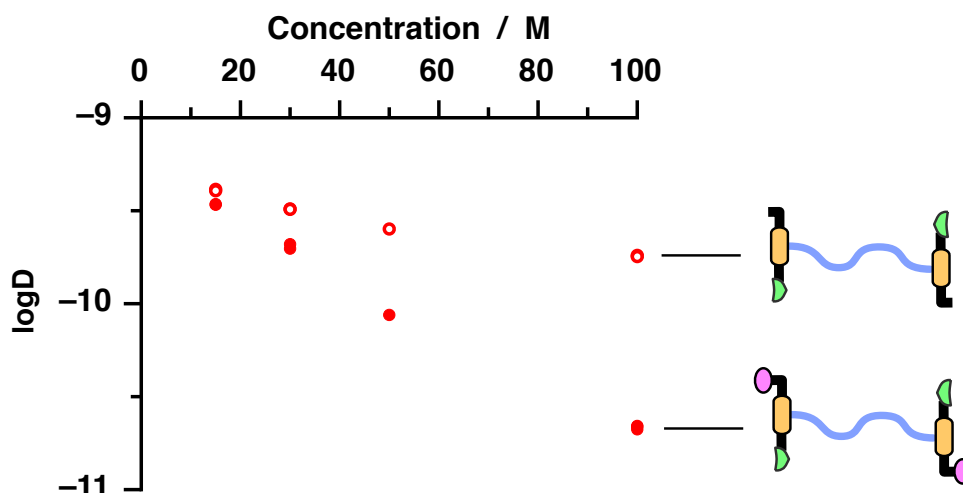
To investigate this effect, a second monomer was designed, *trans*-**165**, in which the linker group was extended from eight to ten methylene groups and synthesised in a route akin to the eight carbon linker reported earlier (Scheme 5.11). The longer linker may either facilitate a greater conformational freedom leading to a greater incidence

of rings instead of chains or have little difference in the ring vs. chain equilibrium and instead lead to longer average polymer size as a result of the longer linker. Whilst we cannot observe this directly, we might expect to see a change in the relative diffusion constant  $D$  difference between the recognition enabled monomer in comparison to the recognition disabled control monomer from which the behaviour may be decoded.

In order to investigate this relative change in diffusion coefficient with this longer monomer, a series of samples of recognition enabled monomer were prepared as described earlier, at concentrations of 15 mM, 30 mM, 50 mM and 100 mM and subjected to analysis by 500.1 MHz  $^1\text{H}$  NMR spectroscopy DOSY analysis. In order to make an accurate comparison, the recognition disabled control monomer of the ten carbon linker was also prepared at these concentrations after reacting 2 equivalents of maleimide **160** with bisnitron **165** to furnish **166** as a mix of *trans* and *cis* isomers.



**Scheme 5.11:** Reagents and conditions (i) NEt<sub>3</sub>, CH<sub>2</sub>Cl<sub>2</sub>, 25 °C, 16 h, 60%, (ii) Rh/C, NH<sub>2</sub>NH<sub>2</sub>·H<sub>2</sub>O, THF, 25 °C, 20 mins, quantitative, (iii) 3 eq. Aldehyde **89**, EtOH, RT, 72 h, 63%, (iv) 2 eq. Maleimide **34**, CDCl<sub>3</sub>, 0 °C, 48 h, quantitative or 2 eq. Maleimide **160**, CDCl<sub>3</sub>, 0 °C, 1 week, quantitative.



**Figure 5.23:** Concentration vs. logD from the results of DOSY NMR spectroscopy experiments with logD of polymer (●) and control monomer (○)

The results of the DOSY analysis of the recognition disabled control monomer (Figure 5.23 open circles) once again revealed a subtle decrease in diffusion coefficient with increasing concentrations. This behaviour was consistent with the previous results and expected to be as a result of amide aggregation at higher concentrations.

The results of the DOSY analysis of the recognition enabled monomer (Figure 5.23 filled circles) revealed once again a decrease in diffusion coefficient with increasing concentration. Comparison of the recognition disabled results to the recognition enabled results once again show that, relative to the recognition disabled control monomer, the diffusion coefficient of the recognition enabled monomer decreases showing that it exists as longer chains than the monomer confirming the expected behaviour for a supramolecular polymer.

A closer look at the relative difference in diffusion coefficient at 100 mM can give insight into the different behaviour of the supramolecular polymer imparted by the longer linker. For the control monomer, logD at 100 mM was  $-9.74$  and for the recognition enabled monomer, logD at 100 mM was  $-10.68$ . This means that the relative difference in logD between the control monomer and the supramolecular polymer was 0.93. In comparison to the eight carbon linker (difference of logD is 0.86 at 100 mM) the ten carbon linker monomer has a greater difference in diffusion coefficient in comparison to the control monomer. This means that average molecular

size of the ten carbon linker monomer must be greater relative to the monomer than the eight carbon linker. This result demonstrates that increasing the length of linker from eight to ten carbon atoms does not decrease the average polymer size as a result of flexibility forming rings and decreasing average size but instead the effect is to increase the average polymer size by increasing the length of the monomer.

## 5.8 Physical Properties of Supramolecular Polymers

Whilst the experimental results provide evidence for the existence of supramolecular polymerisation in solution, in order for the process to have commercial application, the supramolecular polymer must be able to exist and be stable in solid state therefore forming a plastic material.

In order to investigate whether the supramolecular polymer was stable in the solid state we attempted to create a film of the material. In order to achieve this, a 5 mL solution of monomer at 50 mM in  $\text{CHCl}_3$  was poured into a glass petri dish which was covered and the solvent allowed to slowly evaporate at room temperature. This process took roughly 24 h and produced a thin film of a glassy plastic material across the whole surface of the petri dish. For comparison, when a solution of recognition disabled monomer was spread on a surface and the solvent allowed to evaporate as with the recognition enabled monomer, no film was observed but instead the molecule formed a powder.



**Figure 5.24:** Photographs of a sample of polymer film cut from the film formed on the petri dish. Film placed on end of a spatula for size reference.



The polymeric film formed was semitransparent with a pale yellow colouring. The film was thin and brittle but was able to be cut from the petri dish to be observed (Figure 5.24).

One of the desirable properties of a supramolecular polymer is that they may be handled as a solid, dissolved by a solvent into solution, and then for example transported or poured into a new shape mould before removing the solvent to reproduce the polymer thus taking advantage of the dynamic nature of the non-covalent linkers in the supramolecular polymer. In order to attempt to fulfil this property, after producing the thin film of polymer as shown in Figure 5.24, the film was dissolved in chloroform and then poured into a glass petri dish again and the solvent allowed to evaporate overnight. After this time, the thin polymer film was reproduced.

At this stage, the polymeric film was not investigated further but it would be very interesting to attempt to produce thicker films of polymer by attempting to pour a concentrated solution of monomer into a mould to attempt to produce plastic shapes.

## 5.9 Conclusions

The formation of supramolecular polymers requires a repeating monomer of two groups capable of dimerisation by strong, highly directional secondary interactions. After fabrication, the replicating templates discussed previously exist as dimers in solution with a strong association constant of a similar magnitude to groups successfully used for supramolecular polymer formation. Two replicating templates were successfully linked together utilising a recognition bearing bisnitro substrate which could autocatalytically react with a recognition bearing maleimide to form a bitemplate monomer which spontaneously polymerised in solution under the same conditions as the cycloaddition reaction used for its construction.

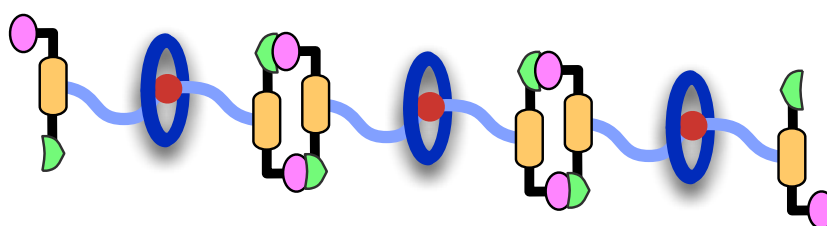
The existence of polymeric structures was confirmed by two methods. Firstly the polymer was demonstrated to be sensitive to a change in environment. Changing the solvent from non-polar chloroform, which allowed dimerisation of monomers, to DMSO, which inhibited dimerisation of monomers, showed a characteristic sharpening of proton resonance signals as a result of the disruption of polymeric

units formed in chloroform now existing as monomers in DMSO. Secondly, DOSY experiments were employed to demonstrate that the average molecular size of the polymer was larger than a monomeric control compound over a range of concentrations and that the size of the polymer increased relative to the monomeric control with an increase in concentration thus following the expected behaviour from a simple supramolecular polymerisation model.

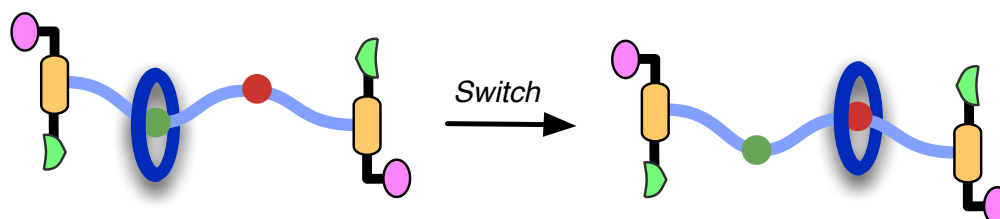
The advent of supramolecular polymers has given rise to a new area of materials science with the potential to create a new class of polymers with unique and interesting properties. The major requirement for the creation of supramolecular polymers is the synthesis of the repeating monomeric unit which can potentially be difficult. Here, the repeating monomeric unit for a supramolecular polymer has been created using autocatalytic self-replication thereby reducing the synthetic burden on monomer formation.

The design of the repeating monomer allows for the linking group, here an alkyl diester, to be changed to explore different properties. One immediate future project should be the creation of a library of different linking groups and a thorough examination of the dependence on linker with their polymeric properties.

(i) **Polyrotaxane**



(ii) **Switchable Polyrotaxane**



**Figure 5.25:** (i) Representation of the formation of a polyrotaxane with the macrocycle (dark blue ring) docked on to a binding site (red circle). (ii) Evolution of the polyrotaxane into a bistable polyrotaxane with two binding sites could give rise to a switchable rotaxane formed by autocatalytic self-replication.

Functionality could also be incorporated into the linker of the bitemplate in the near future such as the incorporation of a rotaxane binding site which could give rise to the formation of polyrotaxanes (Figure 5.25 (i)). Evolution of these structures could lead to creation of a bistable rotaxane (Figure 5.25 (ii)), such as the Stoddart example discussed earlier, which could create a switchable polymeric material formed by autocatalytic self-replication.

With the presentation of a supramolecular polymer formed by self-replication, the dream of applying self-replication to a manufacturing process may well become a more realistic possibility.

# 6

## Overall Conclusions

### 6.1 Conclusions

The field of systems chemistry may conceptually be in its infancy, only having established its identity over the past four years, however the concepts which it embraces have had much longer to evolve.

Self-replicating molecules have appeared in the literature for the past 24 years since von Kiedrowski published the first self-replicating hexanucleotide. Such molecules carry information which allow them to direct reagents to form an identical product. Early examples of self-replicating systems based on DNA suffered from problems of product inhibition which prevented efficient autocatalytic turnover however modification of these systems was to lead to the observation of parabolic growth. Since then, systems based on peptides, RNA and small molecules have appeared which overcome strong product inhibition and almost exponential growth and self-sustained growth have been demonstrated. The investigation of self-replicating molecules may, for the replicator-first theorists, offer an insight into the behaviour of molecules which may have evolved selectively produce themselves from the primordial soup and carry genetic information. Investigation of systems in which multiple replicators are present appear in the literature exploiting the structural similarity of templates to investigate cooperation between network members leading to system level behaviour. Within this thesis, networks of synthetic replicators were created which were shown to coexist, that is, no one replicator completely dominated the product pool. This coexistence of replicators meant that in such systems the Darwinian notation of survival of the fittest was not observed. However, the template directed nature of self-replicating reactions meant that these systems were able to be instructed with a template to direct the formation of itself and thus increase selectivity in the system.

In order to increase the selectivity of products in the systems, we turned to another avenue of systems chemistry which is itself well established, dynamic covalent chemistry. By coupling the reagent pool to a dynamic process, the replicating templates were able to influence the distribution members in the dynamic library, driving it away from equilibrium and thus leading to a greater degree of selectivity of the desired product from the system. In order to achieve greater selectivity from systems of replicators, we must look to move away from kinetically controlled reactions and towards other methods of which to control the reaction such as reactions under thermodynamic control or, as is currently being investigated in our laboratory, the implementation of travelling wave reactions which spatially and temporally separate the products of a system.

The small synthetic molecule approach to reciprocal replication has very little prominence in the literature. The only examples from Rebek and our laboratory suffer from more reactive side products consuming the reagents instead of allowing them to produce an efficient reciprocal replicating system. In order overcome these problems, we designed the reagents for a reciprocal replicating which were of orthogonal reactivity. This meant that only the desired reactions to produce the complementary templates were allowed to proceed and this change meant that any side reactions were extinguished and efficient reciprocal replication was observed. This system also demonstrated the control of both a thermodynamic and a kinetic reaction within the same system as the dynamic imine condensation reaction produced a template capable of directing an irreversible 1,3-dipolar cycloaddition.

If we are to be able to integrate replication processes into the fabrication of higher ordered superstructures, held together by non-covalent recognition, we must seek to attain an in depth understanding about how non-covalent recognition can be used to direct reactions and refine them to optimise both reactivity and selectivity. Numerous examples in the literature use non-covalent recognition to pre-organise reagents into a reactive binary complex with a conformation which enhances the local concentration of the reactive sites leading to an increased rate of reaction and selectivity of the product. In such a system, the maximum rate of reaction is at the start of the reaction, whereas in contrast, in a self-replicating reaction the maximum rate of reaction occurs after a lag period. During this lag period, the self-replicating reaction is not template directed and as such is relatively slow and prone to mistakes.

We postulated that a bifunctional system, designed to take advantage of the advantageous properties of both a binary complex reaction and a self-replicating reaction, would control the selectivity of the reaction at all stages. Implementation of this strategy indeed yielded a system with an increased reactivity and selectivity and demonstrated that more than one recognition mediated behaviour can be rationally programmed into a single system.

Finally, the application of self-replication into an engineering aspect of chemistry was explored. Supramolecular polymers, long chains of repeating monomers held together by reversible non-covalent recognition, are an expanding field of polymer science research with the promise of producing “smart” materials. These polymers would be able to respond to environmental changes as a result of the dynamic interactions that hold them together. This field has blossomed over the past 15 years with hydrogen bonding based supramolecular polymers. We recognised that the template dimers our group produce by self-replication had an association strength comparable to those reported to associate supramolecular polymer monomers. Our contribution to this field was to link two replicating templates together to facilitate their dimerisation resulting in long chain formation. As the self-replicating reaction was undertaken under the same conditions in which the supramolecular polymer forms, this was the first example of a truly self-synthesising nano-material. Further investigation of this exciting application of self-replication to form supramolecular polymers could tune the macroscopic properties of material by altering the chemical composition of the linking units or recognition groups.

Nature, the most complex of systems, has had billions of years to evolve the complex processes which are employed for the tasks it requires to complete. In contrast our knowledge of the synthetic chemistry and non-covalent molecular recognition chemistry we exploit to attempt to mimic natural processes has only been obtained over the past hundred years and fifty years respectively. With such a vast amount of information already gathered, there is little doubt that man’s aspiration to expand our chemical knowledge will continue to produce ever more sophisticated chemical processes in time to come. With a greater knowledge of recognition mediated processes at the molecular level, and with nature as inspiration, the burgeoning fields of systems chemistry and supramolecular chemistry will offer up a wealth of information and methods to design, produce and control functional materials.



# 7

## General Experimental and Characterisation of Compounds

### 7.1 General Experimental Procedures

**Chemicals** were purchased from ABCR GmbH & Co, Alfa Aesar, Apollo Scientific Ltd., Fisher Scientific UK Ltd., TCI UK Ltd., Sigma–Aldrich Company Ltd. or VWR International Ltd. and purified by standard techniques,<sup>213</sup> where necessary. Where appropriate and if not otherwise stated, all non-aqueous reactions were carried out under an inert nitrogen or argon atmosphere with the inert gas passing through a bed of 4 Å molecular sieves and self indicating silica gel. Brine refers to a saturated aqueous solution of sodium chloride.

**Continuous flow hydrogenation** reactions were conducted on a ThalesNano Nanotechnology Inc. H-Cube® Continuous-Flow Hydrogenation Reactor in solutions of HPLC grade methanol (0.2 M) using a 30 mm 10% Pd/C CatCart® catalyst cartridge.

**Anhydrous solvents** were obtained under the following conditions: dry THF was distilled from sodium in a recycling still using benzophenone ketyl as an indicator; dry acetonitrile was distilled from calcium hydride in a recycling still; dry CH<sub>2</sub>Cl<sub>2</sub> was obtained using a MBRAUN GmbH MB SPS-800 solvent purification system, where solvent was dried by passage through filter columns and dispensed under an atmosphere of N<sub>2</sub> or Ar gas; CDCl<sub>3</sub> was neutralised over CaCl<sub>2</sub> filtered and then stored 4 Å activated molecular sieves.

**Analytical Thin Layer Chromatography** (TLC) analysis was performed on MACHEREY-NAGEL GmbH & Co. POLYGRAM® SIL G/UV<sub>254</sub> plates, plastic backed 0.2 mm silica gel plates with fluorescent indicator. Developed plates were air-dried



and visualised under UV lamp ( $\lambda_{\max}$  254 or 366 nm) or after incubating with iodine on sand. Where necessary, thermal development after dipping in: methanolic DNP and sulfuric acid solution, ninhydrin in *n*-butanol or a solution of aqueous potassium permanganate, potassium carbonate and sodium hydroxide, was used to aid visualisation.

**Flash Column chromatography** and silica plugs were carried out on Apollo Scientific Ltd. silica gel 40–63 micron or Silicycle SiliaFash® P60 silica gel (230–400 mesh) eluting with solvents as supplied under a positive pressure of compressed air.

**Melting points** were determined using an Electrothermal 9200 melting point apparatus and are uncorrected.

**Mass spectra** were recorded on a Micromass GCT spectrometer for electron impact ionisation (EI) operating at 70 eV or chemical ionisation (CI) using isobutane as the ionising gas. Electrospray ionisation spectra (ES) were performed on a Micromass LCT spectrometer operating in positive or negative mode from solutions of methanol, acetonitrile or water. *m/z* values are reported in Daltons and followed by their percentage abundance in parentheses.

**Elemental analysis** were performed on a Carlo Erba EA1110 CHNS analyser by the University of St Andrews microanalysis service.

**Molecular modelling** calculations were performed on a Linux workstation using Macromodel (Version 9.5, Schrödinger Inc., 2008) as part of the Maestro program. Each model was initially energy minimised using the MMFFs forcefield with the Polak-Ribiere Conjugate Gradient (PRCG) method and under the GB/SA solvation model for chloroform. The energy minimised structures were then further investigated by running a Monte-Carlo Multiple Minimum (MCOMM) conformational search method to minimise 1000 iterations using the MMFFs forcefield and GB/SA solvation model for chloroform. Transition states were successfully located by semi-empirical calculations using the RM1 method as implemented in MOPAC2009 (version 10.014L) running on a Linux cluster.

**<sup>1</sup>H NMR** spectra were recorded on a Bruker Avance 500 (499.9 MHz), a Varian UNITYplus 500 (500.1 MHz), a Bruker Avance II 400 (400.1 MHz) or a Bruker Avance 300 (300.1 MHz) spectrometer using the deuterated solvent as the lock and the residual solvent as the internal reference in all cases. In the assignment of <sup>1</sup>H NMR spectra the chemical shift information ( $\delta_{\text{H}}$ ) for each resonance signal is given in units of parts per million (ppm) relative to trimethylsilane (TMS) where  $\delta_{\text{H}} \text{ TMS} = 0.00$  ppm. The number of protons (n) for a reported resonance signal are indicated as nH from their integral value and their multiplicity by the symbols s, d, t, q, m and br which denote singlet, doublet, triplet, quartet, multiplet and broad singlet respectively. Their coupling constants (J) are determined by analysis using iNMR© quoted to the nearest 0.1 Hz. Identical coupling constants are averaged in each spectrum and are reported to the nearest 0.1 Hz.

**<sup>13</sup>C NMR** spectra were recorded on a Bruker Avance 500 (125.7 MHz), a Bruker Avance II 400 (100.6 MHz) or a Bruker Avance 300 (75.5 MHz) spectrometer using the CPD or DEPTQ pulse sequences with broadband proton decoupling using the deuterated solvent as the lock and the residual solvent as the internal reference in all cases. The chemical shift information ( $\delta_{\text{C}}$ ) for each resonance signal are given in units of parts per million (ppm) relative to trimethylsilane (TMS) where  $\delta_{\text{C}} \text{ TMS} = 0.00$  ppm. All signals are singlets unless stated, in which case their multiplicity is represented by the symbol d for doublet.

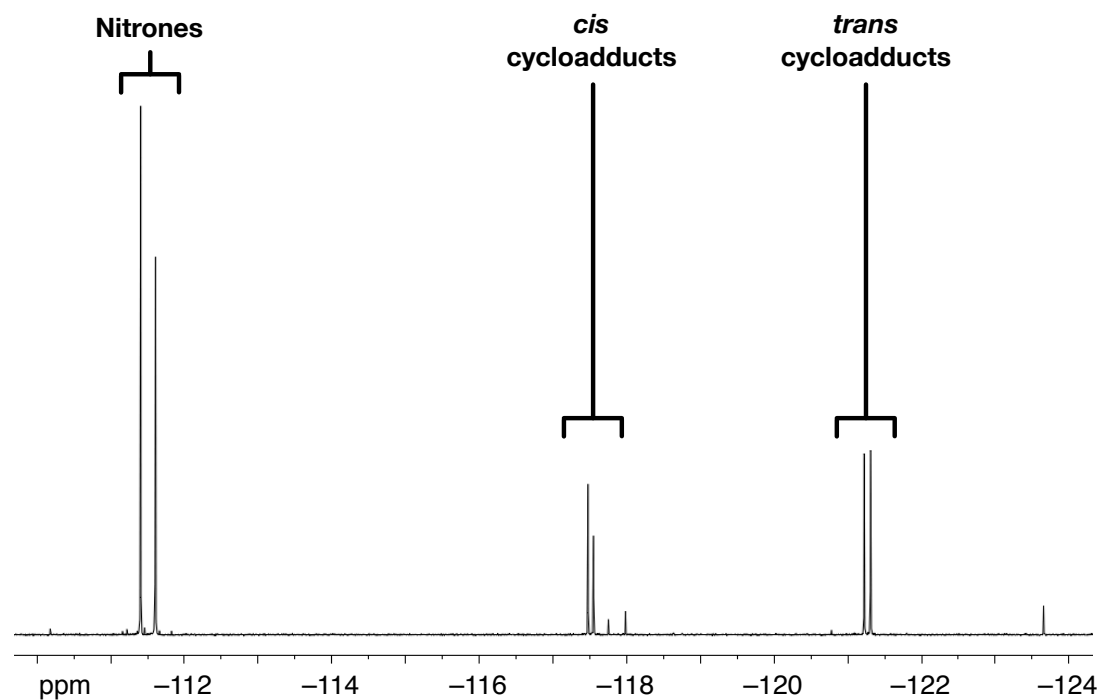
**<sup>19</sup>F NMR** spectra were recorded using a Bruker Avance 500 (470.3 MHz), a Bruker Avance II 400 (376.5 MHz) or a Bruker Avance 300 (282.4 MHz) spectrometer using a broadband proton decoupling pulse sequence with the deuterated solvent as the internal lock. The chemical shift information ( $\delta_{\text{F}}$ ) for each resonance signal are given in units of parts per million (ppm) relative to CCl<sub>3</sub>F where  $\delta_{\text{F}} \text{ CCl}_3\text{F} = 0.00$  ppm.

#### **Preparation of Exchange Medium ( CD<sub>2</sub>Cl<sub>2</sub> / *p*-TSA (sat) )**

99.96 atom % D CD<sub>2</sub>Cl<sub>2</sub> was purchased from Sigma-Aldrich, opened and used immediately. 10 g of CD<sub>2</sub>Cl<sub>2</sub> was stirred vigorously with 0.5 g *para*-toluenesulfonic acid monohydrate under a nitrogen atmosphere for 30 minutes. The suspension was then filtered and the solvent used immediately.

## 7.2 Preparation and Analysis of Two or Three Replicator Experiments in Chapter 2

A standard solution of the two or three nitrones at 20 mM was prepared by dissolving the appropriate nitrones (*m*-nitrone, **68**, *p*-nitrone **52** and *l*-nitrone **83**) in CD<sub>2</sub>Cl<sub>2</sub> / *p*-TSA (sat) which was equilibrated to the reaction concentration (10 °C or 25 °C). Likewise a standard solution of the maleimide (**29** or **72**) at 20 mM was prepared by dissolving the maleimide in CD<sub>2</sub>Cl<sub>2</sub> / *p*-TSA (sat) which was equilibrated to the reaction temperature (10 °C or 25 °C). Pre-synthesised template for doping reactions was individually measured and the reaction initiated by mixing 0.4 mL of the nitrone solution with 0.4 mL of the maleimide solution to give 0.8 mL reaction solution at 10 mM into a vial with or without template, transferred to an NMR spectroscopy tube and sealed with a polyethylene pressure cap to avoid solvent evaporation. In order to assess the exact concentration of added template, the reactions were immediately assayed by <sup>19</sup>F{<sup>1</sup>H} NMR spectroscopy. The reactions were then placed into a thermostatically controlled water bath for the duration of the reaction and assayed by <sup>19</sup>F{<sup>1</sup>H} NMR spectroscopy. The relative concentration of

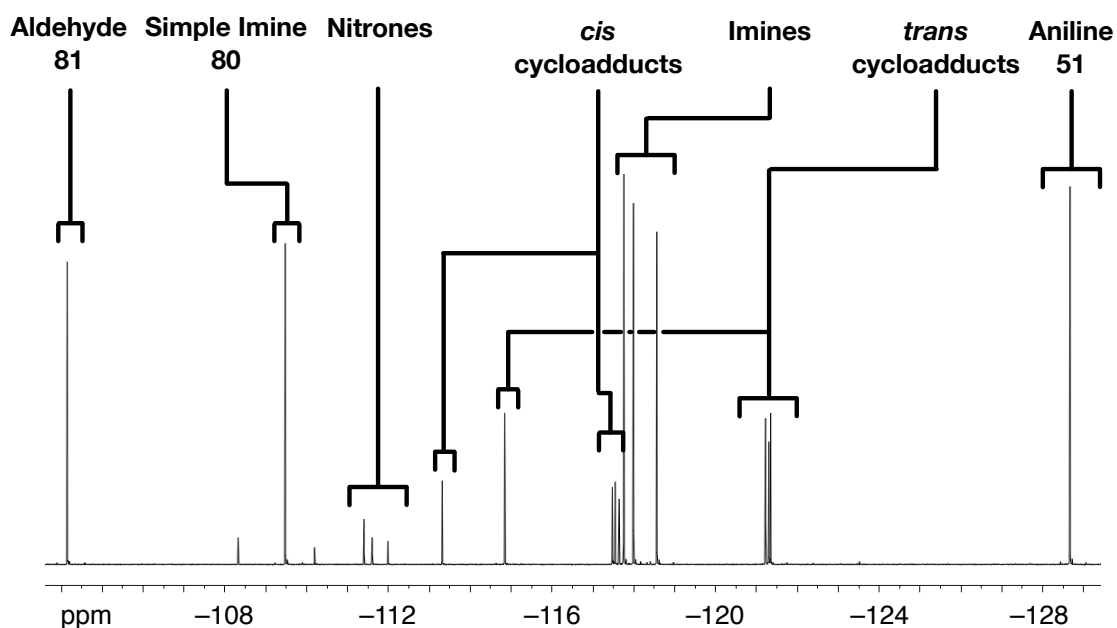


**Figure 7.1:** Partial <sup>19</sup>F{<sup>1</sup>H} NMR spectrum, 376.5 MHz, CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat), 25 °C of a the two replicator experiment at completion showing the characteristic region for nitrone, *cis* and *trans* cycloadduct resonance signals.

each product in each spectra was then calculated using the deconvolution module of iNMR© versions 3.0.1→3.6.2 (current version). The total concentration was determined by summing the integral values for resonance signals for starting materials and products with the exact concentration of each species determined relative to the total concentration.

### 7.3 Preparation and Analysis of Two or Three Replicators Coupled to a Dynamic Process Experiments in Chapter 2

A standard solution of the simple nitrone **65** at 20 mM was prepared by dissolving it in CD<sub>2</sub>Cl<sub>2</sub> / *p*-TSA (sat) which was equilibrated to the reaction concentration (10 °C or 25 °C). Likewise a standard solution of the maleimide with the imines at 20 mM was prepared by dissolving the maleimide and imines in CD<sub>2</sub>Cl<sub>2</sub> / *p*-TSA (sat) which was equilibrated to the reaction temperature (10 °C or 25 °C). Pre-synthesised template for doping reactions was individually measured and the reaction initiated by mixing 0.4 mL of the nitrone solution with 0.4 mL of the maleimide and imine solution to give 0.8 mL reaction solution at 10 mM into a vial with or without template, transferred to an NMR spectroscopy tube and sealed with a polyethylene pressure cap to avoid solvent evaporation. In order to assess the exact concentration of added template, the reactions were immediately assayed by <sup>19</sup>F{<sup>1</sup>H}

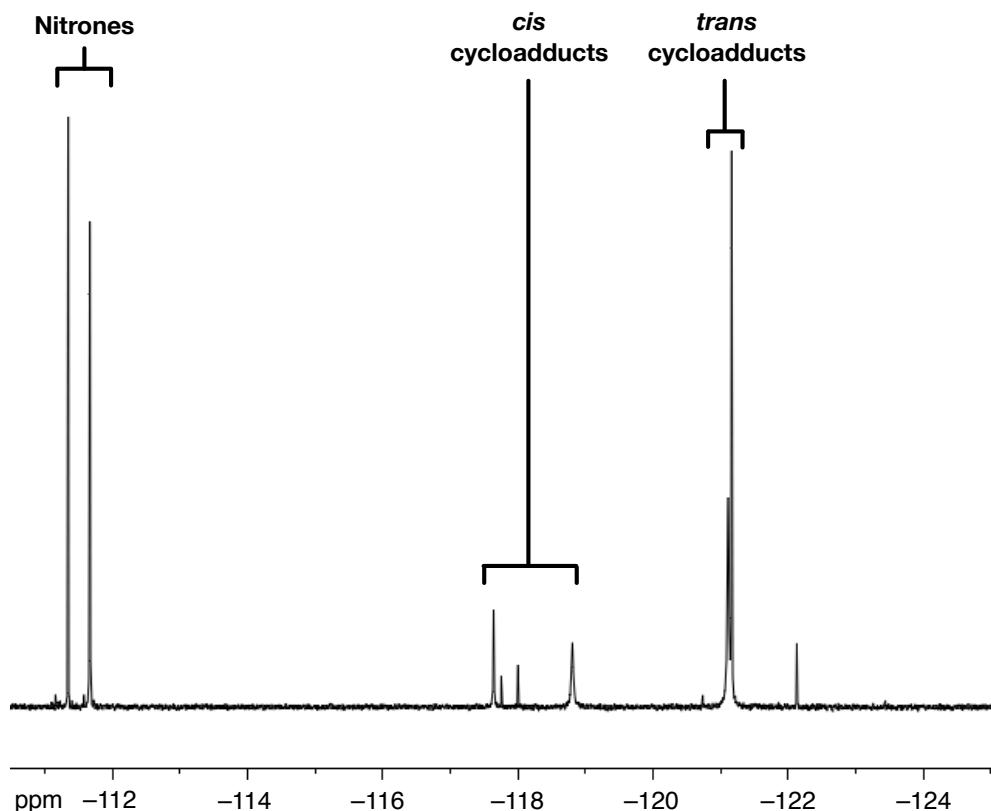


**Figure 7.2:** Partial <sup>19</sup>F {<sup>1</sup>H} NMR spectrum, 376.5 MHz, CD<sub>2</sub>Cl<sub>2</sub>/*p*-TSA (sat), 25 °C of a two replicator coupled to a dynamic reaction experiment at completion showing the characteristic region of resonance signals arising from nitrones, *cis* and *trans* cycloadducts, imines and selected reagents in the system to show their unique position.

NMR spectroscopy. The reactions were then placed into a thermostatically controlled water bath for the duration of the reaction and assayed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy. The relative concentration of each product in each spectra was then calculated using the deconvolution module of iNMR© versions 3.0.1→3.6.2 (current version). The total concentration was determined by summing the integral values for resonance signals for starting materials and products with the exact concentration of each species determined relative to the total concentration.

#### **7.4 Preparation and Analysis of Two or Three Replicators from Simple Components Experiments in Chapter 2**

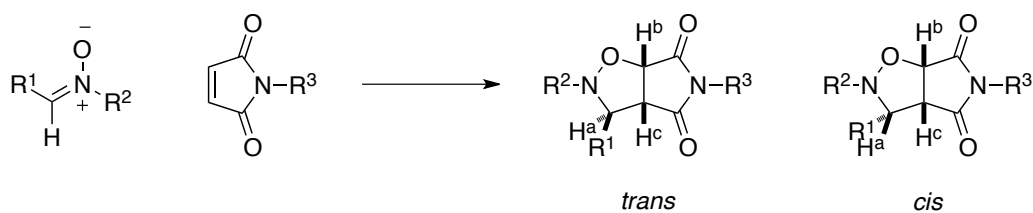
A standard solution of hydroxylamine **46** at 20 mM in  $\text{CD}_2\text{Cl}_2$  / *p*-TSA (sat) equilibrated to the reaction concentration (10 °C or 25 °C) was prepared. Likewise a standard solution of the maleimide with the aldehydes at 20 mM was prepared by dissolving the maleimide and aldehydes in  $\text{CD}_2\text{Cl}_2$  / *p*-TSA (sat) which was equilibrated to the reaction temperature (10 °C or 25 °C). Pre-synthesised template for doping reactions was individually measured and the reaction initiated by mixing 0.4 mL of the hydroxylamine solution with 0.4 mL of the maleimide and aldehyde solution to give 0.8 mL reaction solution at 10 mM into a vial with or without template, transferred to an NMR spectroscopy tube and sealed with a polyethylene pressure cap to avoid solvent evaporation. In order to assess the exact concentration of added template, the reactions were immediately assayed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy. The reactions were then placed into a thermostatically controlled water bath for the duration of the reaction and assayed by  $^{19}\text{F}\{^1\text{H}\}$  NMR spectroscopy. The relative concentration of each product in each spectra was then calculated using the deconvolution module of iNMR© versions 3.0.1→3.6.2 (current version). The total concentration was determined by summing the integral values for resonance signals for starting materials and products with the exact concentration of each species determined relative to the total concentration.



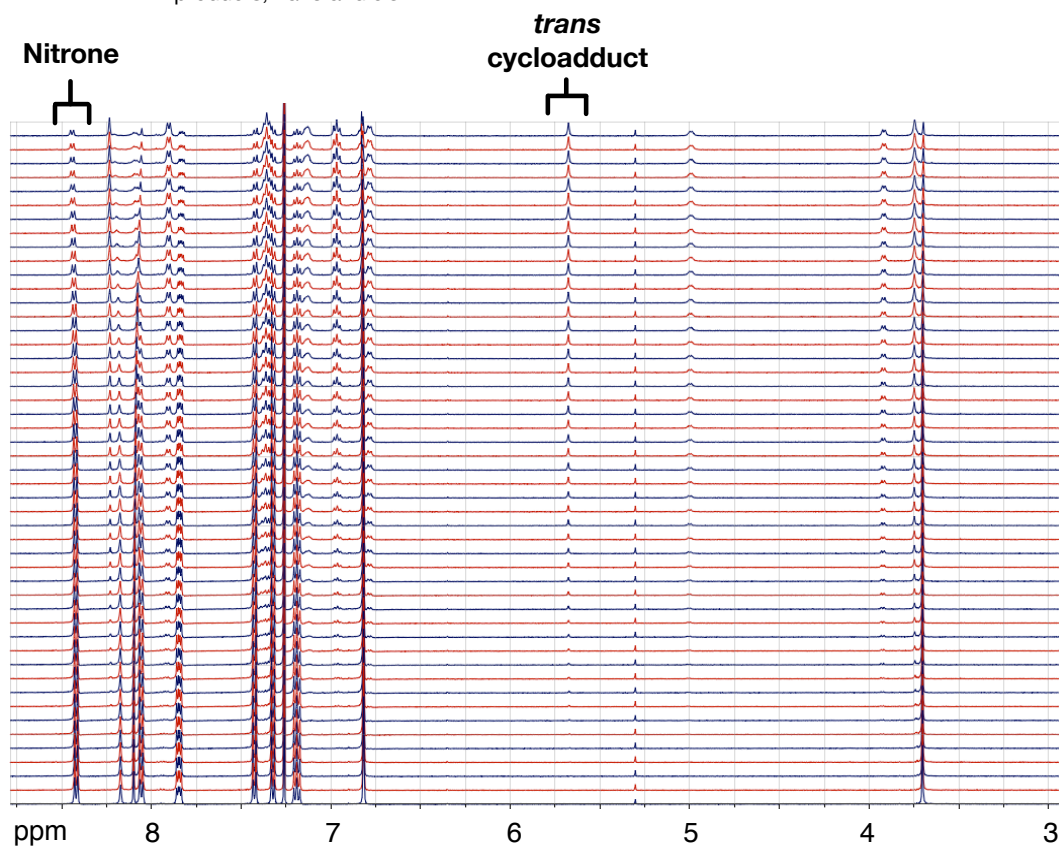
**Figure 7.3:** Partial  $^{19}\text{F}\{^1\text{H}\}$  NMR spectrum, 376.5 MHz,  $\text{CD}_2\text{Cl}_2/p\text{-TSA}$  (sat), 25 °C of a the two replicator simple component experiment after 16 h showing the characteristic region for nitrone, *cis* and *trans* cycloadduct resonance signals. When the reaction reaches completion, no nitrone resonance signals were observed.

## 7.5 Preparation and Kinetic Monitoring of Experiments by $^1\text{H}$ NMR spectroscopy and Deconvolution of Data used in Chapter 2, 3, 4 and 5.

Standard solutions of the appropriate reagents at double the reaction concentration were prepared using fresh  $\text{CDCl}_3$  or  $\text{CD}_2\text{Cl}_2 / p\text{-TSA}$  (sat) and equilibrated to the reaction temperature (0 °C or 10 °C). The reaction was initiated by mixing 0.4 mL of each of the reagent solutions to give a 0.8 mL reaction solution at the reaction concentration (10 mM or 20 mM) into an 8" 5 mm NMR spectroscopy tube (Wilmad 528PP). The tube was sealed with a polyethylene pressure cap to avoid solvent evaporation. The tube was swiftly transferred to a NMR spectrometer (Varian UNITYplus) regulated at the reaction temperature and 500.1 MHz  $^1\text{H}$  NMR spectra was acquired automatically every 30 minutes for a period of 16 hours. The relative concentration of the products was calculated by deconvolution of each of the  $^1\text{H}$  NMR spectra collected using the deconvolution module of iNMR© versions 2.3→3.6.2 (current version). The concentration of reactants and products at each time point was determined by monitoring the consumption of maleimide or nitrone proton signals and the appearance of cycloadduct proton signals.



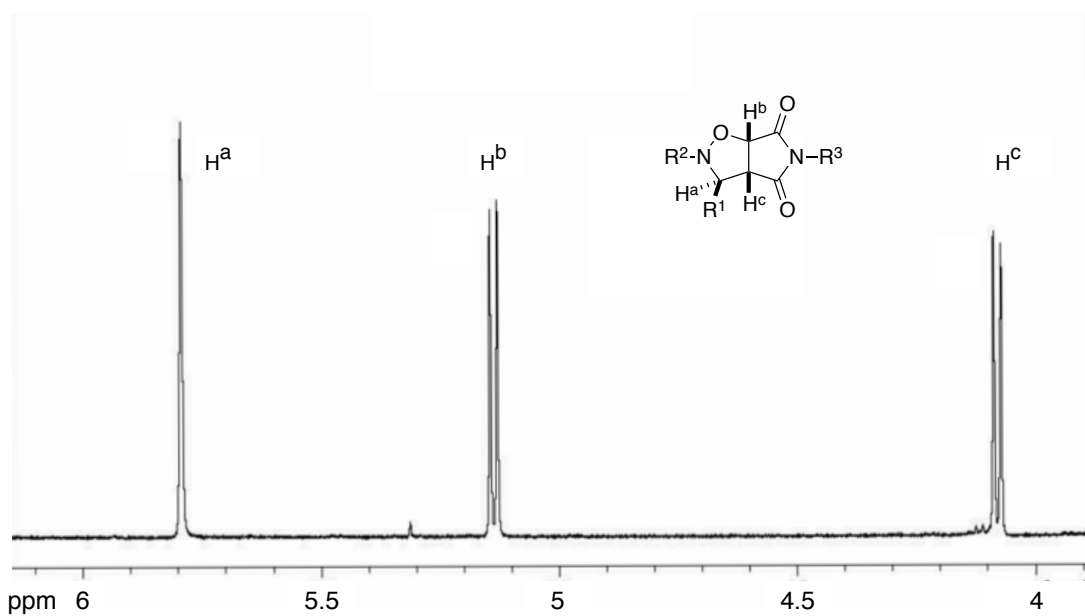
**Scheme 7.1:** 1,3-dipolar cycloaddition reaction of a maleimide with a nitronium ion leads to two diastereoisomeric products, *trans* and *cis*.



**Figure 7.4:** Stacked partial  $^1\text{H}$  NMR spectra, 500.1 MHz,  $\text{CDCl}_3$ ,  $0^\circ\text{C}$  of a kinetic experiment monitoring the depletion of the nitronium resonance signals and the appearance of *cis* and *trans* cycloadduct resonance signals. In this case, the *trans*:*cis* ratio was  $\sim 50:1$  and visualisation of *cis* cycloadducts required magnification.

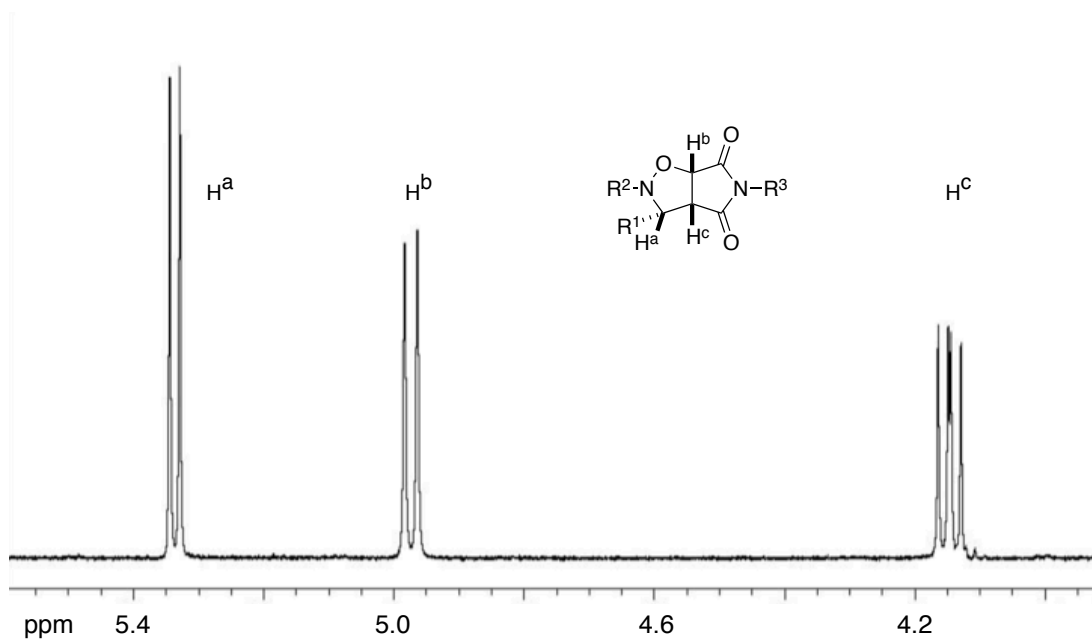
The two diastereoisomeric products of the reaction, *trans* and *cis*, are easily distinguishable by 500.1 MHz  $^1\text{H}$  NMR spectroscopy. The proton signals of the isoxazolidine show different splitting patterns for the two cycloadducts as a result of their different dihedral angles of  $\text{H}^a$  to  $\text{H}^c$ . The *trans* cycloadduct has a dihedral angle of  $\sim 90^\circ$  which, as described by the Karplus equation,<sup>214</sup> leads to a coupling constant of 0 Hz and hence the proton resonance signal for  $\text{H}^a$  is observed as a singlet. The *cis* cycloadduct has a dihedral angle of  $\sim 0^\circ$  which, as described by the Karplus equation, leads to the greatest coupling constant and accordingly the proton resonance signal for  $\text{H}^a$  is observed as a doublet and  $\text{H}^c$  as a double doublet.

### *trans*-cycloadduct



**Figure 7.5:** Partial 500.1 MHz  $^1\text{H}$  NMR spectra,  $\text{CDCl}_3$ ,  $0^\circ\text{C}$  of a *trans* cycloadduct showing the characteristic splitting pattern of resonance signals.  $\text{H}^a$  is a singlet and  $\text{H}^b$  and  $\text{H}^c$  are doublets.

### *cis*-cycloadduct

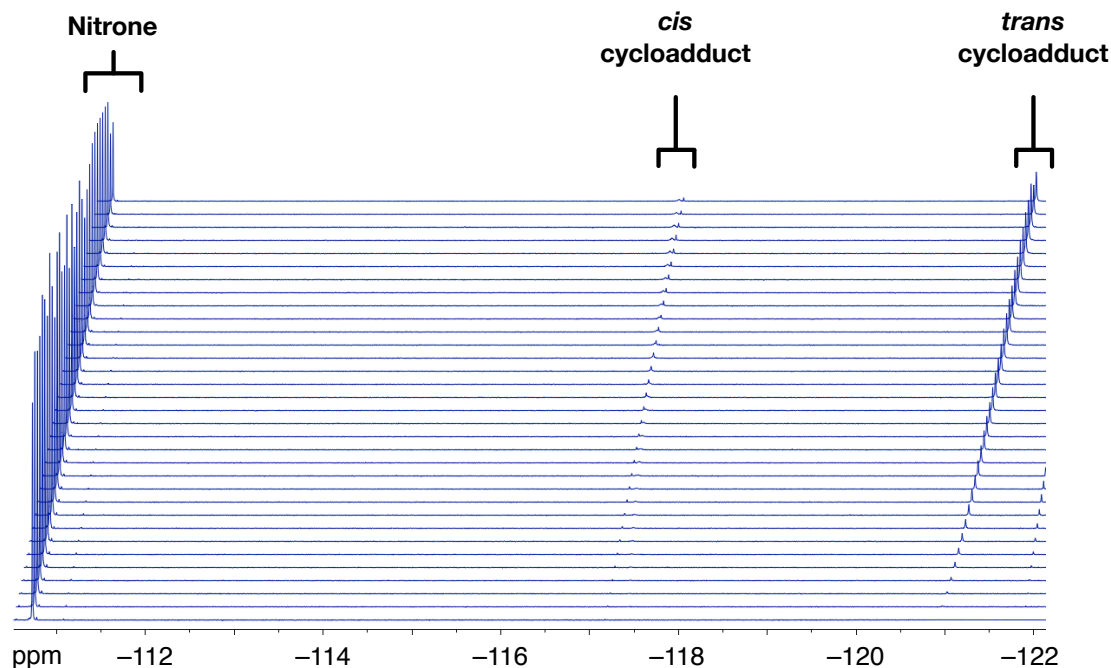


**Figure 7.6:** Partial 500.1 MHz  $^1\text{H}$  NMR spectra,  $\text{CDCl}_3$ ,  $0^\circ\text{C}$  of a *cis* cycloadduct showing the characteristic splitting pattern of resonance signals.  $\text{H}^a$  and  $\text{H}^b$  are doublets and  $\text{H}^c$  is a doublet.



## 7.6 Preparation and Kinetic Monitoring of Experiments by $^{19}\text{F}$ NMR spectroscopy and Deconvolution of Data used in Chapters 2, 3, 4 and 5.

Standard solutions of the appropriate reagents were prepared using fresh  $\text{CDCl}_3$  or  $\text{CD}_2\text{Cl}_2 / p\text{-TSA (sat)}$  and equilibrated to the reaction temperature ( $0\text{ }^\circ\text{C}$  or  $10\text{ }^\circ\text{C}$ ). The reaction was initiated by mixing 0.4 mL of each of the reagent solutions to give a 0.8 mL reaction solution at the reaction concentration (10 mM or 20 mM) into an 8" 5 mm NMR spectroscopy tube (Wilmad 528PP). The tube was then sealed with a polyethylene pressure cap to avoid solvent evaporation. The tube was swiftly transferred to a NMR spectrometer (Bruker Avance 500) regulated at the reaction temperature and a 470.3 MHz  $^{19}\text{F}$  NMR spectra was acquired automatically every 30 minutes for a period of 16 hours. The relative concentration of the products was calculated by deconvolution of each of the  $^{19}\text{F}$  NMR spectra collected using the deconvolution module of iNMR© versions 2.3→3.6.2 (current version).



**Figure 7.7:** Stacked and offset partial  $^{19}\text{F}\{^1\text{H}\}$  NMR spectra, 470.3 MHz,  $\text{CD}_2\text{Cl}_2/p\text{-TSA (sat)}$ ,  $0\text{ }^\circ\text{C}$  of a kinetic experiment monitoring the depletion of the nitron resonance signals and the appearance of *cis* and *trans* cycloadduct resonance signals.

## 7.7 Kinetic Simulation and Fitting

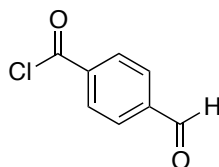
Kinetic simulation and fitting were undertaken using SimFit 32 (2003) and the isosim package incorporated within the software.

An input file, examples of which can be found in the appendix, is written which describes all possible bimolecular reactions and reversible recognition events within the system. The program is instructed to maintain interaction strengths and reaction rates of known processes, for example the association constants between amidopyridines and acids are experimentally determined by titration experiments and these values are used in the model. For reaction rates which are unknown, SimFit is allowed to vary these values  $\pm 1000$  to optimise the fitting of the simulated reaction profile to experimental data using a linear least squares algorithm. For irreversible reactions, one reaction rate is given. For reversible reactions, two reaction rates are used. The forward rate divided by the reverse rate represents the equilibrium constant. Optimised results have a maximum error ( $R^2$ ) of typically  $\pm 3\%$ .

SimFit was kindly provided by Prof. Günter von Kiedrowski, Ruhr-Universität, Bochum.

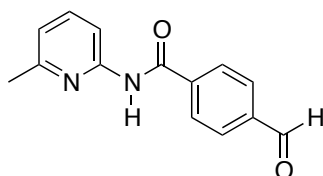
## 7.8 Characterisation of Compounds

### 4-Formylbenzoyl chloride<sup>215</sup>



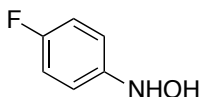
4-carboxybenzoic acid (4 g, 26.7 mmol, 1 eq) was suspended in toluene (100 mL) along with thionyl chloride (4.7 g, 40 mmol, 1.5 eq). The suspension was heated at 110 °C for 16 h after which the solvent was removed under vacuum and the residue dried by high vacuum to yield the acid chloride as a colourless crystalline solid (4.4 g, 26.7 mmol, 99%) which was used immediately without further purification; M.p. 48.6–49.6 °C (lit<sup>215</sup> 48 °C); <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>)  $\delta$  = 10.13 (1H, s, CHO), 8.26 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 2×ArH) and 8.01 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 2×ArH); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>)  $\delta$  = 191.1 (CHO), 167.8 (COCl), 140.5 (ArC), 137.6 (ArC), 131.8 (ArCH) and 129.8 (ArCH).

#### 4-Formyl-N-(6-methylpyridin-2-yl)benzamide<sup>139</sup> (49)



2-Amino-6-methylpyridine (2.4 g, 22 mmol, 1 eq) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (25 mL) and triethylamine (10 mL) and then cooled to 0 °C under N<sub>2</sub>. 4-formylbenzoyl chloride (4.4 g, 26.7 mmol, 1.2 eq) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (50 mL) and then added in a dropwise fashion to the amine solution and after complete addition reaction solution was allowed to rise slowly to RT. The reaction progress was monitored by TLC analysis and upon completion (typically 16 h) the reaction was quenched by addition of water (100 mL) and the product extracted with CH<sub>2</sub>Cl<sub>2</sub> (2 × 100mL). The combined organic fractions were then washed with sodium bicarbonate (aq), dried (MgSO<sub>4</sub>), filtered and the solvent removed to give the crude product which was purified by flash chromatography (cyclohexane:ethyl acetate, 3:2) to yield a colourless powder (4.2 g, 15.8 mmol, 72%); mp.105.5–106.6 °C (lit<sup>139</sup> 104.4–109.6); <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 10.09 (1H, s, HCO), 8.95–8.75 (1H, br, NH), 8.17 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.2 Hz, ArH), 8.06 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.3 Hz, 2×ArH), 7.97 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 2×ArH), 7.65 (1H, t, <sup>3</sup>J<sub>HH</sub> = 7.6 Hz, ArH), 6.94 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.5 Hz, ArH) and 2.41 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ = 191.5 (CHO), 164.6 (NHCO), 157.1 (ArC), 150.6 (ArC), 139.5 (ArC), 139.0 (ArCH), 138.7 (ArC), 130.0 (ArCH), 128.0 (ArCH), 120.0 (ArCH), 111.3 (ArCH), 24.0 (CH<sub>3</sub>); MS (EI<sup>+</sup>) *m/z* 265.03 ([M+Na]<sup>+</sup>, 100) and 241.07 ([M]<sup>+</sup>, 70); HRMS (EI<sup>+</sup>) *m/z* calculated for C<sub>14</sub>H<sub>12</sub>N<sub>2</sub>O<sub>2</sub>Na [MNa<sup>+</sup>] 263.0796, found 263.0791; CHN calculated for C<sub>14</sub>H<sub>12</sub>N<sub>2</sub>O<sub>2</sub> C, 69.99; H, 5.03; N, 11.66; found C, 69.53; H, 4.98; N, 11.40.

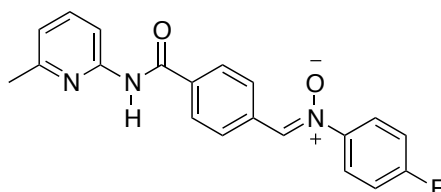
#### N-(4-fluorophenyl)hydroxylamine<sup>216</sup> (46)



4-Fluoronitrobenzene (0.95 g, 6.74 mmol, 1 eq) was dissolved in THF (20 mL) and Rh/C (0.4 g) added before stirring the suspension for for 15 min. Hydrazine monohydrate (0.4 g, 0.39 mL, 8.1 mmol, 1.2 eq) was then added and the reaction monitored by TLC analysis. Upon completion after 30 min the suspension was filtered through celite and the solvent removed to give the crude product as a

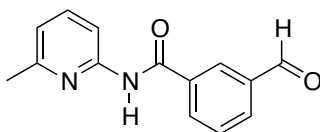
colourless crystalline solid which was dried by high vacuum before use without further purification. (0.85 g, 6.68 mmol, 99%); M.p. 92.5 – 93.2 (lit<sup>216</sup> 100 °C); <sup>1</sup>H {<sup>19</sup>F} NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 7.01 – 6.94 (4H, m, 4×ArH), and 6.83 – 5.68 (2H, br, NH + OH); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ = 158.7 (d, <sup>1</sup>J<sub>CF</sub> = 240 Hz, ArC–F), 145.4 (d, <sup>4</sup>J<sub>CF</sub> = 2.7 Hz, ArC), 116.5 (d, <sup>3</sup>J<sub>CF</sub> = 7.7 Hz, ArCH) and 116.4 (d, <sup>2</sup>J<sub>CF</sub> = 22.7 Hz, ArCH); <sup>19</sup>F {<sup>1</sup>H} NMR (376.5 MHz, CDCl<sub>3</sub>) δ = –121.98 (s, ArF); MS (ES<sup>-</sup>) *m/z* 126.05 ([M–H]<sup>-</sup>, 100 %)

**(Z)-4-fluoro-N-(4-(6-methylpyridin-2-ylcarbamoyl)benzylidene)aniline oxide<sup>155</sup> (52)**



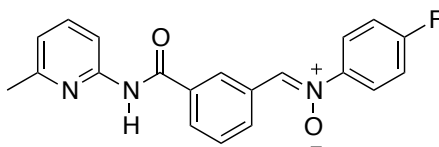
*N*-(6-methyl-2-pyridyl)-4-formyl benzamide **49** (0.57 g, 2.4 mmol, 1 eq) and *N*-(4-fluorophenyl)hydroxylamine **46** (0.3 g, 2.4 mmol, 1 eq) are dissolved together in ethanol (10 mL) and left to stand in darkness to react for 16 h. The precipitate which formed was isolated by filtration and washed with a little cold ethanol to yield 4-((Z)-[(4-fluorophenyl)(oxido)imino]methyl)-*N*-(6-methylpyridin-2-yl)benzamide as a colourless power (0.54 g, 1.5 mmol, 64%); M.p. 217.9–218.9 °C (lit<sup>155</sup> 217.2–218.3 °C); <sup>1</sup>H {<sup>19</sup>F} NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 8.74 (1H, br, NH), 8.45 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.3 Hz, 2×ArH), 8.17 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.2 Hz, ArH), 8.00 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.6 Hz, 2×ArH), 7.95 (1H, s, CH), 7.78 (2H, t, <sup>3</sup>J<sub>HH</sub> = 9.1 Hz, 2×ArH), 7.64 (1H, t, <sup>3</sup>J<sub>HH</sub> = 8.1 Hz, ArH), 7.16 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8 Hz, 2×ArH), 6.93 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.5 Hz, ArH) and 2.45 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ = 164.7 (NHCO), 162.3 (d, <sup>1</sup>J<sub>CF</sub> = 255 Hz, ArC–F), 157.0 (ArC), 150.7 (ArC), 145.1 (d, <sup>4</sup>J<sub>CF</sub> = 3.3 Hz, ArC), 138.8 (ArCH), 135.8 (ArC), 133.7 (ArC), 133.3 (ArCH), 129.0 (ArCH), 127.6 (CH), 123.7 (d, <sup>3</sup>J<sub>CF</sub> = 8.8 Hz, ArCH), 119.7 (ArCH), 116.2 (d, <sup>2</sup>J<sub>CF</sub> = 23.2 Hz, ArCH), 111.2 (ArCH) and 24.0 (CH<sub>3</sub>); <sup>19</sup>F {<sup>1</sup>H} NMR (376.5 MHz, CDCl<sub>3</sub>) δ = –110.09; MS (ES<sup>+</sup>) *m/z* 372.2 ([M+Na]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>20</sub>H<sub>16</sub>N<sub>3</sub>O<sub>2</sub>FNa [M+Na]<sup>+</sup> 372.1124, found 372.1114.

### 3-Formyl-N-(6-methylpyridin-2-yl)benzamide (59)



2-Amino-6-methylpyridine (1.11 g, 10.3 mmol, 1.1 eq) and triethylamine (2.8 g, 28 mmol, 3 eq) were dissolved in CH<sub>2</sub>Cl<sub>2</sub> (15 mL) and cooled to 0 °C under N<sub>2</sub>. A solution of 4-formylbenzoyl chloride (1.57 g, 9.3 mmol, 1 eq) in CH<sub>2</sub>Cl<sub>2</sub> (20 mL) was then added dropwise and after complete addition, the reaction was allowed to rise to RT and stirred for 16 h. NaHCO<sub>3</sub> (aq) (100 mL) was then added to quench the reaction and the product extracted into CH<sub>2</sub>Cl<sub>2</sub> (2 × 100 mL). The organic fractions were dried (MgSO<sub>4</sub>), filtered and the solvent removed under vacuum to give the crude product which was purified by flash chromatography (3:2, cyclohexane: ethyl acetate) to yield aldehyde **59** as a colourless powder (1.5 g, 6.2 mmol, 67%). M.p. 126–128 °C; <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 10.06 (1H, s, CHO), 8.87 (1H, br, NH), 8.40 (1H, t, <sup>4</sup>J<sub>HH</sub> = 1.5 Hz, ArH), 8.22–8.14 (2H, m, 2×ArH), 8.05 (1H, dt, <sup>3</sup>J<sub>HH</sub> = 7.7 Hz and <sup>4</sup>J<sub>HH</sub> = 1.4 Hz, ArH), 7.68–7.62 (2H, m, 2×ArH), 6.93 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.5 Hz, ArH) and 2.42 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ = 191.3 (CHO), 164.4 (CONH), 157.1 (ArC), 150.6 (ArC), 138.9 (ArCH), 136.7 (ArC), 135.4 (ArC), 133.1 (ArCH), 132.7 (ArCH), 129.7 (ArCH), 128.4 (ArCH), 119.9 (ArCH), 111.2 (ArCH) and 24.0 (CH<sub>3</sub>); MS (EI<sup>+</sup>) *m/z* 241 ([M+H]<sup>+</sup>, 100); HRMS (EI<sup>+</sup>) *m/z* calculated for C<sub>14</sub>H<sub>12</sub>N<sub>2</sub>O<sub>2</sub> [M+H]<sup>+</sup> 240.0899, found 240.0888.

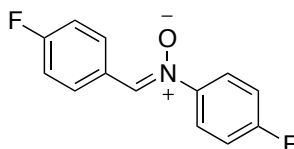
### (Z)-4-fluoro-N-(3-(6-methylpyridin-2-ylcarbamoyl)benzylidene)aniline oxide (58)



3-Formyl-N-(6-methylpyridin-2-yl)benzamide (0.2 g, 0.57 mmol, 1 eq) and N-(4-fluorophenyl)hydroxylamine **46** (0.073 g, 0.57 mmol, 1 eq) were dissolved in ethanol (5 mL) and left to sit in darkness for 2 days. After this time no precipitation was observed and therefore the reaction was cooled to -17 °C for several hours. The precipitate which formed was then isolated by filtration to yield nitrone **58** (0.07 g, 0.20 mmol, 35%) as a pale yellow crystalline solid; M.p. 189.6–191.8 °C; <sup>1</sup>H {<sup>19</sup>F} NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 8.87–8.78 (2H, m, NH and ArH), 8.69 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.9

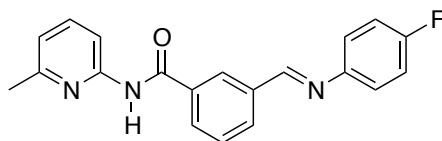
Hz, *ArH*), 8.16 (1H, d,  $^3J_{\text{HH}} = 8.2$  Hz, *ArH*), 8.06–7.99 (1H, m, *ArH*), 7.94 (1H, s, CH), 7.78 (2H, d,  $^3J_{\text{HH}} = 9.1$  Hz,  $2\times\text{ArH}$ ), 7.66–7.55 (2H, m,  $2\times\text{ArH}$ ), 7.17 (2H, d,  $^3J_{\text{HH}} = 9.1$  Hz,  $2\times\text{ArH}$ ), 6.92 (1H, d,  $^3J_{\text{HH}} = 7.5$  Hz, *ArH*) and 2.42 (3H, s,  $\text{CH}_3$ );  $^{13}\text{C}$  NMR (100.6 MHz,  $\text{CDCl}_3$ )  $\delta = 165.1$  (NHCO), 163.8 (d,  $^1J_{\text{CF}} = 253.1$  Hz, ArC), 157.1 (ArC), 145.1 (d,  $^4J_{\text{CF}} = 3$  Hz, ArC), 138.8 (ArCH), 135.0 (ArC), 133.4 (ArCH), 132.2 (ArCH), 131.2 (ArC), 129.7 (ArCH), 129.4 (ArCH), 127.6 (ArCH), 123.7 (d,  $^3J_{\text{CF}} = 9.1$  Hz, ArCH) 119.7 (ArCH), 116.3 (d,  $^2J_{\text{CF}} = 23.4$  Hz, ArCH), 111.1 (ArCH), and 24.1 ( $\text{CH}_3$ );  $^{19}\text{F}$   $\{^1\text{H}\}$  NMR (376.5 MHz,  $\text{CDCl}_3$ )  $\delta = -109.78$  (1F, s, *ArF*); MS ( $\text{ES}^+$ )  $m/z$  372.0 ( $[\text{M}+\text{Na}]^+$ , 10), 350.01( $[\text{M}+\text{H}]^+$ , 100); HRMS ( $\text{ES}^+$ )  $m/z$  calculated for  $\text{C}_{20}\text{H}_{16}\text{N}_3\text{O}_2\text{FNa}$   $[\text{M}+\text{Na}]^+$  372.1124, found 372.1120.

**(Z)-4-fluoro-N-(4-fluorobenzylidene)aniline oxide (65)**



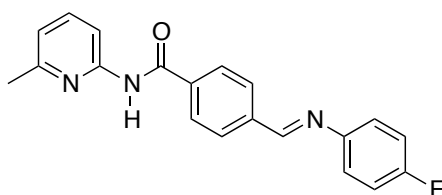
*N*-(4-fluorophenyl)hydroxylamine **46** (0.2 g, 1.57 mmol, 1 eq) and 4-fluorobenzaldehyde (0.2 g, 1.57 mmol, 1 eq) were dissolved in EtOH (10 mL) and left to sit in darkness for 16 h. The resulting precipitate was isolated by filtration and washed with cold ethanol to furnish nitrone **65** (0.25 g, 1.07 mmol, 68 %) as a colourless powder; M.p. 171.2–171.9 °C;  $^1\text{H}$  NMR (300.1 MHz,  $\text{CDCl}_3$ )  $\delta = 8.47$ – $8.39$  (2H, m,  $2\times\text{ArH}$ ), 7.86 (1H, s, CH), 7.81–7.72 (2H, m,  $2\times\text{ArH}$ ), 7.21–7.11 (4H, m,  $4\times\text{ArH}$ );  $^{13}\text{C}$  NMR (100.6 MHz,  $\text{CDCl}_3$ )  $\delta = 163.8$  (d,  $^1J_{\text{CF}} = 254$  Hz, ArCF), 163.2 (d,  $^1J_{\text{CF}} = 251$  Hz, ArC–F), 145.2 (d,  $^4J_{\text{CF}} = 2.9$  Hz, ArC), 133.4 (CH), 131.5 (d,  $^3J_{\text{CF}} = 8.2$  Hz, ArCH), 127.1 (d,  $^4J_{\text{CF}} = 2.1$  Hz, ArCH), 123.8 (d,  $^3J_{\text{CF}} = 8.8$  Hz, ArCH), 116.3 (d,  $^2J_{\text{CF}} = 23.1$  Hz, ArCH), 116.0 (d,  $^2J_{\text{CF}} = 21.7$  Hz, ArCH);  $^{19}\text{F}$   $\{^1\text{H}\}$  NMR (282.6 MHz,  $\text{CDCl}_3$ )  $\delta = -106.86$  (1F, s, *ArF*) and  $-110.74$  (1F, s, *ArF*); MS ( $\text{CI}^+$ )  $m/z$  234.07 ( $[\text{M}+\text{H}]^+$ , 100); HRMS ( $\text{CI}^+$ )  $m/z$  calculated for  $\text{C}_{13}\text{H}_{10}\text{NOF}_2$   $[\text{M}+\text{H}]^+$  234.0730, found 234.0730.

**(E)-3-((4-fluorophenylimino)methyl)-N-(6-methylpyridin-2-yl)benzamide (66)**



4-Fluoroaniline (0.12 g, 1.04 mmol, 1 eq) and aldehyde **59** (0.25 g, 1.04 mmol, 1 eq) were dissolved in ethanol (20 mL) and left to sit for 16 h. The precipitate formed was isolated by filtration and washed with a little cold ethanol to yield imine **66** (0.18 g, 0.53 mmol, 51%) as a pale yellow fine crystalline solid; M.p. 100.2–101.8 °C;  $^1\text{H}$   $\{^{19}\text{F}\}$  NMR (400.1 MHz,  $\text{CDCl}_3$ )  $\delta$  = 8.94 (1H, br, *NH*), 8.47 (1H, s, *CH*), 8.44 (1H, t,  $^3J_{\text{HH}}$  = 1.7 Hz, *ArH*), 8.20 (1H, d,  $^3J_{\text{HH}}$  = 8.2 Hz, *ArH*), 8.10–8.01 (2H, m,  $2\times\text{ArH}$ ), 7.64 (1H, dd,  $^3J_{\text{HH}}$  = 7.6 and  $^3J_{\text{HH}}$  8.1 Hz, *ArH*), 7.58 (1H, t,  $^3J_{\text{HH}}$  = 7.7 Hz, *ArH*), 7.22 (2H, dt,  $^3J_{\text{HH}}$  = 9 Hz and  $^4J_{\text{HH}}$  and 2.8 Hz,  $2\times\text{ArH}$ ), 7.09 (2H, dt,  $^3J_{\text{HH}}$  = 9 Hz and  $^4J_{\text{HH}}$  = 2.8 Hz,  $2\times\text{ArH}$ ), 6.92 (1H, d,  $^3J_{\text{HH}}$  = 7.3 Hz, *ArH*) and 2.45 (3H, s, *CH*<sub>3</sub>);  $^{13}\text{C}$  NMR (100.1 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{C}}$  = 165.1 (CONH), 161.6 (d,  $^1J_{\text{CF}}$  = 250 Hz, ArC–F), 158.7 (CH), 157.0 (ArC), 150.8 (ArC), 147.5 (d,  $^4J_{\text{CF}}$  = 2.9 Hz, ArC), 139.0 (ArCH), 136.7 (ArC), 135.1 (ArC), 132.2 (ArCH), 130.3 (ArCH), 129.5 (ArCH), 127.3 (ArCH), 122.5 (d,  $^3J_{\text{CF}}$  = 8.6 Hz, ArCH), 119.7 (ArCH), 116.1 (d,  $^2J_{\text{CF}}$  = 23 Hz, ArCH), 111.3 (ArCH) and 24.0 (*CH*<sub>3</sub>);  $^{19}\text{F}$   $\{^1\text{H}\}$  NMR (376.5 MHz,  $\text{CDCl}_3$ )  $\delta_{\text{F}}$  = –116.98 (1F, s, *ArF*); MS ( $\text{ES}^+$ )  $m/z$  355.99 ( $[\text{M}+\text{Na}]^+$ , 100), 334.03 ( $[\text{M}+\text{H}]^+$ , 35); HRMS ( $\text{ES}^+$ )  $m/z$  calculated for  $\text{C}_{20}\text{H}_{17}\text{N}_3\text{OF}$   $[\text{M}+\text{H}]^+$  334.1356, found 334.1351.

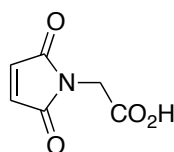
**(E)-4-((4-fluorophenylimino)methyl)-N-(6-methylpyridin-2-yl)benzamide (47)**



4-Fluoroaniline (0.25 g, 2.29 mmol, 1.1 eq) and aldehyde **59** (0.5 g, 2.08 mmol, 1 eq) were dissolved in ethanol (17 mL) and left to sit for 16 h. The precipitate formed was isolated by filtration and washed with a little cold ethanol to yield imine **47** (0.53 g, 1.59 mmol, 76%) as a colourless fine crystalline solid; M.p. 153.8–154.6 °C;  $^1\text{H}$   $\{^{19}\text{F}\}$  NMR (400.1 MHz,  $\text{CDCl}_3$ )  $\delta$  = 8.73 (1H, br, *NH*), 8.48 (1H, s, *CH*), 8.22 (1H,  $^3J_{\text{HH}}$  = 8.19 Hz, *ArH*), 8.06–7.93 (4H, m,  $4\times\text{ArH}$ ), 7.65 (1H, t,  $^3J_{\text{HH}}$  = 7.6 Hz, *ArH*), 7.23 (2H, dt,  $^3J_{\text{HH}}$  = 9.0 Hz and  $^4J_{\text{HH}}$  = 2.8 Hz,  $2\times\text{ArH}$ ), 7.09 (2H, dt,  $^3J_{\text{HH}}$  = 9.0 Hz and  $^4J_{\text{HH}}$  = 2.8

Hz, 2×ArH), 6.95 (1H, d,  $^3J_{\text{HH}} = 7.6$  Hz, ArH), 2.45 (3H, s, CH<sub>3</sub>);  $^{13}\text{C}$  NMR (100.1 MHz, CDCl<sub>3</sub>)  $\delta = 165.0$  (CONH), 161.6 (d,  $^1J_{\text{CF}} = 251$  Hz, ArCF), 158.6 (CH), 157.1 (ArC), 150.8 (ArC), 147.5 (d,  $^4J_{\text{CF}} = 2.9$  Hz, ArC), 139.4 (ArC), 138.9 (ArCH), 136.6 (ArC), 129.1 (ArCH), 127.8 (ArCH), 122.6 (d,  $^3J_{\text{CF}} = 8.1$  Hz, ArCH), 119.8 (ArCH), 116.1 (d,  $^2J_{\text{CF}} = 23.0$  Hz, ArCH), 111.2 (ArCH) and 24.1 (CH<sub>3</sub>);  $^{19}\text{F}$  { $^1\text{H}$ } NMR (376.5 MHz, CDCl<sub>3</sub>)  $\delta = -116.71$  (1F, s, ArF); MS (ES<sup>+</sup>)  $m/z$  356 ([M+Na]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>)  $m/z$  calculated for C<sub>20</sub>H<sub>16</sub>N<sub>3</sub>OFNa [M+Na]<sup>+</sup> 356.1174, found 356.1175

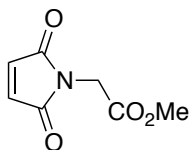
**2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)acetic acid<sup>187</sup> (29)**



Maleic anhydride (1.19 g, 12.1 mmol, 1.2 eq) and glycine (0.76 g, 10.1 mmol, 1 eq) were dissolved in acetic acid (25 mL) and THF (5 mL) and stirred for 16 h. The cloudy solution was treated with ether (50 mL) before the precipitate isolated by filtration and dried. This yielded the intermediate amide (1.43 g, 7.1 mmol, 70%) as a colourless powder which was dissolved in acetonitrile (25 mL) along with ZnBr<sub>2</sub> (1.57 g, 7 mmol, 1 eq) and hexamethyldisilazane (5.6 g, 35 mmol, 5 eq). The suspension was heated to reflux for 1 h and then hot filtered and concentrated. The residue was treated with 0.5 M HCl<sub>(aq)</sub> (25 mL) and the product extracted into chloroform. The organic fraction was then separated, washed with saturated EDTA solution, dried (MgSO<sub>4</sub>), filtered and the solvent removed under vacuum to yield maleimide **29** (1.05 g, 6.8 mmol, 97%) as a colourless powder: M.p. 114.0-116.0 °C (lit<sup>187</sup> 114.0-116.0 °C);  $^1\text{H}$  NMR (300.1 MHz, CDCl<sub>3</sub>)  $\delta = 9.73$  (1H, s, CO<sub>2</sub>H), 6.73 (2H, s, CH), 4.26 (2H, s, CH<sub>2</sub>).  $^{13}\text{C}$  NMR (75.5 MHz, CDCl<sub>3</sub>)  $\delta = 172.9$  (CO<sub>2</sub>H), 170.1 (CO), 134.9 (CH), 38.7 (CH<sub>2</sub>). MS (EI<sup>+</sup>)  $m/z$  155 ([M<sup>+</sup>], 8), 110 (100). Found C 46.61 H 3.34 N 9.01 C<sub>6</sub>H<sub>5</sub>O<sub>4</sub>N requires C 46.46 H 3.25 N 9.03.

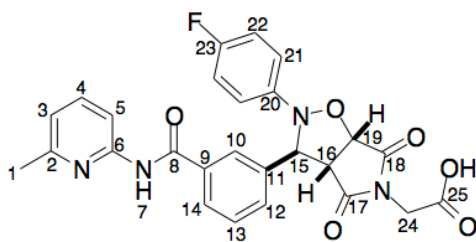


### Methyl 2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)acetate (**60**)



Maleimide **29** (0.4 g, 2.58 mmol, 1 eq) was dissolved in THF (10 mL) and cooled to 0 °C before addition of DCC (0.59 g, 2.84 mmol, 1.1 eq). The solution was then allowed to stir for 1 h before addition of methanol (0.4 g, 13 mmol 5 eq) and subsequently allowing the temperature to rise to RT before stirring for 16 h. The solvent was then removed and the residue dissolved in ether (20 mL). The urea biproduct was removed by filtration before removing the solvent to give the crude product which was purified by flash chromatography (3:1 cyclohexane:ethyl acetate) to yield the purified product (0.3 g, 1.77 mmol, 69%) as a colourless liquid; <sup>1</sup>H NMR (300.1 MHz, CDCl<sub>3</sub>) δ = 6.78 (2H, 2xCH), 4.28 (2H, s, CH<sub>2</sub>), 3.74 (3H, s, CO<sub>2</sub>CH<sub>3</sub>); <sup>13</sup>C NMR (75.5 MHz, CDCl<sub>3</sub>) δ = 169.7 (CO), 167.6 (CO<sub>2</sub>Me), 134.5 (CH), 52.7 (OCH<sub>3</sub>), 38.4 (CH<sub>2</sub>); MS (EI<sup>+</sup>) *m/z* = 169 (M<sup>+</sup>, 42%) 110 (100), 82 (22), 54 (14), 43 (11); HRMS (EI<sup>+</sup>) *m/z* calculated for C<sub>7</sub>H<sub>7</sub>NO<sub>4</sub> [M<sup>+</sup>] 169.0375, Found 169.0382.

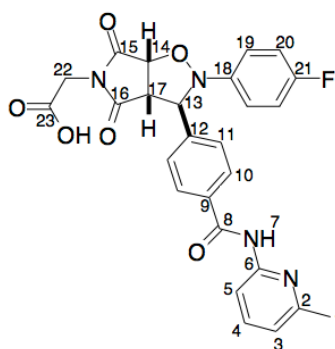
### 2-((3*S*,3*aR*,6*aS*)-2-(4-fluorophenyl)-3-(3-((6-methylpyridin-2-yl)carbamoyl)phenyl)-4,6-dioxotetrahydro-2*H*-pyrrolo[3,4-*d*]isoxazol-5(3*H*)-yl)acetic acid (*trans*-**62**)



Maleimide **29** (0.0062 g, 0.04 mmol, 1 eq) and nitrone **68** (0.014 g, 0.04 mmol, 1 eq) were dissolved in CHCl<sub>3</sub> (2 mL), cooled to 10 °C and allowed to react for 72 h before raising to RT and allowing to react for a further 24 h. The solvent was then removed to yield the crude product as a mixture of *trans* and *cis* isomers 10:1. The *trans* product was then enriched by recrystallisation from methanol and hexane to yield *trans*-**62** as a colourless powder (0.012 g, 0.024 mmol, 60%); M.p. 191.0–194.6; <sup>1</sup>H NMR (499.9 MHz, CDCl<sub>3</sub>) δ = 10.77 (1H, s, NH, H<sub>7</sub>), 8.06 (1H, s, ArH, H<sub>10</sub>), 8.01 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.3 Hz, ArH, H<sub>5</sub>), 7.89 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.8 Hz, ArH, H<sub>14</sub>), 7.74 (1H, t, <sup>3</sup>J<sub>HH</sub> = 362

7.8 Hz, ArH, H<sub>4</sub>), 7.60 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.9 Hz, ArH, H<sub>12</sub>), 7.45 (1H, t, <sup>3</sup>J<sub>HH</sub> = 7.8 Hz, ArH, H<sub>13</sub>), 7.10–6.97 (5H, m, 5×ArH, H<sub>3</sub>, H<sub>21</sub> and H<sub>22</sub>) 5.62 (1H, s, CH, H<sub>15</sub>), 5.53 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.3 Hz, CH, H<sub>19</sub>), 4.33 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.3 Hz, CH, H<sub>16</sub>), 4.00 (1H, d, <sup>2</sup>J<sub>HH</sub> = 17 Hz, HCH, H<sub>24</sub>), 3.95 (1H, d, <sup>2</sup>J<sub>HH</sub> = 17 Hz, HCH, H<sub>24</sub>) and 2.46 (3H, s, CH<sub>3</sub>, H<sub>1</sub>); <sup>13</sup>C NMR (125.7 MHz, CDCl<sub>3</sub>) δ = 174.4 (CO, C<sub>17</sub>), 173.4 (CO, C<sub>18</sub>), 167.5 (CO<sub>2</sub>H, C<sub>25</sub>), 165.6 (CONH, C<sub>8</sub>), 157.7 (d, <sup>1</sup>J<sub>CF</sub> = 239 Hz, ArCF, C<sub>23</sub>), 156.6 (ArC, C<sub>2</sub>), 151.4 (ArC, C<sub>6</sub>), 143.7 (d, <sup>4</sup>J<sub>CF</sub> = 1.9 Hz, ArC, C<sub>20</sub>), 138.5 (ArCH, C<sub>4</sub>), 137.9 (ArC, C<sub>11</sub>), 134.3 (ArC, C<sub>9</sub>), 131.1 (ArCH, C<sub>12</sub>), 128.5 (ArCH, C<sub>13</sub>), 127.7 (ArCH, C<sub>14</sub>), 127.33 (ArCH, C<sub>10</sub>), 119.1 (ArCH, C<sub>3</sub>), 117.3 (d, <sup>3</sup>J<sub>CF</sub> = 7.8 Hz, ArCH, C<sub>21</sub>), 115.2 (d, <sup>2</sup>J<sub>CF</sub> = 22.4 Hz, ArCH, C<sub>22</sub>), 111.6 (ArCH, C<sub>5</sub>), 76.5 (CH, C<sub>19</sub>), 68.6 (CH, C<sub>15</sub>), 55.9 (CH, C<sub>16</sub>), 39.6 (CH<sub>2</sub>, C<sub>24</sub>) and 23.5 (CH<sub>3</sub>, C<sub>1</sub>); <sup>19</sup>F NMR (282.6 MHz, CDCl<sub>3</sub>) δ = -121.2; MS (ES<sup>-</sup>) *m/z* 502.84 ([M-H]<sup>-</sup>, 15), 377.87 (100) and 333.93 (20); HRMS (ES<sup>-</sup>) *m/z* calculated for C<sub>26</sub>H<sub>20</sub>N<sub>4</sub>O<sub>6</sub>F [M-H]<sup>-</sup> 503.1367, found 503.1364.

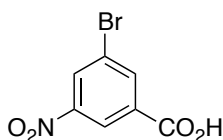
**2-((3*S*,3*aR*,6*aS*)-2-(4-fluorophenyl)-3-(4-((6-methylpyridin-2-yl)carbamoyl)phenyl)-4,6-dioxotetrahydro-2*H*-pyrrolo[3,4-*d*]isoxazol-5(3*H*)-yl)acetic acid (*trans*-**64**)**



Maleimide **29** (0.0062 g, 0.04 mmol, 1 eq) and nitrone **52** (0.014 g, 0.04 mmol, 1 eq) were dissolved in CHCl<sub>3</sub>, cooled to 10 °C and allowed to react for 72 h before raising to RT and allowing to react for a further 24 h. The solvent was then removed to yield the crude product as a mixture of *trans* and *cis* isomers 10:1. The *trans* product was then enriched by recrystallisation from methanol and hexane to yield *trans*-**64** as a colourless powder (0.014 g, 0.028 mmol, 70%); M.p. 218.0–219.7 °C; <sup>1</sup>H NMR (499.9 MHz, d<sub>6</sub>-DMSO) δ = 10.66 (1H, s, NH, H<sub>7</sub>), 8.03–7.92 (3H, m, 3×ArH, H<sub>3</sub> & H<sub>10</sub>), 7.71 (1H, t, <sup>3</sup>J<sub>HH</sub> = 7.1 Hz, ArH, H<sub>4</sub>), 7.52 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8 Hz, 2×ArH, H<sub>11</sub>), 7.09–6.94 (5H, m, 5×ArH, H<sub>5</sub>, H<sub>19</sub> & H<sub>20</sub>), 5.61 (1H, s, CH, H<sub>13</sub>), 5.50 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.2 Hz, CH, H<sub>14</sub>), 4.18 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.2 Hz, CH, H<sub>17</sub>), 4.02 (2H, s, CH<sub>2</sub>, H<sub>22</sub>) and 2.43 (3H, s,

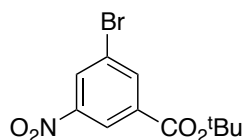
CH<sub>3</sub>, H<sub>1</sub>); <sup>13</sup>C NMR (125.7 MHz, CDCl<sub>3</sub>) δ = 174.3 (CO, C<sub>16</sub>), 173.5 (CO, C<sub>15</sub>), 167.6 (CO<sub>2</sub>H, C<sub>23</sub>), 165.3 (NHCO, C<sub>8</sub>), 157.7 (d, <sup>1</sup>J<sub>CF</sub> = 239 Hz, ArCF, C<sub>21</sub>), 156.5 (ArC, C<sub>2</sub>), 151.4 (ArC, C<sub>6</sub>), 143.4 (d, <sup>4</sup>J<sub>CF</sub> = 2.0 Hz, ArC, C<sub>18</sub>), 141.3 (ArC, C<sub>12</sub>), 138.3 (ArCH, C<sub>4</sub>), 133.6 (ArC, C<sub>9</sub>), 128.1 (ArCH, C<sub>10</sub>), 127.8 (ArCH, C<sub>11</sub>), 119.1 (ArCH, C<sub>5</sub>), 117.4 (d, <sup>3</sup>J<sub>CF</sub> = 7.8 Hz, ArCH, C<sub>19</sub>), 115.2 (d, <sup>2</sup>J<sub>CF</sub> = 22.4 Hz, ArCH, C<sub>20</sub>), 111.7 (ArCH, C<sub>3</sub>), 76.4 (CH, C<sub>17</sub>), 68.3 (CH, C<sub>13</sub>), 55.9 (CH, C<sub>14</sub>), 39.7 (CH<sub>2</sub>, C<sub>22</sub>) and 23.6 (CH<sub>3</sub>, C<sub>1</sub>); <sup>19</sup>F NMR (470.3 MHz, CDCl<sub>3</sub>) δ = -121.2; MS (ES<sup>-</sup>) *m/z* 503.04 ([M-H]<sup>-</sup>, 100); HRMS (ES<sup>-</sup>) *m/z* calculated for C<sub>26</sub>H<sub>20</sub>N<sub>4</sub>O<sub>6</sub>F [M-H]<sup>-</sup> 503.1367, found 503.1367.

### 3-Bromo-5-nitrobenzoic acid<sup>217</sup>



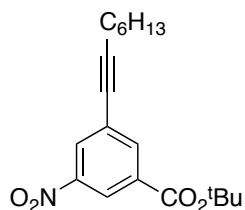
3-Nitrobenzoic acid (11 g, 66 mmol, 1 eq) was dissolved in conc. H<sub>2</sub>SO<sub>4</sub> (40 mL) and heated to 60 °C. *N*-bromosuccinimide (14.1 g, 79.2 mmol, 1.2 eq) was added in 3 portions over 15 minutes. The reaction was stirred for a further 1.5 h before being poured on to ice water (200 mL) and left overnight to complete precipitation. The product was isolated by filtration and dried at 80 °C to yield the product as a colourless powder (16.2 g, 66 mmol, quantitative); M.p. 148.2–149.3 °C (lit<sup>217</sup> 159.0–160.0 °C). <sup>1</sup>H NMR (400.1 MHz, d<sub>6</sub>-DMSO) δ = 8.62 (1H, dd, <sup>4</sup>J<sub>HH</sub> = 1.9 and <sup>4</sup>J<sub>HH</sub> = 2.1 Hz, *ArH*), 8.54 (1H, dd, <sup>4</sup>J<sub>HH</sub> = 1.3 and <sup>4</sup>J<sub>HH</sub> = 1.9 Hz, *ArH*) and 8.41 (1H, dd, <sup>4</sup>J<sub>HH</sub> = 1.3 and <sup>4</sup>J<sub>HH</sub> = 1.9 Hz, *ArH*); <sup>13</sup>C NMR (100.6 MHz, d<sub>6</sub>-DMSO) δ = 164.4 (CO<sub>2</sub>H), 148.7 (ArC), 137.6 (ArCH), 134.2 (ArC), 130. (ArCH), 122.8 (ArCH) and 122.4 ArC); MS (Cl<sup>+</sup>) *m/z* 245.94 ([<sup>79</sup>BrM+H]<sup>+</sup>, 100) and 247.94 ([<sup>81</sup>BrM+H]<sup>+</sup>, 95); HRMS (Cl<sup>+</sup>) *m/z* calculated for [M+H]<sup>+</sup> C<sub>7</sub>H<sub>5</sub>NO<sub>4</sub><sup>79</sup>Br 245.9402, found 245.9405, calculated for [M+H]<sup>+</sup> C<sub>7</sub>H<sub>5</sub>NO<sub>4</sub><sup>81</sup>Br 247.9381, found 247.9386.

### ***tert*-Butyl 3-bromo-5-nitrobenzoate<sup>218</sup>**



3-Bromo-5-nitrobenzoic acid (7.38 g, 30 mmol, 1 eq) was dissolved in dry THF (60 mL) and stirred at RT under N<sub>2</sub>. Di-*t*-butyl dicarbonate (19.6 g, 90 mmol, 3 eq) and dimethylaminopyridine (0.84 g, 6.9 mmol, 0.04 eq) were added and the reaction stirred under N<sub>2</sub> for 60 h before pouring on to ice water (200 mL). The solution was made basic (~pH 10) by addition of Na<sub>2</sub>CO<sub>3</sub> before extraction of the product into CH<sub>2</sub>Cl<sub>2</sub> (3 × 100 mL). The combined organic fractions were dried over MgSO<sub>4</sub>, filtered and the solvent removed to give the crude product which was purified by flash chromatography (10:1, cyclohexane: ethyl acetate) to yield the purified product as a colourless crystalline solid (7.3 g, 24.2 mmol, 81%). M.p. 75.9–76.8 °C (lit<sup>218</sup> 77.0–78.0 °C); <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 8.70 (1H, dd, <sup>4</sup>J<sub>HH</sub> = 1.4 Hz and <sup>4</sup>J<sub>HH</sub> = 2.1 Hz, *ArH*), 8.50 (1H, t, <sup>4</sup>J<sub>HH</sub> = 1.9 Hz, *ArH*), 8.40 (1H, dd, <sup>4</sup>J<sub>HH</sub> = 1.5 Hz and <sup>4</sup>J<sub>HH</sub> = 1.9 Hz, *ArH*) and 1.62 (9H, s, 3×CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ = 162.3 (CO<sub>2</sub>*t*Bu), 148.8 (*ArC*), 138.2 (*ArCH*), 135.3 (*ArC*), 130.0 (*ArCH*), 123.1 (*ArCH*), 123.0 (*ArC*), 83.5 (qC) and 28.1 (3×CH<sub>3</sub>); MS (Cl<sup>+</sup>) *m/z* 304 ([<sup>81</sup>BrM+H]<sup>+</sup>, 70), 302 ([<sup>79</sup>BrM+H]<sup>+</sup>, 90) and 57.07 ([*t*Bu]<sup>+</sup>, 100); HRMS (Cl<sup>+</sup>) *m/z* calculated for C<sub>11</sub>H<sub>13</sub>NO<sub>4</sub><sup>79</sup>Br [M+H]<sup>+</sup> 302.0028, found 302.0027, calculated for C<sub>11</sub>H<sub>13</sub>NO<sub>4</sub><sup>81</sup>Br [M+H]<sup>+</sup> 304.0007, found 304.0002.

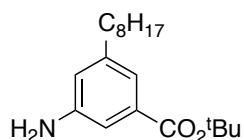
### ***tert*-Butyl 3-nitro-5-(oct-1-ynyl)benzoate**



*tert*-Butyl-3-bromo-5-nitrobenzoate (2 g, 6.6 mmol, 1 eq) was dissolved in vigorously degassed NEt<sub>3</sub> (60 mL) and stirred under N<sub>2</sub>. Octyne (2.87 g, 26.5 mmol, 4 eq), Pd(PPh<sub>3</sub>)<sub>2</sub>Cl<sub>2</sub> (0.42 g, 0.46 mmol, 0.07 eq), Copper Iodide (0.087 g, 0.46 mmol, 0.07 eq) and PPh<sub>3</sub> (0.24 g, 0.92 mmol, 0.14 eq) were added and the solution heated to 80 °C for 16 h. The reaction was then filtered through celite and the solvent removed. The crude material was purified by flash chromatography (15:1,

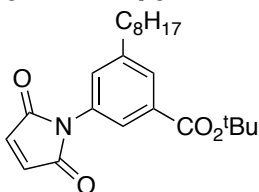
cyclohexane: ethyl acetate) to yield the product as a dark liquid (1.92 g, 5.8 mmol, 88%).  $^1\text{H}$  NMR (400.1 MHz,  $\text{CDCl}_3$ )  $\delta$  = 8.64 (1H, dd,  $^4J_{\text{HH}}$  = 1.5 Hz and  $^4J_{\text{HH}}$  = 2.2 Hz, *ArH*), 8.33 (1H, dd,  $^4J_{\text{HH}}$  = 1.5 Hz and  $^4J_{\text{HH}}$  = 2.2 Hz, *ArH*), 8.25 (1H, t,  $^4J_{\text{HH}}$  = 1.5 Hz, *ArH*), 2.43 (2H, t,  $^3J_{\text{HH}}$  = 7.1 Hz,  $\text{C}\equiv\text{CCH}_2$ ), 1.61 (9H, s,  $3\times\text{CH}_3$ ) 1.51–1.39 (2H, m,  $\text{CH}_2$ ) 1.37–1.27 (6H, m,  $3\times\text{CH}_2$ ) and 0.91 (3H, t,  $^3J_{\text{HH}}$  = 7.1 Hz,  $\text{CH}_3$ );  $^{13}\text{C}$  NMR (100.6 MHz,  $\text{CDCl}_3$ )  $\delta$  = 163.2 ( $\text{CO}_2^t\text{Bu}$ ), 148.2 (*ArC*), 137.9 (*ArCH*), 133.9 (*ArC*), 129.6 (*ArCH*), 126.4 (*ArC*), 123.1, (*ArCH*), 94.7 ( $\text{C}\equiv\text{C}$ ), 82.9 (qC), 78.1 ( $\text{C}\equiv\text{C}$ ), 31.5 ( $\text{CH}_2$ ), 28.8 ( $\text{CH}_2$ ), 28.5 ( $\text{CH}_2$ ), 28.2 ( $3\times\text{CH}_3$ ), 22.4 ( $\text{CH}_2$ ), 19.5 ( $\text{CH}_2$ ) and 14.2 ( $\text{CH}_3$ ); MS ( $\text{Cl}^+$ )  $m/z$  331.18 ( $[\text{M}+\text{H}]^+$ , 100); HRMS: ( $\text{Cl}^+$ )  $m/z$  calculated for  $\text{C}_{19}\text{H}_{26}\text{NO}_4$   $[\text{M}+\text{H}]^+$  332.1862, found 332.1855.

### ***tert*-Butyl 3-amino-5-octylbenzoate**



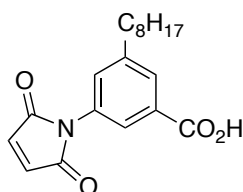
*tert*-Butyl 3-nitro-5-(oct-1-ynyl)benzoate (2 g, 6.03 mmol, 1 eq) was dissolved in HPLC grade MeOH (30 mL) and THF (5 mL) and hydrogenated using an H-Cube, a continuous flow hydrogenation reactor. Using 10% Pd/C as catalyst, the substrate mixture was pumped through the H-Cube at a rate of 2 mL per minute. The hydrogen pressure of the system was set to 50 bar and the temperature to 50 °C. After all of the substrate had passed through the reaction cartridge was washed with methanol (10 mL). The solution was analysed by TLC which showed complete conversion to product. The solvent was then removed under vacuum to yield the amine as a colourless liquid (1.83 g, 5.99 mmol, 99%);  $^1\text{H}$  NMR (400.1 MHz,  $\text{CDCl}_3$ )  $\delta$  = 7.23 (1H, t,  $^4J_{\text{HH}}$  = 1.5 Hz, *ArH*), 7.15 (1H, dd,  $^4J_{\text{HH}}$  = 1.5 and  $^4J_{\text{HH}}$  = 2.2 Hz, *ArH*), 6.70 (1H, dd,  $^4J_{\text{HH}}$  = 1.6 and  $^4J_{\text{HH}}$  = 2.2 Hz, *ArH*), 4.50–3.40 (2H, br,  $\text{NH}_2$ ), 2.54 (2H, t,  $^3J_{\text{HH}}$  = 7.1 Hz, *ArCH}\_2), 1.65–1.52 (11H, m,  $\text{CH}_2$  and  $3\times\text{CH}_3$ ) 1.36–1.19 (10H, m,  $5\times\text{CH}_2$ ) and 0.88 (3H, t,  $^3J_{\text{HH}}$  = 6.9 Hz,  $\text{CH}_3$ );  $^{13}\text{C}$  NMR (100.6 MHz,  $\text{CDCl}_3$ )  $\delta$  = 166.1 ( $\text{CO}_2^t\text{Bu}$ ), 145.6 (*ArC*), 144.3 (*ArC*), 133.0 (*ArC*), 120.4 (*ArCH*), 119.5 (*ArCH*), 113.6, (*ArCH*), 80.8 (qC), 35.9 ( $\text{CH}_2$ ), 31.9 ( $\text{CH}_2$ ), 31.3 ( $\text{CH}_2$ ), 29.5 ( $\text{CH}_2$ ), 29.4 ( $\text{CH}_2$ ), 29.3 ( $\text{CH}_2$ ), 28.2 ( $3\times\text{CH}_3$ ), 22.7 ( $\text{CH}_2$ ) and 14.1 ( $\text{CH}_3$ ); MS ( $\text{Cl}^+$ )  $m/z$  306.24 ( $[\text{M}+\text{H}]^+$ , 45), 305.23 ( $[\text{M}]^+$ , 35) and 250.17 ( $[\text{M}+\text{H}-^t\text{Bu}]^+$ , 100); HRMS ( $\text{Cl}^+$ )  $m/z$  calculated for  $\text{C}_{19}\text{H}_{32}\text{NO}_2$   $[\text{M}+\text{H}]^+$  306.2433, found 306.2438.*

### ***tert*-Butyl 3-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)-5-octylbenzoate**



*tert*-Butyl 3-amino-5-octylbenzoate (0.39 g, 1.6 mmol, 1 eq) and maleic anhydride (0.173 g, 1.76 mmol, 1.1 eq) were dissolved in THF and stirred under N<sub>2</sub> for 16 h. The solvent was then removed and the crude material dissolved in MeCN (10 mL) before addition of ZnBr<sub>2</sub> (0.4 g, 1.76 mmol, 1 eq) and hexamethyldisilazane (1.29 g, 8 mmol, 5 eq). The suspension was then heated to reflux for 1.5 h. After cooling the solvent was removed and saturated EDTA<sub>(aq)</sub> solution (100 mL) added. The product was then extracted into chloroform (3 × 100 mL), the combined organic fractions dried over MgSO<sub>4</sub>, filtered and the solvent removed to give the purified product (0.52 g, 1.35 mmol, 84%) as a colourless liquid; <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 7.82 (1H, t, <sup>4</sup>J<sub>HH</sub> = 1.5 Hz, ArH), 7.74 (1H, t, <sup>4</sup>J<sub>HH</sub> = 1.8 Hz, ArH), 7.29 (1H, t, <sup>4</sup>J<sub>HH</sub> = 1.8 Hz, ArH), 6.87 (2H, s, 2×CH), 2.68 (2H, t, <sup>3</sup>J<sub>HH</sub> = 7.8 Hz, ArCH<sub>2</sub>), 1.68–1.60 (2H, m, CH<sub>2</sub>), 1.57 (9H, s, 3×CH<sub>3</sub>) 1.37–1.18 (10H, m, 5×CH<sub>2</sub>) and 0.88 (3H, t, <sup>3</sup>J<sub>HH</sub> = 7.1 Hz, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ = 169.5 (CO), 165.1 (CO<sub>2</sub><sup>t</sup>Bu), 144.5 (ArC), 134.4 (CH), 133.2 (ArC), 131.2 (ArC), 130.3 (ArCH), 129.2, ArCH), 124.8 (ArCH), 81.6 (qC), 35.8 (CH<sub>2</sub>), 32.0 (CH<sub>2</sub>), 31.3 (CH<sub>2</sub>), 29.5 (CH<sub>2</sub>), 29.4 (CH<sub>2</sub>), 28.3 (CH<sub>2</sub>), 28.3 (<sup>t</sup>but-CH<sub>3</sub>), 22.8 (CH<sub>2</sub>) and 14.3 (CH<sub>3</sub>); MS (ES<sup>+</sup>) *m/z* 408.01 ([M+Na]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>23</sub>H<sub>31</sub>NO<sub>4</sub>Na [M+Na]<sup>+</sup> 408.2151, found 408.2149.

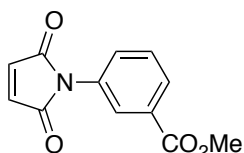
### **3-(2,5-Dioxo-2,5-dihydro-1H-pyrrol-1-yl)-5-octylbenzoic acid (72)**



*tert*-Butyl 3-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)-5-octylbenzoate (0.6 g, 1.56 mmol, 1 eq) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (6 mL) and stirred under N<sub>2</sub>. TFA (3 mL) was then added and the reaction stirred at RT and monitored by TLC analysis. After 1 h the reaction was quenched by addition of water (50 mL) and the product extracted into CH<sub>2</sub>Cl<sub>2</sub> (2 × 50 mL). The combined organics were dried over MgSO<sub>4</sub>, filtered and the solvent removed under vacuum to give the product as a colourless powder. (0.49

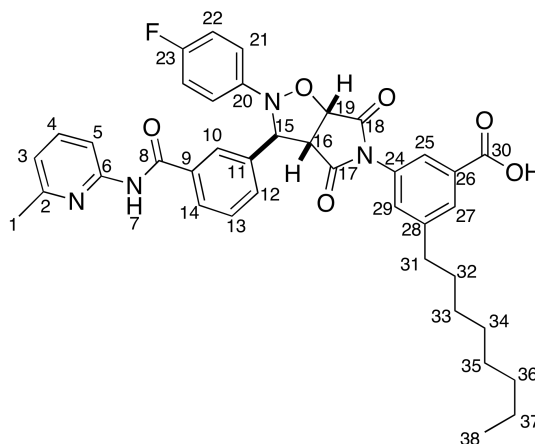
g, 1.5 mmol, 96%); M.p. 127.3–129.0 °C;  $^1\text{H}$  NMR (400.1 MHz,  $\text{CDCl}_3$ )  $\delta$  = 7.93 (2H, m,  $2\times\text{ArH}$ ), 7.41 (1H, t,  $^4J_{\text{HH}}$  = 1.7 Hz,  $\text{ArH}$ ), 6.89 (2H, s,  $2\times\text{CH}$ ), 2.71 (2H, t,  $^3J_{\text{HH}}$  = 7.7 Hz,  $\text{ArCH}_2$ ), 1.75–1.55 (2H, m,  $\text{CH}_2$ ), 1.45–1.10 (10H, m,  $5\times\text{CH}_2$ ) and 0.88 (3H, t,  $^3J_{\text{HH}}$  = 7.1 Hz,  $\text{CH}_3$ );  $^{13}\text{C}$  NMR (100.6 MHz,  $\text{CDCl}_3$ )  $\delta$  = 171.3 ( $\text{CO}_2\text{H}$ ), 169.4 (CO), 144.9 (ArC), 134.4 (CH), 131.6 (ArC), 131.40 (ArCH), 130.4 (ArC), 129.7, (ArCH), 124.2 (ArCH), 35.8 ( $\text{CH}_2$ ), 32.0 ( $\text{CH}_2$ ), 31.3 ( $\text{CH}_2$ ), 29.5 ( $\text{CH}_2$ ), 29.4 ( $\text{CH}_2$ ), 29.3 ( $\text{CH}_2$ ), 22.8 ( $\text{CH}_2$ ) and 14.2 ( $\text{CH}_3$ ); MS ( $\text{Cl}^+$ )  $m/z$  330.17 ( $[\text{M}+\text{H}]^+$ , 100) and 312.16 (30); HRMS ( $\text{Cl}^+$ )  $m/z$  calculated for  $\text{C}_{19}\text{H}_{24}\text{NO}_4$   $[\text{M}+\text{H}]^+$  330.1705, found 330.1707.

### Methyl 3-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)benzoate<sup>156</sup> (73)



Methyl iodide (0.46 g, 3.2 mmol, 2 eq) was added to a stirring solution of 3-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)benzoic acid<sup>156</sup> (0.35 g, 1.6 mmol, 1 eq) and caesium carbonate (0.26 g, 1.6 mmol, 1 eq) in DMSO (15 mL). The mixture was left stirring in the dark under nitrogen for 16 h. The mixture was diluted with water and extracted with DCM. The combined extracts were dried over  $\text{MgSO}_4$ . The solvent was removed and the residue was purified using silica gel column chromatography ( $\text{CH}_2\text{Cl}_2$ ) affording the desired product (0.32 g, 86%). M.p. = 129.6-129.8 °C (lit<sup>156</sup> 129.6-129.8 °C)  $^1\text{H}$  NMR (300.1 MHz,  $\text{CDCl}_3$ )  $\delta$  = 8.01-7.97 (2H, m,  $2\times\text{ArH}$ ), 7.50-7.48 (2H, m,  $2\times\text{ArH}$ ), 6.81 (2H, s,  $2\times\text{CH}$ ), 3.86 (3H, s,  $\text{CH}_3$ );  $^{13}\text{C}$  NMR (100.6 MHz,  $d_6$ -DMSO)  $\delta$  = 169.7 (CO), 165.5 (CO), 134.7 ( $2\times\text{CH}$ ), 132.0 (ArC), 131.3 (ArCH), 130.3 (ArC), 129.4 (ArCH), 128.1 (ArCH), 127.1 (ArCH), 52.3 ( $\text{CH}_3$ ); MS ( $\text{ES}^+$ ):  $m/z$  (%): 254 ( $[\text{M}+\text{Na}]^+$ , 100); HRMS ( $\text{ES}^+$ ): calculated for  $\text{C}_{12}\text{H}_9\text{NO}_4\text{Na}$   $[\text{M}+\text{Na}]^+$  254.0434, found 254.0429.

**3-(2-(4-Fluorophenyl)-3-(3-(6-methylpyridin-2-ylcarbamoyl)phenyl)-4,6-dioxodihydro-2H-pyrrolo[3,4-d]isoxazol-5(3H,6H,6aH)-yl)-5-octylbenzoic acid (*trans*-76)**

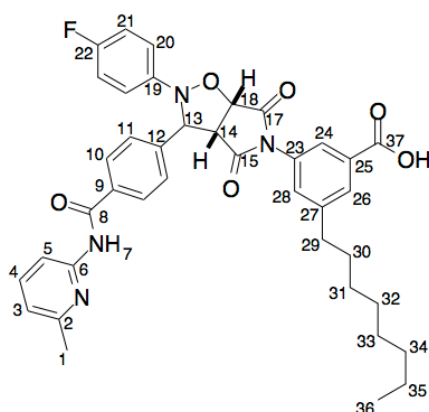


Nitrone **58** (0.011 g, 0.03 mmol, 1 eq) and maleimide **72** (0.010 g, 0.03 mmol, 1 eq) were dissolved in  $\text{CDCl}_3$  (1.5, mL), cooled to 0 °C and allowed to react for 72 h. The solution was then heated to RT and left to react for 24 h before analysis of the product by NMR spectroscopy which showed 99% conversion to product with a *trans*:*cis* ratio of 17:1. The solvent was then removed to yield the crude product as a colourless powder which could be recrystallised from methanol and hexane to give the purified cycloadduct *trans*-**76** (0.01 g, 0.015 mmol, 50%) as a colourless powder; M.p. 129.2 – 132.0 °C;  $^1\text{H}$  NMR (499.9 MHz,  $\text{CDCl}_3$ )  $\delta$  = 11.58–11.28 (1H, b, NH,  $H_7$ ), 8.35 (1H, d,  $^3J_{\text{HH}}$  = 8.4 Hz, ArH,  $H_5$ ), 8.20 (1H, s, ArH,  $H_{10}$ ), 8.07 (1H, d,  $^3J_{\text{HH}}$  = 7.8 Hz, ArH,  $H_{14}$ ), 7.87 (1H, s, ArH,  $H_{27}$ ), 7.80 (1H, t,  $^3J_{\text{HH}}$  = 8.0 Hz, ArH,  $H_4$ ), 7.77 (1H, d,  $^3J_{\text{HH}}$  = 7.4 Hz, ArH,  $H_{12}$ ), 7.58 (1H, t,  $^3J_{\text{HH}}$  = 7.6 Hz, ArH,  $H_{13}$ ), 7.31 (1H, s, ArH,  $H_{25}$ ), 7.09–6.99 (3H, m, 2×ArH,  $H_{22}$  and ArH,  $H_3$ ), 6.74–6.57 (3H, m, ArH $_{29}$  and 2×ArH,  $H_{21}$ ), 5.71 (1H, s, CH,  $H_{15}$ ), 5.20 (1H, d,  $^3J_{\text{HH}}$  = 7.2 Hz, CH,  $H_{16}$ ), 4.56 (1H, b, CH,  $H_{19}$ ), 2.61 (3H, s, ArCH $_3$ ,  $H_1$ ), 2.59 (2H, t,  $^3J_{\text{HH}}$  = 8.1 Hz, CH $_2$ ,  $H_{31}$ ), 1.60–1.51 (2H, m, CH $_2$ ,  $H_{32}$ ), 1.37 – 1.18 (10H, m, 5×CH $_2$ ,  $H_{33}$ ,  $H_{34}$ ,  $H_{35}$ ,  $H_{36}$ ,  $H_{37}$ ) and 0.87 (3H, t,  $^3J_{\text{HH}}$  = 6.7 Hz, Alk-CH $_3$ ,  $H_{38}$ );  $^{13}\text{C}$  NMR (125.7 MHz,  $\text{CDCl}_3$ )  $\delta$  = 175.1 (CO,  $C_{17}$ ), 172.8 (CO,  $C_{19}$ ), 171.0 (CO $_2$ H,  $C_{30}$ ), 167.0 (NHCO,  $C_8$ ), 158.9 (d,  $^1J_{\text{CF}}$  = 243 Hz, ArC-F,  $C_{23}$ ), 155.6 (ArC,  $C_2$ ), 151.7 (ArC,  $C_6$ ), 144.8 (ArC,  $C_{28}$ ), 144.7 (d,  $^4J_{\text{CF}}$  = 2.6 Hz, ArC,  $C_{20}$ ), 140.9 (ArCH,  $C_4$ ), 138.4 (ArC,  $C_{11}$ ), 135.4 (ArCH,  $C_9$ ), 131.8 (ArC,  $C_{26}$ ), 131.6 (ArCH,  $C_{12}$ ), 131.1 (ArC,  $C_{24}$ ), 130.3 (ArCH,  $C_{27}$ ), 130.3 (ArCH,  $C_{29}$ ), 129.6 (ArCH,  $C_{14}$ ), 129.5 (ArCH,  $C_{13}$ ), 125.6 (ArCH,  $C_{10}$ ), 124.4 (ArCH,  $C_{25}$ ), 120.1 (ArCH,  $C_3$ ), 116.3 (d,  $^3J_{\text{CF}}$  = 7.5 Hz, ArCH,  $C_{21}$ ), 116.0 (d,  $^2J_{\text{CF}}$  = 22.8 Hz, ArCH,  $C_{22}$ ), 113.6 (ArCH,  $C_5$ ), 77.6 (CH,  $C_{16}$ ), 70.3 (CH,  $C_{15}$ ), 55.6 (CH,  $C_{19}$ ), 35.8 (CH $_2$ ,  $C_{31}$ ), 32.0 (CH $_2$ ,  $C_{32}$ ), 31.4 (CH $_2$ ,  $C_{33}$ ),



29.5 (CH<sub>2</sub>, C<sub>34</sub>), 29.5 (CH<sub>2</sub>, C<sub>35</sub>), 29.4 (CH<sub>2</sub>, C<sub>36</sub>), 22.8 (CH<sub>3</sub>, C<sub>1</sub>), 22.1 (CH<sub>2</sub>, C<sub>37</sub>) and 14.3 (CH<sub>3</sub>, C<sub>38</sub>); <sup>19</sup>F {<sup>1</sup>H} NMR (470.3 MHz, CDCl<sub>3</sub>) δ = 470.3 Hz, -120.1 (1F, s, ArF); MS (ES<sup>+</sup>) *m/z* 678.91 (MH<sup>+</sup>, 100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>39</sub>H<sub>40</sub>N<sub>4</sub>O<sub>6</sub>F [M+H<sup>+</sup>] 679.2931, found 679.2319.

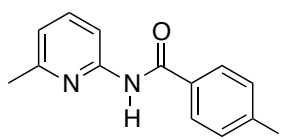
**3-(2-(4-fluorophenyl)-3-(4-(6-methylpyridin-2-ylcarbamoyl)phenyl)-4,6-dioxodihydro-2H-pyrrolo[3,4-d]isoxazol-5(3H,6H,6aH)-yl)-5-octylbenzoic acid (*trans*-74)**



Nitone **52** (0.011 g, 0.03 mmol, 1 eq) and maleimide **72** (0.010 g, 0.03 mmol, 1 eq) were dissolved in CHCl<sub>3</sub> (1.5, mL), cooled to 0 °C and allowed to react for 72 h. The solution was then heated to RT and left to react for 24 h before the solvent was removed under vacuum and the product was analysed by NMR spectroscopy which showed 99% conversion to product with a *trans*:*cis* ratio of 7:1. The *trans* product was then enriched by recrystallisation from ethanol and hexane to give *trans*-**74** (0.014 g, 0.02 mmol, 70%) as a colourless powder; M.p. 209.0 – 209.9 °C; <sup>1</sup>H {<sup>19</sup>F} NMR (499.9 MHz, d<sub>6</sub>-DMSO) δ = 10.75 (1H, s, NH, H<sub>7</sub>), 8.08 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.0 Hz, 2×ArH, H<sub>10</sub>), 8.02 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.3 Hz, ArH, H<sub>5</sub>), 7.77 (1H, s, ArH, H<sub>26</sub>), 7.73 (1H, t, <sup>3</sup>J<sub>HH</sub> = 8.3 Hz, ArH, H<sub>4</sub>), 7.67 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.0 Hz, 2×ArH, H<sub>11</sub>), 7.31 (1H, s, ArH, H<sub>24</sub>), 7.28 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.7 Hz, 2×ArH, H<sub>20</sub>), 7.10 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.7 Hz, 2×ArH, H<sub>21</sub>), 7.04 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.4 Hz, ArH, H<sub>3</sub>), 6.50 (1H, s, ArH, H<sub>28</sub>), 5.92 (1H, s, CH, H<sub>13</sub>), 5.46 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.2 Hz, CH, H<sub>18</sub>), 4.17 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.4 Hz, CH, H<sub>14</sub>), 2.58 (2H, t, <sup>3</sup>J<sub>HH</sub> = 7.6 Hz, ArCH<sub>2</sub>, H<sub>29</sub>), 2.46 (3H, s, ArCH<sub>3</sub>, H<sub>1</sub>), 1.57–1.44 (2H, m, CH<sub>2</sub>, H<sub>30</sub>), 1.34–1.18 (10H, m, 5×CH<sub>2</sub>, H<sub>31</sub>, H<sub>32</sub>, H<sub>33</sub>, H<sub>34</sub> & H<sub>35</sub>) and 0.86 (3H, t, <sup>3</sup>J<sub>HH</sub> = 6.5 Hz, Alk-CH<sub>3</sub>, H<sub>36</sub>); <sup>13</sup>C NMR (125.7 MHz, d<sub>6</sub>-DMSO) δ = 174.3 (CO, C<sub>15</sub>), 173.1 (CO, C<sub>17</sub>), 166.4 (CO<sub>2</sub>H, C<sub>37</sub>), 165.5 (NHCO, C<sub>8</sub>), 158.0 (d, <sup>1</sup>J<sub>CF</sub> = 240 Hz, ArCF, C<sub>22</sub>), 156.6 (ArC,

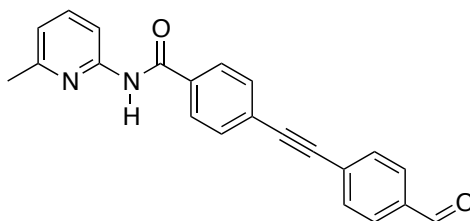
C<sub>2</sub>), 151.5 (ArC, C<sub>6</sub>), 145.2 (d, <sup>4</sup>J<sub>CF</sub> = 2.0 Hz, ArC, C<sub>19</sub>), 143.9 (ArC, C<sub>29</sub>), 142.7 (ArC, C<sub>12</sub>), 138.4 (ArCH, C<sub>4</sub>), 133.6 (ArC, C<sub>9</sub>), 131.7 (ArC, C<sub>23</sub>), 131.6 (ArC, C<sub>25</sub>), 130.2 (ArCH, C<sub>28</sub>), 129.3 (ArCH, C<sub>26</sub>), 128.4 (ArCH, C<sub>10</sub>), 127.1 (ArCH, C<sub>11</sub>), 124.8 (ArCH, C<sub>24</sub>), 119.1 (ArCH, C<sub>3</sub>), 116.2 (d, <sup>3</sup>J<sub>CF</sub> = 7.8 Hz, ArCH, C<sub>20</sub>), 115.7 (d, <sup>2</sup>J<sub>CF</sub> = 22.4 Hz, ArCH, C<sub>21</sub>), 111.7 (ArCH, C<sub>5</sub>), 78.0 (CH, C<sub>18</sub>), 68.4 (CH, C<sub>13</sub>), 56.5 (CH, C<sub>14</sub>), 34.6 (CH<sub>2</sub>, C<sub>29</sub>), 31.3 (CH<sub>2</sub>, C<sub>30</sub>), 30.8 (CH<sub>2</sub>, C<sub>31</sub>), 28.7 (CH<sub>2</sub>, C<sub>32</sub>), 28.7 (CH<sub>2</sub>, C<sub>33</sub>), 28.6 (CH<sub>2</sub>, C<sub>34</sub>), 23.6 (CH<sub>3</sub>, C<sub>1</sub>), 22.1 (CH<sub>2</sub>, C<sub>35</sub>) and 14.0 (CH<sub>3</sub>, C<sub>36</sub>); <sup>19</sup>F {<sup>1</sup>H} NMR (470.3 MHz, d<sub>6</sub>-DMSO) δ = -121.34 (1F, s, ArF); MS (ES<sup>-</sup>) *m/z* 676.96 ([M-H]<sup>-</sup>, 100); HRMS (ES<sup>-</sup>) *m/z* calculated for C<sub>39</sub>H<sub>38</sub>N<sub>4</sub>O<sub>6</sub>F [M-H]<sup>-</sup> 677.2775, found 677.2778.

#### 4-Iodo-N-(6-methylpyridin-2-yl)benzamide



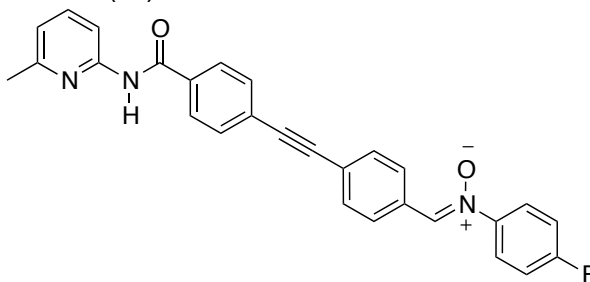
2-Amino-6-methylpyridine (2 g, 18.5 mmol, 1.1 eq) and triethylamine (5.1 g, 50.4 mmol, 3 eq) were dissolved in CH<sub>2</sub>Cl<sub>2</sub> (15 mL) and cooled to 0 °C under N<sub>2</sub>. A solution of 4-iodobenzoyl chloride (4.48 g, 16.8 mmol, 1 eq) in CH<sub>2</sub>Cl<sub>2</sub> (20 mL) was added in a dropwise manner and after complete addition, the reaction was allowed to rise to RT and stirred for 16 h. NaHCO<sub>3</sub> (aq) (100 mL) was added to quench the reaction and the product extracted into CH<sub>2</sub>Cl<sub>2</sub> (2 × 100 mL). The organic fractions were dried (MgSO<sub>4</sub>), filtered and the solvent removed under vacuum to give the crude product which was purified recrystallisation from CH<sub>2</sub>Cl<sub>2</sub> to yield the amide as a colourless powder (4.3 g, 13.0 mmol, 70%). M.p. 118.0–119.6 °C; <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 8.70–8.59 (1H, br, NH), 8.15 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.2 Hz, ArH), 7.82 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.3 Hz, 2×ArH), 7.67–7.60 (3H, m, 3×ArH), 6.92 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.5 Hz, ArH) and 2.43 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ = 164.9 (CONH), 157.0 (ArC), 150.6 (ArC), 138.9 (ArCH), 138.0 (ArCH), 133.8 (ArC), 128.9 (ArCH), 119.7 (ArCH), 111.1 (ArCH), 99.4 (ArC-I) and 24.0 (CH<sub>3</sub>); MS (EI<sup>+</sup>) *m/z* 338 ([M], 23), 310 (57), 231 (82), 108 (100); HRMS (EI<sup>+</sup>) *m/z* calculated for C<sub>13</sub>H<sub>11</sub>IN<sub>2</sub>ONa [M+Na]<sup>+</sup> 360.3824, found 360.9814

#### 4-((4-Formylphenyl)ethynyl)-N-(6-methylpyridin-2-yl)benzamide



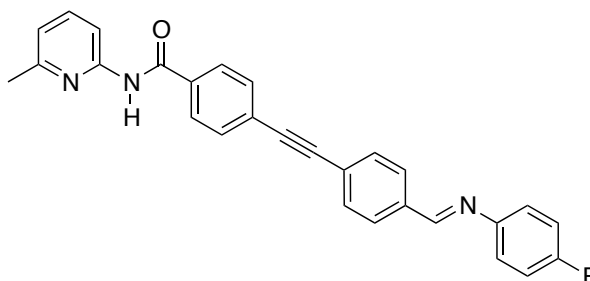
Iodoamide (2.67 g, 7.9 mmol, 1 eq), 4-ethynylbenzaldehyde (2.05 g, 15.8 mmol, 2 eq) and triphenylphosphine (0.2 g, 0.8 mmol, 0.1 eq) were dissolved in triethylamine (100 mL) and the solution vigorously degassed. Pd(P(Ph)<sub>3</sub>)<sub>2</sub>Cl<sub>2</sub> (0.36 g, 0.4 mmol, 0.05 eq) and copper iodide (0.075 g, 0.04 mmol, 0.05 eq) were then added to initiate the reaction which was heated to 50 °C and the reaction progress monitored by TLC analysis. Upon completion (16 h) the solution was cooled and filtered through celite. The solvent was then removed and the residue washed with NH<sub>4</sub>Cl<sub>(aq)</sub> (100 mL), extracted into CH<sub>2</sub>Cl<sub>2</sub> (3 × 100 mL), dried (MgSO<sub>4</sub>), filtered and the crude mixture adsorbed on to MgSO<sub>4</sub> for dry loading on to a chromatography column. The byproducts were first removed by washing through the column with 4:1 cyclohexane:ethyl acetate before eluting the product with a 1:1 mixture. The solvent was then removed to yield the long aldehyde as a pale yellow powder (2.62 g, 7.7 mmol, 97%); M.p. 194.0–194.9 °C; <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 10.01 (1H, s, CHO), 8.74 (1H, br, NH), 8.17 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.2 Hz, ArH), 7.93 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 2×ArH), 7.87 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 2×ArH), 7.70–7.60 (5H, m, 5×ArCH), 6.92 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.5 Hz, ArH) and 2.44 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ = 191.4 (HCO), 164.8 (ArC), 157.0 (ArC), 150.7 (ArC), 138.9 (ArCH), 135.8 (ArC), 134.3 (ArC), 132.3 (ArCH), 132.1 (ArCH), 129.6 (ArCH), 128.9 (ArC), 127.4 (ArCH), 126.4 (ArC), 119.7 (ArCH), 111.1 (ArCH), 92.2 (qC), 91.1 (qC) and 23.6 (CH<sub>3</sub>); MS (Cl<sup>+</sup>) *m/z* 341 ([M+H]<sup>+</sup>, 100), 312 (25), 233 (20); HRMS (Cl<sup>+</sup>) *m/z* calculated for C<sub>22</sub>H<sub>17</sub>N<sub>2</sub>O<sub>2</sub> [M+H]<sup>+</sup> 341.1304, found 341.1290.

**(Z)-4-fluoro-N-(4-((4-(6-methylpyridin-2-yl)carbamoyl)phenyl)ethynyl)benzylidene)aniline oxide (83)**



4-((4-Formylphenyl)ethynyl)-N-(6-methylpyridin-2-yl)benzamide (0.3 g, 0.88 mmol, 1 eq) and *N*-(4-fluorophenyl)hydroxylamine **46** (0.11 g, 0.88 mmol, 1 eq) were dissolved in EtOH (15 mL) and CH<sub>2</sub>Cl<sub>2</sub> (5 mL) and left to sit in darkness at RT for 72 h. The precipitate formed was isolated by filtration and washed with a little cold ethanol to furnish nitrone **83** (0.3 g, 0.66 mmol, 76 %) as a pale yellow crystalline solid. M.p. 214.2–218.8 °C; <sup>1</sup>H {<sup>19</sup>F} NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 8.74 (1H, br, NH), 8.45 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.3 Hz, 2×ArH), 8.17 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.2 Hz, ArH), 8.00 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.6 Hz, 2×ArH), 7.95 (1H, s, CH), 7.78 (2H, t, <sup>3</sup>J<sub>HH</sub> = 9.1 Hz, 2×ArH), 7.64 (5H, m, 5×ArH), 7.16 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8 Hz, 2×ArH), 6.93 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.5 Hz, ArH) and 2.45 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR δ = (100.6 MHz, CDCl<sub>3</sub>) 164.7 (NHCO), 162.3 (d, <sup>1</sup>J<sub>CF</sub> = 255 Hz, ArCF), 157.0 (ArC), 150.7 (ArC), 145.1 (d, <sup>4</sup>J<sub>CF</sub> = 3.3 Hz, ArC), 138.8 (ArCH), 135.8 (ArC), 133.7 (ArC), 133.3 (ArCH), 129.0 (ArCH), 127.6 (CH), 123.7 (d, <sup>3</sup>J<sub>CF</sub> = 8.8 Hz, ArCH), 119.7 (ArCH), 116.2 (d, <sup>2</sup>J<sub>CF</sub> = 23.2 Hz, ArCH), 111.2 (ArCH) and 24.0 (CH<sub>3</sub>); <sup>19</sup>F {<sup>1</sup>H} (376.5 MHz, CDCl<sub>3</sub>) δ = -110.09 (1F, s, ArF); MS (ES<sup>+</sup>) *m/z* 449.99 ([M + H]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>28</sub>H<sub>21</sub>N<sub>3</sub>O<sub>2</sub>F [M+H]<sup>+</sup> 450.1618, found 450.1615.

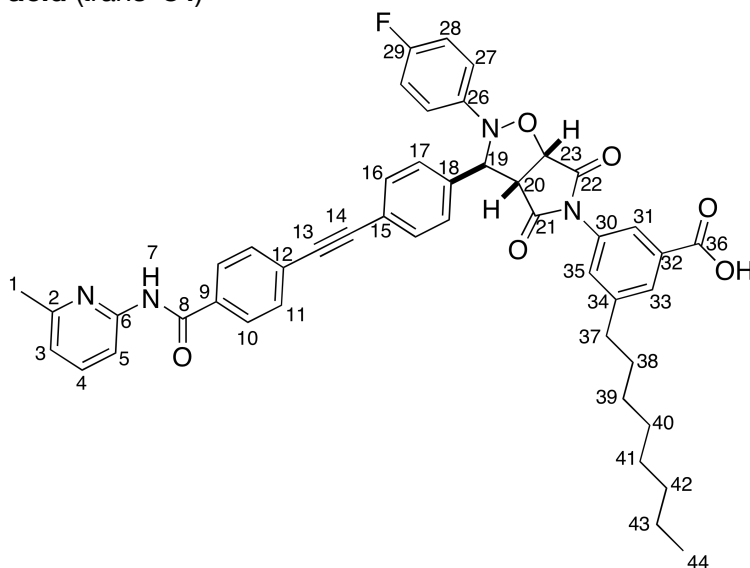
**(E)-4-((4-((4-fluorophenylimino)methyl)phenyl)ethynyl)-N-(6-methylpyridin-2-yl)benzamide (86)**



4-Fluoroaniline (0.065 g, 0.57 mmol, 1 eq) and long aldehyde (0.2 g, 0.57 mmol, 1 eq) were dissolved in ethanol (20 mL) and left to sit for 16 h. The precipitate formed

was isolated by filtration and washed with a little cold ethanol to yield imine **86** (0.12 g, 0.27 mmol, 49%) as a colourless fine crystalline solid; M.p. 188.2–190.4 °C;  $^1\text{H}$   $\{^{19}\text{F}\}$  NMR (400.1 MHz,  $\text{CDCl}_3$ )  $\delta$  = 8.59 (1H, br, *NH*), 8.45 (1H, s, *CH*), 8.17 (1H, d,  $^3J_{\text{HH}}$  = 8.2 Hz, *ArH*), 7.96–7.84 (4H, m, 4×*ArH*), 7.70–7.58 (5H, m, 5×*ArH*), 7.22 (2H, dt,  $^3J_{\text{HH}}$  = 9.0 and  $^4J_{\text{HH}}$  = 2.3 Hz, 2×*ArH*), 7.09 (2H, dt,  $^3J_{\text{HH}}$  = 9.0 and  $^4J_{\text{HH}}$  = 2.3 Hz, 2×*ArH*), 6.93 (1H, d,  $^3J_{\text{HH}}$  = 7.5 Hz, *ArH*) and 2.47 (3H, s,  $\text{CH}_3$ );  $^{13}\text{C}$  NMR (100.6 MHz,  $\text{CDCl}_3$ ) 164.9 (*NHCO*), 161.5 (d,  $^1J_{\text{CF}}$  245 Hz, *ArC*), 159.0 (*ArC*), 159.0 (*CH*), 157.1 (*ArC*), 150.8 (*ArC*), 147.8 (d,  $^4J_{\text{CF}}$  = 2.8 Hz, *ArC*), 138.9 (*ArCH*), 136.2 (*ArC*), 133.0 (*ArCH*), 132.2 (*ArCH*), 132.1 (*ArCH*), 128.8 (*ArCH*), 127.4 (*ArCH*), 127.0 (*ArC*), 125.7 (*ArC*), 122.5 (d,  $^3J_{\text{CF}}$  = 8.2 Hz, *ArCH*), 119.7 (*ArCH*), 116.1 (d,  $^2J_{\text{CF}}$  = 22.3 Hz, *ArCH*), 111.1 (*ArCH*), 92.0 (*C*), 90.9 (*C*) and 24.1 ( $\text{CH}_3$ );  $^{19}\text{F}$   $\{^1\text{H}\}$  NMR (282.4 MHz,  $\text{CDCl}_3$ )  $\delta$  = –116.7 (*ArF*); MS ( $\text{ES}^+$ ) *m/z* 455.9 ( $[\text{M}+\text{Na}]^+$ , 100); HRMS ( $\text{ES}^+$ ) *m/z* calculated for  $\text{C}_{28}\text{H}_{21}\text{N}_3\text{OFNa}$   $[\text{M}+\text{H}]^+$  434.1669, found 434.1655.

**3-(2-(4-fluorophenyl)-3-(4-((4-(6-methylpyridin-2-ylcarbamoyl)phenyl)ethynyl)phenyl)ethynyl)phenyl)-4,6-dioxodihydro-2H-pyrrolo[3,4-d]isoxazol-5(3H,6H,6aH)-yl)-5-octylbenzoic acid (*trans*-**84**)**

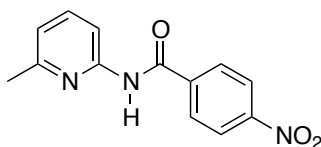


Nitrone **83** (0.014 g, 0.03 mmol, 1 eq) and maleimide **72** (0.010 g, 0.03 mmol, 1 eq) were dissolved in  $\text{CHCl}_3$  (1.5, mL), cooled to 10 °C and allowed to react for 96 h. The solution was then heated to RT and left to react for 48 h before the solvent was removed under vacuum and the product analysed by NMR spectroscopy which showed quantitative conversion to product with a *trans*:*cis* ratio of 20:1. The *trans* product was then enriched by recrystallisation from ethanol and hexane to give

374

*trans*-**84** (0.021 g, 0.027 mmol, 91%) as a colourless powder; M.p. 130.0 – 135.0 °C; <sup>1</sup>H NMR (499.9 MHz, d<sub>6</sub>-DMSO) δ = 10.82 (1H, s, NH, H<sub>7</sub>), 8.07 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 2×ArH, H<sub>10</sub>), 8.01 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.1 Hz, ArH, H<sub>5</sub>), 7.80 – 7.60 (8H, m, 8×ArH, H<sub>4</sub>, H<sub>11</sub>, H<sub>16</sub>, H<sub>17</sub> & H<sub>33</sub>), 7.33 (1H, t, <sup>4</sup>J<sub>HH</sub> = 1.6 Hz, ArH, H<sub>31</sub>), 7.25 (2H, dd, <sup>3</sup>J<sub>HH</sub> 9.4 Hz and <sup>4</sup>J<sub>HF</sub> = 4.5, 2×ArH, H<sub>27</sub>), 7.15–7.00 (3H, m, 3×ArH, H<sub>3</sub> & H<sub>28</sub>), 6.53 (1H, t, <sup>4</sup>J<sub>HH</sub> = 1.5 Hz, ArH, H<sub>35</sub>), 5.95 (1H, s, CH, H<sub>19</sub>), 5.46 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.2 Hz, CH, H<sub>23</sub>), 4.16 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.2 Hz, CH, H<sub>20</sub>), 2.59 (2H, t, <sup>3</sup>J<sub>HH</sub> = 7.5 Hz, PhCH<sub>2</sub>, H<sub>37</sub>), 2.46 (3H, s, ArCH<sub>3</sub>, H<sub>1</sub>), 1.61–1.44 (2H, m, CH<sub>2</sub>, H<sub>38</sub>), 1.33–1.18 (10H, m, 5×CH<sub>2</sub>, H<sub>39</sub>, H<sub>40</sub>, H<sub>41</sub>, H<sub>42</sub> & H<sub>43</sub>) and 0.86 (3H, t, <sup>3</sup>J<sub>HH</sub> = 6.6 Hz, Alk-CH<sub>3</sub>, H<sub>44</sub>); <sup>13</sup>C NMR (125.7 MHz, d<sub>6</sub>-DMSO) δ = 174.2 (CO, C<sub>21</sub>), 173.1 (CO, C<sub>22</sub>), 166.3 (CO<sub>2</sub>H, C<sub>36</sub>), 165.1 (NHCO, C<sub>8</sub>), 158.0 (d, <sup>1</sup>J<sub>CF</sub> = 240 Hz, ArCF, C<sub>29</sub>), 156.6 (ArC, C<sub>2</sub>), 151.4 (ArC, C<sub>6</sub>), 145.0 (d, <sup>4</sup>J<sub>CF</sub> = 1.7 Hz, ArC, C<sub>26</sub>), 143.9 (ArC, C<sub>34</sub>), 139.8 (ArC, C<sub>18</sub>), 138.4 (ArCH, C<sub>4</sub>), 134.0 (ArC, C<sub>12</sub>), 131.8 (ArC, C<sub>30</sub>), 131.7 (ArCH, C<sub>16</sub>), 131.6 (ArC, C<sub>32</sub>), 131.2 (ArCH, C<sub>11</sub>), 130.2 (ArCH, C<sub>35</sub>), 129.3 (ArCH, C<sub>33</sub>), 128.4 (ArCH, C<sub>10</sub>), 127.7 (ArCH, C<sub>17</sub>), 125.5 (ArC, C<sub>9</sub>), 124.8 (ArCH, C<sub>31</sub>), 121.4 (ArC, C<sub>15</sub>), 119.2 (ArCH, C<sub>3</sub>), 116.3 (d, <sup>3</sup>J<sub>CF</sub> = 7.8 Hz, ArCH, C<sub>27</sub>), 115.6 (d, <sup>2</sup>J<sub>CF</sub> = 22.6 Hz, ArCH, C<sub>28</sub>), 111.7 (ArCH, C<sub>5</sub>), 91.2 (qC, C<sub>14</sub>), 89.1 (qC, C<sub>13</sub>), 77.8 (CH, C<sub>23</sub>), 68.5 (CH, C<sub>19</sub>), 56.3 (CH, C<sub>20</sub>), 34.6 (CH<sub>2</sub>, C<sub>37</sub>), 31.2 (CH<sub>2</sub>, C<sub>38</sub>), 30.7 (CH<sub>2</sub>, C<sub>39</sub>), 28.7 (CH<sub>2</sub>, C<sub>40</sub>), 28.6 (CH<sub>2</sub>, C<sub>41</sub>), 28.5 (CH<sub>2</sub>, C<sub>42</sub>), 23.6 (ArCH<sub>3</sub>, C<sub>1</sub>), 22.1 (CH<sub>2</sub>, C<sub>43</sub>) and 13.9 (Alk-CH<sub>3</sub>, C<sub>44</sub>); <sup>19</sup>F {<sup>1</sup>H} NMR (470.3 MHz, d<sub>6</sub>-DMSO-d<sub>6</sub>) δ = -121.31 (1F, s, ArF); MS (ES<sup>+</sup>) *m/z* 778.91 ([M+H]<sup>+</sup>, 35) and 340.96 (100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>47</sub>H<sub>44</sub>N<sub>4</sub>O<sub>6</sub>F [M+H]<sup>+</sup> 779.3245, found 779.3234.

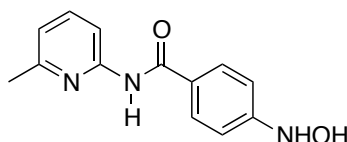
***N*-(6-methylpyridin-2-yl)-4-nitrobenzamide<sup>219</sup> (123)**



2-Amino-6-methylpyridine (2.56 g, 22 mmol, 1.1 eq) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (25 mL) and triethylamine (10mL) and cooled to 0 °C under N<sub>2</sub>. A solution of 4-nitrobenzoylchloride (4g, 20 mmol, 1 eq) in CH<sub>2</sub>Cl<sub>2</sub> (50 mL) was then added in a dropwise fashion to the amine solution. The temperature of the reaction was allowed to rise slowly to room temperature and was stirred until TLC analysis showed the reaction had gone to completion (typically 16 h). The reaction was quenched by addition of water (100 mL) and the product extracted into CH<sub>2</sub>Cl<sub>2</sub> (2 × 100 mL), dried (MgSO<sub>4</sub>), and purified by recrystallisation (EtOAc:cyclohexane) to yield *N*-(6-

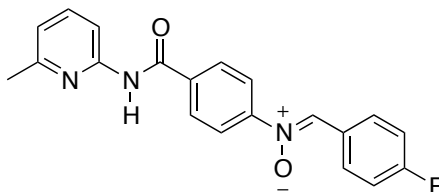
methylpyridin-2-yl)-4-nitrobenzamide as a pale yellow crystalline solid (4.98 g, 19 mmol, 97%); M.p. 127.5–129.0 °C (lit<sup>219</sup> 130–131 °C) <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ = 8.45–8.55 (1H, br, NH), 8.28 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8 Hz, 2×ArH), 8.10 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8 Hz, ArH), 8.01 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8 Hz, 2×ArH), 7.62 (1H, t, <sup>3</sup>J<sub>HH</sub> = 8 Hz, ArH), 6.91 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.0 Hz, ArH) and 2.45 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz, d<sub>6</sub>-DMSO) δ = 164.5 (NHCO), 156.1 (ArC–NO<sub>2</sub>), 150.7 (ArC), 149.3 (ArC), 139.6 (ArC), 139.3 (ArCH), 129.6 (ArCH), 123.4 (ArCH), 119.7 (ArCH), 112.0 (ArCH) and 23.1 (CH<sub>3</sub>); MS (ES<sup>+</sup>) *m/z* 280.12 ([M+Na]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>13</sub>H<sub>12</sub>N<sub>3</sub>O<sub>3</sub> [M+H]<sup>+</sup> 258.0879, found 258.0872.

#### 4-(Hydroxyamino)-N-(6-methylpyridin-2-yl)benzamide (124)



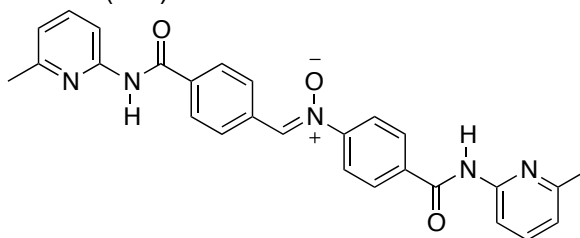
*N*-(6-methylpyridin-2-yl)-4-nitrobenzamide **123** (1g, 3.9 mmol, 1 eq) was dissolved in an ethanol:water solution (20 mL, 10:1) before the addition of bismuth chloride (0.25 g, 0.8 mmol, 0.2 eq). The reaction flask was placed in a water bath to maintain the temperature (25 °C) whilst potassium borohydride (0.32 g, 5.9 mmol, 1.5 eq) was added in portions over 5 minutes. After addition was complete, the suspension was stirred for a further 15 minutes before neutralising excess borohydride by adjusting to pH 7 by addition of 0.5 M HCl (aq). The solution was then filtered through a thin layer of celite and the product extracted into CH<sub>2</sub>Cl<sub>2</sub> (3 × 100mL), dried (MgSO<sub>4</sub>), filtered and the solvent removed to yield 4-(hydroxyamino)-*N*-(6-methylpyridin-2-yl) benzamide (0.8 g, 3.3 mmol, 84%) as a colourless crystalline solid which was used without further purification. M.p. decomposed >120 °C; <sup>1</sup>H NMR (300.1 MHz, CDCl<sub>3</sub>) δ = 8.95 – 9.10 (1H, br, NH), 8.02 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8 Hz, ArH), 7.63 (1H, t, <sup>3</sup>J<sub>HH</sub> = 8 Hz, ArH), 7.50 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8 Hz, 2×ArH), 6.88 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8 Hz, ArH) and 6.81 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8 Hz, 2×ArH); <sup>13</sup>C NMR (75.5 MHz, d<sub>6</sub>-DMSO) δ = 165.5 (CONH), 156.4 (ArC), 155.1 (ArC), 151.9 (ArC), 138.2 (ArCH), 129.2 (ArCH), 123.8 (ArC), 118.5 (ArCH), 111.5 (ArCH), 111.3 (ArCH) and 23.6 (CH<sub>3</sub>); MS (ES<sup>+</sup>) *m/z* 266.14 ([M+Na]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>13</sub>H<sub>13</sub>N<sub>3</sub>O<sub>2</sub>Na [M+Na]<sup>+</sup> 266.0905, found 266.0897

**(Z)-N-(4-fluorobenzylidene)-4-(6-methylpyridin-2-ylcarbamoyl)aniline oxide (121)**



4-(Hydroxyamino)-*N*-(6-methylpyridin-2-yl)benzamide **124** (0.2 g, 0.82 mmol, 1 eq) and 4-fluorobenzaldehyde (0.1 g, 0.82 mmol, 1 eq) were dissolved in ethanol (20 mL) and left to stand in darkness for 16 h. The precipitate was isolated by filtration, washed with a little cold ethanol and dried to yield 4-[(*Z*)-(4-fluorobenzylidene)(oxido)amino]-*N*-(6-methylpyridin-2-yl)benzamide as a pale yellow powder (0.17 g, 0.47 mmol, 57%). M.p. 200 – 201.5 °C;  $^1\text{H}$   $\{^{19}\text{F}\}$  NMR (400.1 MHz,  $\text{CDCl}_3$ )  $\delta$  = 8.46 – 8.54 (1H, br, *NH*), 8.42 (2H, m,  $2 \times \text{ArH}$ ), 8.08 (1H, d,  $^3J_{\text{HH}} = 8$  Hz, *ArH*), 7.90 (1H, s, *NCH*), 7.87 (2H, d,  $^3J_{\text{HH}} = 8$  Hz  $2 \times \text{ArH}$ ), 7.69 (1H, t,  $^3J_{\text{HH}} = 8$  Hz, *ArH*), 7.12 (2H, d,  $^3J_{\text{HH}} = 8$  Hz,  $2 \times \text{ArH}$ ), 6.88 (1H, d,  $^3J_{\text{HH}} = 8$  Hz, *ArH*) and 2.41 (3H, s,  $\text{CH}_3$ );  $^{13}\text{C}$  NMR (75.5 MHz,  $\text{d}_6\text{-DMSO}$ )  $\delta$  = 164.8 (*NHCO*), 162.7 (d,  $^1J_{\text{CF}} = 250$  Hz, *ArC-F*), 156.6 (*ArC*), 151.4 (*ArC*), 150.2 (*ArC*), 138.4 (*ArCH*), 135.2 (*ArC*), 133.2 (*CH*), 131.6 (d,  $^3J_{\text{CF}} = 8.5$  Hz, *ArCH*), 129.0 (*ArCH*), 127.7 (d,  $^4J_{\text{CF}} = 3.2$  Hz, *ArC*), 121.3 (*ArCH*), 119.2 (*ArCH*), 115.6 (d,  $^2J_{\text{CF}} = 22$  Hz, *ArCH*), 111.7 (*ArCH*) and 23.6 ( $\text{CH}_3$ );  $^{19}\text{F}$   $\{^1\text{H}\}$  NMR (376.5 MHz,  $\text{CDCl}_3$ )  $\delta$  = -105.7; MS ( $\text{Cl}^+$ )  $m/z$  350.13 ( $[\text{M}+\text{H}]^+$ , 20), 334.13 (100) and 226.07 (30); HRMS ( $\text{Cl}^+$ )  $m/z$  calculated for  $\text{C}_{20}\text{H}_{17}\text{N}_3\text{O}_2\text{F}$   $[\text{MH}^+]$  350.1305, found 350.1302.

**(Z)-4-(6-methylpyridin-2-ylcarbamoyl)-*N*-(4-(6-methylpyridin-2-ylcarbamoyl)benzylidene)aniline oxide (120)**

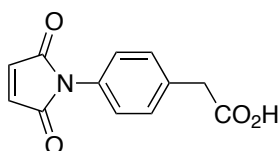


4-(Hydroxyamino)-*N*-(6-methylpyridin-2-yl)benzamide **124** (0.3 g, 1.2 mmol, 1 eq) and *N*-(6-methyl-2-pyridyl)-4-formyl benzamide **49** (0.3 g, 1.2 mmol, 1 eq) were dissolved in ethanol (10 mL) and left to stand in darkness for 16 h. The precipitate which formed was isolated by filtration, washed with cold ethanol and dried to yield the crude product which was recrystallised from ethanol to furnish (*Z*)-4-(6-



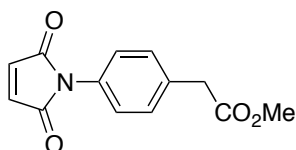
methylpyridin-2-ylcarbamoyl)-*N*-(4-(6-methylpyridin-2-ylcarbamoyl) benzylidene) aniline oxide as a pale yellow powder (0.3 g, 0.65 mmol, 54%); M.p. Decomposed >250 °C; <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 8.58 – 8.69 (2H, br, 2×NH), 8.48 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8 Hz, 2×ArH), 8.12 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8 Hz, 2×ArH), 7.96 (4H, m, 4×ArH), 7.88 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8 Hz, 2×ArH), 7.59 (2H, t, <sup>3</sup>J<sub>HH</sub> = 8 Hz, 2×ArH), 6.88 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8 Hz, 2×ArH) and 2.42 (6H, s, 2×CH<sub>3</sub>); <sup>13</sup>C NMR (75.5 MHz, d<sub>6</sub>-DMSO) δ = 165.2 (NHCO), 164.8 (NHCO), 156.7 (ArC), 156.6 (ArC), 151.5 (ArC), 151.4 (ArC), 150.3 (ArC), 138.5 (ArCH), 138.4 (ArCH), 135.5 (CH), 133.7 (ArCH), 133.7 (ArC), 129.1 (ArCH), 128.7 (ArCH), 128.2 (ArCH), 121.5 (ArCH), 119.3 (ArCH), 119.2 (ArCH), 111.8 (ArCH), 111.7 (ArCH) and 23.6 (CH<sub>3</sub>); MS (ES<sup>+</sup>) *m/z* 488.26 ([M+Na]<sup>+</sup>, 100) and 466.29 ([M+H]<sup>+</sup>, 20); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>27</sub>H<sub>24</sub>N<sub>5</sub>O<sub>3</sub> [M+H]<sup>+</sup> 466.1879, found 466.1884.

**2-(4-(2,5-Dioxo-2,5-dihydro-1H-pyrrol-1-yl)phenyl)ethanoic acid<sup>139</sup> (34)**



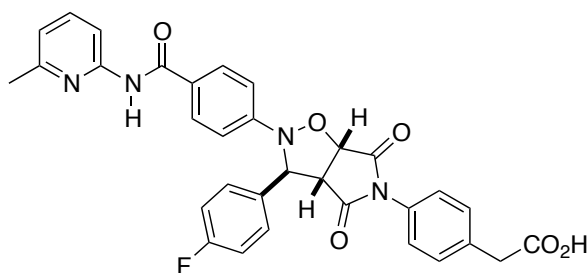
Maleic anhydride (6.05 g, 40 mmol, 1 eq) and 4-aminophenyl acetic acid (3.92 g, 40 mmol, 1 eq) were dissolved in acetic acid (200 mL) and stirred under N<sub>2</sub> at RT for 2 h before refluxing for 2 h. The solvent was then removed under vacuum and the residue dissolved in 5% acetic acid in CH<sub>2</sub>Cl<sub>2</sub>. The solution was passed through a plug of silica eluted with 5% acetic acid in CH<sub>2</sub>Cl<sub>2</sub>. The fractions containing product as identified by tlc were combined and the solvent removed under vacuum. The residue was then purified by recrystallisation from chloroform to yield 2-(4-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)phenyl)ethanoic acid (3.28 g, 14 mmol, 36%) as a yellow powder; M.p. 160–162 °C (lit<sup>139</sup> 151.7–153.1 °C); <sup>1</sup>H NMR (400.1 MHz, d<sub>6</sub>-DMSO) δ = 12.70 (1H, br, CO<sub>2</sub>H), 7.36 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 2×ArH), 7.26 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 2×ArH), 7.18 (2H, s, 2×CH) and 3.63 (2H, s, CH<sub>2</sub>); <sup>13</sup>C NMR (100.6 MHz, D<sub>6</sub>-DMSO) δ = 172.5 (CO<sub>2</sub>H) 169.3 (CO), 134.7 (ArC–N), 134.6 (CH), 130.0 (ArC), 129.9 (ArCH), 126.6 (ArCH) and 40.1 (CH<sub>2</sub>); MS (EI<sup>+</sup>) *m/z* 254 ([M+Na]<sup>+</sup>, 100); HRMS (EI<sup>+</sup>) *m/z* calculated for C<sub>12</sub>H<sub>9</sub>NO<sub>4</sub> [M+Na]<sup>+</sup> 254.0429, found 254.0434.

**Methyl 2-(4-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)phenyl)ethanoate (105)**



2-(4-(2,5-Dioxo-2,5-dihydro-1H-pyrrol-1-yl)phenyl)ethanoic acid **34** (1 g, 4.3 mmol, 1 eq) was dissolved in THF (20 mL) and stirred at room temperature under N<sub>2</sub>. Caesium carbonate (1.04 g, 3.2 mmol, 0.75 eq) was then added followed by dropwise addition of a solution of methyl iodide (0.66 g, 4.7 mmol, 1.1 eq) in THF (10 mL). The solution was stirred for 24 h at RT after which the solvent was removed under vacuum and the residue washed with water (100 mL), the product extracted into CH<sub>2</sub>Cl<sub>2</sub> (3 × 100 mL), dried (MgSO<sub>4</sub>), filtered and the solvent removed under vacuum to yield the crude product which was purified by flash chromatography (3:1 cyclohexane: ethyl acetate) to yield the product as a colourless solid (0.3 g, 1.3 mmol, 30%); M.p. 84.0–86.5 °C (lit<sup>139</sup> 86.7–87.6 °C); <sup>1</sup>H NMR (300.1 MHz, CDCl<sub>3</sub>) δ = 7.38 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8 Hz, 2×ArH), 7.31 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8 Hz, 2×ArH), 6.83 (2H, s, 2×CH), 3.65 (3H, s, OCH<sub>3</sub>) and 3.57 (2H, s, CH<sub>2</sub>); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ = 171.6 (CO<sub>2</sub>H), 169.6 (CO), 134.3 (ArCH), 133.9 (ArC), 130.3 (ArC), 130.2 (ArCH), 126.2 (CH), 52.3 (OCH<sub>3</sub>) and 40.9 (CH<sub>2</sub>); MS (ES<sup>+</sup>) *m/z* 268 ([M+Na]<sup>+</sup>, 100); HRMS (CI<sup>+</sup>) *m/z* calculated for C<sub>13</sub>H<sub>10</sub>NO<sub>4</sub> [M+H]<sup>+</sup> 245.0688 found 245.0694.

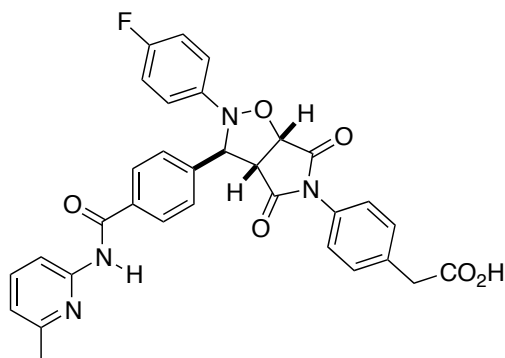
**2-(4-((3R,3aS,6aR)-3-(4-fluorophenyl)-2-(4-((6-methylpyridin-2-yl)carbamoyl)phenyl)-4,6-dioxotetrahydro-2H-pyrrolo[3,4-d]isoxazol-5(3H)-yl)phenyl)acetic acid (*trans*-122)**



Nitron **121** (0.0055 g, 0.016 mmol, 1 eq) and maleimide **34** (0.0037 g, 0.016 mmol, 1 eq) were dissolved in CDCl<sub>3</sub> (0.8 mL) and incubated at 0 °C for several days before removing the solvent under vacuum to yield the crude product (0.0094 g, 0.016 mmol, quantitative), a colourless powder, as a mixture of diastereomers (7:1, *trans*:*cis*); M.p. 215.9–220 °C; <sup>1</sup>H NMR (499.9 MHz, d<sub>6</sub>-DMSO): 12.62–12.16 (1H, br,

CO<sub>2</sub>H), 10.58 (1H, s, NH), 8.00–7.95 (3H, m, 3×ArH), 7.71 (1H, t, <sup>3</sup>J<sub>HH</sub> = 7.1 Hz, ArH), 7.64–7.59 (2H, m, 2×ArH), 7.29–7.20 (6H, m, 6×ArH), 7.01 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.1 Hz, ArH), 6.66 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.3 Hz, 2×ArH), 6.01 (1H, s, CH), 5.51 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.2 Hz, CH), 4.14 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.2 Hz, CH), 3.57 (2H, s, CH<sub>2</sub>) and 2.44 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100.1 MHz, d<sub>6</sub>-DMSO) δ = 174.1 (CO), 173.1 (CO), 172.3 (CO<sub>2</sub>H), 164.9 (NHCO), 161.7 (d, <sup>1</sup>J<sub>CF</sub> = 245 Hz, ArCF), 156.4 (ArC), 151.5 (ArC), 151.2 (ArC), 138.4 (ArCH), 135.8 (ArC), 134.9 (d, <sup>4</sup>J<sub>CF</sub> = 3.0 Hz, ArC–F), 130.0 (ArCH), 129.9 (ArC), 129.4 (d, <sup>3</sup>J<sub>CF</sub> = 8.3 Hz, ArCH), 129.3 (ArCH), 127.4 (ArC), 126.1 (ArCH), 118.5 (ArCH), 117.0 (ArCH), 115.4 (d, <sup>2</sup>J<sub>CF</sub> = 21.4 Hz, ArCH), 113.7 (ArCH), 111.7 (ArCH), 77.7 (CH), 67.4 (CH), 56.3 (CH), 40.1 (CH<sub>2</sub>) and 22.5 (CH<sub>3</sub>); <sup>19</sup>F NMR (376.5 MHz, d<sub>6</sub>-DMSO) δ = –114.8; MS (ES<sup>+</sup>) *m/z* 602.64 ([M+Na]<sup>+</sup>, 20) and 580.95 ([M+H]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>32</sub>H<sub>26</sub>N<sub>4</sub>O<sub>6</sub> [M+H]<sup>+</sup> 581.1836, found 581.1827.

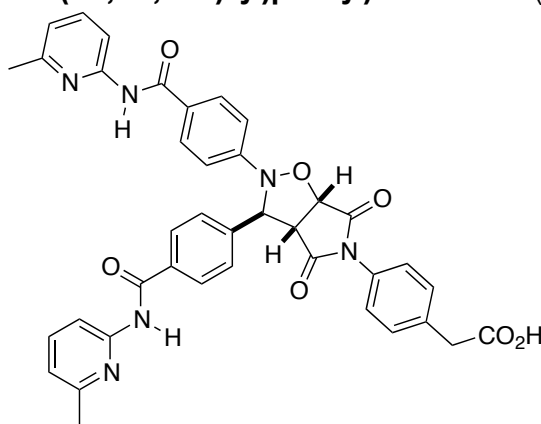
**2-(4-(2-(4-fluorophenyl)-3-(4-(6-methylpyridin-2-ylcarbonyl)phenyl)-4,6-dioxodihydro-2H-pyrrolo[3,4-d]isoxazol-5(3H,6H,6aH)-yl)phenyl)acetic acid<sup>155</sup> (trans-55)**



Nitrone **52** (0.017 g, 0.048 mmol, 1 eq) and maleimide **34** (0.011 g, 0.048 mmol, 1 eq) were dissolved in CHCl<sub>3</sub> (2.4 mL) and incubated at –10 °C for several days before removing the solvent under vacuum to yield the product as a colourless powder (0.028 g, 0.048 mmol, quantitative); M.p. 203.0–205.0 °C (lit<sup>155</sup> 203.0–205.9 °C); <sup>1</sup>H NMR (400 MHz, d<sub>6</sub>-DMSO) δ = 12.64 – 12.23 (1H, br, OH), 10.73 (1H, s, NH), 8.07 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.2 Hz, 2×ArH), 8.01 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.1 Hz, ArH), 7.72 (1H, t, <sup>3</sup>J<sub>HH</sub> = 7.4 Hz, ArH), 7.65 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.2 Hz, 2×ArH), 7.30 (2H, d, <sup>3</sup>J<sub>HH</sub> = 7.8 Hz, 2×ArH), 7.24 (2H, d, <sup>3</sup>J<sub>HH</sub> = 9.1 Hz, 2×ArH), 7.11 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.7 Hz, 2×ArH), 7.02 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.4 Hz, ArH), 6.69 (2H, d, <sup>3</sup>J<sub>HH</sub> = 7.8 Hz, 2×ArH), 5.95 (1H, s, CH), 5.45 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.2 Hz, CH), 4.17 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.0 Hz, CH), 3.61 (2H, s, CH<sub>2</sub>) and 2.45 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100 MHz, d<sub>6</sub>-DMSO) δ = 174.3 (CO), 173.2 (CO), 172.4 (CO<sub>2</sub>H), 165.4

(NHCO), 157.9 (d,  $^1J_{CF} = 239$  Hz, ArCF), 156.6 (ArC), 151.4 (ArC), 145.1 (d,  $^4J_{CF} = 1.9$  Hz, ArC–F), 142.6 (ArC), 138.3 (ArCH), 135.8 (ArC), 133.6 (ArC), 130.0 (ArCH), 129.9 (ArC), 128.3 (ArCH), 127.1 (ArCH), 126.0 (ArCH), 119.1 (ArCH), 116.3 (d,  $^3J_{CF} = 7.9$  Hz, ArCH), 115.6 (d,  $^2J_{CF} = 22.6$  Hz, ArCH), 111.7 (ArCH), 77.7 (CH), 68.5 (CH), 56.4 (CH), 40.1 (CH<sub>2</sub>) and 23.6 (CH<sub>3</sub>);  $^{19}\text{F}$  NMR (376.5 MHz, d<sub>6</sub>-DMSO)  $\delta = -121.6$ ; MS (ES<sup>+</sup>)  $m/z$  581.19 ([M+H]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>)  $m/z$  calculated for C<sub>32</sub>H<sub>26</sub>N<sub>4</sub>O<sub>6</sub>F [M+H]<sup>+</sup> 581.1836, found 581.1862.

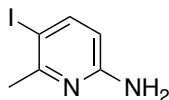
**2-(4-(2,3-bis(4-(6-methylpyridin-2-ylcarbamoyl)phenyl)-4,6-dioxodihydro-2H-pyrrolo[3,4-d]isoxazol-5(3H,6H,6aH)-yl)phenyl)acetic acid (*trans*-126)**



Maleimide **34** (0.17 g, 0.7 mmol, 1 eq) and nitrone **120** (0.32 g, 0.7 mmol, 1 eq) were dissolved in CHCl<sub>3</sub> (20 mL) at 0 °C and stirred under N<sub>2</sub> for 4 days. Upon completion of the reaction, the solvent was removed under vacuum to yield the crude product which was purified by recrystallisation in ethanol to yield 2-(4-(2,3-bis(4-(6-methylpyridin-2-ylcarbamoyl)phenyl)-4,6-dioxodihydro-2H-pyrrolo[3,4-d]isoxazol-5(3H,6H,6aH)-yl)phenyl)acetic acid as a colourless powder (0.29 g, 0.4 mmol 60%); M.p. 274 – 276 °C;  $^1\text{H}$  NMR (400.1 MHz, d<sub>6</sub>-DMSO)  $\delta = 10.72 - 10.8$  (1H, br, NH), 10.5 – 10.6 (1H, br, NH), 8.10 (2H, d,  $^3J_{\text{HH}} = 8$  Hz, 2×ArH), 8.06 (4H, m, 4×ArH), 7.69 (4H, m, 4×ArH), 7.34 (2H, d,  $^3J_{\text{HH}} = 8$  Hz, 2×ArH), 7.26 (2H, d,  $^3J_{\text{HH}} = 8$  Hz, 2×ArH), 7.01 (2H, t,  $^3J_{\text{HH}} = 8$  Hz, 2×ArH), 6.65 (2H, d,  $^3J_{\text{HH}} = 8$  Hz, 2×ArH), 6.12 (1H, s, CH), 5.53 (1H, s, CH), 4.19 (1H, d,  $^3J_{\text{HH}} = 9$  Hz, CH), 3.35 (4H, s, 2×CH<sub>2</sub>) and 2.40 (6H, s, 2×CH<sub>3</sub>);  $^{13}\text{C}$  NMR (100.1 MHz, d<sub>6</sub>-DMSO)  $\delta = 174.2$  (CO), 173.0 (CO), 172.3 (CO<sub>2</sub>H), 165.4 (NHCO), 164.9 (NHCO), 156.6 (ArC), 156.5 (ArC), 151.6 (ArC), 151.5 (ArC), 151.4 (ArC), 142.4 (ArC), 138.3 (ArCH), 138.2 (ArCH), 135.8 (ArC), 133.6 (ArC), 130.0 (ArCH), 129.9 (ArC), 129.3 (ArCH), 128.3 (ArCH), 127.5 (ArC), 127.0 (ArCH), 126.1 (ArCH), 119.1 (ArCH), 118.8 (ArCH), 113.6 (ArCH), 111.7 (ArCH), 77.8 (CH), 67.7

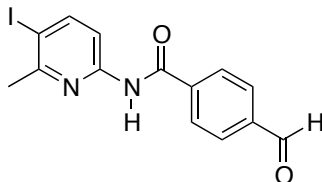
(CH), 56.3 (CH), 38.6 (CH<sub>2</sub>) and 23.6 (CH<sub>3</sub>); MS (ES<sup>+</sup>) *m/z* 697.27 ([M+H]<sup>+</sup>, 80), 233.61 (100) and 120.05 (20); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>39</sub>H<sub>33</sub>N<sub>6</sub>O<sub>7</sub> [M+H]<sup>+</sup> 697.2415, found 697.2411.

#### 5-Iodo-6-methylpyridin-2-amine<sup>220</sup> (**128**)



2-Amino-6-picoline (5.4 g, 50 mmol, 1 eq) was dissolved in acetic acid (50 mL) and concentrated sulfuric acid (7.5 mL) and stirred at 60 °C for 0.5 h before addition of periodic acid (1.63 g, 7.1 mmol, 0.14 eq) and iodine (5.4 g, 21.3 mmol, 0.43 eq). The reaction solution was then stirred at 60 °C for 20 h and upon completion the solvent was removed under vacuum and the residue dissolved in water (100 mL). The pH was adjusted to 14 with 5M KOH (aq) to yield a precipitate which was extracted into CH<sub>2</sub>Cl<sub>2</sub> (2 × 100 mL). The combined organic fractions were dried (MgSO<sub>4</sub>), filtered and the solvent removed under vacuum to yield the crude product which was recrystallised in CH<sub>2</sub>Cl<sub>2</sub> to yield the purified product (8.6 g, 38 mmol, 76%) as a colourless powder; M.p. 98.8–100.2 °C (lit<sup>220</sup> 100–102 °C); <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 7.66 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.7 Hz, ArH), 6.09 (1H, dd, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz and <sup>4</sup>J<sub>HH</sub> = 0.6 Hz, ArH) 4.75 – 4.35 (1H, br, NH<sub>2</sub>) and 2.52 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz CDCl<sub>3</sub>) δ = 158.1 (ArC), 157.6 (ArC), 147.4 (ArCH), 108.0 (ArCH), 80.6 (ArC) and 28.5 (CH<sub>3</sub>); MS (ES<sup>+</sup>) *m/z* 234.9 ([M+H]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>6</sub>H<sub>8</sub>N<sub>2</sub>I [M+H]<sup>+</sup> 234.9732 found 234.9734.

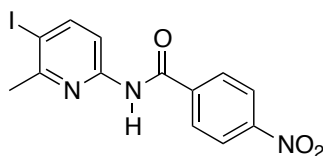
#### 4-Formyl-N-(5-iodo-6-methylpyridin-2-yl)benzamide (**129**)



5-Iodo-6-methylpyridin-2-amine **128** (6.56 g, 28 mmol, 1.05 eq) and triethylamine (5.4 g, 7.5 mL, 53 mmol, 2 eq) were dissolved in CH<sub>2</sub>Cl<sub>2</sub> (20 mL) and stirred under N<sub>2</sub> at 0 °C. A solution of 4-formylbenzoyl chloride (4.5 g, 26.7 mmol, 1 eq) in CH<sub>2</sub>Cl<sub>2</sub> (30 mL) was then added in a dropwise manner and the reaction solution slowly allowed to rise to RT and stir for 48 h. The reaction was quenched by addition of saturated

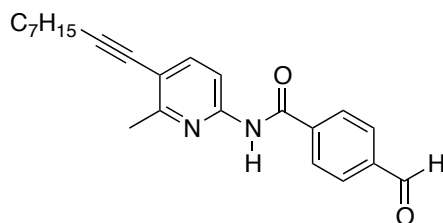
sodium bicarbonate (aq) (100 mL) and the product extracted into CH<sub>2</sub>Cl<sub>2</sub> (3 × 100 mL). The combined organic fractions were washed with 1M HCl (aq), dried (MgSO<sub>4</sub>), filtered and the solvent removed to give the crude product which was purified by recrystallisation in CH<sub>2</sub>Cl<sub>2</sub>. This yielded the product (9.52 g, 26 mmol, 98%) as a colourless powder; M.p. 180–181.5 °C; <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 10.12 (1H, s, CHO), 8.50 (1H, br, NH), 8.11 – 8.05 (3H, m, 3×ArH), 8.02 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 2×ArH), 7.95 (1H, dd, <sup>3</sup>J<sub>HH</sub> = 8.6 Hz and <sup>4</sup>J<sub>HH</sub> = 0.5 Hz, ArH) and 2.65 (3 H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz d<sub>6</sub>-DMSO) δ = 192.9 (CHO), 165.4 (NHCO), 157.9 (ArC), 151.1 (ArC), 147.9 (ArCH), 138.9 (ArC), 138.1 (ArC), 129.2 (ArCH), 128.8 (ArCH), 113.8 (ArCH), 89.4 (ArC-I) and 27.86 (CH<sub>3</sub>); MS (ES<sup>-</sup>) *m/z* 365 ([M-H]<sup>-</sup>, 100); HRMS (ES<sup>-</sup>) *m/z* calculated for C<sub>14</sub>H<sub>10</sub>N<sub>2</sub>O<sub>2</sub>I [M-H]<sup>-</sup> 364.9787 found 364.9790.

#### ***N*-(5-iodo-6-methylpyridin-2-yl)-4-nitrobenzamide (132)**



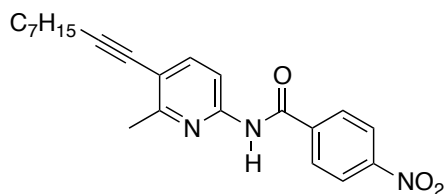
5-Iodo-6-methylpyridin-2-amine **128** (2 g, 8.5 mmol, 1.1 eq) and triethylamine (5 mL) were suspended in Toluene (50 mL) and stirred under N<sub>2</sub> at -10 °C. A solution of 4-nitrobenzoyl chloride (1.4 g, 7.8 mmol, 1 eq) in Toluene (100 mL) was then added over 3 h in a drop wise manner after which the reaction solution was slowly allowed to rise to RT and stir for 16 h. The solvent was then removed under vacuum before addition of saturated sodium bicarbonate (aq) (100 mL). The product was extracted into CH<sub>2</sub>Cl<sub>2</sub> (3 × 100 mL), dried (MgSO<sub>4</sub>), filtered and the solvent removed under vacuum to give the crude product which was purified by recrystallisation from CH<sub>2</sub>Cl<sub>2</sub>. This yielded the product (1.85 g, 4.8 mmol, 62%) as a colourless powder; M.p. 203.5–205.2 °C; <sup>1</sup>H NMR (400.1 MHz, DMSO-d<sub>6</sub>) δ = 11.05 (1H, br, NH), 8.29 (2H, dt, <sup>3</sup>J<sub>HH</sub> = 9 Hz and <sup>4</sup>J<sub>HH</sub> = 2.3 Hz, 2×ArH), 8.17 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.6 Hz, ArH), 8.11 (2H, dt, <sup>3</sup>J<sub>HH</sub> = 9 Hz and <sup>4</sup>J<sub>HH</sub> = 2.3 Hz, 2×ArH), 7.74 (1H, dd, <sup>3</sup>J<sub>HH</sub> = 8.6 Hz and <sup>4</sup>J<sub>HH</sub> = 0.5 Hz, ArH) and 2.57 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz, d<sub>6</sub>-DMSO) δ = 165.5 (CONH), 158.8 (ArC), 151.3 (ArC), 149.9 (ArC), 148.7 (ArCH), 140.1 (ArC), 130.0 (ArCH), 124.1 (ArCH), 114.5 (ArCH), 90.3 (ArC) and 28.5 (CH<sub>3</sub>); MS (CI<sup>+</sup>) *m/z* 383.98 ([M+H]<sup>+</sup>, 100), 354.01 (55) and 237.12 (55); HRMS (CI<sup>+</sup>) *m/z* calculated for C<sub>13</sub>H<sub>11</sub>N<sub>3</sub>O<sub>3</sub>I [M+H]<sup>+</sup> 383.9845 found 383.9843.

#### 4-Formyl-N-(6-methyl-5-(non-1-ynyl)pyridin-2-yl)benzamide (130)



4-Formyl-N-(5-iodo-6-methylpyridin-2-yl)benzamide **129** (2 g, 5.5 mmol, 1 eq), nonyne (2.03 g, 16.4 mmol, 3 eq), copper iodide (0.073 g, 0.38 mmol, 0.07 eq) and triphenyl phosphine (0.1 g, 0.76 mmol, 0.14 eq) were dissolved in triethylamine 100 mL and degassed vigorously with nitrogen gas. Pd(P(Ph<sub>3</sub>)<sub>2</sub>)Cl<sub>2</sub> (0.269 g, 0.38 mmol, 0.07 eq) was then added and the reaction heated to 50 °C, stirred under N<sub>2</sub> and the reaction progress monitored by TLC analysis. Upon completion the solvent was removed under vacuum and the residue washed with saturated NH<sub>4</sub>Cl (aq) solution. The product was extracted into CH<sub>2</sub>Cl<sub>2</sub> (3 × 100 mL), dried (MgSO<sub>4</sub>), filtered and the solvent removed to give the crude product which was purified by flash column chromatography using 5:1 cyclohexane: ethyl acetate as elutant yielding the purified product (1.8 g, 4.9 mmol, 90%) as a colourless powder; M.p. 74.5–76.5 °C; <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 10.10 (1H, s, CHO), 8.67 (1H, br, NH), 8.14 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, ArH), 8.06 (2H, dt, <sup>3</sup>J<sub>HH</sub> = 8.2 Hz and <sup>4</sup>J<sub>HH</sub> = 1.9 Hz, 2×ArH) 7.99 (2H, dt, <sup>3</sup>J<sub>HH</sub> = 8.2 Hz and <sup>4</sup>J<sub>HH</sub> = 1.9 Hz, 2×ArH) 7.70 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, ArH), 2.54 (3 H, s, CH<sub>3</sub>), 2.45 (2H, t, <sup>3</sup>J<sub>HH</sub> = 7 Hz, CH<sub>2</sub>), 1.66 – 1.55 (2H, m, CH<sub>2</sub>), 1.50 – 1.40 (2H, m, CH<sub>2</sub>), 1.38 – 1.24 (6H, m, 3×CH<sub>2</sub>) and 0.89 (3H, t, <sup>3</sup>J<sub>HH</sub> = 7 Hz, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ = 191.5 (CHO), 164.5 (CONH), 159.0 (ArC), 148.8 (ArC), 141.8 (ArCH), 139.4 (ArC), 138.7 (ArC), 130.1 (ArCH), 128.1 (ArCH), 116.6 (ArC), 111.5 (ArCH), 96.5 (Ph-C≡C), 77.4 (Ph-C≡C), 31.9 (CH<sub>2</sub>), 29.0 (CH<sub>2</sub>), 28.9 (CH<sub>2</sub>), 28.8 (CH<sub>2</sub>), 23.1 (CH<sub>3</sub>), 22.7 (CH<sub>2</sub>), 19.7 (CH<sub>2</sub>) and 14.2 (CH<sub>3</sub>); MS (ES<sup>-</sup>) *m/z* 361.23 ([M-H]<sup>-</sup>, 100); HRMS (ES<sup>-</sup>) *m/z* calculated for C<sub>23</sub>H<sub>25</sub>N<sub>2</sub>O<sub>2</sub> [M-H]<sup>-</sup> 361.1916 found 361.1912.

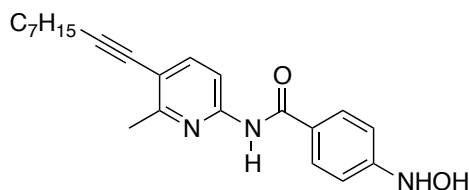
***N*-(6-methyl-5-(non-1-ynyl)pyridin-2-yl)-4-nitrobenzamide (133)**



*N*-(5-iodo-6-methylpyridin-2-yl)-4-nitrobenzamide **132** (4.66 g, 1.2 mmol, 1 eq), nonyne (4.6 g, 3.7 mmol, 3 eq), copper iodide (0.16 g, 0.085 mmol, 0.07 eq) and triphenyl phosphine (0.44 g, 0.17 mmol, 0.14 eq) were dissolved in triethylamine (150 mL) and degassed vigorously with nitrogen gas. Pd(P(Ph<sub>3</sub>)<sub>2</sub>Cl<sub>2</sub>) (0.6 g, 0.085 mmol, 0.07 eq) was then added and the reaction heated to 50 °C, stirred under N<sub>2</sub>. The reaction progress was monitored by TLC analysis and upon completion (48 h) the solvent was removed under vacuum and the residue washed with saturated NH<sub>4</sub>Cl (aq) solution. The product was extracted into CH<sub>2</sub>Cl<sub>2</sub> (3 × 100 mL), dried (MgSO<sub>4</sub>), filtered and the solvent removed to give the crude product which was purified by flash column chromatography using 5:1 cyclohexane: ethyl acetate as elutant to yield the purified product (3 g, 0.79 mmol, 66%) as a pale orange powder; M.p. 72.0–73.4 °C; <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 8.79 (1H, br, NH), 8.30 (2H, dt, <sup>3</sup>J<sub>HH</sub> = 9 Hz and <sup>4</sup>J<sub>HH</sub> = 2.3 Hz, 2×ArH), 8.10 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, ArH), 8.05 (2H, dt, <sup>3</sup>J<sub>HH</sub> = 9 Hz and <sup>4</sup>J<sub>HH</sub> = 2.3 Hz, 2×ArH), 7.69 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, ArH), 2.51 (3H, s, CH<sub>3</sub>), 2.44 (2H, t, <sup>3</sup>J<sub>HH</sub> = 7 Hz, CH<sub>2</sub>), 1.66 – 1.55 (2H, m, CH<sub>2</sub>), 1.49 – 1.40 (2H, m, CH<sub>2</sub>), 1.38 – 1.23 (6H, m, 3×CH<sub>2</sub>) and 0.88 (3H, t, <sup>3</sup>J<sub>HH</sub> = 7 Hz, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ = 163.5 (CONH), 159.1 (ArC), 150.1 (ArC), 148.5 (ArC), 141.8. (ArCH), 139.8 (ArC), 128.5 (ArCH), 124.1 (ArCH), 116.7 (ArC), 111.1 (ArCH), 96.7 (Ph–C≡C), 77.3 (Ph–C≡C), 31.9(CH<sub>2</sub>), 29.0 (CH<sub>2</sub>), 28.9 (CH<sub>2</sub>), 28.8 (CH<sub>2</sub>), 23.2 (CH<sub>3</sub>), 22.7 (CH<sub>2</sub>), 19.7 (CH<sub>2</sub>) and 14.2 (CH<sub>3</sub>); MS (ES<sup>+</sup>) *m/z* 402.15 ([M+Na]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>22</sub>H<sub>25</sub>N<sub>3</sub>O<sub>3</sub>Na [M+Na]<sup>+</sup> 402.1794 found 402.1810.

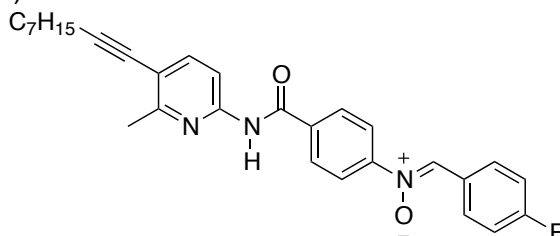


#### 4-(Hydroxyamino)-N-(6-methyl-5-(non-1-ynyl)pyridin-2-yl)benzamide (134)



*N*-(6-methyl-5-(non-1-ynyl)pyridin-2-yl)-4-nitrobenzamide **133** (0.44 g, 1.2 mmol, 1 eq) was dissolved in THF. 5% Rh/C (0.17g) was added to form a suspension which was stirred for 0.5 h before addition of hydrazine hydrate (0.070 mL, 1.4 mmol, 1.2 eq). The reaction was monitored by TLC analysis and upon consumption of starting materials, the reaction was filtered through a thin layer of celite and the solvent removed to give the crude product which was purified by column chromatography with a 3:1 cyclohexane: ethyl acetate solvent system. This yielded the product (0.41 g, 1.1 mmol, 93 %) as a white crystalline solid; M.p. 136.5–138.0 °C; <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 10.1–9.65 (1H, br, NH), 9.33 (1H, s, OH), 8.08 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, ArH), 7.70 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, ArH), 7.49 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.6 Hz, 2×ArH), 7.00 (1H, br, NH), 6.79 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.6 Hz, 2×ArH), 2.59 (3H, s, CH<sub>3</sub>), 2.45 (2H, t, <sup>3</sup>J<sub>HH</sub> = 7 Hz, CH<sub>2</sub>), 1.67 – 1.58 (2H, m, CH<sub>2</sub>), 1.51 – 1.42 (2H, m, CH<sub>2</sub>), 1.38 – 1.23 (6H, m, 3×CH<sub>2</sub>) and 0.90 (3H, t, <sup>3</sup>J<sub>HH</sub> = 7 Hz, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz CDCl<sub>3</sub>) δ = 166.7 (CONH), 158.4 (ArC), 153.0 (ArC), 149.2 (ArC), 142.4 (ArCH), 128.5 (ArCH), 128.1 (ArC), 116.1 (ArC), 114.2 (ArCH), 112.0 (ArCH) 96.6 (Ph-C≡C), 77.1 (Ph-C≡C), 31.9(CH<sub>2</sub>), 29.1 (CH<sub>2</sub>), 28.9 (CH<sub>2</sub>), 28.8 (CH<sub>2</sub>), 22.8 (CH<sub>2</sub>), 22.5 (CH<sub>3</sub>), 19.7 (CH<sub>2</sub>) and 14.2 (CH<sub>3</sub>); MS (ES<sup>+</sup>) *m/z* 388.15 ([M+Na]<sup>+</sup>, 10), 366.17 ([M+H]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>22</sub>H<sub>27</sub>N<sub>3</sub>O<sub>2</sub>Na [M+Na]<sup>+</sup> 388.2001 found 388.1996.

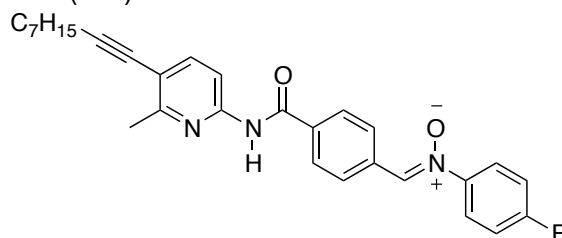
#### 4-[(*Z*)-(4-fluorobenzylidene)(oxido)amino]-N-(6-methyl-5-(non-1-ynyl)pyridin-2-yl)benzamide (135)



4-(hydroxyamino)-*N*-(6-methyl-5-(non-1-ynyl)pyridin-2-yl)benzamide **134** (0.25 g, 0.68 mmol, 1 eq) and 4-fluorobenzaldehyde (0.085 g, 0.68 mmol, 1 eq) were dissolved in ethanol (10 mL) and allowed to react for 48 h in the absence of light.

The product nitron precipitated from the reaction solution was isolated by filtration and washed with a little cold ethanol. This yielded the product (0.25 g, 0.55 mmol, 80%) as a pale yellow powder and was used without further purification; M.p. 207.9–208.7 °C;  $^1\text{H}$   $\{^{19}\text{F}\}$  NMR (400.1 MHz,  $\text{CDCl}_3$ ):  $\delta$  = 8.63 (1H, br, *NH*) 8.47 (2H, d,  $^3J_{\text{HH}}$  = 9 Hz, 2×*ArH*), 8.13 (1H, d,  $^3J_{\text{HH}}$  = 8.5 Hz, *ArH*) 8.04 (2H, d,  $^3J_{\text{HH}}$  = 9 Hz, 2×*ArH*), 7.97 (1H, s, *CH*), 7.91 (2H, dt,  $^3J_{\text{HH}}$  = 8.8 Hz and  $^4J_{\text{HH}}$  = 2.2 Hz, 2×*ArH*), 7.70 (1H, d,  $^3J_{\text{HH}}$  = 8.5 Hz, *ArH*), 7.18 (2H, d,  $^3J_{\text{HH}}$  = 9 Hz, 2×*ArH*), 2.56 (3H, s,  $\text{CH}_3$ ), 2.45 (2H, t,  $^3J_{\text{HH}}$  = 7 Hz,  $\text{CH}_2$ ), 1.67 – 1.57 (2H, m,  $\text{CH}_2$ ), 1.51 – 1.42 (2H, m,  $\text{CH}_2$ ), 1.39 – 1.22 (6H, m, 3× $\text{CH}_2$ ) and 0.89 (3H, t,  $^3J_{\text{HH}}$  = 7 Hz,  $\text{CH}_3$ );  $^{13}\text{C}$  NMR (100.6 MHz  $\text{CDCl}_3$ )  $\delta$  = 164.0 (d,  $^1J_{\text{CF}}$  = 254 Hz, *ArC-F*), 163.9 (*CONH*), 159.0 (*ArC*), 151.2 (*ArC*), 148.7 (*ArC*), 141.7 (*ArCH*), 135.6 (*ArC*), 134.0 (*CH*), 131.7 (d,  $^3J_{\text{CF}}$  = 8.6 Hz, *ArCH*), 128.4 (*ArCH*), 126.8 (d,  $^4J_{\text{CF}}$  = 3.6 Hz, *ArC*), 122.2 (*ArCH*), 116.4 (*ArC*), 116.0 (d,  $^2J_{\text{CF}}$  = 22 Hz, *ArCH*), 110.8 (*ArCH*), 96.3 (*C*) 77.4 (*C*), 31.8 ( $\text{CH}_2$ ), 28.9 ( $\text{CH}_2$ ) 28.9 ( $\text{CH}_2$ ), 28.8 ( $\text{CH}_2$ ), 23.2 ( $\text{CH}_3$ ), 22.7 ( $\text{CH}_2$ ), 19.6 ( $\text{CH}_2$ ) and 14.1 ( $\text{CH}_3$ );  $^{19}\text{F}$   $\{^1\text{H}\}$  NMR (376.5 MHz,  $\text{CDCl}_3$ )  $\delta$  = -106.09 (1F, s, *Ar-F*); MS ( $\text{ES}^+$ )  $m/z$  494.12 ( $[\text{M}+\text{Na}]^+$ , 100); HRMS ( $\text{ES}^+$ )  $m/z$  calculated for  $\text{C}_{29}\text{H}_{30}\text{N}_3\text{O}_2\text{FNa}$   $[\text{M}+\text{Na}]^+$  494.2220 found 494.2234.

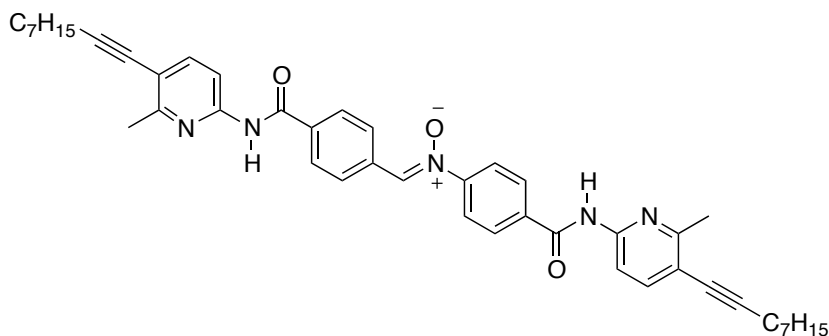
**4- $\{(Z)\text{-}[(4\text{-fluorophenyl})(\text{oxido})\text{imino}]\text{methyl}\}\text{-}N\text{-}(6\text{-methyl-5-(non-1-ynyl)pyridin-2-yl)benzamide (131)$**



*N*-(4-fluorophenyl)hydroxylamine **46** (0.1 g, 0.79 mmol, 1 eq) and 4-formyl-*N*-(6-methyl-5-(non-1-ynyl)pyridin-2-yl)benzamide **130** (0.29 g, 0.79 mmol, 1 eq) were dissolved in ethanol (8 mL) and left to react in the absence of light for 16 h. The precipitate formed was isolated by filtration and purified by column chromatography using 3:2 cyclohexane: ethyl acetate with 1% triethylamine to elute. This yielded the product (0.21 g, 0.45 mmol, 56%) as a colourless powder; M.p. 170.9–171.2 °C;  $^1\text{H}$   $\{^{19}\text{F}\}$  NMR (400.1 MHz,  $\text{CDCl}_3$ )  $\delta$  = 8.69 (1H, br, *NH*), 8.46 (2H, d,  $^3J_{\text{HH}}$  = 8.4 Hz, 2×*ArH*), 8.14 (1H, d,  $^3J_{\text{HH}}$  = 8.5 Hz, *ArH*), 8.00 (2H, d,  $^3J_{\text{HH}}$  = 8.6 Hz, 2×*ArH*), 7.95 (1H, s, *CH*), 7.78 (2H, d,  $^3J_{\text{HH}}$  = 8.9 Hz, 2×*ArH*), 7.68 (1H, d,  $^3J_{\text{HH}}$  = 8.5 Hz, *ArH*), 7.16 (2H, d,  $^3J_{\text{HH}}$  = 8.9 Hz, 2×*ArH*), 2.54 (3H, s,  $\text{CH}_3$ ), 2.44 (2H, t,  $^3J_{\text{HH}}$  = 7 Hz,  $\text{CH}_2$ ), 1.69 –

1.56 (2H, m, CH<sub>2</sub>), 1.51 – 1.40 (2H, m, CH<sub>2</sub>), 1.39 – 1.21 (6H, m, 3×CH<sub>2</sub>) and 0.89 (3H, t, <sup>3</sup>J<sub>HH</sub> = 7 Hz, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ = 164.47 (CONH), 163.3 (d, <sup>1</sup>J<sub>CF</sub> = 251 Hz, ArCF), 158.91 (ArC), 148.90 (ArC), 145.14 (d, <sup>4</sup>J<sub>CF</sub> = 3.4 Hz, ArC), 141.58 (ArCH), 135.59 (ArC), 133.77 (ArC), 133.23 (CH), 129.05 (ArCH), 127.58 (ArCH), 123.74 (d, <sup>3</sup>J<sub>CF</sub> = 8.9 Hz, ArCH) 116.20 (d, <sup>2</sup>J<sub>CF</sub> = 23.4 Hz, ArCH), 116.19 (ArC), 110.84 (ArCH), 96.14 (C), 77.42 (C), 31.78 (CH<sub>2</sub>), 28.93 (CH<sub>2</sub>), 28.84 (CH<sub>2</sub>), 28.76 (CH<sub>2</sub>), 23.15 (CH<sub>3</sub>), 22.67 (CH<sub>2</sub>), 19.62 (CH<sub>2</sub>) and 14.13 (CH<sub>3</sub>); <sup>19</sup>F {<sup>1</sup>H} NMR (376.5 MHz, CDCl<sub>3</sub>) δ = -110.03 (1 F, s, Ar-F); MS (ES<sup>+</sup>) *m/z* 494.13 ([M+Na]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>29</sub>H<sub>30</sub>N<sub>3</sub>O<sub>2</sub>FNa [M+Na]<sup>+</sup> 494.2220 found 494.2227.

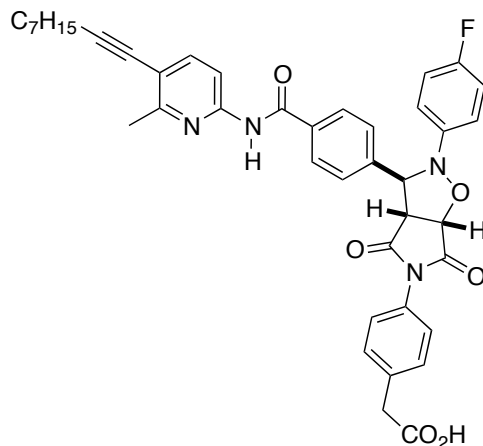
**(Z)-4-(6-methyl-5-(non-1-ynyl)pyridin-2-ylcarbamoyl)-N-(4-(6-methyl-5-(non-1-ynyl)pyridin-2-ylcarbamoyl)benzylidene)aniline oxide (136)**



4-(Hydroxyamino)-*N*-(6-methyl-5-(non-1-ynyl)pyridin-2-yl)benzamide **134** (0.075 g, 0.21 mmol, 1 eq) and 4-formyl-*N*-(6-methyl-5-(non-1-ynyl)pyridin-2-yl)benzamide **130** (0.075 g, 0.21 mmol, 1 eq) were dissolved in ethanol (6 mL) and allowed to react for 48 h in the absence of light. The product nitron precipitated from the reaction solution and was isolated by filtration and washed with a little cold ethanol to yield the product (0.07 g, 0.10 mmol, 50 %) as a pale yellow powder and was used without further purification; M.p. 219–221.5 °C; <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 8.62–8.57 (2H, br, 2×NH), 8.53 (2 H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 2×ArH), 8.19 – 8.12 (2H, m, 2×ArH), 8.10–8.01 (5H, m, 4×ArH + CH), 7.94 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 2×ArH), 7.71 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, ArH), 7.70 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, ArH), 2.58 (6H, s, 2×CH<sub>3</sub>), 2.46 (4H, t, <sup>3</sup>J<sub>HH</sub> = 7 Hz, 2×CH<sub>2</sub>), 1.67–1.57 (4H, m, 2×CH<sub>2</sub>), 1.51–1.41 (4H, m, 2×CH<sub>2</sub>), 1.40–1.25 (12H, m, 6×CH<sub>2</sub>) and 0.90 (6H, t, <sup>3</sup>J<sub>HH</sub> = 7 Hz, 2×CH<sub>3</sub>); <sup>13</sup>C NMR (125.7 MHz, CDCl<sub>3</sub>) δ = 164.7 (CONH), 164.2 (CONH), 159.4 (ArC), 159.4 (ArC), 151.7 (ArC), 149.5 (ArC), 149.3 (ArC), 141.7 (ArCH), 141.7 (ArCH), 136.3 (ArC), 136.1 (ArC), 134.2 (CH), 133.9 (ArC), 129.4 (ArCH), 128.8 (ArCH), 127.9 (ArCH), 122.5 (ArCH), 116.6

(ArC), 116.4 (ArC), 111.0 (ArCH), 110.9 (ArCH), 96.6 (C), 96.5 (C), 77.6 (C), 77.5 (C) 32.1 (CH<sub>2</sub>), 29.2 (CH<sub>2</sub>), 29.2 (CH<sub>2</sub>), 29.1 (CH<sub>2</sub>), 23.3 (CH<sub>3</sub>), 23.00 (CH<sub>2</sub>), 19.84 (CH<sub>2</sub>) and 14.22 (CH<sub>3</sub>); MS (ES<sup>+</sup>) *m/z* 710.24 ([M+H]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>45</sub>H<sub>52</sub>N<sub>5</sub>O<sub>3</sub> [M+H]<sup>+</sup> 710.4070 found 710.4075.

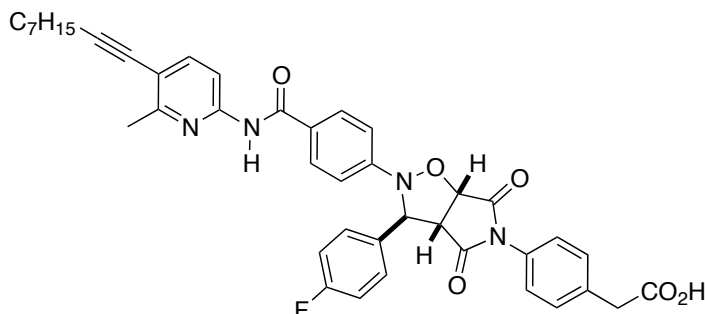
**2-(4-(2-(4-Fluorophenyl)-3-(4-(6-methyl-5-(non-1-ynyl)pyridin-2-ylcarbamoyl)phenyl)-4,6-dioxodihydro-2H-pyrrolo[3,4-d]isoxazol-5(3H,6H,6aH)-yl)phenyl)acetic acid (*trans*-141)**



Nitrone **131** (0.0075 g, 0.016 mmol, 1 eq) and maleimide **34** (0.0037 g, 0.016 mmol, 1 eq) were dissolved in CDCl<sub>3</sub> (0.8 mL) incubated at 0 °C and allowed to react for 48 h. The chloroform was removed under vacuum to give the product cycloadduct (0.011 g, 0.016 mmol, quantitative) as a colourless powder and was used without further purification; M.p. 200–204 °C; <sup>1</sup>H {<sup>19</sup>F} NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 11.19–10.96 (1H, br, NH), 8.23 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.7 Hz, ArH), 7.89 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.4 Hz, 2×ArH), 7.76 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.7 Hz, ArH), 7.33 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.4 Hz, 2×ArH), 7.29 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 2×ArH), 7.14–7.08 (2H, m, 2×ArH), 6.98–6.91 (2H, m, 2×ArH), 6.81 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.4 Hz, 2×ArH), 5.63 (1H, s, CH), 4.96 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.5 Hz, CH), 3.88 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.5 Hz, CH), 3.71 (2H, s, CH<sub>2</sub>) 2.55 (3H, s, CH<sub>3</sub>), 2.47 (2H, t, <sup>3</sup>J<sub>HH</sub> = 7 Hz, CH<sub>2</sub>), 1.69–1.58 (2H, m, CH<sub>2</sub>), 1.52–1.42 (2H, m, CH<sub>2</sub>), 1.41–1.27 (6H, m, 3×CH<sub>2</sub>) and 0.91 (3H, t, <sup>3</sup>J<sub>HH</sub> = 7 Hz CH<sub>3</sub>); <sup>13</sup>C NMR (125.7 MHz CDCl<sub>3</sub>) δ = 176.6 (CO<sub>2</sub>H), 174.8 (CO), 173.4 (CO), 169.6 (CONH), 159.0 (d, <sup>1</sup>J<sub>CF</sub> = 242.5 Hz, ArC–F), 158.0 (ArC), 150.0 (ArC), 144.6 (d, <sup>4</sup>J<sub>CF</sub> = 2.1 Hz, ArC), 143.2 (ArCH), 142.1 (ArC), 135.5 (ArC), 133.8 (ArC), 130.8 (ArCH), 129.9 (ArC), 129.0 (ArCH), 126.8 (ArCH), 126.0 (ArCH), 116.7 (ArC), 116.5 (d, <sup>3</sup>J<sub>CF</sub> = 7.7 Hz, ArCH), 116.1 (d, <sup>2</sup>J<sub>CF</sub> = 22.6 Hz, ArCH), 113.0 (ArCH), 97.3 (C), 77.2 (CH), 76.7 (C), 69.5 (CH), 57.2 (CH), 40.1 (CH<sub>2</sub>), 31.9 (CH<sub>2</sub>), 29.1 (CH<sub>2</sub>), 29.0 (CH<sub>2</sub>), 28.8 (CH<sub>2</sub>), 22.8 (CH<sub>2</sub>), 20.9 (CH<sub>3</sub>), 19.7 (CH<sub>2</sub>) and

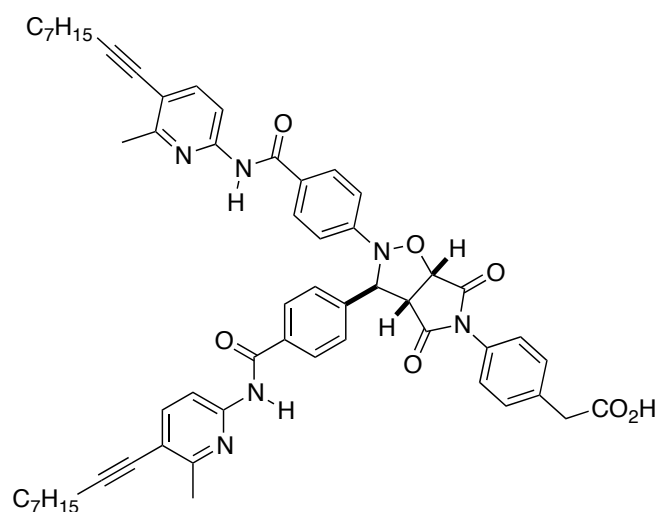
14.3 (CH<sub>3</sub>); HRMS (MALDI) *m/z* calculated for C<sub>41</sub>H<sub>40</sub>FN<sub>4</sub>O<sub>6</sub> [M+H]<sup>+</sup> 703.2932 found 703.2950.

**2-(4-(3-(4-Fluorophenyl)-2-(4-(6-methyl-5-(non-1-ynyl)pyridin-2-ylcarbamoyl)phenyl)-4,6-dioxodihydro-2H-pyrrolo[3,4-d]isoxazol-5(3H,6H,6aH)-yl)phenyl)acetic acid (*trans*-139)**



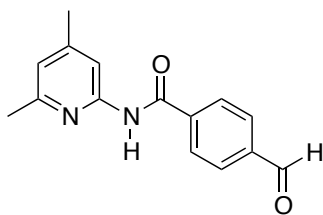
A solution of nitrone **135** (0.075 g, 0.016 mmol, 1 eq) and maleimide **34** (0.037 g, 0.016 mmol 1 eq) in CHCl<sub>3</sub> (8 mL) was cooled to 0 °C and reacted for 96 h. The chloroform was removed under vacuum to give the crude product as a mixture of *trans* and *cis* isomers. The *trans* isomer was enriched by recrystallisation from ethanol to give the product (0.045 g, 0.0064 mmol, 40%) as a colourless powder. M.p. 204–208 °C; <sup>1</sup>H {<sup>19</sup>F} NMR (400.1 MHz, CD<sub>2</sub>Cl<sub>2</sub>) δ = 9.57 (1H, br, NH), 8.04 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.6 Hz and <sup>5</sup>J<sub>HH</sub> = 0.4 Hz, ArH), 7.68 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.9 Hz, 2×ArH), 7.61 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.6 Hz, ArH), 7.37 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.8 Hz, 2×ArH), 7.28 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.6 Hz, 2×ArH), 6.96 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.9 Hz, 2×ArH), 6.94 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.6 Hz, 2×ArH), 6.89 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 2×ArH), 5.83 (1H, s, CH), 5.15 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.5 Hz, CH), 3.95 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.5 Hz, CH), 3.61 (2H, s, CH<sub>2</sub>), 2.48 (3H, s, CH<sub>3</sub>), 2.38 (2H, t, <sup>3</sup>J<sub>HH</sub> = 7 Hz, CH<sub>2</sub>), 1.59–1.49 (2H, m, CH<sub>2</sub>), 1.44–1.33 (2H, m, CH<sub>2</sub>), 1.30–1.19 (6H, m, 3×CH<sub>2</sub>) and 0.81 (3H, t, <sup>3</sup>J<sub>HH</sub> = 7 Hz, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz CD<sub>2</sub>Cl<sub>2</sub>) δ = 175.6 (CO<sub>2</sub>H), 174.1 (CO), 172.9 (CO), 165.6 (CONH), 165.8 (d, <sup>1</sup>J<sub>CF</sub> = 247 Hz, ArC–F), 143.0 (ArCH), 135.9 (ArC), 134.2 (ArC), 134.0 (d, <sup>4</sup>J<sub>CF</sub> = 3.4 Hz, ArC), 130.6 (ArCH), 130.3 (ArC), 129.6 (ArCH), 129.4 (d, <sup>3</sup>J<sub>CF</sub> = 8.1 Hz, ArCH), 128.1 (ArC), 126.2 (ArCH), 116.4 (ArC), 116.0 (d, <sup>2</sup>J<sub>CF</sub> = 21.5 Hz, ArCH), 114.9 (ArCH), 112.7 (CH), 97.2 (C), 77.1 (CH), 76.8 (C), 69.2 (CH), 57.2 (CH) 40.8 (CH<sub>2</sub>), 32.1 (CH<sub>2</sub>), 29.2 (CH<sub>2</sub>), 29.2 (CH<sub>2</sub>), 29.0 (CH<sub>2</sub>), 23.0 (CH<sub>2</sub>), 21.5 (CH<sub>3</sub>), 19.8 (CH<sub>2</sub>) and 14.2 (CH<sub>3</sub>); HRMS (MALDI) *m/z* calculated for C<sub>41</sub>H<sub>40</sub>FN<sub>4</sub>O<sub>6</sub> [M+H]<sup>+</sup> 703.2932 found 703.2931.

**2-(4-(2,3-Bis(4-(6-methyl-5-(non-1-ynyl)pyridin-2-ylcarbamoyl)phenyl)-4,6-dioxodihydro-2H-pyrrolo[3,4-d]isoxazol-5(3H,6H,6aH)-yl)phenyl)acetic acid (trans-143)**



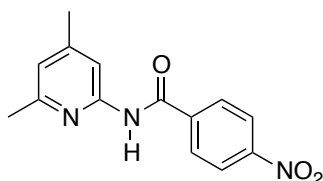
Nitrone **136** (0.0114 g, 0.016 mmol, 1 eq) and maleimide **34** (0.0037 g, 0.016 mmol, 1 eq) were dissolved in  $\text{CDCl}_3$  (0.8 mL), cooled to 0 °C and reacted for 48 h. The chloroform was removed under vacuum to give the product (0.015 g, 0.0159 mmol, 99%) which was used without further purification; M.p. 208.5– 210 °C;  $^1\text{H}$  NMR (499.9 MHz,  $\text{CDCl}_3$ )  $\delta$  = 10.96–10.74 (1H, br, *NH*), 8.54–8.44 (1H, br, *NH*), 8.28 (1H, d,  $^3J_{\text{HH}}$  = 8.6 Hz), 8.10 (1H, d,  $^3J_{\text{HH}}$  = 8.6 Hz, *ArH*), 7.86 (2H, d,  $^3J_{\text{HH}}$  = 8.2 Hz, 2×*ArH*), 7.80 (2H, d,  $^3J_{\text{HH}}$  = 8.6 Hz, 2×*ArH*), 7.75 (1H, d,  $^3J_{\text{HH}}$  = 8.6 Hz, *ArH*), 7.63 (1H, d,  $^3J_{\text{HH}}$  = 8.6 Hz, *ArH*), 7.38 (2H, d,  $^3J_{\text{HH}}$  = 8 Hz, 2×*ArH*), 7.31 (2H, d,  $^3J_{\text{HH}}$  = 7.8 Hz, 2×*ArH*), 7.22 (2H, d,  $^3J_{\text{HH}}$  = 8.7 Hz, 2×*ArH*), 6.8 (2H, d,  $^3J_{\text{HH}}$  = 7.6 Hz, 2×*ArH*), 5.84 (1H, s, *CH*), 5.05 (1H, d,  $^3J_{\text{HH}}$  = 7 Hz, *CH*), 3.95 (1H, d,  $^3J_{\text{HH}}$  = 7 Hz, *CH*), 3.72 (2H, s, *CH*<sub>2</sub>), 2.57 (3H, s, *CH*<sub>3</sub>), 2.53 (3H, s, *CH*<sub>3</sub>), 2.48–2.41 (4H, m, 2×*CH*<sub>2</sub>), 1.66–1.57 (4H, m, 2×*CH*<sub>2</sub>), 1.50–1.41 (4H, m, 2×*CH*<sub>2</sub>), 1.38–1.25 (12H, m, 6×*CH*<sub>3</sub>) and 0.94–0.84 (6H, m, 2×*CH*<sub>3</sub>);  $^{13}\text{C}$  NMR (125.7 MHz  $\text{CD}_2\text{Cl}_2$ )  $\delta$  = 176.9 (*CO*<sub>2</sub>*H*), 174.6 (*CO*), 173.1 (*CO*), 166.1 (*CONH*), 165.0 (*CONH*), 158.7 (*ArC*), 158.4 (*ArC*), 151.9 (*ArC*), 150.3 (*ArC*), 149.8 (*ArC*), 142.9 (*ArCH*), 142.1 (*ArCH*), 142.0 (*ArC*), 136.1 (*ArC*), 134.2 (*ArC*), 130.9 (*ArCH*), 130.2 (*ArC*), 129.4 (*ArCH*), 128.9 (*ArCH*), 128.6 (*ArC*), 127.1 (*ArCH*), 126.2 (*ArCH*), 116.7 (*ArC*), 116.2 (*ArC*), 114.7 (*ArCH*), 112.7 (*ArCH*), 111.6 (*ArCH*), 97.2 (*C*), 96.8 (*C*), 77.7 (*CH*), 77.3 (*C*), 76.9 (*C*), 69.2 (*CH*), 57.2 (*CH*), 40.4 (*CH*<sub>2</sub>), 32.1 (*CH*<sub>2</sub>), 29.3 (*CH*<sub>2</sub>), 29.2 (*CH*<sub>2</sub>), 29.2 (*CH*<sub>2</sub>), 29.1 (*CH*<sub>2</sub>), 29.1 (*CH*<sub>2</sub>), 29.1 (*CH*<sub>2</sub>), 23.0 (*CH*<sub>2</sub>), 22.5 (*CH*<sub>3</sub>), 21.3 (*CH*<sub>3</sub>), 19.8 (*CH*<sub>2</sub>) and 14.2 (*CH*<sub>3</sub>); HRMS (MALDI) *m/z* calculated for  $\text{C}_{57}\text{H}_{61}\text{N}_6\text{O}_7$  [*M*+*H*]<sup>+</sup> 941.4601 found 941.4628.

### ***N*-(4,6-dimethylpyridin-2-yl)-4-formylbenzamide (89)**



2-Amino-4, 6-dimethylpyridine (2.95 g, 24.2 mmol, 1 eq) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (10 mL) and triethylamine (10 mL) and cooled to 0 °C. A solution of 4-formylbenzoyl chloride (4.5 g, 26.7 mmol, 1.1 eq) in CH<sub>2</sub>Cl<sub>2</sub> (90 mL) was added dropwise and after complete addition, the reaction was allowed to rise slowly to RT. The reaction progress was monitored by TLC analysis and upon completion (16 h) was quenched by addition of Na<sub>2</sub>CO<sub>3</sub> (aq) (100 mL). The organic layer was then separated, washed with 1M HCl (aq) (100 mL), dried (MgSO<sub>4</sub>), filtered and the solvent removed to give the crude product which was purified by flash chromatography (3:2, cyclohexane: ethyl acetate) to yield the aldehyde (4.6 g, 18.2 mmol, 75%) as a colourless powder; M.p. 124.2 – 126 °C; <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 10.08 (1H, s, CHO), 8.87–8.70 (1H, br, NH), 8.04 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 2×ArH), 8.01 (1H, s, ArH), 7.96 (2H, d <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 2×ArH), 6.77 (1H, s, ArH), 2.36 (3H, s, CH<sub>3</sub>) and 2.35 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (125.7MHz, CDCl<sub>3</sub>) δ = 191.5 (CHO), 164.6 (NHCO), 156.6, (ArC), 150.6 (ArC), 150.5 (ArC), 139.7 (ArC), 138.6 (ArC), 130.0 (ArCH), 128.0 (ArCH), 121.2 (ArCH), 111.9 (ArCH), 23.8 (Ar-CH<sub>3</sub>) and 21.4 (Ar-CH<sub>3</sub>); MS (ES<sup>-</sup>) *m/z* 253.0 ([M-H]<sup>-</sup>, 100); HRMS (ES<sup>-</sup>) *m/z* calculated for C<sub>15</sub>H<sub>13</sub>N<sub>2</sub>O<sub>2</sub> [M-H]<sup>-</sup> 253.0977, found 253.0982.

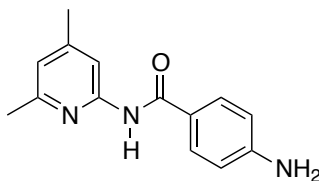
### ***N*-(4,6-dimethylpyridin-2-yl)-4-nitrobenzamide<sup>219</sup>**



4,6-Dimethylpyridin-2-amine (2.1 g, 19.6 mmol, 1.2 eq) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (10 ml) and triethylamine (10 mL), cooled to 0 °C and stirred under N<sub>2</sub>. A solution of 4-nitrobenzoyl chloride (3 g, 16.3 mmol, 1 eq) in CH<sub>2</sub>Cl<sub>2</sub> (50 ml) was added dropwise to the amine solution upon complete addition, the reaction mixture was allowed to warm to room temperature and stirred for 16 h. The reaction was quenched by addition of (sat) NaHCO<sub>3</sub> (aq) (100 mL), the organic layer was separated, washed with

1M HCl<sub>(aq)</sub>, dried (MgSO<sub>4</sub>), filtered and the solvent removed under vacuum to yield the crude product which was recrystallised from CH<sub>2</sub>Cl<sub>2</sub> to yield the product (4.01 g, 14.7 mmol, 90%) as a pale yellow crystalline solid; M.p. = 216.0–218.5 °C (lit<sup>219</sup> 156.0–158.0 °C); <sup>1</sup>H NMR (400.1 MHz; d<sub>6</sub>-DMSO) δ = 11.02 (1H, s, NH), 8.30 (2H, m, ArH), 8.20 (2H, m, ArH), 7.86 (1H, s, ArH), 6.90 (1H, s, ArH), 2.40 (3H, s, CH<sub>3</sub>) and 2.31 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz; d<sub>6</sub>-DMSO) δ = 165.6 (ArC) 164.2 (ArC), 156.1 (ArC), 151.0 (ArC), 149.0 (ArC), 139.7 (ArC), 130.5 (ArCH), 129.3 (ArCH), 123.5 (ArCH), 123.1 (ArCH), 120.2 (ArCH), 112.1 (ArCH), 23.1 (CH<sub>3</sub>) and 20.6 (CH<sub>3</sub>); MS (ES<sup>+</sup>) *m/z* 272 ([M+H]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>14</sub>H<sub>14</sub>N<sub>3</sub>O<sub>3</sub> 272.1035, found 272.1037.

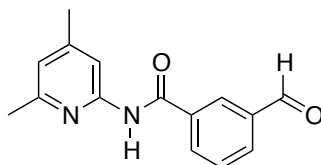
#### 4-Amino-N-(4,6-dimethylpyridin-2-yl)benzamide (88)



*N*-(4,6-dimethylpyridin-2-yl)-4-nitrobenzamide (2.30 g, 8.5 mmol, 1 eq) was dissolved in methanol:THF (10:1) (25 ml). This was added to Pd/C 10 % wt (0.45 g), a hydrogen balloon fitted and the reaction mixture stirred vigorously with the balloon being refreshed when required. The reaction progress was monitored by TLC analysis and upon completion the reaction mixture was filtered through celite the solvent removed under vacuum to give the crude product as a brown gum which was washed with NaHCO<sub>3</sub> (aq), extracted into CH<sub>2</sub>Cl<sub>2</sub>, dried (MgSO<sub>4</sub>), filtered and the solvent removed under vacuum. the product was then recrystallised from EtOAc/DCM and hexane and yielded the amine (2.07 g, 8.4 mmol, 99%) as a pale brown solid; M.p. = 184.2–184.9 °C (lit<sup>221</sup> 184–185.5 °C); <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>): δ = 8.45 (1H, s, NH), 8.04 (1H, s, ArH), 7.76 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.7 Hz, 2×ArH), 6.73 (1H, s, ArH), 6.69 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.7 Hz, 2×ArH), 4.07 (2H, br, NH<sub>2</sub>), 2.42 (3H, s, CH<sub>3</sub>) and 2.34 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100.1 MHz, CDCl<sub>3</sub>) δ = 165.4 (CO), 156.2 (ArC), 151.3 (ArC), 150.3 (ArC), 150.1 (ArC), 129.2 (ArCH), 123.6 (ArC), 120.2 (ArCH), 114.2 (ArCH), 111.6 (ArCH), 23.8 (CH<sub>3</sub>) and 21.3 (CH<sub>3</sub>); MS (ES<sup>+</sup>) *m/z* 265 ([M+Na]<sup>+</sup>, 100) and 242 ([M+H]<sup>+</sup>, 28); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>14</sub>H<sub>15</sub>N<sub>3</sub>ONa [M+Na]<sup>+</sup> 264.1113, found 264.1117.

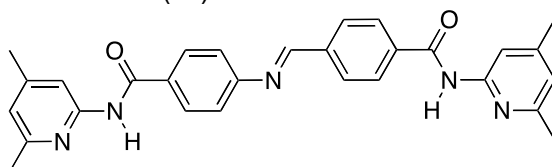


***N*-(4,6-dimethylpyridin-2-yl)-3-formylbenzamide (90)**



2-Amino-4, 6-dimethylpyridine (0.54 g, 4.4 mmol, 1 eq) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (10 mL) and triethylamine (10 mL) and cooled to 0 °C. A solution of 3-formylbenzoyl chloride (0.82 g, 4.8 mmol, 1.1 eq) in CH<sub>2</sub>Cl<sub>2</sub> (30 mL) was added dropwise and after complete addition, the reaction was allowed to rise slowly to RT. The reaction progress was monitored by TLC analysis and upon completion (16 h) was quenched by addition of Na<sub>2</sub>CO<sub>3</sub> (aq) (100 mL). The organic layer was then separated, washed with 1M HCl (aq) (100 mL), dried (MgSO<sub>4</sub>), filtered and the solvent removed to give the crude product which was purified by flash chromatography (3:2, cyclohexane: ethyl acetate) to yield the aldehyde (0.92 g, 3.6 mmol, 82%) as a colourless powder; M.p. = 114.5–116.8 °C; <sup>1</sup>H NMR (400.1 MHz; CDCl<sub>3</sub>): δ 10.10 (1H, s, ArCHO), 8.64 (1H, br, NH) 8.43 (1H, dd, <sup>4</sup>J<sub>HH</sub> = 1.3 Hz & 1.5 Hz, ArH), 8.23 (1H, ddd, <sup>3</sup>J<sub>HH</sub> = 7.7 Hz <sup>4</sup>J<sub>HH</sub> = 1.3 Hz & 1.5 Hz, ArH), 8.08 (1H, ddd, <sup>3</sup>J<sub>HH</sub> = 7.7 Hz <sup>4</sup>J<sub>HH</sub> = 1.3 Hz & 1.5 Hz, ArH), 8.05 (1H, s, ArH), 7.70 (1H, dd, <sup>3</sup>J<sub>HH</sub> = 7.7 Hz & 7.7 Hz, ArH), 6.81 (1H, s, ArH), 2.45 (3H, s, CH<sub>3</sub>) and 2.38 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz; CDCl<sub>3</sub>): δ 191.2 (ArCH), 164.4 (ArC), 156.5 (ArC), 150.5 (ArC), 150.3 (ArC), 136.6 (ArC), 135.5 (ArC), 133.0 (ArCH), 132.5 (ArCH), 129.6 (ArCH), 128.4 (ArCH), 120.9 (ArCH), 111.8 (ArCH), 23.6 (CH<sub>3</sub>) and 21.3 (CH<sub>3</sub>); MS (ES<sup>+</sup>) *m/z* 308.91 (100), 293.97 ([M+K]<sup>+</sup>, 20) and 276.93 ([M+Na]<sup>+</sup>, 50); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>15</sub>H<sub>14</sub>N<sub>2</sub>O<sub>2</sub>Na [M+Na]<sup>+</sup> 277.0953, found 277.0947.

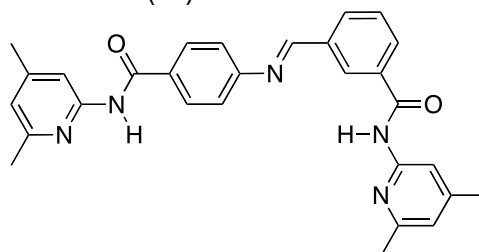
***(E)*-*N*-(4,6-dimethylpyridin-2-yl)-4-((4-((4,6-dimethylpyridin-2-yl)carbamoyl)benzylidene)amino)benzamide (97)**



Amine **88** (0.64 g, 2.5 mmol, 1 eq) was dissolved in ethanol (10 mL) before addition of aldehyde **89** (0.6 g, 2.5 mmol, 1 eq). The solution was then left to sit overnight before isolating the precipitated product by filtration. The product was isolated as a yellow crystalline solid (1.03 g, 2.16 mmol, 86 %). M.p. = 222.3–222.5 °C; <sup>1</sup>H NMR

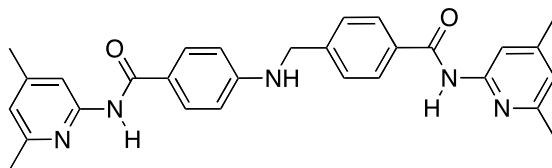
(400.1 MHz; CDCl<sub>3</sub>)  $\delta$  8.70 (1H, br, NH), 8.67 (1H, br, NH), 8.52 (1H, s, Ar-NCH-Ar), 8.03–8.07 (6H, m, ArH), 7.98–8.01 (2H, m, ArH), 7.28–7.32 (2H, m, ArH), 6.79 (1H, s, ArH), 6.78 (1H, s, ArH), 2.43 (3H, s, CH<sub>3</sub>), 2.42 (3H, s, CH<sub>3</sub>), 2.37 (3H, s, CH<sub>3</sub>) and 2.36 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz; CDCl<sub>3</sub>)  $\delta$  165.0 (CO), 164.8 (CO), 160.2 (CH), 156.46 (ArC), 156.35 (ArC), 154.9 (ArC), 150.8 (ArC), 150.6 (ArC), 150.2 (ArC), 150.2 (ArC), 138.8 (ArC), 137.0 (ArC), 132.0 (ArC), 129.3 (ArCH), 128.5 (ArCH), 127.7 (ArCH), 121.1 (ArCH), 120.9 (ArCH), 120.6 (ArCH), 111.74 (ArCH), 111.64 (ArCH), 23.9 (CH<sub>3</sub>), 23.9 (CH<sub>3</sub>) 21.3 (CH<sub>3</sub>) and 21.3 (CH<sub>3</sub>); MS (ES<sup>+</sup>)  $m/z$  499.86 ([M+Na]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>)  $m/z$  calculated for C<sub>29</sub>H<sub>28</sub>N<sub>5</sub>O<sub>2</sub> [M+H]<sup>+</sup> 478.2238, found 478.2237, calculated for C<sub>29</sub>H<sub>27</sub>N<sub>5</sub>O<sub>2</sub>Na [M+Na]<sup>+</sup> 500.2062, found 500.2061.

**(E)-N-(4,6-dimethylpyridin-2-yl)-3-(((4-((4,6-dimethylpyridin-2-yl)carbamoyl)phenyl)imino)methyl)benzamide (98)**



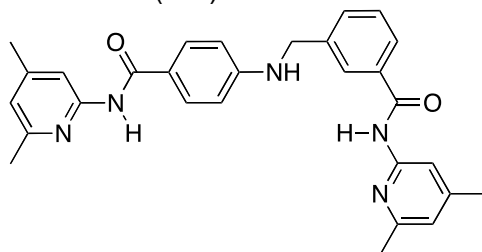
Amine **88** (0.30 g, 1.23 mmol, 1 eq) was dissolved in ethanol (5 mL) aldehyde **90** (0.31 g, 1.23 mmol, 1 eq) and allowed to sit overnight in darkness. The product precipitated from the solution and was isolated by filtration to yield the product an off-white solid powder (0.2 g, 0.4 mmol, 34%): M.p. = 150.3–150.6 °C; <sup>1</sup>H NMR (400.1 MHz; CDCl<sub>3</sub>)  $\delta$  = 8.74 (1H, br, NH), 8.73 (1H, br, NH), 8.49 (1H, t, <sup>4</sup>J<sub>HH</sub> = 1.7 Hz, ArH), 8.46 (1H, s, Ar-NCH-Ar), 8.14–8.08 (4H, m, ArH), 8.05–8.01 (2H, m, ArH), 7.64 (1H, t, <sup>3</sup>J<sub>HH</sub> = 7.8 Hz, ArH), 7.34–7.29 (2H, m, ArH), 6.81 (1H, t, <sup>4</sup>J<sub>HH</sub> = 0.63 Hz, ArH), 6.80 (1H, t, <sup>4</sup>J<sub>HH</sub> = 0.67 Hz, ArH), 2.464 (3H, s, CH<sub>3</sub>), 2.459 (3H, s, CH<sub>3</sub>), 2.388 (3H, s, CH<sub>3</sub>) and 2.387 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>)  $\delta$  = 165.1 (NHCO), 164.8 (NHCO), 160.4 (CH), 156.7 (ArC), 156.6 (ArC), 155.1 (ArC) 151.0 (ArC), 150.8 (ArC), 150.4 (ArC), 150.4 (ArC), 136.5 (ArC), 135.3 (ArC), 132.6 (ArC), 132.1 (ArCH), 130.7 (ArCH), 129.6 (ArCH), 128.7 (ArCH), 127.4 (ArCH), 121.3 (ArCH), 121.0 (ArCH), 120.8 (ArCH), 111.8 (ArCH), 111.7 (ArCH), 24.0 (CH<sub>3</sub>), 21.5 (CH<sub>3</sub>); MS (CI<sup>+</sup>)  $m/z$  478.24 ([M+H]<sup>+</sup>, 100). HRMS (ES<sup>+</sup>)  $m/z$  calculated for C<sub>29</sub>H<sub>28</sub>N<sub>5</sub>O<sub>2</sub> [M+H]<sup>+</sup> 478.2238, found 478.2250.

***N*-(4,6-dimethylpyridin-2-yl)-4-(((4-((4,6-dimethylpyridin-2-yl)carbamoyl)benzyl)amino)benzamide) (103)**



Imine **97** (0.1 g, 0.21 mmol, 1 eq) was dissolved in dry THF (3 mL). Sodium borohydride (0.016 g, 0.42 mmol, 2 eq) was added to the reaction mixture and stirred at RT and monitored by TLC analysis. Upon completion, (typically 7 h) CH<sub>2</sub>Cl<sub>2</sub> (50 mL) was added to quench the reaction and washed with water (10 mL) and saturated ammonium chloride solution (40 mL). The aqueous and organic layers were separated and the CH<sub>2</sub>Cl<sub>2</sub> was dried over MgSO<sub>4</sub>, filtered and the CH<sub>2</sub>Cl<sub>2</sub> removed under reduced pressure to obtain the product as a colourless solid (0.1 g, 0.21 mmol, quantitative). M.p. = 116.0–118.0 °C; <sup>1</sup>H NMR (400.1 MHz; CDCl<sub>3</sub>) δ 8.58 (1H, br, NH), 8.55 (1H, br, NH), 8.03–8.06 (2H, m, ArH), 7.87–7.93 (2H, m, ArH), 7.76–7.81 (2H, m, ArH), 7.47 (1H, s, ArH), 7.44 (1H, s, ArH), 6.77 (1H, s, ArH), 6.73 (1H, s, ArH), 6.61–6.66 (2H, m, ArH), 4.67 (1H, t, *J* = 5.8 Hz, Ar-NH-CH<sub>2</sub>-Ar), 4.48 (2H, d, *J* = 5.8 Hz, Ar-NH-CH<sub>2</sub>-Ar), 2.42 (3H, s, CH<sub>3</sub>), 2.42 (3H, s, CH<sub>3</sub>), 2.35 (3H, s, CH<sub>3</sub>) and 2.34 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz; CDCl<sub>3</sub>) δ 165.40 (C), 165.34 (C), 156.6 (C), 156.3 (C), 151.4 (C), 151.04 (C), 150.92 (C), 150.39 (C), 150.24 (C), 143.1 (C), 133.8 (C), 129.3 (ArCH), 127.8 (ArCH), 127.6 (ArCH), 123.1 (C), 120.8 (ArCH), 120.3 (ArCH), 112.3 (ArCH), 111.79 (ArCH), 111.64 (ArCH), 47.5 (CH<sub>2</sub>), 23.9 (CH<sub>3</sub>) and 21.5 (CH<sub>3</sub>); MS (ES<sup>+</sup>) *m/z* 501.96 ([M+Na]<sup>+</sup>, 100) and 480.00 ([M+H]<sup>+</sup>, 35); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>29</sub>H<sub>30</sub>N<sub>5</sub>O<sub>2</sub> [M+H]<sup>+</sup> 480.2400, found 480.2391.

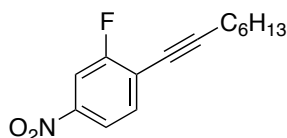
***N*-(4,6-dimethylpyridin-2-yl)-3-(((4-((4,6-dimethylpyridin-2-yl)carbamoyl)phenyl)amino)methyl)benzamide (104)**



Imine **104** was synthesized using the same experimental procedure as imine **103** to yield the product after purification by column chromatography was a colourless powder (0.12g, 59 %); M.p. = 172.2–172.5 °C: <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 8.58

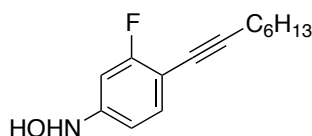
(s, 2H), 8.05 (d,  $J = 3.8$ , 2H), 7.95 (t,  $J = 2.1$ , 1H), 7.84 (dt,  $J = 7.7$ , 1.2, 1H), 7.81–7.79 (m, 2H), 7.56 (d,  $J = 7.8$ , 1H), 7.48 (t,  $J = 7.7$ , 1H), 6.78 (s, 1H), 6.73 (s, 1H), 6.67–6.65 (m, 2H), 4.62 (t,  $J = 5.5$ , 1H), 4.49 (d,  $J = 5.7$ , 2H), 2.43 (s, 3H), 2.43 (s, 3H), 2.37 (s, 3H), 2.34 (s, 3H);  $^{13}\text{C}$  NMR (100.6 MHz,  $\text{CDCl}_3$ )  $\delta = 165.5$  (CO), 165.4 (CO), 156.5 (ArC), 156.2 (ArC), 151.4 (ArC), 151.0 (ArC), 150.8 (ArC), 150.4 (ArC), 150.3 (ArC), 139.6 (ArC), 135.2 (ArC), 131.1 (ArCH), 129.4 (ArCH), 129.3 (ArCH), 126.3 (ArCH), 126.3 (ArCH), 123.0 (ArC), 120.9 (ArCH), 120.2 (ArCH), 112.3 (ArCH), 111.8 (ArCH), 111.6 (ArCH), 47.6 ( $\text{CH}_2$ ), 23.92 ( $\text{CH}_3$ ), 23.91 ( $\text{CH}_3$ ), 21.48 ( $\text{CH}_3$ ) and 21.46 ( $\text{CH}_3$ ); MS ( $\text{ES}^+$ )  $m/z$  501.97 ( $[\text{M}+\text{Na}]^+$ , 100) and 480.01 ( $[\text{M}+\text{H}]^+$ , 40); HRMS ( $\text{ES}^+$ )  $m/z$  calculated for  $\text{C}_{29}\text{H}_{30}\text{N}_5\text{O}_2$   $[\text{M}+\text{H}]^+$  480.2400 found 480.2380.

### 2-Fluoro-4-nitro-1-(oct-1-yn-1-yl)benzene



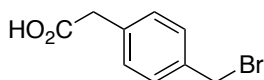
1-Bromo-2-fluoro-4-nitrobenzene (5 g, 22.7 mmol, 1 eq), octyne (10.02 g, 90.9 mmol, 4 eq),  $\text{Pd}(\text{P}(\text{Ph}_3)_2)\text{Cl}_2$  (1.43 g, 1.59 mmol, 0.07 eq),  $\text{CuI}$  (0.30 g, 1.59 mmol, 0.07 eq) and  $\text{PPh}_3$  (0.83g, 3.18 mmol, 0.14 eq) were dissolved in triethylamine (100 mL) which was vigorously degassed with  $\text{N}_2$ . The reaction was heated to  $80\text{ }^\circ\text{C}$  and the progress monitored by TLC analysis. Upon completion (16 h) the reaction was filtered through celite and the solvent removed. The residue was washed with  $\text{NH}_4\text{Cl}$  (aq) (100 mL) and the product extracted into  $\text{CH}_2\text{Cl}_2$  ( $3 \times 100$  mL), the organic fractions separated, dried ( $\text{MgSO}_4$ ), filtered and the solvent removed under vacuum. The crude product was then purified by flash chromatography (cyclohexane) to furnish the purified product as an orange oil (3.3 g, 13.4 mmol, 59%);  $^1\text{H}$  NMR (400.1 MHz,  $\text{CDCl}_3$ )  $\delta = 7.97$  (1H, m, ArH), 7.93 (1H, m, ArH), 7.53 (1H, m, ArH), 2.49 (2H, t,  $^3J_{\text{HH}} = 7.1$  Hz,  $\text{CH}_2$ ), 1.67–1.60 (2H, m,  $\text{CH}_2$ ), 1.50–1.43 (2H, m,  $\text{CH}_2$ ), 1.35–1.30 (4H, m,  $2 \times \text{CH}_2$ ) and 0.90 (3H, t,  $^3J_{\text{HH}} = 7.0$  Hz,  $\text{CH}_3$ );  $^{13}\text{C}$  NMR (100.6 MHz,  $\text{CDCl}_3$ )  $\delta = 162.4$  (d,  $^1J_{\text{CF}} = 254.7$  Hz, ArC–F), 147.3 (d,  $^3J_{\text{CF}} = 8.2$  Hz, ArC– $\text{NO}_2$ ), 134.0 (d,  $^4J_{\text{CF}} = 2.0$  Hz, ArCH), 120.1 (d,  $^2J_{\text{CF}} = 16.4$  Hz, ArC), 119.2 (d,  $3J_{\text{CF}} = 3.7$  Hz, ArCH), 111.4 (d,  $^2J_{\text{CF}} = 26.4$  Hz, ArCH), 102.5 (d,  $^3J_{\text{CF}} = 3.6$  Hz, Ar– $\text{C}\equiv\text{C}$ ), 73.0 (Ar– $\text{C}\equiv\text{C}$ ), 31.4 ( $\text{CH}_2$ ), 28.7 ( $\text{CH}_2$ ), 28.4 ( $\text{CH}_2$ ), 22.7 ( $\text{CH}_2$ ), 19.9 ( $\text{CH}_2$ ) and 14.2 ( $\text{CH}_3$ );  $^{19}\text{F}$   $\{^1\text{H}\}$  NMR (376.5 MHz,  $\text{CDCl}_3$ ):  $\delta_{\text{F}} -107.3$ ; MS ( $\text{ES}^+$ )  $m/z$  250 ( $[\text{M}+\text{H}]^+$ , 100); HRMS ( $\text{ES}^+$ )  $m/z$  calculated for  $\text{C}_{14}\text{H}_{17}\text{NO}_2\text{F}$   $[\text{M}+\text{H}]^+$  250.1243, found 250.1242.

### ***N*-(3-fluoro-4-(oct-1-yn-1-yl)phenyl)hydroxylamine**



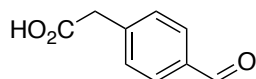
2-Fluoro-4-nitro-1-(oct-1-yn-1-yl)benzene (0.5 g, 2 mmol, 1 eq) was dissolved in THF (20 mL) and stirred with Rh/C (0.1 g) for 15 min. Hydrazine hydrate (0.15 mL, 3 mmol, 1.5 eq) was then added and the reaction progress monitored by TLC analysis. Upon completion (20 min) the suspension was filtered through celite and the solvent removed to yield the hydroxylamine (0.5 g, 2 mmol, quantitative) as a colourless liquid:  $^1\text{H}$   $\{^{19}\text{F}\}$  NMR (400.1 MHz,  $\text{CDCl}_3$ ):  $\delta$  = 7.26 (1H, d,  $^3J_{\text{HH}} = 8.1$  Hz, *ArH*), 6.80 (1H, s, *OH*), 6.75 (1H, dd,  $^3J_{\text{HH}} = 8.1$  Hz and  $^4J_{\text{HH}} = 2.1$  Hz, *ArH*), 6.62 (1H, dd,  $^3J_{\text{HH}} = 8.1$  Hz and  $^4J_{\text{HH}} = 2.1$  Hz, *ArH*), 5.17 (1H, d,  $^4J_{\text{HH}} = 2.2$  Hz, *NH*), 2.42 (2H, t,  $^3J_{\text{HH}} = 7.1$  Hz,  $\text{CH}_2$ ), 1.62 (2H, m,  $\text{CH}_2$ ), 1.49–1.42 (2H, m,  $\text{CH}_2$ ), 1.32 (4H, m,  $2 \times \text{CH}_2$ ) and 0.90 (3H, t,  $^3J_{\text{HH}} = 7.0$  Hz,  $\text{CH}_3$ );  $^{19}\text{F}$   $\{^1\text{H}\}$  NMR (376.5 MHz,  $\text{CDCl}_3$ ):  $\delta_{\text{F}}$  -110.2 (*ArF*).

### **2-(4-(Bromomethyl)phenyl)acetic acid**<sup>222</sup>



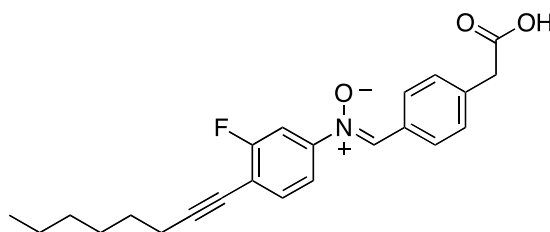
A solution of *p*-tolylacetic acid (6.30 g, 42.0 mmol, 1 eq) and bromine (7.35 g, 46.0 mmol, 1.1 eq) in chlorobenzene (50 mL) was stirred for 2 hours under irradiation of a 60 watt tungsten lamp. The occurring precipitate was filtered, washed with copious amounts of hexanes and dried to yield the desired product as a white solid (6.03 g, 62%); M.p. 176.8–178.8 °C (lit<sup>222</sup> 180 °C)  $^1\text{H}$  NMR (400.1 MHz,  $d_6$ -DMSO):  $\delta$  = 7.38 (2H,  $^3J_{\text{HH}} = 8.2$  Hz,  $2 \times \text{ArH}$ ), 7.24 (2H, d,  $^3J_{\text{HH}} = 8.2$  Hz,  $2 \times \text{ArH}$ ), 4.69 (2H, s,  $\text{CH}_2$ ) and 3.57 (2H, s,  $\text{CH}_2$ );  $^{13}\text{C}$  NMR (75.5 MHz,  $d_6$ -DMSO)  $\delta$  = 172.5 ( $\text{CO}_2\text{H}$ ), 136.3 (*ArC*), 135.2 (*ArC*), 129.7 (*ArCH*), 129.2 (*ArCH*), 40.3 ( $\text{CH}_2$ ) and 34.4 ( $\text{CH}_2$ ); MS ( $\text{EI}^+$ )  $m/z$  229 ( $[\text{BrM}^+]$ , 5), 227 ( $[\text{BrM}^+]$ , 5), 149 (100). HRMS ( $\text{ES}^+$ ) calculated for  $\text{C}_9\text{H}_9\text{O}_2^{81}\text{Br}$  [ $\text{M}^+$ ] 229.9765, found 229.9761, calculated for  $\text{C}_9\text{H}_9\text{O}_2^{79}\text{Br}$  227.9786, found 227.9793.

## 2-(4-Formylphenyl)acetic acid <sup>197</sup> (**100**)



(4-Bromomethylphenyl)acetic acid (2.00 g, 8.8 mmol, 1 eq) and hexamethylene tetramine (1.90 g, 26.2 mmol, 3 eq) were dissolved in an EtOH/H<sub>2</sub>O mixture (1:1, 10 mL) and heated under reflux for 4 hours. While still heating under reflux, concentrated HCl (4 mL) was then added to the solution. After 30 minutes the reaction mixture was left to cool to room temperature and diluted in water (30 mL). The reaction mixture was then extracted using CH<sub>2</sub>Cl<sub>2</sub> (2 x 100 mL). The organic fractions were combined and dried over MgSO<sub>4</sub>, before the filtrate is concentrated under reduced pressure to afford **100** as a colorless solid (0.80 g, 4.4 mmol, 50%); M.p. 125-127 °C (lit<sup>140</sup> 125.0–127.0 °C); <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 10.00 (1H, s, CHO), 7.86 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.0 Hz, 2×ArH), 7.46 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.0 Hz, 2×ArH), 3.74 (2H, s, CH<sub>2</sub>); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ = 192.1 (CHO), 176.5 (CO<sub>2</sub>H), 140.3 (ArC), 135.6 (ArC), 130.29 (ArCH), 130.18 (ArCH) and 41.2 (CH<sub>2</sub>); IR (KBr plates) 2617, 1726, 1660. MS (ES<sup>+</sup>) *m/z* 164 ([M<sup>+</sup>], 80).

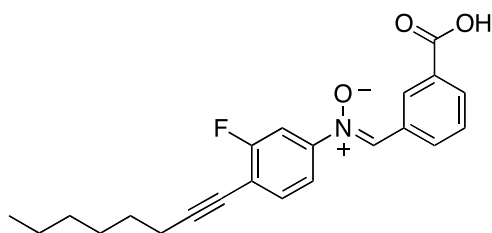
## (Z)-N-(4-(carboxymethyl)benzylidene)-3-fluoro-4-(oct-1-yn-1-yl)aniline oxide (**92**)



*N*-(3-fluoro-4-(oct-1-yn-1-yl)phenyl)hydroxylamine (0.42 g, 1.78 mmol, 1 eq) and aldehyde **100** (0.29 g, 1.78 mmol, 1 eq) were dissolved in ethanol (15 mL) and left to sit in darkness for 16 h. The precipitate which formed was isolated by filtration, washed with cold ethanol and dried to yield nitron **92** (0.35 g, 0.93 mmol, 52%) as a colourless crystalline solid; M.p. 159.9–160.3 °C; <sup>1</sup>H {<sup>19</sup>F} NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 8.26 (2H, <sup>3</sup>J<sub>HH</sub> = 8.3 Hz, 2×ArH), 7.93 (1H, s, CH), 7.58 (1H, dd, <sup>3</sup>J<sub>HH</sub> = 9.5 and <sup>4</sup>J<sub>HH</sub> = 1.6 Hz, ArH), 7.51–7.44 (2H, m, 2×ArH), 7.33 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.3 Hz, 2×ArH), 3.63 (2H, s, CH<sub>2</sub>), 2.47 (2H, t, <sup>3</sup>J<sub>HH</sub> = 7.1 Hz, CH<sub>2</sub>), 1.74–1.57 (2H, m, CH<sub>2</sub>), 1.54–1.41 (2H, m, CH<sub>2</sub>), 1.41–1.28 (4H, m, CH<sub>2</sub>) and 0.89 (3H, t, <sup>3</sup>J<sub>HH</sub> = 7.1 Hz, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ = 174.7 (CO<sub>2</sub>H), 162.7 (d, <sup>1</sup>J<sub>CF</sub> = 253.0 Hz, ArC–F), 148.1 (d,

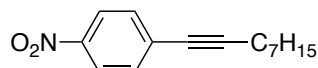
$^3J_{CF} = 8.4$  Hz, ArC), 138.1 (ArC), 136.6 (CH), 133.9 (d,  $^4J_{CF} = 2.2$  Hz, ArCH), 130.1 (ArCH), 129.8 (ArCH), 128.9 (ArC), 117.0 (d,  $^3J_{CF} = 3.6$  Hz, ArCH), 115.0 (d,  $^2J_{CF} = 16.3$  Hz, ArC), 110.0 (d,  $^2J_{CF} = 26.2$  Hz, ArCH), 99.0 (d,  $^3J_{CF} = 3.2$  Hz, Ar-C≡C), 73.1 (Ar-C≡C), 41.5 (CH<sub>2</sub>), 31.5 (CH<sub>2</sub>), 28.7 (CH<sub>2</sub>), 28.5 (CH<sub>2</sub>), 22.7 (CH<sub>2</sub>), 19.8 (CH<sub>2</sub>) and 14.2 (CH<sub>3</sub>); MS (ES<sup>+</sup>) *m/z* 404 ([M+Na]<sup>+</sup>, 33) and 382 ([M+H]<sup>+</sup>, 100). HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>23</sub>H<sub>25</sub>NO<sub>3</sub>F [M+H]<sup>+</sup> 382.1818, found 382.1813.

**(Z)-N-(3-carboxybenzylidene)-3-fluoro-4-(oct-1-yn-1-yl)aniline oxide (91)**



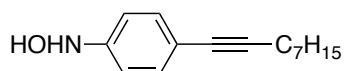
*N*-(3-fluoro-4-(oct-1-yn-1-yl)phenyl)hydroxylamine (0.34 g, 1.46 mmol, 1 eq) and 3-carboxybenzaldehyde (0.22 g, 1.46 mmol, 1 eq) were dissolved in ethanol (10 mL) and left to sit in darkness for 16 h. The precipitate which formed was isolated by filtration, washed with cold ethanol and dried to yield nitronium **91** (0.23 g, 0.62 mmol, 43%) as a pale yellow crystalline solid; M.p. 185.4–185.8 °C; <sup>1</sup>H {<sup>19</sup>F} NMR (NMR (400.1 MHz; CDCl<sub>3</sub>) δ 9.10 (1H, t, *J* = 1.6 Hz, ArH), 8.76 (1H, s, CH), 8.69 (1H, m, ArH), 8.05 (1H, m, ArH), 7.97 (1H, dd,  $^3J_{HH} = 8.3$  Hz and  $^4J_{HH} = 2.2$  Hz, ArH), 7.85 (1H, m, ArH), 7.66 (1H, dd,  $^3J_{HH} = 8.1$  Hz and  $^4J_{HH} = 1.9$  Hz, ArH), 7.64 (1H, dd,  $^3J_{HH} = 7.8$  Hz and  $^4J_{HH} = 1.5$  Hz, ArH), 2.51 (2H, t,  $^3J_{HH} = 6.3$  Hz, CH<sub>2</sub>), 1.52–1.62 (2H, m, CH<sub>2</sub>), 1.38–1.48 (2H, m, CH<sub>2</sub>), 1.26–1.35 (4H, m, 2×CH<sub>2</sub>) and 0.89 (3H, t,  $^3J_{HH} = 7.0$  Hz, CH<sub>3</sub>); <sup>13</sup>C NMR (75.5 MHz, d<sub>6</sub>-DMSO): δ = 166.9 (CO<sub>2</sub>H), 161.6 (d,  $^1J_{CF} = 249.4$  Hz, ArC-F), 148.2 (d,  $^3J_{CF} = 8.6$  Hz, ArC), 133.7 (d,  $^4J_{CF} = 1.8$  Hz, ArCH), 133.6 (CH), 132.9 (ArCH), 131.4 (ArCH), 131.1 (ArC), 131.0 (ArC), 130.1 (ArCH), 128.9 (ArCH), 117.5 (d,  $^3J_{CF} = 3.4$  Hz, ArCH), 113.1 (d,  $^2J_{CF} = 16.3$  Hz, ArC), 109.4 (d,  $^2J_{CF} = 26.3$  Hz, ArCH), 98.6 (d,  $^3J_{CF} = 3.2$  Hz, Ar-C≡C), 73.1 (Ar-C≡C), 30.7 (CH<sub>2</sub>), 27.9 (CH<sub>2</sub>), 27.8 (CH<sub>2</sub>), 22.0 (CH<sub>2</sub>), 18.9 (CH<sub>2</sub>) and 13.9 (CH<sub>3</sub>). <sup>19</sup>F {<sup>1</sup>H} NMR (282.4 MHz, d<sub>6</sub>-DMSO): δ<sub>F</sub> -109.9 (ArF); MS (ES<sup>-</sup>) *m/z* 366 ([M-H]<sup>-</sup>, 100); HRMS (ES<sup>-</sup>) *m/z* calculated for C<sub>22</sub>H<sub>21</sub>NO<sub>3</sub>F [M-H]<sup>-</sup> 366.1505, found 366.1515.

### 1-Nitro-4-(non-1-yn-1-yl)benzene



4-iodonitrobenzene (5 g, 20 mmol, 1 eq) was dissolved in a mixture of THF (150 mL) and triethylamine (50 mL) and degassed vigorously. To the solution was added nonyne (9.9 g, 80 mmol, 4 eq), Pd(P(Ph<sub>3</sub>)<sub>2</sub>)Cl<sub>2</sub> (0.7 g, 1 mmol, 0.05 eq), CuI (0.19 g, 1 mmol, 0.05 eq) and PPh<sub>3</sub> (0.58 g, 2 mmol, 0.1 eq). The solution was then heated to 50 °C for 16 h and upon completion was filtered through celite and the solvent removed under vacuum. The crude product was then extracted into CH<sub>2</sub>Cl<sub>2</sub> (100 mL), washed with NH<sub>4</sub>Cl(aq) (100 mL), dried over MgSO<sub>4</sub>, filtered and the solvent removed to give the crude product which was purified by column chromatography using a gradient of ethyl acetate in cyclohexane to elute the product obtained after drying as a colourless oil (3.67 g, 15 mmol, 75%): <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 8.15 (2H, <sup>3</sup>J<sub>HH</sub> = 8.9 Hz, 2×ArH), 7.51 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.9 Hz, 2×ArH), 2.44 (2H, t, <sup>3</sup>J<sub>HH</sub> = 7.1, CH<sub>2</sub>), 1.60–1.64 (2H, m, CH<sub>2</sub>), 1.51–1.39 (2H, m, CH<sub>2</sub>), 1.39–1.24 (6H, m, 3×CH<sub>2</sub>) and 0.93–0.86 (m, 3H, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ = 146.7 (ArC), 132.4 (ArCH), 131.4 (ArC), 123.6 (ArCH), 97.0 (Ar–C≡C), 79.4 (Ar–C≡C), 31.9 (CH<sub>2</sub>), 29.1 (CH<sub>2</sub>), 28.9 (CH<sub>2</sub>), 28.6 (CH<sub>2</sub>), 22.8 (CH<sub>2</sub>), 19.7 (CH<sub>2</sub>) and 14.2 (CH<sub>3</sub>); MS (ES<sup>+</sup>) *m/z* 268 ([M+Na]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>15</sub>H<sub>19</sub>NO<sub>2</sub>Na [M+Na]<sup>+</sup> 268.1313, found 268.1310.

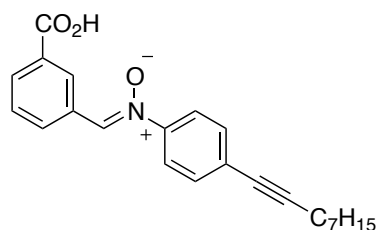
### *N*-(3-fluoro-4-(oct-1-yn-1-yl)phenyl)hydroxylamine



1-Nitro-4-(non-1-yn-1-yl)benzene (0.61 g, 2.5 mmol, 1 eq) was dissolved in THF (20 mL) and Rh/C (0.1 g) added. The suspension was then stirred under N<sub>2</sub> at RT for 30 mins before addition of hydrazine hydrate (0.19 mL, 3.75 mmol, 1.5 eq). The solution was allowed to stir at RT and the reaction progress monitored by TLC analysis. Upon completion, typically 1 h, the reaction solution was filtered through celite and then the solvent removed under vacuum and the product dried by high vacuum to yield the hydroxylamine (0.58 g, 2.5 mmol, quantitative) as a pale yellow liquid which was used without further purification. <sup>1</sup>H NMR (300.1 MHz, CDCl<sub>3</sub>) δ = 7.31 (2H, <sup>3</sup>J<sub>HH</sub> = 8.7 Hz, 2×ArH), 6.90 (2H, <sup>3</sup>J<sub>HH</sub> = 8.7 Hz, 2×ArH), 2.38 (2H, t, <sup>3</sup>J<sub>HH</sub> = 7.1 Hz, CH<sub>2</sub>), 1.61–1.57 (2H, m, CH<sub>2</sub>), 1.48–1.37 (2H, m, CH<sub>2</sub>), 1.38–1.24 (6H, m, 3×CH<sub>2</sub>), 0.91–0.85 (3H, m, CH<sub>3</sub>).

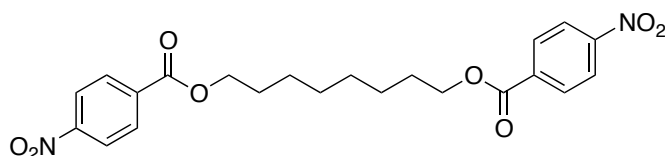


**(Z)-N-(3-carboxybenzylidene)-4-(non-1-yn-1-yl)aniline oxide (113)**



*N*-(3-fluoro-4-(oct-1-yn-1-yl)phenyl)hydroxylamine (0.58 g, 2.50 mmol, 1 eq) and 3-carboxybenzaldehyde (375 mg, 2.50 mmol, 1 eq) were dissolved in ethanol (15 mL) and the solution stirred in darkness for 16 h. After removal of the ethanol, the crude solid was recrystallised using CH<sub>2</sub>Cl<sub>2</sub> and petroleum ether to yield the desired product as a colourless solid (0.48 g, 1.28 mmol, 52%); M.p. 175.2–175.5 °C; <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 9.00 (1H, d, <sup>3</sup>J<sub>HH</sub> = 8.0 Hz, *ArH*), 8.77 (1H, s, *ArH*), 8.18 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.8 Hz, *ArH*), 8.03 (1H, s, *ArH*), 7.74 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.6 Hz, 2×*ArH*), 7.61 (1H, dd, <sup>3</sup>J<sub>HH</sub> = 8.0 and <sup>3</sup>J<sub>HH</sub> = 7.8 Hz, *ArH*), 7.50 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.6 Hz, 2×*ArH*), 2.43 (2H, t, <sup>3</sup>J<sub>HH</sub> = 7.1 Hz, CH<sub>2</sub>), 1.67–1.56 (2H, m, CH<sub>2</sub>), 1.49–1.40 (2H, m, CH<sub>2</sub>), 1.38–1.25 (6H, m, 3×CH<sub>2</sub>), 0.93–0.86 (3H, m, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ = 170.6 (CO<sub>2</sub>H), 147.3 (*ArC*), 134.3 (CH), 133.4 (*ArCH*), 132.5 (*ArCH*), 132.5 (*ArCH*), 131.5 (*ArCH*), 130.9 (*ArC*), 130.2 (*ArC*), 129.3 (*ArCH*), 126.7 (*ArC*), 121.7 (*ArCH*), 93.6 (*Ar-C≡C*), 79.6 (*Ar-C≡C*), 31.9 (CH<sub>2</sub>), 29.1 (CH<sub>2</sub>), 29.0 (CH<sub>2</sub>), 28.7 (CH<sub>2</sub>), 22.8 (CH<sub>2</sub>), 19.6 (CH<sub>2</sub>) and 14.2 (CH<sub>3</sub>); MS (ES<sup>+</sup>) *m/z* 386 ([M+Na]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>23</sub>H<sub>25</sub>NO<sub>3</sub>Na [M+Na]<sup>+</sup> 386.1732, found 386.1731.

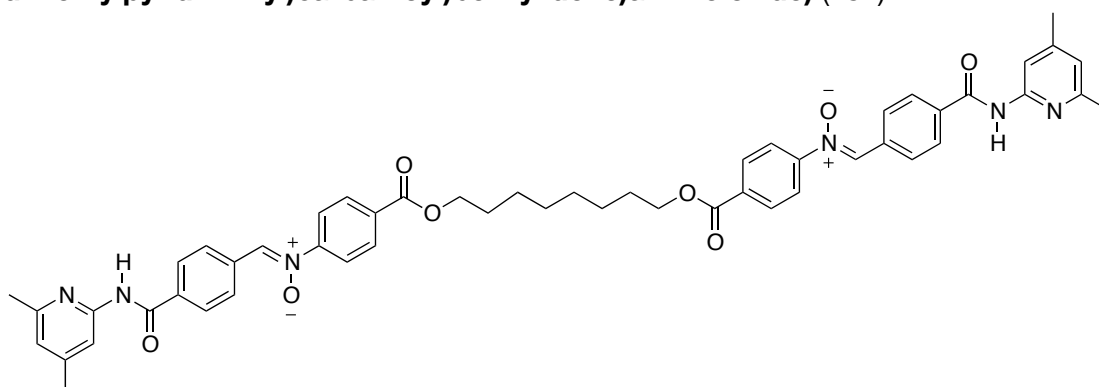
**Octane-1,8-diyl bis(4-nitrobenzoate) (150)**



1,8-Octanediol (1.58 g, 10.8 mmol, 1 eq) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (10 mL) and triethylamine (10 mL) and cooled to 0 °C under N<sub>2</sub>. A solution of 4-nitrobenzoyl chloride (6 g, 32 mmol, 3 eq) in CH<sub>2</sub>Cl<sub>2</sub> (50 mL) was added dropwise and upon complete addition, the reaction was allowed to rise to RT and the progress monitored by TLC analysis. Upon completion Na<sub>2</sub>CO<sub>3(aq)</sub> (100 mL) was added and stirred vigorously. The organic layer was then separated, dried (MgSO<sub>4</sub>), filtered and the solvent removed to give the crude product which was purified by flash chromatography (3:1, cyclohexane: ethyl acetate) to yield the purified dinitro (2.5 g, 402

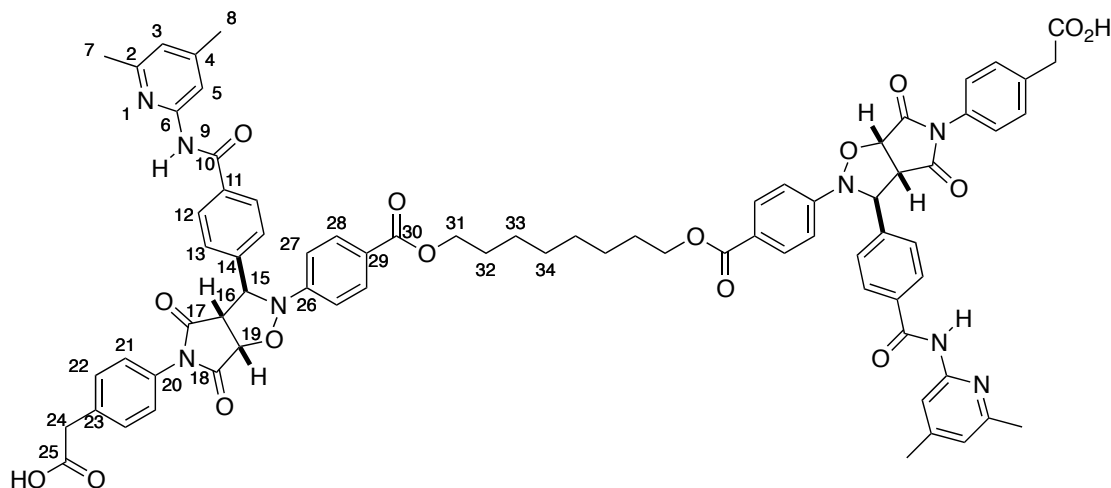
5.6 mmol, 52%) as a colourless crystalline solid; M.p. 101.2–102.7 °C; <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 8.27 (4H, d, <sup>3</sup>J<sub>HH</sub> = 9 Hz, 4×ArH), 8.18 (4H, d, <sup>3</sup>J<sub>HH</sub> = 9 Hz, 4×ArH), 4.36 (4H, t, <sup>3</sup>J<sub>HH</sub> = 6.6 Hz, 2×CH<sub>2</sub>), 1.79 (4H, m, 2×CH<sub>2</sub>) and 1.42 (8H, m, 4×CH<sub>2</sub>); <sup>13</sup>C NMR (100.6 MHz, d<sub>6</sub>-DMSO) δ = 164.8 (CO), 150.5 (ArC), 135.8 (ArC), 130.7 (ArCH), 123.5 (ArCH), 66.0 (OCH<sub>2</sub>), 29.2 (CH<sub>2</sub>), 28.6 (CH<sub>2</sub>) and 26.9 (CH<sub>2</sub>); MS (ES<sup>+</sup>) *m/z* 466.98 ([M+Na]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>22</sub>H<sub>24</sub>N<sub>2</sub>O<sub>8</sub>Na [M+Na]<sup>+</sup> 467.1430, found 467.1419.

**(*NZ,N'Z*)-4,4'-((octane-1,8-diylbis(oxy))bis(carbonyl))bis(*N*-(4-((4,6-dimethylpyridin-2-yl)carbamoyl)benzylidene)aniline oxide) (152)**



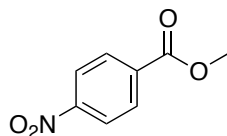
Dihydroxylamine **151** (0.9 g, 2.16 mmol, 1 eq) and aldehyde **89** (2.2 g, 8.64 mmol, 4 eq) were dissolved in ethanol (144 mL) with the aid of sonication and left to sit in darkness for 72 h. The precipitate which formed was isolated by filtration, washed with a little cold ethanol and dried to yield the bisnitronium (1.3 g, 1.5 mmol, 69%) as a colourless powder; M.p. decomposed > 180 °C; <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 8.74 (2H, s, 2×NH) 8.51 (4H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 4×ArH) 8.18 (2H, s, 2×ArH), 8.16 (2H, s, 2×CH) 8.05 (8H, m, 8×ArH) 7.86 (4H, d, <sup>3</sup>J<sub>HH</sub> = 8.7 Hz, 4×ArH), 6.80 (2H, s, 2×ArH) 4.36 (4H, t, <sup>3</sup>J<sub>HH</sub> = 6.5 Hz, 2×CH<sub>2</sub>), 2.46 (6H, s, 2×CH<sub>3</sub>), 2.38 (6H, s, 2×CH<sub>3</sub>), 1.80 (4H, m, 2×CH<sub>2</sub>) and 1.54–1.36 (8H, m, 4×CH<sub>2</sub>); <sup>13</sup>C NMR (100.6 MHz, d<sub>6</sub>-DMSO) δ = 165.4 (CO), 164.7 (CONH), 156.2 (ArC), 151.9 (ArC), 151.0 (ArC), 150.6 (ArC), 136.1 (ArC), 134.1 (CH), 133.7 (ArC), 132.3 (ArC), 130.9 (ArCH), 129.4 (ArCH), 127.8 (ArCH), 122.0 (ArCH), 121.0 (ArCH), 112.0 (ArCH), 65.8 (OCH<sub>2</sub>), 29.3 (CH<sub>2</sub>), 28.8 (CH<sub>2</sub>), 26.1 (CH<sub>2</sub>), 23.7 (CH<sub>3</sub>) and 21.6 (CH<sub>3</sub>); MS (ES<sup>+</sup>) *m/z* 910.63 ([M+Na]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>52</sub>H<sub>52</sub>N<sub>6</sub>O<sub>8</sub>Na [M+Na]<sup>+</sup> 911.3744, found 911.3727. CHN analysis calculated C 70.25% H 5.9% N 9.45% found C 69.4% H 5.38% N 9.03%.

**2,2'-(((3*R*,3*aS*,3'*R*,3*a'S*,6*aR*,6*a'R*)-2,2'-(((octane-1,8-diylbis(oxy))bis(carbonyl))bis(4,1-phenylene))bis(3-(4-((4,6-dimethylpyridin-2-yl)carbonyl)phenyl)-4,6-dioxotetrahydro-2*H*-pyrrolo[3,4-*d*]isoxazole-5,2(3*H*)-diyl))bis(4,1-phenylene)) diacetic acid (*trans*-158)**



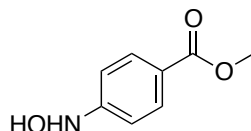
Bisnitron **152** (0.071 g, 0.08 mmol, 1 eq) and maleimide **35** (0.037 g, 0.16 mmol, 2 eq) were dissolved in CDCl<sub>3</sub>, cooled to 0 °C and reacted for 40 h before allowing the temperature to rise to RT and react for a further 6 h. The reaction was analysed by NMR and showed quantitative conversion to the *trans*-**158** product. The solvent was then removed to yield *trans*-**158** (0.108 g, 0.08 mmol, 99%) as a pale yellow solid; M.p. decomposed > 250 °C; <sup>1</sup>H NMR (400 MHz, d<sub>6</sub>-DMSO) δ = 10.64 (2H, s, 2×NH, H<sub>9</sub>), 8.05 (4H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 4×ArH, H<sub>28</sub>), 7.87 (2H, s, 2×ArH, H<sub>5</sub>), 7.82 (4H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 4×ArH, H<sub>27</sub>), 7.66 (4H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 4×ArH, H<sub>12</sub>), 7.31 (4H, d, <sup>3</sup>J<sub>HH</sub> = 8.7 Hz, 4×ArH, H<sub>21</sub>), 7.23 (4H, d, <sup>3</sup>J<sub>HH</sub> = 8.7 Hz, 4×ArH, H<sub>22</sub>), 6.87 (2H, s, 2×ArH, H<sub>3</sub>), 6.61 (4H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 4×ArH, H<sub>13</sub>), 6.07 (2H, s, 2×CH, H<sub>15</sub>), 5.51 (2H, d, <sup>3</sup>J<sub>HH</sub> = 7.4 Hz, 2×CH, H<sub>19</sub>), 4.35–4.16 (6H, m, 2×CH, H<sub>16</sub> and 2×CH<sub>2</sub>, H<sub>31</sub>), 3.57 (4H, s, 2×CH<sub>2</sub>, H<sub>24</sub>), 2.39 (6H, s, 2×CH<sub>3</sub>, H<sub>7</sub>), 2.30 (6H, s, 2×CH<sub>3</sub>, H<sub>8</sub>), 1.67 (4H, m, 2×CH<sub>2</sub>, H<sub>32</sub>) and 1.35 (8H, m, 4×CH<sub>2</sub>, H<sub>33</sub> and H<sub>34</sub>); <sup>13</sup>C NMR (100.6 MHz, d<sub>6</sub>-DMSO) δ = 174.1 (CO, C<sub>17</sub>), 173.0 (CO<sub>2</sub>H, C<sub>25</sub>), 172.4 (CO, C<sub>18</sub>), 165.3 (CO<sub>2</sub>R, C<sub>30</sub>), 165.2 (CONH, C<sub>10</sub>), 156.2 (ArC, C<sub>2</sub>), 152.3 (ArC, C<sub>29</sub>), 151.5 (ArC, C<sub>6</sub>), 149.0 (ArC, C<sub>4</sub>), 142.0 (ArC, C<sub>14</sub>), 135.9 (ArC, C<sub>23</sub>), 133.8 (ArC, C<sub>11</sub>), 130.4 (ArCH, C<sub>27</sub>), 130.0 (ArCH, C<sub>22</sub>), 129.9 (ArC, C<sub>26</sub>), 128.3 (ArCH, C<sub>28</sub>), 127.1 (ArCH, C<sub>12</sub>), 126.0 (ArCH, C<sub>13</sub>), 123.4 (ArC, C<sub>20</sub>), 120.1 (ArCH, C<sub>3</sub>), 114.0 (ArCH, C<sub>21</sub>), 112.3 (ArCH, C<sub>5</sub>), 77.8 (CH, C<sub>19</sub>), 67.8 (CH, C<sub>15</sub>), 64.4 (CH<sub>2</sub>, C<sub>31</sub>), 56.2 (CH, C<sub>16</sub>), 40.2 (CH<sub>2</sub>, C<sub>24</sub>), 29.0 (CH<sub>2</sub>, C<sub>34</sub>), 28.2 (CH<sub>2</sub>, C<sub>32</sub>), 25.4 (CH<sub>2</sub>, C<sub>33</sub>), 23.4 (CH<sub>3</sub>, C<sub>8</sub>) and 20.8 (CH<sub>3</sub>, C<sub>7</sub>); HRMS (MALDI) *m/z* calculated for C<sub>76</sub>H<sub>71</sub>N<sub>8</sub>O<sub>16</sub> [M+H]<sup>+</sup> 1351.4983, found 1351.4672

### Methyl 4-nitrobenzoate<sup>223</sup> (153)



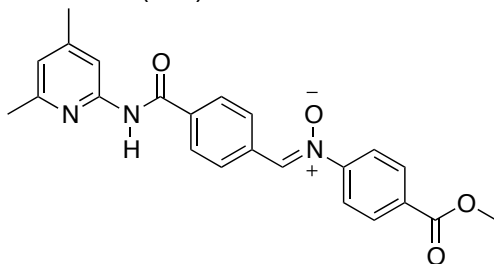
Methanol (6.9 g, 216 mmol, 10 eq) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (10 mL) and triethylamine (10 mL) and cooled to 0 °C under N<sub>2</sub>. A solution of 4-nitrobenzoyl chloride (4 g, 21.6 mmol, 1 eq) in CH<sub>2</sub>Cl<sub>2</sub> (50 mL) was then added drop wise and upon complete addition the reaction was allowed to rise to RT and monitored by TLC analysis. Upon completion, the reaction was quenched by addition of NaHCO<sub>3</sub> (aq) (100 mL), the organic fraction separated, dried (MgSO<sub>4</sub>), filtered and the solvent removed under vacuum to yield Methyl 4-nitrobenzoate (3.8 g, 21.2 mmol, 98%) as a colourless crystalline solid; M.p. 94.0–94.9 °C (lit<sup>223</sup> 94–96 °C); <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 8.24 (2H, d, <sup>3</sup>J<sub>HH</sub> = 9.1 Hz, 2×ArH) 8.15 (2H, d, <sup>3</sup>J<sub>HH</sub> = 9.1 Hz, 2×ArH) and 3.94 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ = 165.2 (CO<sub>2</sub>CH<sub>3</sub>), 150.5 (ArC), 135.5 (ArC), 130.7 (ArCH), 123.5 (ArCH) and 52.9 (CO<sub>2</sub>CH<sub>3</sub>); MS (EI<sup>+</sup>, *m/z*) 181.04 ([M], 10), 150.01 ([M–21]<sup>–</sup>, 100); HRMS (EI<sup>+</sup>) *m/z* calculated for C<sub>8</sub>H<sub>7</sub>NO<sub>4</sub> [M] 181.0375, found 181.0370.

### Methyl 4-(hydroxyamino)benzoate (154)



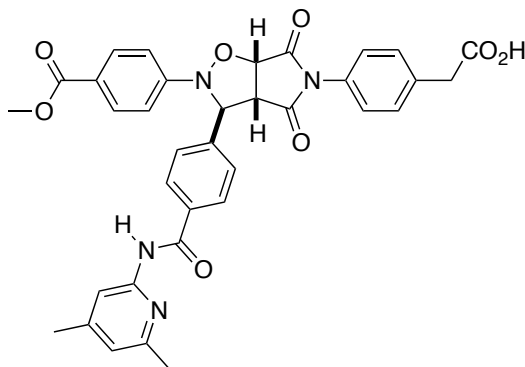
Methyl 4-nitrobenzoate **153** (0.5 g, 2.76 mmol, 1 eq) was dissolved in THF (10 mL) and stirred with Rh/C (0.2 g) for 15 min before addition of hydrazine hydrate (0.21 mL, 4.14 mmol, 1.5 eq). The reaction was monitored by TLC and upon completion the suspension was filtered through celite and the solvent removed to yield the hydroxylamine as a colourless crystalline solid (0.45 g, 2.70 mmol, 98%) which was used immediately without further purification; M.p. 118.2–119.1 °C (lit<sup>224</sup> 118–120 °C); <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 8.96 (1H, s, NH), 8.65 (1H, s, OH), 7.78 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.8 Hz, 2×ArH), 6.84 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.8 Hz, 2×ArH) and 3.77 (3H, s, OCH<sub>3</sub>); <sup>13</sup>C NMR (75.5 MHz, d<sub>6</sub>-DMSO) δ = 166.2 (CO), 156.0 (ArC), 130.4 (ArCH), 119.1 (ArC), 111.2 (ArCH) and 51.4 (OCH<sub>3</sub>).

**(Z)-N-(4-((4,6-dimethylpyridin-2-yl)carbamoyl)benzylidene)-4-(methoxycarbonyl)aniline oxide (155)**



Hydroxylamine **154** (0.1 g, 0.6 mmol, 1 eq) and aldehyde **89** (0.18 g, 0.72 mmol, 1.2 eq) were dissolved in ethanol (10 mL) and left to sit in darkness for 16 h. The precipitate which formed was isolated by filtration, washed with cold ethanol and dried to yield the nitronium (0.095g , 0.24 mmol, 39%) as a pale yellow crystalline solid; M.p. 189.0–191.7 °C;  $^1\text{H NMR}$  (400.1 MHz,  $\text{CDCl}_3$ )  $\delta$  = 8.80–8.70 (1H, br, NH), 8.48 (2H, d,  $^3J_{\text{HH}}$  = 8.4 Hz, 2×ArH), 8.16 (2H, d,  $^3J_{\text{HH}}$  = 8.8 Hz, 2×ArH), 8.20–8.05 (4H, m, 3×ArH and CH), 7.86 (2H, d,  $^3J_{\text{HH}}$  = 8.8 Hz, 2×ArH), 6.76 (1H, s, ArH), 3.94 (3H, s, OCH<sub>3</sub>), 2.39 (3H, s, ArCH<sub>3</sub>) and 2.34 (3H, s, ArCH<sub>3</sub>);  $^{13}\text{C NMR}$  (100.6 MHz,  $\text{CDCl}_3$ )  $\delta$  = 165.8 (CO<sub>2</sub>Me), 164.7 (CONH), 156.6 (ArC), 151.9 (ArC), 150.8 (ArC), 150.4 (ArC), 136.2 (ArC), 134.1 (CH), 133.6 (ArC), 131.9 (ArC), 130.9 (ArCH), 129.3 (ArCH), 127.7 (ArCH), 121.9 (ArCH), 120.9 (ArCH), 111.9 (ArCH), 52.6 (OCH<sub>3</sub>), 23.8 (CH<sub>3</sub>) and 21.4 (CH<sub>3</sub>); MS (ES<sup>+</sup>)  $m/z$  425.77 ([M+Na]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>)  $m/z$  calculated for C<sub>23</sub>H<sub>21</sub>N<sub>3</sub>O<sub>4</sub>Na [M+Na]<sup>+</sup> 426.1430, found 426.1431.

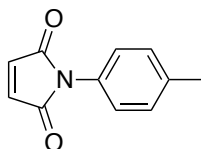
**2-(4-((3R,3aS,6aR)-3-(4-((4,6-dimethylpyridin-2-yl)carbamoyl)phenyl)-2-(4-(methoxycarbonyl)phenyl)-4,6-dioxotetrahydro-2H-pyrrolo[3,4-d]isoxazol-5(3H)-yl)phenyl)acetic acid (trans-156)**



Nitronium **155** (0.016 g, 0.04 mmol, 1 eq) and maleimide **34** (0.009 g, 0.04 mmol, 1 eq) were dissolved in  $\text{CHCl}_3$  (2 mL) cooled to 0 °C and allowed to react for 48 h. The solvent was then removed to give the product cycloadduct as a colourless powder

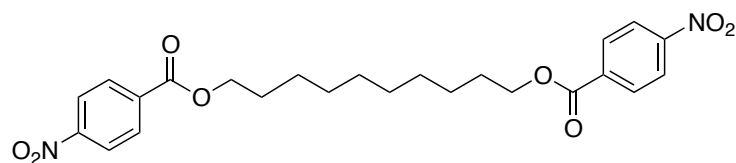
(0.025 g, 0.04 mmol, 99%) which was used without further purification. M.p. = 201.7–207 °C; <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 11.44–11.00 (1H, br, NH), 8.23 (1H, s, ArH), 7.92 (4H, m, 4×ArH), 7.33 (4H, m, 4×ArH), 7.19 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.2 Hz, 2×ArH), 6.81 (2H, s, 2×ArH), 6.78 (1H, s, ArH), 5.79 (1H, s, CH), 5.01 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.3 Hz, CH), 3.93 (1H, d, <sup>3</sup>J<sub>HH</sub> = 7.3 Hz, CH), 3.87 (3H, s, OCH<sub>3</sub>), 3.72 (2H, s, CH<sub>2</sub>), 2.46 (3H, s, CH<sub>3</sub>) and 2.40 (3H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ = 177.0 (CO<sub>2</sub>H), 174.5 (CO), 173.0 (CO), 166.6 (CO<sub>2</sub>Me), 166.5 (CONH), 154.9 (ArC), 152.8 (ArC), 152.3 (ArC), 151.4 (ArC), 141.6 (ArC), 136.0 (ArC), 134.1 (ArC), 131.3 (ArCH), 130.8 (ArCH), 129.8 (ArC), 129.1 (ArCH), 126.6 (ArCH), 125.8 (ArCH), 125.6 (ArC), 121.2 (ArCH), 114.0 (ArCH), 113.9 (ArCH), 77.4 (CH), 68.8 (CH), 57.1 (CH), 52.1 (CH<sub>3</sub>), 40.6 (CH<sub>2</sub>), 21.8 (2×CH<sub>3</sub>); MS (ES<sup>+</sup>) *m/z* 634.65 ([M+H]<sup>+</sup>, 100), HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>35</sub>H<sub>31</sub>N<sub>4</sub>O<sub>8</sub> [M+H]<sup>+</sup> 635.2142, found 635.2142.

**1-(*p*-Tolyl)-1*H*-pyrrole-2,5-dione<sup>225</sup> (160)**



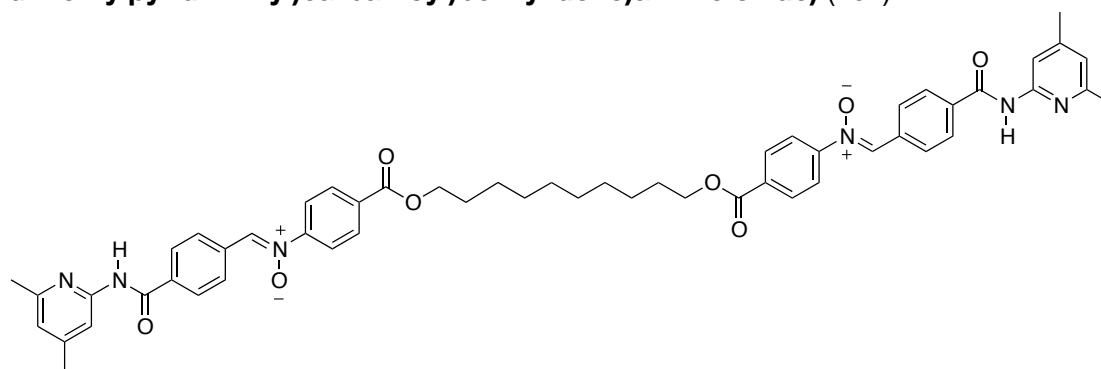
*p*-Toluidine (0.4 g, 3.76 mmol, 1 eq) and maleic anhydride (0.37 g, 3.76 mmol, 1 eq) were dissolved together in THF and stirred at RT for 16 h. The solvent was then removed under vacuum to yield the crude product as a colourless powder which was used immediately without purification. Zinc bromide (0.85 g, 3.76 mmol, 1 eq) and hexamethyldisilazane (4 ml, 18.8 mmol, 5 eq) were added to a solution of (*Z*)-4-oxo-4-(*p*-tolylamino)but-2-enoic acid (0.77 g, 3.76 mmol, 1 eq) in acetonitrile (40 mL). The reaction mixture was refluxed at 70 °C for an hour, before being quenched with water (50 mL). The organic layer was treated with conc. HCl (5 mL) and the aqueous layer was extracted with DCM (3 × 200 mL). The combined organic layers were dried over MgSO<sub>4</sub>, filtered and concentrated in vacuum to furnish a yellow solid (0.5 g, 2.67 mmol, 71%); M.p. 146.0–148.2 °C (lit<sup>225</sup> 158–160 °C); <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 7.27 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.4 Hz, 2×ArH), 7.21 (2H, d, <sup>3</sup>J<sub>HH</sub> = 8.4 Hz, 2×ArH), 6.85 (2H, s, 2×CH), 2.38 (3H, s, 3H); <sup>13</sup>C NMR (75.5 MHz, CDCl<sub>3</sub>) δ = 169.7 (CO), 138.1 (ArC), 134.2 (CH), 129.8 (ArCH), 128.5 (ArC), 126.1 (ArCH), 21.2 (CH<sub>3</sub>); MS (CI<sup>+</sup>) *m/z* = 188 ([M+H]<sup>+</sup>, 100); HRMS (CI<sup>+</sup>) *m/z* calculated for C<sub>11</sub>H<sub>10</sub>NO<sub>2</sub> [M+H]<sup>+</sup> 188.0712, found 188.0715.

### Decane-1,10-diyl bis(4-nitrobenzoate) (162)



1,10-Decanediol (2.96 g, 17 mmol, 1 eq) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (10 mL) and triethylamine (10 mL) and cooled to 0 °C under N<sub>2</sub>. A solution of 4-nitrobenzoyl chloride (12.6 g, 67.7 mmol, 4 eq) in CH<sub>2</sub>Cl<sub>2</sub> (90 mL) was added dropwise and after complete addition the reaction was allowed to rise to RT and the reaction progress monitored by TLC analysis. Upon completion, Na<sub>2</sub>CO<sub>3</sub> (aq) (100 mL) was added and the solution stirred for 1 h. The organic layer was separated, dried (MgSO<sub>4</sub>), filtered and the solvent removed to give the crude product which was purified by recrystallisation from chloroform and hexane to yield the product (4.33 g, 9.16 mmol, 60%) as a colourless crystalline solid; M.p. 89.2–89.7 °C; <sup>1</sup>H NMR (400.1 MHz, CDCl<sub>3</sub>) δ = 8.29 (4H, d, <sup>3</sup>J<sub>HH</sub> = 9.1 Hz, 4×ArH), 8.20 (4H, d, <sup>3</sup>J<sub>HH</sub> = 9.1 Hz, 4×ArH), 4.39 (4H, t, <sup>3</sup>J<sub>HH</sub> = 6.6 Hz, 2×CH<sub>2</sub>), 1.80–1.62 (4H, m, 2×CH<sub>2</sub>) and 1.55 (12H, m, 6×CH<sub>2</sub>); <sup>13</sup>C NMR (75.5 MHz, CDCl<sub>3</sub>) δ = 164.7 (CO), 150.5 (ArC), 135.9 (ArC), 130.7 (ArCH), 123.5 (ArCH), 66.1 (OCH<sub>2</sub>), 29.5 (CH<sub>2</sub>), 29.2 (CH<sub>2</sub>), 28.6 (CH<sub>2</sub>) and 26.0 (CH<sub>2</sub>); MS (ES<sup>+</sup>) *m/z* 495.10 ([M+Na]<sup>+</sup>, 100); HRMS (ES<sup>+</sup>) *m/z* calculated for C<sub>24</sub>H<sub>28</sub>N<sub>2</sub>O<sub>8</sub>Na [M+Na]<sup>+</sup> 495.1743 found 495.1742.

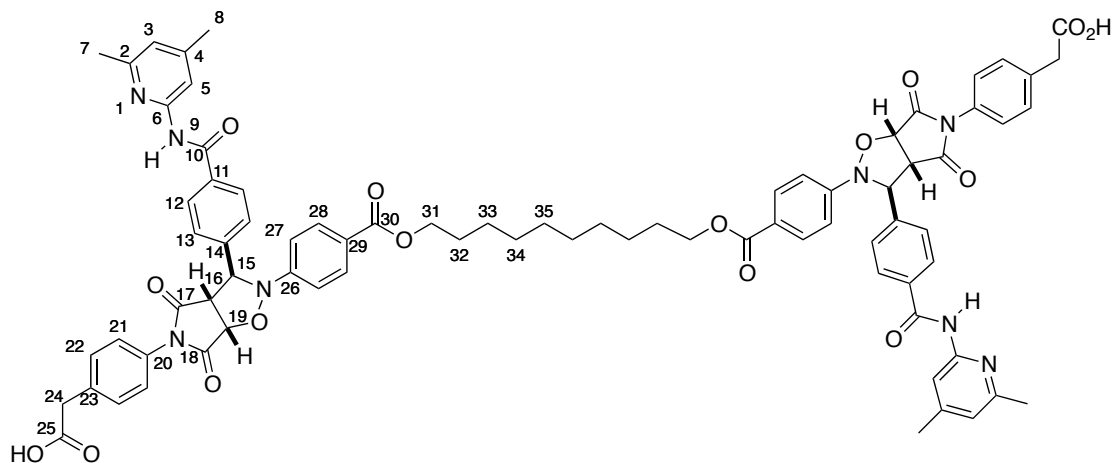
### (*NZ,N'Z*)-4,4'-((decane-1,10-diylbis(oxy))bis(carbonyl))bis(N-(4-((4,6-dimethylpyridin-2-yl)carbamoyl)benzylidene)aniline oxide) (164)



Aldehyde **89** (1.83 g, 7.2 mmol, 3 eq) was dissolved in ethanol (160 mL) with the aid of gentle heating and sonication. Dihydroxylamine **163** (0.93 g, 2.1 mmol, 1 eq) was then added and the reaction left to sit in darkness for 72 h. The precipitate which formed was isolated by filtration, washed with a little cold ethanol and dried to yield

the bisnitron (1.3 g, 1.5 mmol, 63%) as a colourless powder; M.p. 196–198.3 °C; <sup>1</sup>H NMR (300.1 MHz, CDCl<sub>3</sub>) δ = 9.00–8.75 (2H, s, 2×NH) 8.48 (4H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 4×ArH) 8.16 (4H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 4×ArH), 8.06–7.98 (8H, m, 6×ArH and 2×CH) 7.86 (4H, d, <sup>3</sup>J<sub>HH</sub> = 8.7 Hz, 4×ArH), 6.77 (2H, s, 2×ArH) 4.34 (4H, t, <sup>3</sup>J<sub>HH</sub> = 6.5 Hz, 2×CH<sub>2</sub>), 2.40 (6H, s, 2×CH<sub>3</sub>), 2.35 (6H, s, 2×CH<sub>3</sub>), 1.83–1.71 (4H, m, 2×CH<sub>2</sub>) and 1.49–1.27 (12H, m, 6×CH<sub>2</sub>); <sup>13</sup>C NMR (75.5 MHz, CDCl<sub>3</sub>) δ = 165.7 (CO), 165.1 (CONH), 156.8 (ArC), 152.1 (ArC), 151.1 (ArC), 150.8 (ArC), 136.5 (ArC), 134.4 (CH), 133.9 (ArC), 132.6 (ArC), 131.1 (ArCH), 129.6 (ArCH), 128.1 (ArCH), 122.2 (ArCH), 121.3 (ArCH), 112.3 (ArCH), 66.1 (OCH<sub>2</sub>), 29.8 (CH<sub>2</sub>), 29.6 (CH<sub>2</sub>), 29.1 (CH<sub>2</sub>), 26.4 (CH<sub>2</sub>) 24.1 (CH<sub>3</sub>) and 21.8 (CH<sub>3</sub>); HRMS (MALDI) *m/z* calculated for C<sub>54</sub>H<sub>56</sub>N<sub>6</sub>O<sub>8</sub> [M+H]<sup>+</sup> 917.4238, found 917.4247.

**2,2'-(((3R,3aS,3'R,3a'S,6aR,6a'R)-2,2'-(((decane-1,10-diylbis(oxy))bis(carbonyl))bis(4,1-phenylene))bis(3-(4-((4,6-dimethylpyridin-2-yl)carbamoyl)phenyl)-4,6-dioxotetrahydro-2H-pyrrolo[3,4-d]isoxazole-5,2(3H)-diyl))bis(4,1-phenylene)) diacetic acid (*trans*-165)**



Bisnitron **164** (0.069 g, 0.08 mmol, 1 eq) and maleimide **34** (0.037 g, 0.16 mmol, 2 eq) were dissolved in CDCl<sub>3</sub> (4 mL), cooled to 0 °C and reacted for 40 h before allowing the temperature to rise to RT and react for a further 6 h. The reaction was analysed by NMR and showed quantitative conversion to the *trans*-**165** product. The solvent was then removed to yield *trans*-**165** (0.106 g, 0.08 mmol, 99%) as a pale yellow solid; M.p. decomposed > 250 °C; <sup>1</sup>H NMR (400.1 MHz, d<sub>6</sub>-DMSO) δ = 12.88–12.17 (2H, br, 2×CO<sub>2</sub>H) 10.64 (2H, s, 2×NH), 8.07 (4H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 4×ArH), 7.89 (2H, s, 2×ArH), 7.83 (4H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 4×ArH), 7.67 (4H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 4×ArH), 7.33 (4H, d, <sup>3</sup>J<sub>HH</sub> = 8.7 Hz, 4×ArH), 7.25 (4H, d, <sup>3</sup>J<sub>HH</sub> = 8.7 Hz, 4×ArH), 6.85 (2H, s, 2×ArH), 6.64 (4H, d, <sup>3</sup>J<sub>HH</sub> = 8.5 Hz, 4×ArH), 6.08 (2H, s, 2×CH),



5.53 (2H, d,  $^3J_{\text{HH}} = 7.4$  Hz, 2×CH), 4.34–4.14 (6H, m, 2×CH and 2 ×CH<sub>2</sub>), 3.59 (4H, s, 2×CH<sub>2</sub>), 2.39 (6H, s, 2×CH<sub>3</sub>, H<sub>7</sub>), 2.29 (6H, s, 2×CH<sub>3</sub>), 1.77–1.57 (4H, m, 2×CH<sub>2</sub>) and 1.44–1.15 (12H, m, 6×CH<sub>2</sub>); <sup>13</sup>C NMR (100.6 MHz, d<sub>6</sub>-DMSO) δ = 174.1 (CO, C<sub>17</sub>), 173.0 (CO<sub>2</sub>H, C<sub>25</sub>), 172.3 (CO, C<sub>18</sub>), 165.3 (CO<sub>2</sub>R, C<sub>30</sub>), 165.2 (CONH, C<sub>10</sub>), 156.1 (ArC, C<sub>2</sub>), 152.3 (ArC, C<sub>29</sub>), 151.6 (ArC, C<sub>6</sub>), 149.0 (ArC, C<sub>4</sub>), 142.1 (ArC, C<sub>14</sub>), 135.8 (ArC, C<sub>23</sub>), 133.8 (ArC, C<sub>11</sub>), 130.5 (ArCH, C<sub>27</sub>), 129.9 (ArCH, C<sub>22</sub>), 129.9 (ArC, C<sub>26</sub>), 128.4 (ArCH, C<sub>28</sub>), 127.1 (ArCH, C<sub>12</sub>), 126.0 (ArCH, C<sub>13</sub>), 123.5 (ArC, C<sub>20</sub>), 120.1 (ArCH, C<sub>3</sub>), 114.0 (ArCH, C<sub>21</sub>), 112.3 (ArCH, C<sub>5</sub>), 77.8 (CH, C<sub>19</sub>), 67.9 (CH, C<sub>15</sub>) 64.4 (CH<sub>2</sub>, C<sub>31</sub>), 56.3 (CH, C<sub>16</sub>), 40.1 (CH<sub>2</sub>, C<sub>24</sub>), 28.9 (CH<sub>2</sub>, C<sub>34</sub>), 28.7 (CH<sub>2</sub>, C<sub>35</sub>) 28.2 (CH<sub>2</sub>, C<sub>32</sub>), 25.5 (CH<sub>2</sub>, C<sub>33</sub>), 23.4 (CH<sub>3</sub>, C<sub>8</sub>) and 20.9 (CH<sub>3</sub>, C<sub>7</sub>); HRMS (MALDI) *m/z* calculated for C<sub>80</sub>H<sub>79</sub>N<sub>8</sub>O<sub>16</sub> [M+H]<sup>+</sup> 1378.5301, found 1379.5258.

# 8

## References

1. Nicolaou, K. C.; Yang, Z.; Liu, J. J.; Ueno, H.; Nantermet, P. G.; Guy, R. K.; Claiborne, C. F.; Renaud, J.; Couladouros, E. A.; Paulvannan, K.; Sorensen, E. J. *Nature* **1994**, *367*, 630-634.
2. Nicolaou, K. C.; Sorensen, E. J. *Classics in Total Synthesis*; VCH Publishers, Inc.: New York, **1996**.
3. Whitesides, G. M.; Boncheva, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4769-4774.
4. Reinhoudt, D. N.; Crego-Calama, M. *Science* **2002**, *295*, 2403-2407.
5. Desiraju, G. R. *Nature* **2001**, *412*, 397-400.
6. Menger, F. M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4818-4822.
7. Fang, L.; Olson, M. A.; Benitez, D.; Tkatchouk, E.; Goddard, W. A., 3rd; Stoddart, J. F. *Chem. Soc. Rev.* **2010**, *39*, 17-29.
8. de Silva, A. P.; Fox, D. B.; Moody, T. S. In *Stimulating Concepts in Chemistry*, 1st ed.; Stoddart, J. F., Vögtle, F., Shibasaki, M., Eds.; Wiley-VCH: Weinheim, 2000; Vol. 1, p 307-316.
9. Fabbrizzi, L.; Poggi, A. *Chem. Soc. Rev.* **1995**, *24*, 197-202.
10. Breslow, R. *Artificial Enzymes*; Weiley-VCH, **2005**.
11. Motherwell, W. B.; Bingham, M. J.; Six, Y. *Tetrahedron* **2001**, *57*, 4663-4686.
12. Breslow, R.; Zhang, B. *J. Am. Chem. Soc.* **1992**, *114*, 5882-5883.
13. Kirby, A. J. *Enzyme Mimics*; 1st ed.; Wiley-VCH: Weinheim, **2000**.
14. Anslyn, E. V.; Dougherty, D. A. *Modern Physical Organic Chemistry*, 1st ed.; University Science Books, **2006**.
15. Prins, L. J.; Reinhoudt, D. N.; Timmerman, P. *Angew. Chem. Int. Ed.* **2001**, *40*, 2382-2426.
16. Beijer, F. H.; Kooijman, H.; Spek, A. L.; Sijbesma, R. P.; Meijer, E. W. *Angew. Chem. Int. Ed.* **1998**, *37*, 75-78.
17. Beijer, F. H.; Sijbesma, R. P.; Kooijman, H.; Spek, A. L.; Meijer, E. W. *J. Am. Chem. Soc.* **1998**, *120*, 6761-6769.
18. Bell, D. A.; Díaz, S. G.; Lynch, V. M.; Anslyn, E. V. *Tetrahedron. Lett.* **1995**, *36*, 4155-4158.
19. Blight, B. A.; Camara-Campos, A.; Djurdjevic, S.; Kaller, M.; Leigh, D. A.; McMillan, F. M.; McNab, H.; Slawin, A. M. Z. *J. Am. Chem. Soc.* **2009**, *131*, 14116-14122.
20. Boucher, E.; Simard, M.; Wuest, J. D. *J. Org. Chem.* **1995**, *60*, 1408-1412.
21. Ducharme, Y.; Wuest, J. D. *J. Org. Chem.* **1988**, *53*, 5787-5789.
22. Hunter, C. A. *Angew. Chem. Int. Ed.* **2004**, *43*, 5310-5324.
23. Zimmerman, S. C.; Murray, T. J. *Tetrahedron. Lett.* **1994**, *35*, 4077-4080.

24. Autumn, K.; Sitti, M.; Liang, Y. A.; Peattie, A. M.; Hansen, W. R.; Sponberg, S.; Kenny, T. W.; Fearing, R.; Israelachvili, J. N.; Full, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12252-12256.
25. Murthy, P. S. *J. Chem. Ed.* **2006**, *83*, 1010-1013.
26. Villalonga, R.; Cao, R.; Fragoso, A. *Chem. Rev.* **2007**, *107*, 3088-3116.
27. Pedersen, C. J. *J. Am. Chem. Soc.* **1967**, *89*, 7017-7036.
28. Pedersen, C. J. *Science* **1988**, *241*, 536-540.
29. Cram, D. J. *Angew. Chem. Int. Ed. Engl.* **1988**, *27*, 1009-1020.
30. Lehn, J.-M. *Angew. Chem. Int. Ed. Engl.* **1988**, *27*, 89-112.
31. Hunter, C. A.; Anderson, H. L. *Angew. Chem. Int. Ed.* **2009**, *48*, 7488-7499.
32. Kyogoku, Y.; Rich, A.; Lord, R. C. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *57*, 258-264.
33. Pranata, J.; Wierschke, S. G.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 2810-2819.
34. Jorgensen, W. L.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008-2010.
35. Murray, T. J.; Zimmerman, S. C. *J. Am. Chem. Soc.* **1992**, *114*, 4010-4011.
36. Sartorius, J.; Schneider, H.-J. *Chem. Eur. J.* **1996**, *2*, 1446-1452.
37. Djurdjevic, S.; Leigh, D. A.; McNab, H.; Parsons, S.; Teobaldi, G.; Zerbetto, F. *J. Am. Chem. Soc.* **2006**, *129*, 476-477.
38. Sijbesma, R. P.; Meijer, E. W. *Chem. Commun.* **2003**, 5-16.
39. Park, T.; Todd, E. M.; Nakashima, S.; Zimmerman, S. C. *J. Am. Chem. Soc.* **2005**, *127*, 18133-18142.
40. Park, T.; Zimmerman, S. C.; Nakashima, S. *J. Am. Chem. Soc.* **2005**, *127*, 6520-6521.
41. Todd, E. M.; Quinn, J. R.; Park, T.; Zimmerman, S. C. *Israel J Chem* **2005**, *45*, 381-389.
42. Vaillancourt, L.; Simard, M.; Wuest, J. D. *J. Org. Chem.* **1998**, *63*, 9746-9752.
43. Gallant, M.; Phan Viet Minh, T.; Wuest, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 721-723.
44. Garcia-Tellado, F.; Goswami, S.; Chang, S. K.; Geib, S. J.; Hamilton, A. D. *J. Am. Chem. Soc.* **1990**, *112*, 7393-7394.
45. Corbin, P. S.; Zimmerman, S. C. *J. Am. Chem. Soc.* **2000**, *122*, 3779-3780.
46. Fischer, E. *Ber. Dtsch. Chem. Ges.* **1894**, *27*, 2985-2993.
47. Leskovac, V. *Comprehensive Enzyme Kintics*; Kluwer Academic / Plenum Publishers: New York, **2003**.
48. Heppel, L. A.; Hilmoie, R. J. *J. Biol. Chem.* **1951**, *188*, 665-676.
49. Koshland, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1958**, *44*, 98-104.
50. Page, M. I. *Enzyme Mechanisms*; 1st ed.; Royal Society of Chemistry: London, **1987**.
51. Menger, F. M. *Biochemistry* **1992**, *31*, 5368-5373.
52. Pauling, L. *Nature* **1948**, *161*, 707-709.
53. Kirby, A. J. *Advances in Physical Organic Chemistry*; Academic Press: London, **1980**; Vol. 17.
54. Page, M. I. *Chem. Soc. Rev.* **1973**, *2*, 295-323.
55. Page, M. I.; Jencks, W. P. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 1678-1683.
56. Krishnamurthy, V. M.; Semetey, V.; Bracher, P. J.; Shen, N.; Whitesides, G. M. *J. Am. Chem. Soc.* **2007**, *129*, 1312-1320.
57. Cacciapaglia, R.; Di Stefano, S.; Mandolini, L. *Acc. Chem. Res.* **2004**, *37*, 113-122.
58. Bruice, T. C.; Pandit, U. K. *J. Am. Chem. Soc.* **1960**, *82*, 5858-5865.
59. Bruice, T. C.; Lightstone, F. C. *Acc. Chem. Res.* **1998**, *32*, 127-136.
60. Storm, D. R.; Koshland, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1970**, *66*, 445-452.

61. Koshland, D. E. *J. Theor. Biol.* **1962**, *2*, 75.
62. Jencks, W. P. *Catalysis in Chemistry and Enzymology*; McGraw-Hill: New York, **1969**; Vol. 16.
63. Menger, F. M. *Acc. Chem. Res.* **1985**, *18*, 128-134.
64. Lightstone, F. C.; Bruice, T. C. *J. Am. Chem. Soc.* **1996**, *118*, 2595-2605.
65. Anderson, S.; Anderson, H. L. In *Templated Organic Synthesis*; Stang, P. J., Diederich, F., Eds.; Wiley-VCH Verlag GmbH: 2000, p 1-38.
66. Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737-738.
67. Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 964-967.
68. Busch, D. H.; Stephenson, N. A. *Coord. Chem. Rev.* **1990**, *100*, 119-154.
69. Seidel, F. *Chem. Ber.* **1926**, *59B*, 1894-1908.
70. Melson, G. A.; Busch, D. H. *Proc. Chem. Soc., London*, **1963**, 223-224.
71. Busch, D. H. *Rec. Chem. Prog.* **1964**, *25*, 106-126.
72. Busch, D. H.; Burke Jnr, J. A.; Jicha, D. C.; Thompson, M. C.; Morris, M. L. *Chem. Eng. News.* **1962**, *40*, 57.
73. Busch, D. H.; Burke Jnr, J. A.; Jicha, D. C.; Thompson, M. C.; Morris, M. L. *Adv. Chem. Ser.* **1962**, *37*, 125-142.
74. Busch, D. H. *Adv. Chem. Ser.* **1963**, *37*, 1-18.
75. Anderson, S.; Anderson, H. L.; Sanders, J. K. M. *Acc. Chem. Res.* **1993**, *26*, 469-475.
76. Sauvage, J. P.; Dietrich-Buchecker, C. O.; Kern, J. M. *J. Am. Chem. Soc.* **1984**, *106*, 3043-3045.
77. Dietrich-Buchecker, C. O.; Sauvage, J. P.; Kern, J. M. *J. Am. Chem. Soc.* **1984**, *106*, 3043-3045.
78. Aucagne, V.; José, B.; Crowley, J. D.; Goldup, S. M.; Hänni, K. D.; Leigh, D. A.; Lusby, P. J.; Ronaldson, V. E.; Slawin, A. M. Z.; Viterisi, A.; Walker, D. B. *J. Am. Chem. Soc.* **2007**, *129*, 11950-11963.
79. Goldup, S. M.; Leigh, D. A.; Long, T.; McGonigal, P. R.; Symes, M. D.; Wu, J. *J. Am. Chem. Soc.* **2009**, *131*, 15924-15929.
80. Crowley, J. D.; Hänni, K. D.; Leigh, D. A.; Slawin, A. M. Z. *J. Am. Chem. Soc.* **2010**, *132*, 5309-5314.
81. Vickers, M. S.; Beer, P. D. *Chem. Soc. Rev.* **2007**, *36*, 211-225.
82. Dichtel, W. R.; Miljanic, O. S.; Zhang, W.; Spruell, J. M.; Patel, K.; Aprahamian, I.; Heath, J. R.; Stoddart, J. F. *Acc. Chem. Res.* **2008**, *41*, 1750-1761.
83. Anelli, P. L.; Ashton, P. R.; Ballardini, R.; Balzani, V.; Delgado, M.; Gandolfi, M. T.; Goodnow, T. T.; Kaifer, A. E.; Philp, D.; Pietraszkiewicz, M.; Prodi, L.; Reddington, M. V.; Slawin, A. M. Z.; Spencer, N.; Stoddart, J. F.; Vicent, C.; Williams, D. E. *J. Am. Chem. Soc.* **1992**, *114*, 193-218.
84. Ashton, P. R.; Goodnow, T. T.; Kaifer, A. E.; Reddington, M. V.; Slawin, A. M. Z.; Spencer, N.; Stoddart, J. F.; Vicent, C.; Williams, D. J. *Angew. Chem. Int. Ed. Engl.* **1989**, *28*, 1396-1399.
85. Kelly, T. R.; Zhao, C.; Bridger, G. J. *J. Am. Chem. Soc.* **1989**, *111*, 3744-3745.
86. Kelly, T. R.; Bridger, G. J.; Zhao, C. *J. Am. Chem. Soc.* **1990**, *112*, 8024-8034.
87. Orgel, L. E. *Nature* **1992**, *358*, 203-209.
88. Blackmond, D. G. *Angew. Chem. Int. Ed.* **2009**, *48*, 386-390.
89. Weber, A. *Orig. Life. Evol. B.* **2007**, *37*, 105-111.
90. Wächtershäuser, G. *Chem. Biodivers.* **2007**, *4*, 584-602.
91. von Kiedrowski, G. *Angew. Chem. Int. Ed. Engl.* **1986**, *25*, 932-935.
92. von Kiedrowski, G.; Wlotzka, B.; Helbing, J.; Matzen, M.; Jordan, S. *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 423-426.

93. von Kiedrowski, G.; Wlotzka, B.; Helbing, J. *Angew. Chem. Int. Ed. Engl.* **1989**, *28*, 1235-1237.
94. Kiedrowski, G. V. *Bioorg. Chem. Front.* **1993**, 113-146.
95. Vidonne, A.; Philp, D. *Eur. J. Org. Chem.* **2009**, *2009*, 593-610.
96. Severin, K.; Lee, D. H.; Martinez, J. A.; Ghadiri, M. R. *Chem. Eur. J.* **1997**, *3*, 1017-1024.
97. Lee, D. H.; Granja, J. R.; Martinez, J. A.; Severin, K.; Ghadiri, M. R. *Nature* **1996**, *382*, 525-528.
98. O'Shea, E.; Klemm, J.; Kim, P.; Alber, T. *Science* **1991**, *254*, 539-544.
99. Dawson, P.; Muir, T.; Clark-Lewis, I.; Kent, S. *Science* **1994**, *266*, 776-779.
100. Yao, S.; Ghosh, I.; Zutshi, R.; Chmielewski, J. *J. Am. Chem. Soc.* **1997**, *119*, 10559-10560.
101. Yao, S.; Ghosh, I.; Zutshi, R.; Chmielewski, J. *Angew. Chem. Int. Ed.* **1998**, *37*, 478-481.
102. Issac, R.; Chmielewski, J. *J. Am. Chem. Soc.* **2002**, *124*, 6808-6809.
103. Li, X.; Chmielewski, J. *J. Am. Chem. Soc.* **2003**, *125*, 11820-11821.
104. Gilbert, W. *Nature* **1986**, *319*, 618-618.
105. Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R. *Cell* **1982**, *31*, 147-157.
106. Joyce, G. F. *Nature* **2002**, *418*, 214-221.
107. Paul, N.; Joyce, G. F. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12733-12740.
108. Rogers, J.; Joyce, G. F. *RNA* **2001**, *7*, 395-404.
109. Voytek, S. B.; Joyce, G. F. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 15288-15293.
110. Joyce, G. F. *Angew. Chem. Int. Ed.* **2007**, *46*, 6420-6436.
111. Kim, D. E.; Joyce, G. F. *Chem. Biol.* **2004**, *11*, 1505-1512.
112. Lincoln, T. A.; Joyce, G. F. *Science* **2009**, *323*, 1229-1232.
113. Askew, B.; Ballester, P.; Buhr, C.; Jeong, K. S.; Jones, S.; Parris, K.; Williams, K.; Rebek, J. *J. Am. Chem. Soc.* **1989**, *111*, 1082-1090.
114. Nowick, J. S.; Feng, Q.; Tjivikua, T.; Ballester, P.; Rebek, J. *J. Am. Chem. Soc.* **1991**, *113*, 8831-8839.
115. Menger, F. M.; Eliseev, A. V.; Khanjin, N. A. *J. Am. Chem. Soc.* **1994**, *116*, 3613-3614.
116. Conn, M. M.; Wintner, E. A.; Rebek, J. *J. Am. Chem. Soc.* **1994**, *116*, 8823-8824.
117. Menger, F. M.; Eliseev, A. V.; Khanjin, N. A.; Sherrod, M. J. *J. Org. Chem.* **1995**, *60*, 2870-2878.
118. Wintner, E. A.; Tsao, B.; Rebek, J. *J. Org. Chem.* **1995**, *60*, 7997-8001.
119. Reinhoudt, D. N.; Rudkevich, D. M.; de Jong, F. *J. Am. Chem. Soc.* **1996**, *118*, 6880-6889.
120. Rotello, V.; Hong, J. I.; Rebek, J. *J. Am. Chem. Soc.* **1991**, *113*, 9422-9423.
121. Park, T. K.; Feng, Q.; Rebek, J. *J. Am. Chem. Soc.* **1992**, *114*, 4529-4532.
122. Kamioka, S.; Ajami, D.; Rebek, J. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 541-544.
123. Kamioka, S.; Ajami, D.; Rebek, J. *Chem. Commun.* **2009**, 7324-7326.
124. Ouellet, S. G.; Walji, A. M.; Macmillan, D. W. C. *Acc. Chem. Res.* **2007**, *40*, 1327-1339.
125. Terfort, A.; von Kiedrowski, G. *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 654-656.
126. Wang, B.; Sutherland, I. O. *Chem. Commun.* **1997**, *16*, 1495-1496.
127. Kindermann, M.; Stahl, I.; Reimold, M.; Pankau, W. M.; von Kiedrowski, G. *Angew. Chem. Int. Ed.* **2005**, *117*, 6908-6913.

128. Bennes, R. I.; Philp, D.; Spencer, N.; Kariuki, B. M.; Harris, K. D. M. *Org. Lett.* **1999**, *1*, 1087-1090.
129. Booth, C. A.; Philp, D. *Tetrahedron. Lett.* **1998**, *39*, 6987-6990.
130. Pearson, R. J.; Kassianidis, E.; Philp, D. *Tetrahedron. Lett.* **2004**, *45*, 4777-4780.
131. Pearson, R. J.; Kassianidis, E.; Slawin, A. M. Z.; Philp, D. *Org. Biomol. Chem.* **2004**, *2*, 3434-3441.
132. Philp, D.; Robertson, A. *Chem. Commun.* **1998**, 879-880.
133. Robertson, A.; Philp, D.; Spencer, N. *Tetrahedron* **1999**, *55*, 11365-11384.
134. Domingo, L. R.; Aurell, M. J.; Arnó, M.; Sáez, J. *J. Org. Chem.* **2007**, *72*, 4220-4227.
135. Quayle, J. M.; Slawin, A. M. Z.; Philp, D. *Tetrahedron. Lett.* **2002**, *43*, 7229-7233.
136. Allen, V. C.; Philp, D.; Spencer, N. *Org. Lett.* **2001**, *3*, 777-780.
137. Pearson, R. J.; Kassianidis, E.; Slawin, A. M. Z.; Philp, D. *Chem. Eur. J.* **2006**, *12*, 6829-6840.
138. Kassianidis, E.; Pearson, R. J.; Philp, D. *Chem. Eur. J.* **2006**, *12*, 8798-8812.
139. Kassianidis, E.; Philp, D. *Angew. Chem. Int. Ed.* **2006**, *45*, 6344-6348.
140. del Amo, V.; Slawin, A. M. Z.; Philp, D. *Org. Lett.* **2008**, *10*, 4589-4592.
141. Achilles, T.; von Kiedrowski, G. *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 1198-1201.
142. Stahl, I.; von Kiedrowski, G. *J. Am. Chem. Soc.* **2006**, *128*, 14014-14015.
143. Wilcox, C. *Frontiers in Supramolecular Chemistry and Photochemistry*; VCH: Weinheim: New York, Basel, Cambridge, **1991**.
144. Santos, M. A.; Afonso, A.; Marques, M. M.; Wilcox, C. *Supramol. Chem.* **2000**, *11*, 201 - 215.
145. Eigen, M. *Naturwissenschaften* **1971**, *58*, 465-523.
146. Severin, K.; Lee, D. H.; Martinez, J. A.; Vieth, M.; Ghadiri, M. R. *Angew. Chem. Int. Ed.* **1998**, *37*, 126-128.
147. Lee, D. H.; Severin, K.; Yokobayashi, Y.; Ghadiri, M. R. *Nature* **1997**, *390*, 591-594.
148. Lee, D. H.; Severin, K.; Yokobayashi, Y.; Ghadiri, M. R. *Nature* **1998**, *394*, 101-101.
149. Darwin, C.; *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*; 5th ed.; John Murray: London, **1890**.
150. Bada, J. *Nature* **1995**, *374*, 594-595.
151. Saghatelian, A.; Yokobayashi, Y.; Soltani, K.; Ghadiri, M. R. *Nature* **2001**, *409*, 797-801.
152. del Amo, V.; Philp, D. *Chem. Eur. J.* **2010**, *16*, 13304-13318.
153. Xu, S.; Giuseppone, N. *J. Am. Chem. Soc.* **2008**, *130*, 1826-1827.
154. Turega, S. M.; Lorenz, C.; Sadownik, J. W.; Philp, D. *Chem. Commun.* **2008**, 4076-4078.
155. Sadownik, J. W.; Philp, D. *Angew. Chem. Int. Ed.* **2008**, *47*, 9965-9970.
156. Sadownik, J. W., University of St. Andrews, 2009.
157. Pasteur, L. C. R. *Séanc. Acad. Sci. Paris* **1848**, *26*, 535.
158. Noorduyn, W. L.; Izumi, T.; Millemaggi, A.; Leeman, M.; Meekes, H.; Van Enckevort, W. J.; Kellogg, R. M.; Kaptein, B.; Vlieg, E.; Blackmond, D. G. *J Am Chem Soc* **2008**, *130*, 1158-1159.
159. Ostwald, W. *Lehrbuch der Allgemeinen Chemie*; Engelmann: Leipzig, Germany **1896**, Vol 2, part 1.

160. Carnall, J. M. A.; Waudby, C. A.; Belenguer, A. M.; Stuart, M. C. A.; Peyralans, J. J. P.; Otto, S. *Science* **2010**, *327*, 1502-1506.
161. Prasanna de Silva, A.; McClenaghan, N. D. *J. Am. Chem. Soc.* **2000**, *122*, 3965-3966.
162. de Silva, A. P.; Uchiyama, S. *Nature Nanotech.* **2007**, *2*, 399-410.
163. Credi, A.; Balzani, V.; Langford, S. J.; Stoddart, J. F. *J. Am. Chem. Soc.* **1997**, *119*, 2679-2681.
164. Brown, G. J.; de Silva, A. P.; Weir, S. M. *Encyclopedia of Supramolecular Chemistry* **2004**, 893 - 900.
165. Kay, E. R.; Leigh, D. A.; Zerbetto, F. *Angew. Chem. Int. Ed.* **2007**, *46*, 72-191.
166. Serrelli, V.; Lee, C.-F.; Kay, E. R.; Leigh, D. A. *Nature* **2007**, *445*, 523-527.
167. Green, J. E.; Wook Choi, J.; Boukai, A.; Bunimovich, Y.; Johnston-Halperin, E.; Delonno, E.; Luo, Y.; Sheriff, B. A.; Xu, K.; Shik Shin, Y.; Tseng, H.-R.; Stoddart, J. F.; Heath, J. R. *Nature* **2007**, *445*, 414-417.
168. Choi, J. W.; Flood, A. H.; Steuerman, D. W.; Nygaard, S.; Braunschweig, A. B.; Moonen, N. N. P.; Laursen, B. W.; Luo, Y.; Delonno, E.; Peters, A. J.; Jeppesen, J. O.; Xu, K.; Stoddart, J. F.; Heath, J. R. *Chem. Eur. J.* **2006**, *12*, 3-3.
169. Adleman, L. *Science* **1994**, *266*, 1021-1024.
170. Adleman, L. *Sci. Am.* **1998**, *279*, 54-61.
171. Benenson, Y.; Gil, B.; Ben-Dor, U.; Adar, R.; Shapiro, E. *Nature* **2004**, *429*, 423-429.
172. Margolin, A. A.; Stojanovic, M. N. *Nat. Biotechnol.* **2005**, *23*, 1374-1376.
173. Stojanovic, M. N.; Mitchell, T. E.; Stefanovic, D. *J. Am. Chem. Soc.* **2002**, *124*, 3555-3561.
174. Penchovsky, R.; Breaker, R. R. *Nat. Biotechnol.* **2005**, *23*, 1424-1433.
175. Li, Y.; Breaker, R. R. *Curr. Op. Struc. Biol.* **1999**, *9*, 315-323.
176. Taylor, S. K.; Pei, R.; Moon, B. C.; Damera, S.; Shen, A.; Stojanovic, M. N. *Angew. Chem. Int. Ed.* **2009**, *48*, 4394-4397.
177. Kolpashchikov, D. M.; Stojanovic, M. N. *J. Am. Chem. Soc.* **2005**, *127*, 11348-11351.
178. Stojanovic, M. N.; Stefanovic, D. *Nat. Biotechnol.* **2003**, *21*, 1069-1074.
179. Macdonald, J.; Li, Y.; Sutovic, M.; Lederman, H.; Pendri, K.; Lu, W.; Andrews, B. L.; Stefanovic, D.; Stojanovic, M. N. *Nano. Lett.* **2006**, *6*, 2598-2603.
180. Stojanovic, M. N.; Stefanovic, D. *J. Am. Chem. Soc.* **2003**, *125*, 6673-6676.
181. Lederman, H.; Macdonald, J.; Stefanovic, D.; Stojanovic, M. N. *Biochemistry* **2006**, *45*, 1194-1199.
182. Stojanovic, M. N.; Semova, S.; Kolpashchikov, D.; Macdonald, J.; Morgan, C.; Stefanovic, D. *J. Am. Chem. Soc.* **2005**, *127*, 6914-6915.
183. Ashkenasy, G.; Jagasia, R.; Yadav, M.; Ghadiri, M. R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 10872-10877.
184. Kennan, A. J.; Haridas, V.; Severin, K.; Lee, D. H.; Ghadiri, M. R. *J. Am. Chem. Soc.* **2001**, *123*, 1797-1803.
185. Ashkenasy, G.; Ghadiri, M. R. *J. Am. Chem. Soc.* **2004**, *126*, 11140-11141.
186. von Kiedrowski, G.; SimFit 32, Bochum, Germany; : 2003.
187. Allen, V. C.; Robertson, C. C.; Turega, S. M.; Philp, D. *Org. Lett.* **2010**, *12*, 1920-1923.
188. Oxley, P. W.; Adger, B. M.; Sasse, M. J.; Forth, M. A. *Org. Synth.* **1993**, *8*, 16-18.
189. Rajesh, K.; Somasundaram, M.; Saiganesh, R.; Balasubramanian, K. K. *J. Org. Chem.* **2007**, *72*, 5867-5869.
190. Shapiro, R. *Sci. Am.* **2007**, *296*, 46-53.

191. Shapiro, R. *Q. Rev. Biol.* **2006**, *81*, 105-125.
192. Meselson, M.; Stahl, F. W. *Proc. Natl. Acad. Sci. U.S.A.* **1958**, *44*, 671-682.
193. Pieters, R. J.; Huc, I.; Rebek, J. *Tetrahedron* **1995**, *51*, 485-498.
194. Pieters, R. J.; Huc, I.; Rebek, J. *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 1579-1581.
195. Kassianidis, E.; Philp, D. *Chem. Commun.* **2006**, 4072-4074.
196. Rowan, S. J.; Cantrill, S. J.; Cousins, G. R. L.; Sanders, J. K. M.; Stoddart, J. F. *Angew. Chem. Int. Ed.* **2002**, *41*, 898-952.
197. Wood, E. A., University of St. Andrews, 2007.
198. Huck, J., University of St. Andrews, 2011.
199. Carroll, F. I.; Robinson, T. P.; Brieady, L. E.; Atkinson, R. N.; Mascarella, S. W.; Damaj, M. I.; Martin, B. R.; Navarro, H. A. *J. Med. Chem.* **2007**, *50*, 6383-6391.
200. Moore, G. *Electronics* **1965**, *38*, 114-117.
201. The International Technology Roadmap for Semiconductors (ITRS): process integration, d., and structures. (Semiconductor Industry Association, San Jose, California, 2009).
202. De Greef, T. F. A.; Smulders, M. M. J.; Wolfs, M.; Schenning, A. P. H. J.; Sijbesma, R. P.; Meijer, E. W. *Chem. Rev.* **2009**, *109*, 5687-5754.
203. Brunsveld, L.; Folmer, B. J.; Meijer, E. W.; Sijbesma, R. P. *Chem. Rev.* **2001**, *101*, 4071-4098.
204. Cordier, P.; Tournilhac, F.; Soulie-Ziakovic, C.; Leibler, L. *Nature* **2008**, *451*, 977-980.
205. Burattini, S.; Colquhoun, H. M.; Fox, J. D.; Friedmann, D.; Greenland, B. W.; Harris, P. J. F.; Hayes, W.; Mackay, M. E.; Rowan, S. J. *Chem. Commun.* **2009**, 6717-6719.
206. de Greef, T. F.; Meijer, E. W. *Nature* **2008**, *453*, 171-173.
207. Lange, R. F. M.; Van Gurp, M.; Meijer, E. W. *J. Polym. Sci. A1.* **1999**, *37*, 3657-3670.
208. Sijbesma, R. P.; Beijer, F. H.; Brunsveld, L.; Folmer, B. J. B.; Hirschberg, J. H. K. K.; Lange, R. F. M.; Lowe, J. K. L.; Meijer, E. W. *Science* **1997**, *278*, 1601-1604.
209. Fox, J. D.; Rowan, S. J. *Macromolecules* **2009**, *42*, 6823-6835.
210. Kolomiets, E.; Lehn, J.-M. *Chem. Commun.* **2005**, 1519-1521.
211. Keressebaum, A.; Salnikov, G. *DOSY and Diffusion by NMR* Rheinstetten, Germany, **2006**.
212. Cohen, Y.; Avram, L.; Frish, L. *Angew. Chem. Int. Ed.* **2005**, *44*, 520-554.
213. Armarego, W. L. F.; Chai, C. L. L. *Purification of Laboratory Chemicals*; 5th ed.; Elsevier: Oxford, **2003**.
214. Williams, D. E.; Fleming, I. *Spectroscopic Methods in Organic Chemistry*; 5th ed.; McGraw-Hill Companies, **1995**.
215. Simonis, H. *Chem. Ber.* **1912**, *45*, 1584-1592.
216. Peltier, G. *Bull. Soc. Chim. Fr* **1969**, 264.
217. Scarborough, H. A.; Walters, W. A. *J. Chem. Soc.* **1927**, *1*, 1133-1141.
218. Novartis AG; Novartis Pharma GmbH, WO2004/5281A1, 2004.
219. Sun, W.-H.; Zhang, W.; Gao, T.; Tang, X.; Chen, L.; Li, Y.; Jin, X. *J. Organomet. Chem.* **2004**, *689*, 917-929.
220. Carroll, F. I.; Robinson, T. P.; Brieady, L. E.; Atkinson, R. N.; Mascarella, S. W.; Damaj, M. I.; Martin, B. R.; Navarro, H. n. A. *J. Med. Chem.* **2007**, *50*, 6383-6391.
221. Moffett, R. B.; Robert, A.; Skaletzky, L. L. *J. Med. Chem.* **1971**, *14*, 963-968.



222. Gea, A.; Farcy, N.; Russel, N. R. I.; Martins, J. C.; De Clercq, P. J.; Madder, A. *Eur. J. Org. Chem.* **2006**, *2006*, 4135-4146.
223. Yoshida, M.; Katagiri, Y.; Zhu, W.-B.; Shishido, K. *Org. Biomol. Chem.* **2009**, *7*, 4062-4066.
224. Trisler, J. C.; McKinney, S. K. *Tetrahedron Lett.* **1977**, *18*, 3125-3128.
225. Balasubramanian, V.; Bhatia, V. G.; Wagh, S. B. *Tetrahedron* **1983**, *39*, 1475-1485.

# 9

## Appendixes



## A1. Typical SimFit input file for a Bimolecular control reaction with Explanation of commands

```

*=====
* Bimolecular Reaction of Nitron (52) with Maleimide Ester (105)
* 10 mM, 0 °C, 16 h
*
* Results
* ktrans = 6.8095E-5
* kcis   = 2.2514E-5
*=====
DIM (2)
REACTION (A + B --> TRANS ) CONSTANT ( 1, 7.0668E-5, 1, 1, 1000)
REACTION (A + B --> CIS   ) CONSTANT ( 2, 2.0668E-5, 2, 1, 1000)
REACTION (COMPILE)
REACTION (SHOW)
CONSTANT (SHOW)
DEFINE (1, TRANS, P, 1) SCALE (3,1)
DEFINE (2, CIS , P, 4) SCALE (3,1)
SELECT (TRANS, CIS)
READ (SRCONTROL)
REACTION (DOC)
CONSTANT (DOC)
TIME (SEC)
WIN (0, 60000, 15000, 200, 0, 10e-3, 2e-3, 3e-4)
ASSIGN (OBS, TRANS = TRANS )
ASSIGN (OBS, CIS   = CIS   )
ASSIGN (SPEC,  A = #10e-3 )
ASSIGN (SPEC,  B = #10e-3 )
CHOOSE (EXP1)
INTEG (STIFF, 1E-9, 1, 0.075, 100, 50)
PLOT (OBS, RES)
* Optimise rate constants using simplex
SIMPLEX (PLOT)
SIMPLEX (PLOT)
SIMPLEX (PLOT)
SIMPLEX (PLOT)
SIMPLEX (PLOT)
NEWTON (PLOT)
PLOT (OBS, RES)
PLOT (FILE)

```

Lines beginning with a \* are for reference only.

Number of constants in model allowed to vary

Each unique reaction is described

Accompanied by the rate constant for the process

Simulation is allowed to vary the rate constant by this magnitude to optimise fitting

Experimental results file

Length of reaction (x axis) and concentration (y axis)

Species involved

Initial concentrations

## A2. Typical SimFit input file for a Self-Replicating reaction

```
*****
* Maleimide (34) + Nitron (52)
*****
*
* This model only allows the dimerisation constant
* and the rate constant in the ternary complex to vary.
*
* Bimolecular rates from experiments
* Ka = 1020 M-1 from experiment
*
* Results
*
* 1.679% RMS
* K off 1.39E+03
* Ka Duplex 7.21E5 Mol-1
* TRANSAB-->TransTrans 1.063E-3
* EM 19.4 Mol
*****

DIM (2)

* Bimolecular reactions

REACTION (A + B --> TRANS ) CONSTANT ( 1, 6.8095E-5, 0)
REACTION (A + B --> CIS ) CONSTANT ( 2, 2.2514E-5, 0)

* Formation of binary complexes with Trans including product duplexes
* Only [trans*trans] is stable beyond one Pyr*COOH association

REACTION (A + TRANS ==> ATRANS ) CONSTANT ( 3, 1E9, 0) CONSTANT ( 4, 9.8E5,0)
REACTION (B + TRAN ==> BTRANS ) CONSTANT ( 5, 1E9, 0) CONSTANT ( 6, 9.8E5,0)
REACTION (A + BTRANS ==> ABTRANS ) CONSTANT ( 7, 1E9, 0) CONSTANT ( 8, 9.8E5,0)
REACTION (B + ATRANS ==> ABTRANS ) CONSTANT ( 9, 1E9, 0) CONSTANT (10, 9.8E5,0)
REACTION (A + B ==> AB ) CONSTANT (11, 1E9, 0) CONSTANT (12, 9.8E5,0)

REACTION (TRANS + TRANS ==> TRANSTRANS) CONSTANT (13, 1E9, 0) CONSTANT (14, 1.00E
+3, 1, 1, 1000)

* Ternary complex reaction

REACTION (ABTRANS --> TRANSTRANS) CONSTANT (15, 2E-4, 2, 1, 1000)

*Bimolecular Reactions of Trans Complexes

REACTION (AB + A --> TRANS + A ) CONSTANT (16, 6.8095E-5, 0)
REACTION (AB + B --> TRANS + B ) CONSTANT (17, 6.8095E-5, 0)
REACTION (AB + ATRANS --> TRANS + A + TRANS ) CONSTANT (18, 6.8095E-5, 0)
REACTION (AB + BTRANS --> TRANS + B + TRANS ) CONSTANT (19, 6.8095E-5, 0)
REACTION (AB + ABTRANS--> TRANS + A + B + TRANS ) CONSTANT (20, 6.8095E-5, 0)
REACTION (AB + AB --> TRANS + A + B ) CONSTANT (21, 6.8095E-5, 0)

REACTION (AB + A --> CIS + A ) CONSTANT (22, 2.2514E-5, 0)
REACTION (AB + B --> CIS + B ) CONSTANT (23, 2.2514E-5, 0)
REACTION (AB + ATRANS --> CIS + A + TRANS ) CONSTANT (24, 2.2514E-5, 0)
REACTION (AB + BTRANS --> CIS + B + TRANS ) CONSTANT (25, 2.2514E-5, 0)
REACTION (AB + ABTRANS--> CIS + A + B + TRANS ) CONSTANT (26, 2.2514E-5, 0)
REACTION (AB + AB --> CIS + A + B ) CONSTANT (27, 2.2514E-5, 0)

* Reversible interactions with cis

REACTION (A + CIS ==> ACIS ) CONSTANT ( 28, 1E9, 0) CONSTANT ( 29, 9.8E5,0)
REACTION (B + CIS ==> BCIS ) CONSTANT ( 30, 1E9, 0) CONSTANT ( 31, 9.8E5,0)
REACTION (A + BCIS ==> ABCIS ) CONSTANT ( 32, 1E9, 0) CONSTANT ( 33, 9.8E5,0)
REACTION (B + ACIS ==> ABCIS ) CONSTANT ( 34, 1E9, 0) CONSTANT ( 35, 9.8E5,0)

* Bimolecular Reactions of Cis Complexes
```

```

REACTION (A + BCIS --> CIS + CIS ) CONSTANT (36, 2.2514E-5, 0)
REACTION (B + ACIS --> CIS + CIS ) CONSTANT (37, 2.2514E-5, 0)
REACTION (A + ABCIS --> CIS + CIS + A ) CONSTANT (38, 2.2514E-5, 0)
REACTION (B + ABCIS --> CIS + CIS + B ) CONSTANT (39, 2.2514E-5, 0)

REACTION (COMPILE)
REACTION (SHOW)
CONSTANT (SHOW)
DEFINE (1, TRANS, P, 1) SCALE (3,1)
DEFINE (2, CIS , P, 4) SCALE (3,1)
READ (SR)
SELECT (TRANS, CIS)
REACTION (DOC)
CONSTANT (DOC)
TIME (SEC)
WIN (0, 60000, 20000, 200, 0, 10e-3, 5e-3, 3e-4)
ASSIGN (OBS, TRANS = TRANS + ATRANS + BTRANS + ABTRANS + 2 TRANSTRANS)
ASSIGN (OBS, CIS = CIS + ACIS + BCIS + ABCIS)
ASSIGN (SPEC, A = #10e-3 )
ASSIGN (SPEC, B = #10e-3 )
CHOOSE (EXP1)
INTEG (STIFF, 1E-9, 1, 0.075, 100, 50)
PLOT (OBS, RES)

* Optimise rate constant using simplex

SIMPLEX (PLOT)
SIMPLEX (PLOT)
SIMPLEX (PLOT)
SIMPLEX (PLOT)
SIMPLEX (PLOT)

PLOT (OBS, RES)
PLOT (FILE)

```

### A3. Four Product System Simulation SimFit Input File (All Rate Equal)

```
*=====
* Simulation of 4 product system
*=====
* A = A
* B = B
* C = C
* D = D
* E = AB
* F = BD
* G = CD
* H = AC
*=====

MODE (ISOSIM)

* Reactions

REACTION (A + B      --> E      , 3E-02)
REACTION (A + C      --> H      , 3E-02)
REACTION (D + B      --> F      , 3E-02)
REACTION (D + C      --> G      , 3E-02)

REACTION ( COMPILE )
REACTION ( SHOW      )

* Initial Concentrations

INIT      ( A , 0.01000 , 1)
INIT      ( B , 0.01000 , 2)
INIT      ( C , 0.01000 , 3)
INIT      ( D , 0.01000 , 4)
INIT      ( E , 0.00000 , 5)
INIT      ( F , 0.00000 , 6)
INIT      ( G , 0.00000 , 7)
INIT      ( H , 0.00000 , 8)

INTEG      (STIFF, 1E-9, 8, 0.05, 200, 100)
NUMPLOT    ( 100 )
TIME       ( SEC )
WIN        (0, 2400000, 200000, 0.1, 0, 10E-3, 2E-3, 1E-4)
PLOT       (FILE)
PLOT
```

#### A4. Four Product System Simulation SimFit Input File (AB Rate Enhanced)

```
*=====
* Simulation of 4 product system
*=====
* A = A
* B = B
* C = C
* D = D
* E = AB
* F = BD
* G = CD
* H = AC
*=====

MODE (ISOSIM)

* Reactions

REACTION (A + B      --> E      , 3E-01)
REACTION (A + C      --> H      , 3E-02)
REACTION (D + B      --> F      , 3E-02)
REACTION (D + C      --> G      , 3E-02)

REACTION ( COMPILE )
REACTION ( SHOW      )

* Initial Concentrations

INIT      ( A , 0.01000 , 1)
INIT      ( B , 0.01000 , 2)
INIT      ( C , 0.01000 , 3)
INIT      ( D , 0.01000 , 4)
INIT      ( E , 0.00000 , 5)
INIT      ( F , 0.00000 , 6)
INIT      ( G , 0.00000 , 7)
INIT      ( H , 0.00000 , 8)
INTEG     (STIFF, 1E-9, 8, 0.05, 200, 100)
NUMPLOT   ( 100 )
TIME      ( SEC )
WIN       (0, 2400000, 200000, 0.1, 0, 10E-3, 2E-3, 1E-4)
PLOT      (FILE)
PLOT
```





## **A6. Publications**

# A simple network of synthetic replicators can perform the logical OR operation

Victoria C. Allen, Craig C. Robertson, Simon M. Turega, and Douglas Philp

*EaStCHEM and Centre for Biomolecular Sciences, School of Chemistry, University of  
St Andrews, North Haugh, St. Andrews, Fife KY16 9ST, United Kingdom*

Organic Letters (2010) Vol. 12, No. 9 pp. 1920-1923

(doi: 10.1021/ol100404g)

Owing to copyright restrictions, the electronic  
version of this thesis does not contain the text of  
this article