

LEABHARLANN CHOLÁISTE NA TRÍONÓIDE, BAILE ÁTHA CLIATH | TRINITY COLLEGE LIBRARY DUBLIN
Ollscoil Átha Cliath | The University of Dublin

Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Towards A Comprehensive System For The Recognition Of Double Helical DNA: Components for Base Pair Recognition in the Major Groove

By

Mary Rose Kelleher

Trinity College Dublin, 2001

A thesis submitted to the University of Dublin in fulfilment of the requirements for the degree of Doctor of Philosophy.

TRINITY COLLEGE
22 MAY 2002
LIBRARY DUBLIN
Thesis
6822

Declaration

This thesis has not been submitted as an exercise for a degree at this or any other University. Except where otherwise stated, the experimental work described herein has been carried out by the author alone. I agree that the library may lend or copy this thesis upon request.

Mary Rose Kelleher

Mary Rose Kelleher

Acknowledgments

I would like to thank my supervisor Professor Tony Davis for all his help and encouragement during the past four years. I would also like to thank him for presenting the opportunity to work for a year in the University of Bristol with funding. I am very grateful to Trinity Foundation, Enterprise Ireland and Limerick County Council for earlier financial assistance.

I am indebted to both the Chemistry Department, Trinity College Dublin, and the School of Chemistry, University of Bristol on every level for providing the means and the facilities for completing this project. I am particularly grateful to TCD for the multitude of opportunities presented to me through my association with the Chemistry Department. I would like to thank the Heads of Department Professor John Kelly for his support throughout my College career, and Professor John Corish for the possibility to work in the department. Of the other academic and technical staff who have all helped me along the way, special thanks must go to Dr. John O'Brien for introducing me to the Black Art of NMR in the mystical surroundings of Room 0.14; Fred and Ed for their supreme stores service; Colm and Noel for the ever-efficient solvent service, and Theresa for the fast sourcing of chemicals. In the School of Chemistry in Bristol, I am grateful to Paul and Brian in Stores for always spelling my name in full on the whiteboard along with their excellent service; Martin West and Martin Pearce for endeavouring to keep N405 at 20 °C; Dr. Martin Murray and his team for their accessible super-NMR service and Dr. Ken MacNeil for his ever-vigilant Mass Spectrometry. and Lab. Safety service. I am also very grateful to the mysterious person in Latvia who sent me the much-needed Russian article (in English) by e-mail without asking for reward.

I would like to take this opportunity to show my appreciation for my friends and colleagues in the APD Research Group (also known as the Backlab/N405/N314). Special thanks is given to Dr. Gavin Dunne for quietly and determinedly paving the way in many aspects of the chemistry used in this project, and never complaining about bench space. I would like to especially thank Jens, Trini, Julie and Laura for their true friendship and support, particularly in the last year of change, and helping to provide me with a balanced outlook on life. I would also like to thank Nieves, Tríona and Andrew for their help and friendship while in the group, and even now that they have moved on to pastures greener, and also John Clare for his invaluable computer support. I am also grateful to the other members of my group (Alan, Bea, John Barry, Adrian B, Shay, Killian, Fionn, Amy, Larry, Greg, Trevor, Eileen, Mary, Michael, Lise, Adrian R, Theo and Rich) for their varied contributions to my working environment and wish them all the very best in their bright futures. I would also like to thank the other postgraduate and postdoctoral workers in both universities for their welcome input into my working and social life. Thanks must also be given to Dr. Damien Scully for proof-reading my various drafts.

I would like to thank my 'normal' friends and acquaintances who have helped to make the last four years endurable at the worst of times, and enjoyable at every other time. Last, but by no means least, I would like to thank my family for their solid love and support during this time and always.

For My Family.

"The best books are those that tell you what you know already."

George Orwell
-Nineteen Eighty Four

Summary

This thesis describes the attempted synthesis of sequence selective double-stranded DNA-binding ligands. It had been intended to provide a comprehensive DNA recognition system *via* two different tricyclic heterocyclic motifs. The (5,5,5) pyrrolo[2′,3′:4,5]thieno[3,2-*d*]thiazole system **1.52** was proposed to recognise A-T/T-A base pairs, and the functionalised (6,4,6)-tricyclic heterocyclic system **1.56** would recognise G-C/C-G base pairs.

Initial efforts attended to the synthesis of a series of *N*-carbamoyl squaramides **2.8-2.22** with the purpose of biological testing. A second series of thiourea analogues **3.2-3.17** were produced in preparation for the intermediate (6,4)-bicyclic systems **4.4** and **4.7**. Methylation of these thiourea squaramides resulted in unstable solids that invariably decomposed into red tars. Aminoguanidine derivatives **4.5** and **4.6** were synthesised in an attempt to yield the intermediate bicyclic (6,4)-system in a one-pot synthesis with diisopropyl squarate **2.5**. These one-pot attempts either failed to affect **2.5** or resulted in a polymeric tar. The synthesis of the thionated squaric acid derivatives **4.28**, **4.30**, **4.31** and **4.32** was successfully achieved. However, the inevitable hydrolysis of the ensuing dimethylated thiocarbonyl compound to the thiosquarate **4.33** dissuaded further efforts on this (6,4,6)-system.

Attempts to access the pyrrolo[2′,3′:4,5]thieno[3,2-*d*]thiazole system **1.52** were then commenced. Oxidative cyclisation of the 1-acetyl-3-(3-thienyl)-thiourea **5.5** yielded the desired iodothieno[3,2-*d*]thiazole **5.8**. It was found to be necessary to block the acetamido nitrogen of **5.8** with a 2,4-dimethoxybenzyl protecting group. Further derivatisation to the enamine **5.17** was attempted by a variety of methods, and success was achieved on one occasion to yield **5.17** as a mixture of *cis* and *trans* isomers. The previously completed pyrrolo[2′,3′:4,5]thieno[3,2-*d*]thiazole system **1.53** was resynthesised for synthetic comparisons. It was proposed that **1.52** could have been obtained *via* a one-pot synthesis from **5.16** using Pd catalysis, but attempts yielded polymeric tars, or did not affect the starting material **5.16**.

The C-G/G-C recognising methylthienothiazole system **6.1** was simultaneously pursued. Synthesis to the thioacetylthiophene intermediate **6.12** was successfully completed, but time did not allow completion of this effort.

Abbreviations

A	Adenine
Ac	Acetyl
AcOH	Acetic acid
app	Apparent
aq.	Aqueous
Ac_2O	Acetic anhydride
^{t}Boc	^{t}Butoxycarbonyl
$(^t\text{Boc})_2O$	Di-^{t}Butyl dicarbonate
bp	Base pair
b.p.	Boiling point
br	Broad
C	Cytosine
CCl_4	Carbon tetrachloride
$CDCl_3$	Chloroform-*d*
$CHCl_3$	Chloroform
CH_3CN	Acetonitrile
d	Day
d	Doublet
δ	Chemical shift
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DCM	Dichloromethane
DEAD	Diethyl azodicarboxylate
decomp.	Decomposed
DEPT	Distortionless Enhancement by Polarisation Transfer
DIAD	Diisopropyl azodicarboxylate
DIBAL-H	Diisobutylaluminum hydride
DIPEA	*N,N*-Diisopropylethylamine
DMF	*N,N*-Dimethylformamide
DMSO	Dimethylsulphoxide
DMAP	4-(*N,N*-dimethylamino)pyridine
DNA	Deoxyribonucleic acid
equiv.	Equivalent
Et	Ethyl
EtOAc	Ethyl Acetate
EtOH	Ethanol
FAB MS	Fast Atom Bombardment Mass Spectroscopy
G	Guanine
h	Hour
Hp	Hydroxypyrrole
HPLC	High Performance Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
Hz	Hertz
Im	Imidazole
IR	Infrared
J	Coupling constant

K_a	Association constant
K_D	Dimerisation constant
KO^tBu	Potassium *tert*-butoxide
K_2CO_3	Potassium carbonate
LC-MS	Liquid chromatography-mass spectrometry
lit.	Literature
m	Multiplet
MALDI-TOF	Matrix assisted, laser desorption/ionisation-time of flight
Me	Methyl
MeOH	Methanol
min	Minute
m.p.	Melting point
m/z	mass charge ratio
NaH	Sodium hydride
NaOAc	Sodium acetate
NaOMe	Sodium methoxide
NIS	*N*-Iodosuccinimide
NOE	Nuclear Overhauser effect
NMR	Nuclear Magnetic Resonance
ODN	Oligodeoxyribonucleotide
$PdCl_2(PPh_3)_2$	Dichlorobis(triphenylphosphine)palladium(II)
$Pd(OAc)_2$	Palladium acetate
Ph	Phenyl
PPh_3	Triphenylphosphine
PNA	Peptide Nucleic Acid
ppm	Parts per million
p-TsOH	*para*-Toluenesulphonic acid
Py	Pyrrole
quant.	Quantitative
Ref.	Reference
R_f	Retention factor
R.S.	Relative Specificity
RT	Room Temperature
s	Singlet
T	Thymine
TFA	Trifluoroacetic acid
TFO	Triplex Forming Oligonucleotide
THF	Tetrahydrofuran
TLC	Thin Layer Chromatography
TOCSY	Total Correlation Spectroscopy
UV	Ultraviolet

TABLE OF CONTENTS

Title page (i)
Declaration (ii)
Acknowledgements (iii)
Summary (vi)
Abbreviations (vii)
Table of Contents (ix)

CHAPTER 1: INTRODUCTION

1.1 Sequence Selective Binding of Double Helical DNA 1

1.2 Sequence Specific Binding of Double Stranded DNA by Intercalation 4

1.2.1 Sequence specific binding of double stranded DNA by simple intercalation 5
1.2.2 Sequence specific binding of double helical DNA by major groove binding metallointercalators 6

1.3 Sequence Selective Recognition of Double Helical DNA by Minor Groove Binding Polyamides 10

1.3.1 Sequence selective minor groove binding of DNA by hairpin motif polyamides 13
1.3.2 Sequence specific minor groove binding of DNA by extended polyamides 14
1.3.3 Formation of a complete pairing code by minor groove binding polyamides 20
1.3.4 Sequence selective minor groove recognition by second generation polyamides 23

1.4 Sequence Selective Recognition of Double Helical DNA by Major Groove Binding Triplex Forming Ligands 29

1.4.1 Recognition of double stranded DNA *via* the pyrimidine motif 31
1.4.1.1 Replacement of cytosine by pH independent pyrimidine-like analogues 31
1.4.1.2 Replacement of cytosine by pH independent purine analogues 38
1.4.2 Recognition of double stranded DNA *via* the purine motif 40
1.4.3 Recognition of double stranded DNA by novel base analogues 43
1.4.3.1 C-G base pair recognition by extended heterocyclic systems 43
1.4.3.2 G-C base pair recognition by extended heterocyclic systems 51
1.4.3.3 T-A base pair recognition by extended heterocyclic systems 53

1.5 Aims of this Project 56

CHAPTER 2: The *N*-Carbamoyl Squaramide System

2.1 Preamble 65

2.2 Synthesis of *N*-Carbamoyl Squaramide 1.55a-b 66

2.3 Synthesis of *N*-Carbamoyl Squaramide 2.8-2.22 68

CHAPTER 3: The Bicyclo[4.2.0]octa-1(6)-ene-7,8-dione System

3.1 Preamble 76

3.2 Synthesis of Hydrazino Thiourea Derivatives 3.3, 3.5, 3.7 and 3.9 77

3.3 Synthesis of Ethoxycarbonylmethylamino Thiourea Derivatives 3.14, 3.15, 3.16 and 3.17 82

3.4 Attempted Methylation of Thiourea Derivatives 84

CHAPTER 4: Alternative Routes to (6,4,6)-systems

4.1 Preamble 88

4.2 Synthesis of Aminoguanidine Derivatives 90

4.3 Attempted One-Pot Synthesis of Intermediate (6,4)-Bicyclic Systems 93

4.3.1 Attempted One-Pot Synthesis of Bicyclic Systems **4.4** and **4.7** 93
4.3.2 Synthesis of Thionate Squarate Derivatives **4.28**, **4.30**, **4.31** and **4.32** 96

CHAPTER 5: The *N*-2,4-Dimethoxybenzyl Protected Pyrrolo[2′,3′:4,5]thieno [3,2-*d*]thiazole System

5.1 Preamble 101

5.2 Synthesis of the 2-Acetamidothieno[3,2-*d*]thiazole System 5.7 102

5.3 Synthesis of Enamine Derivatives 5.12 and 5.17 106

5.4 Attempted Synthesis of Enamine Derivative 5.15 110

5.4.1 Attempted *N*-Vinylation of Thienothiazole **5.16** 110
5.4.2 Attempted Synthesis of (3,3,3)-Triethoxypropanal **5.22** 112

5.5 Synthesis of 2,4-Dimethoxybenzyl-protected 5-Ethoxycarbonylmethylpyrrolo [2′,3′:4,5] thieno [3,2-*d*]thiazole 5.26 113

5.6 Attempted One-Pot Pd-Catalysed Synthesis of 2,4-Dimethoxybenzyl Protected 5-Ethoxycarbonylpyrrolo [2′,3′:4,5] thieno [3,2-*d*]thiazole 5.27 114

CHAPTER 6: Alternative Thieno[3,2-*b*]pyrrole Systems

6.1 Preamble 117

6.2 Attempts towards the Synthesis of 6-Amino-5-ethoxycarbonyl-2-methylpyrrolo[2',3':4,5]thieno[3,2-*d*]thiazole 6.1 118

6.2.1 Attempted Synthesis of Substituted 3,4-Dinitro-2-thienylcyanoacetate **6.3** 118
6.2.2 Attempted Synthesis of 2-Methylthieno[3,2-*d*]thiazole **6.12** 119

CHAPTER 7: Experimental 123

CHAPTER 8: References 165

APPENDIX A: NMR Reference Codes 171

APPENDIX B: X-Ray Crystallography Tables 172

Chapter One

INTRODUCTION

1.1 Sequence Selective Binding of Double Helical DNA

The ever-increasing frequency of cancer, mutations in viral strains (e.g. the 'flu' virus) and other diseases that are caused by the expression of particular genes that are present in, or incorporated into, the human genome have been prominent areas of intensive research in recent years. As a result, the modern drugs used in the treatment of these diseases, as well as many of the more important antibiotics, numerous antiparasitic agents, many antineoplastic drugs and most of the antiviral compounds exert their varied actions on the different phases of nucleic acid function.

The nucleic acids, deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are, respectively, the molecules that preserve hereditary information, and that transcribe and translate it in a way that allows the synthesis of all the varied proteins in the cell. The large number of RNA molecules that are produced by a single gene in the cell have long been targeted by 'antisense' drugs, which prevent the translation of the RNA into harmful proteins, which leads to continually increasing drug dosages and, consequently, toxicity and other undesirable side-effects such as gastrointestinal disturbances, hair loss and amnesia. It would be much more efficient and clinically preferable to generate a drug which would target the specific sequence in the DNA that produces the harmful proteins as a result of its expression.

DNA has the 3-dimensional structure of a flexible ladder wrapped helically around a central axis. The two rails of the ladder are chains of alternating deoxyribose sugar rings and phosphate groups. The rungs are purine-pyrimidine base pairs. There are two types of bicyclic purine rings: adenine and guanine, and two kinds of monocyclic pyrimidines, cytosine and thymine (see **Figure 1.1**). The base pairs are held together by hydrogen bonds. Adenine normally is paired with thymine by two hydrogen bonds, and guanine with cytosine *via* three bonds (**Figure 1.2**).

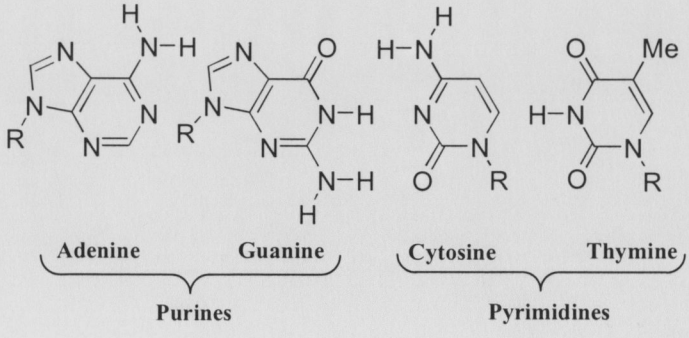


Figure 1.1 DNA Bases.

DNA, the vehicle of inheritance, carries two kinds of genetic information in its double helical structure. DNA not only encodes the genetic message itself, but also instructions for the selective expression of the message. A gene is a blueprint for a protein that is encoded in a particular sequence of base pairs in DNA, and as there are only two copies of each gene in a diploid cell, it would be advantageous to exploit this characteristic and focus on molecules that exert their biological activity directly on DNA.

B-DNA, which is the active configuration found within the high humidity conditions of the cell, has a secondary structure such that there are ten layers of perpendicularly stacked bases for every complete turn of the helix. The asymmetric attachment of the base pairs to the phosphate backbones forms a major and a minor groove within the double helix, which wind around the helix, with the top and bottom edges forming the floor of the grooves (**Figure 1.3**).

Within these grooves, DNA contains two channels of information. The minor groove is narrow, is normally relatively unoccupied and offers itself as a poor candidate for information readout, with two hydrogen bond acceptor groups on A-T base pairs, and a G-C pair differs only by the intrusion of an NH_2 donor between the acceptors (**Figure 1.2**).

The major groove offers approximately twice the information content of the minor groove in terms of hydrogen bond acceptance and donation sites for molecular recognition. Four different hydrogen-bonding patterns are possible. An A-T pair offers an acceptor/donor/acceptor pattern, and the reverse pattern holds for T-A. C-G offers a donor/acceptor/acceptor pattern. The additional hydrogen bond donor and acceptor groups possessed by the bases, i.e. those not used in specific Watson-Crick base pairing in both grooves of the double helix, may be exploited for sequence specific recognition.

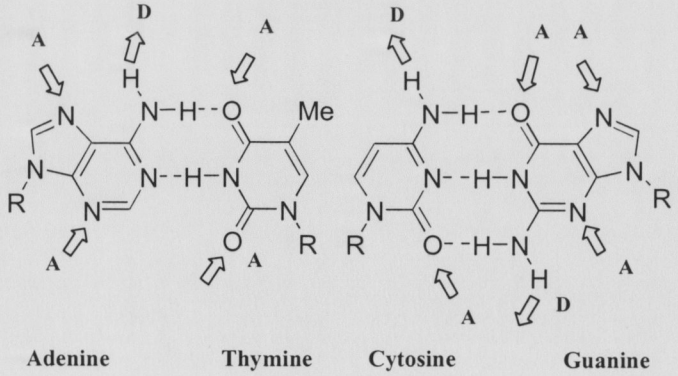


Figure 1.2 Watson-Crick Base Pairing.

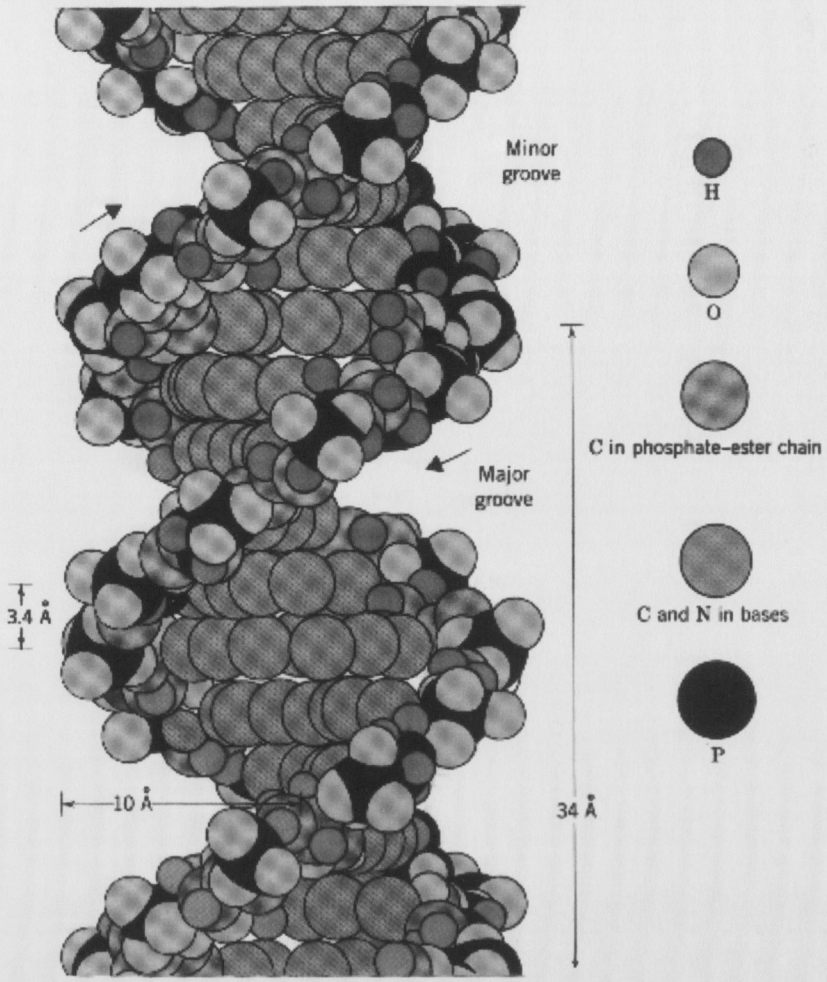


Figure 1.3 3-Dimensional Structure of B-DNA showing the major and minor grooves.[1.1]

While every cell in the human body will contain exactly the same DNA, not every tissue will express the same set of genes. The human genome, that is the sum of all the genes in the human cell, is estimated to contain 100,000 genes, or 2.9 billion base pairs, but most cells only express about 10,000 of these genes at any one time.

Thus mechanisms must exist by which certain genes are switched on and others remain dormant. It is known that gene expression involves specific regulatory proteins called transcription factors. The binding of the transcription factor to a promoter site, just upstream of the beginning of the gene to be copied, permits the transcription of that gene. If the promoter site is blocked, however, by the presence of a DNA binding molecule, for example, then effective binding of the transcription factor is prohibited and the gene will remain untranscribed. In order for a ligand to recognise a single sequence of DNA, the sequence of the gene must be of sufficient length that, on a statistical basis, it should occur only once. A general DNA effector

capable of recognising a unique sequence in the human genome would need to cover a site of 14-16 base pairs in length.

As only a small proportion of the DNA present function as genes (only about 3% actively, and about 70% of this 3% is thought to be expressed), and is therefore accessible for inhibition by the drug, it may be possible to inhibit gene expression with molecules that recognise shorter base pair sequences, which in themselves would be of considerable benefit in cutting down side effects.

If the ligand is to sequence selectively bind to double helical DNA, it must interact with the bases. Access to the bases, however, can only be had through either the major or minor grooves, which are on the external surface of the helix. This means that specific binding can be achieved only *via* one or more of three motifs: intercalation between base pairs, binding in the minor groove and binding in the major groove.

Although these non-covalent binding motifs observe different recognition modes, there is frequent overlap between the categories of the forces of interaction involved; hydrogen bonding, electrostatic, hydrophobic, thermodynamic, steric and π-π interactions play a role of varying importance in all three binding modes.

The current interest in DNA recognition has lead to the investment of a great deal of intellectual effort and organic synthesis in the molecular recognition of double stranded DNA. It is quite remarkable that although great progress has been made in this area, a general recognition scheme of any desired sequence of DNA is still elusive. The following pages offer a concise summary of the most recent and relevant publications in three sections, dealing with intercalation, minor and major groove recognition, respectively. The chapter concludes with a short draft of the aims of my project: *Towards a comprehensive system for the recognition of double helical DNA: components for base pair recognition in the major groove.*

1.2 Sequence Selective Binding of Double Stranded DNA by Intercalation

DNA intercalators comprise of a diverse collection of small aromatic planar compounds that insert (intercalate) between two adjacent base pairs. This action involves elongating the DNA helix by 3.4 Å and simultaneously unwinding the double strand to fit the ligand. The highly stable, yet reversibly bound complex is therefore formed not through hydrogen bonding with the groove functionalities, but through π-π and hydrophobic interactions.

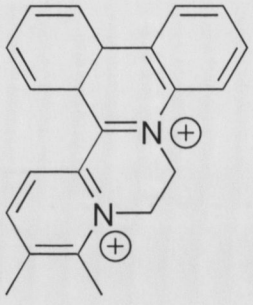
Although intercalators bind to B-DNA with high affinity, their intercalative mode of binding has suffered from a poor degree of sequence selectivity, due to the non-specific nature of the forces involved. Some progress has been made in the development of site-specific

metallointercalators which incorporate the use of hydrogen-bonding, van der Waals forces, electrostatic interaction and steric bulk to direct the ligand to a desired sequence, usually of high polarisability or twistability.[1,2] A short review in the recent developments in this area follows.

1.2.1 Sequence Specific Binding of Double Stranded DNA by Simple Intercalation

Contribution of Orellana *et al.*

Compound **1.1**, a *N,N'*-dialkylated phenanthridinium dication, was designed, synthesised and its binding affinity (as the association constant, K_a) with polynucleotides of varying base sequence determined by visible absorption titrations at 25 °C;[1,3] the results of which are shown in **Table 1.1**.



1.1

Table 1.1 Association constants K_a (M^{-1}) for interaction of **1.1** with calf thymus DNA and polynucleotides in pH 5.5 sodium phosphate buffer at 25 °C.

Polynucleotide	K_a (M^{-1})
Poly $[d(G-C)]_2$	2.42×10^5
Calf Thymus DNA	1.9×10^5
Poly $[d(A-T)]_2$	6.3×10^6
Poly(dG)•poly(dC)	2.6×10^6
Poly(dA)•poly(dT)	1.8×10^6

A preference of **1.1** for alternating G-C base pairs was observed as the poly$[d(G-C)]_2$ oligomer was bound with an almost four fold higher affinity over poly$[d(A-T)]_2$. Calf thymus DNA, which possessed approximately 42 % G-C base pairs, was bound with an affinity

intermediate between the former two. The remaining order concluded with poly(dG)•poly(dC) binding with higher affinity than poly(dA)•poly(dT).

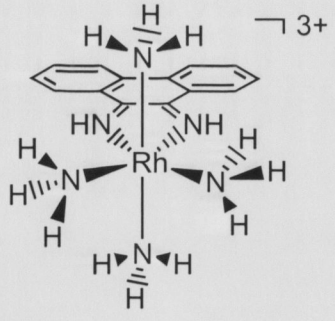
The strong binding preference of **1.1** for alternating G-C sequences has been attributed to the higher polarity of the G-C base pair (compared to the A-T base pair) which formed a complex of greater stability with the polarisable **1.1**.

1.2.2 Sequence Specific Binding of Double Stranded DNA by Major Groove Binding Metallointercalators

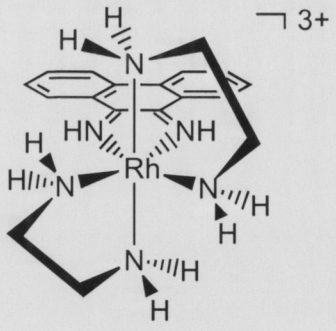
The modest number of strategies available for interfering with DNA transcription, a process which takes place in the major groove, prompted interest in the use of metallointercalators as small molecules which would preferentially recognise predetermined regions of double stranded DNA through direct hydrogen bonding and/or specific van der Waals contacts between auxiliary ligands on the metallointercalator and the DNA bases in the major groove. A class of intercalator now exists that has the ability to sequence specifically bind DNA.

Contribution of Barton *et al.*

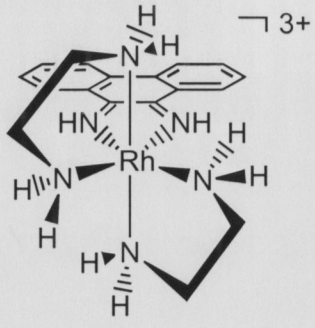
The series of rhodium polyamine complexes **1.2**, **1.3** and **1.4**, all of which contain the 9,10-phenanthrenequinone diimine (*phi*) ligand, were synthesised and their binding affinities for different base pair sites tested.[1.4]



1.2



1.3

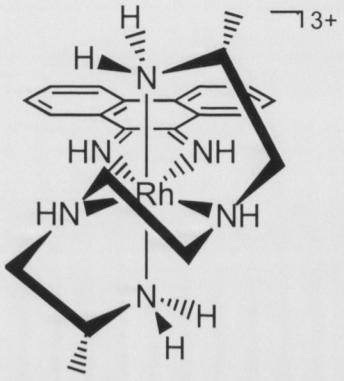


1.4

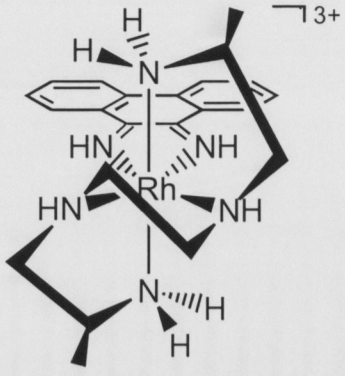
Complexes **1.2** and **1.3** were observed to cleave strongly at 5′-GC-3′ sites, with $K_a \sim 10^6$ M^{-1}. Such specificity was ascribed to the hydrogen bonding contributions between axial amines on the complex and the C-6 carbonyl of guanine. Complex **1.4**, the enantiomer of **1.3**, was observed by cleavage studies to show less sequence selectivity by binding 5′-TX-3′ steps (X = A, G, C, T), but preference was shown for 5′-TA-3′ sites. As both enantiomers contained similar hydrogen bond donor amine sites, enantioselectivity was not accounted for by hydrogen

bonding; but instead by van der Waals contacts between the methylene groups on the backbone of **1.4** and the methyl groups of thymine on the floor of the major groove. Support for this theory was obtained by replacing thymine with uracil in 5′-TA-3′ steps in further photocleavage studies, where **1.3** enantioselectivity was not observed.

With this knowledge at hand, **1.5-(R,R)** and **1.5-(S,S)** were synthesised. It was thought that a 5′-TGCA-3′ site could be targeted by this design through the incorporation of two methyl groups at either end of the original 5′-GC-3′ recognising design of **1.3**, which form van der Waals interactions with the thymine methyl groups.



1.5-(R,R)



1.5-(S,S)

Photoinduced DNA cleavage studies were carried out on **1.5-(R,R)** and **1.5-(S,S)** to investigate their binding selectivities.[1.5] **1.5-(S,S)** showed no improvement in selectivity on the original **1.3** design. **1.5-(R,R)**, however, bound the target site with a binding constant of 9×10^7 M^{-1}.[1.6] A high-resolution NMR structure[1.6] and an X-ray crystal structure (the first crystallographically characterised high resolution view of a metallointercalator bound to double-stranded DNA) of **1.5-(R,R)** intercalated through the major groove[1.7] both verified that the complex binds the target sequence 5′-TGCA-3′ *via* the major groove, and all molecular model-predicted contacts were observed.

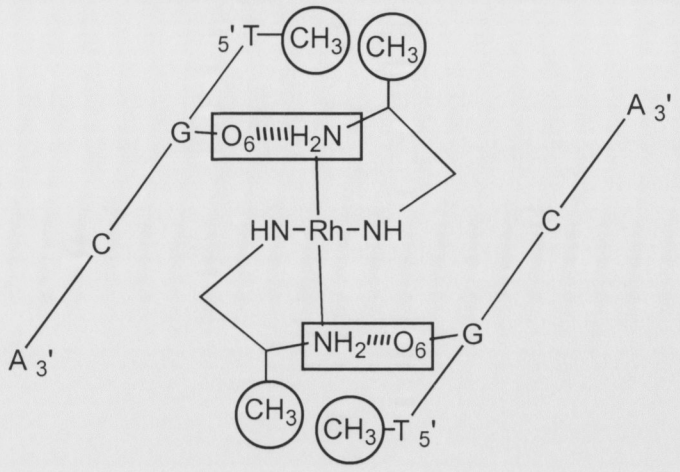
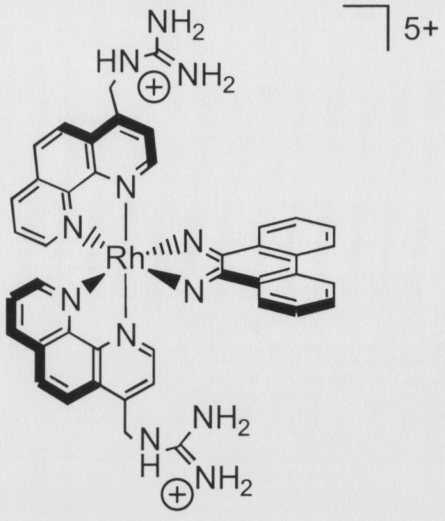


Figure 1.4 Schematic of **1.5-(R,R)** bound to 5′-TGCA-3′ showing proposed hydrogen bonds and van der Waals contacts.

The crystal structure of **1.5-(R,R)** co-crystallised with the eight base pair oligomer 5′-G-dIU-TGCAAC-3′ (dIU = 5-iodo-deoxyuridine) was solved at 1.2 Å resolution. Detailed interactions in sugar pucker, backbone conformation, π-stacking and sequence-specific interactions were obtained from five independent intercalated DNA complexes in the asymmetric unit of the crystal. Metallointercalation resulted in the minimum perturbation of DNA local structure in this case. The crystal structure exhibited good base pair stacking, a deeply intercalated *phi* ligand and the B-form conformation of the deoxyribose sugars at the site of intercalation. In other words, the intercalator resembled an inserted base pair.

In an effort to combine direct readout with shape selection in DNA recognition, **1.6**, a metallointercalator incorporating two asymmetric 4-(guanidylmethyl)-1,10-phenanthroline (MGP) ligands, was designed to target a 5′-CATATG-3′ site.[1.8, 1.9] Modelling studies predicted that the *phi*-ligand preferentially inserts in the major groove between DNA base pairs, and forms a rigid complex which orients the chiral MGP ligands in such a way that favourable hydrogen bond contacts can be made between the guanidinium groups and the N-7 and C-6 carbonyl of guanine at the first and sixth base pair positions.



1.6

Photocleavage studies showed that the target 5′-CATATG-3′ site was indeed bound specifically and with high affinity.[1.8] However, it was found that the recognition elements direct specific binding only if the target double strand is significantly unwound, so that it practically resembles a ladder. A DNA unwinding assay proved that a necessary 70 ° unwinding of the target sequence was observed to achieve optimal guanidinium-guanine contact. The inherent 'twistability' of A-T segments permitted this unwinding without energy penalty, and the complete loss of recognition incurred on replacing the sequence with a 5′-CACGTG-3′ site showed how this flexibility was as important as the hydrogen bond capabilities. NMR studies[1.10] augmented the results obtained.

A significant development in the medicinal application of metallointercalators has been achieved with **1.6**. **1.6** has been used to site-specifically inhibit transcription factor binding.[1.11] Competition studies with yeast Activator Protein 1 (yAP-1) showed that a binding domain, containing both yAP-1 and **1.6** binding sites, was successfully competed for by **1.6**, at a concentration of 120 nM.

1.3 Sequence Selective Recognition of Double Helical DNA by Minor Groove Binding Polyamides

The minor groove of B-DNA presents itself as an attractive target for molecular recognition. It is both narrow and deep, and so its overall shape establishes the standards by which the A-T and G-C base pair information can be obtained from the floor of the groove. As such, small molecules that bind in the minor groove comply with the following rules set by the structural criteria of the minor groove.

A minor groove binder must have an overall shape complementarity with the groove. It follows that such binders are often crescent shaped or are pliable enough to follow the helical shape of the groove. It must also fit compactly into the narrow groove, thus providing an enhanced stabilisation by hydrophobic van der Waals contacts between the binder and the walls of the groove, and also acquiring the required base pair information from the floor of the groove by hydrogen bonding. Electrostatic attraction dictates that the binder should at least have a neutral charge, or preferably a positive charge to provide a generalised electrostatic stabilisation with the negatively charged backbone.

It has long been recognised that an organised layer of water molecules within the minor groove acts as the prime stabilising influence on B-DNA.[1.12] The spine of hydration provides bifurcated (three-centred) hydrogen bonds with the free phosphate oxygens on the backbones and free nitrogen and oxygen atoms on the edge of the base pairs, which help stabilise the structure. The entropy effect of displacing the water of hydration from the minor groove of the binder and also providing bifurcated hydrogen bonds with the DNA backbone and base pairs are factors that contribute to its selective binding. The binder must ultimately be capable of preferentially identifying the particular message associated with each of the canonical base pairs: A-T, T-A, C-G and G-C.

The current interest in minor groove binding originates from examination of natural products such as netropsin and distamycin A (**Figure 1.5a** and **b**).[1.13] Both bind preferably and reversibly to $(A\text{-}T)_4$ and $(A\text{-}T)_5$ sequences, respectively, within the minor groove. They bind through the required formation of bifurcated hydrogen bonds of the amide hydrogens with the adenine N-3 and the thymine O-2 atoms, as demonstrated in the case of netropsin in **Figure 1.6**.

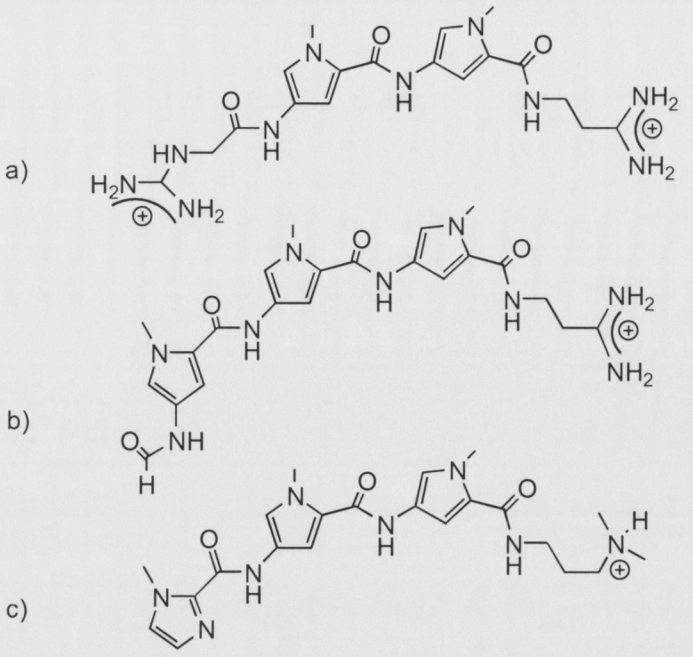


Figure 1.5 Structure of natural and synthetic DNA minor groove binders: a) netropsin; b) distamycin A and c) 1-methyl-imidazole-2-carboxamide-netropsin.

Their binding specificities for A-T rich regions in DNA can be explained in two ways: firstly, A-T base pairs are associated with a narrow minor groove and hence present a tighter fit for these elongated molecules; and secondly, the presence of the N-2 amino group of guanine serves as a major steric block, preventing the N-methyl pyrrolecarboxamide chain from docking fully to the floor of the minor groove in G-C rich segments. It is also known that the minor groove of A-T rich DNA possesses a high negative electrostatic potential, which is likely to influence the binding of these cationic species to DNA.

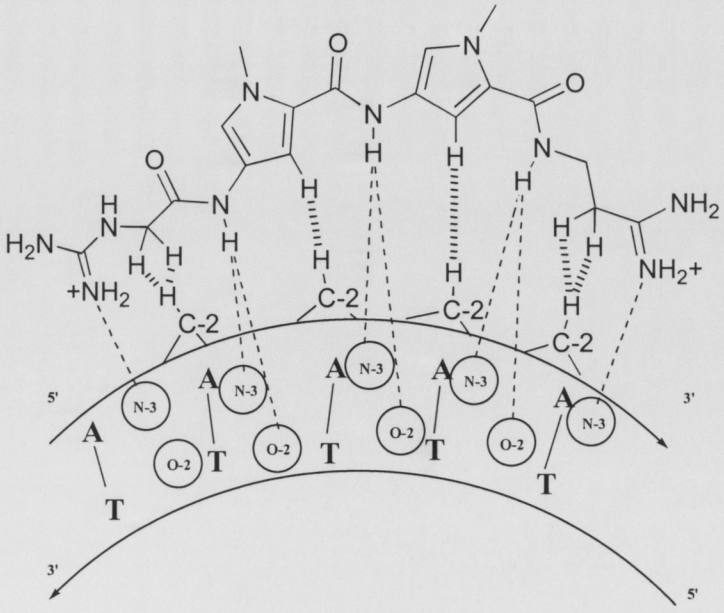


Figure 1.6 Interactions between netropsin and DNA (from crystal structure and NMR). Lown representation.

While it has been identified that regions of DNA that are unusually rich in A-T base pairs play key roles in the structure and function of the eukaryotic genome (i.e. that found in humans), it is necessary to achieve recognition of mixed sequences containing all four Watson-Crick base pairs in order to reach the ideal goal of targeting any desired sequence of double helical DNA.

It was proposed that the replacement of one or more of the pyrrole groups of netropsin by imidazole should alter the A-T preference and thus extend the recognition pattern to include G-C base pairs.[1.14] This is achieved through the specificity of the imidazole ring for guanine, which is related to the fact that the nitrogen lone pair of imidazole points to the floor of the minor groove and forms a specific hydrogen bond with the exocyclic amino group of guanine. Initial attempts, however, found that these lexitropsins (netropsin analogues) bound to DNA with less specificity and affinity than that of the parent netropsin, while recognising guanine in the target sequence.

The approach to the rational design of lexitropsins was advanced by the discovery by the group of Dervan in Caltech that the replacement of the terminal N-methyl pyrrolecarboxamide of distamycin A with N-methylimidazole-2-carboxamide formed N-methylimidazole-2-carboxamide-netropsin (**Figure 1.5c**), a ligand which bound the target 5′-TGTCA-3′ sequence in the minor groove in an unexpected fashion.[1.15] The study found that this ligand exhibited a particular preference for cooperative 2:1 binding, in which the lexitropsin adopted a side-by-side antiparallel dimeric motif, over 1:1 binding. The bulkier fit of the dimer apparently provided a better fit for the sequence-dependent width of the minor groove in sequences containing A-T and G-C base pairs, as G-C containing regions have inherently wider minor grooves.

As anticipated, the C-G/G-C selection is attained by the protruding guanine NH_2 group, which participates in specific hydrogen bonding with the imidazole N-3, with neighbouring pyrrole rings apparently positioning the imidazole for effective interaction and specific molecular recognition.

This outcome embarked these workers on the arduous task of finding a general solution for the selective recognition of significantly longer A-T rich, G-C rich and mixed A-T/G-C rich sequences in the minor groove of DNA. The next section [1.3.1-1.3.4] deals with a brief synopsis of the work achieved in this area.

1.3.1 Sequence Selective Minor Groove Binding of DNA by Hairpin Motif Polyamides

Contribution of Dervan *et al.*

The 2:1 peptide-DNA model (proposed by the complex formed by the imidazole-containing homodimer with the 5′-TGACT-3′ DNA tract) encouraged the use of a dimer containing two different peptides to recognise longer sequences. But it was immediately recognised that the formation of homodimers as well as heterodimers was possible within this set-up, along with each individual peptide being capable of molecular recognition.[1.15]

In order to attain the desired heterodimeric motif, it was deemed necessary to attach the two different peptides covalently in an antiparallel side-by-side manner (which allowed the elongated molecule to 'fold back' upon itself to form the dimer). The result resembled a U-shaped hairpin, hence the term 'hairpin motif'.[1.15]

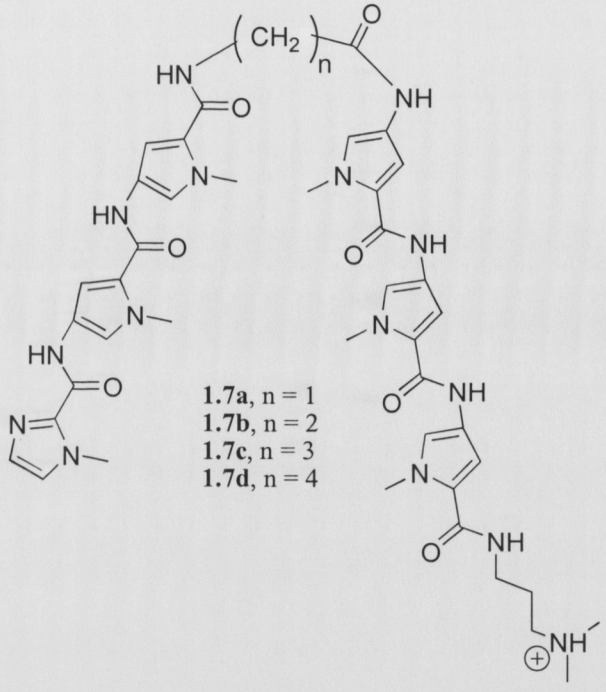
To test whether this new class of oligopeptide would give sequence-selective success, a variety of hexapeptides **1.7a-d** were synthesised. These hexapeptides consisted of linking the originally used 1-methylimidazole-2-carboxamide-netropsin (i.e. imidazole-pyrrole-pyrrole polyamides, or more succinctly, Im-Py-Py) with another three-ring polyamide unit (pyrrole-pyrrole-pyrrole, or Py-Py-Py) *via* a simple amino acid of differing carbon chain length, i.e. glycine (Gly), β-alanine (βala or β), γ-aminobutyric acid (GABA or γ) and 5-aminovaleric acid (Ava).

The binding affinities of these peptides to the target five base pair sites 5′-TTTTT-3′, 5′-TGTTA-3′, and 5′-TGACA-3′ on a 3 -^{32}P-end-labelled 135 base pair EcoRI/BsrBI restriction fragment were obtained *via* quantitative DNase I footprinting, and the results shown in **Table 1.2**.[1.16]

Table 1.2 Apparent first order binding constants K_a (M^{-1}) of peptides **1.7a-d** with target binding sites.

Peptide	Binding Site		
	5′-TTTTT-3′	5′-TGTTA-3′	5′-TGACA-3′
1.7a	$< 5 \times 10^5$	2.2×10^7	$\leq 6 \times 10^6$
1.7b	$< 5 \times 10^5$	$\leq 2 \times 10^6$	$< 5 \times 10^5$
1.7c	$< 5 \times 10^5$	7.6×10^7	3.2×10^6
1.7d	$< 5 \times 10^5$	$< 5 \times 10^5$	$< 5 \times 10^5$

The values gained indicated that the best success was achieved by the GABA-linked peptide, **1.7c**, which targeted the 5′-TGTTA-3′ sequence with the largest affinity (7.6×10^7 M^{-1}). Augmented specificity was also achieved by this peptide in obtaining a binding affinity of 3.2×10^6 M^{-1} for the single mismatch site 5′-TGACA-3′. The improvement in sequence specificities of assorted antiparallel side-by-side oligopeptides for DNA sequences was such that an estimation of increased affinity of hairpin linked oligopeptides versus the unattached peptide of at least two orders of magnitude was given. An examination of the position of the 'turn peptide' (γ) showed minimal effects on the specificity and affinity of the polyamides, indicating a new degree of flexibility within the 2:1 motif.[1.17] More recent studies have shown that substitution of the prochiral α-position of the γ-turn residue to provide (*R*)-2,4-diaminobutyric acid yielded chiral linked hairpins with enhanced DNA binding sequence specificity and orientation preference.[1.18]



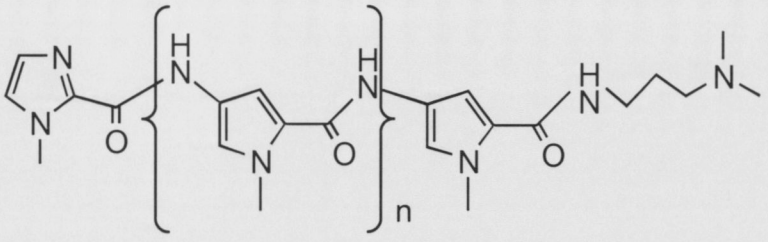
1.7a, n = 1
1.7b, n = 2
1.7c, n = 3
1.7d, n = 4

1.3.2 Sequence Selective Minor Groove Binding of DNA by Extended Polyamides

Contribution by Dervan *et al.*

In an attempt to elucidate the extent to which the 2:1 binding motif could be applied to longer binding sites, and also to determine its limitations, a series of six polyamides **1.8a-f** containing three to eight rings were synthesised. The binding site size was determined by MPE.Fe(II)

footprinting, and the apparent first order binding affinity and sequence specificities of each polyamide in binding to their respective five to ten base pair (bp) sites as antiparallel dimers was determined by quantitative DNase I footprint titration experiments,[1.19] and the results obtained shown in **Table 1.3**.



1.8a, n = 1
1.8b, n = 2
1.8c, n = 3
1.8d, n = 4
1.8e, n = 5
1.8f, n = 6

Table 1.3 Equilibrium association constants K_a (M^{-1}) and binding specificity for **1.8a-f** and DNA interactions.

Polyamide	Rings	Binding Site Size, bp	Hatch	Mismatch	Specificity
1.8a	3	5	1.3×10^5	$< 2 \times 10^4$	> 6.5
1.8b	4	6	8.5×10^6	1.6×10^6	5.3
1.8c	5	7	4.5×10^7	7.9×10^6	5.7
1.8d	6	8	5.3×10^7	$< 2 \times 10^7$	> 2.7
1.8e	7	9	4.7×10^7	1.7×10^7	2.8
1.8f	8	10	$< 2 \times 10^6$	$< 2 \times 10^6$	≈ 1

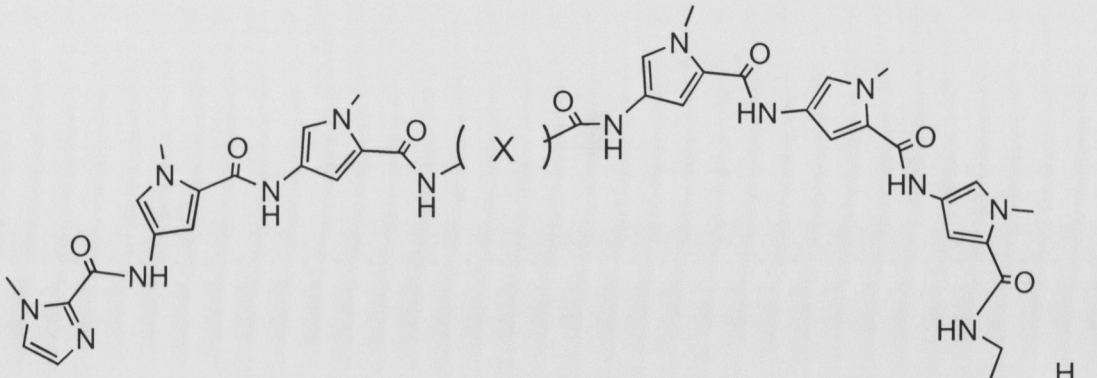
The results of the analysis showed that these pyrrole-imidazole containing polyamides of three to seven rings could specifically recognise sequences up to nine bp long, thus setting a new lower limit on the sequence length that could be targeted by imidazole-pyrrole polyamides.

Also, it was found that, as the length of polyamide dimer having the general sequence $ImPy_{2\text{-}6}$ increased beyond five rings (corresponding to a seven base pair binding site), the DNA binding affinity ceased to increase with polyamide length. This indicated that the N-methylimidazole (Im) and N-methylpyrrole (Py) residues failed to maintain the appropriate base-pair register across the entire length of the polyamide-DNA complex.

The structural basis for this observed drop in affinity was provided by X-ray crystal structures of a four-ring homodimer (imidazole-imidazole-pyrrole-pyrrole) in complex with the DNA sequence 5′-CCAGGCCTGG-3′, which revealed a perfect match of polyamide rise-per-residue with the pitch of the DNA double strand, but an overwound ligand curvature.[1.20]

These results suggested that extension of the 2:1 dimer:DNA motif to sequences longer than nine bp required a new class of polyamide. It was predicted that a more flexible amino acid spacer in place of an imidazole or pyrrole subunit would permit the antiparallel dimer to reset the register to allow sustained affinity and specificity in longer DNA sequences.

The discovery that β-alanine provided the desired 'spring-like' quality was achieved serendipitously through the study of the binding conformations formed by the four polyamides **1.9a-d** with the nine bp sequence 5′-TGTTAAACA-3′ and the two thirteen bp targets 5′-A_5GACA_5-3′ and 5′-ATATAGACATATA-3′. The polyamides synthesised were of the form ImPyPy-X-PyPyPy-Dp, where X = pyrrole (Py), glycine (G), β-alanine (β) and γ-aminobutyric acid (γ), and Dp = dimethylaminopropylamino group. The binding affinities of each polyamide dimer was determined by quantitative DNase I footprinting titration experiments for the three target sites on a 281 bp restriction fragment.[1.21]



1.9a: X = pyrrole
1.9b: X = CH_2
1.9c: X = $(CH_2)_2$
1.9d: X = $(CH_2)_3$

The association constants obtained for **1.9a**, **1.9b**, and **1.9d** to a nine bp segment were approximately the same (~1 X 10^8 M^{-1}). However, the binding affinity of the β-alanine linked peptide **1.9c** was around eight times higher, implying that β-alanine residues reset the optimum fit of the polyamide dimer with the DNA helix at long binding sites. The binding of the polyamide dimer to a target 5′-TGTATATCA-3′ was directly characterised by NMR, and restrained molecular modelling studies indicated similar results of highest affinity and also sequence specificity for the β/β complex.[1.22]

The original study of these polyamides on the nine bp site and the two thirteen base pair sites uncovered other interesting results. All polyamides bound to all sequences tested to varying extents. It was found that the nine base pair target was bound *via* the conventional 2:1 motif, in an "overlapped" binding mode. However, the thirteen base pair sequences were bound by an extended conformation. Here, the intermolecular dimer is formed by the two ImPyPy

units of each of the two polyamides in the dimer binding a central 5′-AGACA-3′ sequence in a 2:1 motif, and the PyPyPy fragment of the same polyamides bind the flanking A-T tracts in the similar 1:1 binding mode adopted by DNA complexes of distamycin, which overall gives a "slipped" motif (**Figure 1.7**).

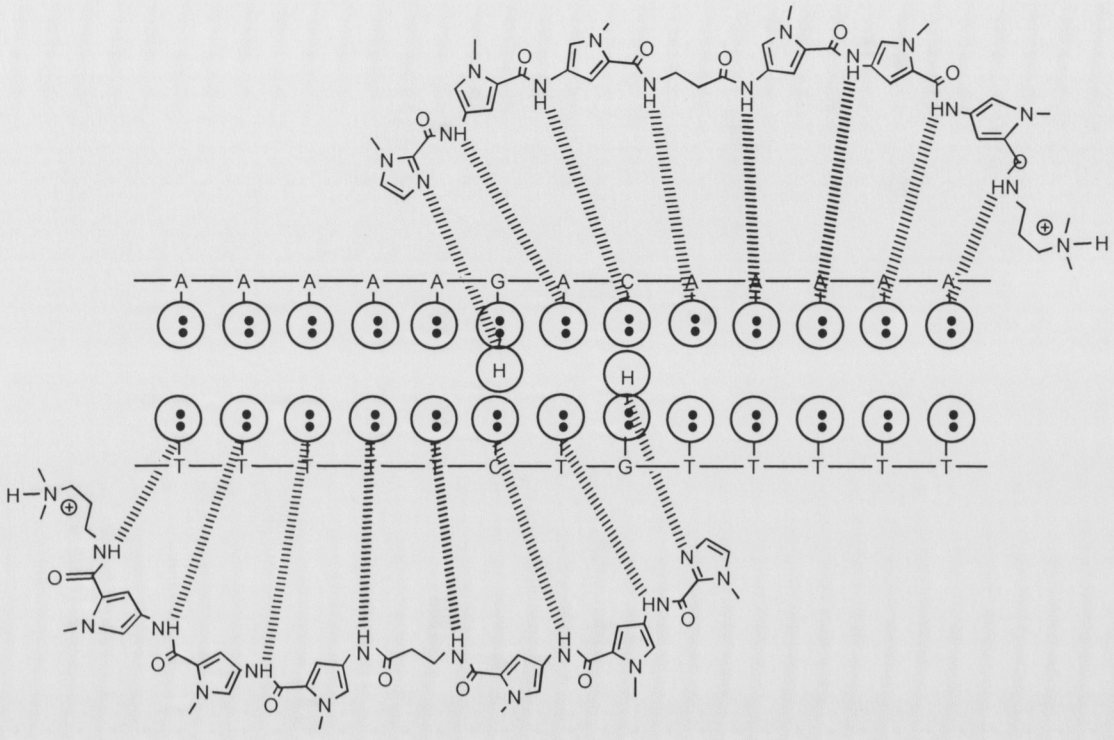


Figure 1.7 Schematic of 'slipped' binding mode of **1.9c** in the minor groove of a DNA duplex. Circles with double dots represent lone pairs on the N-3 of purines and the C-2 carbonyl of pyrimidine. Circles containing a H represent the exocyclic amino hydrogen of guanine.

It is observed that in the nine base pair "overlapped" and the thirteen base pair "slipped" binding sites, the G-C and C-G base pairs are separated by one and five A-T base pairs, respectively. Further complications in binding modes are possible as well, as the potential binding sites of these polyamides of ten, eleven and twelve base pair sites can also give rise to a "partially slipped" motif. Here, the G-C and C-G base pairs are separated by two, three and four A-T base pairs, respectively. It is obvious that polyamides that are capable of binding DNA in several ways are less specific.

Apart from the multiplicity in binding modes now available when pursuing this approach of matching longer DNA sequences with larger synthetic polyamides, another consideration must be given, as the upper limit of polyamide size with regard to efficient cell permeation is not known. Perhaps a more biomimetic method would be to target longer DNA sequences with polyamides of moderate size. A possible means of doing so would be through the use of cooperative hairpin dimers.[1,23] These extended polyamides self-associate in a pre-

organised manner in which the ligands slip sideways with respect to each other to permit recognition of other sequences, and so doubles the binding site size possible, without increasing the molecular weight of the binder.

Molecules **1.10** and **1.11** were then synthesised to test this hypothesis. Both molecules contain the required Im and Py ring sequences for molecular recognition, a γ turn to achieve the hairpin motif, and a β residue to alleviate curvature strain, but contain a different number of Py residues. The design is such that each polyamide is self-complementary, and so can form the self-associated cooperative dimer required. The binding affinities of both molecules were then determined by DNase I footprinting titration experiments on a 245 base pair restriction fragment containing the three target sites 5′-AGCAGCTGCT-3′, 5′-AGATGCTGCA-3′ (single mismatch site) and the twelve bp site 5′-AAGCAGCTGCTT-3′ (double mismatch site). It was found that **1.10** bound the target site with an association constant of 1.9×10^8 M^{-1}, and the single base pair mismatch site with a nine-fold lower affinity. **1.11** was found to bind the twelve bp site with a ten-fold higher affinity than **1.10**.

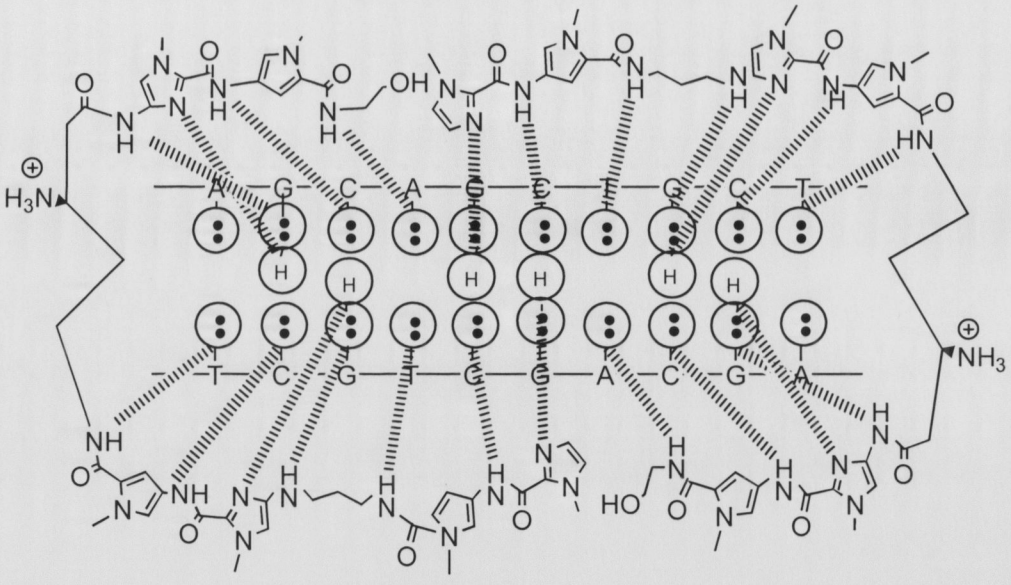
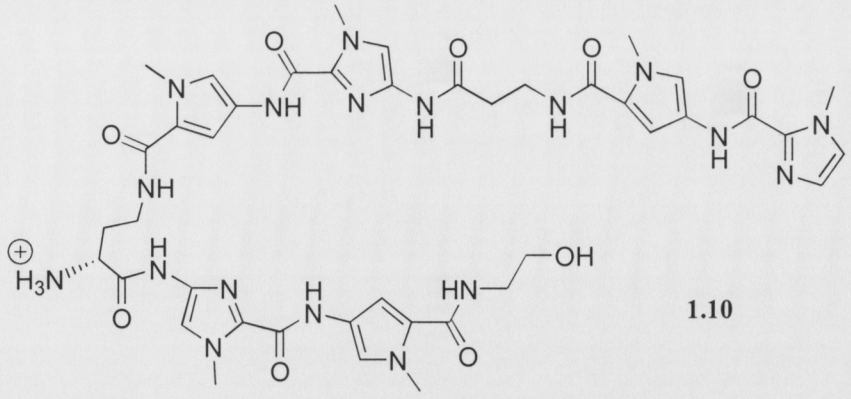
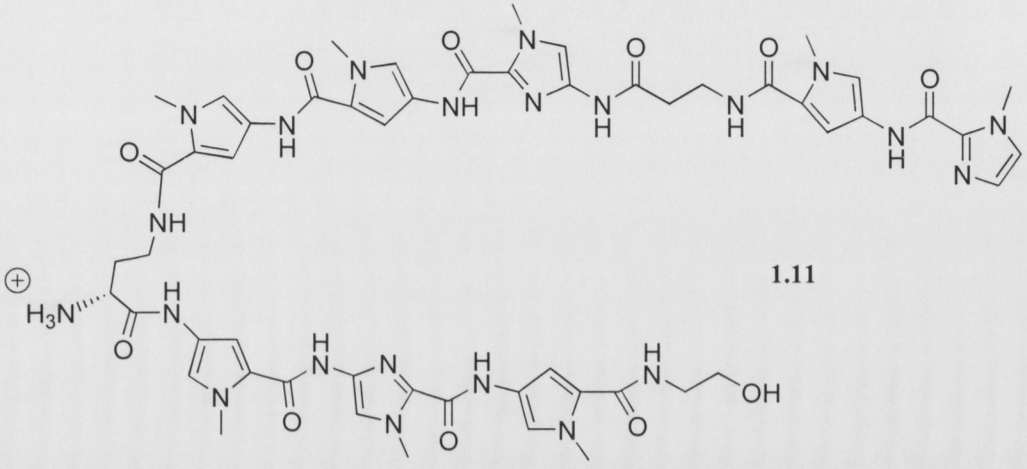


Figure 1.8 Schematic of a hairpin dimer of **1.10** binding in the minor groove of a DNA double helix. Circles with dots represent lone pairs on N-3 of purines and C-2 carbonyl of pyrimidines. Circles containing a H represents the exocyclic amino hydrogen of guanine.

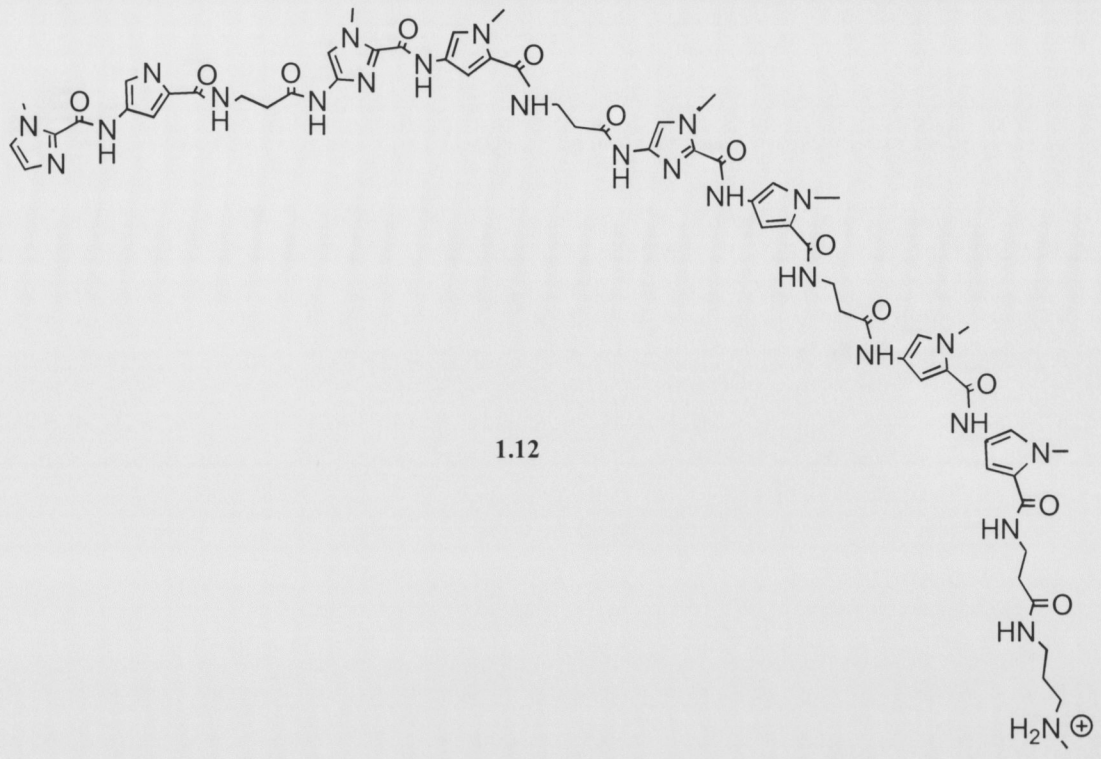


1.10



1.11

A milestone was reached a few years ago in the development of chemical approaches to sequence selective DNA recognition. An eight ring polyamide **1.12** was designed to form a cooperative antiparallel dimer, which self-associates to recognise a sixteen base pair segment 5′-ATAAGCAGCTGCTTTT-3′ *via* the "slipped" motif.[1.24] As a run of sixteen base pairs represents the minimum single site size present within the human genome (based on the statistic number of base pairs present), recognition of this site size offers at last a real opportunity to interfere with gene expression. The precise binding site location of the polyamide on the target site were determined by quantitative DNase I footprint titration experiments on a 245 base pair restriction fragment. The results obtained revealed that target binding was achieved with an association constant of greater than 3.5×10^{10} M^{-1}. Further evidence provided by affinity cleavage studies showed that the polyamide-DNA complex is bound by a fully overlapped core, containing the central 5′-GCAGCT-3′ tract and six Im-Py pairs, while the flanking A/T base pairs are bound in the 1:1 motif by the remaining two β and two pyrrole residues on each polyamide.



1.12

While this preliminary result offered much in the way of encouragement, further investigation revealed that mismatch sites were recognised by the polyamide with only ten to twenty fold lower affinities relative to the target binding site. Efforts are now being employed to achieve optimum sequence specificity at this level.

1.3.3 Formation of a Complete Pairing Code by Minor Groove Binding Polyamides

Contribution by Dervan *et al.*

It was recognised early on in this work on minor groove DNA recognition by extended imidazole-pyrrole polyamides that pairing rules could be developed which assist the design of these sequence selective binders in a rational manner. An antiparallel pairing of imidazole opposite pyrrole (Im/Py pair) distinguishes G-C from C-G and both of these from A-T/T-A base pairs. A Py/Py pair recognises both A-T and T-A from G-C/C-G, but does not distinguish between A-T and T-A.

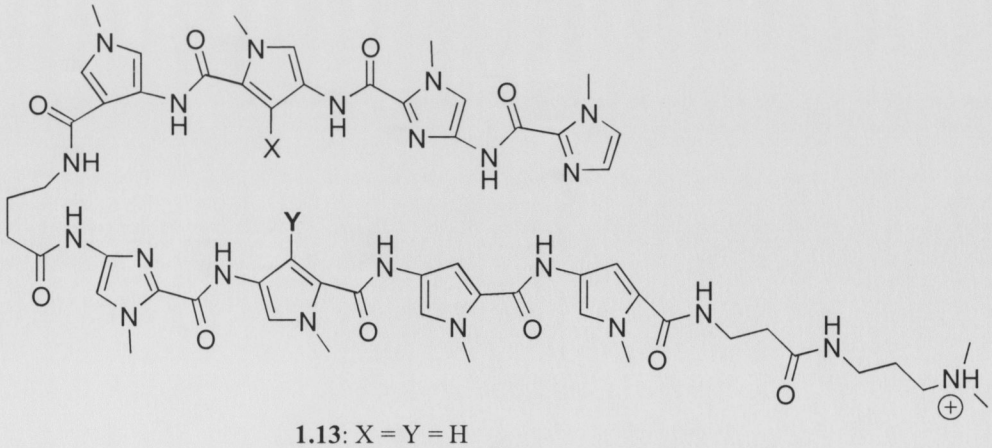
This lack of discrimination between A-T/T-A base pairs, due to the almost identical placement of the hydrogen bond accepting N-3 of adenine and O-2 of thymine in the minor groove, constituted a severe constraint in the development of the polyamide approach to minor groove recognition.

The realisation that single atom substitution at position-3 on the aromatic amino acid could complement small structural differences at the edges of the base pairs in the minor groove led the way forward to overcoming the obstacle to A-T/T-A discrimination. Examination of molecular models of the A-T base pair revealed that discrimination may be achieved by means of shape selection of an asymmetric cleft on the floor of the minor groove formed by the O-2 of thymine and C-2 of adenine. Discrimination also arises through the fact that the adenine nitrogen is in possession of only one lone pair of electrons, and so can only accept one hydrogen bond, while the two lone pairs on the O-2 of thymine are capable of forming two specific hydrogen bonds with two suitable donors.

The replacement of the C3-H with a C3-OH on the pyrrole ring to give 3-hydroxypyrrole (Hp) was proposed as an alternative aromatic amino acid to break the A-T/T-A degeneracy when paired with another pyrrole to form a Hp/Py pair. The bulky hydroxyl group would preferentially select the O-2 of thymine over the sterically hindered adenine N-3, as it is blocked by the C2-H. Further stabilisation of the hydroxypyrrole ring interaction with thymine is achieved by the formation of two specific hydrogen bonds with the thymine O-2 lone pairs and the 3-hydroxy and 4-carboxamido groups on Hp.

Experimental support of the Hp/Py hypothesis was obtained through the synthesis of polyamides **1.13**, **1.14** and **1.15**. A 250 base pair DNA restriction fragment containing the target sequences 5′-TGGACA-3′ and 5′-TGGTCA-3′ was used to determine equilibrium dissociation constants (K_D) of these hairpin polyamides by quantitative DNase I footprinting.[1.25]

Interest was in differentiating between the A-T base pair and the T-A base pair, respectively, in the fourth position of each target DNA tract, and the results displayed in **Table 1.4**. Polyamide **1.13**, designed to locate a Py/Py pair in the proximity of the A-T/T-A fourth position, showed no outright preference for either target site, as was expected, and bound each sequence moderately, with only a two fold partiality for T-A.



1.13: X = Y = H
1.14: X = H; Y = OH
1.15: X = OH; Y = H

Table 1.4 Equilibrium dissociation constants K_D (nM) of polyamides **1.13**, **1.14**, and **1.15** and target nucleotides.

Polyamide		5′-TGGTCA-3′	5′-TGGACA-3′
1.13	Py/Py	0.077	0.15
1.14	Py/Hp	15	0.83
1.15	Hp/Py	0.48	37

However, a dramatic difference was observed with polyamides **1.14** and **1.15**, where Hp/Py pairs were placed opposite the A-T/T-A base pairs, but in inverse order in each case. Polyamide **1.14**, with the pair order Hp/Py, was found to have a seventy-seven fold difference between dissociation binding constants in favour of the target site 5′-TGGACA-3′ relative to the 5′-TGGACA-3′ site. The Py/Hp pair arrangement of polyamide **1.15** favours the 5′-TGGTCA-3′ site with a K_D of 15 nM as compared to 0.83 nM for the other target site. A crystal structure of the polyamide•DNA complex ImHpPyPy-β-Dp•5′-CCAGTACTGG-3′ was obtained to verify the structural basis for why the Hp/Py pair discriminates for T-A over A-T.[1.26]

Further work was carried out on these polyamides to elucidate their preference for A-T/T-A base pairs over G-C/C-G pairs. As well as the original six base pair target sites, 5′-TGGCCA-3′ and 5′-TGGGCA-3′ were added to investigate the relative binding affinities of the sixteen possible complexes (although twelve cases are identified in reality as the Py/Py pair is degenerate). Equilibrium association constants were obtained from quantitative DNase I footprint titrations on an appropriate plasmid containing all four six base pair recognition sites, giving the results exhibited in **Table 1.5**.[1.27]

Table 1.5 Equilibrium association constants K_a (M^{-1}) of polyamide pairs and the target oligomers.

Polyamide Pair	5′-TGGTCA-3′	5′-TTGACA-3′	5′-TGGCCA-3′	5′-TGGGCA-3′
Py/Py	4.8×10^9	2.8×10^9	2.8×10^8	9.0×10^7
Hp/Py	9.8×10^7	1.1×10^9	3.3×10^7	2.5×10^7
Py/Hp	1.6×10^9	8.1×10^7	7.9×10^7	5.5×10^7

The results above confirmed the selectivities expected. The near-equal affinities of the Py/Py pair for T-A/A-T was as before, while the dramatic difference in affinity when placed opposite G-C/C-G pairs is likely to be explained by the steric presence of the exocyclic amino group of guanine preventing the polyamide fitting deep within the minor groove. The Hp/Py containing polyamide was found to achieve the highest binding affinity of all, with preferential

binding to the 5'-TGGTCA-3' tract, and a greater than twenty fold lower binding constant found for the other three target sites.

This ultimately implies that the Hp/Py pair is the prime recognition element of the T-A base pair in polyamides to date. Likewise, the Py/Hp pairing selects the A-T containing site with an eleven fold affinity relative to the T-A containing site. It appears that a complete pairing code for the recognition of all four Watson-Crick base pairs in the minor groove of double helical DNA is now available.

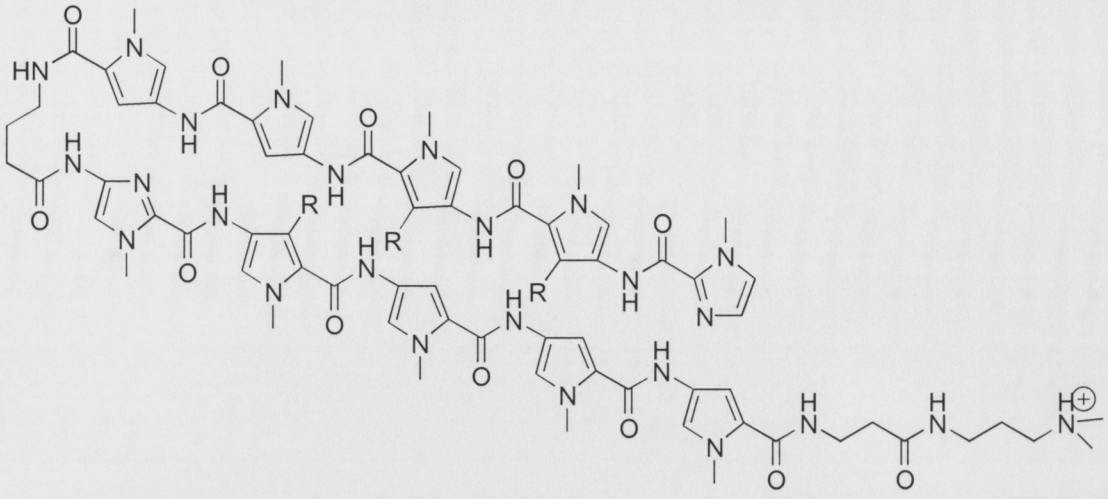
1.3.4 Sequence Selective Minor Groove Recognition by Second Generation Polyamides

Contribution of Dervan *et al.*

While it was acknowledged that all the pieces were by now in place for complete and selective recognition of pre-determined DNA sequences, the results obtained in terms of binding affinity and sequence specificity of several target sites implied that there is much scope for design improvement.

The identification of the Hp/Py pair represented a major leap forward in terms of DNA recognition. However, it had been noted that replacement of a single Py/Py pair by Hp/Py resulted in a five-fold destabilisation of an eight-ring hairpin polyamide for an identical match site. The incorporation of multiple Hp/Py pairs would be necessary to achieve recognition of any chosen sequence.

A ten ring hairpin polyamide **1.16** was designed to bind selectively to a seven base pair target site 5'-TGTTACA-3' and its binding affinity determined by quantitative DNase I footprinting titration experiments on a 361 base pair restriction fragment.[1.28] The mismatch sites on 5'-TGTATCA-3' were also examined. The results revealed that the polyamide bound the target site with a 94-fold greater affinity, again confirming the T-A/A-T discrimination. More importantly, the polyamide containing three Hp/Py pairs displayed only an eleven-fold decrease in stability relative to polyamide **1.17** containing Py/Py pairs instead of Hp/Py pairs at these points.



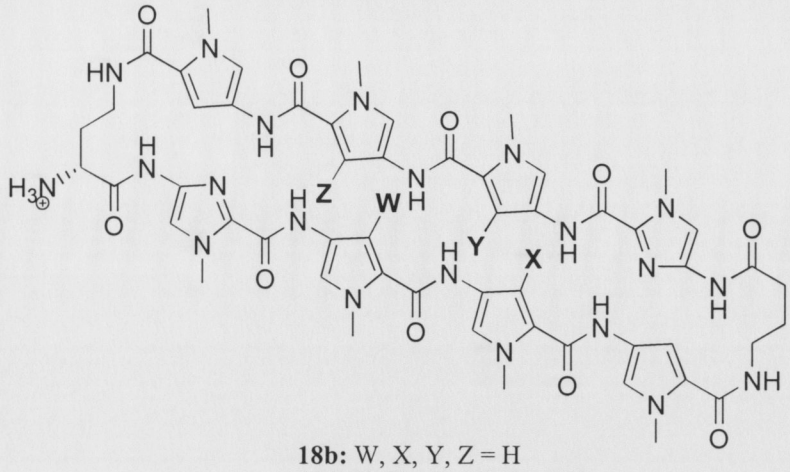
1.16: R = OH
1.17: R = H

Further efforts to enhance the binding properties of polyamides containing Hp/Py pairings were employed through the synthesis of four eight-ring cyclic polyamides **1.18b**, **1.18d**, **1.18f** and **1.18h** (cyclic polyamides are formed by the covalent connection of the carboxy and amino ends of a hairpin polyamide) and their corresponding hairpin analogues **1.18a**, **1.18c**, **1.18e** and **1.18g**, respectively, as controls. Equilibrium association constants (K_a) were determined by quantitative DNase I footprinting for the eight polyamides on a restriction fragment containing the three six base pair target sites 5′-AGAACT-3′, 5′-AGTACT-3′ and 5′-AGATCT-3′, and the results shown on **Table 1.6**.[1.29]

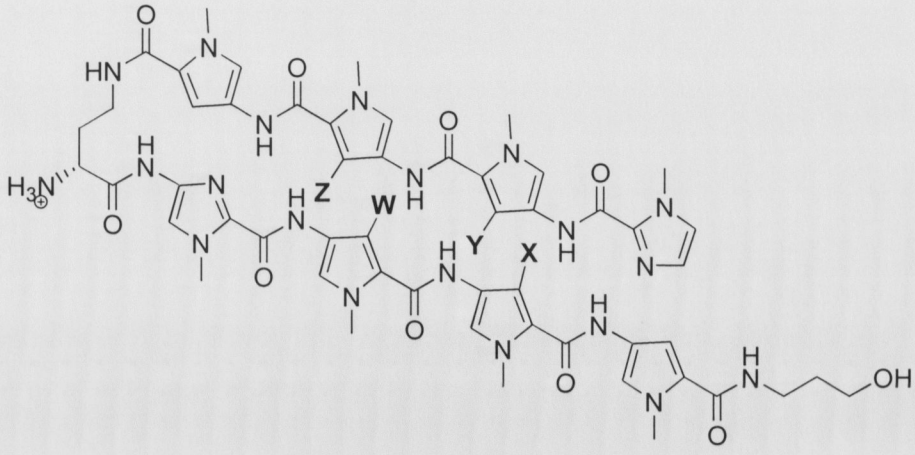
Table 1.6 Equilibrium association constants K_a (M^{-1}) and relative specificities of polyamides **1.18a-h** and the target DNA tracts.

Polyamide	5′-AGAACT-3′	R.S.	5′-AGTACT-3′	R.S.	5′-AGATCT-3′	R.S.
1.18a	1.0×10^{10}	2.5	2.1×10^{10}	5.4	4.0×10^{9}	1
1.18b	4.0×10^{10}	8.0	7.0×10^{10}	14	5.0×10^{9}	1
1.18c	4.4×10^{6}	4.4	1.4×10^{6}	1.4	$\leq 1.0 \times 10^{6}$	1
1.18d	7.5×10^{7}	75	1.4×10^{7}	14	$\leq 1.0 \times 10^{6}$	1
1.18e	2.6×10^{7}	2.6	7.0×10^{8}	70	$\leq 1.0 \times 10^{7}$	1
1.18f	3.5×10^{8}	3.5	3.2×10^{9}	320	$\leq 1.0 \times 10^{7}$	1
1.18g	2.0×10^{7}	1	5.4×10^{9}	2.7	2.7×10^{7}	1.4
1.18h	3.5×10^{7}	2.5	1.4×10^{7}	1	6.2×10^{7}	4.4

where R.S. stands for Relative Specificity



18b: W, X, Y, Z = H
18d: Y, Z = H; W, Z = OH
18f: X, Z = H; W, Y = OH
18h: W, Y = H; X, Z = OH



18a: W, X, Y, Z = H
18c: Y, Z = H; W, Z = OH
18e: X, Z = H; W, Y = OH
18g: W, Y = H; X, Z = OH

The target sites bound were match sites for polyamides **1.18a** and **1.18b**, while one of the three binding sites represented a match, and the other two sites were single or double base pair mismatches with respect to the Hp/Py pairing for polyamides **1.18c-h**.

The parent polyamides **1.18a** and **1.18b** showed a slight preference for the 5′-AGTACT-3′ sequence, with the cyclic polyamide **1.18b** binding with a higher affinity. Polyamides **1.18c** and **1.18d**, which contained two contiguous Hp/Py pairs and so bound 5′-AGAACT-3′ only as a match site, bound the target site with a five- to seventy five-fold preference, and again, cyclic polyamide **1.18d** achieved a sixteen fold binding increase on its hairpin analogue.

Target site 5′-AGTACT-3′ was bound by a ten-fold and 320-fold affinity in preference to its single and double mismatch sites, respectively, by cyclic polyamide **1.18f**, where a staggered Hp residue pattern towards the 5′-direction (relative to the DNA helix) was displayed. Here again, **1.18f** afforded a higher association constant than its hairpin version, **1.18e**. An

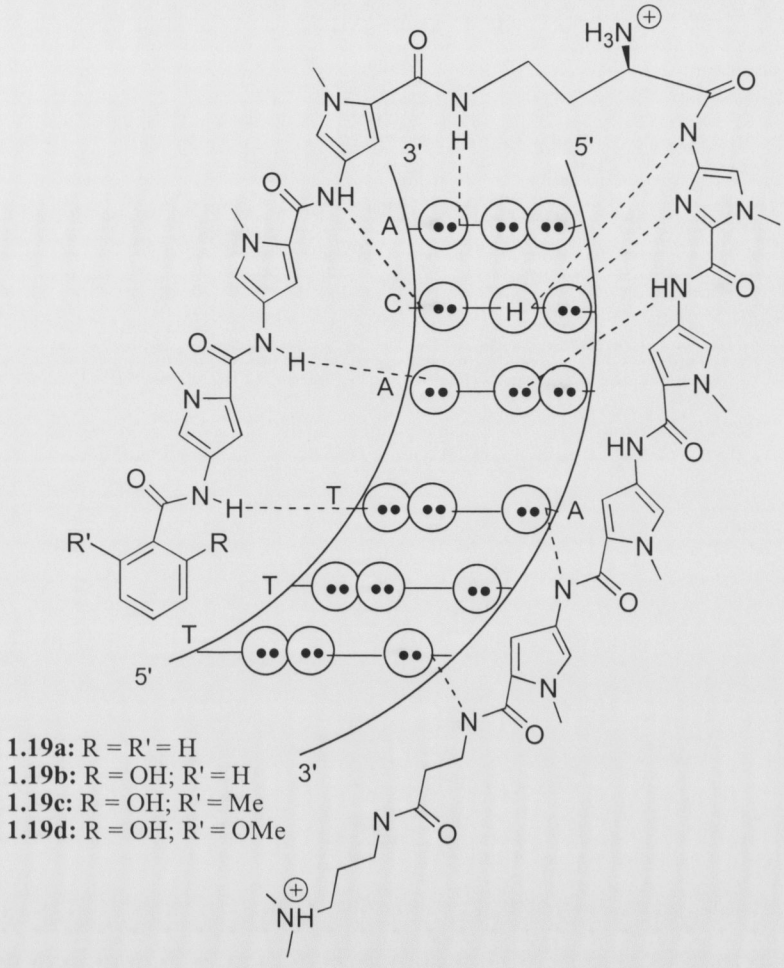
anomaly existed in polyamide **1.18g**, where the target site of 5′-AGATCT-3′ was bound with a nearly-equal and a two-fold decreased affinity than its single and double mismatch sites, respectively. The cyclic twin **1.18h**, however, bound preferentially to the target site, and with increased affinity.

Another limitation in the use of the Hp/Py pair is that evidence has shown that a terminal Hp/Py pair does not display the same specificity, selectivity and affinity characteristics for A-T/T-A base pairs, thus restricting success for Hp/Py to internal positions. A hypothesis based on earlier success found from six-membered aromatic rings[1.15] resulted in the synthesis of four eight-ring hairpin polyamides **1.19a-d**. Four DNA target sites of the form 5′-TNTACA-3′ (where N = T, A, G, C) were used in quantitative DNase I footprint titration experiments with the four polyamides to determine their binding affinities, as seen in **Table 1.7**.[1.30]

Table 1.7. Equilibrium association constant K_a (M^{-1}) of polyamides **1.19a-d** with the four target oligonucleotides.

Polyamide	5′-TTTACA-3′	5′-TATACA-3′	5′-TGTACA-3′	5′-TCTACA-3′
1.19a	1.9×10^{11}	1.2×10^{11}	3.5×10^{10}	1.7×10^{10}
1.19b	1.9×10^{10}	1.9×10^{10}	6.3×10^{9}	3.7×10^{9}
1.19c	1.0×10^{10}	1.0×10^{10}	1.5×10^{9}	$\leq 1 \times 10^{8}$
1.19d	2.6×10^{10}	8.4×10^{9}	4.7×10^{8}	$\leq 1 \times 10^{8}$

It appeared as if there is no observable discrimination between A-T and T-A, and G-C and C-G base pairs for polyamides **1.19a** and **1.19b**, and a slight preference for G-C over C-G in polyamide **1.19c** for the target sites used. However, a three-fold specificity for 5′-TTTACA-3′ (T-A) over 5′-TATACA-3′ (A-T) was observed for polyamide **1.19d**, as well as sequences 5′-TGTACA-3′ and 5′-TCTACA-3′ (placing G-C and C-G base pairs, respectively, at the relevant positions) afforded a decreased affinity of fifty and 200-fold. This implied a 2-hydroxy-6-methylbenzamide (Hb-3)/Py preference of T > A >> G > C, and so provides another avenue for exploration. It is noteworthy that the complete recognition code relies on the 2:1 binding motif. A comparable pairing code does not exist for 1:1 binding as of yet. A study by Dervan and Urbach elucidated that Hp/Py does indeed select a T-A base pair *via* this 1:1 motif,[1.31] but structural studies are pending.

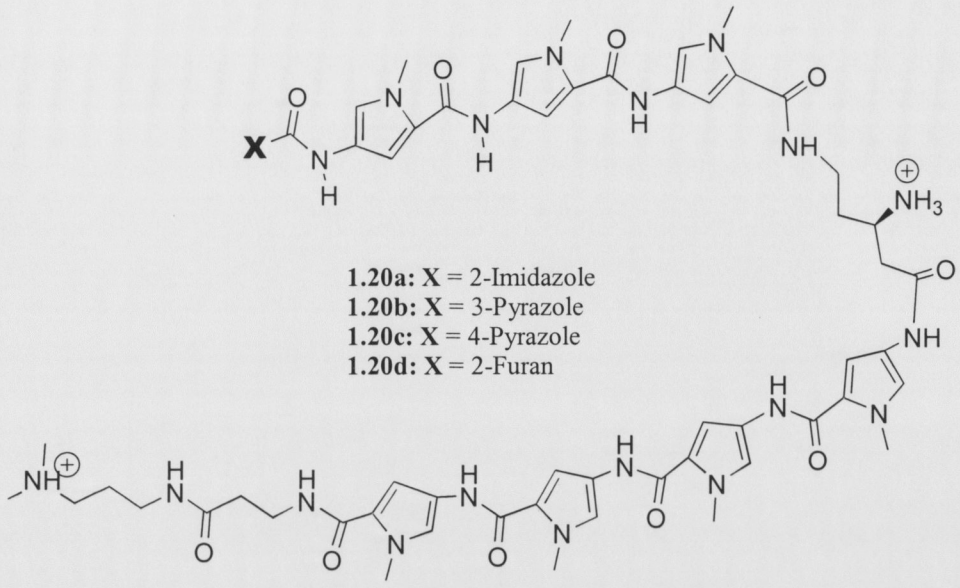


1.19a: R = R' = H
1.19b: R = OH; R' = H
1.19c: R = OH; R' = Me
1.19d: R = OH; R' = OMe

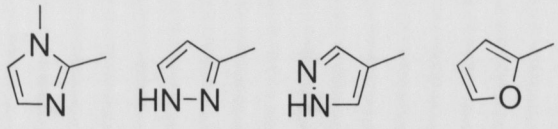
Dervan has continued to expand the repertoire of DNA recognition elements by varying established design features to uncover their roles, if any, and also modifying these features, if necessary. For instance, the discovery that the N-methyl substituent is necessary to attain optimum specificity,[1.32] and although the β-alanine subunit proves itself to be useful in resetting the register in a variety of locations along the hairpin polyamide,[1.33] the A-T/T-A degeneracy involved in the β/β pairing in the hairpin motif is solved using an isoserine/β pair.[1.34]

His present course of action dwells on the search for new heterocyclic systems to expand his recognition reserve even further. The ability of the nitrogen lone pair of imidazole to form a specific hydrogen bond with the protruding amine of guanine as a result of its directional capacity inspired the examination of other five-membered heterocycles which could position an electron-rich heteroatom towards the floor of the groove. A quantitative DNase I footprinting titration study on a 283 bp DNA plasmid containing four six-bp oligomers 5′-TNCCTA-3′ (N = G, C, A, T) was carried out on four eight-ring polyamides **1.20a-d**. Their binding affinities and specificities were thence obtained.[1.35] The most significant result obtained was by ligand **1.20c**, containing the 3-pyrazole carboxylic acid (3-Pz)/Py pair, and the target site 5 -TGCCTA-3 with

K_a = 1.7 X 10^{11} M^{-1}. This result displays a preference of a 3-Pz/Py pair for G-C base pairs with an enhanced affinity and specificity as compared with Im/Py pairs.

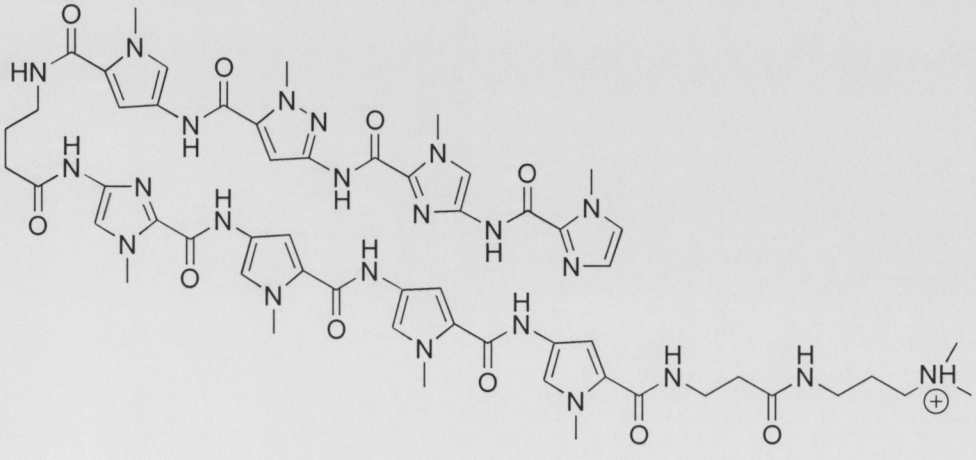


1.20a: X = 2-Imidazole
1.20b: X = 3-Pyrazole
1.20c: X = 4-Pyrazole
1.20d: X = 2-Furan



1.20a **1.20b** **1.20c** **1.20d**

A study of N-methylpyrazole (Pz)/Py pairs by means of quantitative DNase I footprinting titration experiments to elucidate association constants, and binding sites with MPE•Fe(II) footprinting, showed that an increased binding specificity for T-A over A-T base pairs is achieved by Pz/Py pairs relative to Py/Py pairs in targeting the molecule **1.21** with test sites 5′-TGGTCA-3′ (K_a = 2.0 X 10^9 M^{-1}) and 5′-TGGACA-3′ (K_a = 1.0 X 10^9 M^{-1})[1.36] and a 100-fold increase in specificity over C-G and G-C mismatch sites (K_a = 2 X 10^7 M^{-1}).



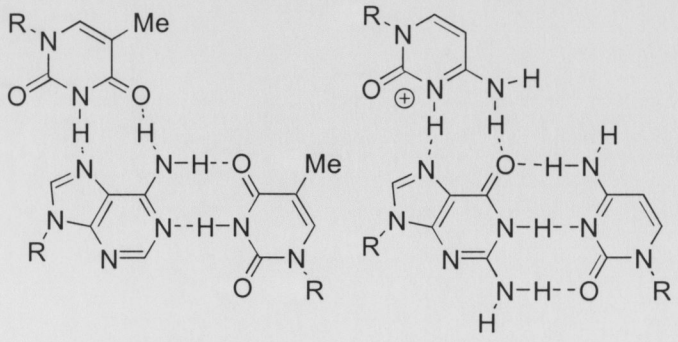
1.21

1.4 Sequence Selective Recognition of Double Helical DNA by Major Groove Binding Triplex Forming Ligands

Specific recognition of double helical B-form DNA has long been predicted to occur preferentially in the major groove. Control proteins and promoter repressors appear to bind largely *via* this groove. The structural factors that promote the recognition preferences are the easy access of the base pairs due to the wide nature of the groove and the wealth of information in terms of hydrogen bonding sites on the floor of the helix groove. These features clarify why the major groove should be the preferred target of DNA-binding molecules of very high or absolute specificity, and why this area has been the subject of an abundance of synthetic chemistry over the past few decades.

The only nucleic acid binding ligands known to date that achieve such a high degree of specificity are nucleic acids themselves. Soon after the elucidation of the double helix itself, it was shown that a triple helix could be formed by polynucleotides.[1.37] Thirty years later, Dervan and Hélène showed that short synthetic DNA oligonucleotide strands could bind sequence selectively in the major groove of duplex DNA, and hence could be used as DNA readout agents.[1.38] The triple helix pattern was found to form by two motifs. The two motifs are differentiated by the orientation and base composition of the third strand, but in both motifs, the third strand lies in the major groove, where it undergoes specific interactions with the substituents on the duplex bases. These motifs are called the pyrimidine and the purine motifs.

In the pyrimidine motif, pyrimidine oligodeoxyribonucleotides bind parallel to the purine Watson-Crick strand through the formation of two specific hydrogen bonds to the purine Watson-Crick base. The hydrogen bonding interactions involved in triple-helix formation are often referred to as Hoogsteen hydrogen bonds. Specificity is derived from thymine recognition of A-T base pairs and protonated cytosine recognition of G-C base pairs, as seen in **Figure 1.9** (in the following, the notation X•Z-Y denotes a triplet in which the third strand base X interacts with the Z-Y base pair forming hydrogen bonds to base Z).



T•A-T **C^+•G-C**

Figure 1.9 T•A-T and C^+•G-C Base Triplets.

Protonation of the N-3 of cytosine can only take place at about pH 5.0, thus placing a severe practical limitation on C^+•G-C formation in biological systems. Although C^+•G-C triplets have long been shown to be more stable than T•A-T triplets,[1.39] contiguous C^+•G-C triplets are destabilising due to charge repulsion, which places another practical constraint on the triplex approach. Another important aspect of the pyrimidine motif is the stabilisation of the triplex due to isomorphism. That is, the three C -1 atom positions on each of the T•A-T and C^+•G-C triplets are superimposable.

A third strand of purine rich oligodeoxyribonucleotides binds in the major groove of DNA in an antiparallel fashion, hence its name "the purine motif". Specificity is derived from G recognition of G-C base pairs (G•G-C) and A or T recognition of A-T base pairs (A•A-T or T•A-T). The thymine recognition of the A-T base pair in the purine motif differs from the pyrimidine motif by rotating the thymine base 180 ° with respect to the recognition sequence used in the pyrimidine motif.

The advantages of isomorphism are not seen in the purine motif (**Figure 1.10**). The G•G-C and A•A-T triplets are not isomorphous, and as a result, A•A-T triplets are only present in homopolymers. Also G•G-C and T•A-T are not isomorphous. In fact, both of these latter triplets are capable of forming both pyrimidine and purine binding patterns, although thymine prefers to bind A-T *via* the pyrimidine motif, and guanine adopts the purine motif in the G •G-C triplet. A poly (dG-T) third strand will therefore adopt a purine motif only if there are sufficient G residues in the third strand. Importantly, the purine motif is pH independent. As such, it would appear to be the more feasible approach to *in vivo* applications. However, the presence of monovalent cations in physiological conditions promotes the formation of quadruplexes instead of triplexes in G-rich sequences.[1.40]

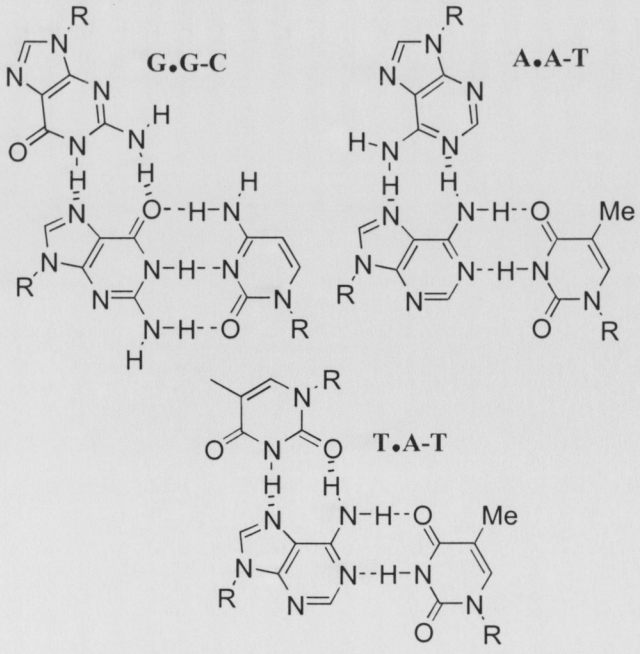


Figure 1.10 G•G-C, A•A-T and T•A-T Base Triplets in the Purine Motif.

And so, despite their superb sequence recognition properties, there are several outstanding problems that must be overcome if triplex-based technologies are to be of use in the future. Along with the pH-dependency of C^+•G-C triplets, the low stability of triplexes compared with their duplex counterparts, poor cellular uptake and nuclease instability, both pyrimidine and purine triple helical motifs provide limited recognition codes and specify mostly purine tracts of double helical DNA. Apart from the requirement to maintain the backbone geometry, this limitation arises from the recognition mechanism that involves the purine's two remaining Hoogsteen hydrogen bonding sites in the major groove of the DNA duplex; pyrimidines having only one hydrogen bonding site vacant cannot be efficiently bound.

Current approaches have led to some success due to the overrepresentation of poly(purine-pyrimidine) sequences in eukaryotic genomes. However, the fact that most of these sequences contain purine-pyrimidine inversions necessitates the extension of the triplet recognition code to include all possible base sequences. This requires that these obstacles at the very least are overcome.

Ideally, as well as providing hydrogen bonding opportunities with the Watson-Crick base pair, the binding moiety should achieve optimal van der Waals and electrostatic interactions, π-stacking and tautomerisation. Also, not only should the desired binder bind the target base pair with the highest affinity, but discrimination against the other three base pairs is also of equal importance. And so, although an association of nucleic acids based on sequence recognition is not inherently suited for maximal binding affinity, the apparently very high degree of specificity of triple helical binding may be ideal for molecular intervention. Sections 1.4.1-1.4.3 deal with the more recent developments in this area.

1.4.1 Recognition of Double Stranded DNA via the Pyrimidine Motif

1.4.1.1 Replacement of Cytosine by pH Independent Pyrimidine-Like Analogues

The necessity to protonate cytosine in order to form two Hoogsteen hydrogen bonds with guanine restricts the formation of C^+•G-C triplets to conditions of low pH. As physiological pH only in extreme circumstances ventures outside the 7.2-7.3 range, a solution was needed for *in vivo* applications. A potential remedy in the form of 5-methylcytosine (^{m5}C) was offered. It was found to bind double stranded DNA with greater affinities over an extended pH range, and resulted in a more stable triple helix.[1.41, 1.42]

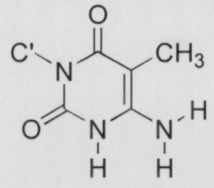
However, in order to associate as the ^{m5}C•G-C triplet, protonation of N-3 was still necessary. As such, the same charge repulsion problems observed with C^+•G-C triplets arose in the recognition of contiguous G-C tracts using 5-methylcytosine. It appeared that the enhanced

stability was not due to the pK_a increase (over 0.2 units), but is attributed to the entropic effect of displacing the water molecules from the major groove by methylating the 5-position of cytosine.[1.42] These methyl groups provide additional hydrophobic interactions through the formation of a spine of methyls in the major groove.

The rationale then followed by a number of research groups was to prepare synthetic nucleosides e.g. pyrimidine-like ring systems, that offer the hydrogen bonding pattern of protonated cytosine, i.e. two specific hydrogen bonds formed with O-6 and N-7 of guanine, while maintaining the appropriate backbone geometry, and ideally, incorporating a 5-methyl group, as it has been shown to enhance stability in the ^{m5}C•G-C triplet.

Contribution of McLaughlin *et al.*

The pyrimidine analogue **1.22**, a system designed to retain the desirable methyl group so successful in ^{m5}C, and also provide a hydrogen atom at the N-3 position to allow recognition of guanine in a pH-independent fashion, was synthesised and incorporated into a series of 15mer oligonucleotides d($T_3ZT_4ZTZTZT_2$) (Z = C, ^{m5}C, **1.22**) and their binding affinities with a 25mer DNA duplex embedded with the 15mer purine sequence d($A_3GA_4GAGAGA_2$) tested by UV thermal denaturation studies over the range of pH values 6.4 to 8.5, with the presence or absense of the triplex binding ligand spermine.[1.43] The results are seen in **Table 1.8**.



1.22

Table 1.8 Triplex Transition Temperature T_m values for 15mers containing C, ^{m5}C and **1.22**.

		Triplex Transition Temperature T_m (°C)				
Z	**[Spermine]**	**pH 6.4**	**pH 7.0**	**pH 7.5**	**pH 8.0**	**pH 8.5**
C	-	35.4	25.0	19.5	12.1	10.8
C	0.5 mM	45.1	32.6	26.9	15.8	-
^{m5}C	-	42.9	34.6	33.0	28.7	26.3
^{m5}C	0.5 mM	54.6	42.6	37.9	27.2	14.4
1.22	-	26.9	27.0	27.5	27.3	26.3
1.22	0.5 mM	39.1	40.7	39.7	38.0	30.8

Unsurprisingly, the T_m values obtained for C- and ^{m5}C-containing oligomers decreased with increasing pH, with spermine providing higher T_m values (and hence enhanced stability) due to the shielding effect against the charged phosphate backbones. The triplex containing **1.22** never achieved the stability of the ^{m5}C-containing triplexes, as seen from the T_m values, but seemed relatively stable at all pH tested, and hence appeared pH-independent. To examine its superiority as an uncharged moiety in contiguous tracts, a second study was conducted. Two sequences $d(T_3ZT_4ZT_2)$ (Z = ^{m5}C, **1.22**) were prepared and their binding affinities with a duplex containing five contiguous G-C base pairs (and six G-C base pairs in total) investigated. Neither of the sequences resulted in significant T_m values. This may be explained by charge repulsion between neighbouring ^{m5}C bases, and the bulky O-6 carbonyl atom providing steric hindrance, or poor base stacking by analogues **1.22**. Success was eventually attained by a final 15mer nucleotide containing alternating ^{m5}C and **1.22** residues. A T_m value was obtained up to pH 8.0, indicating triplex formation and stability.

Contribution of Fox *et al.*

The triple helix formation at the sequence $d(G_6A_6)$•$d(T_6C_6)$ was examined using a series of modified nucleotides based on C_6T_6, where C is replaced by a variety of isostructural cytosine derivatives **1.23** (**Figure 1.11**), ^{m5}C, and sequence variations of the α- and β-anomers of **1.23**: $(\alpha\beta)_3$, $\alpha_3\beta_3$ and $\beta_3\alpha_3$. Unnatural α-oligonucleotides are known to resist nuclease attack and hence their interest in this study. The study involved comparing the binding of the varied **1.23**-containing nucleotides with C_6T_6 and $^{m5}C_6T_6$ at a range of pH from 5.0 to 7.0, temperatures, presence of naphthylquinoline (NQ) and oligonucleotide concentration, as in **Table 1.9**.[1.44]

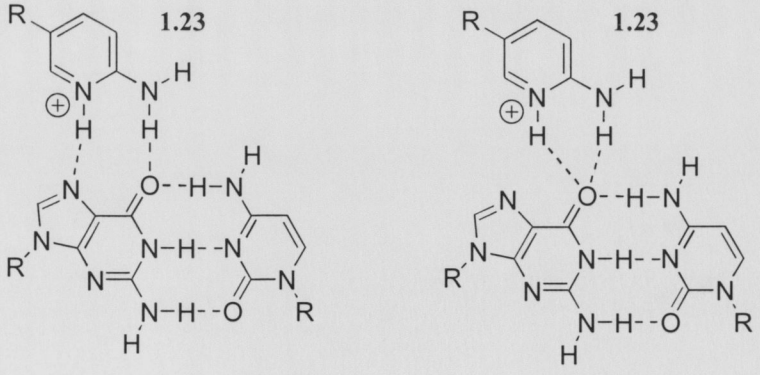


Figure 1.11 Proposed hydrogen bonding schemes between **1.23** and a G-C base pair.

Table 1.9 Summary of binding interactions between target site G_6A_6•T_6C_6 and the **1.23**-derivatised oligonucleotides under a range of experimental conditions.

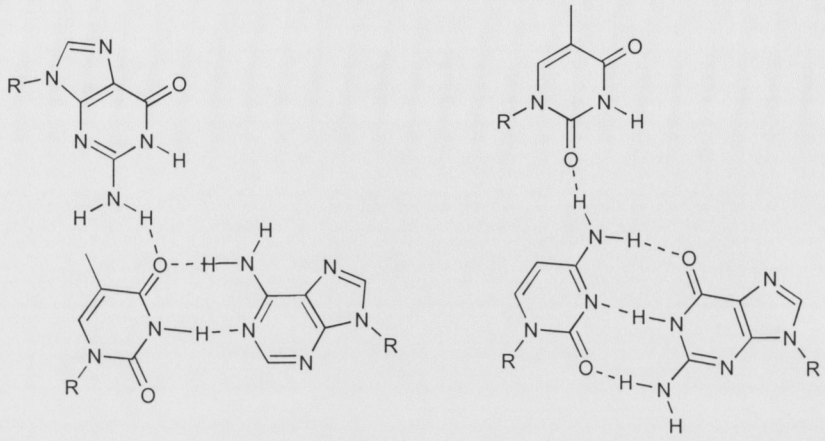
Experimental Conditions	Oligonucleotides						
	C_6	$^{5m}C_6$	β_6	α_6	$(\alpha\beta)_3$	$\alpha_3\beta_3$	$\beta_3\alpha_3$
10μM ODN pH 5.0-6.5	X	X	✓	X	X	X	X
10μM ODN pH 7.0-7.5	X	X	X	X	X	X	X
10μM ODN pH 5.0, 4 °C	✓	✓	✓	W	W	✓	✓
50μM ODN pH 5.0	X	✓	✓	✓	✓	✓	✓
10μM ODN pH 5.0, 10μM NQ	✓	✓	✓	X	✓	✓	W
10μM ODN pH 6.5, 10 μM	X	X	✓	X	X	W	X

(✓) represents binding, (X) indicates no binding, (W) implies possible (weak) binding.

The oligonucleotide containing β_6 (**1.23b$_6$**) consistently formed the most stable triplex, indicating that the dramatic improvement in binding affinity and complex stability at various pH was as a result of replacing cytosine with **1.23b**; a surprising outcome as it represented a contiguous run of six **1.23**•G-C triplets. The results found that the stability was greater than that obtained with 5-methylcytosine. This stability was attributed to an increased degree of protonation of N-3 (by raising the pK_a) on this occasion. ODNs containing the α-anomer were found to have the weakest binding overall, suggesting that the β-anomer was more effective in this case. Molecular modelling studies indicated that weak binding of the α-anomer was due to steric restrictions.

Several studies have been conducted on all possible base triplet combinations of natural DNA bases to elucidate the most stable associations.[1.45, 1.46] The base triplets G•T-A and T•C-G (**Figure 1.12**) emerge constantly as the most stable triplets *via* the pyrimidine motif. They are inherently more weakly bound than C^+•G-C and T•A-T due to only one hydrogen bond being formed with the third strand base, and being non-isomorphous, are less stable. T•C-G can be accommodated with both the pyrimidine and purine motifs as T is recognised *via* both motifs. However, G•T-A is only formed by way of the pyrimidine motif, and so there is no natural

system known to date that recognises T-A in the purine motif. Realistically, however, it will be necessary to recognise purine tracts containing more than one pyrimidine interruption, and much effort is being spent in this area.



G.T-A **T.C-G**

Figure 1.12 Base Triplets G•T-A and T•C-G.

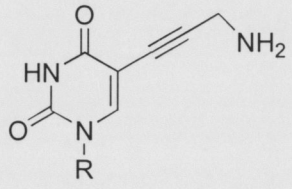
Contribution of Fox *et al.*

A series of oligonucleotides $d(T_2CT_4ZT_2CT_4)$ (Z = A, G, C, T) were prepared and their binding affinities with duplexes containing the target sites $d(A_2GA_4XA_2GA_4)$ (X = T, C) measured using DNase I footprinting titrations in the presence or absence of the triplex binding ligand naphthylquinoline (NQ).[1.47] For the recognition of isolated pyrimidine sites on the duplex, triplets G•T-A, T•T-A, C•T-A, T•C-G and C•C-G showed moderate stability, although a reduction of approximately 30-fold in stability relative to the canonical C^+•G-C and T•A-T was observed, as an oligonucleotide concentration of 1 μM was required for footprinting in the absence of NQ. NQ provides additional stability in triplexes, and so, footprints were observed for all eight combinations up to 0.3 μM.

Experiments carried out on similar tracts containing up to four contiguous T-A and C-G base pairs uncovered important information. As the number of adjacent pyrimidine sites increased to two, then three triplets, both G•T-A and T•C-G were observed to be the most stable triplets. However, three contiguous pyrimidines required the presence of NQ for stability, and four contiguous sites could not be stabilised.

A possible means of improving the triplex formation of G•T-A and T•C-G triads is to compensate for poor pyrimidine recognition by increasing the stability of the flanking T•A-T and C^+•G-C triplets. G•T-A is known to be stabilised by neighbouring T•A-T triplets, due to the presence of another hydrogen bond to thymine in the next T•A-T triplet. Contiguous C^+•G-C triplets are destabilising as a result of charge repulsion, but isolated C^+•G-C triads are more

stable than the isomorphous T•A-T triplet. It is reasoned that increasing the number of these isomorphous triplets between the isolated G•T-A (or T•C-G) triplets should see an improvement in triplex stability. Fox *et al.* carried out a DNase I footprinting study on how different arrangements of these triplets affect triplex stabilities on duplexes containing several (up to 25 %) pyrimidine interruptions. Analogue **1.24** was substituted instead of T to optimise stability in the third strands used to target 22mers containing the sequences $(A_3Y)_5AA$, $(AGAY)_5AA$ and $((AYGA)_5AA$ (Y = T or C) in the presence or absence of NQ.[1.48]



1.24

Results revealed the importance of the number and relative positions of the canonical triplets with respect to the nonisomorphous triplets. Placing five A-T triplets between each G•T-A induced triplex formation between $(T_5G)_3T_4$ and $(A_5T)_3A_4$, and addition of NQ produced a footprint at A_3T and A_5T at lower [ODN]. Replacement of a central T•A-T triplet in a block of three T•A-T triplets with a C^+•G-C resulted in a complex, without the need for NQ. However, placement of this C^+•G-C triplet on the 5′-side of G•T-A resulted in a destabilisation with respect to the effect observed with having a T•A-T triad on the 5′-side, which provided an extra inter-triplet hydrogen bond interaction. T•C-G triplets were found to be less stable in this study, as NQ was required to obtain footprints for A_5C, and an AGAC fragment required the highest oligonucleotide concentration. The position of C^+•G-C again was least favourable at the 5′-side of T•C-G. However, the strongest T•C-G complex involved C^+•G-C at the 3′-side.

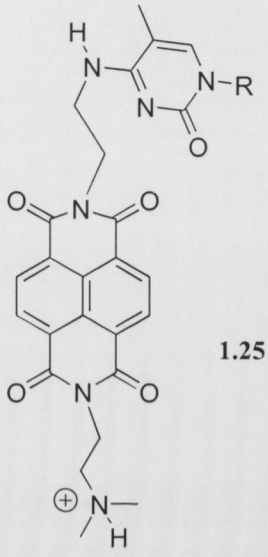
The modified oligonucleotides containing **1.24** were observed to form stable structures consisting of **1.24**•A-T triplets, due to the positive charge counteracting part of the large negative influence of the phosphates in the $(A_3T)_5A_2$ target site. Triplexes containing adjacent C^+•G-C and **1.24**•A-T triads were less stable than the corresponding unmodified nucleotides due to electrostatic repulsion between positively charged species.

Contribution of McLaughlin and Gianolio

To address the additional destabilising influences found with the provision of extra hydrogen bonding sites by analogue residues, a system which provided only one hydrogen bond with the available N-4 of the C-G base pair, and used intercalation as the mode of stabilisation was

devised and synthesised. Adduct **1.25**, an analogue of ^{m5}C (which possesses an ideal hydrogen bond acceptor system at N-3) containing a naphthalene-based intercalator, was incorporated into a 15mer oligonucleotide $d(T_3CT_4CTZTCT_2)$ (Z = **1.25**), and its selectivity for four 25mer duplexes containing the target sites $d(A_3GA_4GAXAGA_2)$-$d(T_3CT_4CTYTCT_2)$ (X-Y = A-T, T-A, C-G, G-C) examined using UV thermal denaturation studies, and compared with an analogous unmodified oligonucleotide, where Z = C.[1.49] The results are displayed in **Table 1.10**.

a)



1.25

b)

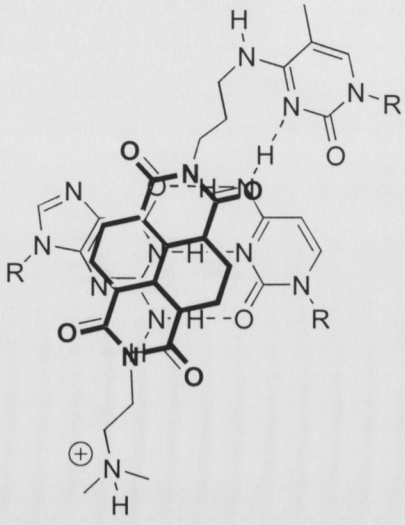


Figure 1.12 a) Structure of adduct **1.25**; b) Potential hydrogen bonding and intercalative interactions between **1.25** and a G-C base pair.

Table 1.10 Triplex Transition Temperature T_m (°C) for various triplets containing **1.25** and C.

Base Triplet	Triplex Transition Temperature T_m (°C)		
	pH 6.4	pH 7.0	pH 7.5
X-Y = G-C			
Z = C	30	22	15
Z = **1.25**	28	22	—
X-Y = C-G			
Z =C	13	—	—
Z = **1.25**	36	30	27
X-Y = A-T			
Z = C	9	—	—
Z = **1.25**	26	19	11
X-Y = T-A			
Z = C	—	—	—
Z = **1.25**	26	22	broad

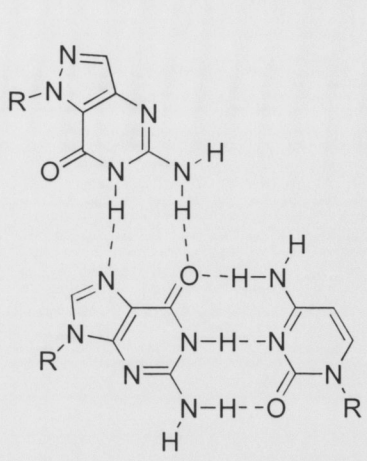
T_m values decreased with increasing pH. Also unsurprising was the obvious preference of the tract containing C at the pre-selected position for G-C base pairs at lower pH. The analogue tract containing adduct **1.25** was observed to bind all four base pairs with moderate affinity, but the desired preference for C-G base pairs was upheld. A possible explanation is due to the modes of binding. All four base pairs were recognised by the analogue tract through intercalation. However, C-G base pairs were additionally bound by an interresidue hydrogen bond, thus aiding their selectivity.

1.4.1.2 Replacement of Cytosine by pH independent Purine Analogues

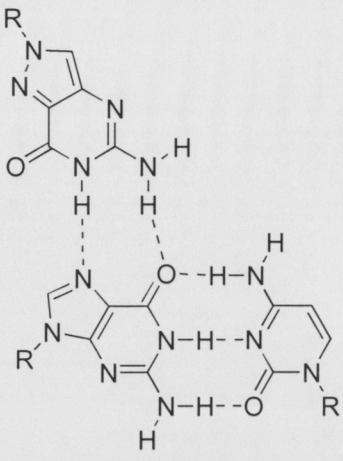
Contribution by Dervan and Koh

Two purine-like cytosine analogues **1.26** and **1.27** were designed, synthesised and incorporated into oligonucleotides to assess their sequence selectivities by means of DNA affinity cleavage studies.[1.50] Four different oligonucleotides d(T_7ZT_7) (Z = **1.26**, **1.27**, C, Φ; where Φ is an abasic residue used as a control) were prepared and examined with 30mer duplexes containing the

target sites $d(A_7XA_7)$-$d(T_7YT_7)$ (X-Y = A-T, T-A, C-G, G-C) at conditions of 35 °C and pH 7.4. Results revealed the most intense cleavage for triplexes containing the combinations of C•G-C and **1.26**•G-C, while **1.27**•G-C was observed to display moderate cleavage patterns.



1.26•G-C



1.27•G-C

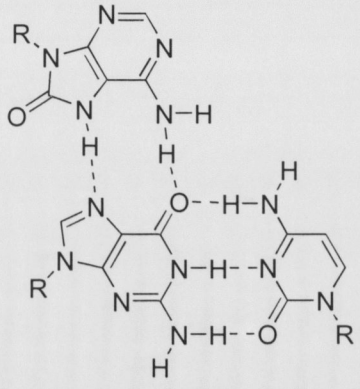
To determine the affinity of DNA tracts containing many **1.26**-residues, a study of four nucleotides $d(T_5(ZT)_5)$ (Z = **1.26**, **1.27**, C, ^{m5}C) binding to a DNA plasmid containing the 15mer purine sequence $d(A_5(GA)_5)$ was conducted over the pH range 6.2 to 7.8. Binding was not observed above pH 7.0 for C-containing tracts and pH 7.8 for ^{m5}C-containing oligonucleotides. The analogue nucleotides containing **1.26** and **1.27** showed contrasting behaviour. The **1.26**-containing tract bound all base pairs under all conditions of the study, while the **1.27**-containing nucleotide was not observed to bind at any pH tested. The nonisomorphism of T•A-T and **1.27**-G-C base triplets was offered as an explanation.

A study of a duplex containing six contiguous G-C base pairs was then carried out. The $(G)_6$ tract was targeted by three oligonucleotides containing $d(C_6)$, $d(^{m5}C_6)$ or $d(\mathbf{1.26}_6)$ sequences, respectively, over the pH range 6.2-7.4 at 30 °C, and the sequence specificities determined by affinity cleavage experiments. The C_6- and $^{m5}C_6$-containing nucleotides both gave cleavage at pH 6.2, but both decrease in intensity above pH 6.6. Even up to pH 7.4 (30 °C), the oligodeoxyribonucleotide containing **1.26** gave observable cleavage.

Contribution of Miller *et al.*

The known tendency of 8-substituted adenosine analogues to exist primarily in the *syn* conformation led to the use of 8-oxoadenine (**1.28**) as a protonated cytosine substitute.[1.51] The similarity of **1.28** to C^+ as a bidentate hydrogen donor that offers the identical hydrogen bonding pattern required for G-C base pair recognition led to a comparison study of oligodeoxyribonucleotides containing psoralen derivatives of **1.28** and C.[1.52]

UV melting experiments were used as the means to examine the triple helix formation of the two oligonucleotides d($TZ_3TZ_3T_5ZT$) (Z = **1.28**, C) with the 28mer DNA duplex containing the target purine sequence d($AG_3AG_3A_5GA$) at pH 6.2 and 7.5. The melting studies resulted in a temperature profile for the two pH systems. At pH 6.2, the triplex dissociation temperature of the C-containing oligomer is 33 °C, compared to 28 °C for the **1.28**-containing tract, implying a more stable triplex. However, at pH 7.5, triplex formation was not observed with the cytosine-containing tract. Unexpectedly, the **1.28**-containing oligonucleotide was found to have a melting curve corresponding to two transition temperatures at pH 7.5, 27 °C for triplex dissociation and 73 °C for duplex dissociation.



1.28•G-C

1.4.2 Recognition of Double Stranded DNA via the Purine Motif

The pH independence of the purine motif has offered the promise of a more viable approach to triplex formation and double helix recognition, as the requirement of protonation is not necessary in the formation of two reverse Hoogsteen hydrogen bonds between the third strand and the Watson-Crick base pairs of G•G-C, T•A-T and A•A-T base triplets.[1.53] The fact that isomorphism is not observed between combinations of these triads leads to an overall stability that is dependent on the content of G-residues in the third strand. However, G-rich oligonucleotides display a characteristic that hampers their *in vivo* application.

At physiological concentrations of monovalent metal ions, such as Na^+ and K^+, nucleotides with a high G content self-associate into tetrad structures (**Figure 1.13**). These quadruplexes are formed through eight intermolecular hydrogen bonds and are stabilised by co-ordination of the four O-6 atoms of guanine with these monovalent cations, and also π-stacking interactions between two associating tetrad structures. Formation of these quadruplexes reduces the number of oligonucleotides available for triplex formation. The presence of these metal ions

not only encourages quadruplex formation, but also induces triplex dissociation. If this approach is to be realistically used as an *in vivo* method, quadruplex formation must be inhibited.

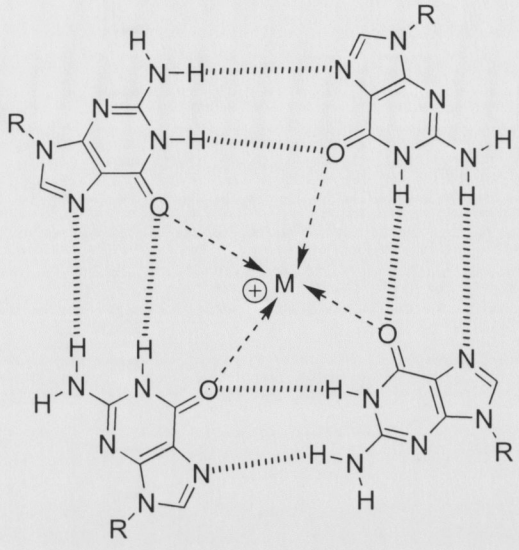
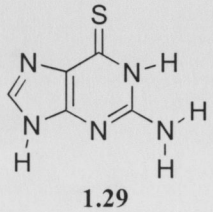


Figure 1.13 Structure of H-bonded G-quadruplex, with an internally co-ordinated monovalent cation M^+.

In a move to disrupt the ion chelation and the multiple hydrogen bonding interactions associated with quadruplex formation, the sulphur-containing guanine analogue **1.29** was postulated as a possible solution. The replacement of the oxygen atom at position-6 with the larger sulphur atom would not interfere with the reverse Hoogsteen hydrogen bonding. But the lower electronegativity of sulphur would make it a less efficient hydrogen bond acceptor. The prerequisite of an eight-coordinate chelation of K^+ (for example) to the carbonyl oxygen atoms in two stacked quartets may not be fulfilled in **1.29** due to the inherently lower affinity of sulphur for K^+. Also, the bulkier sulphur atom may prevent quartet formation on steric grounds. To examine if these theoretical advantages could succeed in practice, and provide an assessment of its success, two similar studies were carried out on **1.29** and offered different perspectives on similar conclusions.



1.29

Contribution of Revankar *et al.*

A series of oligonucleotides of the form $d(G_4TG_6T_3G_2TG_2TG_2T_2GTG_2TG_2TG_2TG)$ was prepared with G being substituted to varying extents by **1.29**. These nucleotides were then tested for binding affinity with a duplex containing the 36-base pair purine oligomer $d(GAG_2AG_2AG_2C_2GAG_2AG_2AG_2A_3G_6CG_4)$ using a gel shift assay, and the dissociation constants (K_D) obtained under conditions of buffer or salt.[1.54] The unsubstituted TFO gave a K_D of 6×10^{-9} M in buffer, while 50 mM KCl conditions raised its value to 1×10^{-6} M.

Oligonucleotides substituted to various degrees with **1.29** instead of guanine found salt independent values of K_D as the percentage of **1.29** increased; 3×10^{-8} M for approximately 50 % substitution in both buffer and salt conditions. While the requirement of triplex formation in the presence of metal (salt) ions was adhered to in this study, it was noted that an increased K_D value relative to the unmodified oligonucleotide was observed (at 66 % substitution, $K_{Dbuffer} = K_{Dsalt} = 1 \times 10^{-6}$ M), indicating a lower stability in **1.29**•G-C triplets than in G•G-C base triads.

Contribution of Maher *et al.*

A duplex of short length containing the purine sequence $d(G_3AG_3AG_2AG_3)$ was targeted by two oligomers d(GZGTGZGTGTZGTGZG) (Z = G, **1.29**) to examine the triplex forming ability of both oligomers in the presence or absence of K^+ or Na^+.[1.55] The dissociation constants (K_D) obtained as a result of gel mobility shift titrations are in **Table 1.11**.

Table 1.11 Triplex dissociation constants K_D (M) of G- and **1.29**-containing oligonucleotides in the presence or absence of M^+.

Z	10 mM M^+	K_D (M)
G	—	1.4×10^{-7}
G	Na^+	2.1×10^{-7}
G	K^+	5.3×10^{-6}
1.29	—	2.2×10^{-7}
1.29	Na^+	2.0×10^{-7}
1.29	K^+	1.9×10^{-7}

G-containing oligonucleotides were influenced by the presence of Na^+ and K^+ and triplex formation was observed to decrease. However, **1.29**-containing TFOs seemed resistant to ionic influence even up to a K^+ concentration of 200 mM, while triplexes involving G-containing

ODNs were eliminated by 50 mM K^+, indicating a high stability for **1.29**•G-C triplets with **1.29**-substitution up to four residues. NMR and optical melting experiments were conducted on **1.29**-containing DNA duplexes and quadruplexes in a further study by Bolton and co-workers,[1.56] and verified that the presence of **1.29** in duplexes causes a decrease in stability of double helical DNA due to the poor ability of sulphur to act as a hydrogen bond acceptor. Of greater importance perhaps was the verification of the ability of **1.29** to completely inhibit quadruplex formation.

1.4.3 Recognition of Double Stranded DNA by Novel Base Analogues

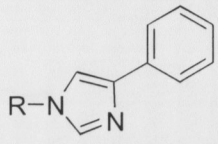
The omnipresent restriction to homopurine tract recognition, due to the low stability of pyrimidine-bound triplets with only one hydrogen bond available, and the steric clash potential of the 5-methyl group of thymine, has prompted efforts of major groove recognition into three main areas: alternate triplex motifs, strand switching strategies[1.57] and new base analogues. The former two presently seem only to limit destabilisation rather than increase triplex stabilisation. Base analogues, however, may succeed in achieving selectivity, specificity and triplex stability as a result of their use in a novel approach to base pair recognition.

The most promising way to date to achieve a general triplex-mediated DNA recognition scheme is to use extended heterocyclic systems that can span the entire width of the major groove, and thereby make use of all the possible hydrogen bonding sites on the major groove edge of each of the four Watson-Crick base pairs.

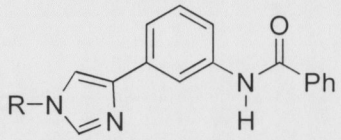
1.4.3.1 C-G Base Pair Recognition by Extended Heterocyclic Systems

Contribution of Dervan *et al.*

This strategy was first observed in the design (from model building) and synthesis of systems **1.30** and **1.31**. C-G base pair recognition within the triplex motif had not achieved success up to now, and hence nonnatural base **1.30** was designed to accept a hydrogen bond from the C-4 amino group of cytosine; and **1.31** planned to be of sufficient length to interact with the entire major groove edge of the C-G base pair and form two hydrogen bonds. Both systems involve two aromatic rings joined by a flexible single bond, which allowed system rotation to achieve the best interaction geometry. Other base pairs were meant to be discriminated against on steric grounds.



1.30



1.31

Affinity cleavage studies were conducted on a 30mer DNA duplex containing the purine segment d(A_7XA_7)-d(T_7YT_7) (X-Y = A-T, T-A, C-G, G-C) with a single variable position, targeted by a series of 15-base pair tracts d(T_7ZT_7) (Z = C, T, **1.30**, **1.31**) at pH 7.4 and 30 °C.[1.58] Analysis of results found that the most intense cleavage was observed for base triplets C•G-C, T•A-T, **1.31**•T-A and **1.31**•C-G. This implies that **1.31** favours pyrimidine-purine base pairs over purine-pyrimidine base pairs, due perhaps to a shape-selective mechanism. The stability of these triplets is neighbour-dependent, with flanking T•A-T triplets providing more stability than C^+•G-C triads. No discrimination was observed however between C-G and T-A base pairs.

The binding mode of **1.31** with the Watson-Crick base pair T-A was investigated and elucidated through a later study of an intramolecular triplex formed from the oligonucleotide d(AGATAGAACCCCTTCTATCTTATATCT**1.31**TCTT), designed to form a **1.31**•T-A base triplet.[1.59] The ^{1}H and NOESY spectra indicated that it formed the desired triplex. However, further examination of the NOESY spectrum uncovered the fact that crosspeaks expected for T_4/A_5, T_{16}/A_{17} and $T_{27}/\mathbf{1.31}_{28}$ were missing, as were the NOE crosspeaks between the base protons on **1.31** and the $T_{27}/A_3/T_{18}$ triplet. The presence of other unexpected crosspeaks, along with the other evidence presented, implied an intercalative mode of binding of **1.31** between the $T_{16}/A_5/T_{29}$ triplet and the A_{17}/T_4 base pair. This unexpected result was then the subject of an exhaustive NMR study,[1.60] which concluded that its preference for pyrimidine-purine base pairs may be as a result of increased stacking interaction due to better overlap between **1.31** and C-G and **1.31** and T-A.

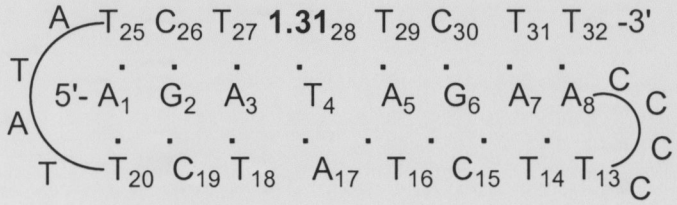
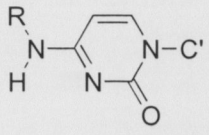


Figure 1.14 Schematic of the proposed folding of the 32-base intramolecular triplex with the numbering system used in text. Hydrogen bonds are indicated by (•).

Contribution of Miller *et al.*

A series of N^4-modified cytosines **1.32a-d** with side-chains of varying length were synthesised and included in oligomers d($CTTCT_6XT_4$) (Z = C, **1.32a-d**) for targeting the 15-base pair duplex d($GAAGA_6YA_4$)-d(T_4ZT_6CTTC) (X-Y = A-T, T-A, C-G, G-C) and their affinities measured by UV thermal denaturation over the temperature range 0-60 °C.[1.61] These results are shown in **Table 1.12**.



1.32a: R = $(CH_2)_3CH_3$
1.32b: R = $(CH_2)_3COOH$
1.32c: R = $(CH_2)_3NH_2$
1.32d: R = $(CH_2)_3NHC(O)CH_3$

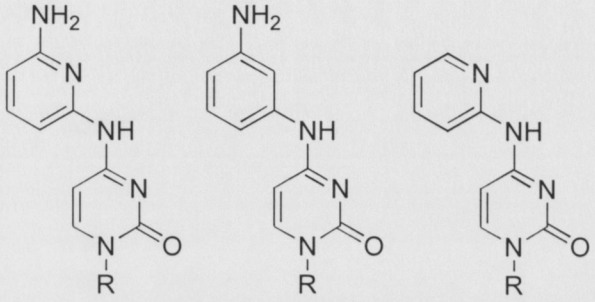
Table 1.12 Triplet dissociation temperature T_m (°C) of Triplet X-Y•Z.

Base Substitution			T_m (°C)
X	**Y**	**Z**	
C	G	C	<0
C	G	**1.32a**	<0
C	G	**1.32b**	<0
C	G	**1.32c**	10
C	G	**1.32d**	20
T	A	**1.32d**	<0
A	T	**1.32d**	<0
G	C	**1.32d**	8
G	C	C	32

Only **1.32c** and **1.32d** were observed to form triplexes, with both preferentially selecting C-G base pairs. Further examination of the molecular models indicated that success was achieved due to the long 3-aminopropyl and 3-acetamidopropyl side-chains which extended across the major groove and enabled hydrogen bonding between the amino or amido NH on the 3-aminopropyl or 3-acetamidopropyl chains, respectively, and the C-6 carbonyl oxygen of G of the target C-G base pair.

It was then reasoned that binding could be improved by providing a rigid platform on which to position any hydrogen bonding sites instead of the weakly binding flexible propyl chain. As such, N^4-(aryl)deoxycytosine derivatives **1.33a-c** were hypothesised.[1.62] The pyridine ring of **1.33a** was shown by molecular modelling studies to traverse the major groove, thus

placing its 6-amino group within H-bonding distance of the carbonyl O-6 or N-7 of guanine. The tautomerisation of **1.33a** to the imine would provide a further hydrogen bond acceptor to interact with the C4-amino group on cytosine, and in fact, UV and NMR experiments indicate that this tautomerisation was indeed the case. Aromatic base stacking also provided an additional stabilising influence.



1.33a **1.33b** **1.33c**

The novel bases **1.33a-c** were synthesised, inserted into oligomers d($^{m5}CT_2{}^{m5}CT_6ZT_4$) (Z = **1.33a-c**) and UV melting experiments carried out to determine their binding affinities to the duplex tracts d($GA_2GA_6XA_4$)-d($T_4YT_6CT_2C$) (X-Y = A-T, T-A, C-G, G-C). The results obtained are detailed in **Table 1.13**.

Table 1.13 Triplex dissociation temperatures T_m (°C) of triplexes X-Y•Z containing base analogues **1.33a-c**.

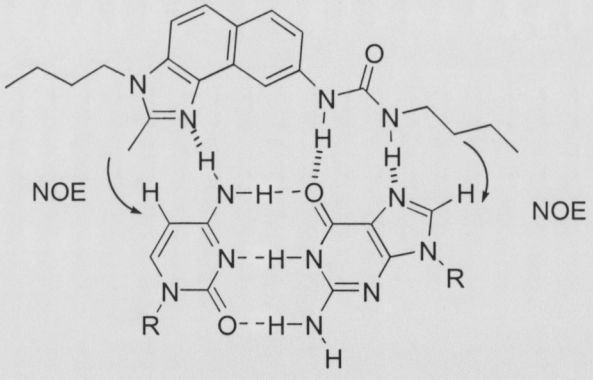
		T_m (°C)		
X	**Y**	**Z = 1.33a**	**Z = 1.33b**	**Z = 1.33c**
A	T	32	29	—
T	A	23	23	<5
G	C	20	18	—
C	G	25, 38	23, 36	<5

System **1.33c** failed to form a triplex. More hopeful were the results obtained for analogues **1.33a** and **1.33b**, which concluded that the interaction between the O-6 of guanine and the 6-amino group on the novel base was essential for triplex formation. A biphasic transition was observed with **1.33a** and C-G, indicating the formation of two interconvertible triplexes (depending on the heating technique used), and thus the presence of two different binding modes, probably hydrogen bonding and intercalation, as found in the case of system **1.31**. As a similar two-temperature result was obtained with **1.33b** and C-G, the behaviour was not

associated with simply the pyridinyl ring. The high degree of binding of these systems to A-T base pairs may also be due to the hydrogen bonding pattern offered by the imino tautomer. In fact, moderately stable triplexes were formed by **1.33a** and **1.33b** with all four base pairs. A possible explanation is that additional stabilisation was obtained by the triplet from interresidue hydrogen bonding with neighbouring base pairs.

Contribution of Zimmerman and Schmitt

Cytosine-guanine recognition through hydrogen bonding with both bases was the driving force behind the design and synthesis of novel structure **1.34**. ^{1}H NMR titration experiments were used to elucidate its binding to a C-G base pair in chloroform-*d*.[1.63] This study marked the beginning of a new approach to the examination of triplex formation in solution, as studies prior to this were conducted only in aqueous conditions due to the inherent water-solubility of DNA. A reverse titration was carried out on the 1:1 mixture of tri-*O*-acetylguanosine and tri-*O*-acetylcytidine with **1.34**, and K_a, the association constant, obtained was 160 M^{-1}. Difference NOE experiments on a 1:1:1 9-octylguanine:1-propylcytosine:**1.34** elucidated the binding pattern to be that shown in **Figure 1.14**, where the addition of DMSO-d_6 separated overlapping signals to show weak intermolecular interactions between the methyl group of **1.34** and H-5 on cytosine and the β-methylene protons on the butyl group and the H-8 of guanine.

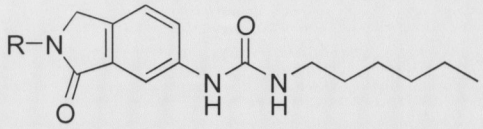


1.34•C-G

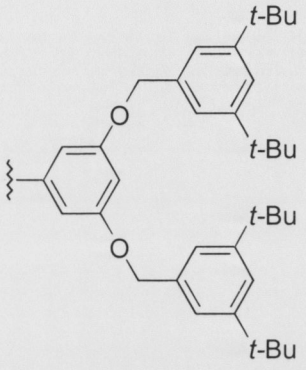
Figure 1.15 Hydrogen bonding scheme of **1.34** and C-G base pair.

The relatively low association constant obtained as a result of the poor binding of **1.34** with C-G base pairs (K_a = 160 M^{-1}) initiated a molecular modelling study of ligands that are more complementary in design. The (hexylureido)-analogues **1.35** and **1.36** that resulted were synthesised and ^{1}H NMR binding titrations carried out on each of the systems being titrated with a 1:1 mixture of tri-*O*-acetylguanosine and tri-*O*-acetylcytidine.[1.64] The association constants

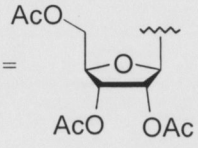
determined for **1.35**•C-G and **1.36**•C-G were 1016 M^{-1} and 736 M^{-1}, respectively, indicating a ten-fold increase in stability on **1.34**.



1.35: R =



1.36: R =



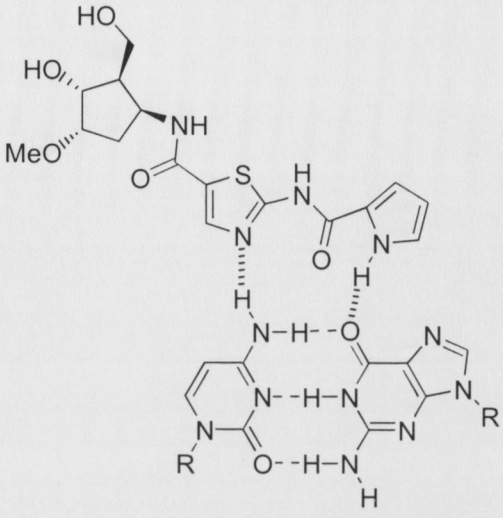
Dimerisation studies were carried out on both **1.35** and **1.36** to determine if self-association would be of concern at higher concentrations for these systems. While **1.35** exhibited low dimerisation properties (K_{dimer} = 18 M^{-1}), the increased hydrogen-bond acceptor potential of the lactam carbonyl of **1.36** promoted self-association, and a lower concentration of **1.36** (0.13 mM instead of 1.9 mM) was necessary.

The hexyl urea group was shown to be important for binding as its replacement with an amide group resulted in weaker binding (K_a = 70 M^{-1}) and so a third hydrogen bond may be necessary for stabilised complexation, as proved by observed downfield shifts of the H-5 of cytosine hydrogen bonded to the imide or lactam carbonyl oxygen in **1.35** or **1.36**. Reverse titrations of $CDCl_3$ solutions of C-G with **1.35** and **1.36** gave additional evidence for these complexes, with K_a of 800 M^{-1} for **1.35**•C-G, and thus supports the theoretical 1:1 binding model. However, a similar model could not be assigned to the **1.36**•C-G titration, due possibly to C-G base pair dissociation on increasing **1.36** concentrations. **1.36** has been incorporated into an oligonucleotide, and its binding affinity for C-G base pairs on duplex DNA ascertained, but publication of these results are pending.

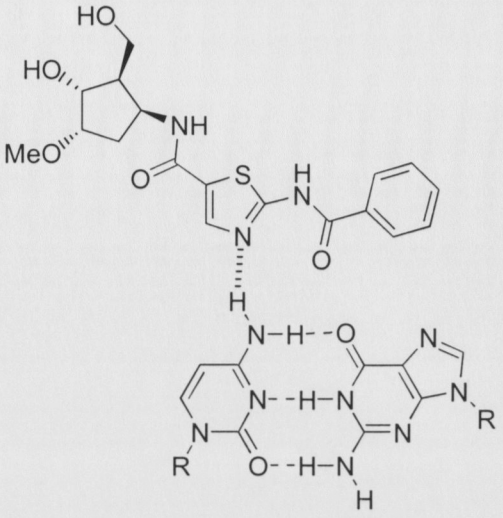
Contribution of Dervan *et al.*

Experience gained to date on these base analogues promoted continuing efforts to find heterocyclic systems which selectively recognise C-G or T-A Watson-Crick base pairs *via* the pyrimidine motif. Molecular modelling suggested systems **1.37** and **1.38** for interaction with C-

G base pairs. These base analogues were proposed to form two and one hydrogen bonds, respectively (see **Figure 1.16**).



1.37•C-G



1.38•C-G

Figure 1.16 Hydrogen bonding schemes of C-G base pairs with **1.37** and **1.38**.

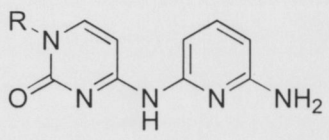
The binding affinities and sequence selectivities of two oligomers containing **1.37** and **1.38** d($TT^{m5}CT^{m5}CZT^{m5}C^{m5}C^{m5}CTTT$) (Z = **1.37**, **1.38**) with a 314 bp DNA plasmid containing all four possible 15 base pair duplex sites d(AAGAGAXAGGAAA) (X = A, G, C, T) were assessed using DNase I quantitative footprinting titrations.[1.65] The experimental results are reported in **Table 1.14**.

Table 1.14 Equilibrium association constants K_T (M^{-1}) for the eight triplets of **1.37** and **1.38**.

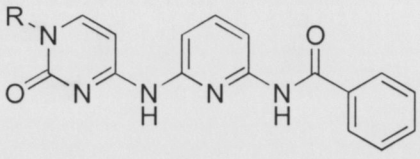
X	Z = 1.37	Z = 1.38
C	9.2×10^5	2.4×10^6
T	1.6×10^6	2.6×10^6
G	1.4×10^5	1.0×10^5
A	1.5×10^5	1.9×10^5

Although the oligonucleotides containing **1.37** and **1.38** bound across C-G inversions, they were found to bind with similar affinities, implying that the predicted hydrogen bonding patterns were not the primary influence for the interaction. Likewise, T-A inversions were bound with affinities similar to both **1.37**- and **1.38**-containing oligomers, thus exhibiting a comparable selectivity pattern to **1.31**. It is likely that an intercalative binding mode existed for **1.37** and **1.38**.

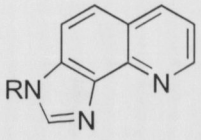
To obtain an overall view of how well these systems compare with contemporary base systems, further DNase I footprinting experimental studies were conducted on a number of extended heterocyclic units; some new, some already reported. Those systems obtained or synthesised were that of Miller's group, **1.33a**; its N-benzoylated analogue, **1.39**; the imidazoloquinoline, **1.40**; the 8-aminoimidazoloquinoline, **1.41**; and the deoxyribose derivative of the nucleobase **1.34** reported by Zimmerman and colleagues, **1.42**. Each system was then incorporated into a separate oligomer of the same sequence as that originally containing **1.37** and **1.38** and the footprinting analysis carried out. The results are summarised in **Table 1.15**.



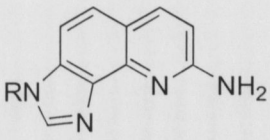
1.33a



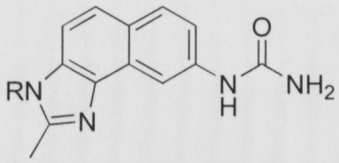
1.39



1.40



1.41



1.42

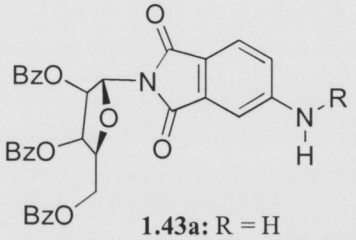
Table 1.15 Equilibrium association constants K_T (M^{-1}) for the triplexes containing systems **1.34**, **1.37**, **1.38**, **1.39**, **1.40**, **1.41 and 1.42**.

Compounds	Base Pairs			
	C-G	T-A	G-C	A-T
1.37	9.2×10^5	1.6×10^6	1.4×10^5	1.5×10^5
1.38	2.4×10^6	2.6×10^6	1.0×10^5	1.9×10^5
1.39	8.5×10^6	5.8×10^6	5.7×10^5	8.8×10^5
1.34	$\sim 1 \times 10^5$	$<1 \times 10^5$	7.0×10^5	$<1 \times 10^5$
1.40	1×10^5	1×10^5	$<1 \times 10^5$	1×10^5
1.41	1×10^5	1×10^5	2.6×10^5	1×10^5
1.42	$<1 \times 10^5$	$<1 \times 10^5$	1×10^5	$<1 \times 10^5$

Results predicted that the similar binding affinity of **1.39** (compared with that of **1.37** and **1.38**) should suggest a similar intercalative binding mechanism. The shorter aromatic systems of **1.34**, **1.40**, **1.41** and **1.42** resulted in lower affinities for all four base pairs, with the likelihood that no hydrogen bonding was observed in aqueous conditions.

Contribution of Weisz and Lengeler

The phthalimide derivatives **1.43a-d** were hypothesised to selectively recognise C-G base pairs through the formation of two hydrogen bonds with the C-4 amino group of cytosine and the C-6 carbonyl oxygen of guanine. Molecular modelling also proposed that triplets formed by these analogues with C-G would be isomorphous with the canonical triplets C^{+}•G-C and T•A-T, encouraging stability through lack of backbone distortion in the third strand nucleotide.



1.43a: R = H
1.43b: R = -CO-C_6H_5
1.43c: R = -CO-NH_2
1.43d: R = -CO-NH-C_4H_9

The synthesised molecules were then subjected to ^{1}H NMR titration studies in CD_2Cl_2, with each analogue being titrated with a 1:1 mixture of 3 ,5 -di-*O*-(triisopropylsilyl)protected 2'-deoxycytidine and 2'-deoxyguanosine at low temperature and the 4-^{15}N-amino labelled proton NH_b of cytosine followed to observe hydrogen bond formation.[1.66]

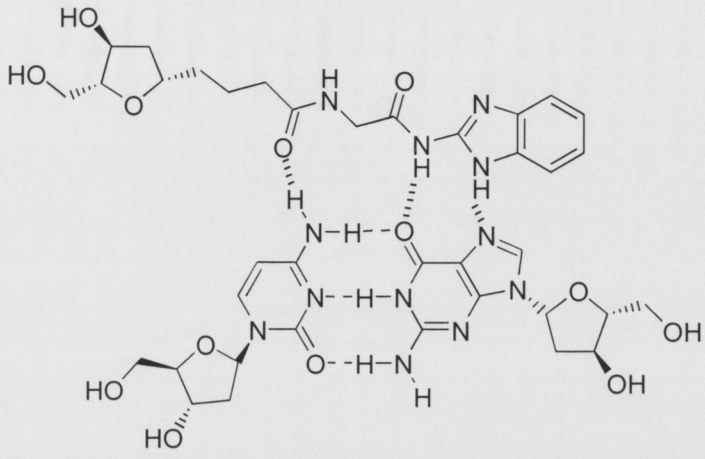
All four derivatives were observed to induce a downfield shift ($\Delta\delta$) in the NH_b proton signal, indicating the formation of hydrogen bonds. This occurred without disrupting the C-G Watson-Crick hydrogen bonding. However, **1.43c** and **1.43d**, having extra urea hydrogens, bound most strongly with the C-G base pairs, exhibiting a chemical shift (δ) of ~6.6 and 6.5 ppm, respectively. Although the loss of hydrogen bonds should be associated with a better steric fit, both **1.43a** and **1.43b** were found to complex less strongly with C-G, with **1.43a** being a slightly stronger complexing agent. The lowest δ of ~5.6 ppm displayed by **1.43b** may be due to its bulky aryl group, and perhaps a stronger amide hydrogen bond preventing the formation of an optimal NH_b-O hydrogen bond.

1.4.3.2 G-C Base Pair Recognition by Extended Heterocyclic Systems

Contribution of Maeda *et al.*

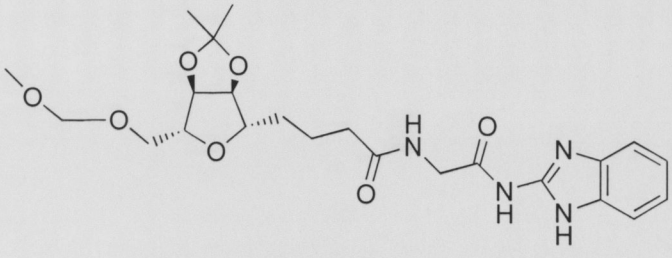
The lack of success achieved in the formation of unnatural triplets prompted a semi-empirical molecular modelling study to design **1.44**, a system aimed to select G-C base pairs specifically and form triple-helical structures *via* the usual Hoogsteen hydrogen bonds with guanine and an

extra hydrogen bond with cytosine. **1.44** was also found from molecular modelling to be isomorphous with T•A-T and C^+•G-C, thus aiding triplex formation through minimal backbone distortions.



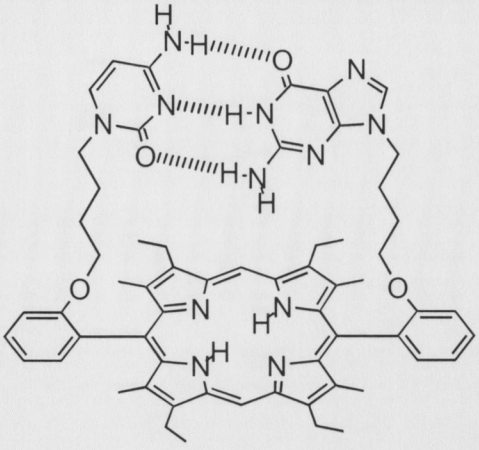
1.44•C-G

For synthetic purposes, the protected form of 2'-deoxy-2-**1.44**, **1.45**, was prepared, and ^{1}H NMR titration experiments used to examine its binding with each of the four natural deoxynucleotide bases in chloroform-*d* at 25 °C and obtain an association constant (K_a) value for each.[1.67] Adenosine and thymidine titrations did not produce a binding result. Cytidine addition to the $CDCl_3$ solution of **1.45** produced K_a of 2.6 X 10^3 M^{-1} (approximately half that of the natural G-C base pair); and binding was in a 1:1 mode. However, a 2:1 mode was observed in guanosine:**1.45** binding, and K_a was calculated to be 3 X 10^2 M^{-1}. These results imply specificity of **1.45** for cytidine and guanosine.

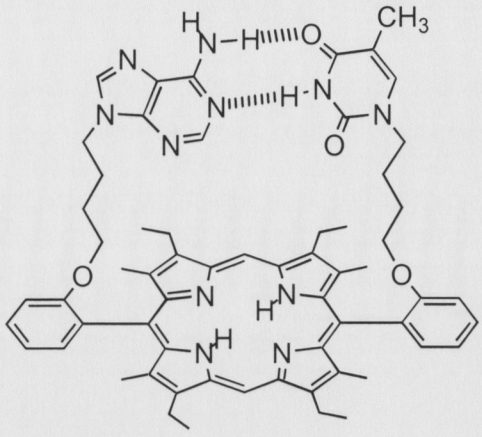


1.45

As the triplex association constant was indeterminate by ^{1}H NMR titration studies, an intermolecular NOE experiment was necessary. Model systems **1.46** and **1.47** were prepared as a control for stable base pairs, with the C-G and T-A base pairs fixed in close proximity with each other. A 1:1 mixture of **1.45**:**1.46** gave an NOE spectrum corresponding to the formation of a **1.45**•C-G base triplet. A 1:1 mixture of **1.45**:**1.47** failed to give a similar result.[1.67]



1.46



1.47

1.4.3.3 T-A Base Pair Recognition by Extended Heterocyclic Systems

Despite the best efforts of research groups worldwide, a general solution sequence-specific recognition scheme for C-G base pairs is elusive. But even less success has been achieved in the pursuit of T-A base pair binders. Recognising T-A is more of an immediate challenge, in fact, as natural bases do not recognise T-A base pairs *via* the purine motif. The recognition of the T-A base pair *via* the major groove would depend on a method of avoiding steric clash with the 5-methyl group of thymine, and also the possibility of forming a hydrogen bond with the C-4 carbonyl oxygen of thymine. A system that could also accept a hydrogen bond from the C-6 amino group on adenine would be even more desirable.

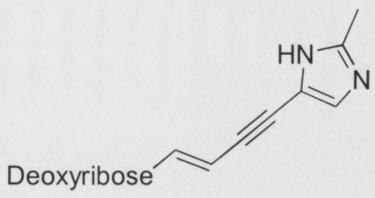
There have only been two reports that show promise: one of which was a study of molecular models of potential T-A base pair ligands, the other provides a synthetic base analogue that was successfully employed for the recognition of a T-A inversion.

Contribution of Richards and Rothman

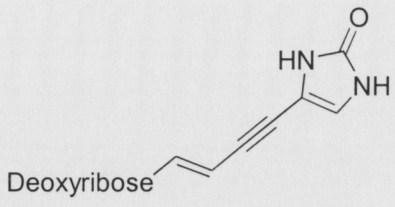
The four nonnatural nucleotide bases **1.48a-d** were designed with the use of molecular modelling as suitable candidates for T-A base pair recognition. They were incorporated into 11mer oligonucleotides $d(T_5ZT_5)$ (Z = **1.48a-d**) and the binding of these 11mer oligomers to a 11 base pair DNA double stranded segment $d(A_5XA_5)$-$d(T_5YT_5)$ (X-Y = T-A, G-C) examined in both A-type and B-type DNA conformations with molecular modelling.[1.68]

The most significant results were obtained for the systems involving **1.48c** and **1.48d**, with both showing different conformational preferences. The oligomer containing **1.48d** was shown to prefer binding with a T-A base pair rather than a G-C pair in B-DNA, exhibiting a binding affinity of approximately halfway between those of G•T-A and T•A-T base triads. The

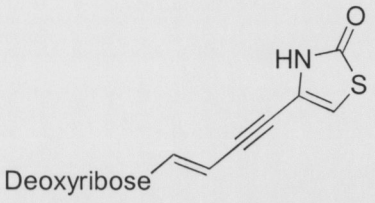
1.48c-containing nucleotide gave a modelling result in which preference for the T-A base pair in A-DNA was observed, but with a binding affinity approximately as low as that seen with G•T-A triplets. Follow-up synthetic and binding studies have not as yet been published.



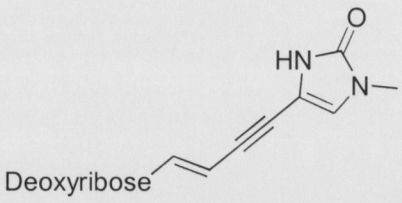
1.48a



1.48b



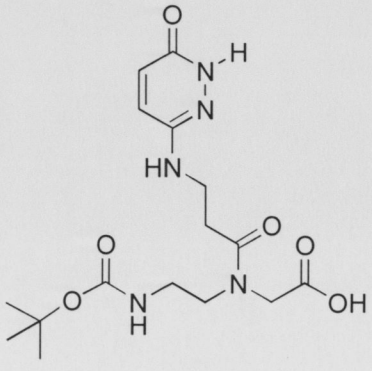
1.48c



1.48d

Contribution of Nielsen *et al.*

Novel base analogue **1.49** was designed by molecular modelling, synthesised and included in a bis-decamer-linked bis-PNA H-T^{iso}CTZT^{iso}CZTTT-eg-eg-eg-TTTACTACT-NH_2 (where ^{iso}C = pseudocytosine, eg = PNA linker, and Z= **1.49**, Φ (a 8-amino-3,6-dioxaoctanoic acid linker), G) and the thermal stability of these PNAs with their respective DNA target sites examined.[1.69] The systems were set up to place an adenine opposite each of the two, e.g. **1.49**, units, thus recognising T in the DNA tract, and a T opposite each T, and a C opposite each ^{iso}C for standard DNA A and G recognition. The results are given in **Table 1.16**.



1.49

Table 1.16 Triplex dissociation temperature T_m (°C) of PNA-DNA complexes.

bis-PNA	T_m (°C)
1.49	57.5
Φ	47.5
G	46.0

The results show that **1.49** bound more strongly to thymine than Φ (no-base) or G bound to thymine, by approximately 5 °C per unit, which made the **1.49**•T-A triplet more stable than the known G•T-A triplet. An analogous study of the bis PNA H-T^{iso}CT**1.49**T^{iso}C**1.49**TTT-eg-eg-eg-TTTGCTGCT-NH_2 , where **1.49** was opposite a G and so would recognise DNA cytosines, gave a T_m of 42 °C, thus exhibiting the preference of **1.49** for thymine over cytosine.

A further study was carried out on the bis-PNA H-T^{iso}CT**1.49**T^{iso}C**1.49**TTT-eg-eg-eg-TTTACTACT-NH_2 to investigate its sequence selectivity for oligonucleotide sequences containing various mismatches, including U, which would give some assessment of any steric clash of **1.49** with T, as U lacks the methyl group on the 5-position. The results of these thermal stability studies are summarised in **Table 1.17**.

Table 1.17 Triplex dissociation temperature T_m (°C) of bis-PNA-DNA complexes involving mismatches.

Oligonucleotide **5′-dCGCAGAXAGYAAACGC-3′**		T_m (°C)
X	**Y**	
T	T	57.0
A	T	45.0
G	T	40.5
C	T	40.5
U	U	63.5
A	U	49.5
G	U	44.5
C	U	44.5

The results show that the steric clash with the 5-methyl group of thymine is not entirely resolved by system **1.49**, as the bis-PNA segment bound with a stronger affinity ($\Delta T \approx 3$ °C per unit) to the U-containing tracts.

1.5 Aims of This Project

Previous work within the Davis group has focussed upon a *de novo* approach to producing a nonnatural system capable of recognising any random chosen sequence of double helical DNA, based on the behaviour of triplex-forming oligonucleotides. Preliminary computer based molecular modelling predicted that the (5,5,5)-tricyclic heterocyclic systems **1.50** and **1.51** should have the appropriate shape and hydrogen bonding capabilities (for T-A/A-T and C-G/G-C base pair recognition, respectively) to span the *entire width* of the major groove. They should link through either end of the system to a repetitive backbone (PNA[1.70] or analogous), and make simultaneous contacts with both strands. It was intended to differentiate between A-T and T-A base pairs, and C-G and G-C base pairs by varying the linking point on each system. The use of point **X** to attach **1.50** or **1.51** to the backbone should selectively recognise T-A and C-G base pairs, respectively, while point **Y** attachment should preferentially bind A-T and G-C base pairs, respectively (**Figure 1.17**).

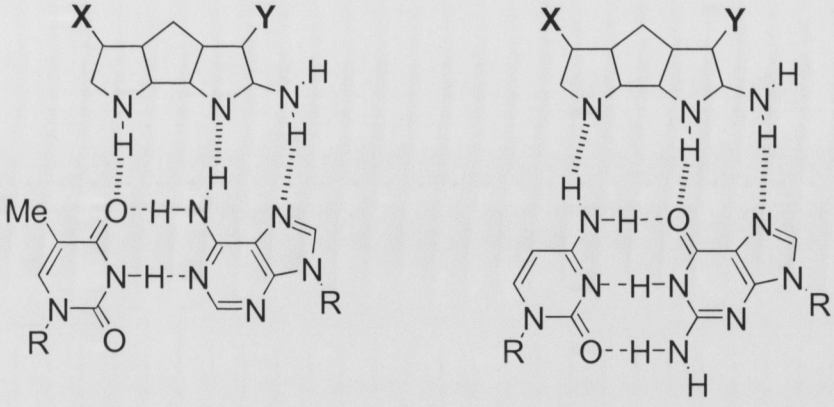
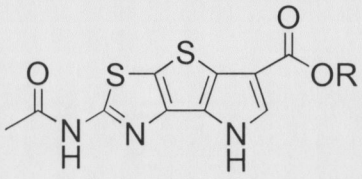


Figure 1.17 Proposed hydrogen bonding schemes for the **1.50**•T-A and **1.51**•C-G triplets.

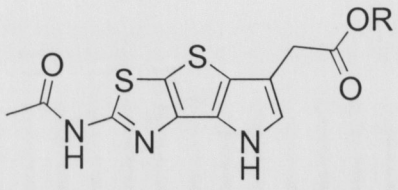
A combination of computer-based literature searches, molecular modelling and preliminary synthetic work produced molecule **1.52** for T-A base pair recognition.[1.71] The *N*-acetyl and ester groups were proposed to alleviate the electron rich nature of the system which seemed likely to cause instability. The *N*-acetyl group should also provide the added bulk required to fill the major groove, as well as enhance triplex stability through π-stacking and van der Waals contacts. Incorporation of the ligand into a backbone was to be mediated by the use of an ethoxycarbonylmethyl group, which would be obtained from the eventual transformation of the electron withdrawing ester group.



1.52

1.52 contains a thiazole ring, which results in the loss of a point of attachment **X** or **Y**. Molecular modelling studies showed that the attachment of **1.52** to the backbone *via* the acetyl group would provide an adequate alternative linkage point.

Previous work within the group had previously attempted to synthesise **1.52**,[1.71] but the synthetic difficulties encountered in trying to make this compound had veered efforts towards the production of **1.53**, as discussed in Chapter 5.



1.53

The long synthetic effort that had culminated with the synthesis of the (5,5,5)-tricycle **1.53** reached a disappointing but foreseen conclusion. The deprotected product appeared to be unstable. Electron rich systems are known to be susceptible to oxidative reactions that result in their decomposition into polymeric tars. An electron-withdrawing ester group placed directly in conjugation with the ring system is considered to reduce the electron density of the ring.[1.72] This hypothesis has resulted in a renewed effort towards the development of system **1.52**. An investigation of **1.52**•T-A formation in NMR binding study conditions would then be carried out. The next step would be to attach **1.52** to a backbone, and use affinity cleavage, DNase I footprint titration experiments, UV melting studies *etc.* to examine its behaviour with duplex DNA.

While a general recognition scheme has been originally proposed using the (5,5,5)-tricyclic system, a second tricyclic motif had been suggested by molecular modelling that would similarly occupy the major groove, and also offer the required hydrogen bonding patterns for recognition of all four base pairs. This idea had originated also from earlier work[1.71] with the squaramide unit **1.55**. Ballester, Costa *et al.* had previously shown that the distinctive hydrogen bond donor and acceptor capabilities on the squaramide unit **1.54** could be exploited for molecular recognition,[1.73] as shown for carboxylate recognition in **Figure 1.18**.

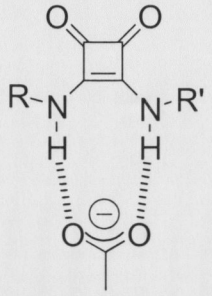
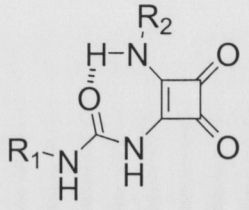


Figure 1.18 Carboxylate hydrogen bonding recognition scheme by squaramide **1.54** unit.

As a novel extension of the 3,4-monoalkyldisubstituted-squaramide theme, compound **1.55** possesses an ureidyl functionality, which predisposes the molecule to form an intramolecular hydrogen bond between the urea carbonyl and the amino proton.[1.74]



1.55

The displayed hydrogen-bonding pattern suggests that they might recognise C-G/G-C base pairs in organic solvent through interaction with the exocyclic amino group of cytosine and the N-7 and O-6 hydrogen acceptor groups of guanine, as shown in **Figure 1.19**.

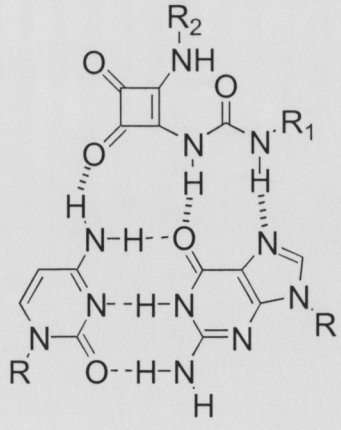


Figure 1.19 Proposed hydrogen bonding scheme for the **1.55**•C-G triplet.

The hydrogen-bonding pattern of **1.55** may also be compared with that of guanine, and so may recognise cytosine via conventional Watson-Crick hydrogen bonding interactions (**Figure 1.20**). This proposes that **1.55** might be capable of biological activity.

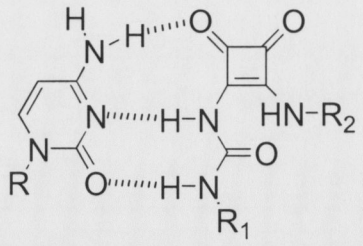


Figure 1.20 Proposed hydrogen bonding scheme of C-**1.55** pair.

The initial plan, therefore, was to synthesise a series of carbamoyl squaramides of the type **1.55** with varying alkyl and aryl substituents on the amino and urea functionalities. These molecules would then be tested for pharmacological activity in collaboration with an industrial partner.

The imposed rigidity presented by the (7,4)-framework of **1.55** was observed to be influential in suitably positioning the hydrogen bonding moieties, and so inspired the proposal of a (6,4,6)-tricyclic heterocyclic system for base pair recognition. Molecular modelling suggested that the proposed system **1.56** (**Figure 1.21**) would possess the required geometrical and hydrogen bonding characteristics to bind in the major groove of C-G/G-C base pairs and form a stable triplet with the complementary base pair.

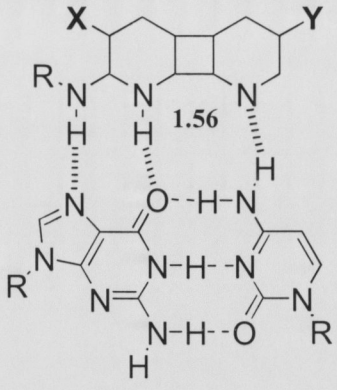
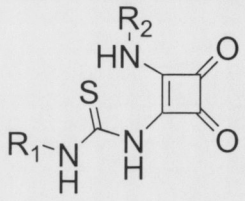
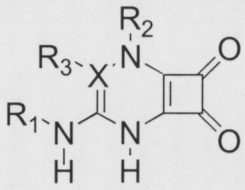


Figure 1.21 Proposed hydrogen bonding scheme for **1.56**•G-C triplet.

It was intended to use the squaric cyclobutenedione unit as the foundation stone upon which this system would be built. It was envisaged that the thiourea derivative **1.57** would provide a more synthetically facile route to forming the intermediate (6,4)-bicyclic structure **1.58**.



1.57



1.58

Various substituents would be considered to modulate the electron-rich nature of the system, as well as provide points of attachment for possible incorporation into oligonucleotides (should monomeric binding prove successful) and control the solubility properties and crystallinity of the products. The formation of **1.56**•C-G in non-polar solvents e.g. $CDCl_3$ would then be examined to complement the results obtained from the previous study of **1.55**•C-G systems.

On completion of both of these projects, it should be possible to recognise C-G/G-C, and T-A/A-T base pairs, and potentially any desired sequence of double-helical DNA. The medicinal and commercial benefits should be quite substantial.

[1.1] Picture courtesy of T.W. G. Solomons, *Organic Chemistry*, 5th Ed. **1992**, Wiley Books, 1146.

[1.2] S. Neidle, L. H. Pearl and J. V. Skelly, *Biochem. J.* **1987**, *243*, 1-13.

[1.3] G. Colmenarejo, F. Montero and G. Orellana, *Anti-Cancer Drug Design,* **1997**, *12*, 239-260.

[1.4] A. H. Krotz, L. Y. Kuo, T. P. Shields, and J. K. Barton, *J. Am. Chem. Soc.* **1993**, *115*, 3877-3882.

[1.5] A. H. Krotz, B. P. Hudson and J. K. Barton, *J. Am. Chem. Soc.* **1993**, *115*, 12577-12578.

[1.6] B. P. Hudson, C. M. Dupureur and J. K. Barton, *J. Am. Chem. Soc.* **1995**, *117*, 9379-9380.

[1.7] C. L. Kielkopf, K. E. Erkkila, B. P. Hudson, J. K. Barton and D. C. Rees, *Nat. Struct. Biol.* **2000**, *7(2)*, 117-121.

[1.8] R. H. Terbrueggen and J. K. Barton, *Biochemistry* **1995**, *34*, 8227-8234.

[1.9] R. H. Terbrueggen, T. W. Johann and J. K. Barton, *Inorg. Chem.* **1998**, *37*, 6874-6883.

[1.10] S. J. Franklin and J. K. Barton, *Biochemistry* **1998**, *37*, 16093-16105.

[1.11] D. T. Odom, C. S. Parker and J. K. Barton, *Biochemistry* **1999**, *38*, 5155-5163.

[1.12] R. E. Dickerson, *Scientific American* **1983**, *249*, 86-102.

[1.13] a) M. Coll, J. Aymami, G. A. van der Marel, J. H. van Boom, A. Rich, and A. H.-J. Wang, *Biochemistry* **1989**, *28*, 310-320; b) M. Coll, C. A. Frederick, A. H.-J. Wang, and A. Rich, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 8385-8389.

[1.14] J. W. Lown, K. Krowicki, U. G. Bhat, A. Skorobogaty, B. Ward and J. C. Dabrowiak, *Biochemistry* **1986**, *25*, 7408-7416.

[1.15] W. S. Wade, M. Mrksich and P. B. Dervan, *J. Am. Chem. Soc.* **1992**, *114*, 8783-8794.

[1.16] M. Mrksich, M. E. Parks and P. B. Dervan, *J. Am. Chem. Soc.* **1994**, *116*, 7983-7988.

[1.17] M. E. Parks, E. E. Baird and P. B. Dervan, *J. Am. Chem. Soc.* **1996**, *118*, 6153-6159.

[1.18] D. M. Herman, E. E. Baird and P. B. Dervan, *J. Am. Chem. Soc.* **1998**, *120*, 1382-1391.

[1.19] J. J. Kelly, E. E. Baird and P. B. Dervan, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 6981-6985.

[1.20] C. L. Kielkopf, E. E. Baird, P. B. Dervan and D. C. Rees, *Nat. Struct. Biol.* **1998**, *5(2)*, 104-109.

[1.21] J. W. Trauger, E. E. Baird, M. Mrksich and P. B. Dervan, *J. Am. Chem. Soc.* **1996**, *118*, 6160-6166.

[1.22] R. Peláez Lamamie de Clairac, C. J. Seel, B. H. Geierstanger, M. Mrksich, E. E. Baird, P. B. Dervan and D. E. Wemmer, *J. Am. Chem. Soc.* **1999**, *121*, 2956-2964.

[1.23] J. W. Trauger, E. E. Baird and P. B. Dervan, *Angew. Chem. Int. Ed.* **1998**, *37*, 1421-1423.

[1.24] J. W. Trauger, E. E. Baird and P. B. Dervan, *J. Am. Chem. Soc.* **1998**, *120*, 3534-3535

[1.25] S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird and P. B. Dervan, *Nature* **1998**, *391*, 468-471.

[1.26] C. L. Kielkopf, S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan and D. C. Rees, *Science* **1998**, *282*, 111-115.

[1.27] A. R. Urbach, J. W. Szewczyk, S. White, J. M. Turner, E. E. Baird and P. B. Dervan, *J. Am. Chem. Soc.* **1999**, *121*, 11621-11629.

[1.28] S. White, J. M. Turner, J. W. Szewczyk, E. E. Baird and P. B. Dervan, *J. Am. Chem. Soc.* **1999**, *121*, 260-261.

[1.29] C. Melander, D. M. Herman and P. B. Dervan, *Chem. Eur. J.* **2000**, *6(24)*, 4487-4497.

[1.30] U. Ellervik, C. C. C. Wang, and P. B. Dervan, *J. Am. Chem. Soc.* **2000**, *122*, 9354-9360.

[1.31] A. R. Urbach and P. B. Dervan, *Proc. Nat. Acad. Sci. USA* **2001**, *98(8)*, 4343-4348.

[1.32] R. E. Bremer, J. W. Szewczyk, E. E. Baird and P. B. Dervan, *Bioorg. Med. Chem.* **2000**, *8*, 1947-1955.

[1.33] C. C. C. Wang, U. Ellervik and P. B. Dervan, *Bioorg. Med. Chem.* **2001**, *9*, 653-657.

[1.34] P. E. Florancig, S. E. Swalley, J. W. Trauger and P. B. Dervan, *J. Am. Chem. Soc.* **2000,** *122*, 6342-6350.

[1.35] Z.-Y. J. Zhan and P. B. Dervan, *Bioorg. Med. Chem.* **2001**, *9*, 2467-2474.

[1.36] D. H. Nyugen, J. W. Swewczyk, E. E. Baird and P. B. Dervan, *Bioorg. Med. Chem.* **2001**, *9*, 7-17.

[1.37] G. Felsenfeld, D. R. Davies and A. Rich, *J. Am. Chem. Soc.* **1957**, *79*, 2023-2024.

[1.38] a) H. E. Moser and P. B. Dervan, *Science* **1987**, *238*, 645-650; b) T. Le Doan, L. Perrouault, D. Praseuth, N. Habhoub, J.-L. Decout, N. T. Thuong, J. Lhomme and C. Hélène, *Nucleic Acids Res.* **1987**, *15*, 7749-7760.

[1.39] M. D. Keppler and K. R. Fox, *Nucleic Acids Res.* **1997**, *25(22)*, 4644-4649.

[1.40] W. M. Olivas and L. J. Maher III, *Biochemistry* **1995**, *34*, 278-284.

[1.41] J. S. Lee, M. L. Woodsworth, L. J. P. Latimer and A. R. Morgan, *Nucleic Acids Res.* **1984**, *12* (*16*), 6603-6614.

[1.42] T. J. Povsic and P. B. Dervan, *J. Am. Chem. Soc.* **1989**, *111*, 3059-3061; b) L. E. Xodo, G. Manzini, F. Quadrifoglio, G. A. van der Marel and J. H. van Boom, *Nucleic Acids Res.* **1992**, *19(20)*, 5625-5631.

[1.43] G. Xiang, R. Bogacki and L. W. McLaughlin, *Nucleic Acids Res.* **1996**, *24 (10)*, 1963-1970.

[1.44] S. A. Cassidy, P. Slickers, J. O. Trent, D. C. Capaldi, P. D. Roselt, C. B. Reese, S. Neidle and K. R. Fox, *Nucleic Acids Res.* **1997**, *25(24)*, 4891-4898.

[1.45] L. C. Griffin and P. B. Dervan, *Science* **1989**, *245*, 967-971.

[1.46] G. C. Best and P. B. Dervan, *J. Am. Chem. Soc.* **1995**, *117*, 1187-1193.

[1.47] D. M. Gowers and K. R. Fox, *Nucleic Acids Res.* **1997**, *25(19)*, 3787-3794.

[1.48] D. M. Gowers, J. Bijapur, T. Brown and K. R. Fox, *Biochemistry* **1999**, *38*, 13747-13748.

[1.49] D. A. Gianolio and L. W. McLaughlin, *J. Am. Chem. Soc.* **1999**, *121*, 6334-6335.

[1.50] J. Sung Koh and P. B. Dervan, *J. Am. Chem. Soc.* **1992**, *114*, 1470-1479.

[1.51] P. S. Miller, P. Bhan, C. D. Cushman and T. L. Trapane, *Biochemistry* **1992**, *31*, 6788-6793.

[1.52] P. S. Miller, G. Bi, S. A. Kipp, V. Fok and R. K. Delong, *Nucleic Acids Res.* **1996**, *24(4)*, 730-736.

[1.53] W. A. Greenberg and P. B. Dervan, *J. Am. Chem. Soc.* **1995**, *117*, 5016-5022.

[1.54] T. S. Rao, R. H. Durland, D. M. Seth, M. A. Myrick, V. Bodepudi and G. R. Revankar, *Biochemistry* **1995**, *34*, 765-772.

[1.55] W. M. Olivas and L. J. Maher III, *Nucleic Acids Res.* **1995**, *23*, 1936-1941.

[1.56] V. M. Marathias, M. J. Sawicki and P. H. Bolton, *Nucleic Acids Res.* **1999**, *27(14)*, 2860-2867.

[1.57] P. Brodin, J.-S. Sun, J.-F. Mouscadet and C. Auclair, *Nucleic Acids Res.* **1999**, *27(15)*, 3029-3034.

[1.58] L. C. Griffin, L. L. Kiessling, P. A. Beal, P. Gillespie and P. B. Dervan, *J. Am. Chem. Soc.* **1992**, *114*, 7976-7982.

[1.59] K. M. Koshlap, P. Gillespie, P. B. Dervan and J. Feigon, *J. Am. Chem. Soc.* **1993**, *115*, 7908-7909.

[1.60] E. Wang, K. M. Koshlap, P. Gillespie, P. B. Dervan and J. Feigon, *J. Mol. Biol.* **1996**, *257*, 1052-1069.

[1.61] C.-Y. Huang, C. D. Cushman and P. S. Miller, *J. Org. Chem.* **1993**, *58*, 5048-5049.

[1.62] C.-Y. Huang, G. Bi and P. S. Miller, *Nucleic Acids Res.* **1996**, *24(13)*, 2606-2613.

[1.63] S. C. Zimmerman and P. Schmitt, *J. Am. Chem. Soc.* **1995**, *117*, 10769-10770.

[1.64] E. Mertz, S. Mattei and S. C. Zimmerman, *Org. Lett.* **2000**, *2(19)*, 2931-2934.

[1.65] T. E. Lehmann, W. A. Greenberg, D. A. Liberles, C. K. Wada and P. B. Dervan, *Helv. Chim. Acta* **1997**, *80*, 2002-2022.

[1.66] D. Lengeler and K. Weisz, *Tetrahedron Lett.* **2001**, *42*, 1479-1481.

[1.67] S. Sasaki, S. Nakashima, F. Nagatsugi, Y. Tanaka, M. Hisatome and M. Maeda, *Tetrahedron Lett.* **1995**, *36(52)*, 9521-9524.

[1.68] J. H. Rothman and W. G. Richards, *Biopolymers* **1996**, *39*, 795-812.

[1.69] A. B. Eldrup, O. Dahl and P. E. Nielsen, *J. Am. Chem. Soc.* **1997**, *119*, 11116-11117.

[1.70] a) M. Egholm, O. Buchardt, P. E. Nielsen, and R. H. Berg, *J. Am. Chem. Soc.* **1992**, *114*, 1895-1897; b) P. E. Nielsen and G. Haaima, *Chem. Soc. Rev.* **1997**, 73-78.

[1.71] G. P. Dunne, Ph.D. Thesis, University of Dublin, **1999**.

[1.72] G. P. Dunne, Postdoctoral Report, University of Dublin, **2000**.

[1.73] a) S. Tomàs, R. Prohens, M. Vega, M. C. Rotger, P. M. Deyà, P. Ballester, and A. Costa, *J. Org. Chem.* **1996**, *61*, 9394-9401; b) R. Prohens, S. Tomàs, J.Morey, P. M. Deyà, P. Ballester, and A. Costa, *Tetrahedron Lett.* **1998**, *39*, 1063-1066.

[1.74] A. P. Davis, S. M. Draper, G. Dunne and P. Ashton, *Chem. Commun.* **1999**, 2265-2266.

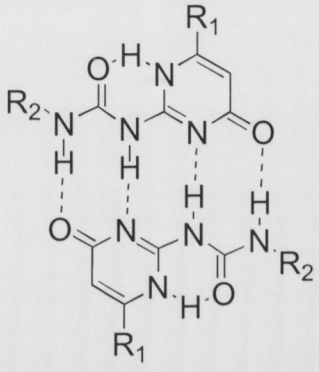
Chapter Two

The *N*-Carbamoyl Squaramide System

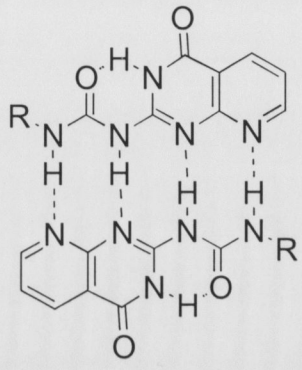
2.1 Preamble

The aim of this project, as stated previously, was to synthesise tricyclic heterocyclic systems which would recognise double stranded DNA through non-covalent and steric interactions. The primary recognition feature of these ligands would be the series of hydrogen bonding interactions that are possible between each ligand and their complementary base pair.

Such intermolecular hydrogen bonding interactions have been observed to play key roles in supramolecular self-assembly, as in the self-complementary 'DDAA' systems **2.1** and **2.2** studies by the groups of Meijer and Zimmerman,[2.1] respectively.

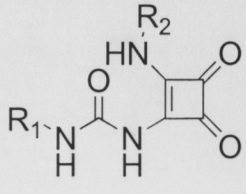


2.1

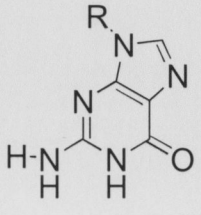


2.2

Ballester, Costa *et al.* had shown that the squaramide unit **1.54** displayed a distinctive H-bonding array for molecular recognition.[1.74] Previous work within the Davis group had identified that the novel *N*-carbamoyl squaramide unit **1.55** was capable of self-association.[1.71] A similarity between the hydrogen bonding array offered by squaramide **1.55** and that offered by guanine derivatives **2.3** was observed, as shown in **Figure 2.1**. This would suggest that such squaramide systems might be capable of biological activity.



1.55



2.3

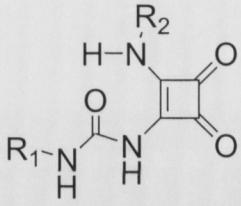
Figure 2.1 Structural comparisons between squaramide unit **1.55** and guanine derivative **2.3**.

The diverse chemistry of squaric acid (3,4-dihydroxy-3-cyclobutene-1,2-dione, **2.4**) derivatives has received a lot of attention since the first synthesis of squaric acid **2.4** in 1959.[2.2, 2.3] However, a considerable amount of synthetic work pertaining to these derivatives has been patented, due to their success as smooth muscle relaxants and anti-hypertensive preparations. This evidence would support the argument that amino *N*-carbamoyl squaramides may be worthy of pharmacological studies.

This chapter outlines the previous work carried out within the group on squaramide systems **1.55**, followed by a detailed description of the syntheses of a variety of alkyl/arylamino and hydrazino functionalised phenylcarbamoyl squaramides, intended for biological activity. It was intended to follow the synthetic routes elucidated previously[1.71] towards the syntheses of these *N*-carbamoyl squaramides, with modifications where necessary.

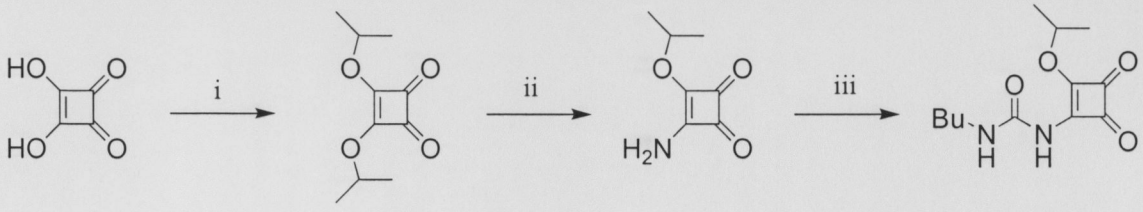
2.2 Synthesis of *N*-Carbamoyl Squaramides 1.55a-b

Contribution of Gavin Dunne



1.55a: $R_1 = R_2 = Bu$
1.55b: $R_1 = Bu$, $R_2 =$ Cyclohexyl

Squaramide systems **1.55a-b** were synthesised as G-C recognising ligands. The syntheses originated from the production of diisopropyl squarate **2.5** from squaric acid **2.4**, based on a modification of the method of Liebeskind *et al.*[2.4] **2.5** was aminated based on a procedure of Liebeskind *et al.*[2.5] to produce the ester squaramide **2.6**, as shown in **Scheme 2.1**.



2.4 **2.5** **2.6** **2.7**

Scheme 2.1 *Reagents, conditions (and yields)*: (i) 2-Propanol, benzene, Dean and Stark apparatus, 100 °C, 48 h; (ii) $NH_{3(g)}$, MeOH, THF, RT, 20 min (75 %); (iii) BuNCO, Et_3N, CH_3CN, RT, 18 h (64 %).

Butyl isocyanate and triethylamine were reacted with ester squaramide **2.6** to produce *N*-carbamoyl squaramide **2.7**. Condensation of **2.7** with butylamine yielded carbamoyl squaramide **1.55a**, while condensation of **2.7** with cyclohexylamine in a similar methodology produced system **1.55b**.

The nucleoside recognition properties of units **1.55a** and **1.55b** were then investigated by ^{1}H NMR binding studies in CD_3CN. However, these systems appeared to self-associate strongly under the conditions of the experiment. Various ^{1}H NMR dilution studies in CD_3CN were then conducted, and it was concluded that K_D (the dimerisation constant) was in the region of 10^3 M^{-1} for each system. An X-ray crystal structure of system **1.55b** confirmed that it existed as a centrosymmetric dimeric entity, apparently held together in the solid state by four hydrogen bonds from the two urea NH to the C-1 carbonyl group, as seen in **Figure 2.2**. An intramolecular H-bond from the amino proton of the cyclohexylamino group to the urea carbonyl oxygen was observed, which preorganised the squaramide molecule into a (7,4)-bicyclic framework capable of self-association via the 'DDAA' hydrogen bonding motif.

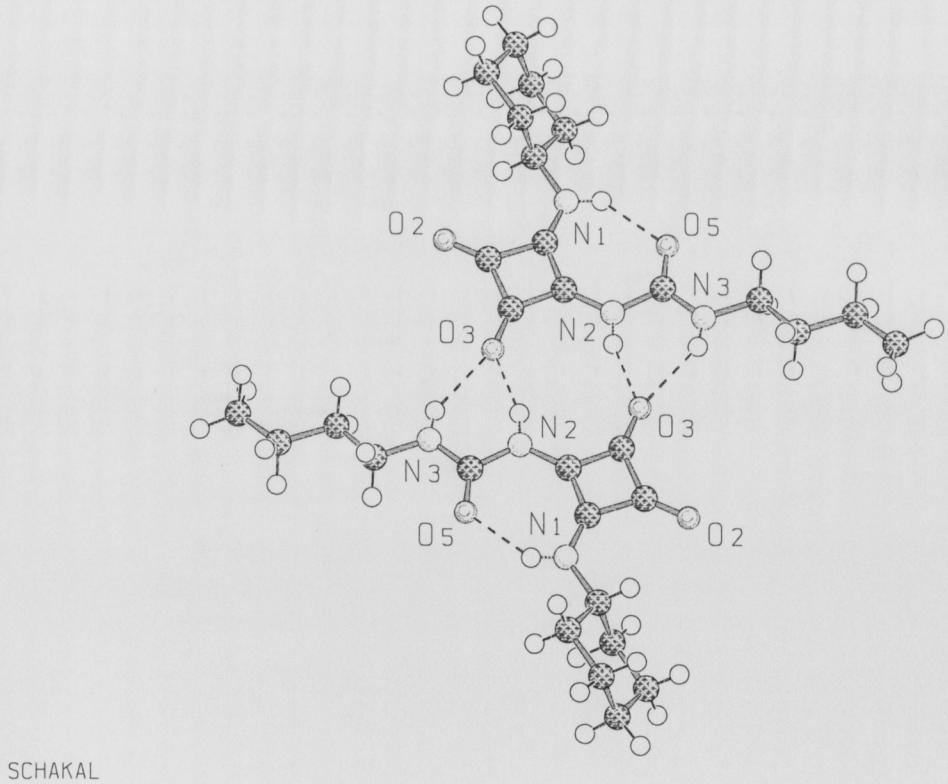
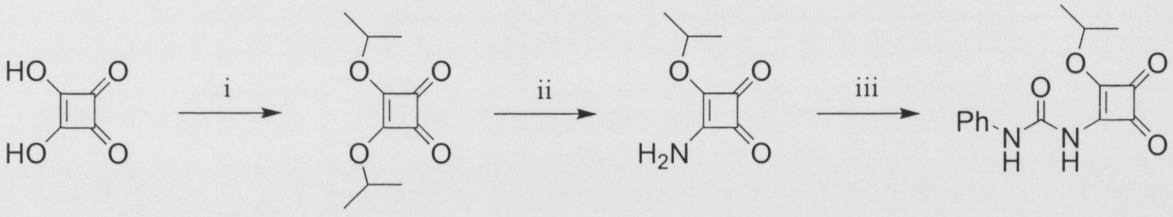


Figure 2.2 Crystal structure of **1.55b**. Selected distances (Å) are N(2)H·· O(3) 1.937, N(3)H·· O(3) 2.108, N(1)H·· O(5) 2.120.

2.3 Synthesis of *N*-Carbamoyl Squaramides 2.8-2.22

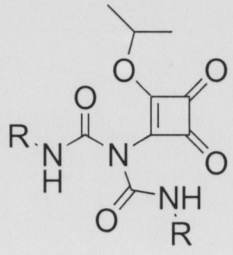
In an analogous synthetic approach to that discussed in Section 2.2, a series of *N*-carbamoyl squaramides were then synthesised. The synthesis of the phenylcarbamoyl squaramide **2.8** is detailed in **Scheme 2.2**. The diisopropyl squarate **2.5** was obtained from **2.4** by a modification of Liebeskind *et al.*[2.4] It was found during this work that the use of *p*-TsOH as catalyst resulted in shorter reaction times (1 d instead of 2 d). The low purity of commercially available **2.5** prompted its preparation throughout this project. **2.5** was aminated based on a procedure of Liebeskind *et al.*[2.5] to produce the isopropoxysquaramide **2.6**.



2.4 **2.5** **2.6** **2.8**

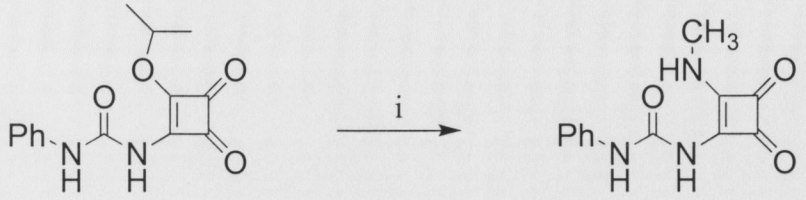
Scheme 2.2 *Reagents, conditions (and yields)*: (i) 2-Propanol, benzene, Dean and Stark apparatus, 100 °C, 24 h (94%); (ii) $NH_{3(g)}$, MeOH, THF, RT, 10 min (89 %); (iii) PhNCO, DBU, CH_3CN, RT, 24 h (72 %).

The phenylcarbamoyl squaramide **2.8** was obtained by the treatment of excess **2.6** with phenyl isocyanate and DBU in acetonitrile in a modification of the method used by Dunne.[1.74] Lower yields were obtained if a 1:1 ratio of squaramide **2.6** and isocyanate were used. It was proposed previously[1.71] that **2.6** tended to undergo reaction twice at the amino position with the isocyanate to form a compound of hypothesised structure **2.9**. The use of DBU as base was found to result in faster reaction times than those reactions where DIPEA was used as base, and acetonitrile also promoted faster and cleaner reactions than analogous THF experiments.



2.9

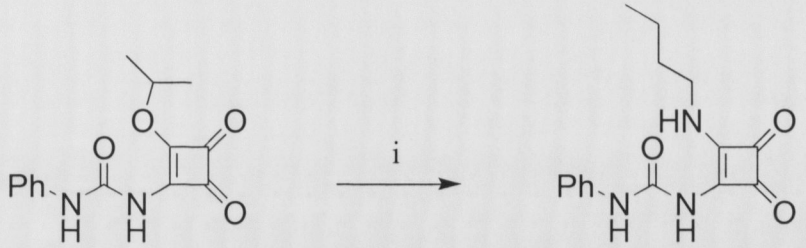
A series of amino and hydrazino phenylcarbamoyl squaramides were then prepared from **2.8** by condensation of the ester squaramide with a variety of primary amines and hydrazines. **Scheme 2.3** details the preparation of carbamoyl squaramide **2.10** from methylamine and **2.8**.



2.8 **2.10**

Scheme 2.3 *Reagents, conditions (and yield)*: (i) Methylamine, DCM, RT, overnight (96 %).

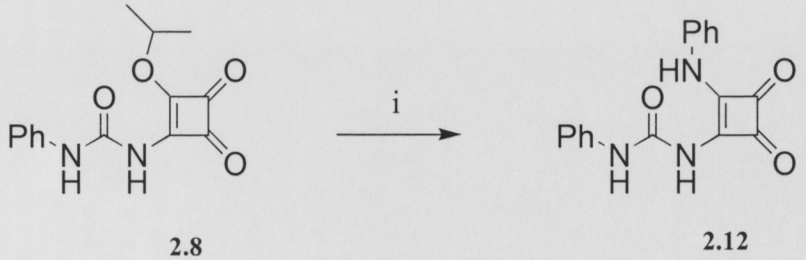
Squaramide **2.11** was synthesised from **2.8** and butylamine, as shown in **Scheme 2.4**.



2.8 **2.11**

Scheme 2.4 *Reagents, conditions (and yield)*: (i) Butylamine, DCM, RT, overnight (47 %).

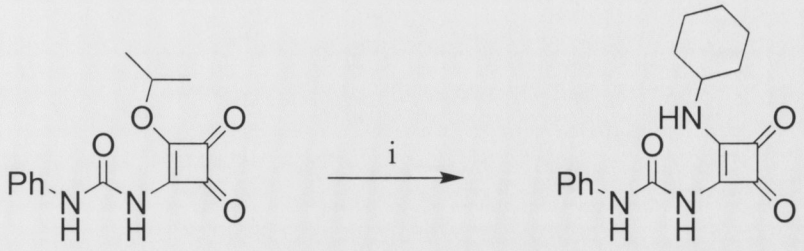
Aniline and **2.8** yielded analogue **2.12**, shown in **Scheme 2.5**.



2.8 **2.12**

Scheme 2.5 *Reagents, conditions (and yield)*: (i) Aniline, DCM, RT, overnight (78 %).

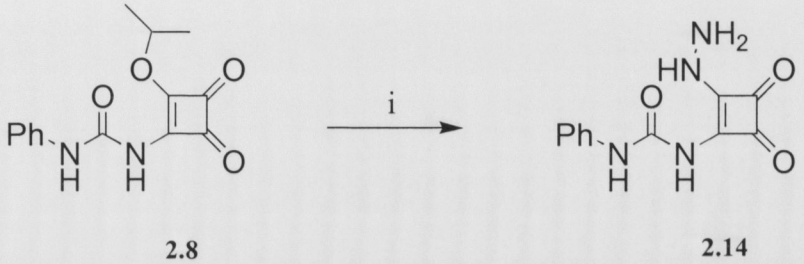
Condensation of **2.8** with cyclohexylamine under similar reaction conditions yielded **2.13**, shown in **Scheme 2.6**.



2.8 **2.13**

Scheme 2.6 *Reagents, conditions (and yield)*: (i) Cyclohexylamine, DCM, RT, overnight (52 %).

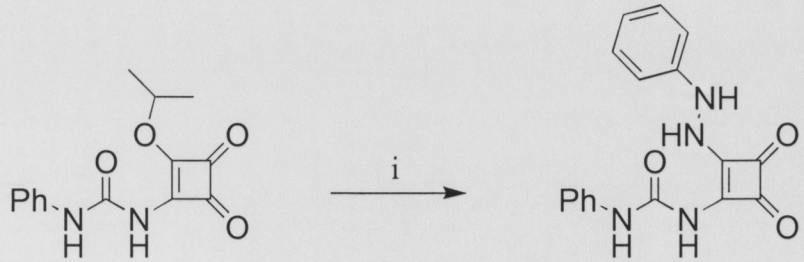
Hydrazine hydrate and **2.8** condensed in a THF solution to form **2.14**, as in **Scheme 2.7**.



2.8 **2.14**

Scheme 2.7 *Reagents, conditions (and yield)*: (i) Hydrazine hydrate, THF, RT, overnight (quant.).

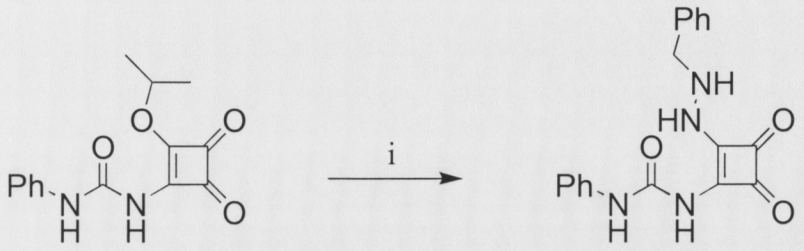
The two carbonyls on the cyclobutenedione ring are known to behave as carboxylic acid carbonyl groups, and so do not condense with phenylhydrazine.[2.3] As such, the reaction of **2.8** with phenylhydrazine to form **2.15** was possible (**Scheme 2.8**). However, this experiment was prone to over-reaction, resulting in the formation of a red tar.



2.8 **2.15**

Scheme 2.8 *Reagents, conditions (and yield)*: (i) Phenylhydrazine, THF, RT, 3 h (72 %).

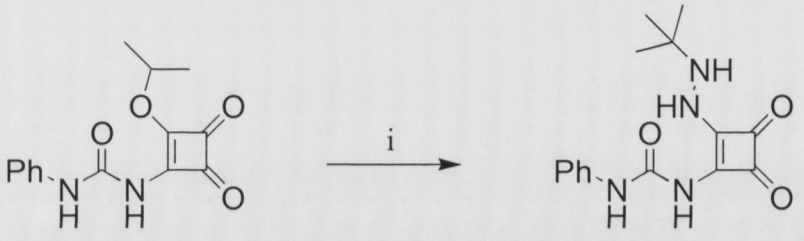
In **Scheme 2.9**, the reaction of **2.8** with benzylhydrazine to yield **2.16** is shown.



2.8 **2.16**

Scheme 2.9 *Reagents, conditions (and yield)*: (i) Benzylhydrazine dihydrochloride, DIPEA, CH_3CN, RT, overnight (47 %).

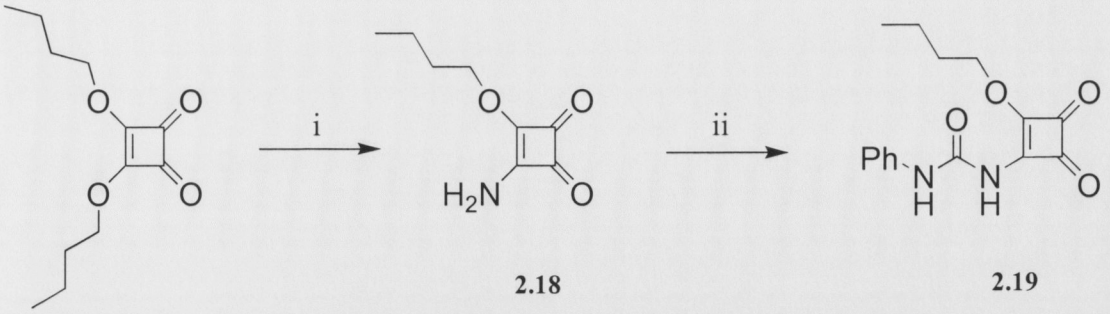
Scheme 2.10 details the condensation of *tert*-butylhydrazine with **2.8** in CH_3CN to form **2.17**.



2.8 **2.17**

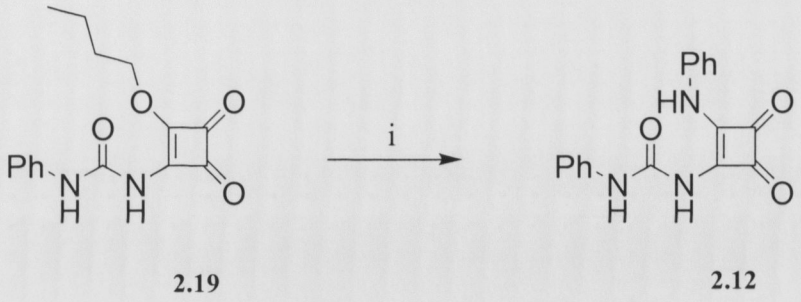
Scheme 2.10 *Reagents, conditions (and yield)*: (i) *tert*-Butylhydrazine hydrochloride, DIPEA, CH_3CN, 50 °C, 3 h (33 %).

In a parallel methodology, a selection of these amino and hydrazino phenylcarbamoyl squaramides were prepared from starting material **2.19**, which was obtained from amination of dibutyl squarate, to form ester squaramide **2.18**, which was then reacted with phenyl isocyanate and DBU, as in the synthesis of **2.8**. The synthetic outline is displayed in **Scheme 2.11**.



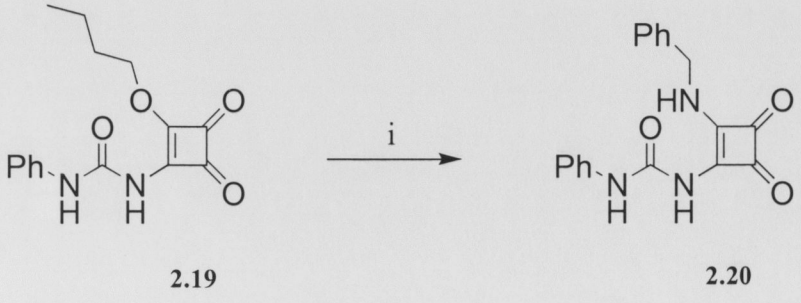
Scheme 2.11 *Reagents, conditions (and yields)*: (i) $NH_{3(g)}$, MeOH, THF, RT, 10 min (92 %); (ii) PhNCO, DBU, CH_3CN, RT, 24 h (86 %).

2.12 was prepared from **2.19** (**Scheme 2.12**) in an analogous manner to its preparation from **2.8** and aniline.



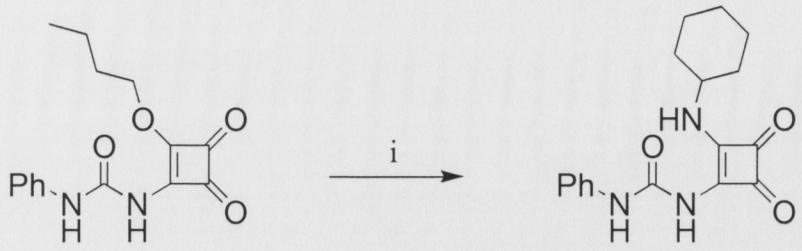
Scheme 2.12 *Reagents, conditions (and yield)*: (i) Aniline, DCM, RT, overnight (37 %).

Benzylamine was added to a DCM solution of **2.19** to produce **2.20** in a similar manner, as outlined in **Scheme 2.13**.



Scheme 2.13 *Reagents, conditions (and yield)*: (i) Benzylamine, DCM, RT, overnight (51 %).

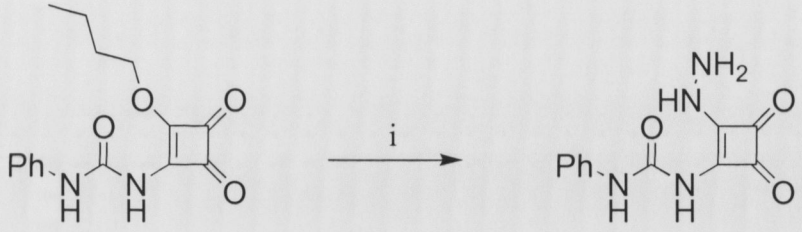
Scheme 2.14 details the synthesis of **2.13** from parent squaramide **2.19** and cyclohexylamine.



2.19 **2.13**

Scheme 2.14 *Reagents, conditions (and yield)*: (i) Cyclohexylamine, DCM, RT, overnight (82 %).

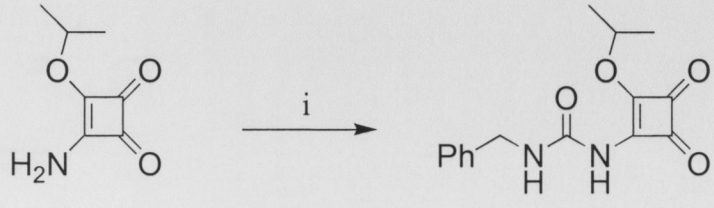
Hydrazine hydrate was added to a CH_3CN solution of **2.19** to form **2.14** as before, shown below in **Scheme 2.15**.



2.19 **2.14**

Scheme 2.15 *Reagents, conditions (and yield)*: (i) Hydrazine hydrate, CH_3CN, RT, overnight (91 %).

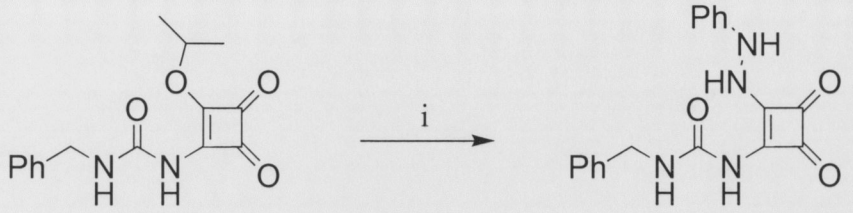
To extend the series of carbamoyl squaramides even further, squaramide **2.7** was reacted with benzyl isocyanate and DBU to yield ester carbamoyl squaramide **2.21**, seen in **Scheme 2.17**.



2.7 **2.21**

Scheme 2.16 *Reagents, conditions (and yield)*: (i) Benzyl isocyanate, DBU, CH_3CN, RT, 150 min (65 %).

The ester carbamoyl squaramide **2.21** was then reacted with phenylhydrazine (**Scheme 2.17**) to yield squaramide **2.22**. This conversion was also prone to over-reaction, with the formation of a similar red tar as before.



2.21 **2.22**

Scheme 2.17 *Reagents, conditions (and yield)*: (i) Phenylhydrazine, DCM, ethanol, RT, 2 h (18 %).

It had been intended to test these compounds for pharmacological activity in collaboration with an industrial partner. Unfortunately, such a collaboration could not be initiated before the completion of this project. It is anticipated that some very interesting results will be obtained when these compounds are tested.

[2.1] (a) R. P. Sijbesma, F. H. Beijer, L. Brunsveld, B. J. B. Folmer, J. H. K. K. Hirschberg, R. F. M. Lange, J. K. L. Lowe, and E. W. Meijer, *Science* **1997**, *278*, 1601-1604; (b) F. H. Beijer, R. P. Sijbesma, H. Kooijman, A. L. Spek, and E. W. Meijer, *J. Am. Chem. Soc.* **1998**, *120*, 6761-6769; (c) P. S. Corbin and S. C. Zimmerman, *J. Am. Chem. Soc.* **1998**, *120*, 9710-9711.

[2.2] (a) G. Maahs and P. Hegenberg, *Angew. Chem. Int. Ed.* **1966**, *5(10)*, 888-893; (b) A. H. Schmidt and W. Ried, *Synthesis* **1978**, 1-22.

[2.3] A. H. Schmidt, *Synthesis* **1980**, 961-994.

[2.4] L. S. Liebeskind, R. W. Fengl, K. R. Wirtz and T. T. Shawe, *J. Org. Chem.* **1988**, *52*, 2482-2488.

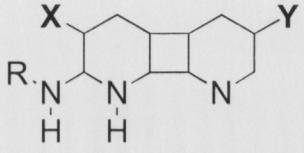
[2.5] L. S. Liebeskind, M. S. Yu and R. W. Fengl, *J. Org. Chem.* **1993**, *58*, 3543-3549.

Chapter Three

The Bicyclo[4.2.0]octa-1(6)-ene-7,8-dione Systems

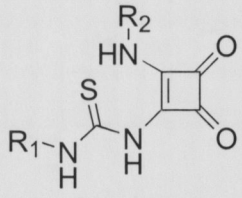
3.1 Preamble

The observation that the (7,4)-framework (provided by the *N*-carbamoyl squaramides described in Chapter 2) bore an intriguing resemblance to guanine inspired the proposal of a novel DNA binding ligand. It is known that binding can be improved by the provision of a rigid platform on which to position any hydrogen bonding sites, instead of a weakly binding flexible chain. As such, a rigid (6,4,6)-tricyclic system, such as **1.56**, should ideally organise and maintain the required H-bonding interactions for C-G base pair recognition. Molecular modelling suggested that the proposed system **1.56** would possess the required geometrical and hydrogen bonding characteristics to bind the major groove of C-G/G-C base pairs, and form a stable triplet with the complementary base pair.

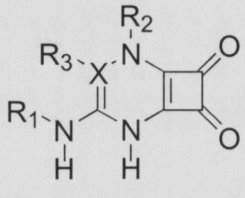


1.56

It was intended to use the squaric cyclobutenedione unit as the foundation stone upon which this system would be built. It was envisaged that the thiourea derivative **1.57** would provide a more synthetically facile route to forming the intermediate (6,4)-bicyclic structure **1.58**.



1.57



1.58

Cyclisation would be achieved through nucleophilic attack at the carbon base of an alkylated thiocarbonyl group. Thiocarbonyl sulphur atoms are relatively easy to alkylate, due to the resonance interaction that exists between the non-bonding electron pair on nitrogen and the carbon-sulphur π-bond, which leads to a negatively charged sulphur atom with pronounced nucleophilicity,[3.1] as seen in **Figure 3.1**.

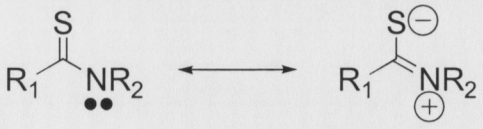


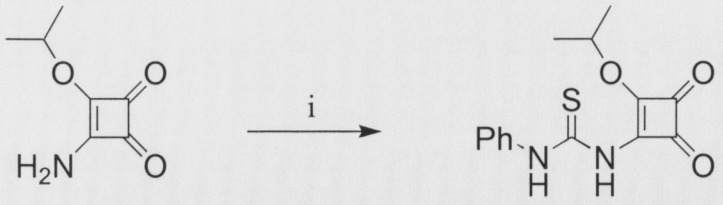
Figure 3.1 Resonance forms of the thiocarbonyl group.

The nucleophilic group intended for use in the desired intramolecular cyclisation was the hydrazine group. However, the resultant cyclised product would have a nitrogen atom with no free reaction sites at the position on the bicyclic system intended for attachment of the ring system to the backbone (i.e. in **1.58**, X = N, R_3 is redundant). While symmetry is not a requirement of the ligand (the target base pairs are not symmetrical), the elimination of a point of attachment at this nitrogen centre may prove to be of importance. An ethoxycarbonylmethylamino group was consequently considered as a suitable ligand for cyclisation, as the acidic CH_2 group next to the ester group should aid cyclisation, and the ester group provided a means of linking the ligand to a PNA backbone (i.e. in **1.58**, X = C, R_3 = COOEt).

This chapter describes the synthesis of hydrazino thiourea squaramides **3.3**, **3.5**, **3.7** and **3.9** in Section 3.2. Section 3.3 gives an account of the synthesis of the ethoxycarbonylmethylamino analogues **3.14**, **3.15**, **3.16** and **3.17**. The attempts to methylate these thiourea derivatives are detailed in Section 3.4.

3.2 Synthesis of Hydrazino Thiourea Derivatives 3.3, 3.5, 3.7 and 3.9

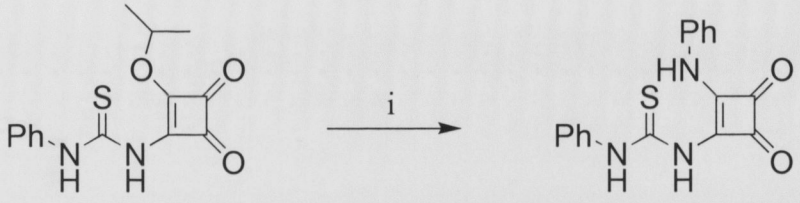
The synthesis of thiourea squaramide **3.1** is detailed in **Scheme 3.1**. The ester phenylthiourea squaramide **3.1** was synthesised from **2.6** by reaction with phenyl isothiocyanate and DBU in acetonitrile. The synthetic requirement of DBU in these carbamoyl/thiourea squaramide reactions was elucidated at this point, as DIPEA did not induce thiourea squaramide formation. It was noted that these reactions did not proceed as cleanly as their *N*-carbamoyl analogues, as several extra spots were observed by TLC analysis.



2.6 **3.1**

Scheme 3.1 *Reagents, conditions (and yield)*: (i) Phenyl isothiocyanate, DBU, CH_3CN, RT, overnight (95 %).

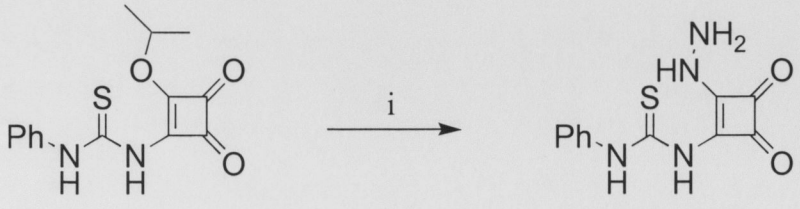
Model system **3.2**, an insoluble white solid, was synthesised from aniline and **3.1**, which should add an interesting variant to the carbamoyl squaramides synthesised previously for biological testing. It should be noted that an intramolecular hydrogen bond is not possible here, and so this ultimately results in a less rigid framework than the squaramides discussed in Chapter 2. This reaction strategy is shown in **Scheme 3.2**.



3.1 **3.2**

Scheme 3.2 *Reagents, conditions (and yield)*: (i) Aniline, DCM, ethanol, RT, overnight (73 %).

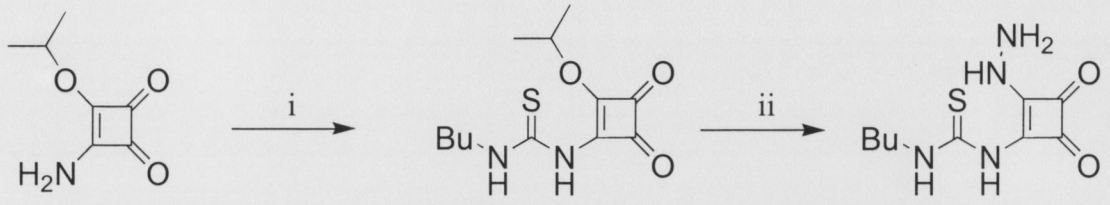
Hydrazine hydrate was condensed with **3.1** to form the desired intermediate **3.3**, as shown in **Scheme 3.3**.



3.1 **3.3**

Scheme 3.3 *Reagents, conditions (and yield)*: (i) Hydrazine hydrate, DCM, RT, 2 h (quant.).

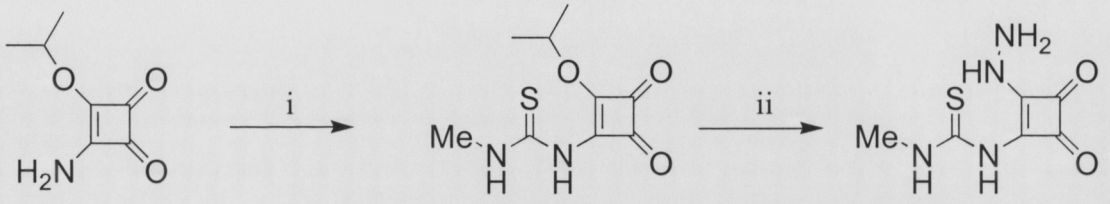
Hydrazino squaramide **3.5** was prepared as a means of expanding the range of possible exocyclic alkyl- and arylamino substituents on the completed tricyclic system. It was synthesised from hydrazination of **3.4**, which was prepared by reaction of **2.5** with butyl isothiocyanate and DBU, as detailed in **Scheme 3.4**.



2.6 **3.4** **3.5**

Scheme 3.4 *Reagents, conditions (and yield)*: (i) Butyl isothiocyanate, DBU, CH_3CN, RT, overnight (68 %); (ii) Hydrazine hydrate, DCM, RT, overnight (quant.).

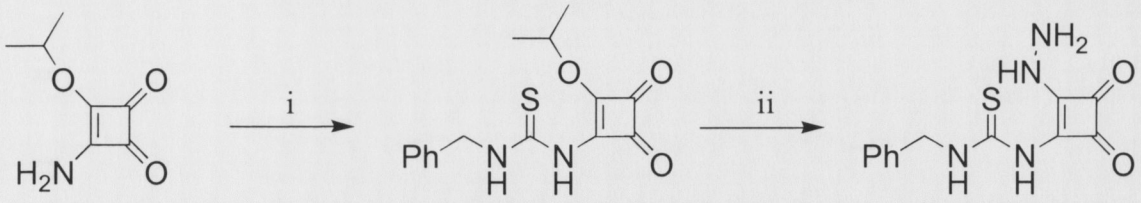
Similarly, the methylthiourea variant **3.7** was prepared, as shown in **Scheme 3.5**, from reaction of **3.6** with hydrazine hydrate. **3.6** was prepared from the reaction of methyl isothiocyanate and DBU with **2.6**.



2.6 **3.6** **3.7**

Scheme 3.5 *Reagents, conditions (and yield)*: (i) Methyl isothiocyanate, DBU, CH_3CN, RT, overnight (55 %); (ii) Hydrazine hydrate, DCM, RT, 3 h (82 %).

A fourth hydrazino thiourea derivative **3.9** was prepared from **2.6** by reaction with benzyl isothiocyanate and DBU to produce **3.8**, then reaction of **3.8** with hydrazine hydrate yielded **3.9**. **Scheme 3.6** details this reaction strategy.

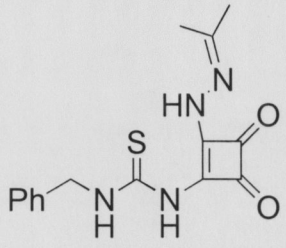


2.6 **3.8** **3.9**

Scheme 3.6 *Reagents, conditions (and yield)*: (i) Benzyl isothiocyanate, DBU, CH_3CN, RT, 90 min (81 %); (ii) Hydrazine hydrate, DCM, RT, overnight (89 %).

However, a problem with these hydrazino thiourea squaramides **3.3**, **3.5**, **3.7** and **3.9** as convenient synthetic intermediates to the bicyclic system was encountered from an early stage in the project. Although **3.5** was soluble enough in THF/DCM for chromatographic purification, polar compounds **3.3**, **3.7** and **3.9** were only moderately soluble in THF, EtOH, MeOH and CH_3CN, which compromised their purification, yield determination and use in further reaction steps.

Moreover, acetone appeared to be the solvent of choice in the dissolution of hydrazino derivatives **3.3**, **3.5**, **3.7** and **3.9**. However, it was very quickly discovered that these hydrazino derivatives decomposed from yellow solids into red tars in acetone, especially if heat was applied. It was apparent that these labile hydrazino derivatives immediately formed the hydrazone in the presence of acetone, and the resultant yellow hydrazone derivatives quickly decomposed, especially at elevated temperatures. The formation of the hydrazone of **3.9**, **3.10**, was confirmed by X-ray crystallography. The X-ray crystal structure obtained is shown in **Figure 3.2**.



3.10

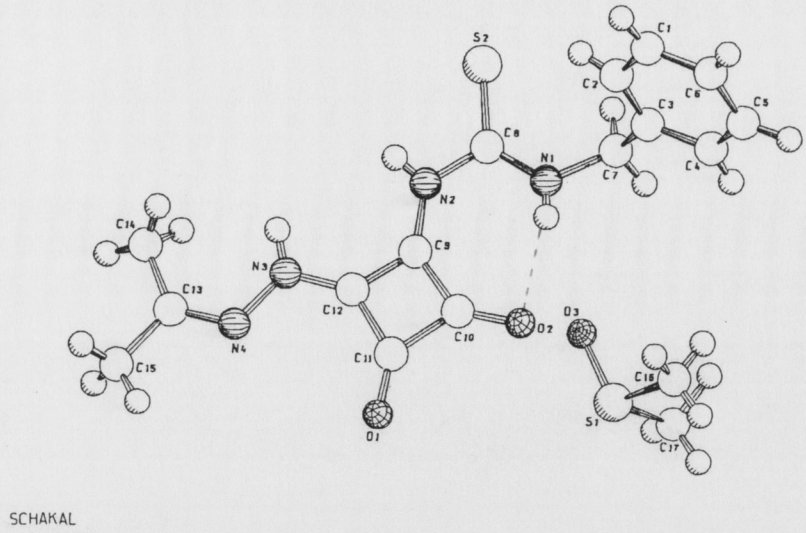
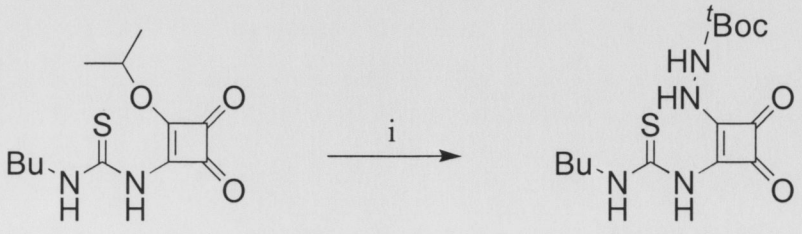


Figure 3.2 Crystal structure of hydrazone **3.10** with first layer solvent (DMSO) molecules. Selected distance (Å) is N1(H)·· O(2) 2.141.

The crystallographic evidence for an intramolecular hydrogen bond between the thiourea N1(H) and the ring carbonyl O2, as displayed in **Figure 3.2**, suggested that the cyclisation of these thiourea derivatives might not be as facile as originally thought. It appeared as if the intramolecular bond would prevent the thiocarbonyl group from facilitating cyclisation, by rigidly positioning the thiocarbonyl group away from the hydrazone, and by extrapolation, the free hydrazine. As this intramolecular hydrogen bond would be present in every thiourea derivative, and would not be affected by subsequent methylation, it indicated at this stage that there was a possibility that cyclisation might not occur. However, as empirical evidence might prove otherwise, this convincing argument did not dissuade efforts towards the production of system **1.58** via the original synthetic route at this point.

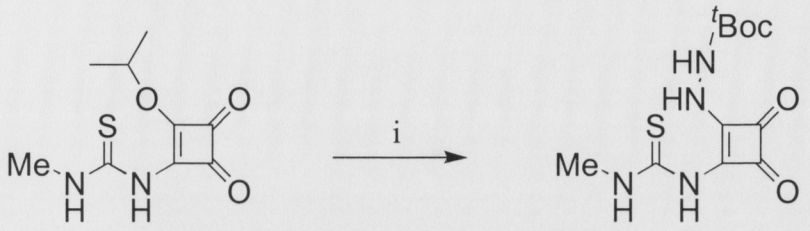
As a precaution against undesirable hydrazone formation, the ^{t}Boc-protected version of **3.5**, **3.11** was obtained by condensing *tert*-butyl carbazate with **3.4**, as shown in **Scheme 3.7**.



3.4 **3.11**

Scheme 3.7 *Reagents, conditions (and yield)*: (i) *tert*-Butyl carbazate, CH_3CN, RT, overnight (87 %).

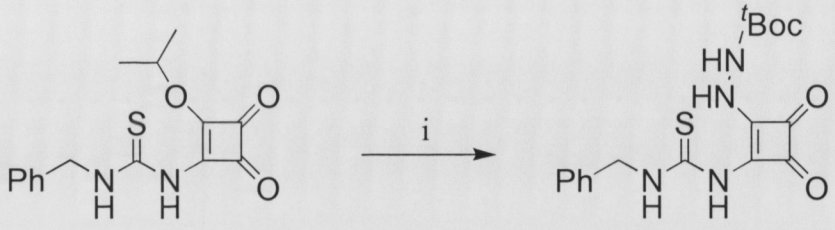
Similarly, **3.12** was formed from the reaction of **3.6** with *tert*-butyl carbazate, seen in **Scheme 3.8**.



3.6 **3.12**

Scheme 3.8 *Reagents, conditions (and yield)*: (i) *tert*-Butyl carbazate, CH_3CN, RT overnight, then 80 °C for 5 min (73 %).

Scheme 3.9 details the reaction of **3.8** with *tert*-butyl carbazate with heating to form **3.13**.

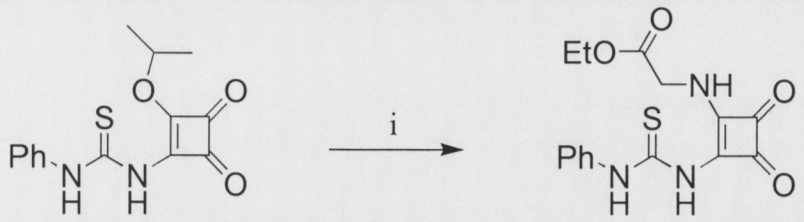


3.8 **3.13**

Scheme 3.9 *Reagents, conditions (and yield)*: (i) *tert*-Butyl carbazate, CH_3CN, RT overnight, then 80 °C for 5 min (59 %).

3.3 Synthesis of Ethoxycarbonylmethylamino Thiourea Derivatives 3.14, 3.15, 3.16 and 3.17

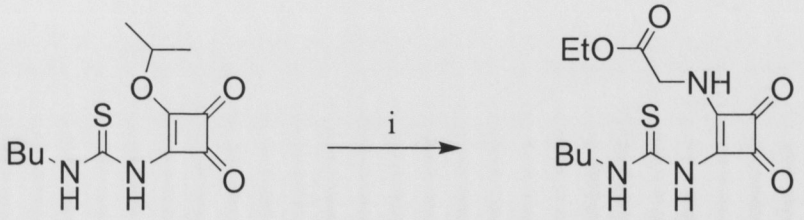
With the aim of synthesising a tricyclic system containing two linkage points to both backbone strands as originally designed, the ethoxycarbonylmethylamino derivative **3.14** was prepared from reaction of **3.1** with glycine ethyl ester hydrochloride and DIPEA, as shown in **Scheme 3.10**.



3.1 **3.14**

Scheme 3.10 *Reagents, conditions (and yield)*: (i) Glycine ethyl ester hydrochloride, DIPEA, THF, RT overnight (69 %).

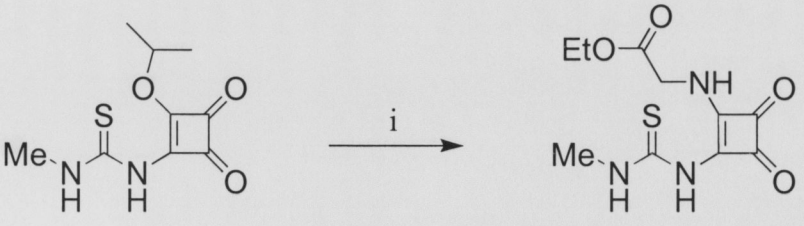
Scheme 3.11 outlines the similar methodology used in the synthesis of **3.15** from **3.4** and glycine ethyl ester.



3.4 **3.15**

Scheme 3.11 *Reagents, conditions (and yield)*: (i) Glycine ethyl ester hydrochloride, DIPEA, THF, RT overnight (85 %).

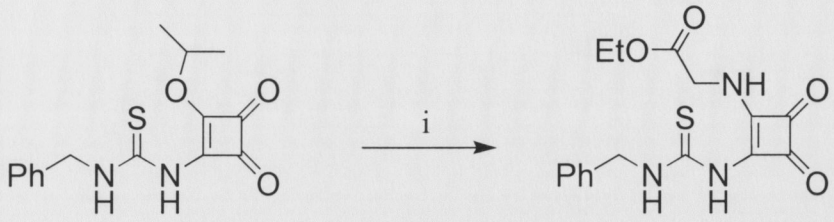
Ethoxycarbonylmethylamino analogue **3.16** was synthesised in a similar manner by the reaction of glycine ethyl ester with **3.6** (**Scheme 3.12**).



3.6 **3.16**

Scheme 3.12 *Reagents, conditions (and yield)*: (i) Glycine ethyl ester hydrochloride, DIPEA, THF, RT overnight (94 %).

The benzylthiourea analogue **3.17** was prepared from the reaction of **3.8** and glycine ethyl ester, as shown in **Scheme 3.13**.

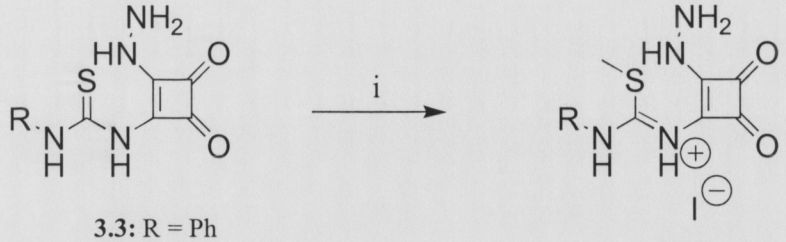


3.8 **3.17**

Scheme 3.13 *Reagents, conditions (and yield)*: (i) Glycine ethyl ester hydrochloride, DIPEA, THF, RT overnight (78 %).

3.4 Attempted Methylation of Thiourea Derivatives

The next step in the cyclisation process of these thiourea squaramides was to methylate the thiourea thiocarbonyl group. It was expected that some degree of difficulty would be encountered in methylating these thiourea squaramides, due to the richness in functionalities that were available for methylation on the hydrazino squaramides. However, it was hoped that reaction would occur preferentially at the thiocarbonyl sulphur atom.



3.3: R = Ph
3.5: R = Bu
3.7: R = Me
3.9: R = $PhCH_2$

3.18

Scheme 3.14 *Reagents and conditions*: (i) Iodomethane, solvent, overnight.

The general scheme of methylation initially used is outlined in **Scheme 3.14**. The methylated product **3.18** were found to be unstable, and invariably decomposed from a bright yellow precipitate to a red tar on filtration, with the accompanied smell of turnips, which could indicate the loss of methyl mercaptan. Different reaction conditions were used to elucidate whether solvent or time were factors in the decomposition. DCM, THF and CH_3CN permitted the initial formation of the yellow solid product. Nucleophilic solvents such as MeOH and

EtOH invariably decomposed the product *in situ*, and resulted in an indeterminate brown solid. The reaction time was not found to be a determining factor, as any solids salvaged from the reaction mixtures were starting materials.

In addition to the evidence provided by the crystal structure of hydrazone **3.10** in **Figure 3.2**, another explanation for the non-isolation of cyclised product was speculatively offered by molecular modelling (**Figure 3.3**). An energy minimisation experiment (Macromodel V5.0, AMBER forcefield, GB/SA solvation in chloroform) on **3.9** found that, in this case, π-stacking between the benzyl phenyl ring and the cyclobutenedione ring was observed in one of the lowest energy conformations. This interaction acted to point the hydrazine group away from the thiocarbonyl carbon, and so possibly prevented the formation of the required geometry for reaction. In which case, any other nucleophiles present, e.g. an alkoxide or hydroxide ion, may easily have reacted with the methylated thiocarbonyl to displace the thiol and caused further reactions to lead to decomposition. Steric factors were therefore assumed to play a large role in all these methylated-product decompositions, as the lability of the hydrazine group towards hydrazone formation in these squaramides indicated that cyclisation should occur easily if the hydrazino group could interact with the methylated thiocarbonyl carbon. While aromatic π-stacking between the squaramide and the benzyl group would not be observed in the case of the other thiourea derivatives, it at least proposes a reason of why one particular case failed.

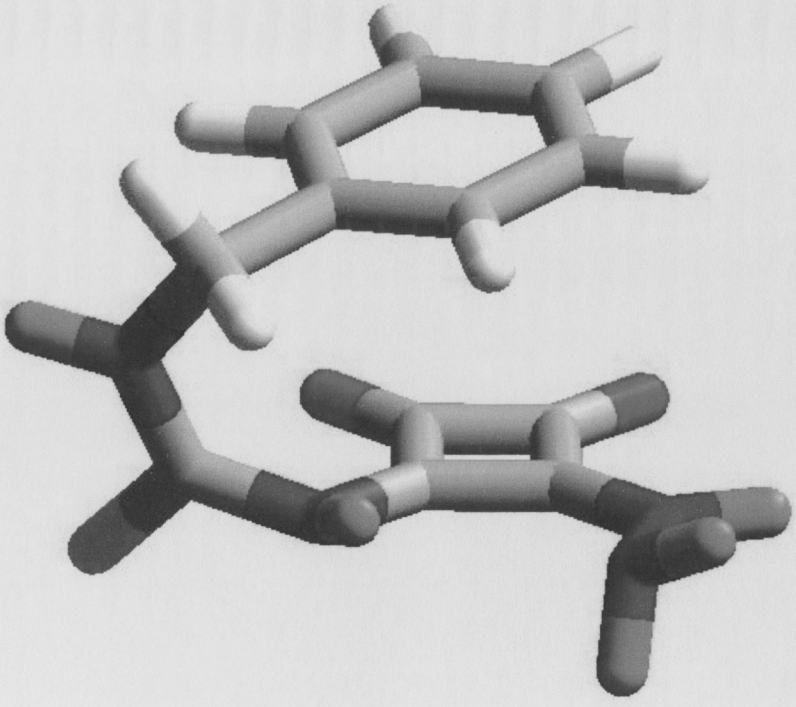
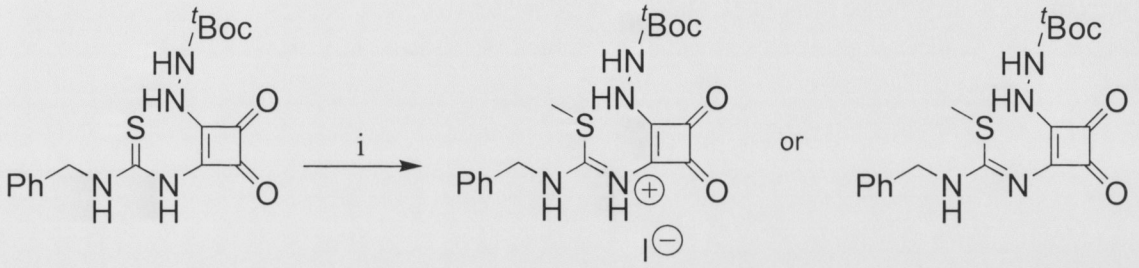


Figure 3.3 Ball and stick representation of the seventh lowest energy conformations of **3.9** obtained from molecular modelling.

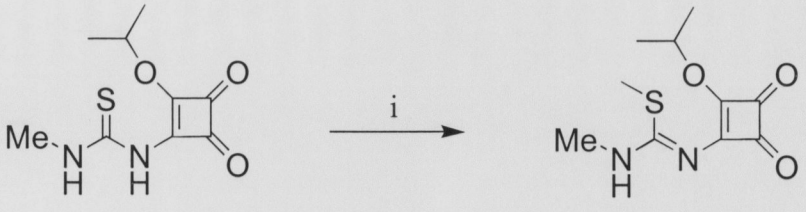
There was also a possibility that the nitrogen atoms on the hydrazine group were labile towards methylation. The ^{t}Boc-protected hydrazino derivative **3.13** was consequently used as a methylation target. Methylation was attempted with iodomethane, both with and without NaOMe as base, as detailed in **Scheme 3.15**. Invariably, a red intractable reaction mixture resulted, presumably as a consequence of the concomitant decomposition of the methylated products **3.19** or **3.20**.



3.13 **3.19** **3.20**

Scheme 3.15 *Reagents and conditions*: (i) Iodomethane, DCM, RT, overnight; or iodomethane, NaOMe, DCM, RT, overnight.

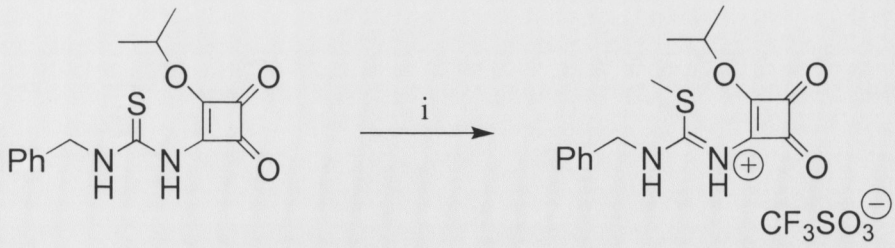
Methylation of the ester squaramides **3.6** and **3.8**, followed by reaction with hydrazine hydrate might provide the flexibility required for cyclisation. The attempts to methylate **3.6**, using CH_3I and NaOMe, are detailed in **Scheme 3.16**. An unstable yellow solid of hypothesised structure **3.21** resulted, and even the addition of hydrazine hydrate to the reaction mixture could not prevent its decomposition.



3.6 **3.21**

Scheme 3.16 *Reagents and conditions*: (i) Iodomethane, NaOMe, DCM, RT, overnight.

The use of methyl triflate as a methylating agent at –78 °C and 0 °C proved to be too vigorous when reacted with **3.8** and an indeterminate solid resulted, probably originating from methylated compound **3.22** (**Scheme 3.17**).



3.8 **3.22**

Scheme 3.17 *Reagents and conditions*: (i) Methyl triflate, DCM, -78 °C or 0 °C, 1 h.

The attempts to methylate the glycine ester derivatives **3.14-3.17** suffered from a similar fate. Methylations with CH_3I in the absence of base resulted in solutions containing inconclusive products. The use of base (NaOMe) to deprotonate the acidic CH_2 group invariably resulted in black tars.

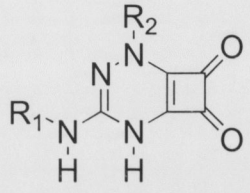
[3.1] E. Schaumann, *The Chemistry of Double Bonded Functional Groups,* **1987**, Wiley, 1270-1367.

Chapter Four

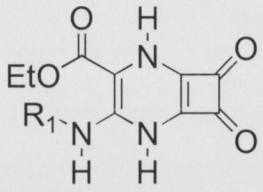
Alternative Routes to (6,4,6)-systems

4.1 Preamble

The inability to access alkylated thiourea derivatives **3.18-3.22** seemed to severely limit the possibility of securing the (6,4)-bicyclic intermediates **4.1** and **4.2** required. However, the ease with which the nucleophilic addition/elimination reactions occurred at the cyclobutenedione ring, as in the facile displacement of the isopropoxy group by the various amines and hydrazines discussed in Chapter 2, suggested that a similar nucleophilic addition with a specific bidentate ligand should form the desired bicycle in one step.

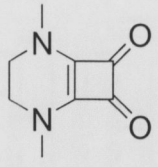


4.1

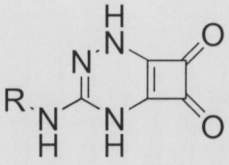


4.2

Hünig *et al.* reported the synthesis of bicyclic system **4.3** from **2.5** and *N,N'*-dimethylethylenediamine in refluxing ethanol.[4.1] It should be possible to obtain the desired bicyclic system **4.4** through the condensation of **2.5** with the aminoguanidine derivatives **4.5** and **4.6** in a similar procedure.

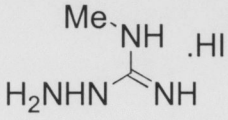


4.3

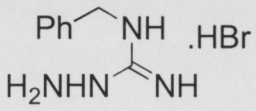


4.4

This hypothesis led to the synthesis of these aminoguanidine derivatives. The attempted reactions of these aminoguanidine derivatives **4.5** and **4.6** and various other nucleophilic substituents with diisopropyl squarate **2.5** are described in Section 4.2.

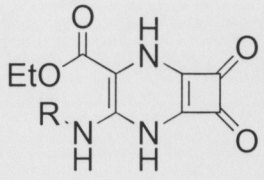


4.5

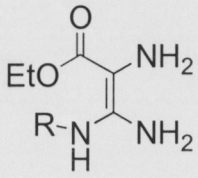


4.6

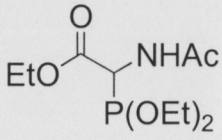
Similarly, it should be possible to access the desired bicycle **4.7** through the condensation of squarate ester **2.5** with the triaminoethene derivative **4.8**. It was proposed that **4.8** could be synthesised *via* a Wittig-type reaction from the phosphonate **4.9**, previously reported by Kober and Steglich,[4.2] and a substituted urea. Section 4.3.1 deals with the attempted synthesis of alkene **4.8**.



4.7



4.8



4.9

It was noted that the formation of the other six-membered ring of the desired (6,4,6)-tricyclic system would require a different approach due to the presence of the relatively inert ring carbonyl groups on that side of the cyclobutenedione ring. These carbonyl groups are known to behave as carboxylic acid carbonyl groups, which allowed the synthesis of the phenylhydrazino carbamoyl derivative **2.15** described in Chapter 2. As such, it became apparent that the transformation of these oxocarbons into their sulphur analogues would be necessary to achieve further reaction at these ring carbons.

Müller *et al.* published a procedure for the regioselective monothionation of squaric acid derivatives using Lawesson's reagent (2,4-bis(4-methoxyphenyl)-1,3,2,4-dithiaphosphetane 2,4-disulfide, **Figure 4.1**).[4.3] From NMR and X-ray data, they confirmed that thionation occurred preferentially on the non-adjacent carbonyl to form the even more polarised vinylogous thioamide or thioester unit, respectively. However, thionation of dialkyl squarates was not achieved at this point.

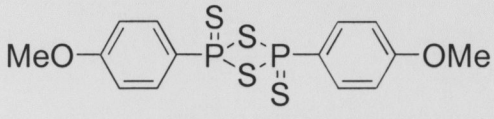
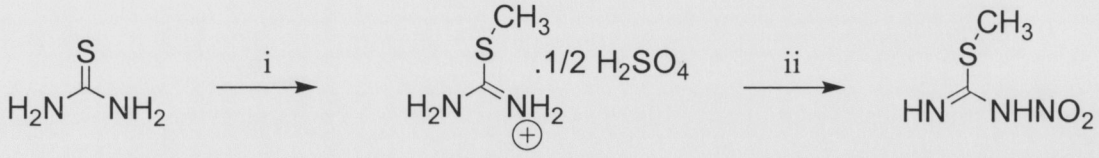


Figure 4.1 Lawesson's reagent.

Section 4.3.2 gives an account of the synthesis of the thiosquarate derivatives **4.28**, **4.30**, **4.31**, and **4.32**.

4.2 Synthesis of Aminoguanidine Derivatives

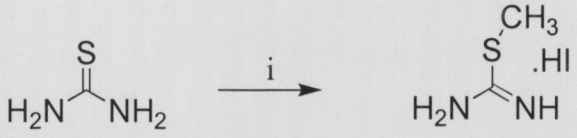
The successful syntheses of aminoguanidine derivatives **4.5** and **4.6**, and the attempted syntheses of aminoguanidine derivatives **4.15** and **4.16** are described. **Scheme 4.1** outlines the synthesis of 2-methyl-1-nitro-2-thiopseudourea **4.11**. Isothiourea sulphate **4.10** was prepared by the method of Shildneck and Windus.[4.4] **4.10** was nitrated by the procedure of Fishbein and Gallaghan[4.5] to produce nitroisothiourea **4.11**.



4.10 **4.11**

Scheme 4.1 *Reagents, conditions (and yields)*: (i) Dimethyl sulfate, H_2O, reflux 1 h (80 %); (ii) HNO_3, H_2SO_4, -10 °C (42 %).

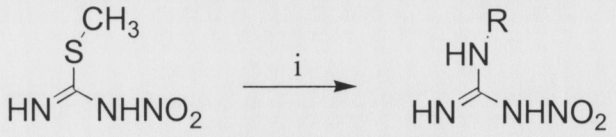
A similar nitration reaction was carried out on isothiourea hydroiodide **4.12**, formed from the reaction of thiourea with iodomethane in the method of Clapp *et al.*[4.6] (**Scheme 4.2**). However, sinister purple fumes emerged violently from the reaction vessel on the careful addition of **4.12** to the nitration mixture, and a black precipitate of intractable solid resulted.



4.12

Scheme 4.2 *Reagents, conditions (and yield)*: (i) Iodomethane, ethanol, H_2O, reflux, overnight (63 °%).

Butylnitroguanidine **4.13** was prepared by treatment of **4.11** with butylamine and **4.14** also prepared by treatment of **4.11** with benzylamine, in a modification of the procedure by Fishbein and Gallaghan,[4.5] with the characteristic displacement of methyl mercaptan (**Scheme 4.3**).

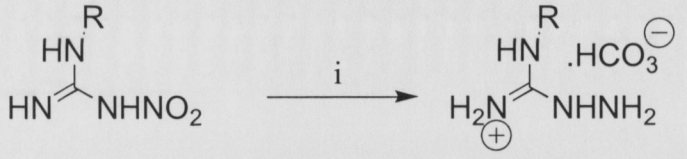


4.11

4.13: R = Bu
4.14: R = $PhCH_2$

Scheme 4.3 *Reagents, conditions (and yields)*: (i) RNH_2 (R = Bu, $PhCH_2$), overnight (81, 90 %, respectively).

The attempted synthesis of **4.15** and **4.16** from the reduction of **4.13** and **4.14**, respectively, with zinc metal and acetic acid is detailed in **Scheme 4.4**, which followed a modified procedure of Shriner and Neumann.[4.7] However, the isolation of aminoguanidine derivatives seems to depend on the counterion involved. Finnegan *et al.* discovered that the hydroiodides of various substituted aminoguanidines had to be converted to the nitrate or picrate to isolate a solid product.[4.8] It is likely that the bicarbonate counterion associated with this procedure may have contributed to the inability to isolate the benzylaminoguanidine derivative **4.16**, and also the extremely low yield of butylaminoguanidine **4.15** obtained (8 %).

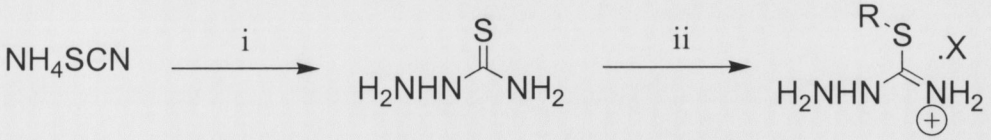


4.13: R = Bu
4.14: R = $PhCH_2$

4.15: R = Bu
4.16: R = $PhCH_2$

Scheme 4.4 *Reagents and conditions*: (i) Zn metal, AcOH, 5-15 °C, 8 h, then 40 °C, 40 min, $NaHCO_3$.

Another route to the desired aminoguanidine derivatives was offered in just two steps from thiosemicarbazide **4.17**, which was obtained by the method of Scott *et al.*[4.9] To investigate whether the counterion played a decisive role in the isolation of these aminoguanidine derivatives, **4.17** was methylated with iodomethane by the method of Emilsson[4.10] to yield **4.18**, and also **4.17** was methylated by ethyl bromide by the method of Jensen *et al.*[4.11] to yield **4.19** (**Scheme 4.5**).

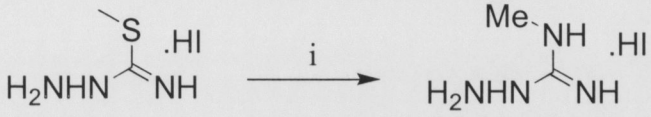


4.17

4.18: R = CH_3, X = I
4.19: R = C_2H_5, X = Br

Scheme 4.5 *Reagents, conditions (and yields)*: (i) Hydrazine hydrate, H_2O, reflux, 3 h (21 %); (ii) CH_3I or C_2H_5Br, ethanol, reflux, 2 h (75, 85 %, respectively).

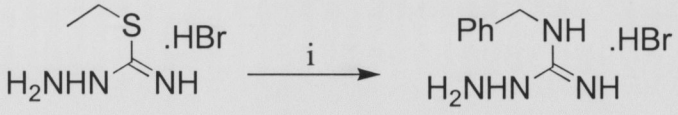
Methylaminoguanidine **4.5** was successfully synthesised as the hydroiodide by reaction of **4.18** with methylamine (**Scheme 4.6**) by the method of Kirsten and Smith.[4.12]



4.18 **4.5**

Scheme 4.6 *Reagents, conditions (and yield)*: (i) Methylamine, H_2O, reflux, 30 min (81 %).

Similarly, benzylaminoguanidine **4.6** was synthesised as the hydrobromide by the procedure of Heinisch[4.13] from benzylamine and **4.19**, as shown in **Scheme 4.7**. Another confirmation of the counterion theory was obtained in the attempted synthesis of benzylaminoguanidine as the iodide, as the desired product **4.6** was not obtained in this case.



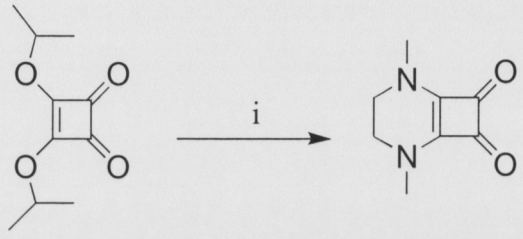
4.19 **4.14**

Scheme 4.7 *Reagents, conditions (and yield)*: (i) Benzylamine, H_2O, reflux, 30 min (62 %).

4.3 Attempted One-Pot Synthesis of Intermediate (6,4)-Bicyclic Systems

4.3.1 Attempted One-Pot Synthesis of Bicyclic Systems 4.4 and 4.7

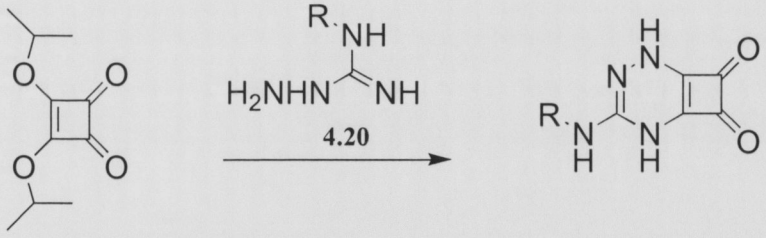
4.3 was synthesised from **2.4** and *N,N'*-dimethylethylenediamine in a modification of the method by Hünig *et al.*,[4.1] as shown in **Scheme 4.8**.



2.5 **4.3**

Scheme 4.8 *Reagents, conditions (and yield)*: (i) *N,N'*-Dimethylethylenediamine, ethanol, reflux, 4 h (25 %).

The facile synthesis of bicyclic **4.3** from two easily accessible starting materials inspired the development of system **4.4** from diisopropyl squarate **2.5** and an aminoguanidine derivative **4.20**, as in **Figure 4.2**.



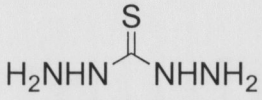
2.5 **4.4**

Figure 4.2 Proposed Scheme for formation of bicycle **4.4** from diisopropyl squarate **2.5** and aminoguanidine derivative **4.20**.

Refluxing **4.5** and **2.5** in acetic acid for one hour in a modified method of Heinisch[4.13] resulted in a deep red solution from which **2.5** was reclaimed. When the same reactants were refluxed in pyridine, **2.5** was reclaimed with a material that was baseline by TLC. A red tar resulted when the reagents were refluxed in aqueous NaOH, and H_2O. A stirred solution of **2.5** and **4.5** in concentrated H_2SO_4 resulted in an indeterminate red precipitate of negligible yield.

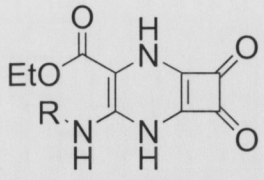
The nitrogen atoms on the aminoguanidine derivatives may arguably not be nucleophilic enough to displace the isopropoxy group from the cyclobutenedione ring. The question arose as to whether nucleophilic attack on **2.5** could be achieved by the precursor to the aminoguanidine derivative, with further methylation or hydrazination, for example, of the resultant product to form the desired bicyclic system.

Thiosemicarbazide **4.17** was refluxed with **2.5** in ethanol for four hours. The resultant solution was analysed by TLC, with the starting material **2.5** being the main spot observed. A white solid precipitated on cooling (presumably **4.17**). However, the reaction mixture turned red after 4 days. The methylated thiosemicarbazide **4.18** was refluxed in AcOH. A deep red solution resulted with mainly starting material **2.5** present by TLC. Thiocarbohydrazide **4.21** was refluxed with **2.5** in ethanol, and again only **2.5** and **4.21** were observed to be the main compounds present on cooling.



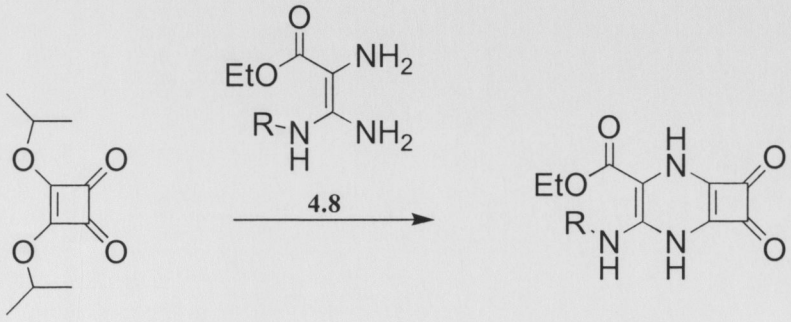
4.21

As it was proving unreasonably difficult to synthesise bicycle **4.4**, a similar methodology was envisaged to proceed more smoothly in the synthesis of bicycle **4.7**.



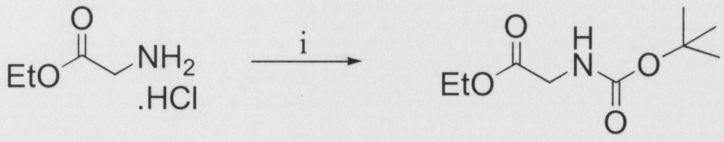
4.7

It was proposed to synthesise **4.7** from **2.5** and the novel triaminoethene derivative **4.8** in a one-pot reaction, as before.



2.5 **4.8** **4.7**

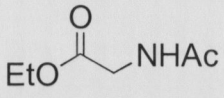
The speculative synthesis of the novel compound **4.8** was initially commenced with the ^{t}Boc-protected glycine ethyl ester **4.22**. **4.22** was synthesised from glycine ethyl ester in the method of Tarbell *et al.* (**Scheme 4.9**).[4.14] However, purification of the **4.22** product by distillation (104-105 °C/15 mm) was unattainable at that time, and so efforts were instead concentrated on producing the acetyl-protected glycine ethyl ester **4.23**.



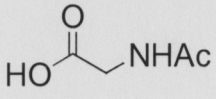
4.22

Scheme 4.9 *Reagents and conditions*: (i) $NaHCO_3$, NaCl, di-*tert*-butyl dicarbonate.

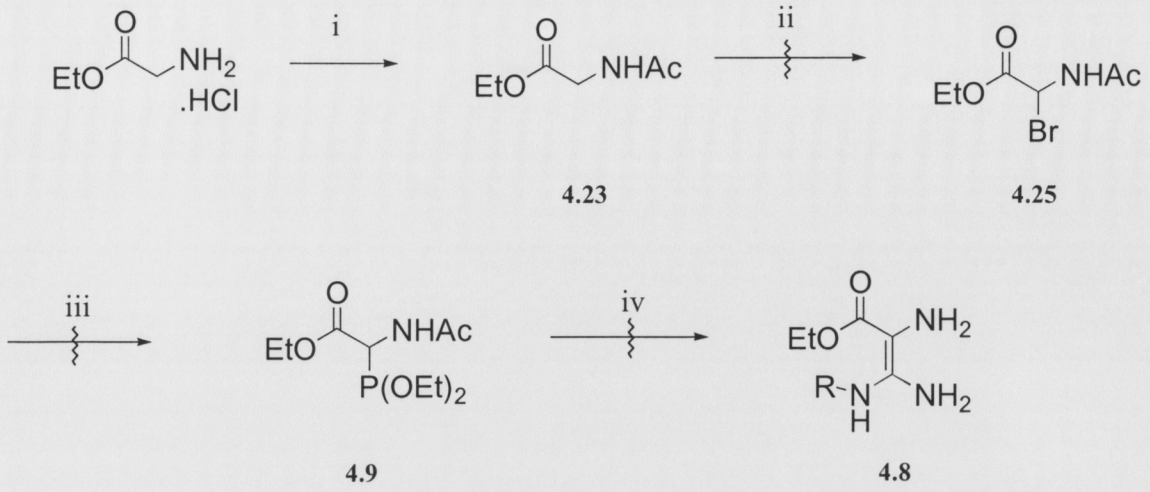
It was initially attempted to obtain **4.23** by the esterification of acetylglycine **4.24**, which was obtained from glycine by the method of Herbst and Shemin.[4.15] The esterification of **4.24** by a modification of the procedure of Hanby *et al.*[4.16] was not achieved even after several days. An alternative route to **4.23** from glycine ethyl ester, in the method of Nandi and Robinson,[4.17] was found to yield the desired compound **4.23** reliably, quickly and in good yield. The attempted synthesis of **4.8** from precursor **4.23** is outlined in **Scheme 4.10**.



4.23



4.24



Scheme 4.10 *Reagents and conditions*: (i) Ac_2O, NaOH, H_2O (55 %); (ii) Br_2, AIBN, HCOOMe, hν; (iii) $P(OEt)_3$, DCM.

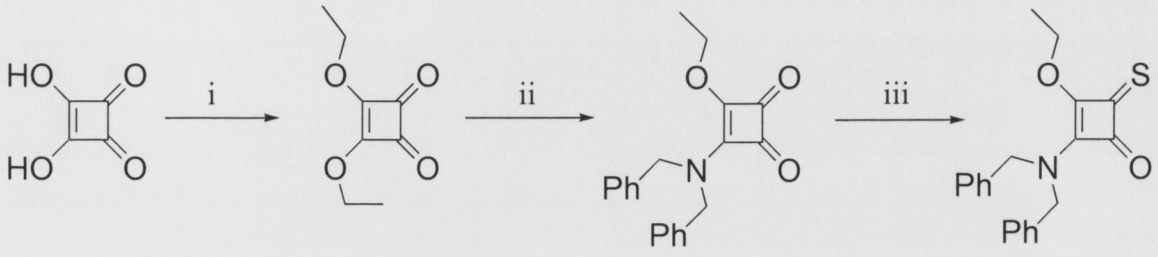
Kober and Steglich had reported a method for obtaining the phosphonate ester **4.9** *via* bromination of **4.23** and subsequent reaction of the brominated product **4.25** with triethyl phosphite yielded phosphonate **4.9**. They did not report any methods for purification of the brominated product. Instead it was used as crude in the subsequent reaction with triethyl phosphite. A repeat of the procedure of Kober and Steglich to brominate **4.23**, using methyl formate instead of the required carbon tetrachloride, yielded a red mixture, supposedly containing product **4.25** which was then reacted with triethyl phosphite. The resulting oil was suspended in ethyl acetate, and a white solid was obtained after one week. Unfortunately, analysis of this compound showed it to be starting material **4.23**.

4.3.2 Synthesis of Thionated Squarate Derivatives 4.28, 4.30, 4.31 and 4.32

Attention was then turned to the synthesis of the six-membered ring on the opposite side of the cyclobutenedione ring. This side contained two relatively unreactive ring carbonyls. It was decided that further reaction at these ring carbons would be difficult to achieve without first derivatising the carbonyl groups, perhaps by thionation to the thiocarbonyl, then methylation of the thiocarbonyl followed by nucleophilic attack by the required nucleophile to yield the desired result. It was also envisaged that the derivatisation of these ring carbonyls would eliminate the intramolecular hydrogen bond that had been hindering cyclisation on the thiourea side of the ring. Thionation should be achieved by the method of Müller *et al.*[4.3] from Lawesson's reagent.

Thiocyclobut-2-enones **4.28** and **4.30** were prepared as model systems from a modification of literature procedures to assess whether thionation was a viable method here.

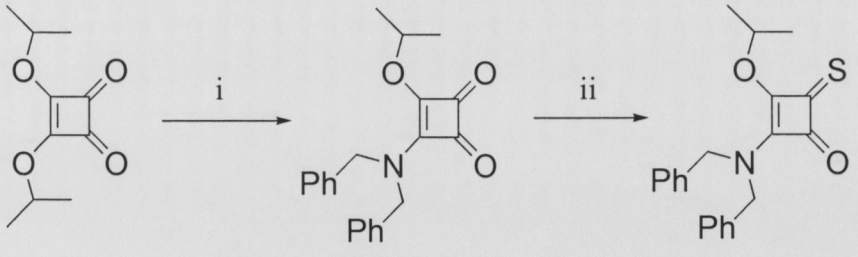
Scheme 4.11 shows the synthesis of **4.27** from diethyl squarate **4.26**, which was obtained from squaric acid **2.4** in the method of Ohno *et al.*.[4.18] Nucleophilic substitution on **4.26** with an equimolar quantity of dibenzylamine, as in the procedure of Thorpe,[4.19] produced the squaramide **4.27**. Thionation of **4.27** with half an equivalent of Lawesson's reagent formed the thiosquaramide **4.28**.



2.4 **4.26** **4.27** **4.28**

Scheme 4.11 *Reagents, conditions (and yields)*: (i) Ethanol, benzene, Dean and Stark apparatus, 15 h (44 %); (ii) Dibenzylamine, ethanol, RT, overnight (83 %); (iii) Lawesson's reagent (0.5 equiv.), DCM, RT, overnight (63 %).

Thiosquaramide **4.30** was similarly produced from diisopropyl squarate **2.5** in two steps, shown in **Scheme 4.12**. Reaction of **2.5** with dibenzylamine formed squaramide **4.29** in a modification of the procedure of Thorpe, which produced **4.30** on reaction with Lawesson's reagent.

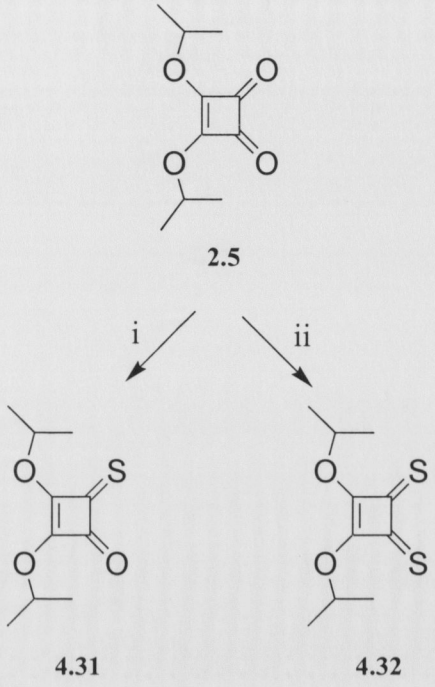


2.5 **4.29** **4.30**

Scheme 4.12 *Reagents, conditions (and yields)*: (i) Dibenzylamine, ethanol, RT, overnight (99 %); (iii) Lawesson's reagent (0.5 equiv.), DCM, RT, overnight (42 %).

Scheme 4.13 shows the successful thionation and dithionation of diisopropyl squarate **2.5,** using half an equivalent of Lawesson's reagent to form thiosquarate ester **4.31**, and one equivalent to form thiosquarate **4.32**, respectively. It was discovered that the basewashing process outlined in the original procedure of Müller (yet disputed earlier in the discussion part of

his paper) was detrimental to the integrity of the thiosquarate product. When the reaction mixtures were purified only by flash column chromatography after the considerable reaction times involved, reasonable yields were obtained of the thiosquarate **4.31** and the unstable dithiosquarate **4.32**.

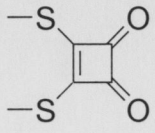


2.5

4.31 **4.32**

Scheme 4.13 *Reagents, conditions (and yields)*: (i) Lawesson's reagent (0.5 equiv.), DCM, RT, 5 d (62 %); (ii) Lawesson's reagent (1.0 equiv.), DCM, RT, 7 d (65 %).

Attempts to thionate the methylthiourea and benzylthiourea derivatives **3.6** and **3.8**, respectively, resulted in orange solids which could not be identified conclusively. Another obstacle in the formation of the desired six-membered ring was revealed in literature.[4.20] Apparently, the products obtained from the methylation of thiosquaramides are hygroscopic, and prefer to displace the associated amino functionalities rather than eliminate methyl mercaptan. In particular, dimethylated thiosquaramides hydrolyse to form the *S,S'*-dimethyl 1,2-dithiosquarate **4.33**.



4.33

The methylating agent quoted in the publication by West[4.20] was 'magic methyl' ($FSO_3CH_3 \bullet SbF_5$). The more commonly available methyl triflate ($CF_3SO_2CH_3$, which might not have been available for standard use for these pre-1980 methylations) may have successfully achieved methylation. It may also have been possible to effectively avoid hydrolysis by providing an efficient nucleophile (e.g. amine) to aid in further cyclisation. However, the overall lack of success achieved in this project deterred further investigation into these systems, and this effort was reluctantly left aside at this point.

[4.1] H. Ehrhardt, S. Hünig and H. Pütter, *Chem. Ber.* **1977**, *110*, 2506-2523.

[4.2] R. Kober and W. Steglich, *Liebigs Ann. Chem.* **1983**, 599-609.

[4.3] M. Müller, M. J. Heileman, H. W. Moore, E. Schaumann and G. Adiwidjaja, *Synthesis* **1997**, 50-52.

[4.4] P. R. Shildneck and W. Windus, *Org. Synth. Coll. 2.* 411-412.

[4.5] L. Fishbein and J. A. Gallaghan, *J. Am. Chem. Soc.* **1954**, *76*, 1877-1879.

[4.6] L. Long Jr., R. C. Clapp, F. H. Bissett and T. Hasselstrom, *J. Org. Chem.* **1961**, 85-88.

[4.7] R. L. Shriner and F. W. Neumann, *Org. Synth. Coll. 3.* 73-75.

[4.8] W. G. Finnegan, R. A. Henry and E. Lieber, *J. Am. Chem. Soc.* **1952**, 779-791.

[4.9] E. S. Scott, E. E. Zeller and L. F. Audrieth, *J. Org. Chem.* **1954**, 749-752.

[4.10] H. Emilsson, *J. Heterocyclic Chem.* **1989**, *26*, 1077-1081.

[4.11] K. A. Jensen, A. Holm and S. Rachlin, *Acta Chem. Scanda.* **1966**, *20*, 2795-2806.

[4.12] G. W. Kirsten and G. B. L. Smith. *J. Am. Chem. Soc.* **1936**, *58,* 800-801.

[4.13] L. Heinisch, *Chem. Ber.* **1967**, *100*, 893-904.

[4.14] D. S. Tarbell, Y. Yamamoto and B. M. Pope, *Proc. Natl. Acad. Sci. USA* **1972**, *69(3)*, 730-732.

[4.15] R. M. Herbst and D. Shemin, *Org. Synth. Coll. 1.* 11-12.

[4.16] W.E. Hanby, S. G. Waley and J. Watson, *J. Chem. Soc.* **1950**, 3239-3249.

[4.17] P. K. Nandi and D. R. Robinson, *J. Am. Chem. Soc.* **1972**, *94(4)*, 1299-1308.

[4.18] M. Ohno, Y. Yamamoto, Y. Shirasaki and S. Eguchi, *J. Chem. Soc. Perkin Trans. 1.* **1993**, *2*, 263-271.

[4.19] J. E. Thorpe, *J. Chem. Soc. (B)* **1968**, 435-436.

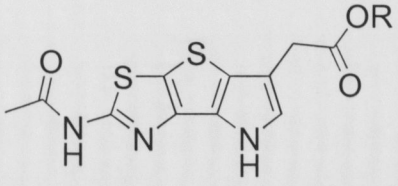
[4.20] R. West, *Oxocarbons*, **1980**, Academic Press, 15-41.

Chapter Five

The *N*-2,4-Dimethoxybenzyl Protected Pyrrolo[2',3':4,5]thieno [3,2-*d*]thiazole System

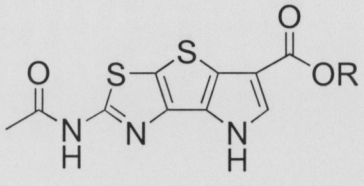
5.1 Preamble

As the (6,4,6)-tricyclic project was proving unproductive, it was considered to be more constructive at that point to pursue an alternative binding system. Success had recently been achieved in the synthesis of the A-T base pair recognising system 5-ethoxycarbonylmethylpyrrolo[2′,3′:4,5]thieno[3,2-*d*]thiazole system **1.53**.[1.72] Unfortunately, **1.53** had proved to be unstable when all of its protecting groups were removed, presumably due to the electron-rich nature of the π-system. This result was not unexpected. It had originally been intended to produce the 5-ethoxycarbonyl analogue **1.52**, which placed the electron-withdrawing ester group directly in conjugation with the ring system, and so would pull the electron density out of the electron-rich tricyclic system. However, the production of system **1.52** was found to be synthetically elusive, and efforts towards the production of tricycle **1.53** were concentrated on instead.



1.53

The instability of tricycle **1.53** confirmed the necessity for an electron-withdrawing group on the ring system to deter the oxidative processes that led to its decomposition. A renewed effort towards the synthesis of 5-ethoxycarbonylpyrrolo[2′,3′:4,5]thieno[3,2-*d*]thiazole **1.52** was then commenced.



1.52

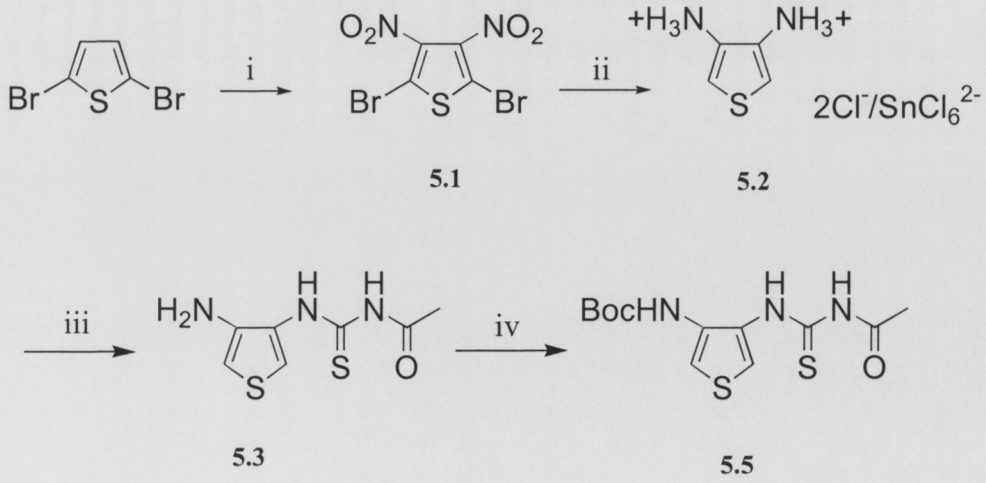
The synthetic routes followed in the efforts towards **1.52** are based closely on previous methods,[1.71] with any modifications made discussed in detail. It was anticipated that a full investigation of the unsuccessful syntheses towards **1.52**, as well as the involvement of new routes would yield the desired result. Section 5.2 deals with the synthesis of bicyclic intermediate 5-iodothieno[3,2-*d*]thiazole **5.8**.

Of the many routes described in literature for the synthesis of a pyrrole ring, it was thought that reaction of the amino group to form an enamine, and subsequent cyclisation *via* an intramolecular Heck reaction should be synthetically feasible. The successful attempts to access the enamine derivatives **5.12** and **5.17** are described in Section 5.3. However, the lack of reproducibility found in the synthesis of desired enamine **5.17** led to an intensive effort towards obtaining **5.17** from a variety of different reagents and reaction conditions, as detailed in Section 5.4.1.

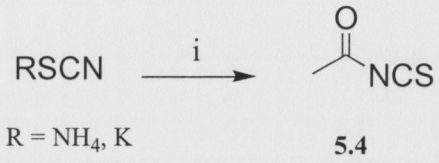
Section 5.4.2 gives an account of the attempted synthesis of 3,3,3-triethoxypropanal **5.22**, a preferred carbonyl component for enamine formation. The resynthesis of the 2,4-dimethoxybenzyl-protected version of the tricyclic system **1.53**, **5.26** for synthetic purposes is described in Section 5.5. It was felt that Pd catalysis could be used to aid the formation of the enamine, and simultaneously provide the optimum conditions for subsequent ring cyclisation. These efforts are recalled in Section 5.6.

5.2 Synthesis of the 2-Acetamidothieno[3,2-*d*]thiazole System 5.8

Scheme 5.1 exhibits the synthesis of the acetylthiourea derivative **5.5**. The dibromothiophene was bisnitrated following the procedure of Easton.[5.1] The dibromodinitro product **5.1** was reduced by the method of Outurquin and Paulmier.[5.2] Despite the difficulty in elucidating the exact constitution of the mixed hexachlorostannate chloride/ammonium salt **5.2**, it was possible to prepare the acetylthiourea derivative **5.3** from **5.2** and acetyl isothiocyanate **5.4**. **5.4** was synthesised from a modified procedure of Takamizawa *et al.*[5.3] (**Scheme 5.2**).



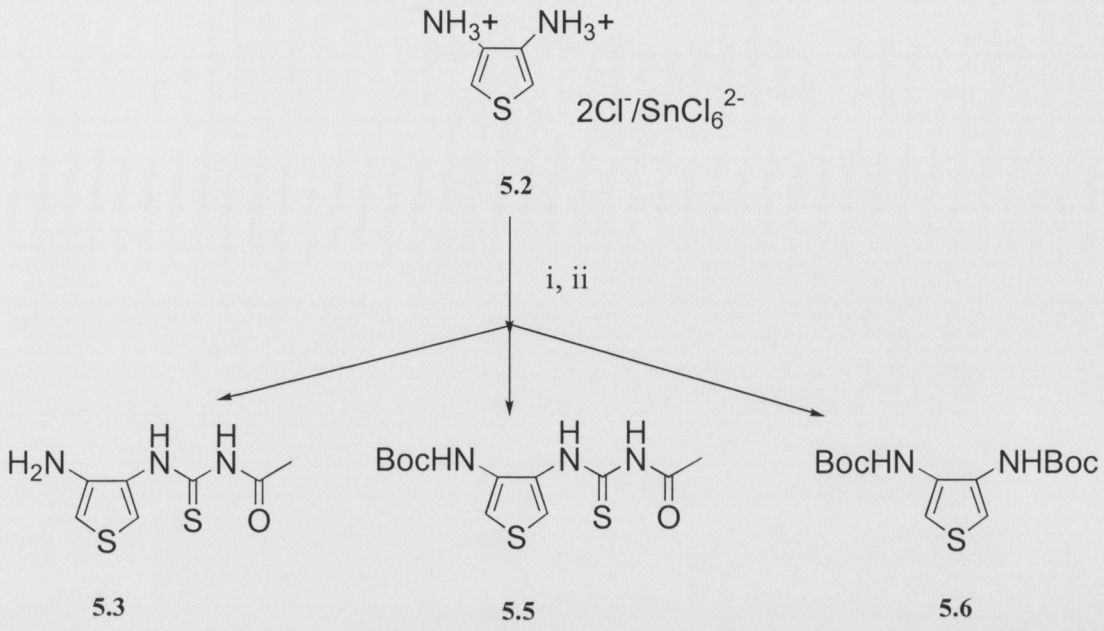
Scheme 5.1 *Reagents and conditions (and yields)*: (i) H_2SO_4 (conc.), fuming H_2SO_4 and fuming HNO_3, <30 °C (50 %); (ii) Sn granules, HCl (conc.), 4 h, <30 °C, then overnight at 4 °C; (iii) **5.4**, DCM, 3 d (40 % from **5.4**); (iv) Di-*tert*-butyl dicarbonate, DCM, 40 °C, 24 h (37 % from **5.4**).



Scheme 5.2 *Reagents and conditions (and yield)*: (i) Acetyl chloride, dry acetone, 20 min (65 %).

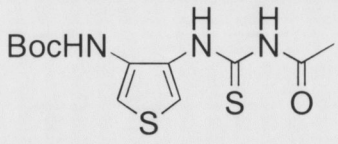
The chromatographic purification of **5.3** was found to simplify if the residue was dissolved in THF, and the insoluble tin solids removed by filtration. The procedure of Gavin Dunne, in which the residue was dissolved in a 4:1 DCM/EtOAc solution (50 mL) and the undissolved solids removed by filtration and washed with methanol, when repeated, was found to be ineffective, as the filtered solids dissolved in methanol to leave only the tin salts in any case. Unfortunately, the separation of the acetylthiourea **5.3** from a more polar white solid by flash column chromatography was mostly unachievable, and so the crude mixture was usually carried forward to the next step. As such, the acetylthiourea derivative **5.3** was usually reacted with di-*tert*-butyl dicarbonate to form **5.5**, typically in a one-pot synthesis from **5.2**.

However, the reactivity of **5.3** towards di-*tert*-butyl dicarbonate was found to be remarkably low. A second addition of di-*tert*-butyl dicarbonate was usually made after overnight stirring at RT, then refluxing for a further 15 hours usually yielded a comparatively low yield of **5.5**, with still a considerable quantity of unreacted **5.3** reclaimed. This reaction also yielded the di-^{t}Boc-protected diaminothiophene **5.6** as a by-product, produced from the reaction of unreacted diaminothiophene **5.2** and di-*tert*-butyl dicarbonate, shown in **Scheme 5.3**.

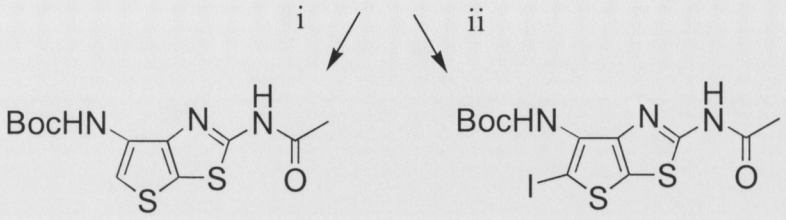


Scheme 5.3 *Reagents and conditions*: (i) **5.4**, DCM, 3 d; (ii) Di-*tert*-butyl dicarbonate, DCM, 40 °C, 24 h.

The oxidative cyclisation of **5.5** to form the acetylamino thienothiazole **5.7** and the iodothienothiazole **5.8** was performed with *N*-iodosuccinimide (NIS), as shown in **Scheme 5.4**. NIS was found to give the most efficient cyclisation results. The previously tried method of using bromine in $CHCl_3$ was found to yield three main products, including 6-aminothienothiazole **5.9**, and so the use of NIS was favoured in the present study. Two equivalents of NIS were required for cyclisation to form thienothiazole **5.7**. It was anticipated that an intramolecular Heck-type reaction would be carried out on the bicycle to form the pyrrole ring, and so an extra equivalent of NIS was used to iodinate the 5-position of thienothiazole **5.7** to produce 5-iodothienothiazole **5.8**.

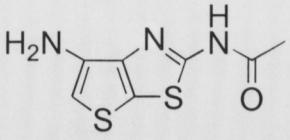


5.5



5.7 **5.8**

Scheme 5.4 *Reagents, conditions (and yields)*: (i) *N*-Iodosuccinimide (2 equiv.), $CHCl_3$, 4 h (53 %); (ii) *N*-Iodosuccinimide (3 equiv.), $CHCl_3$, 4 h (81 %).



5.9

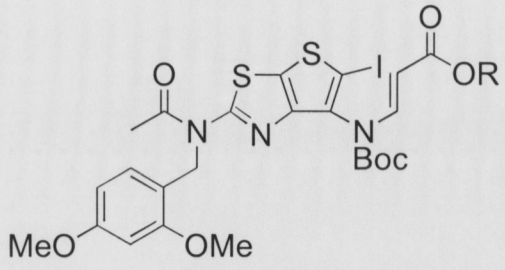
The previous purification procedure in the synthesis of **5.8** involved additions and subsequent evaporations of chlorinated solvent to the residue to remove the red iodine impurities. This method was time consuming and generally inefficient, with the added inconvenience of coating the inner glassware of the rotary evaporator with the volatile iodine impurities. It was also usually necessary to eliminate any residual iodinated impurities chromatographically with DCM until the fractions and the silica gel were no longer pink, after which time the desired compounds could be obtained. The entire purification process alone invariably took longer than eight hours.

It was found that a sodium thiosulfate wash did not adversely affect the iodinated product, and so simplified the purification process by removing any inorganic impurities firsthand. This simple procedure reduced the total purification time required by at least an hour, when all factors are taken into consideration.

It was also noted that the thienothiazole and iodothienothiazole products **5.7** and **5.8** were only moderately soluble in $CHCl_3$. A relatively large volume of $CHCl_3$ was therefore needed to ensure dissolution of the products, as small volumes often resulted in the formation of an insoluble dark purple precipitate and a correspondingly low yield of products. Another factor that influenced product yields was that the *N*-iodosuccinimide reagent was found to lose its efficiency over time, requiring recrystallisation (dioxane/CCl_4) after a couple of months stored at 4 °C.

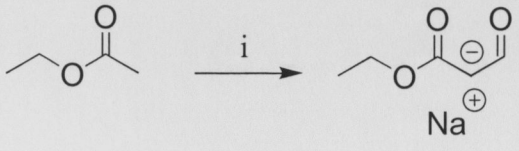
5.3 Synthesis of Enamine Derivatives 5.12 and 5.17

As in the synthesis of analogue system **1.53**,[1.71] the completion of the thienothiazole framework required the formation of the pyrrole ring to obtain the desired ethoxycarbonylpyrrolo[2′,3′:4,5]thieno[3,2-*d*]thiazole ring system. The approach taken was to obtain the enamine **5.10** by a variety of tried and new methods, then to cyclise the ring by a Pd-catalysed intramolecular Heck reaction.



5.10

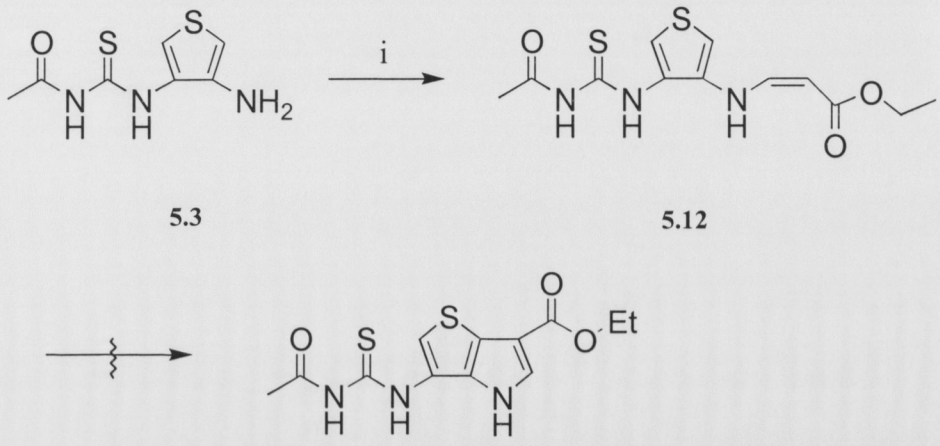
The commonest procedure for enamine formation is the reaction of a carbonyl compound with an amine.[5.4] The obvious choice of carbonyl compound to yield enamine **5.10** is the formylacetate **5.11** prepared from ethyl acetate by the method of Robins *et al.*[5.5] **Scheme 5.5** outlines the synthesis of **5.11**, which was handled as the sodium salt due to the notorious instability of these formylacetate compounds.



5.11

Scheme 5.5 *Reagents, conditions (and yields)*: (i) Ethyl formate, sodium sand†, toluene, RT, 2 d. †Sodium metal, toluene, vigorous stirring with reflux, then 0 °C with stirring.

Reaction of **5.11** (converted to the free formylester *in situ* by the addition of AcOH) with the thienothiazole **5.7** did not occur, with starting material **5.7** reclaimed from an otherwise polymeric tar. However, success was achieved, albeit in low yields, from the reaction of **5.11** with the 3-(4-aminothienyl)thiourea **5.3** under similar reaction conditions to form enamine **5.12**, as detailed in **Scheme 5.6**. However, this reaction suffered from poor reproducibility. A previous attempt to cyclise **5.12** with bromine as the oxidative agent to form the thienopyrrole **5.13** had resulted in an intractable reaction mixture. Unfortunately, a repeat of this reaction with NIS was not possible.



Scheme 5.6 *Reagents, conditions (and yield)*: (i) **5.11**, AcOH, dry ethanol, reflux, 30 min, then RT overnight (40 %).

As before, it was realised that the thienothiazole ring system must be produced before the enamine could be formed. Another route to the enamine was then considered. *N*-vinylation of the amino functionality should be possible through nucleophilic substitution with the α,β-unsaturated ester, ethyl *cis*-3-iodoacrylate.

It had been previously found that alkylation of the thienothiazole **5.7** occurred preferentially at the acetamido nitrogen to form the enamine compound **5.14** as the major product, while the desired enamine **5.15**, formed *via* alkylation of the 6-amino group, was simultaneously produced as the minor product,[1.71] as shown in **Figure 5.1**.

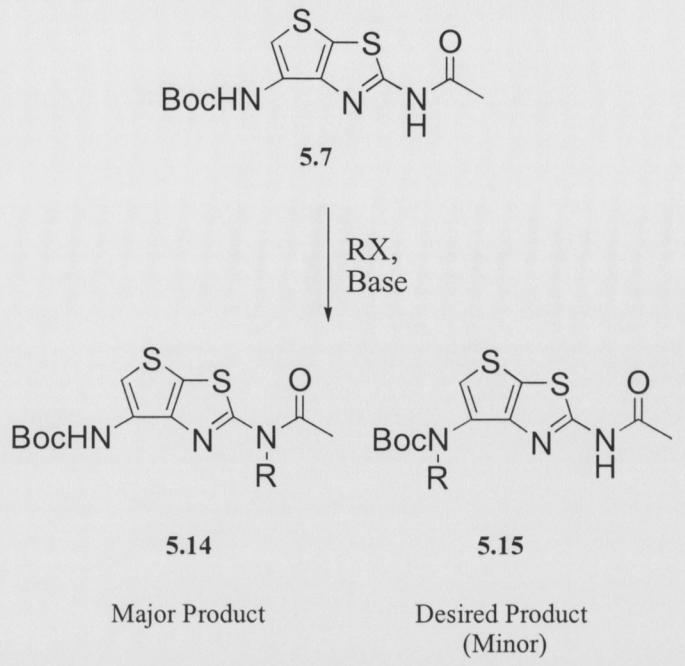
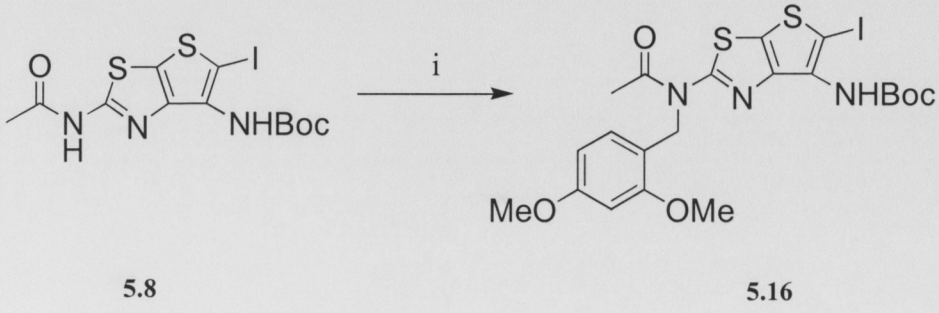


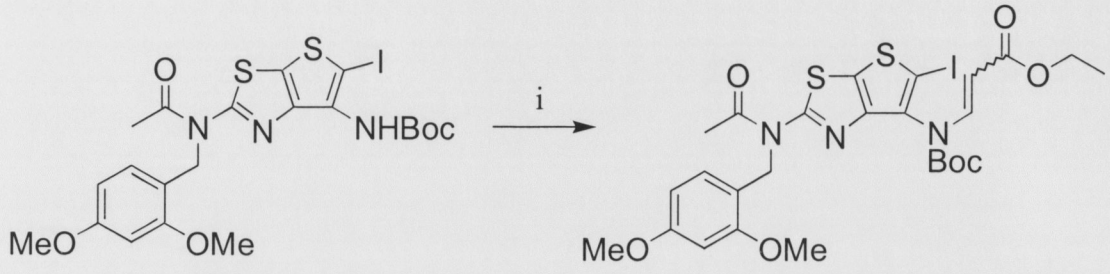
Figure 5.1 Alkylation products of **5.7**

It was deemed necessary to block the acetamido nitrogen with the use of a protecting group. The 2,4-dimethoxybenzyl group had been found to be the protecting group of choice, as it could verify the regiospecificity of the protection using difference NOE experiments, as well as being relatively easy to remove. A considerable effort had been previously spent on the investigation of many benzyl derivatives as possible protecting groups,[1.72] but removal of these groups could not be achieved. **Scheme 5.7** shows the addition of the protecting group to iodothienothiazole **5.8** by means of a low-yielding Mitsunobu reaction. Attempts to form the 2,4-dimethoxybenzyl-protected product **5.16** *via* the corresponding mesylate[5.6] from 2,4-dimethoxybenzyl alcohol resulted in insoluble product (presumably polymerisation product) and untransformed starting material **5.8**.



Scheme 5.7 *Reagents, conditions (and yield)*: (i) 2,4-Dimethoxybenzyl alcohol, DIAD, PPh_3, THF (36 %).

The reaction of ethyl iodoacrylate with protected thienothiazole **5.16** had previously been unsuccessful with starting material being reclaimed. **Scheme 5.8** outlines the successful synthesis of the 6-enamino derivative **5.17** from the 2,4-dimethoxybenzyl-protected **5.16**, using the relatively vigorous conditions of sodium hydride as base, and DMF as solvent. Both NMR and LC-MS data show the presence of both the *cis* and *trans* isomers of **5.17**.



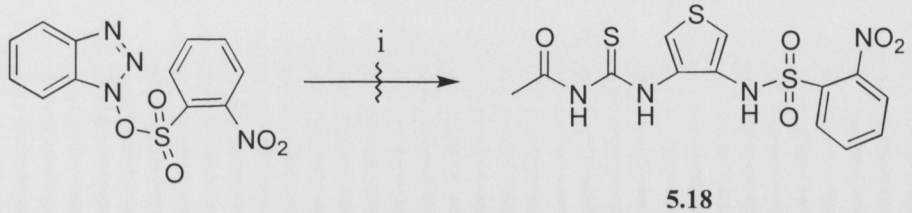
5.16 **5.17**

Scheme 5.8 *Reagents, conditions (and yield)*: (i) Ethyl iodoacrylate, NaH, DMF, RT, overnight (21%).

Unfortunately, subsequent trials of this reaction have never met with success. The R_f of the desired product **5.17** is higher than the starting material **5.16**, and TLC analyses of many crude reaction mixtures indicated its presence. However, the subsequent purification process generally resulted in reclaiming starting material **5.16**. A typical purification procedure involved evaporation of the DMF solvent under reduced pressure at 80 °C, repeated addition and subsequent evaporation of toluene (1 mL) to remove residual DMF (the toluene emulsions resembled wallpaper paste), then flash column chromatography was carried out on the crude residue, using a 0-2 % EtOAc/DCM eluent gradient (DCM was also the solvent used to load the residue onto the column).

The organic-insoluble sodium hydride-containing impurities were removed as part of the chromatographic process, by remaining on top of the silica gel. Efforts to remove these solid impurities (and also the residual DMF) by aqueous work-up failed, as an inseparable light brown emulsion formed on addition of water to the DCM suspension of the crude mixture. It may have been possible that the conditions of solvent evaporation (80 °C, 20 mm Hg) might have beneficially influenced the reaction. However, repeating the reaction at 80 °C and atmospheric pressure did not affect the starting material, nor did varying the rotary evaporation conditions.

Wensbo *et al.*[5.7] reported that the use of strong bases such as sodium hydride or potassium *tert*-butoxide, in analogous *N*-allylation reactions of *o*-halothienylamines, led to the rapid decomposition of the allylating agent ethyl bromocrotonate. It is not unreasonable to suppose that a similar decomposition occurs with iodoacrylate. In an effort to eliminate the necessity for such aggressive deprotonation conditions, an attempt was made, as seen in **Scheme 5.9**, to synthesise *o*-nitrobenzylsulphonyl-protected **5.18** from **5.3** and 1-(2-nitrobenzenesulphonyloxy)-benzotriazole, which was synthesised by Laura Siracusa.[5.8] The *o*-nitrobenzenesulphonyl group is more electron-withdrawing than the *tert*-butoxycarbonyl group, and so should facilitate the formation of the anion. Unfortunately, this reaction resulted in the formation of a vivid green solution that ultimately formed a black tar.



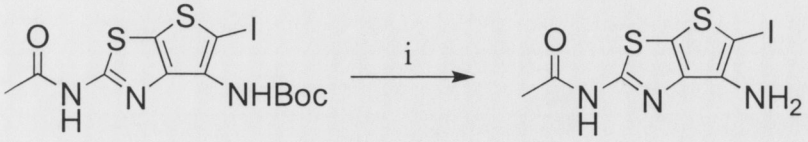
5.18

Scheme 5.9 *Reagents and conditions*: (i) **5.3**, dry THF, 1 h.

5.4 Attempted Synthesis of Enamine Derivative 5.17

5.4.1 Attempted *N*-Vinylation of Thienothiazole 5.16

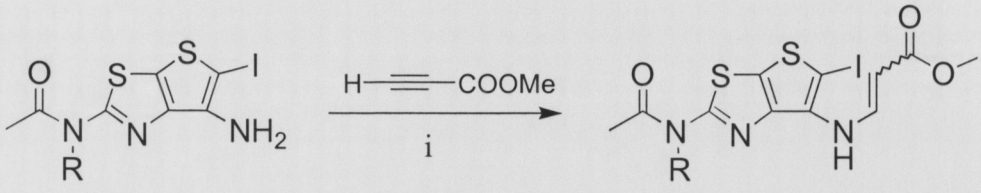
The general lack of success in enamine production again prompted exploration into other avenues to enamine formation. The reaction of primary amines and alkynes to yield enamines is well documented. This route would be particularly desirable as it would eliminate the necessity of the low-yielding Mitsunobu reaction. **Scheme 5.10** outlines the deprotection of ^{t}Boc-protected iodothienothiazole **5.8** by the standard procedure (TFA) to yield the 5-aminothienothiazole **5.19**. **5.19** was then reacted with the activated alkyne methyl propiolate by a variety of procedures.



5.8 **5.19**

Scheme 5.10 *Reagents and conditions*: (i) TFA, DCM, RT, 1 h.

Grigg and Savic[5.9] reported the reaction of aliphatic amines with methyl propiolate in DMF at 85 °C to form the desired enamine. **Scheme 5.11** outlines a general procedure for the synthesis of a desired product **5.21** from a general starting material **5.20** and methyl propiolate. However, when this procedure was applied to aminothienothiazole **5.19**, it failed to affect the starting material. The use of DIPEA as base in CH_3CN and THF left **5.19** similarly untouched. The use of sodium hydride as base, both at RT and 100 °C, resulted in polymeric tars. A possible explanation may be that the electron-withdrawing ester group of methyl propiolate facilitates the loss of the acetylenic proton, which leads to carbanions which complicate the reaction.[5.10] This was seen in reactions where various excesses of methyl propiolate were used (from two to ten equivalents), all of which resulted in complex reaction mixtures, tars and indeterminate precipitates. However, a tar was obtained even with one equivalent of methyl propiolate.



5.20 **5.21**

Scheme 5.11 *Reagents and conditions*: (i) solvent, RT/100 °C/reflux, base/no base, overnight.

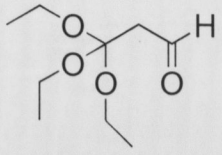
A tar was obtained also when the deprotection by-products in the formation of **5.19** were first removed by an aqueous base wash, before reaction with the alkyne. One of the contributing factors to the failure of these reactions was possibly the instability of the 6-aminothienothiazole product **5.19**. It was found to decompose from a white solid to a brown tar in a few days, and it is very possible that the reaction conditions of these propiolate additions may have encouraged decomposition.

The reaction of anilines with methyl propiolate in methanol at RT was published by Heinkel *et al.*[5.11] Previous work concluded that neither the reaction conditions of **5.19** with ethyl propiolate in RT nor refluxing methanol resulted in product,[1.71] and so these conditions were not repeated with methyl propiolate in this study.

Hénin *et al.* reported the formation of an enamine with a carbolinic amino group and methyl propiolate in refluxing AcOH.[5.12] The reaction of ^{t}Boc-protected **5.8** with methyl propiolate in refluxing AcOH resulted in a polymeric tar.

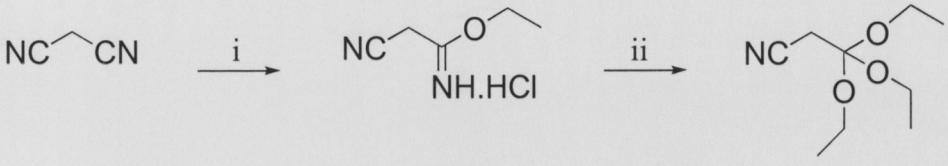
5.4.2 Attempted Synthesis of (3,3,3)-Triethoxypropanal 5.22

It had been appreciated that the instability of the formyl acetate **5.11** may have contributed considerably to the unsatisfactory success record achieved in the formation of enamine **5.12**. Perhaps a more convenient derivative of the formylacetate ester **5.11** is provided by the orthoester **5.22**.



5.22

The orthoester functionality acts as an ester protecting group, which can easily be removed in the presence of acid when required. It is not possible to access the orthoester directly from the ester, but instead *via* a Pinner synthesis, as detailed in **Scheme 5.12**.



5.23 **5.24**

Scheme 5.12 *Reagents, conditions (and yields)*: (i) $HCl_{(g)}$, diethyl ether, dry ethanol, 0 °C, 24 h (68 %); (ii) Dry ethanol, RT, 3 d (56 %).

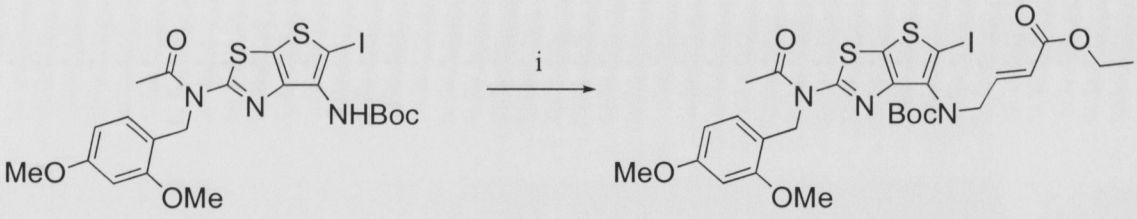
Malononitrile was treated with anhydrous HCl and EtOH by a modified method of McElvain and Schroeder[5.13] to form the imidic ester hydrochloride **5.23**. Alcoholysis of **5.23** *via* the Pinner reaction from a modification of the procedure of deWolfe[5.14] resulted in the

orthocyanoacetate **5.24**. The reduction of the cyano group to the aldehyde with DIBAL-H at –78 °C, and the subsequent hydrolysis of the intermediate imine with NaOH and H_2O, by the method of Dreyer and Metcalf,[5.15] (and the purification technique of vacuum distillation involved) have never resulted in the desired compound. It appeared that the hydrolysis of the intermediate imine was not trivial and the protecting group was removed during the course of the reaction.

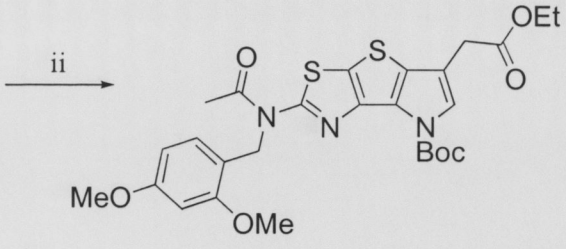
5.5 Synthesis of 2,4-Dimethoxybenzyl-protected 5-Ethoxycarbonylmethylpyrrolo[2′,3′:4,5]thieno[3,2-*d*]thiazole 5.26

It was considered constructive at this stage to resynthesise the 2,4-dimethoxybenzyl-protected tricyclic system **5.26**, as previously synthesised,[1.72] to provide synthetic comparisons.

There has been some literature precedence for the production of 5-ethoxycarbonylmethylthieno[3,2-*d*]pyrrole systems *via* the alkylation of *N*-^{t}Boc-protected *o*-iodothienyl amines. Wensbo *et al.* achieved allylation of these amines with ethyl 4-bromocrotonate, with subsequent Pd-catalysed ring cyclisation to form *N*-^{t}Boc-protected thienopyrroles in a one-pot synthesis.[5.7]



5.16 **5.25**



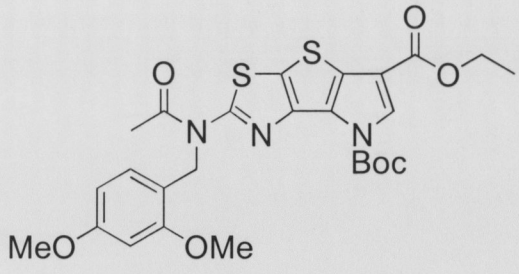
5.26

Scheme 5.13 *Reagents, conditions (and yields)*: (i) Ethyl 4-bromocrotonate, NaH, DMF, RT, 2 h (47 %); (ii) $Pd(OAc)_2$, PPh_3, DIPEA, THF, 65 °C, 2 h (75 %).

Scheme 5.13 shows the synthesis of the 2,4-dimethoxybenzyl protected 5-ethoxycarbonylmethylpyrrolo[2′,3′:4,5]thieno[3,2-*d*]thiazole **5.26**. The enamine intermediate **5.25** was produced from **5.16** by a modification of the method of Wensbo *et al.*[5.7] using sodium hydride (instead of K_2CO_3) as base and DMF as solvent, and resulted in faster reaction times (usually in two hours instead of two days). The decomposition of ethyl bromocrotonate in the presence of sodium hydride was rarely observed. Palladium catalysed cyclisation formed the desired tricycle **5.26**, using the modified conditions of $Pd(OAc)_2$, PPh_3 and DIPEA in CH_3CN at 65 °C.

5.6 Attempted One-Pot Pd-Catalysed Synthesis of 2,4-Dimethoxybenzyl-protected 5-Ethoxycarbonylpyrrolo [2′,3′:4,5] thieno[3,2-*d*]thiazole 5.27

Palladium catalysis offers a convenient means of adding a multiple bonded functionality to an unsaturated system. The 2,4-dimethoxybenzyl protected tricycle **5.18** was successfully cyclised with the aid of $Pd(OAc)_2$, PPh_3 and DIPEA. It might be possible therefore, to add a double bonded moiety to the amino group on the thienyl ring to form an enamine using Pd catalysis. Wensbo *et al.* obtained thienopyrroles from a one-pot synthesis by subsequent addition of $Pd(OAc)_2$, PPh_3 and DIPEA to the allylation reaction mixture.[5.7] It seems reasonable to propose that the Pd catalysed reaction conditions for enamine formation could be manipulated to coincide with the Heck cyclisation conditions required to produce the desired tricyclic system **5.27**.



5.27

The single success achieved with enamine formation from ethyl iodoacrylate made it the prime candidate for initial trials. The 2,4-dimethoxybenzyl thienothiazole **5.16** was reacted with iodoacrylate and DIPEA in a THF solution containing catalytic quantities of $Pd(OAc)_2$ and PPh_3, both at RT and 60 °C. In both cases, the reaction mixtures darkened within a few minutes, and a polymeric tar resulted, with excess iodoacrylate being the only compound salvaged from the mixture.

The use of KO^{t}Bu as a base using similar conditions at RT resulted in a tar plus unreacted starting materials. The use of 2-(dicyclohexylphosphino)-biphenyl as catalyst at 60 °C also yielded a tar.

Pd/Cu catalysis has been used for terminal alkyne coupling to aryl and vinyl halides in the Sonogashira reaction.[5.16] Paulmier *et al.* observed a similar addition of the triple bond directly to the thienyl ring of halothiophenes in Pd-catalysed alkyne coupling reactions.[5.17] It was with some trepidation, therefore, that the Pd-catalysed reaction of methyl propiolate and iodothienothiazole **5.8** was approached. Unfortunately, reaction in DMF with DIPEA, $Pd(OAc)_2$ and PPh_3, both at RT and 95 °C, resulted in complex reaction mixtures. The Pd-catalysed reaction of 2,4-dimethoxybenzyl protected thienothiazole **5.16** and methyl propiolate in THF at 60 °C brought a similar conclusion.

Reisch and Dittmann had formed an enamine from *N*-vinylation on acridinones with methyl propiolate, assisted by $PdCl_2(PPh_3)_2$.[5.18] The reflux of **5.16** and methyl propiolate in Et_3N as base/solvent and $Pd(OAc)_2$ resulted in a dark red solution from which starting material was reclaimed.

[5.1] R. Mozingo, S. A. Harris, D. E. Wolf, C. E. Hoffhine Jr., N. R. Easton and K. Folkers, *J. Am. Chem. Soc.* **1945**, *67*, 2092-2095.

[5.2] F. Outurquin and C. Paulmier, *Bull. Soc. Chim. Fr.* **1983**, *5-6 (II)*, 153-158.

[5.3] A. Takamizara, K. Hirai and K. Matsui, *Bull. Chem. Soc. Jpn.* **1963**, *36*, 1214-1220.

[5.4] S. F. Dyke, *The Chemistry of Enamines*, **1973**, Cambridge University Press, 1.

[5.5] R. K. Robins, G. R. Revankar, D. E. O'Brien, R. H. Springer, T. Novinson, A. Albert, K. Senga, J. P. Miller and D. G. Streeter, *J. Heterocyclic Chem.* **1985**, *22*, 601-634.

[5.6] P. Nussbaumer, K. Baumann, T. Dechat and M. Harasek, *Tetrahedron* **1991**, *47(26)*, 4591-4602.

[5.7] D. Wensbo, U. Annby and S. Gronowitz, *Tetrahedron* **1995**, *51(37)*, 10323-10342.

[5.8] L. Siracusa, Postdoctoral Report, University of Bristol, **2001**.

[5.9] R. Grigg and V. Savic, *Chem. Commun.* **2000**, 873-874.

[5.10] P. Sanna, P. A. Sequi, S. Piras and G. Pagletti, *Heterocycles* **1995**, *41(11)*, 2459-2474.

[5.11] N. D. Heinkel, P. D. Kennewell and V. B. Fish, *J. Heterocyclic Chem.* **1969**, *6*, 77-81.

[5.12] J. Hénin, J. Vercauteren, C. Mangenot, B. Hénin, J. M. Nuzillard and J. Guilhem, *Tetrahedron* **1999**, *55*, 9817-9828.

[5.13] S. M. McElvain and J. P. Schroeder, *J. Am. Chem. Soc.* **1949**, *71*, 40-46.

[5.14] R. W. deWolfe, *Synthesis* **1974**, 153-172.

[5.15] G. B. Dreyer and B. W. Metcalf, *Tetrahedron* **1988**, *29(52)*, 6885-6888.

[5.16] K. Sonogashira, Y. Tohda, and N. Hagihara, *Tetrahedron. Lett.* **1975**, 4467-4470.

[5.17] D. Brugier, F. Outurquin and C. Paulmier, *Tetrahedron* **2000**, *56*, 2985-2993.

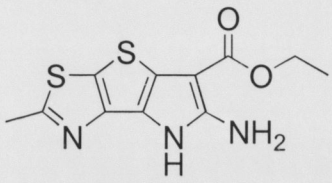
[5.18] J. Reisch and S. Dittmann, *J. Heterocyclic Chem.* **1992**, *29*, 1857-1858.

Chapter Six

Alternative Thieno[3,2-*b*]pyrrole Systems

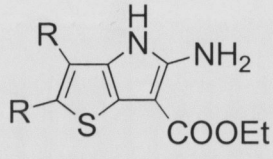
6.1 Preamble

On the commencement of the (5,5,5)-tricyclic project to synthesise the A-T base pair binding system **1.52**, it had also intended to produce an analogous system that would recognise C-G/G-C base pairs through its ADD/DDA H-bonding pattern, and so provide a comprehensive DNA recognition system. It had been envisaged that the required H-bonding pattern could be achieved from a variant of system **1.52**, by replacement of the acetyl group by a methyl group, and the placement of a H-bond donor amino group at position-6 on the tricyclic system, to produce the C-G recognising 6-amino-5-ethoxycarbonyl-2-methylpyrrolo[2′,3′:4,5]thieno[3,2-*d*]thiazole **6.1**.



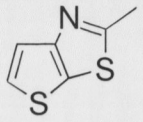
6.1

Two approaches to the synthesis of **6.1** were taken. Initially, it had been planned to obtain **6.1** *via* the aminothieno[3,2-*b*]pyrrole intermediate **6.2**. Attempts towards the production of intermediate **6.2** are detailed in Section 6.2.1.

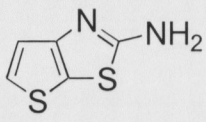


6.2

It was then realised that **6.1** could be obtained *via* the methylthieno[3,2-*d*]thiazole intermediate **6.12**. The procedures followed towards the synthesis of **6.1** were modifications of the methods of Paulmier.[6.1, 6.2] Paulmier also published a procedure for obtaining the aminothienothiazole **6.13**. It was envisaged that if a reliable and convenient method for accessing **6.13** could be achieved, then this procedure could possibly in part replace the laborious and low-yielding synthesis of the acetamidothieno[3,2-*d*]thiazole intermediate **5.8**. The attempts made towards the completion of the thienothiazole intermediates are discussed in detail in Section 6.2.2.



6.12



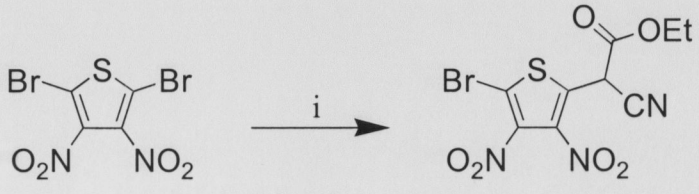
6.13

6.2 Attempts towards the Synthesis of 6-Amino-5-ethoxycarbonyl-2-methylpyrrolo[2′,3′:4,5]thieno[3,2-*d*]thiazole 6.1

6.2.1 Attempted Synthesis of Substituted 3,4-Dinitro-2-thienylcyanoacetate 6.3

The formation of 2-amino-3-ethoxycarbonylindole from 2-nitrophenylcyanoacetate treated with anhydrous HCl and $SnCl_2$ has been reported by Snyder *et al.*[6.3] A more thorough literature search uncovered an obscure Russian article on the synthesis of thienopyrrole from *o*-halonitropyrroles, reported by Shedov *et al.*[6.4]

It seemed possible to use the dibromodinitrothiophene starting material **5.1** as the required reactant with ethyl cyanoacetate. The reaction of the thiophene **5.1** with the sodium derivative of ethyl cyanoacetate in a modified procedure of Shedov *et al.* resulted in the formation of a bright red opaque reaction mixture, assumed to contain the anion of **6.3**. However, addition of aqueous HCl did not affect the reaction mixture. Chromatographic purification of the mixture yielded mainly starting material **5.1**. Attempts to acid wash the organic mixture resulted in two visually inseparable dark red solvent layers which contained intractable mixtures from which product, if formed, could not be isolated.

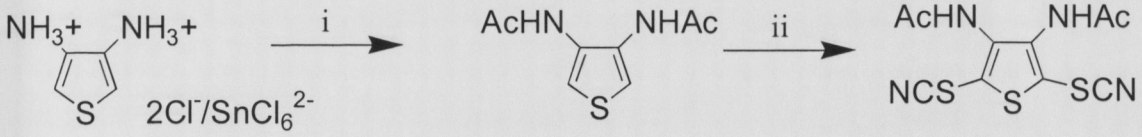


5.1 6.3

6.2.2 Attempted synthesis of 2-Methylthieno[3,2-*d*]thiazole 6.12

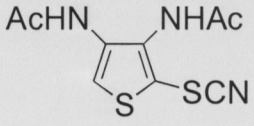
It was envisaged that the construction of the 6-amino-5-ethoxycarbonyl-2-methylpyrrolo[2',3':4,5]thieno[3,2-*d*]thiazole **6.1** could be completed by obtaining the thienothiazole structure **6.12** in the first instance. The pyrrole ring could be formed by subsequent nitration, reduction to the amine, enamine formation and Pd-catalysed ring closure as before. Although less elegant in its approach, this synthetic strategy might provide a more solid route to tricyclic formation.

It would be even more convenient to access a diaminothiophene derivative in the first instance. **Scheme 6.1** outlines the preparation of the bisacetylaminothieno bisthiocyanate **6.5** in two steps from the tin salt **5.2.** The free amine was liberated and reacted with acetic anhydride to form the diacetylaminothiophene **6.4** in the method of Easton.[5.1] Reaction of **6.4** with potassium thiocyanate and bromine yielded the bisthiocyanate **6.5** in the procedure of Paulmier.[6.2] Attempts to obtain the monothiocyanate **6.6** did not meet with success, as **6.5** was obtained as a precipitated product (the procedure specified required a two-fold excess of thiocyanate for mono-substitution) and subsequent continuous extraction of the reaction mixture yielded an intractable mixture. Further reaction of **6.5** seemed impractical at this point. Nucleophilic substitution at only one thiocyanate site seemed improbable and the time and effort involved in achieving mono-substitution could not be justified at this point in the project.



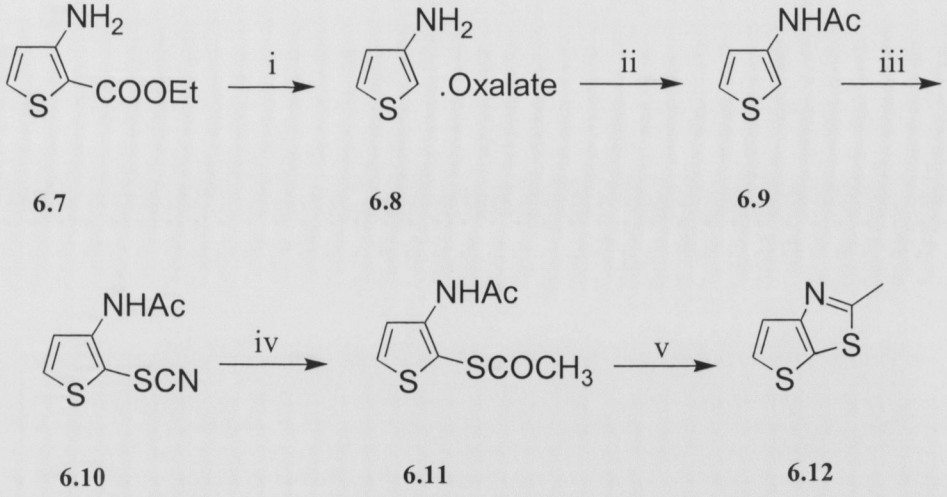
5.2 **6.4** **6.5**

Scheme 6.1 *Reagents, conditions (and yields)*: (i) Ac_2O, $NaHCO_3$, RT, overnight (86 %); (ii) NH_4SCN, Br_2, 0-10 °C, 1 h (64 %).



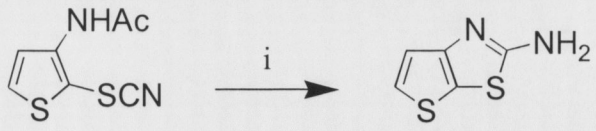
6.6

Scheme 6.2 details the attempted synthesis of 2-methylthieno[3,2-*d*] thiazole **6.12**. Methyl-3-aminothiophene-2-carboxylate **6.7** was hydrolysed to the acid in refluxing aqueous NaOH, then decarboxylation formed 3-aminothiophene **6.8** as the oxalate salt in a modification of the method of Barker *et al.*.[6.5] A globule of unreacted starting material was observed until full hydrolysis had occurred, which usually took longer than the time specified in the original procedure (30 min). The aminothiophene **6.8** was then liberated and converted to the acetamidothiophene **6.9** with acetic anhydride in the method of Campaigne *et al.*.[6.6] An earlier attempt to obtain **6.9** from 2-bromothiophene, sodium amide and liquid ammonia at –33 °C, and further derivatisation with acetic anhydride in a modified procedure of Reinecke *et al.*[6.7] was unsuccessful. The thienylthiocyanate **6.10** was obtained from the reaction of **6.9** with potassium thiocyanate and bromine in the method of Paulmier.[6.2] Thioacetylthiophene **6.11** was then produced from **6.10** by reaction of the sodium derivative obtained from the reaction of **6.12** with sodium sulfide with acetic anhydride in the procedure of Paulmier.[6.2]



Scheme 6.2 *Reagents, conditions (and yields)*: (i) NaOH, H_2O, reflux, 50 min, then HCl, then 2-propanol, oxalic acid, 45 min, 38 °C; (ii) NH_3 (soln.), Ac_2O; (iii) Br_2, NH_4SCN, MeOH, RT, 2 h (61 %); (iv) $Na_2S.9H_2O$, H_2O, RT, 2 h, then Ac_2O, 0-10 °C, 30 min (30 %); (v) Ethyl benzoate, 200 °C, 45 min.

The reaction to form methylthienothiazole **6.12** from thioacetylthiophene **6.11**, in the method of Paulmier,[6.2] did not go to completion. A temperature of 200 °C was required for reaction, which was attained with the use of a sand bath, and ethyl benzoate as solvent. Time did not allow further investigation. Similarly, an attempt to form the acetylamino thienothiazole **6.13** from thienylthiocyanate **6.10,** in the method of Paulmier,[6.2] (**Scheme 6.3**) did not yield a product at 200 °C, and time did not allow a repeated attempt.



6.10 **6.13**

Scheme 6.3 *Conditions*: (i) Ethyl benzoate, 200 °C.

[6.1] C. Paulmier, *Bull. Soc. Chim. Fr.* **1980**, 3-4(*II*), 151-156.

[6.2] C. Paulmier, *Bull. Soc. Chim. Fr.* **1979**, 11-12(*II*), 592-598.

[6.3] H. R. Snyder, E. P. Merica, C. G. Force and E. G. White, *J. Am. Chem. Soc.* **1958**, *80*, 4622-4625.

[6.4] V. I. Shedov, V. K. Vasil'era, Y. I. Trofimkin and A. H. Griner, *Chem. Hetero. Cmpds (Russ.)* **1975**, 1473-1475. This article was provided courtesy of their Latvian publisher.

[6.5] J. M. Barker, P. R. Huddleston and M. L. Wood, *Synth. Commun.* **1995**, *25(23),* 3729-3734.

[6.6] E. Campaigne and P. A. Monroe, *J. Am. Chem. Soc.* **1954**, *76*, 2447-2450.

[6.7] M. Reinecke, H. W. Adickes and C. Pyun, *J. Org. Chem.* **1971**, *36(18),* 2690-2692.

Chapter Seven

Experimental

Experimental

General:

^{1}H and ^{13}C NMR spectra were recorded on Bruker MSL-300, Bruker DPX-400, JEOL Delta/GX400 and JEOL ecp-400 spectrometers at 300 MHz and 75.5 MHz, and 400.14 MHz and 100.61 MHz, respectively, with the corresponding deuterated solvents noted. The chemical shifts are reported in parts per million relative to the residual non-deuterated solvent peak (δ_H $CHCl_3$ 7.27; δ_H [D_5]DMSO 2.50; and δ_C $CHCl_3$ 77.0 and δ_C [D_5]DMSO 39.5 central peak) or tetramethylsilane (δ_H TMS 0.0, δ_C TMS 0.0) as internal standard, where stated. *J* values are given in Hz. The DEPT and NOE techniques were used to assign different carbon atoms. The ^{13}C spectra are partially unassigned due to the ambiguity of the cyclobutenedione signal assignments. Chemical shifts are reported: value (splitting pattern, coupling constant (where applicable), number of protons; and assignment). Splitting pattern is designated as follows: s, singlet; app d, apparent doublet; d, doublet; dd, double doublet; t, triplet; q, quartet; qt, quintet; sept, septet; m, multiplet; and br, broad. Melting points were obtained using open-ended capillary tubes on an Electrothermal 9100 melting block or a Sanyo Gallenkamp digital melting point apparatus and are uncorrected.

FT-IR spectra were carried out using a Nujol mull on sodium chloride plates and were measured using a Perkin-Elmer FT-IR PARAGON 1000 spectrometer or a Mattson Genesis II FT-IR spectrometer; or as neat samples using a Perkin-Elmer One spectrometer. Mass spectra (low and high resolution) were obtained using a VG Analytical Autospec instrument equipped with a standard FAB positive source. Only the most significant molecular ions and/or base peaks in MS are reported. Matrix-assisted, laser desorption/ionisation-time of flight mass spectrometry was carried out without matrix elsewhere. LC-MS analyses were run with MassLynx v.3.3 software on a Waters 600 controller, which was connected to a 996 photodiode array detector, using a C_8 reverse phase column, eluting with water/acetonitrile gradient (95:5 5:95) with 1 % trifluoroacetic acid with run times of 20 min, flow rate 1 mL/min, coupled with a Micromass electrospray 600 spectrometer set in positive mode. Elemental analyses were carried out in the Microanalytical Laboratories of the Department of Chemistry, University College Dublin, and the School of Chemistry, University of Bristol. X-ray diffraction studies were carried out by Sylvia Draper, Department of Chemistry, Trinity College Dublin.

All reagents were purchased from Aldrich, Sigma, Fluka, Lancaster or Acros Organics, and were used without further purification, unless stated otherwise. All solvents were fractionally distilled before use and dried if necessary using standard techniques.[7.1] Reactions were followed by thin layer chromatography (TLC) on DC-Alufolien Kieselgel $60F_{254}$ 0.2 mm

plates and compounds were visualised by UV fluorescence, ninhydrin dip, potassium permanganate dip and phosphomolybdic acid dip. The reaction mixtures were evaporated in a vacuum rotary evaporator (Büchi *Rotavapor* R-114) using the vacuum of a water aspirator or diaphragm pump. This process is referred to below as "evaporated/removed/distilled *in vacuo*" or "under reduced pressure". Flash column chromatography refers to the technique described by Still.[7.2] The height of the silica gel 60 (400-230 mesh) in all cases was 15 cm.

All air and moisture sensitive reactions were carried out under nitrogen or argon atmospheres, as indicated, in oven-dried glassware. Reaction mixture temperatures were measured externally. Where solvent systems were used, the ratios are v/v unless otherwise stated.

3,4-Diisopropyl-3-cyclobutene-1,2-dione 2.5

This procedure is a modified from that of Liebeskind et al.[2.4]

To a suspension of 3,4-dihydroxy-3-cyclobutene-1,2-dione **2.4** (5.00 g, 0.044 mol) in a Dean and Stark apparatus was added 2-propanol (20 mL) and *p*-toluenesulphonic acid (0.940 g, 4.9 mmol) and the suspension refluxed at 90 °C for 24 h until a clear solution had formed and all the water formed had collected in the abstractor. The solvents were removed under reduced pressure. The clear residue was taken up with ethyl acetate (100 mL) and washed with sat. aqueous $NaHCO_3$ (10 mL), dried ($MgSO_4$), then concentrated *in vacuo*. The colourless clear residue was cooled under vacuum to form **2.5** as a hard white solid (8.160 g, 94 %). R_f = 0.26 (DCM); m.p. 43-45 °C (literature[2.4] 43-44 °C); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.36 (sept, *J* = 6.3 Hz, 2 H; C*H*), 1.47 (d, *J* = 6.3 Hz, 12 H; CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 188.36 (C=O), 181.43 (C=O), 175.83 (C), 173.04 (C), 74.87 (*C*H), 21.09 (*C*H_3); IR (neat): ν = 2987, 2937, 2875, 1806, 1736, 1582 cm^{-1}.

3-Amino-4-isopropoxy-3-cyclobutene-1,2-dione 2.6

This procedure is modified from that of Dunne.[1.74]

To a solution of 3,4-diisopropoxy-3-cyclobutene-1,2-dione **2.5** (1.580 g, 8 mmol) in THF (30 mL) and methanol (20 mL) was bubbled ammonia gas at a fast rate until the onset of a cloudy suspension had occurred. The solvent was then removed as quickly as possible under reduced pressure. The residue was dissolved in THF (100 mL), the solution filtered, then concentrated *in vacuo* to yield **2.6** (1.100 g, 89 % yield) as a fluffy white solid. R_f = 0.56 (EtOAc); m.p. 168-173 °C (decomp.) (literature[1.74] 173-175 °C (decomp.)); ^{1}H NMR (400 MHz, [D_6]DMSO): δ = 8.28 (b, 2 H; N*H*), 5.19 (sept, *J* = 6.0 Hz, 1 H; C*H*), 1.37 (d, *J* = 6.0 Hz, 6 H; CH_3); ^{13}C NMR (100.6 MHz, [D_6]DMSO): δ = 188.36 (C=O), 181.43 (C=O), 175.83 (C), 173.04 (C), 74.87 (*C*H), 21.09 (*C*H_3); IR (Nujol): ν = 3312, 3104, 2952-2852, 1808, 1714, 1699, 1618 cm^{-1}.

1-(2-Isopropoxy-3,4-dioxo-cyclobuten-1-yl)-3-phenyl-urea 2.8

This procedure is modified from that of Dunne.[1.74]

To a suspension of 3,4-diisopropoxy-3-cyclobutene-1,2-dione **2.5** (0.910 g, 5.88 mmol) in acetonitrile (10 mL) under Ar were quickly added phenyl isocyanate (0.32 mL, 2.94 mmol) and DBU (0.44 mL, 2.94 mmol) with stirring. The golden solution was left stirring at RT for 24 h; then concentrated *in vacuo* and purified by flash column chromatography (DCM/THF 30:1) to give **2.8** as a white solid (0.580 g, 72 % yield). R_f = 0.4 (DCM/THF 30:1); m.p. 154-157 °C; ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 9.27 (s, 1 H; Ph-N*H*), 8.21 (b, 1 H; N*H*-CO), 7.47 (d, *J* = 8.0 Hz, 2 H; Ar-C*H*), 7.32 (t, *J* = 7.5 Hz, 2 H; Ar-C*H*), 7.06 (t, *J* = 6.0 Hz, 1 H; Ar-C*H*), 5.40 (sept, *J* = 6.0 Hz, 1 H; C*H*), 1.40 (d, *J* = 6.0 Hz, 6 H; C*H*$_3$); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 186.78 (C=O), 180.95 (C=O), 169.12 (C), 148.8 (C), 138.24 (Ar-*C*), 128.85 (Ar-*C*), 123.15 (Ar-*C*), 118.7 (Ar-*C*), 77.68 (*C*H), 22.54 (*C*H$_3$); IR (neat): ν = 3293, 3098, 3042, 2985, 2887, 1805, 1712, 1683, 1636, 1588, 1565 cm^{-1}; elemental analysis calcd for $C_{14}H_{14}N_2O_4$ (274.3): C 61.31, H 5.14, N 10.21; found C 61.06, H 4.89, N 10.16.

1-(2-Methylamino-3,4-dioxo-cyclobuten-1-yl)-3-phenyl-urea 2.10

To a solution of **2.8** (0.232 g, 0.85 mmol) in DCM (5 mL) was added methylamine (40 % in H_2O, 0.3 mL, 0.01 mol) with stirring at RT overnight, then filtered and washed with acetone (4 x 5 mL) and DCM (4 x 5 mL) to afford **2.10** as a white solid (0.200 g, 96 % yield). R_f = 0.23 (DCM/EtOAc 4:1); m.p. >300 °C (decomp.); ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 9.61 (b, 1 H; N*H*), 7.43 (b, 1 H; N*H*), 7.41 (d, *J* = 7.5 Hz, 2 H; Ar-C*H*), 7.31 (t, *J* = 8.0 Hz, 2 H; Ar-C*H*), 7.00 (t, *J* = 6.0 Hz, 1 H; Ar-C*H*), 3.23 (s, 3 H; C*H*$_3$); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 184.0 (C=O), 170.0 (C=O), 163.5 (C=O), squaramide quaternary carbons not detected, 139.0 (Ar-*C*), 129.2 (Ar-*C*), 122.5 (Ar-*C*), 118.0 (Ar-*C*), 30.6 (*C*H$_3$); IR (Nujol): ν = 3185, 2921-2338, 1796, 1659, 1615 cm^{-1}; elemental analysis calcd for $C_{12}H_{11}N_3O_3$ (245.2): C 58.77, H 4.52, N 17.13; found C 58.90, H 4.53, N 16.93.

1-(2-Butylamino-3,4-dioxo-cyclobuten-1-yl)-3-phenyl-urea 2.11

To a solution of **2.8** (100 mg, 0.37 mmol) in DCM (10 mL) was added butylamine (0.36 mL, 3.7 mmol) with stirring. The precipitate obtained was purified by the procedure of **2.10** above to yield **2.11** (50 mg, 47 % yield) as a white powder. R_f = 0.33 (EtOAc); m.p. 296 ºC (decomp.); ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 9.57 (s, 1 H; N*H*), 7.62 (b, 1 H; N*H*), 7.43 (d, *J* = 8.0 Hz, 2 H; Ar-C*H*), 7.31 (t, *J* = 8.0 Hz, 2 H; Ar-C*H)* 7.00 (t, *J* = 7.5 Hz, 1 H; Ar-C*H*), 3.60 (q, *J* = 6.0 Hz, 2 H; CH_2), 1.53 (m, *J* = 7.5 Hz, 2 H; CH_2), 0.89 (t, *J* = 7.5 Hz, 3 H; CH_3); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 183.90 (C=O), 180.07 (C=O), 169.21 (C), 163.44 (C=O), 139.01 (Ar-*C*), 129.24 (Ar-*C*), 122.48 (Ar-*C*), 117.90 (Ar-*C*), 43.31 (*C*H$_2$), 32.57 (*C*H$_2$), 18.57 (*C*H$_2$), 13.40 (*C*H$_3$); IR (neat): ν = 3267, 3139, 2957, 2860, 1794, 1720, 1680, 1585 cm^{-1}; elemental analysis calcd for $C_{15}H_{17}N_3O_3$ (287.3): C 62.71, H 5.96, N 14.63; found C 62.59, H 5.91, N 14.92.

1-Phenyl-3-(3,4-dioxo-2-phenylamino-cyclobuten-1-yl)-urea 2.12

Route A:

To a solution of **2.8** (0.160 g, 0.6 mmol) in DCM (10 mL) was added aniline (0.53 mL, 6 mmol) with stirring. The solution was left stirring at RT overnight, then concentrated *in vacuo*. Ethyl acetate (20 mL) was added to the residue and the precipitate formed filtered and washed with acetone (4 X 5 mL) and DCM (4 X 5 mL) under suction to yield **2.12** as a white powder (0.140 g, 78 % yield).

Route B:

To a solution of **2.19** (0.250 g, 0.9 mmol) in DCM (12 mL) was added aniline (0.84 mL, 9 mmol) with stirring. The procedure followed was that of **Route A** above to yield **2.12** as a white powder (0.100 g, 37 %). R_f = 0.23 (4:1 DCM/THF); m.p. > 300 ° C (decomp.); ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 9.97 (b, 2 H; N*H*), 7.49 (d, *J* = 8.0 Hz, 4 H; Ar-C*H*), 7.36 (t, *J* = 7.5 Hz, 4 H; Ar-C*H*), 7.07 (t, *J* = 7.0 Hz, 2 H; Ar-C*H*); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 181.55 (C=O), 180.71 (C=O), 166.33 (C), 165.33 (C=O), 138.51 (Ar-*C*), 129.29 (Ar-*C*), 129.00 (Ar-*C*), 123.20 (Ar-*C*), 118.44 (Ar-*C*); IR (Nujol): ν = 3141, 2951-2852, 1797, 1672, 1602 cm^{-1}; elemental analysis calcd for $C_{17}H_{13}N_3O_3$ (307.3): C 66.44, H 4.26, N 13.67; found C 66.24, H 4.26, N 13.57.

1-(2-Cyclohexylamino-3,4-dioxo-cyclobuten-1-yl)-3-phenyl-urea 2.13

Route A:

To a solution of **2.8** (177 mg, 0.65 mmol) in DCM (10 mL) was added cyclohexylamine (0.74 mL, 0.5 mmol) with stirring. The procedure followed was that of **2.10** above to afford **2.13** as a white powder (105 mg, 52 % yield).

Route B:

To a solution of **2.19** (100 mg, 0.35 mmol) in DCM (10 mL) was added cyclohexylamine (0.74 mL, 0.5 mmol) with stirring. The procedure followed was that of **2.10** above to afford **2.13** as a white powder (90 mg, 82 %). M.p. 280-300 °C (decomp.); ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 9.58 (b, 1 H; N*H*), 8.54 (b, 1 H; N*H*), 7.44 (d, *J* = 7.5 Hz, 2 H; Ar-C*H*), 7.34 (t, *J* = 7.0 Hz, 2 H; Ar-C*H*), 7.01 (t, *J* = 6.5 Hz, 1 H; Ar-C*H*), 3.87 (s, 1 H; Cy-C*H*), 1.94 (b, 2 H; C*H*), 1.73 (b, 2 H; Cy-C*H*), 1.56 (app d, *J* = 12.0 Hz, 1 H; Cy-C*H*), 1.34 (m, *J* = 12.0, 4.0 Hz, 4 H; Cy-C*H*), 1.21 (s, 1 H; Cy-C*H*); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 179.93 (C=O), 168.41 (C=O), 163.52 (C=O), squaramide quaternary carbons not detected, 139.00 (Ar-*C*), 129.23 (Ar-*C*), 122.46 (Ar-*C*), 117.86 (Ar-*C*), 52.48 (Cy-*C*), 33.52 (Cy-*C*), 24.65 (Cy-*C*), 23.90 (Cy-*C*); IR (Nujol): ν = 3162, 2921-2852, 1754, 1664, 1603, 1548, 1500 cm^{-1}.

1-(2-Hydrazino-3,4-dioxo-cyclobuten-1-yl)-3-phenyl-urea 2.14

Route A:

To a solution of **2.8** (0.660 mg, 2.18 mmol) in THF (10 mL) was added hydrazine hydrate (1.06 mL, 0.022 mol) with stirring. The procedure followed was that of **2.10** above to yield **2.14** as a cream solid (0.535 g, 100 %).

Route B:

To a solution of **2.19** (0.583 g, 2.02 mmol) in acetonitrile (15 mL) was added hydrazine hydrate (0.5 mL, 0.01 mol) with stirring. The procedure followed was that of **2.10** above to yield **2.14** as a cream solid (0.450 g, 91 %). R_f = 0.56 (EtOAc); m.p. 144 °C (decomp.); ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 7.49 (d, *J* = 7.5 Hz, 2 H; Ar-C*H*), 7.29 (t, *J* = 8.0 Hz, 2 H; Ar-C*H*), 6.99 (t, *J* = 7.5 Hz, 1 H; Ar-C*H*), 5.14 (b, 2 H; N*H*), 3.38 (b, 2 H; N*H*); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 170.88 (C=O), 169.86 (C=O), squaramide quaternary carbons not detected, 128.97 (Ar-*C*), 122.31 (Ar-*C*), 118.33 (Ar-*C*); IR (Nujol): ν = 3328, 3192, 2952-2852, 1800, 1666, 1641, 1595 cm^{-1}.

1-Phenyl-3-(2-phenylhydrazino-3,4-dioxo-cyclobuten-1-yl)-urea 2.15

To a solution of **2.8** (0.260 g, 0.95 mmol) in THF (10 mL) and ethanol (5 mL) was added phenylhydrazine (0.9 mL, 9.1 mmol) with stirring. The solution was left stirring at RT for 3 h. The procedure followed was that of **2.10** above to afford **2.15** as a white solid (0.220 g, 72 % yield). R_f = 0.22 (DCM/EtOAc 4:1); m.p. 228 °C (decomp.); ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 10.19 (b, 1 H; N*H*), 9.29 (b, 1 H; N*H*), 8.25 (b, 1 H; N*H*), 7.48 (d, *J* = 8.0 Hz, 2 H; Ar-C*H*), 7.32 (t, *J* = 8.0 Hz, 2 H; Ar-C*H*), 7.20 (t, *J* = 7.5 Hz, 2 H; Ar-C*H*), 7.07 (t, *J* = 7.5 Hz, 1 H; Ar-C*H*), 6.81 (d, *J* = 6.0 Hz, 3 H; Ar-C*H*); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): carbonyls not detected, δ = 158.95 (C), 150.31 (C), 148.68 (Ar-*C*), 137.70 (Ar-*C*), 128.87 (Ar-*C*), 123.49 (Ar-*C*), 119.46 (Ar-*C*), 119.09 (Ar-*C*), 118.09 (Ar-*C*), 112.48 (Ar-*C*); IR (Nujol): ν = 3296, 2923-2852, 1807, 1722, 1684, 1629, 1600 cm^{-1}; elemental analysis calcd for $C_{17}H_{14}N_4O_3$ (322.3): C 63.35, H 4.38, N 17.38; found C 63.15, H 4.48, N 17.38.

1-(2-Benzylhydrazino-3,4-dioxo-cyclobuten-1-yl)-3-phenyl-urea 2.16

To a suspension of **2.8** (0.250 g, 0.91 mmol) in acetonitrile (10 mL) were added benzylhydrazine dihydrochloride (0.355 g, 1.82 mmol) and DIPEA (0.31 mL, 1.82 mmol) with stirring. A solution formed immediately, with the release of gaseous HCl, and after a few minutes, formed a suspension, which was allowed to stir at RT overnight. The procedure followed was that of **2.10** above to afford **2.16** (0.145 g, 47 % yield) as a yellow powder. R_f = 0.19 (DCM/EtOAc 9:1); m.p. 226 °C (decomp.); ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 10.22 (b, 1 H; N*H*), 9.31 (b, 1 H; N*H*), 9.10 (b, 1 H; N*H*), 8.79 (b, 1 H; N*H*), 7.29 (m, 9 H; Ar-C*H*), 7.05 (m, 1 H; Ar-C*H*), 4.87 (s, 2 H; C*H*$_2$); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 183.78 (C=O), 180.19 (C=O), 168.51 (C), 163.52 (C=O), 138.77 (Ar-*C*), 128.97 (Ar-*C*), 128.44 (Ar-*C*), 127.46 (Ar-*C*), 122.31 (Ar-*C*), 117.76 (Ar-*C*), 46.02 (*C*H$_2$); IR (neat): ν = 3279, 3205, 2955, 2867, 1805, 1716, 1683, 1599, 1544, 1519 cm^{-1}.

1-(2-*tert*-Butylhydrazino-3,4-dioxo-cyclobuten-1-yl)-3-phenyl-urea 2.17

To a suspension of **2.8** (0.290 g, 1 mmol) in acetonitrile (10 mL) was added *tert*-butylhydrazine hydrochloride (0.626 g, 5 mmol) and DIPEA (0.87 mL, 5 mmol) with stirring. The clear yellow solution was then heated to 50 °C and stirred for 3 h. The orange solution was allowed to cool overnight. The precipitate formed was collected and washed as in the procedure of **2.10** above to yield **2.17** as a white solid (0.100 g, 33 % yield). R_f = 0.33 (DCM/EtOAc 4:1); m.p. 239 °C (decomp.); ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 10.12 (b, 1 H; N*H*), 9.42 (b, 1 H; N*H*), 8.68 (b, 1 H; N*H*), 7.45 (d, *J* = 7.5 Hz, 2 H; Ar-C*H*), 7.33 (t, *J* = 8.0 Hz, 2 H; Ar-C*H*), 7.08 (t, *J* = 7.5 Hz, 1 H; Ar-C*H*), 5.14 (b, 2 H; N*H*), 1.06 (s, 9 H; C(C*H*$_3$)$_3$); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 182.82 (C=O), 171.95 (C=O), 157.60 (C=O), squaramide quaternary carbons not detected, 150.66 (*C*(CH$_3$)), 137.80 (Ar-*C*), 128.93 (Ar-*C*), 123.93 (Ar-*C*), 119.06 (Ar-*C*), 26.28 (C(*C*H$_3$)$_3$); IR (Nujol): ν = 3266, 3199, 2922-2852, 1798, 1720, 1682, 1606 cm^{-1}.

3-Amino-4-butoxy-3-cyclobutene-1,2-dione 2.18

To a solution of 3,4-dibutoxy-3-cyclobutene-1,2-dione (1.80 mL, 1.890 g, 8.3 mmol) in THF (30 mL) and methanol (20 mL) was bubbled gaseous ammonia at a fast rate until a cloudy suspension had formed. The procedure then followed was that of **2.6** above to yield **2.18** as a white solid (1.273 g, 90 % yield). R_f = 0.23 (DCM/EtOAc 4:1); m.p. 164-165 °C; ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 8.28 (b, 2 H; N*H*$_2$), 4.60 (t, *J* = 6.5 Hz, 2 H; C*H*$_2$), 1.69 (m, 2 H; C*H*$_2$), 1.35 (m, 2 H; C*H*$_2$), 0.91 (t, *J* = 7.5 Hz, 3 H; C*H*$_3$); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 190.05 (C=O), 183.18 (C=O), 177.91 (C), 174.28 (C), 72.16 (O-*C*H$_2$), 31.48 (*C*H$_2$), 18.07 (*C*H$_2$), 13.47 (*C*H$_3$); IR (Nujol): ν = 3312, 3117, 2952-2858, 1806, 1712, 1700, 1616 cm^{-1}; elemental analysis calcd for $C_8H_{11}NO_3$ (169.2): C 56.80, H 6.55, N 8.28; found C 56.81, H 6.39, N 8.13.

1-(2-Butoxy-3,4-dioxo-cyclobuten-1-yl)-3-phenyl-urea 2.19

To a suspension of 3-amino-4-butoxy-3-cyclobutene-1,2-dione **2.18** (0.430 g, 2.54 mmol) in acetonitrile (15 mL) under Ar were quickly added phenyl isocyanate (0.14 mL, 1.27 mmol) and DBU (0.19 mL, 1.27 mmol) with stirring. The solution was allowed to stir for 24 h at RT, then the solvents were removed *in vacuo* and the residue purified by flash column chromatography (DCM/THF 30:1) to give **2.19** as a white solid (0.308 g, 86 %). R_f = 0.5 (DCM/THF 30:1); m.p. 178-185 °C (decomp.); ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 9.83 (s, 1 H; N*H*), 8.06 (b, 1 H; N*H*), 7.59 (d, *J* = 8.0 Hz, 2 H; Ar-C*H*), 7.33 (m, $J_{o,m}$ = 8.5, $J_{o,p}$ = 1.5 Hz, 2 H; Ar-C*H*), 7.15 (m, *J* = 7.0 Hz, 1 H; Ar-C*H*), 4.86 (t, *J* = 6.5 Hz, 2 H; CH_2-O), 1.76 (m, *J* = 6.5 Hz, 2 H; CH_2), 1.46 (m, *J* = 7.0 Hz, 2 H; CH_2), 0.99 (t, *J* = 7.5 Hz, 3 H; CH_3); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 182.23 (C=O), squaramide quaternary carbons not detected, 164.73 (C=O), 135.18 (Ar-*C*), 127.25 (Ar-*C*), 122.85 (Ar-*C*), 111.40 (Ar-*C*), 73.74 (*C*H), 72.34 (O-*C*H$_2$), 30.10 (*C*H$_2$), 16.64 (*C*H$_2$), 11.72 (*C*H$_3$); IR (neat): ν = 3210, 3051, 2955, 2871, 2669, 2436, 1814, 1717, 1694, 1582, 1541 cm^{-1}; elemental analysis calcd for $C_{15}H_{16}N_2O_4$ (288.3): C 62.49, H 5.59, N 9.72; found C 62.57, H 5.29, N 9.09.

1-(2-Benzylamino-3,4-dioxo-cyclobuten-1-yl)-3-phenyl-urea 2.20

To a solution of **2.19** (200 mg, 0.7 mmol) in DCM (10 mL) was added benzylamine (0.75 mL, 7 mmol) with stirring. The procedure followed was that of **2.10** above to yield **2.20** as a cream solid (120 mg, 51 %). M.p. 276-279 °C (decomp.); ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 8.24 (s, 3 H; N*H*), 7.46 (d, *J* = 7.0 Hz, 2 H; Ar-C*H*), 7.42 (d, *J* = 5.0 Hz, 4 H; Ar-C*H*), 7.32 (m, 3 H; Ar-C*H*), 7.02 (t, *J* = 7.0 Hz, 1 H; Ar-C*H*), 4.87 (s, 2 H; CH_2); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 183.78 (C=O), 180.19 (C=O), 168.51 (C) 163.52 (C=O), 138.77 (Ar-*C*), 138.13 (Ar-*C*), 128.44 (Ar-*C*), 127.36 (Ar-*C*), 122.31 (Ar-*C*), 117.76 (Ar-*C*), 47.06 (*C*H$_2$); IR (neat): ν = 3174, 2974, 2872, 1796, 1654, 1607, 1557 cm^{-1}; elemental analysis calcd for $C_{18}H_{15}N_3O_3$ (321.3): C 67.28, H 4.71, N 13.08; found C 67.79, H 4.74, N 12.76.

1-Benzyl-3-(2-isopropoxy-3,4-dioxo-cyclobuten-1-yl)-urea 2.21

To a solution of 3-amino-4-isopropoxy-3-cyclobutene-1,2-dione **2.6** (0.720 g, 4.64 mmol) in acetonitrile (15 mL) were added benzyl isocyanate (0.55 mL, 0.59 g, 4.45 mmol) and DBU (0.7 mL, 4.68 mmol). The solution was stirred at RT for 150 min. The solvents were removed *in vacuo* and the residue purified by flash column chromatography in DCM/EtOAc (2:1) to yield 0.880 g (65 % yield) of **2.21** as a white waxy solid. R_f = 0.6 (DCM/EtOAc 2:1); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.82 (b, 1 H; N*H*), 8.19 (b, 1 H; N*H*), 7.34 (m, 5 H; Ar-C*H*), 5.47 (sept, *J* = 6.5 Hz, 1 H; C*H*-O), 4.49 (d, *J* = 6.0 Hz, 2 H; Ph-CH_2), 1.48 (d, *J* = 6.0 Hz, 6 H; CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 181.55 (C=O), 167.54 (C), 152.39 (C=O), 137.63 (Ar-*C*), 128.75 (Ar-*C*), 127.66 (Ar-*C*), 127.38 (Ar-*C*), 80.24 (*C*H), 44.83 (Ph-*C*H_2), 22.82 (*C*H_3).

1-Benzyl-3-(2-phenylhydrazino-3,4-dioxo-cyclobuten-1-yl)-urea 2.22

To a solution of **2.21** (0.530 g, 1.84 mmol) in DCM/EtOH (50:50, 10 mL) stirred at RT was added phenylhydrazine (1.81 mL, 0.018 mol). The resulting suspension was allowed to stir for 2 h. The solid was then removed by filtration, and the filtrate was allowed to react further overnight. The precipitate obtained was again filtered off and the two precipitates combined and washed with DCM/EtOH to yield **2.22** as a white solid (0.110 g, 18 %). M.p. 233 °C (decomp.); ^{1}H NMR (400 MHz, [D_6]DMSO): δ = 10.37 (b, 1 H; N*H*), 9.13 (b, 1 H; N*H*), 8.20 (b, 1 H; N*H*), 7.45 (b, 1 H; N*H*), 7.32 (m, 5 H; Ar-C*H*), 7.21 (m, 2 H; Ar-C*H*), 6.77 (m, 3 H; Ar-C*H*), 4.37 (d, *J* = 5.5 Hz; 2 H; Ph-CH_2); ^{13}C NMR (100.6 MHz, [D_6]DMSO): δ = 183.39, 182.54, 180.76, 178.99 (C), 159.98 (C=O), 138.82 (Ar-*C*), 129.00 (Ar-*C*), 128.45 (Ar-*C*), 127.09 (Ar-*C*), 112.54 (Ar-*C*), 43.26 (*C*H_2); elemental analysis calcd for $C_{18}H_{16}N_4O_3$ (336.4): C 64.28, H 4.79, N 16.66; found C 64.09, H 4.84, N 16.48.

1-(2-Isopropoxy-3,4-dioxo-cyclobuten-1-yl)-3-phenyl-thiourea 3.1

To a suspension of 3-amino-4-isopropoxy-3-cyclobutene-1,2-dione **2.6** (0.150 g, 1 mmol) in acetonitrile (10 mL) were added phenyl isothiocyanate (115 μL, 1 mmol) and DBU (0.15 mL, 1 mmol) simultaneously with stirring at RT under Ar. The solution was left stirring overnight, then concentrated *in vacuo* and purified by flash column chromatography (DCM) to yield **3.1** (0.280 g, 95 % yield) as a white solid. R_f = 0.5 (DCM); m.p. 172 °C (decomp.); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 11.01 (b, 1 H; N*H*), 8.41 (b, 1 H; N*H*), 7.74 (d, *J* = 8.0 Hz, 2 H; Ar-C*H*), 7.42 (t, *J* = 8.0 Hz, 2 H; Ar-C*H*), 7.31 (t, *J* = 7.5 Hz, 1 H; Ar-C*H*), 5.48 (sept, *J* = 6.0 Hz, 1 H; C*H*), 1.54 (d, *J* = 6.5 Hz, 6 H; CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 187.46 (C=O), 180.21 (C=O), 174.83 (C=S), 137.61 (Ar-*C*), 128.45 (Ar-*C*), 126.50 (Ar-*C*), 123.30 (Ar-*C*), 80.93 (*C*H), 22.37 (CH_3); IR (neat): ν = 3136, 2935, 2871, 1800, 1728, 1627, 1575, 1516 cm^{-1}; elemental analysis calcd for $C_{14}H_{14}N_2O_3S$ (290.3): C 57.92, H 4.86, N 9.65; found C 57.70, H 4.89, N 9.57.

1-Phenyl-3-(2-phenylamino-3,4-dioxo-cyclobuten-1-yl)-thiourea 3.2

To a stirred solution of **3.1** (0.370 g, 1.27 mmol) in DCM/EtOH (50:50, 15 mL) was added aniline (1.16 mL, 12.7 mmol) and the mixture reacted at RT overnight. The solid was collected and washed with DCM/EtOH (50:50) to give **3.2** as a lemon yellow powder (0.300 g, 73 % yield). R_f = 0.37 (DCM/EtOAc 9:1); m.p. 217 °C (decomp.); IR (neat): ν = 3288, 3071, 2903, 1796, 1701, 1605, 1622, 1594, 1563, 1533, 1500 cm^{-1}; elemental analysis calcd for $C_{17}H_{13}N_3O_2S$ (323.4): C 63.14, H 4.05, N 12.99; found C 62.84, H 4.07, N 12.87.

1-(2-Hydrazino-3,4-dioxo-cyclobuten-1-yl)-3-phenyl-thiourea 3.3

To a suspension of **3.1** (0.202 g, 0.7 mmol) in DCM (20 mL) was added hydrazine hydrate (0.1 mL, 20 mmol) with stirring. The orange solution was left stirring for 2 h, by which time all of the starting material had reacted. The solution was then concentrated in vacuo and hot ethanol (20 mL) added to precipitate **3.3** as a fine yellow powder (0.170 g, 100 % yield). R_f = 0.22 (DCM/EtOAc 9:1); m.p. 234-237 °C (decomp.); ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 11.38 (b, 2 H; N*H*), 10.81 (b, 1 H; N*H*), 9.64 (b, 1 H; N*H*), 7.69 (d, *J* = 8.1 Hz, 2 H; Ar-C*H*), 7.41 (t, *J* = 8.0 Hz, 2 H; Ar-C*H*), 7.24 (t, *J* = 7.5 Hz, 1 H; Ar-C*H*); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 170.88 (C), 169.82 (C), carbonyls not detected, 128.62 (Ar-*C*), 125.92 (Ar-*C*), 123.47 (Ar-*C*); IR (neat): ν = 3294, 3234, 2991, 1800, 1698, 1651, 1599, 1568, 1518 cm^{-1}.

1-Butyl-3-(2-isopropoxy-3,4-dioxo-cyclobuten-1-yl)-thiourea 3.4

To a stirring suspension of 3-amino-4-isopropoxy-3-cyclobutene-1,2-dione **2.6** (1.770 g, 0.011 mol) in acetonitrile (15 mL) at RT were added butyl isothiocyanate (1.38 mL, 0.011 mol) and DBU (1.65 mL, 0.011 mol). The solution was allowed to stir at RT overnight. It was then concentrated *in vacuo* and the residue purified by flash column chromatography (DCM) to yield **3.4** as a white powder (1.990 g, 68 %). R_f = 0.4 (DCM); m.p. 147-151 °C (decomp.); ^{1}H NMR (400 MHz, CDCl$_3$): δ = 9.34 (b, 1 H; N*H*), 8.52 (b, 1 H; N*H*), 5.43 (sept, *J* = 6.0 Hz, 1 H; C*H*-O), 3.64 (m, *J* = 5.0 Hz, 2 H; C*H*$_2$-N), 1.69 (m, *J* = 7.0 Hz, 2 H; C*H*$_2$), 1.51 (d, *J* = 6.0 Hz, 6 H; C*H*$_3$), 1.41 (m, *J* = 7.0 Hz, 2 H; C*H*$_2$CH$_3$), 0.96 (t, *J* = 7.5 Hz, 3 H; C*H*$_3$); ^{13}C NMR (100.6 MHz, CDCl$_3$): δ = 187.24 (C=O), 181.93 (C=O), 180.14 (C), 176.48 (C), 165.50 (C=S), 80.70 (*C*H), 46.83 (*C*H$_2$), 29.77 (*C*H$_2$), 22.48 (*C*H$_2$), 19.78 (*C*H$_3$), 13.32 (*C*H$_3$); IR (Nujol): ν = 2922-2852, 1795, 1714, 1588 cm^{-1}; elemental analysis calcd for C$_{12}$H$_{18}$N$_2$O$_3$S (270.3): C 53.31, H 6.71, N 10.36, S 11.86; found C 53.31, H 6.65, N 10.26, S 11.97.

1-Butyl-3-(2-hydrazino-3,4-dioxo-cyclobuten-1-yl)-thiourea 3.5

Hydrazine hydrate (0.3 mL, 6 mmol) was added to a stirring solution of **3.4** (0.143 g, 0.53 mmol) in DCM (10 mL), to quickly form a suspension which was allowed to stir at RT overnight, then purified by flash column chromatography (DCM/THF 30:1) to afford **3.5** as a white powder (0.128 g, 100 %). R_f = 0.2 (DCM/THF 30:1); m.p. 198 °C (decomp.); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 9.98 (b, 1 H; N*H*), 9.62 (b, 1 H; N*H*), 9.37 (b, 1 H; N*H*), 3.66 (m, *J* = 4.5 Hz, 2 H; C*H*$_2$-N), 1.71 (m, *J* = 7.5 Hz, 2 H; C*H*$_2$), 1.44 (m, *J* = 7.5 Hz, 2 H; C*H*$_2$), 0.97 (t, *J* = 7.0 Hz, 3 H; C*H*$_3$); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 180.34 (C=O), 176.57 (C=O), 164.07 (C), 161.22 (C), 158.13 (C=S), 46.66 (*C*H$_2$), 29.74 (*C*H$_2$), 19.69 (*C*H$_2$), 16.49 (*C*H$_3$); IR (Nujol): ν = 3196, 2920-2852, 1798, 1694, 1600 cm^{-1}.

1-(2-Isopropoxy-3,4-dioxo-cyclobuten-1-yl)-3-(methyl)thiourea 3.6

To a suspension of 3-amino-4-isopropoxy-3-cyclobutene-1,2-dione **2.6** (1.050 g, 6.73 mmol) in acetonitrile (10 mL) were added methyl isothiocyanate (0.493 g, 6.75 mmol) and DBU (101 μL, 6.75 mmol) with stirring, and the solution allowed to stir overnight at RT. The solvent was removed *in vacuo* and the residue purified by flash column chromatography (DCM) to yield **3.6** as a white solid (0.847 g, 55 %). R_f = 0.17 (DCM); m.p. 187-189 °C (decomp.); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 9.43 (b, 1 H; N*H*), 8.61 (b, 1 H; N*H*), 6.25 (sept, *J* = 6.5 Hz, 1 H; C*H*-O), 3.25 (d, *J* = 4.5 Hz, 3 H; C*H*$_3$-N), 1.53 (d, *J* = 6.0 Hz, 6 H; C*H*$_3$; ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 187.71 (C=O), 182.56 (C=O), 180.54 (C), 178.02 (C), 165.76 (C=S), 81.12 (*C*H), 33.52 (*C*H$_3$-N), 22.80 ((*C*H$_3$)$_3$C); IR (neat): ν = 3070, 3007, 2935, 1795, 1708, 1583, 1528 cm^{-1}; elemental analysis calcd for $C_9H_{12}N_2O_3S$ (228.26): C 47.36, H 5.30, N 12.27, S 14.05; found C 47.12, H 5.22, N 12.08, S 13.85.

1-(2-Hydrazino-3,4-dioxo-cyclobuten-1-yl)-3-methyl-thiourea 3.7

To a solution of **3.6** (0.370 g, 1.62 mmol) in DCM (20 mL) was added hydrazine hydrate (80 μL, 1.62 mmol). The suspension was allowed to stir for 3 h; then the precipitated solid collected and washed with DCM to yield 0.265 g (82 %) of **3.7** as a bright custard yellow solid. R_f = 0.42 (EtOAc); m.p. 222 °C (decomp.); ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 9.83 (b, 1 H; N*H*), 9.66 (b, 1 H; N*H*), 5.40 (b, 1 H; N*H*), 1 H; C*H*-O), 3.06 (s, 3 H; C*H*$_3$); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 179.75 (C=O), 177.51 (C=O), 169.61 (C) 157.31 (C=S), 32.67 (*C*H$_3$); IR (neat): ν = 3299, 3192, 2998, 2940, 1797, 1681, 1639, 1539, 1512 cm^{-1}; MS(FAB): *m/z* = 201, calcd for $C_6H_8N_4O_2S$ + H: 201.

1-Benzyl-3-(2-isopropoxy-3,4-dioxo-cyclobut-1-enyl)-thiourea 3.8

To a stirred suspension of **2.6** (1.055 g, 6.75 mmol) in acetonitrile (20 mL) was added benzyl isothiocyanate (0.9 mL, 6.79 mmol), then DBU (1.015 mL, 6.75 mmol). The resulting orange solution was allowed to stir at RT for 90 min. The solvents were removed *in vacuo* and the residue purified by flash column chromatography (DCM/THF 30:1) to yield 1.660 g (81 % yield) of **3.8** as a lemon white solid. R_f = 0.4 (DCM/EtOAc 49:1); m.p. 191-192 °C (decomp.); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 9.68 (b, 1 H; N*H*), 8.52 (b, 1 H; N*H*), 7.36 (m, 5 H; Ar-C*H*), 5.47 (sept, *J* = 6.5 Hz, 1 H; C*H*-O), 4.88 (d, *J* = 5.5 Hz, 2 H; Ph-CH_2), 1.48 (d, *J* = 6.5 Hz, 6 H; CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 187.43 (C=O), 182.27 (C=O), 180.46 (C), 177.08 (C), 165.61 (C=S), 135.85 (Ar-*C*), 128.80 (Ar-*C*), 128.20 (Ar-*C*), 128.07 (Ar-*C*), 81.13 (Ar-*C*), 51.09 (Ph-*C*H_2), 22.77 (*C*H_3); IR (neat): ν = 3136, 2976, 2916, 1794, 1715, 1591, 1583, 1529 cm^{-1}; elemental analysis calcd for $C_{15}H_{16}N_2O_3S$ (304.4): C 59.19, H 5.30, N 9.20, O 15.77, S 10.53; found C 59.20, H 5.32, N 9.21, S 10.66.

1-Benzyl-3-(2-hydrazino-3,4-dioxo-cyclobuten-1-yl)-thiourea 3.9

To a stirring solution of **3.8** (0.560 g, 1.84 mmol) in DCM (10 mL) was added hydrazine hydrate (90 μL, 2.97 mmol), and the resulting mixture allowed to stir at RT overnight. The solvent was removed *in vacuo* and the yellow residue dissolved in hot ethanol (10 mL). On cooling, the yellow precipitate was collected and washed with ethanol to yield 0.450 g (89 %) of **3.9** as a bright yellow solid when dry. R_f = 0.46 (EtOAc); m.p. 159 °C (decomp.); ^{1}H NMR; (400 MHz, [D_6]DMSO): δ = 11.01 (b, 1 H; N*H*); 10.75 (b, 1 H; N*H*); 10.16 (b, 1 H; N*H*); 9.47 (b, 1 H; N*H*); 7.36 (d, *J* = 4 Hz, 4 H; Ar-C*H*); 7.30 (qt, *J* = 3.5 Hz, 1 H, Ar-C*H*); 4.83 (d, *J* = 5.0 Hz, 2 H, Ph-CH_2); ^{13}C NMR (100.6 MHz, [D_6]DMSO): δ = 182.47 (*C*), 180.67 (*C*), 177.80 (*C*), 166.15 (*C*), 157.79 (C=S), 137.01 (Ar-*C*), 128.52 (Ar-*C*), 127.67 (Ar-*C*), 127.42 (Ar-*C*), 49.11 (Ph-*C*H_2), IR (neat): ν = 3272, 3229, 2956, 1802, 1686, 1646, 1560, 1510 cm^{-1}; MS(FAB): *m/z* = 277; calcd for $C_{15}H_{16}N_4O_2S$ + H: 277.

1-(2-*tert*-Butoxycarbonylhydrazino-3,4-dioxo-cyclobuten-1-yl)-3-butyl-thiourea 3.11

To a suspension of **3.4** (0.420 g, 1.55 mmol) in acetonitrile (10 mL) was added *tert*-butyl carbazate (1.20 g, 8 mmol) and the suspension allowed to stir at RT overnight. The solution was concentrated *in vacuo* and the resulting residue purified by column chromatography (DCM/EtOAc 1:1) to yield 0.550 g (87 %) of **3.11** as a cream yellow solid. R_f = 0.51 (DCM/EtOAc 4:1); m.p. 130-132 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.22 (b, 1 H; N*H*), 3.61 (q, *J* = 7.0 Hz, 2 H; -C*H*$_2$-N), 1.65 (m, *J* = 7.0 Hz, 2 H; -C*H*$_2$-), 1.48 (s, 9 H; (C*H*$_3$)$_3$C), 1.40 (m, *J* = 7.0 Hz, 2 H; -C*H*$_2$-), 0.96 (t, *J* = 7.0 Hz, 3 H; -CH$_2$C*H*$_3$); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 180.75 (C), 176.77 (C), 154.37 (C=S), 76.40 (*C*(CH$_3$)$_3$), 29.71 (*C*H$_2$), 27.69 (C*H*$_2$), 19.60 (*C*H$_2$), 13.19 (*C*H$_3$); IR (neat): ν = 3283, 2961, 2935, 2875, 1806, 1730, 1701, 1626, 1574, 1538 cm^{-1}; elemental analysis calcd for $C_{14}H_{22}N_4O_4S$ (342.4): C 49.11, H 6.48, N 16.38; found C 49.00, H 6.61, N 16.33.

1-(2-*tert*-Butoxycarbonylhydrazino-3,4-dioxo-cyclobuten-1-yl)-3-methyl-thiourea 3.12

An acetonitrile suspension (10 mL) of **3.6** (0.400 g, 1.75 mmol) was stirred at RT, and *tert*-butyl carbazate (0.231 g, 1.75 mmol) added. The reaction mixture was allowed to stir overnight. The suspension was then heated to 80 °C for a few minutes to form a solution. The solvent was removed at reduced pressure and DCM added to the residue until no more precipitate formed. The solid was collected and washed with DCM to yield **3.12** as a white solid (0.383 g, 73 %). R_f = 0.5 (DCM/EtOAc 3:2); m.p. 190-191 °C; ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 10.66 (b, 1 H; N*H*), 9.70 (b, 1 H; N*H*), 9.41 (b, 1 H; N*H*), 3.04 (d, 3 H; C*H*$_3$-N), 1.42 (s, 9 H; (C*H*$_3$)$_3$C); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 182.50 (C=O), 181.52 (C=O), 178.27 (C), 170.31 (C), 157.43 (N-C=O), 80.47 ((CH$_3$)$_3$*C*), 32.45 (*C*H$_3$-N), 27.97 ((*C*H$_3$)$_3$C); IR (neat): ν = 3282, 3228, 3079, 3005, 1801, 1739, 1705, 1608, 1574, 1516 cm^{-1}; elemental analysis calcd for $C_{11}H_{16}N_4O_4S$ (300.3): C 43.99, H 5.37, N 18.65; found C 44.12, H 5.50, N 18.84.

1-Benzyl-3-(2-*tert*-butoxycarbonylhydrazino-3,4-dioxo-cyclobuten-1-yl)-thiourea 3.13

To a solution of **3.8** (0.490 g, 1.61 mmol) in acetonitrile (15 mL) was added *t*-butyl carbazate (1.060 g, 8.05 mmol) and the suspension allowed to stir at RT overnight. The suspension was then heated to boiling point until a solution formed. The solution was allowed to cool and the resulting solid precipitate collected and washed with acetonitrile to yield 0.360 g (59 %) of **3.13** as a white solid. R_f = 0.23 (DCM/THF 30:1); m.p. 194-195 °C (decomp.); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 10.47 (b, 1 H; N*H*), 9.85 (b, 1 H; N*H*), 9.34 (b, 1 H; N*H*), 8.86 (b, 1 H; N*H*), 7.08 (m, 5 H; Ar-C*H*), 4.66 (d, 2 H; Ph-CH_2), 1.29 (s, 9 H; C(CH_3)$_3$); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 182.39 (C=O), 180.65 (C=O), 177.91 (C), 169.8 (C), 136.09 (Ar-*C*), 128.00 (Ar-*C*), 127.25 (Ar-*C*), 126.99 (Ar-*C*), 80.94 (*C*(CH$_3$)$_3$), 49.47 (Ph-*C*H$_2$), 27.67 (C(*C*H$_3$)$_3$); IR (neat): ν = 3316, 3288, 3191, 3059, 2973, 2929, 1801, 1751, 1716, 1614, 1568, 1540 cm^{-1}; elemental analysis calcd for $C_{17}H_{20}N_4O_4S$ (376.4): C 54.24, H 5.36, N 14.88, O 17.00, S 8.52; found C 54.54, H 5.50, N 14.89.

1-(2-Ethoxycarbonylmethylamino-3,4-dioxo-cyclobuten-1-yl)-3-phenyl-thiourea 3.14

To a solution of **3.1** (0.446 g, 1.53 mmol) in THF (10 mL) was added glycine ethyl ester hydrochloride (0.360 g, 2.59 mmol) and DIPEA (0.45 mL, 2.59 mmol). The solution was left stirring at RT overnight, the solvent removed *in vacuo* and the residue purified by flash column chromatography (DCM/EtOAc 3:2) to afford the *title compound* **3.14** as a yellow solid (0.350 g, 69 %). R_f = 0.5 (DCM/EtOAc 2:1); m.p. 202 °C (decomp.); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 11.10 (b, 1 H; N*H*), 10.64 (b, 1 H; N*H*), 8.43 (t, *J* = 5.5 Hz, 1 H; CH$_2$-N*H*), 7.34 (d, *J* = 6.0 Hz, 2 H; Ar-C*H*), 7.00 (t, *J* = 6.0 Hz, 2 H; Ar-C*H*), 6.86 (t, *J* = 7.0 Hz, 1 H; Ar-C*H*), 4.15 (d, *J* = 6.0 Hz, 2 H; NH-CH_2), 3.88 (q, *J* = 7.0 Hz, 2 H; CH_2CH$_3$), 0.94 (t, *J* = 7.0 Hz, 3 H; CH$_2$CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 183.0 (C=O), 182.12 (C=O), 178.59 (*C*OOEt), 168.90 (C), 168. 25 (C), 158.22 (C=S), 61.09 (CH_2COOEt), 44.84 (*C*H$_2$CH$_3$), 32.17 (*C*H$_3$N), 13.62 (CH$_2$*C*H$_3$); IR (neat): ν = 3300, 3184, 3016, 2980, 1792, 1732, 1668, 1643, 1583, 1568, 1521 cm^{-1}; elemental analysis calcd for $C_{15}H_{15}N_3O_4S$ (333.4): C 54.04, H 4.54, N 12.60; found C 53.82, H 4.59, N 12.29.

1-Butyl-3-(2-ethoxycarbonylmethylamino-3,4-dioxo-cyclobuten-1-yl)-thiourea 3.15

To a solution of **3.4** (0.510 g, 1.88 mmol) in THF (10 mL) was added glycine ethyl ester hydrochloride (0.262 g, 1.88 mmol) and DIPEA (0.33 mL, 1.88 mmol). The solution was allowed to stir overnight at RT, the solvent removed *in vacuo* and the residue purified by flash column chromatography (DCM/EtOAc 3:2) to yield 0.500 g (85 %) of **3.15** as a white solid. R_f = 0.26 (DCM/EtOAc 10:1); m.p. 170-172 °C (decomp.); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 10.74 (b, 1 H; N*H*), 9.06 (b, 2 H; N*H*), 4.56 (d, *J* = 6.0 Hz, 2 H; C*H*$_2$-NH), 4.26 (q, *J* = 7.0 Hz, 2 H; C*H*$_2$), 3.62 (q, *J* = 6.0 Hz, 2 H; -$CO_2CH_2CH_3$), 1.64 (qt, *J* = 7.0 Hz, -C*H*$_2$-), 1.40 (m, *J* = 7.0 Hz, 2 H; -CH_2CH_3), 1.27 (t, *J* = 7.0 Hz, 3 H; -$CO_2CH_2CH_3$), 0.97 (t, *J* = 7.0 Hz, 3 H; -CH_2CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 183.13 (C=O), 177.53, 168.35 (*C*O_2Et), 159.51 (C=S), 62.13 (EtO_2CCH_2-), 45.54 ($CO_2CH_2CH_3$), 30.23 (*C*H$_2$), 20.06 (*C*H$_2$), 14.12 ($CO_2CH_2CH_3$), 13.65 (-CH_2CH_3); IR (neat): ν = 3298, 3088, 2962, 2932, 2871, 1805, 1740, 1709, 1699, 1589, 1547 cm^{-1}; elemental analysis calcd for $C_{13}H_{19}N_3O_4S$ (313.4): C 49.83, H 6.11, N 13.41; found C 49.77, H 6.05, N 13.50.

1-(2-Ethoxycarbonylmethylamino-3,4-dioxo-cyclobuten-1-yl)-3-methyl-thiourea 3.16

To a solution of **3.1** (0.450 g, 1.97 mmol) in THF (10 mL) was added glycine ethyl ester hydrochloride (0.275 g, 1.97 mmol) and DIPEA (0.34 mL, 1.97 mmol). The solution was allowed to stir at RT overnight, the solvent removed *in vacuo* and the residue purified by flash column chromatography (DCM/EtOAc 2:1) to yield 0.500 g (94 %) of **3.16** as a white solid. R_f = 0.5 (DCM/EtOAc 2:1); m.p. 199-200 °C (decomp.); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 10.29 (b, 1 H; N*H*), 9.21 (b, 1 H; N*H*), 8.15 (b, 1 H; CH_2-N*H*), 3.94 (d, *J* = 7.0 Hz, 2 H; C*H*$_2$-NH), 3.68 (q, *J* = 7.0 Hz, 2 H; -C*H*$_2$CH$_3$), 2.62 (s, 3 H; C*H*$_3$-N), 0.76 (t, *J* = 7 Hz, 3 H; -CH_2CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 183.00 (C=O), 182.18 (C=O), 178.59 (*C*OOEt), 168.90 (C), 168.25 (C), 158.66 (C=S), 61.09 (*C*H$_2$$CO_2$Et), 44.84 (-C*H*$_2CH_3$), 32.17 (*C*H$_3$-N), 13.62 (-$CH_2CH_3$); IR (neat): ν = 3285, 3192, 3077, 2980, 2942, 1806, 1749, 1708, 1621, 1570, 1547 cm^{-1}; elemental analysis calcd for $C_{10}H_{13}N_3O_4S$ (271.3): C 44.27, H 4.83, N 15.49; found C 44.51, H 5.10, N 15.71.

1-Benzyl-3-(2-ethoxycarbonylmethylamino-3,4-dioxo-cyclobuten-1-yl)-thiourea 3.17

To a solution of **3.8** (0.180 g, 0.59 mmol) in THF (10 mL) was added glycine ethyl ester hydrochloride (0.083 g, 0.59 mmol) and DIPEA (103 μL, 0.59 mmol). The solution was allowed to stir overnight at RT, the solvent removed *in vacuo* and the residue purified by flash column chromatography (DCM/THF 25:1) to yield **3.17** (0.160 g, 78 %) as a white solid. R_f = 0.6 (DCM/THF 24:1); m.p. 190-192 °C (decomp.); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 10.83 (s, 1 H; N*H*), 8.95 (br s, 2 H; -CH_2-N*H* and –NH), 7.30 (m, 5 H; Ar-C*H*), 4.82 (d, *J* = 6.0 Hz, 2 H; Ph-CH_2), 4.56 (d, *J* = 6.0 Hz, 2 H; CH_2COOEt), 4.25 (q, *J* = 7.0 Hz, 2 H; -COOC$H_2$$CH_3$), 1.31 (t, *J* = 7.0 Hz, 3 H; CH_2CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 184.82 (C=O), 183.03 (C=O), 177.82 (*C*O_2Et), 171.44 (C), 168.35 (C), 159.38 (C=S), 135.87 (Ar-*C*), 129.2 (Ar-*C*), 128.12 (Ar-*C*), 127.66 (Ar-*C*), 125.49 (Ar-*C*), 62.15 (NH-CH_2), 49.94 (Ph-CH_2), 45.52 (-CH_2CH_3), 14.11 (-CH_2CH_3); IR (neat): ν = 3303, 3321, 3060, 2985, 1811, 1745, 1707, 1626, 1586, 1542 cm^{-1}; elemental analysis calcd for $C_{16}H_{17}N_3O_4S$ (347.4): C 55.32, H 4.93, N 12.10; found C 55.08, H 4.90, N 11.93.

1,2,3,4-Tetrahydro-1,4-dimethyl-cyclobuta[b]pyrazine-5,6-dione 4.3

This procedure is modified from that of Hünig et al.[4.13]

Two solutions in two dropping funnels containing **2.5** (1.00 g, 5.04 mmol) and ethanol (40 mL) and *N,N'*-dimethylethylenediamine (0.54 mL, 5.04 mmol) in ethanol (40mL) were added simultaneously in a dropwise manner to refluxing ethanol (60 mL) over 4 h. The orange solution was concentrated *in vacuo* until a precipitate was observed. Ether (100 mL) was added to the mixture, and the mixture cooled on ice for 1 h. The solid was collected and dried to yield 0.560 g (67%) of a white solid. The solid was dissolved in chloroform and fractionally crystallised with isopropanol to yield 0.210 g (25 %) of the *title compound* **4.3** as white needles. M.p. 274-276 °C (decomp.) (literature[4.13] 235 °C (decomp.)); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 3.41 (s, 2 H; CH_2), 3.23 (s, 3 H; CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 180.44 (C=O), 167.47 (C), 47.28 (*C*H_2), 37.67 (*C*H_3); IR (neat): ν = 3255, 2933, 1792, 1655, 1549 cm^{-1}.

α-Methyl-γ-aminoguanidinium iodide 4.5

This procedure is modified from that of Kirsten and Smith.[4.11]

S-methylisothiosemicarbazide iodide **4.18** (2.33 g, 0.01 mol) was dissolved in water (6 mL) and methylamine (40 % in water, 1.03 g, 1.15 mL, 0.033 mol) added with stirring. The mixture was connected to a condenser which was attached to a trap and the solution refluxed for 30 min. After cooling, the solution was evaporated to dryness and ethanol (6 mL) added to the orange residue. The solution was heated, filtered, and allowed to cool. Ether was added until no more precipitate formed (*ca.* 12 mL). The solid product was collected by filtration, washed with anhydrous ether and dried in a desiccator to afford 1.75 g (81 %) of **4.5** as a white shiny solid. M.p. 117-119 °C (literature[4.11] 120-121°C); ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 8.49 (b, 1 H; N*H*), 7.38 (b, 1 H; N*H*), 7.20 (b, 2 H; NH_2), 4.63 (b, 2 H; NH_2), 2.72 (s, 3 H; CH_3); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 158.32 (C=N), 27.54 (*C*H$_3$); IR (neat): ν = 3327, 3297, 3258, 3139, 2930, 1647, 1580 cm^{-1}.

Benzylaminoguanidine hydrobromide 4.6

This procedure is modified from that of Heinisch.[4.12]

Benzylamine (1.65 mL, 0.015 mol) was added to a stirred solution of *s*-ethyl isothiosemicarbazide bromide **4.19** (2.70 g, 5.02 mmol) in water (10 mL). The resulting solution was refluxed for 40 min. The water was removed *in vacuo*, ethanol (10 mL) added, the mixture heated and hot-filtered. The solution was allowed to cool, ether added, and the resulting solid collected, washed with ether, and dried under vacuum to produce 0.763 g (62 % yield) of **4.6** as an off-white coloured solid. M.p. 122-125 °C; ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 8.69 (b, 1 H; N*H*), 8.25 (b, 1 H; N*H*), 7.92 (b, 1 H; N*H*), 7.40 (m, 5 H; Ar-C*H*), 4.74 (s, 1 H; N*H*), 4.42 (s, 2 H; Ph-CH_2), 4.05 (b, 1 H; N*H*); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 159.64 (C=N), 137.32 (Ar-*C*), 128.88 (Ar-*C*), 128.47 (Ar-*C*), 127.37 (Ar-*C*), 47.24 (*C*H$_2$); IR (neat): ν = 3365, 3265, 3212, 2992, 2992, 2884, 2675, 2560, 2338, 2017, 1662, 1633, 1589 cm^{-1}.

S-Methyl isothiourea sulfate 4.10

This procedure is modified from that of Shildneck and Windus.[4.3]

In a one litre round-bottomed flask were mixed thiourea (80.0 g, 1.05 mmol) and water (35 mL). Dimethyl sulphate (100 mL, 1.06 mol) was added and a reflux condenser which was connected to a trap immediately fitted to the flask. The suspension was stirred and heated gently until reaction commenced. Heat was removed until the initially vigorous reaction had subsided, then the mixture was refluxed at 150 °C for 1 h, during which time a solution formed, followed by the formation of a precipitate. The mixture was allowed to cool overnight. Ethanol (100 mL) was added to the flask, and the precipitate filtered and washed with ethanol (50 mL) to yield 104.5 g (79 % yield). The mother liquor was concentrated *in vacuo* and ethanol (60 mL) added to the paste after cooling to yield another 6.2 g of white solid to give a total yield of 110.7 g (80 %) of **4.10** as a white crystalline solid. M.p. 255 °C (decomp.) [literature[4.3] 244 °C (decomp.)]; IR (neat): ν = 3184, 3005, 2802, 1678, 1563 cm^{-1}.

2-Methyl-1-nitro-2-thiopseudourea 4.11

This procedure is modified from that of Fishbein and Gallaghan.[4.4]

2-methyl-2-thiopseudouronium sulphate **4.10** (20.00 g, 0.72 mol) was added portionwise over a period of 20 min to a nitrating mixture consisting of nitric acid (fuming, 20 mL) and sulphuric acid (99 %, 60 mL). The nitration was carried out at -10 °C while half the reacting material was being added, then the remainder was added at 0-5 °C, during which time a yellow solution had formed. The solution was cooled to 0 °C, then poured in two portions onto ice (850 g). The fine white precipitate was filtered with a glass fritted funnel, washed with water (150 mL) and air-dried. The white paste was recrystallised from water/ethanol (2:1) to yield 12.93 g (42 %) of **4.11** as a brilliant white solid. M.p. 172-175 °C (literature[4.4] 163-164 °C); ^{1}H NMR (400 MHz, [D_6]DMSO): δ = 9.06 (b, 2 H; N*H*), 2.32 (s, 3 H; CH_3); ^{13}C NMR (100.6 MHz, [D_6]DMSO): δ = 172.62 (C=N), 13.69 (*C*H_3); IR (neat): ν = 3373, 3277, 3099, 1740, 1626 cm^{-1}.

S-Methyl isothiouronium iodide 4.12

This procedure is modified from that of Clapp et al.[4.5]

Thiourea (1.52 g, 0.02 mol) and iodomethane (1.25 mL, 0.02 mol) were dissolved in ethanol (10 mL) and water (10 mL) and the colourless solution refluxed overnight. The resulting deep red solution was then concentrated *in vacuo*. Acetone (10 mL) was added to the orange residue and the resulting white solid collected and washed with acetone to yield the *title compound* **4.12** of 2.75 g (63 %). M.p. 126-129 °C; IR (neat): ν = 3305, 3257, 3158, 3094, 2910, 2683, 1634, 1524 cm^{-1}.

1-Butyl-3-nitroguanidine 4.13

This procedure is modified from that of Fishbein and Gallaghan.[4.4]

2-methyl-1-nitro-2-thiopseudourea **4.11** (11.93 g, 0.088 mol) was added to butylamine (9 mL, 0.166 mol) with stirring in a flask connected to a trap. The evolution of methyl mercaptan was instantaneous. The mixture was allowed to stir at RT overnight. The resulting paste was dissolved in hot ethanol (70 mL) to form a pink solution, and water (70 mL) added. The resulting precipitate was collected, washed with ethanol/water (1:1) and dried to yield 11.36 g (81 %) of **4.13** as a white crystalline material. M.p. 87-89 °C (literature[4.4] 89.5-90.5 °C); ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 8.43 (b, 1 H; N*H*), 7.82 (b, 1 H; N*H*), 3.06 (t, *J* = 6.5 Hz, 2 H; C*H*$_2$-N), 1.40 (m, *J* = 7.0 Hz, 2 H; C*H*$_2$), 1.24 (m, *J* = 7.5 Hz, 2 H; C*H*$_2$), 0.81 (m, *J* = 7.2 Hz, 3 H; C*H*$_3$); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 159.26 (C=N), 40.10 (*C*H$_2$-N), 30.21 (*C*H$_2$), 19.38 (*C*H$_2$), 13.55 (C*H*$_3$); IR (neat): ν = 3386, 3306, 3156, 2961, 2946, 2875, 1648, 1592, 1549 cm^{-1}.

1-Benzyl-3-nitroguanidine 4.14

This procedure is modified from that of Fishbein and Gallaghan.[4.4]

2-methyl-1-nitro-2-thiopseudourea **4.11** (10.86 g, 80.3 mmol) was added to benzylamine (10 mL, 0.091 mol) with stirring in a flask connected to a trap. Methyl mercaptan was evolved almost immediately. The paste formed was allowed to stir overnight. Ethanol/water (1:1, 500 mL) was added to the paste, and the suspension heated to boiling to form a solution. The solution was allowed to cool, and the white crystalline solid collected, washed with ethanol/water (1:1) to give 14.10 g of the *title compound* **4.14** (90 % yield). M.p. 187-189 °C; ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 8.80 (b, 1 H; N*H*), 7.85 (b, 1 H; N*H*), 7.20 (m. 5 H; Ar-C*H*), 4.29 (s, 2 H; Ph-C*H*$_2$); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 159.38 (C=N), 128.54 (Ar-*C*), 127.32 (Ar-*C*), 127.21 (Ar-*C*), 43.81 (Ph-*C*H$_2$); IR (neat): ν = 3372, 3288, 3158, 1651, 1588, 1576 cm^{-1}.

Butylaminoguanidine hydrogen carbonate 4.15

This procedure is modified from that of Shriner and Neumann.[4.6]

1-Butyl-3-nitroguanidine **4.13** (10.00 g, 0.068 mol) and Zn dust (22.50 g, 0.344 mol) were thoroughly ground together using a mortar and pestle, and water (20 mL) added to form a thick grey paste. The paste was transferred to a beaker in an icebath. A solution of glacial acetic acid (5 mL, 0.075 mol) in water (10 mL) was cooled in an icebath; and the ice-cooled paste added to the solution with stirring over a period of 8 h. with the temperature of the reaction mixture kept to 5-15 °C. Ice was added if the mixture became too hot or too thick to stir. The mixture was then heated to 40 °C for 5 min when reduction was shown to be complete using Mohr's Salt (ferrous ammonium sulfate. Red colouration indicated incompletion, green precipitate showed completion). The solution was immediately filtered and the resultant cake sucked as dry as possible. The residue was triturated with water (25 mL) and the residue filtered off, repeating the trituration process twice. Ammonium chloride (7 g) was dissolved in the mother liquor, then $NaHCO_3$ (7.50 g, 0.089 mol) added over a period of 10 min. The solution was then refrigerated overnight. The resulting precipitate was collected, triturated with ammonium chloride solution (5 %, 20 mL), washed with water, ethanol, ether, and air-dried to yield 0.70 g (8 %) of **4.15** as a white solid. M.p. >300 °C; IR (neat): ν =3475, 1650, 1536 cm^{-1}.

Thiosemicarbazide 4.17

This procedure is modified from that of Scott et al.[4.8]

Ammonium thiocyanate (100.0 g, 1.31 mol) was dissolved in a solution containing hydrazine hydrate (98 %, 25 mL, 0.76 mol) and water (25 mL), and the resulting solution refluxed under Ar for 3 h. The solution was cooled slightly and the coagulated sulphur removed by filtration. The filtrate was allowed to cool overnight. The precipitate formed was collected and recrystallised from 50 % ethanol (200 mL) to yield 12.2 g of **4.17** as a white crystalline solid. Additional refluxing and evaporation of the mother liquors yielded a further 2.2 g of product. Total yield 14.4 g (21 %). M.p. 185-189 °C (literature[4.8] 180-181 °C); ^{1}H NMR (400 MHz, [D_6]DMSO): δ = 8.59 (b, 1 H; N*H*), 7.50 (b, 1 H; N*H*), 7.18 (b, 1 H; N*H*), 4.47 (b, 2 H; NH_2); ^{13}C NMR (100.6 MHz, [D_6]DMSO): δ = 180.65 (C=S); IR (neat): ν = 3361, 3257, 3169, 1643, 1618, 1530 cm^{-1}.

S-Methyl isothiosemicarbazide hydroiodide 4.18

This procedure is modified from that of Emilsson.[4.9]

Thiosemicarbazide **4.17** (9.10 g, 0.10 mol) was treated with iodomethane (6.25 mL, 0.10 mol) in refluxing ethanol (10 mL) for 2 h, during which time a pale yellow solution formed. The solution was allowed to cool, the crystalline precipitate collected, washed with ethanol and recrystallised from ethanol to afford 17.50 g (75 %) of **4.18** as a white solid. M.p. 135-140 °C (literature[4.9] 140 °C); ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 10.43 (b, 1 H; N*H*), 8.89 (b, 2 H; N*H*$_2$), 5.20 (b, 1 H; N*H*), 2.56 (s, 3 H; C*H*$_3$); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 167.61 (C=N), 13.35 (*C*H$_3$); IR (neat): ν = 3337, 3292, 3230, 3195, 3135, 2998, 1644, 1608 cm^{-1}.

S-Ethyl isothiosemicarbazide hydrobromide 4.19

This procedure is modified from that of Beyer et al.[4.10]

Thiosemicarbazide **4.17** (9.10 g, 0.10 mol) was suspended in ethanol (250 mL) and ethyl bromide (17.5 mL, 0.11 mol) added to the stirred suspension, which was refluxed until a clear solution formed (*ca.* 2 h). The solution was allowed to cool overnight. The resulting suspension was concentrated *in vacuo* to *ca.* 50 mL, ether (50 mL) added, and the precipitate filtered and washed with ether to yield 1.71 g (85 %) of **4.19** as a white solid. M.p. 125-128 °C (literature[4.10] 119-120 °C); ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 10.66 (b, 1 H; N*H*), 9.17 (b, 1 H; N*H*), 5.20 (b, 1 H; N*H*), 3.17 (d, *J* = 6.5 Hz, 2 H; -C*H*$_2$CH$_3$), 1.25 (t, *J* = 6.5 Hz, 3 H; -CH$_2$C*H*$_3$); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 165.95 (C=N), 24.93 (-*C*H$_2$CH$_3$), 14.40 (CH$_2$*C*H$_3$); IR (neat): ν = 3297, 3248, 3197, 3109, 2978, 2932, 2870, 2673, 1646, 1617, 1548 cm^{-1}.

Thiocarbohydrazide 4.21

This procedure is modified from that of Audrieth et al. [7.3]

To a vigorously stirred solution of hydrazine hydrate (98 %, 60 mL, 1.24 mol) in water (40 mL) was added carbon disulfide (15 mL, 0.25 mol) in a dropwise fashion, and the resultant light green suspension stirred as strongly as possible until a temperature rise was observed. The mixture was then refluxed for 30 min and the molasses coloured solution cooled on ice for 30 min. The crystalline precipitate formed was filtered off, washed with ethanol, then ether and the solid air-dried. The mother liquor was then refluxed for a further 30 min and the solid formed collected by filtration and washed as before. This procedure was repeated twice more, until no further solid formed. The three samples (yields 16.5 g, 3.1 g, 0.5 g, respectively) were combined and recrystallised from water to yield 18.05 g (68 % yield) of **4.21** as a colourless crystalline solid. M.p. 165-167 °C (decomp.) (literature[7.3] 180-182 (decomp.)); ^{1}H NMR (400 MHz, [D_6]DMSO): δ = 7.63 (b, 3 H; N*H*), 6.51 (b, 1 H; N*H*), 6.15 (b,1 H; N*H*), 1.46 (s, 1 H; N*H*); ^{13}C NMR (100.6 MHz, [D_6]DMSO, TMS): δ = 181.49 (C=S); IR (neat): ν = 3361, 3261, 3173, 2964, 1641, 1619, 1530, 1486 cm^{-1}.

N-*t*-Butoxycarbonylglycine ethyl ester 4.22

This procedure is modified from that of Tarbell et al.[4.14]

Glycine ethyl ester hydrochloride (5.00 g, 36 mmol) was suspended in chloroform (50 mL) and $NaHCO_3$ (3.01 g, 36 mmol) in water (40 mL) added. Sodium chloride (6 g) was added with stirring, and then a solution of di-*tert*-butyl dicarbonate (7.82 g, 36 mmol) in chloroform was added. The resultant mixture was refluxed for 90 min. After cooling, the two-layer system was transferred to a separating funnel and the chloroform layer separated. The water layer was then extracted with chloroform (7 x 20 mL), the combined extracts plus the original chloroform layer dried ($MgSO_4$) to give **4.22** as an opaque oil. R_f = 0.4 (DCM/EtOAc 25:1); ^{1}H NMR (300 MHz, $CDCl_3$): δ = 5.15 (b, 1 H; N*H*), 4.15 (m, *J* = 9.5 Hz, 2 H; -C*H*$_2$CH$_3$), 3.94 (d, *J* = 7.0 Hz, 2 H; C*H*$_2$-N), 1.41 (s, 9 H; (C*H*$_3$)$_3$C), 1.24 (t, *J* = 9.4 Hz, 3 H; -CH$_2$C*H*$_3$); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 170.23 (C=O), 155.63 (C=O), 85.00, 79.71 (*C*(CH$_3$)$_3$), 61.11 (*C*H$_2$-N), 42.32 (*C*H$_2$), 28.17 (C(*C*H$_3$)$_3$), 27.25, 13.45 (-CH$_2$C*H*$_3$); IR (neat): ν = 3368, 2980, 2936, 1752, 1697, 1513 cm^{-1}.

Acetylglycine ethyl ester 4.23

This procedure is modified from that of Nandi and Robinson.[4.17]

To a –10 °C cooled solution of glycine ethyl ester hydrochloride (34.90 g, 0.25 mol) in aqueous NaOH (2 M, 100 mL, 0.20 mol) was slowly added acetic anhydride (23.5 mL, 0.25 mol) and aqueous NaOH (2 M, 100 mL, 0.20 mol) with stirring, and the resultant solution was allowed to react for 1 h. The solution was then adjusted to pH 7 with solid Na_2CO_3, and the product extracted with ethyl acetate. The solvent was removed under reduced pressure until a solid crystallised. The residue was recrystallised from ether to yield 20.00 g (55 % yield) of **4.23** as a white crystalline solid. M.p. 48-56 °C (literature[4.17] 48 °C); ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 8.25 (b, 1 H; N*H*), 4.10 (q, *J* = 7.5 Hz, 2 H; C*H*$_2$CH$_3$), 3.80 (d, *J* = 6.0 Hz, 2 H; C*H*$_2$NH), 1.87 (s, 3 H; COC*H*$_3$), 1.20 (t, *J* = 7.5 Hz, 3 H; CH$_2$C*H*$_3$); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 169.88 (CO), 169.59 (CO), 60.21 (*C*H$_2$NH), 40.58 (*C*H$_2$CH$_3$), 22.12 (CO*C*H$_3$), 13.97 (CH$_2$*C*H$_3$); IR (neat): ν = 3263, 3084, 2989, 2945, 1738, 1640, 1557 cm^{-1}.

Acetylglycine 4.24

This procedure is modified from that Herbst and Shemin.[4.15]

To a vigorously stirring solution of glycine (37.5 g, 0.5 mol) in water (150 mL) was added acetic anhydride (108 g, 100 mL, 1 mol) in one portion. During 20 min of stirred reaction at RT, the solution became hot and solid began to precipitate. The process of precipitation was completed overnight in a refrigerator. The solid was collected, washed with ice-cold water and dried at 110 °C. The mother liquor was evaporated *in vacuo* to yield a second fraction of white solid, which was recrystallised from water. The total yield of white crystalline acetylglycine **4.24** was 45 g (62 % yield). M.p. 212-223 °C (literature[4.15]. 207-209 °C); ^{1}H NMR: (400 MHz, [D$_6$]DMSO): δ = 12.03 (b, 1 H; COO*H*), 7.70 (b, 1 H; N*H*), 3.30 (d, *J* = 5.5 Hz, 2 H; C*H*$_2$), 1.43 (s, 3 H; C*H*$_3$); ^{13}C NMR: (100.6 MHz, [D$_6$]DMSO): δ = 171.42 (*C*OOH), 169.73 (*C*OCH$_3$), 40.68 (*C*H$_2$), 22.30 (*C*H$_3$); IR (neat): ν = 3349, 2941, 2437, 2260, 1937, 1717, 1581, 1547 cm^{-1}.

3,4-Diethoxy-3-cyclobutene-1,2-dione 4.26

This procedure is modified from that of Ohno et al.[4.18]

Squaric acid **2.4** (3.00 g, 0.026 mol) was added to a solution of *p*-toluenesuphonic acid (0.50 g, 2.6 mmol) in ethanol (10 mL) and benzene (20 mL) over molecular sieves (4Å) and refluxed in a Dean and Stark apparatus overnight. A further addition of ethanol (30 mL) was then added to the mixture and the mixture refluxed for a further 3 h. The mixture was then allowed to cool, and the precipitate removed with filtration. The solvent was removed *in vacuo*. Ethyl acetate (50 mL) was added to the residue and the solution washed with NaOH (2 M, 10 mL). The EtOAc layer was then washed with water (4 X 25 mL), dried ($MgSO_4$) and the solvent evaporated, and the residue purified by distillation to yield the *title compound* **4.26** as a clear pale yellow oil (1.95 g, 45 % yield). R_f = 0.2 (DCM); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.73 (q, *J* = 7.0 Hz, 2 H; O-C*H*$_2$CH$_3$), 1.48 (t, *J* = 7.0 Hz, 3 H; O-CH$_2$C*H*$_3$); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 188.74 (C$_1$, C$_2$), 182.75 (C$_3$, C$_4$), 70.02 (-O*C*H$_2$CH$_3$), 15.10 (-OCH$_2$*C*H$_3$).

4-Dibenzylamino-3-ethoxy-3-cyclobutene-1,2-dione 4.27

This procedure is modified from that of Thorpe.[4.19]

Dibenzylamine (0.56 mL, 2.93 mmol) was added to a stirring ethanolic solution (5 mL) of **4.26** (0.500 g, 2.93 mmol). The solution was continuously stirred overnight at RT, and the residuals removed *in vacuo*. The solid obtained was recrystallised from petroleum ether (40-60) to afford **4.27** as a white waxy solid (0.825 g, 99 %). R_f = 0.26 (DCM/EtOAc 49:1); m.p. 110-112 °C (literature[4.19] 107-108 °C); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.34 (m, 8 H; Ar-C*H*), 7.23 (m, 2 H; Ar-C*H*), 4.82 (q, *J* = 7.0 Hz, 2 H; O-CH_2CH$_3$), 4.77 (s, 2 H; Ph-CH_2), 4.47 (s, 2 H; Ph-CH_2), 1.46 (t, *J* = 7.0 Hz, 3 H; -OCH$_2$CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 188.32 (C=O), 182.27 (C=O), 175.96 (C), 171.54 (C), 134.34 (Ar-*C*), 128.64 (Ar-*C*), 128.52 (Ar-*C*), 127.96 (Ar-*C*), 69.49 (-O*C*H$_2$CH$_3$), 51.04 (Ph-*C*H$_2$), 50.48 (Ph-*C*H$_2$), 15.39 (-OCH$_2$*C*H$_3$); IR (neat): ν = 3027, 2980, 2932, 1801, 1699, 1610, 1582 cm^{-1}; elemental analysis calcd for $C_{20}H_{19}NO_3$ (321.4): C 74.75, H 5.96, N 4.36; found C 74.76, H 6.00, N 4.34.

2-Dibenzylamino-3-ethoxy-4-thioxocyclobut-2-enone 4.28

This procedure is modified from that of Müller et al.[4.2]

4.27 (0.200 g, 0.71 mmol) was dissolved in DCM (5 mL) and Lawesson's Reagent (0.142 g, 0.35 mmol) added. The reaction mixture was stirred at RT overnight, then subjected to flash column chromatography (3:2 DCM/Hexane) to yield the *title compound* **4.28** as a yellow/orange solid (0.163 g, 68 % yield). R_f = 0.1 (DCM/Hexane 3:2); m.p. 85 °C; ^{1}H NMR (300 MHz, $CDCl_3$): δ = 7.39 (m, 6 H; Ar-C*H*), 7.26 (m, 4 H; Ar-C*H*), 5.26 (q, *J* = 9.4 Hz, 2 H; -CH_2CH$_3$), 4.72 (s, 2 H; Ph-CH_2), 4.48 (s, 2 H; Ph-CH_2), 1.47 (t, *J* = 9.4 Hz, 3 H; -CH$_2$CH_3); ^{13}C NMR (75.5 MHz, $CDCl_3$): δ = 206.44 (C=S), 184.07 (C=O), 179.74 (C), 174.45 (C), 135.58 (Ar-*C*), 133.53 (Ar-*C*), 129.06 (Ar-*C*), 128.92 (Ar-*C*), 128.74 (Ar-*C*), 128.61 (Ar-*C*), 128.56 (Ar-*C*), 128.17 (Ar-*C*), 69.95 (-*C*H$_2$CH$_3$), 51.88 (Ph-*C*H$_2$), 51.57 (Ph-*C*H$_2$), 16.10 (-CH$_2$*C*H$_3$); IR (neat): ν = 2925, 1765, 1646, 1602, 1583 cm^{-1}; elemental analysis calcd for $C_{20}H_{19}NO_2S$ (337.43): C 71.17, H 5.68, N 4.15; found C 71.18, H 5.72, N 4.13.

4-Dibenzylamino-3-isopropoxy-3-cyclobutene-1,2-dione 4.29

This procedure is modified from that of Thorpe.[4.19]

To a suspension of **2.5** (0.680 g, 3.43 mmol) in absolute alcohol (5 mL) was added dibenzylamine (0.65 mL, 3.43 mmol). The solution was allowed to stir overnight at RT, and the solvents were then removed *in vacuo*. The solid obtained was subjected to recrystallisation from petroleum ether (40-60) to yield **4.29** as a white waxy solid (0.960 g, 83 %). R_f = 0.34 (DCM); m.p. 119-120 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.37 (m, *J* = 7.0 Hz, 6 H; Ar-C*H*), 7.23 (dd, *J* = 21.0, 7.0 Hz, 4 H; Ar-C*H*), 5.49 (m, *J* = 6.0 Hz, 1 H; C*H*-O), 4.79 (s, 2 H; Ph-CH_2), 4.48 (s, 2 H; Ph-CH_2), 1.44 (d, *J* = 6.0 Hz, CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 188.78 (C=O), 182.55 (C=O), 176.26 (C), 172.23 (C), 134.96 (Ar-*C*), 129.07 (Ar-*C*), 128.96 (Ar-*C*), 128.81 (Ar-*C*), 128.41 (Ar-*C*), 128.21 (Ar-*C*), 51.41 (Ph-*C*H_2), 51.04 (Ph-*C*H_2), 22.92 (*C*H_3); IR (neat): ν = 3027, 2979, 2931, 2871, 1801, 1704, 1607, 1582 cm^{-1}.

2-Dibenzylamino-3-isopropoxy-4-thioxo-cyclobut-2-enone 4.30

This procedure is modified from that of Müller et al.[4.2]

A solution of **4.29** (0.267 g, 0.795 mmol) in DCM (5 mL) was stirred at RT, and Lawesson's Reagent (0.161 g, 0.398 mmol) added. The resulting solution was stirred at RT overnight, and subjected to column chromatography (DCM/Hexane 3:2) to afford 0.175 g (63 %) of **4.30** as an orange solid. R_f = 0.13 (DCM/Hexane 3:2); m.p. 103 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.50 (m, 5 H; Ar-C*H*), 7.27 (m, 5 H; Ar-C*H*), 6.44 (m, *J* = 6.0 Hz, 1 H; C*H*-O), 4.77 (s, 2 H; Ph-CH_2), 4.51 (s, 2 H; Ph-CH_2), 1.48 (d, *J* = 6.0 Hz, CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 206.42 (C=S), 184.39 (C=O), 179.65 (C), 174.86 (C), 133.86 (Ar-*C*), 133.77 (Ar-*C*), 130.11 (Ar-*C*), 129.26 (Ar-*C*), 129.12 (Ar-*C*), 128.99 (Ar-*C*), 128.79 (Ar-*C*), 128.34 (Ar-*C*), 80.08 (*C*H), 52.10 (Ph-*C*H_2), 51.57 (Ph-*C*H_2), 23.30 (*C*H_3); IR (neat): ν = 3063, 2984, 2925, 2867, 1884, 1768, 1703, 1646, 1602, 1584 cm^{-1}; elemental analysis calcd for $C_{21}H_{21}NO_2S$ (351.5): C 71.77, H 6.02, N 3.99; found C 71.46, H 6.09, N 3.95, S 9.47.

2,3-Diisopropoxy-4-thioxo-cyclobut-2-enone 4.31

This procedure is modified from that of Müller et al.[4.2]

Lawesson's Reagent (0.306 g, 0.76 mmol) was added to a stirring solution of **2.5** (0.300 g, 1.51 mmol) in DCM (10 mL) and the reaction mixture allowed to stir at RT for 5 d. The residuals were removed at reduced pressure and the remaining mixture subjected to flash chromatography (DCM/Hexane 3:2) to afford **4.31** as an orange/pink solid (0.200 g, 62 %). R_f = 0.17 (DCM/Hexane 3:2); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.74 (m, J = 6.0 Hz, 1 H; C*H*), 1.48 (m, 12 H; CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 222.55 (C=S), 187.37 (C=O), 184.92 (C), 184.22 (C), 80.23 (*C*H), 78.64 (*C*H), 22.83 (*C*H_3), 22.74 (*C*H_3); elemental analysis calcd for $C_{10}H_{14}O_3S$ (214.3): C 56.05, H 6.59; found C 56.24, H 6.65.

3,4-Diisopropoxy-3-cyclobutene-1,2-dithione 4.32

This procedure is modified from that of Müller et al.[4.2]

To a stirring solution of **2.5** (0.300 g, 1.51 mmol) in DCM was added Lawesson's Reagent (0.612 g, 1.51 mmol). The solution was stirred at RT for 7 d. The solvents were removed under vacuum. Purification by column chromatography (DCM/Hexane 3:2) yielded 0.225 g (65 %) of **4.32** as an orange solid. R_f = 0.37 (DCM/Hexane 3:2); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.94 (m, J = 6.5 Hz, 1 H; C*H*- O), 1.51 (d, J = 6.5 Hz, 6 H; CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 218.08 (C=S), 185.85 (C), 79.46 (*C*H), 22.88 (*C*H_3).

2,5-Dibromo-3,4-dinitrothiophene 5.1

This procedure is modified from that of Easton et al.[5.1]

A nitrating solution consisting of concentrated sulphuric acid (98 %, 110 mL), fuming sulphuric acid (82 mL) and fuming nitric acid (85 mL) was cooled on ice to 0 °C, then 2,5-dibromothiophene (45.7 mL, 100 g, 0.413 mol) was added dropwise to the solution, which was maintained at a reaction temperature of under 30 °C by cooling in an icebath. The mixture was left stirring for a further 30 min, then poured onto ice, and the yellow precipitate collected by filtration and washed with water until the washings were no longer yellow nor acidic. The solid was recrystallised from methanol to yield **5.1** as a yellow powder (67.55 g, 50 % yield). R_f = 0.33 (1:1 DCM/Hexane); m.p. 139-141 °C (literature[5.1] 134-135 °C); ^{13}C NMR (100.6 MHz, [D_6]DMSO, TMS): δ = 139.62 (*C*-NO_2), 116.63 (C-Br); IR (neat): ν = 2852, 1538 cm^{-1}.

3,4-Diaminothiophene hexachlorostannate chloride ammonium salt 5.2

This procedure is modified from that of Outurquin and Paulmier.[5.2]

2,5-dibromo-3,4-dinitrothiophene **5.1** (64.5 g, 0.194 mol) was suspended in concentrated HCl (900 mL), and Sn granules (122 g, 1.028 mol) added portionwise, while maintaining a reaction temperature of less than 30 °C. The mixture was stirred for a further 4 h, then stored overnight in the refrigerator. The brown solid was collected by filtration and washed several times with ether until the washings were clear and the solid no longer appeared wet, then air-dried to yield **5.2** as a mixture of brown hexachlorostannate and hydrochlorate salts of weight 31.20 g. R_f = 0.5 (EtOAc); IR (neat): ν = 3089, 2782, 2565, 1643, 1574, 1515 cm^{-1}.

1-Acetyl-3-(4-amino-thiophen-3-yl)thiourea 5.3

This procedure is modified from that of Dunne.[1.71]

The hexachlorostannate salt **5.2** (3.0 g) was suspended in dry DCM (60 mL) at 0 °C, and triethylamine (4.0 mL, 29 mmol) was added dropwise with stirring. The mixture was allowed to warm to room temperature with a calcium chloride drying tube fitted to the reaction flask. Acetyl isothiocyanate (1 mL, 11.38 mmol) in dry DCM (250 mL) was added dropwise during 2 h, and the reaction mixture stirred for 3 d. The residuals were then removed *in vacuo*. The residue was suspended in THF (50 mL), and the solids filtered and washed with THF. The filtrate was concentrated, and the residue purified by flash column chromatography (DCM/EtOAc 9:1) to yield **5.3** as a light brown solid (0.980 g, 40 % yield from **5.4**). R_f = 0.22 (DCM/EtOAc 9:1); m.p. 180-181 °C; IR (neat): υ = 3251, 2977, 1721, 1688, 1521 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$, TMS): δ = 12.34 (b, 1 H; CON*H*), 9.93 (b, 1 H; CSN*H*), 7.78 (d, J = 3.2 Hz, 1 H; Ar-C*H*), 6.36 (d, J = 3.2 Hz, 1 H; Ar-C*H*), 3.59 (b, 2 H; NH_2), 2.20 (s, 3 H; $COCH_3$); ^{13}C NMR (100.6 MHz, $CDCl_3$, TMS): δ = 177.66 (NH*C*S), 171.64 (NH*C*O), 138.78 (Ar-*C*-NH_2), 128.55 (Ar-*C*), 116.08 (Ar-*C*), 103.28 (Ar-*C*), 24.17 (CH_3).

Acetyl Isothiocyanate 5.4

This procedure is modified from that of Takimizara et al.[5.3]

Potassium thiocyanate (21.0 g, 0.216 mol) was dissolved in dry acetone (200 mL) and acetyl chloride (15 mL, 0.211 mol) added with a syringe. The mixture was allowed to stir at RT for 20 min. The KCl was removed by filtration, and washed with ether. The filtrate was concentrated to a volume of *ca.* 30 mL under reduced pressure at 20 °C. The solution was then distilled under vacuum to yield acetyl isothiocyanate **5.4** (b.p. 50-55 °C/20 Torr) as a yellow lachrymatic liquid (13.655 g, 65 % yield). R_f = 0.43 (Hexane/EtOAc); ^{1}H NMR (400 MHz, $CDCl_3$, TMS): δ = 2.34 (s, 3 H; C*H*$_3$); ^{13}C NMR (100.6 MHz, $CDCl_3$, TMS): δ = 165.44 (C=S), 147.63 (C=O), 27.69 (*C*H$_3$); IR (neat): 1942, 1717 cm^{-1}.

4-[(3-Acetyl-thioureido)-thiophen-3-yl]carbamic acid ^{t}butyl ester 5.5

This procedure is modified from that of Dunne.[1.71]

To a suspension of the hexachlorostannate salt **5.2** (3.0 g) in dry DCM (60 mL) was added triethylamine (4.0 mL, 2.9 mmol) with stirring at 0 °C. The reaction vessel was fitted with a calcium chloride drying tube, and the mixture allowed to warm to RT during 20 min. A solution of acetyl isothiocyanate (1 mL, 1.15 g, 11.38 mmol) in dry DCM (250 mL) was added dropwise to the mixture during 2 h, and the mixture stirred at RT for 3 d. The mixture was reduced *in vacuo,* dry DCM (60 mL) added and di-*tert*-butyl dicarbonate (3.6 g, 16.5 mmol) added to the suspension. The suspension was then refluxed for 18 h under a nitrogen atmosphere. The residuals were removed *in vacuo* and the remaining substance purified by flash column chromatography (24:1 DCM/EtOAc) to yield **5.5** as a yellow solid (1.35 g, 38 % from **5.4**). R_f = 0.2 (24:1 DCM/EtOAc); m.p. 153-154 °C (Ref. 154-156 °C); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 12.03 (b, 1 H; N*H*), 9.28 (b, 1 H; N*H*), 7.60 (s, 1 H; Ar-C*H*), 7.35 (s, 1 H; Ar-C*H*), 6.62 (b, 1 H; N*H*), 2.19 (s, 3 H; C*H*$_3$CO), 1.51 (s, 9 H; (CH$_3$)$_3$C); ^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 178.69 (C=O), 170.75 (C=O), 130.23 (Ar-*C*), 116.92 (Ar-*C*), 112.19 (Ar-*C*), 78.27 (C(CH$_3$)$_3$), 27.83 ((*C*H$_3$)$_3$C), 23.84 (*C*H$_3$); IR (neat): ν = 3193, 2979, 1688, 1531 cm^{-1}; MALDI-TOF-MS [M$^+$ + Na – H] 337.28: calcd 337.05; and also 3,4-di(*t*-butoxycarbonylamino)thiophene **5.6**. R_f = 0.64 (49:1 DCM/EtOAc); m.p. 158-160 °C (literature[7.4] 157-158 °C); ^{1}H NMR (400 MHz, $CDCl_3$, TMS): δ = 7.15 (b, 1 H, N*H*), 6.91(b, 1 H, N*H*), 1.51 (s, 18 H; C*H*$_3$); ^{13}C NMR (100.6 MHz, $CDCl_3$, TMS): δ = 153.71 (C=O), 130.09 (Ar-*C*), 112.93 (Ar-*C*), 81.73 (*C*(CH$_3$)$_3$), 28.23 (C(*C*H$_3$)$_3$); IR (neat): ν = 3377, 3158, 2990, 1726, 1676, 1573, 1506 cm^{-1}.

2-Acetylaminothieno[3,2-*d*]thiazol-6-yl)-carbamic acid ^{t}butyl ester 5.7

This procedure is modified from that of Dunne.[1.71]

5.5 (0.775 g, 2.46 mmol) was dissolved in chloroform (40 mL) and the stirring solution treated with *N*-iodosuccinimide (1.11 g, 4.93 mmol). The mixture was stirred at RT overnight, then the solution washed with water (3 X 20 mL), and the aqueous layers combined and back-extracted with chloroform (2 X 15 mL). The combined organic layers were then washed with brine (20 mL). The chloroform solution was dried over $MgSO_4$, the solvents removed under reduced pressure and the residue purified by flash column chromatography (DCM/EtOAc 4:1) to yield 0.410 g (53 %) of the *title compound* **5.7** as a yellow solid. R_f = 0.31 (DCM/EtOAc 9:1); m.p. 180-181 °C; ^{1}H NMR (400 MHz, [D_6]DMSO, TMS): δ = 8.85 (b, 1 H; N*H*), 7.31 (b, 1 H; Ar-C*H*), 2.17 (s, 3 H; CH_3), 1.44 (s, 9 H; (CH_3)$_3$C); ^{13}C NMR (100.6 MHz, [D_6]DMSO, TMS): δ = 169.02 (CO-NH), 160.11 (Ar-*C*), 153.11 (*C*OO), 148.01 (C=N), 126.64 (Ar-*C*), 112.92 (Ar-*C*), 79.27 (*C*(CH$_3$)$_3$), 28.06 (C(*C*H$_3$)$_3$), 22.50 (*C*H$_3$); IR (neat): ν = 3256, 3070, 2974, 1719, 1667, 1541 cm^{-1}.

2-Acetylamino-5-iodo-thieno[3,2-*d*]thiazol-6-yl)-carbamic acid ^{t}butyl ester 5.8

This procedure is modified from that of Dunne.[1.71]

To a solution of **5.5** (0.935 g, 2.96 mmol) in chloroform (50 mL) was added *N*-iodosuccinimide (2.28 g, 0.01 mol) with stirring. The reactants were stirred at RT overnight. The solution was then washed with saturated aqueous sodium thiosulphate solution (20 mL), water (3 X 20 mL), and the combined aqueous layers back-extracted with chloroform (2 X 15 mL). The combined chloroform layers were then washed with brine (20 mL), dried ($MgSO_4$) and the solvents removed *in vacuo*. The residue was subjected to flash column chromatography (DCM/EtOAc 4:1) to yield **5.8** as a yellow solid (1.050 g, 81 % yield). R_f = 0.2 (DCM/EtOAc 4:1); m.p. 195-201 °C (literature 180-190 °C); ^{1}H NMR (400 MHz, [D_6]DMSO, TMS): δ = 12.31 (s, 1 H; N*H*), 8.65 (b, 1 H; N*H*), 2.15 (s, 3 H; CH_3), 1.42 (s, 9 H; (CH_3)$_3$C); ^{13}C NMR (100.6 MHz, [D_6]DMSO, TMS): δ = 179.37 (C=O), 168.94 (Ar-*C*), 160.85 (C=O), 149.62 (Ar-*C*), 131.68 (Ar-*C*), 126.10 (Ar-*C*), 78.81 (Ar-*C*), 126.10 (Ar-*C*), 78.81 (*C*(CH$_3$)$_3$), 67.03 (Ar*C*-I), 28.13 (C(*C*H$_3$)$_3$), 22.46 (*C*H$_3$); IR (neat): ν = 3359, 3161, 3066, 2990, 1702, 1634, 1562, 1524 cm^{-1}; MALDI-TOF-MS [M^{+} + K – H] 477.38: calcd 476.69.

Sodium Ethylformylacetate 5.11

This procedure is modified from that of Robins et al.[5.5]

Freshly cut sodium metal (1.30 g, 0.058 mol) was added to dry toluene (15 mL), and the suspension refluxed with vigorous stirring, until the sodium had melted, and was finely divided in the solvent. The suspension was quickly cooled in an ice-bath with continuous stirring to produce the sodium sand. The suspension was then allowed to reach RT, and a solution of ethyl formate (4.6 mL, 0.058 mol) in ethyl acetate (10 mL, 0.106 mol) added with stirring. A calcium chloride tube was fitted to the reaction flask, and the suspension was stirred at RT for 2 d to yield the ethyl formylacetate **5.11** as a cream suspension. It was used without further purification.

(Z)-3-{[4-(3-Acetyl-thioureido)-thiophen-3-yl]-amino}-propenoic acid ethyl ester 5.12

This procedure is modified from that of Dunne.[1.71]

The crude suspension of sodium ethyl formylacetate **5.11** in toluene (6 mL) was added to dry ethanol (5 mL). Acetic acid (4 mL) was added, and a solution formed. The solution was added by syringe to acetylthiourea **5.3** (0.181 g, 84 mmol) in dry ethanol (10 mL), and the resulting mixture refluxed under a nitrogen atmosphere for 30 min. The red solution was allowed to cool overnight. The solvents were removed *in vacuo* and acetic acid residues removed by the addition and subsequent evaporation of toluene (1 mL) several times. The residue was suspended in DCM and the solids removed by filtration. The solvent was removed *in vacuo* and the residual compound purified by flash column chromatography (DCM/EtOAc 24:1) to yield 70 mg (40 %) of **5.12** as a yellow solid. R_f = 0.6 (DCM/EtOAc 4:1); ^{1}H NMR (400 MHz, $CDCl_3$, TMS): δ = 12.12 (br s, 1 H; CON*H*CS), 9.51 (d, *J* = 12.0 Hz, CHN*H*), 9.09 (br s, 1 H; N*H*), 7.82 (d, *J* = 3.5 Hz, 1 H; Ar-C*H*), 6.94 (dd, *J* = 12.0, 8.5 Hz, C*H*NH), 6.72 (d, *J* = 3.5 Hz, 1 H; Ar-C*H*), 4.86 (d, *J* = 8.5 Hz, 1 H; C*H*COOEt), 4.18 (q, *J* = 7.0 Hz, 2 H; -CH_2CH$_3$), 2.23 (s, 3 H, COCH_3), 1.30 (t, *J* = 7.0 Hz, 3 H; -CH$_2$CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$, TMS): δ = 179.21 (NH*C*SNH), 171.56 (NH*C*O), 170.23 (NH*C*O), 145.83 (*C*HNH), 136.25 (Ar-*C*), 128.21 (Ar-*C*), 119.15 (Ar-*C*), 106.79 (Ar-*C*), 89.19 (*C*HCOOEt), 59.38 (*C*H$_2$CH$_3$), 24.77 (-CH$_2$*C*H$_3$).

{2-[*N*-Acetyl-*N*-(2,4-dimethoxybenzyl)-amino]-5-iodo-thieno[3,2-*d*]thiazol-6-yl}-carbamic acid ^{t}butyl ester 5.16

This procedure is modified from that of Dunne.[1.72]

To a stirring solution of 5-iodothienothiazole **5.8** (0.450 g, 1.02 mmol), 2,4-dimethoxybenzyl alcohol (0.200 g, 1.19 mmol) and triphenylphosphine (0.400 g, 1.525 mmol) in dry THF (20 mL) was added diisopropyl azodicarboxylate (0.40 mL, 2.03 mmol) in a dropwise manner under an atmosphere of nitrogen at RT. The reaction was continued overnight. The residuals were removed under reduced pressure, and the residue subject to flash column purification (DCM/EtOAc 24:1) to yield the *title compound* **5.16** as a white solid (0.217 g, 36 % yield). R_f = 0.41 (DCM/EtOAc 24:1); m.p. 158-159 °C; ^{1}H NMR (400 MHz, $CDCl_3$, TMS): δ = 6.94 (d, *J* = 8.0 Hz, 1 H; Ar-C*H*), 6.47 (s, 1 H; Ar-C*H*), 6.35 (dd, *J* = 8.5, 2.5 Hz, 1 H; Ar-C*H*), 6.12 (b, 1 H; N*H*), 5.42 (s, 2 H; Ph-CH_2), 3.84 (s, 3 H; CH_3-O), 3.77 (s, 3 H; CH_3-O), 2.35 (s, 3 H; CH_3-CO), 1.44 (s, 9 H; C(CH_3)$_3$); ^{13}C NMR (100.6 MHz, $CDCl_3$, TMS): δ = 171.62 (NCOCH$_3$), 160.35 (thiazolo-*C*2), 158.51 (NH*C*OO), 131.89 (Ar-*C*), 128.92 (Ar-*C*), 116.44 (Ar-*C*), 104.11 (Ar-*C*), 98.42 (Ar-*C*I), 80.92 (*C*(CH$_3$)$_3$), 55.36 (*C*H$_3$-O), 55.31 (*C*H$_3$-O), 46.80 (*C*H$_2$), 28.10 (C(*C*H$_3$)$_3$), 22.57 (*C*H$_3$); IR (neat): ν = 3298, 2971, 2936, 2837, 1707, 1655, 1612, 1587, 1530, 1508 cm^{-1}; elemental analysis[†] calcd for $C_{21}H_{24}IN_3O_5S_2$ (589.5): C 42.79, H 4.10, N 7.13; found C 43.76, H 4.04, N 6.87.

[†] Spurious results are not uncommon with these elemental analyses, due to a problem with the calibration of the equipment involved. The situation was under review at the time of publication.

(E)-4-({(2-[*N*-Acetyl-*N*-(2,4-dimethoxybenzyl)-amino]-5-iodo-thieno[3,2-*d*]thiazol-6-yl}-(*t*butoxycarbonyl)-amino)-prop-2-enoic acid ethyl ester 5.17

DMF (1 mL) was added to a nitrogen-flushed reaction vessel containing 2,4-dimethoxybenzyl thienothiazole **5.16** (83 mg, 0.14 mmol) and sodium hydride (60 % in mineral oil, 20 mg, 0.5 mmol), and the reaction stirred under nitrogen for 30 min. Ethyl *cis*-3-iodoacrylate (45 μL, 0.35 mmol) was added *via* a syringe, and the brown suspension stirred under nitrogen overnight. The DMF was removed under reduced pressure at 80 °C, with residual DMF removed by repeated addition and evaporation of toluene (1 mL). The residue was then purified by flash column chromatography (initially with DCM, then DCM/EtOAc 49:1) to yield 20 mg (21 %) of **5.17** as a yellow solid. R_f = 0.35 (25:1 DCM/EtOAc); ^{1}H NMR (400 MHz, $CDCl_3$, TMS): δ = 8.36 (d, *J* = 14.2 Hz, 1 H; N*H*CH), 6.92 (b, 1 H; Ar-C*H*), 6.42 (d, *J* = 2.2 Hz, 1 H; Ar-C*H*), 6.30 (dd, *J* = 8.2, 2.5 Hz, 1 H; Ar-C*H*), 5.45 (d, *J* = 16.1 Hz, 1 H; C*H*^{a}H^b-Ph), 4.68 (d, *J* = 16.1 Hz, 1 H; CH^a*H*^{b}-Ph), 4.68 (d, J = 14.2 Hz, 1 H; C*H*COOEt), 4.15 (m, 2 H, -C*H*$_2$CH$_3$), 3.85 (s, 3 H; OC*H*$_3$), 3.74 (s, 3 H; OC*H*$_3$), 2.39 (s, 3 H, COC*H*$_3$), 1.26 (partially obscured s and t, *J* = 7.3 Hz, 9 H; C*H*$^a_3$C(CH$^b_3$) and CH$^a_3$C(C*H*$^b_3$)), 0.95 (t, 3 H; -CH$_2$C*H*$_3$); ^{13}C NMR (100.6 MHz, $CDCl_3$, TMS): δ = 171.28 (CH$_3$*C*O), 167.37 (*C*OOEt), 160.78 (thiazolo-*C*2), 147.94 (N*C*OO), 146.82 (*C*HN), 136.79 (Ar-*C*), 123.12 (Ar-*C*), 121.50 (Ar-*C*), 116.18 (Ar-*C*), 98.31 (Ar-*C*), 85.18 (*C*I), 77.05 (*C*(CH$_3$)$_3$), 55.33 (*C*H$_3$-O), 55.31 (*C*H$_3$-O), 50.63 (Ph-*C*H$_2$), 49.33 (*C*H$_2$CH$_3$), 27.70 (C(*C*H$_3$)$_3$), 22.57 (COC*H*$_3$), 14.29 (CH$_2$*C*H$_3$); IR (neat): ν = 2928, 1736, 1707, 1633, 1591, 1508 cm^{-1}; LC-MS: *m/z* (%, retention time): 688 (100, 18.01 min), 688 (100, 18.06 min); *m/z* calcd for $C_{26}H_{30}IN_3O_7S_2$ + H: 688.

Ethyl carboximidate hydrochloride 5.23

This procedure is modified from that of McElvain and Schroeder.[5.13]

Malononitrile (19.8 g, 0.3 mol) was dissolved in dry ether (100 mL) under anhydrous conditions at 0 °C, dry ethanol (17.7 mL, 0.31 mol) added and dry HCl gas bubbled through the solution using a Kipp's Apparatus until an opaqueness appeared in the solution. The flask was stoppered and stored at –18 °C for 24 h. The suspension was then chilled to –78 °C, and filtered. The collected solid was then triturated several times in ice-cold dry ether until the washings were no longer acidic by litmus paper. The solid was dried in a desiccator to yield 22.6 g (51 %) of a pale pink solid. The mother liquor was stored in a sealed flask at –18 °C for 7 d and a further 7.7 g of solid obtained. Total yield of **5.23** was 30.3 g (68 %). ^{1}H NMR (400 MHz, D_2O): δ = 4.66 (s, 2 H; CH_2), 4.13 (q, J = 7.1 Hz, 2 H; CH_2CH_3), 1.15 (t, J = 7.1 Hz, 3 H; CH_2CH_3); ^{13}C NMR: (100.6 MHz, D_2O): δ = 167.23 (C≡N), 115.96 (C=N), 63.80 (CH_2CH_3), 13.26 (CH_2CH_3); IR (neat): ν = 3118, 2987, 2805, 2265, 2006, 1744 cm^{-1}.

3,3,3-Triethoxy-propionitrile 5.24

This procedure is modified from that of deWolfe.[5.14]

Ethyl carboximidate hydrochloride **5.22** (11.496 g, 77.3 mmol) was suspended in dry ethanol (80 mL, 2.17 mmol) and stirred at RT for 3 d. Ether (40 mL) was then added to the suspension and the ammonium chloride filtered off. The ether solution was then basified with ammonia solution, dried over $MgSO_4$ and the solvents removed *in vacuo* at RT. The residue was then cooled to –78 °C and allowed to reach RT. The resultant cyanoacetamide precipitate was filtered off and the residue purified by vacuum distillation (85 °C/10 Torr) to yield **5.21** a clear colourless liquid (8.152 g, 56 % yield). R_f = 0.5 (DCM/EtOAc 49:1); b.p. 125 °C/20 Torr (literature[5.14] 83-84 °C/2 Torr); ^{1}H NMR (400 MHz, $CDCl_3$, TMS): δ = 4.27 (q, J = 7.1 Hz, 2 H; -CH_2CH_3), 3.49 (s, 2 H; CH_2), 1.33 (t, J = 7.1 Hz, 3 H; -CH_2CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$, TMS): δ = 162.90 (C N), 113.13($C(OC_2H_5)$, 62.68 (CH_2), 24.53 (-CH_2CH_3), 13.70 (-CH_2CH_3); IR (neat): υ = 3105, 2985, 2213, 1606 cm^{-1}.

Attempted Synthesis of 3,3,3-triethoxy-propionaldehyde 5.22

This procedure is modified from that of Dreyer and Metcalf.[5.15]

To a stirring solution of orthocyanoacetate **5.24** (5.88 g, 0.03 mol) in dry toluene (60 mL) at –78 °C was added DIBAL-H (1.5 M in toluene, 20 mL, 0.03 mol) over 10 min. The solution was allowed to stir for 90 min. Dry ice-chilled dry ethanol (20 mL) was then rapidly added, and the solution allowed to reach RT over 40 min with stirring. Then water (1.2 mL), NaOH (15 %, 2.15 mL), and then water (3.4 mL) was successively added to the resultant yellow solution with stirring. The suspension was then filtered after a further 20 min. The filtrate was dried ($MgSO_4$) and concentrated *in vacuo*. The resulting residue was distilled at reduced pressure to yield a polymeric tar.

(E)-4-({2-[N-Acetyl-N-(2,4-dimethoxybenzyl)-amino]-5-iodo-thieno[3,2-*d*]thiazol-6-yl}-(*t*butoxycarbonyl)-amino)-but-2-enoic acid ethyl ester 5.25

This procedure is modified from that of Dunne.[1.72]

Sodium hydride (60 %, 17 mg, 0.43 mmol) and (2,4-dimethoxybenzyl)-5-iodothienothiazole **5.15** (83 mg, 141 μmol) were added to a nitrogen-flushed reaction flask and dry DMF (1 mL) added with stirring under nitrogen. The mixture was stirred at 15 min at RT. Ethyl bromocrotonate (75 %, 60 μL, 326 μmol) was added dropwise, and the reaction mixture stirred at RT overnight. TLC analysis revealed that consumption of all the starting material had occurred, as well as the presence of a higher running spot, which stained green with ninhydrin. The residual were removed *in vacuo*, and toluene (1 mL) added repetitively and evaporated to remove DMF residues. The remaining mixture was purified by flash column chromatography (DCM/EtOAc 25:1) to yield **5.25** as a white compound (63 mg, 64 %). R_f = 0.25 (DCM/EtOAc 49:1); ^{1}H NMR (400 MHz, $CDCl_3$, TMS): δ = 6.93 (m, 2 H; Ar-C*H* and C*H*CHCOOEt), 6.46 (d, *J* = 2.0 Hz, 1 H; Ar-C*H*), 6.37 (dd, *J* = 8.5, 2.0 Hz, 1 H; Ar-C*H*), 6.02 (d, *J* = 15.0 Hz, 1 H; C*H*COOEt), 5.52 (d, *J* = 17.0 Hz, 1 H; C*H*^{a}H^b-Ph), 5.34 (d, *J* = 17.0 Hz, 1 H; CH^a*H*^{b}-Ph), 4.39 (d, *J* = 4.5 Hz, 2 H; CH_2NHCOO), 4.13 (m, 2 H; -CH_2CH$_3$), 3.85 (s, 3 H; CH_3-O), 3.76 (s, 3 H; CH_3-O), 2.83 (s, 3 H; CH_3CO), 1.49 (br s, 3 H; C*H*$^a_3$C(CH$^b_3$)$_2$), 1.26 (partially obscured s and t, *J* = 7.0 Hz, 9 H; CH$_2$C*H*$_3$ and CH$^a_3$C(C*H*$^b_3$)$_2$); ^{13}C NMR (100.6 MHz, $CDCl_3$, TMS): δ = 171.57 (CH$_3$*C*O), 166.19 (*C*OOEt), 160.39 (thiazolo-*C*2), 148.23 (N*C*OO), 143.89 (*C*HCH$_2$), 128.46 (Ar-*C*), 122.97 (*C*HCOOEt), 116.60 (Ar-*C*), 104.19 (Ar-*C*), 98.49 (Ar-*C*), 81.10 ((Ar-*C*I), 79.87 (*C*(CH$_3$)$_3$), 60.18 (*C*H$_2$CH), 55.36 (O-*C*H$_3$), 55.33 (O-*C*H$_3$), 49.94 (Ph-*C*H$_2$), 47.19 (*C*H$_2$CH$_3$), 27.99 (C(*C*H$_3$)$_3$), 22.65(CO*C*H$_3$), 14.21 (CH$_2$*C*H$_3$); IR (neat): ν = 2975, 1705, 1661, 1608, 1508 cm^{-1}; elemental analysis calcd for $C_{27}H_{32}IN_3O_7S_2$ (701.6): C 46.22, H 4.60, N 5.99; found C 47.06, H 5.06, N 5.91.

2-[N-Acetyl-N-(2,4-dimethoxybenzyl)-amino]-5-ethoxycarbonylmethyl)-pyrrolo[2′,3′:4,5] thieno[3,2-*d*]thiazole-7-carboxylic acid ^{t}butyl ester 5.26

This procedure is modified from that of Dunne.[1.72]

To a suspension of crotonylamino iodothienothiazole **5.25** (45 mg, 64 μL), $Pd(OAc)_2$ (2 mg, 9 μmol) and triphenylphosphine (4 mg, 15 μmol) in dry acetonitrile (1 mL) was added DIPEA (50 μL, 287 μmol). The mixture was stirred under nitrogen at 65 °C for 18 h. TLC analysis showed that the starting material had been consumed and a major product had formed by UV. The mixture was concentrated *in vacuo* and the residue subjected to flash column chromatography (DCM/EtOAc 49:1) to give the *title compound* **5.26** as a white solid (28 mg, 75 %). R_f = 0.35 (DCM/EtOAc 49:1); ^{1}H NMR (400 MHz, $CDCl_3$, TMS): δ = 7.34 (s, 1 H; C-5 pyrrolo-*H*), 7.20 (d, *J* = 8.5 Hz, 1 H; Ar-C*H*), 6.45 (s, 1 H; Ar-C*H*), 6.37 (d, *J* = 8.5 Hz, 1 H; Ar-C*H*), 5.49 (br s, 2 H; CH_2-Ph), 4.19 (q, *J* = 7.0 Hz, 2 H; CH_2CH_3), 3.85 (s, 3 H; CH_3-O), 2.39 (s, 3 H; CH_3-O), 1.58 (s, 9 H; $C(CH_3)_3$), 1.30 (q, *J* = 7.0 Hz, 3 H; CH_2CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$, TMS): δ = 171.31 (*C*O), 170.49 (*C*O), 160.21 (thiazolo-*C*2), 149.10 (Ar-*C*), 130.43 (Ar-*C*), 129.71 (Ar-*C*), 127.41 (Ar-*C*), 122.26 (Ar-*C*), 116.68 (Ar-*C*), 113.31 (Ar-*C*), 104.16 (Ar-*C*), 98.27 (Ar-*C*), 84.32 ($C(CH_3)_3$), 61.19 (Ph-CH_2), 55.34 (O-CH_3), 55.27 (O-CH_3), 47.04 (CH_3CH_2COO), 32.39 (CH_2COOEt), 28.11 ($C(CH_3)_3$), 22.76 (CH_3CO), 14.24 (CH_2CH_3); IR (neat): ν = 2982, 1743, 1728, 1662, 1619, 1589, 1509 cm^{-1}.

3,4-Diacetamidothiophene 6.4

This procedure is modified from that of Easton et al.[5.1]

The hexachlorostannate salt **5.2** was suspended in water (20 mL) and methanol (10 mL), and $NaHCO_3$ (3.1 g, 37 mmol) added slowly with stirring. Once the effervescence had subsided, acetic anhydride (7 mL, 74 mmol) was added to the mixture. A precipitation started to form immediately. The suspension was allowed to stir at RT overnight. The resulting precipitate was collected by filtration and recrystallised from acetone to yield 1.265 g (86 % yield) of **6.4** as a grey solid. R_f = 0.14 (EtOAc); M.p. 218-219 °C (decomp.) (literature[5.1] 208-209 °C (water)); ^{1}H NMR (400 MHz, [D_6]DMSO): δ = 9.48 (b, 2 H; N*H*), 7.55 (b, 2 H; Ar-C*H*), 2.08 (s, 6 H; CH_3); ^{13}C NMR (100.6 MHz, [D_6]DMSO): δ = 167.94 (C=O), 128.90 (C), 110.64 (C), 23.44 (*C*H_3); IR (neat): ν = 3222, 3160, 3016, 1648, 1585 cm^{-1}.

_N,N'_-Bis(acetylamino)-3,4-thiophenediyl-2,5-bisthiocyanate 6.5

This procedure is modified from that of Paulmier.[6.2]

A stirring solution of 3,4-diacetamidothiophene **6.4** (0.51 g, 2.60 mmol) and ammonium thiocyanate (1.80 g, 0.024 mol) in methanol (50 mL) at 0 °C was added bromine (0.49 mL, 1.5 g, 0.014 mol) dropwise. The reaction mixture was allowed to occur at 10 °C for 1 h. Water (100 mL) was added to the mixture, and then the mixture was allowed to stand at RT overnight. The resulting precipitate was collected by filtration and recrystallised from acetone to yield **6.5** as a grey solid (0.540 g, 64%). R_f = 0.4 (EtOAc); m.p. 232 °C (decomp.) (literature[6.2] 220 °C (decomp.)); ^{1}H NMR (400 MHz, [D$_6$]DMSO): δ = 10.06 (b, 2 H; CH$_3$N*H*), 2.10 (s, 6 H; C*H*$_3$NH); ^{13}C NMR (100.6 MHz, [D$_6$]DMSO): δ = 169.12 (*C*OCH$_3$), 151.51 (C≡N), 135.12 (Ar-*C*), 111.25 (Ar-*C*), 22.72 (CO*C*H$_3$); IR (neat): ν = 3288, 3064, 2157, 1689, 1650, 1553, 1500 cm^{-1}.

3-Aminothiophene oxalate 6.8

This procedure is modified from that of Barker et al.[6.6]

A suspension of methyl 3-aminothiophene-2-carboxylate **6.7** (6.40 g, 0.04 mol) in NaOH (2 M, 20 mL) was refluxed for 50 min. The solution was quickly cooled to 0 °C and concentrated HCl added cautiously until the resulting suspension has reached pH 7. The solid formed was then quickly filtered into a Buchner funnel and pressed as dry as possible with a spatula. The solid was then dissolved in acetone (50 mL), dried (MgSO$_4$), filtered and the solvents removed at reduced pressure at 20 °C. The residue was suspended in 1-propanol (4.00 g, 0.04 mol) added, and the suspension stirred at 38 °C for 45 min. The suspension was diluted with ether (30 mL), filtered, washed with ether and dried to yield 4.9 g of **6.8** as a white/lilac solid. IR (neat): ν = 3303, 2890, 2649, 1917, 1723, 1688, 1598, 1524 cm^{-1}.

3-Acetamidothiophene 6.9

This procedure is modified from that of Campaigne and Monroe.[6.5]

3-aminothiophene oxalic salt **6.8** (0.75 g) was suspended in water (20 mL) and ammonia solution added until a solution had formed, and the solution was basic by litmus paper. The pale yellow solution was extracted with DCM (3 X 15 mL), dried (Na_2SO_4) and concentrated *in vacuo* to yield a yellow/brown oil. ^{1}H NMR (400 MHz, $CDCl_3$, TMS): δ = 7.10 (s, 1 H; Ar-C*H*), 6.61 (s, 1 H; Ar-C*H*), 6.12 (s, 1 H; N*H*), 3.58 (b, 3 H; N*H*); ^{13}C NMR (100.6 MHz, $CDCl_3$, TMS); δ = 145.02 (Ar-*C*-NH_2), 125.17 (Ar-*C*), 121.02 (Ar-*C*), 100.03 (Ar-*C*). Acetic anhydride (10 mL) was added to the oil, the solution swirled for a few seconds and the excess acetic anhydride removed under reduced pressure. The white residue was recrystallised from water to yield 0.385 g (45 % from **6.7**) of the *title compound* **6.9**. R_f = 0.23 (9:1 DCM/EtOAc); m.p. 152-153 °C (literature[6.5] 148 °C); ^{1}H NMR (400 MHz, $CDCl_3$, TMS): δ = 7.75 (s, 1 H; N*H*), 7.54 (d, *J* = 1.2 Hz, 1 H; Ar-C*H*), 7.22 (m, 1 H; Ar-C*H*), 6.99 (d, *J* = 5.1 Hz, 1 H; Ar-C*H*), 2.16 (s, 3 H; CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$, TMS): δ = 167.60 (C=O), 135.54 (Ar-*C*), 124.47 (Ar-*C*), 120.94 (Ar-*C*), 110.26 (Ar-*C*), 23.86 (Ar-*C*); IR (neat): ν = 3261, 3218, 3091, 2929, 1654, 1572, 1528 cm^{-1}.

3-Acetamido-2-thiocyanatothiophene 6.10

This procedure is modified from that of Paulmier.[6.1]

To a solution of 3-acetamidothiophene **6.9** (0.320 g, 2.27 mmol) and ammonium thiocyanate (0.700 g, 9.08 mmol) in methanol (120 mL) was added bromine (0.13 mL, 2.5 mmol) dropwise with stirring at RT. The resultant red solution was allowed to stir at RT for 2 h; then diluted with water (500 mL) and the pH adjusted to 4-5 with sodium acetate. The solution was then extracted with DCM, the DCM washed with water, then dried over $MgSO_4$ and the solvent removed under vacuum. The white solid obtained was recrystallised from $CHCl_3$/Pet. Ether (80-100) to yield 0.275 g (61 %) of **6.10**. R_f = 0.19 (DCM/EtOAc 24:1); m.p. 135-137 °C (literature[6.1] 118 °C); ^{1}H NMR (400 MHz, $CDCl_3$, TMS): δ = 7.93 (d, *J* = 5.6 Hz, 1 H; Ar-C*H*), 7.78 (b, 1 H; N*H*), 7.58 (d, *J* = 5.8 Hz, 1 H; Ar-C*H*), 2.28 (s, 3 H; CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$, TMS): δ = 167.34 (C=O), 143.04 (C=N), 132.41 (Ar-*C*), 123.15 (Ar-*C*), 109.36 (Ar-*C*), 24.16 (*C*H_3); IR (neat): ν = 3338, 3148, 3108, 2163, 1692, 1617, 1565 cm^{-1}.

Thioacetic acid-S-(3-acetylamino-thiophen-2-yl)ester 6.11

This procedure is modified from that of Paulmier.[6.2]

3-acetamido-2-thiocyanatothiophene **6.10** (200 mg, 1.0 mmol) was progressively added to a solution of $Na_2S.9H_2O$ (700 mg, 2.9 mmol) in water (5 mL) with stirring over 1 h. The solution was allowed to stir at RT for a further hour. The solution was then cooled to 0 °C and acetic anhydride added to adjust the pH to 4-5. The resultant suspension was stirred at 0-10 °C for 30 min, after which time the solid was collected, washed with water, dried and recrystallised (Hexane/$CHCl_3$ 90:10) to yield 64 mg (30 %) of **6.11** as white needles. R_f = 0.35 (DCM/EtOAc 9:1); m.p. 108-110 °C (literature[6.2] 102-103 °C); ^{1}H NMR (400 MHz, $CDCl_3$, TMS): δ = 7.95 (d, *J* = 5.6 Hz, 2 H; Ar-*H*), 7.52 (d, *J* = 5.6 Hz, 2 H; Ar-*H*), 7.41 (b, 1 H, N*H*), 7.26 (s, *1* H; Ar-*H*), 2.44 (s, 3 H; SOCH_3), 2.18 (s, 3 H, NHCOCH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$, TMS): δ = 194.53 (C=O), 141.97 (Ar-*C*), 130.48 (Ar-*C*), 122.95 (Ar-*C*), 111.83 (Ar-*C*), 29.62 (SO*C*H_3), 24.37 (NHCO*C*H_3); IR (neat): ν = 3332, 1698, 1682, 1559 cm^{-1}.

[7.1] D.D. Perrin, W.L.F. Armarego, and D.R. Perrin, *Purification of Laboratory Chemicals*, Pergamon Press, Oxford, **1997**.

[7.2] W. Clarke-Still, M. Kahn and A. Mitra, *J. Org. Chem.* **1978**, *43(14)*, 2923-2925.

[7.3] L.F. Audrieth, E. S. Scott and P. S. Kippur, *J. Org. Chem.* **1954**, 733-741.

[7.4] C. Galvez, F. Garcia, J. Garcia and J. Soldevila, *J. Heterocyclic Chem.* **1986**, *23*, 1103-1108.

Chapter Eight

References

8.1 References

1.1 Picture courtesy of T.W. G. Solomons, *Organic Chemistry,* 5th Ed. **1992**, Wiley , Books, 1146.

1.2 S. Neidle, L. H. Pearl and J. V. Skelly, *Biochem. J.* **1987**, *243*, 1-13.

1.3 G. Colmenarejo, F. Montero and G. Orellana, *Anti-Cancer Drug Design* **1997**, *12*, 239-260.

1.4 A. H. Krotz, L. Y. Kuo, T. P. Shields, and J. K. Barton, *J. Am. Chem. Soc.* **1993**, *115*, 3877-3882.

1.5 A. H. Krotz, B. P. Hudson and J. K. Barton, *J. Am. Chem. Soc.* **1993**, *115*, 12577-12578.

1.6 B. P. Hudson, C. M. Dupureur and J. K. Barton, *J. Am. Chem. Soc.* **1995**, *117*, 9379-9380.

1.7 C. L. Kielkopf, K. E. Erkkila, B. P. Hudson, J. K. Barton and D. C. Rees, *Nat. Struct. Biol.* **2000**, *7(2)*, 117-121.

1.8 R. H. Terbrueggen and J. K. Barton, *Biochemistry* **1995**, *34*, 8227-8234.

1.9 R. H. Terbrueggen, T. W. Johann and J. K. Barton, *Inorg. Chem.* **1998**, *37*, 6874-6883.

1.10 S. J. Franklin and J. K. Barton, *Biochemistry* **1998**, *37*, 16093-16105.

1.11 D. T. Odom, C. S. Parker and J. K. Barton, *Biochemistry* **1999**, *38*, 5155-5163.

1.12 R. E. Dickerson, *Scientific American* **1983**, *249*, 86-102.

1.13 a) M. Coll, J. Aymami, G. A. van der Marel, J. H. van Boom, A. Rich, and A. H.-J. Wang, *Biochemistry* **1989**, *28*, 310-320; b) M. Coll, C. A. Frederick, A. H.-J. Wang, and A. Rich, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 8385-8389.

1.14 J. W. Lown, K. Krowicki, U. G. Bhat, A. Skorobogaty, B. Ward and J. C. Dabrowiak, *Biochemistry* **1986**, *25*, 7408-7416.

1.15 W. S. Wade, M. Mrksich and P. B. Dervan, *J. Am. Chem. Soc.* **1992**, *114*, 8783-8794.

1.16 M. Mrksich, M. E. Parks and P. B. Dervan, *J. Am. Chem. Soc.* **1994**, *116*, 7983-7988.

1.17 M. E. Parks, E. E. Baird and P. B. Dervan, *J. Am. Chem. Soc.* **1996**, *118*, 6153-6159.

1.18 D. M. Herman, E. E. Baird and P. B. Dervan, *J. Am. Chem. Soc.* **1998**, *120*, 1382-1391.

1.19 J. J. Kelly, E. E. Baird and P. B. Dervan, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 6981-6985.

1.20 C. L. Kielkopf, E. E. Baird, P. B. Dervan and D. C. Rees, *Nat. Struct. Biol.* **1998**, *5(2)*, 104-109.

1.21 J. W. Trauger, E. E. Baird, M. Mrksich and P. B. Dervan, *J. Am. Chem. Soc.* **1996**,

118, 6160-6166.

1.22 R. Peláez Lamamie de Clairac, C. J. Seel, B. H. Geierstanger, M. Mrksich, E. E. Baird, P. B. Dervan and D. E. Wemmer, *J. Am. Chem. Soc.* **1999**, *121*, 2956-2964.

1.23 J. W. Trauger, E. E. Baird and P. B. Dervan, *Angew. Chem. Int. Ed.* **1998**, *37*, 1421-1423.

1.24 J. W. Trauger, E. E. Baird and P. B. Dervan, *J. Am. Chem. Soc.* **1998**, *120*, 3534-3535.

1.25 S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird and P. B. Dervan, *Nature* **1998**, *391*, 468-471.

1.26 C. L. Kielkopf, S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan and D. C. Rees, *Science* **1998**, *282*, 111-115.

1.27 A. R. Urbach, J. W. Szewczyk, S. White, J. M. Turner, E. E. Baird and P. B. Dervan, *J. Am. Chem. Soc.* **1999**, *121*, 11621-11629.

1.28 S. White, J. M. Turner, J. W. Szewczyk, E. E. Baird and P. B. Dervan, *J. Am. Chem. Soc.* **1999**, *121*, 260-261.

1.29 C. Melander, D. M. Herman and P. B. Dervan, *Chem. Eur. J.* **2000**, *6(24)*, 4487-4497.

1.30 U. Ellervik, C. C. C. Wang, and P. B. Dervan, *J. Am. Chem. Soc.* **2000**, *122*, 9354-9360.

1.31 A. R. Urbach and P. B. Dervan, *Proc. Nat. Acad. Sci. USA* **2001**, *98(8)*, 4343-4348.

1.32 R. E. Bremer, J. W. Szewczyk, E. E. Baird and P. B. Dervan, *Bioorg. Med. Chem.* **2000**, *8*, 1947-1955.

1.33 C. C. C. Wang, U. Ellervik and P. B. Dervan, *Bioorg. Med. Chem.* **2001**, *9*, 653-657.

1.34 P. E. Florancig, S. E. Swalley, J. W. Trauger and P. B. Dervan, *J. Am. Chem. Soc.* **2000,** *122*, 6342-6350.

1.35 Z.-Y. J. Zhan and P. B. Dervan, *Bioorg. Med. Chem.* **2001**, *9*, 2467-2474.

1.36 D. H. Nyugen, J. W. Swewczyk, E. E. Baird and P. B. Dervan, *Bioorg. Med. Chem.* **2001**, *9*, 7-17.

1.37 G. Felsenfeld, D. R. Davies and A. Rich, *J. Am. Chem. Soc.* **1957**, *79*, 2023-2024.

1.38 a) H. E. Moser and P. B. Dervan, *Science* **1987**, *238*, 645-650; b) T. Le Doan, L. Perrouault, D. Praseuth, N. Habhoub, J.-L. Decout, N. T. Thuong, J. Lhomme and C. Hélène, *Nucleic Acids Res.* **1987**, *15*, 7749-7760.

1.39 M. D. Keppler and K. R. Fox, *Nucleic Acids Res.* **1997**, *25(22)*, 4644-4649.

1.40 W. M. Olivas and L. J. Maher III, *Biochemistry* **1995**, *34*, 278-284.

1.41 J. S. Lee, M. L. Woodsworth, L. J. P. Latimer and A. R. Morgan, *Nucleic Acids Res.* **1984**, *12* (*16*), 6603-6614.

1.42 T. J. Povsic and P. B. Dervan, *J. Am. Chem. Soc.* **1989**, *111*, 3059-3061; b) L. E. Xodo, G. Manzini, F. Quadrifoglio, G. A. van der Marel and J. H. van Boom, *Nucleic*

Acids Res. **1992**, *19(20)*, 5625-5631.

1.43 G. Xiang, R. Bogacki and L. W. McLaughlin, *Nucleic Acids Res.* **1996**, *24 (10)*, 1963-1970.

1.44 S. A. Cassidy, P. Slickers, J. O. Trent, D. C. Capaldi, P. D. Roselt, C. B. Reese, S. Neidle and K. R. Fox, *Nucleic Acids Res.* **1997**, *25(24)*, 4891-4898.

1.45 L. C. Griffin and P. B. Dervan, *Science* **1989**, *245*, 967-971.

1.46 G. C. Best and P. B. Dervan, *J. Am. Chem. Soc.* **1995**, *117*, 1187-1193.

1.47 D. M. Gowers and K. R. Fox, *Nucleic Acids Res.* **1997**, *25(19)*, 3787-3794.

1.48 D. M. Gowers, J. Bijapur, T. Brown and K. R. Fox, *Biochemistry* **1999**, *38*, 13747-13748.

1.49 D. A. Gianolio and L. W. McLaughlin, *J. Am. Chem. Soc.* **1999**, *121*, 6334-6335.

1.50 J. Sung Koh and P. B. Dervan, *J. Am. Chem. Soc.* **1992**, *114*, 1470-1479.

1.51 P. S. Miller, P. Bhan, C. D. Cushman and T. L. Trapane, *Biochemistry* **1992**, *31*, 6788-6793.

1.52 P. S. Miller, G. Bi, S. A. Kipp, V. Fok and R. K. Delong, *Nucleic Acids Res.* **1996**, *24(4)*, 730-736.

1.53 W. A. Greenberg and P. B. Dervan, *J. Am. Chem. Soc.* **1995**, *117*, 5016-5022.

1.54 T. S. Rao, R. H. Durland, D. M. Seth, M. A. Myrick, V. Bodepudi and G. R. Revankar, *Biochemistry* **1995**, *34*, 765-772.

1.55 W. M. Olivas and L. J. Maher III, *Nucleic Acids Res.* **1995**, *23*, 1936-1941.

1.56 V. M. Marathias, M. J. Sawicki and P. H. Bolton, *Nucleic Acids Res.* **1999**, *27(14)*, 2860-2867.

1.57 P. Brodin, J.-S. Sun, J.-F. Mouscadet and C. Auclair, *Nucleic Acids Res.* **1999**, *27(15)*, 3029-3034.

1.58 L. C. Griffin, L. L. Kiessling, P. A. Beal, P. Gillespie and P. B. Dervan, *J. Am. Chem. Soc.* **1992**, *114*, 7976-7982.

1.59 K. M. Koshlap, P. Gillespie, P. B. Dervan and J. Feigon, *J. Am. Chem. Soc.* **1993**, *115*, 7908-7909.

1.60 E. Wang, K. M. Koshlap, P. Gillespie, P. B. Dervan and J. Feigon, *J. Mol. Biol.* **1996**, *257*, 1052-1069.

1.61 C.-Y. Huang, C. D. Cushman and P. S. Miller, *J. Org. Chem.* **1993**, *58*, 5048-5049.

1.62 C.-Y. Huang, G. Bi and P. S. Miller, *Nucleic Acids Res.* **1996**, *24(13)*, 2606-2613.

1.63 S. C. Zimmerman and P. Schmitt, *J. Am. Chem. Soc.* **1995**, *117*, 10769-10770.

1.64 E. Mertz, S. Mattei and S. C. Zimmerman, *Org. Lett.* **2000**, *2(19)*, 2931-2934.

1.65 T. E. Lehmann, W. A. Greenberg, D. A. Liberles, C. K. Wada and P. B. Dervan, *Helv. Chim. Acta* **1997**, *80*, 2002-2022.

1.66 D. Lengeler and K. Weisz, *Tetrahedron Lett.* **2001**, *42*, 1479-1481.

1.67 S. Sasaki, S. Nakashima, F. Nagatsugi, Y. Tanaka, M. Hisatome and M. Maeda, *Tetrahedron Lett.* **1995**, *36(52)*, 9521-9524.

1.68 J. H. Rothman and W. G. Richards, *Biopolymers* **1996**, *39*, 795-812.

1.69 A. B. Eldrup, O. Dahl and P. E. Nielsen, *J. Am. Chem. Soc.* **1997**, *119*, 11116-11117.

1.70 a) M. Egholm, O. Buchardt, P. E. Nielsen, and R. H. Berg, *J. Am. Chem. Soc.* **1992**, *114*, 1895-1897; b) P. E. Nielsen and G. Haaima, *Chem. Soc. Rev.* **1997**, 73-78.

1.71 G. P. Dunne, Ph.D. Thesis, University of Dublin, **1999**.

1.72 G. P. Dunne, Postdoctoral Report, University of Dublin, **2000**.

1.73 a) S. Tomàs, R. Prohens, M. Vega, M. C. Rotger, P. M. Deyà, P. Ballester, and A. Costa, *J. Org. Chem.* **1996**, *61*, 9394-9401; b) R. Prohens, S. Tomàs, J.Morey, P. M. Deyà, P. Ballester, and A. Costa, *Tetrahedron Lett.* **1998**, *39*, 1063-1066.

1.74 A. P. Davis, S. M. Draper, G. Dunne and P. Ashton, *Chem. Commun.* **1999**, 2265-2266.

2.1 (a) R. P. Sijbesma, F. H. Beijer, L. Brunsveld, B. J. B. Folmer, J. H. K. K. Hirschberg, R. F. M. Lange, J. K. L. Lowe, and E. W. Meijer, *Science* **1997**, *278*, 1601-1604; (b) F. H. Beijer, R. P. Sijbesma, H. Kooijman, A. L. Spek, and E. W. Meijer, *J. Am. Chem. Soc.* **1998**, *120*, 6761-6769; (c) P. S. Corbin and S. C.

2.2 Zimmerman, *J. Am. Chem. Soc.* **1998**, *120*, 9710-9711.

(a) G. Maahs and P. Hegenberg, *Angew. Chem. Int. Ed.* **1966**, *5(10)*, 888-893; (b) A. H. Schmidt and W. Ried, *Synthesis* **1978**, 1-22.

2.3 A. H. Schmidt, *Synthesis* **1980**, 961-994.

2.4 L. S. Liebeskind, R. W. Fengl, K. R. Wirtz and T. T. Shawe, *J. Org. Chem.* **1988**, *52*, 2482-2488.

2.5 L. S. Liebeskind, M. S. Yu and R. W. Fengl, *J. Org. Chem.* **1993**, *58*, 3543-3549.

3.1 E. Schaumann, *The Chemistry of Double Bonded Functional Groups,* **1987**, Wiley, 1270-1367.

4.1 H. Ehrhardt, S. Hünig and H. Pütter, *Chem. Ber.* **1977**, *110*, 2506-2523.

4.2 R. Kober and W. Steglich, *Liebigs Ann. Chem.* **1983**, 599-609

4.3 M. Müller, M. J. Heileman, H. W. Moore, E. Schaumann and G. Adiwidjaja, *Synthesis* **1997**, 50-52.

4.4 a) F. Arndt, *Ber.* **1921**, *54B*, 2236-2242; b) P. R. Shildneck and Wallace Windus, *Org. Synth. Coll. 2.* 411-412.

4.5 L. Fishbein and J. A. Gallaghan, *J. Am. Chem. Soc.* **1954**, *76*, 1877-1879.

4.6 L. Long Jr., R. C. Clapp, F. H. Bissett and T. Hasselstrom, *J. Org. Chem.* **1961**, 85-88.

4.7 R. L. Shriner and F. W. Neumann, *Org. Synth. Coll. 3.* 73-75.

4.8 W. G. Finnegan, R. A. Henry and E. Lieber, *J. Am. Chem. Soc.* **1952**, 779-791.

4.9 E. S. Scott, E. E. Zeller and L. F. Audrieth, *J. Org. Chem.* **1954**, 749-752.

4.10 H. Emilsson, *J. Heterocyclic Chem.* **1989**, *26*, 1077-1081.

4.11 a) H. Beyer, W. Liebenow and T. Pyl, *Chem. Ber,* **1957**, 1738-1744; b) K. A. Jensen, A. Holm and S. Rachlin, *Acta Chem. Scand,* **1966**, *20(10),* 2795-2806.

4.12 G. W. Kirsten and G. B. L. Smith. *J. Am. Chem. Soc.* **1936**, *58,* 800-801.

4.13 L. Heinisch, *Chem. Ber.* **1967**, *100*, 893-904.

4.14 D. S. Tarbell, Y. Yamamoto and B. M. Pope, *Proc. Natl. Acad. Sci. USA* **1972**, *69(3)*, 730-732.

4.15 R. M. Herbst and D. Shemin, *Org. Synth. Coll. 1.* 11-12.

4.16 W. E. Hanby, S. G. Waley and J. Watson, *J. Chem. Soc.* **1950**, 3239-3249.

4.17 P. K. Nandi and D. R. Robinson, *J. Am. Chem. Soc.* **1972**, *94(4)*, 1299-1308.

4.18 M. Ohno, Y. Yamamoto, Y. Shirasaki and S. Eguchi, *J. Chem. Soc. Perkin Trans. 1.* **1993**, *2*, 263-271.

4.19 J. E. Thorpe, *J. Chem. Soc. (B)* **1968**, 435-436.

4.20 R. West, *Oxocarbons*, **1980**, Academic Press, 15-41.

5.1 R. Mozingo, S. A. Harris, D. E. Wolf, C. E. Hoffhine Jr., N. R. Easton and K. Folkers, *J. Am. Chem. Soc.* **1945**, *67*, 2092-2095.

5.2 F. Outurquin and C. Paulmier, *Bull. Soc. Chim. Fr.* **1983**, *5-6 (II)*, 153-158.

5.3 A. Takamizara, K. Hirai and K. Matsui, *Bull. Chem. Soc. Jpn.* **1963**, *36*, 1214-1220.

5.4 S. F. Dyke, *The Chemistry of Enamines*, **1973**, Cambridge University Press, 1.

5.5 R. K. Robins, G. R. Revankar, D. E. O'Brien, R. H. Springer, T. Novinson, A. Albert, K. Senga, J. P. Miller and D. G. Streeter, *J. Heterocyclic Chem.* **1985**, *22*, 601-634.

5.6 P. Nussbaumer, K. Baumann, T. Dechat and M. Harasek, *Tetrahedron* **1991**, *47(26)*, 4591-4602.

5.7 D. Wensbo, U. Annby and S. Gronowitz, *Tetrahedron* **1995**, *51(37)*, 10323-10342.

5.8 L. Siracusa, Postdoctoral Report, University of Bristol, **2001**.

5.9 R. Grigg and V. Savic, *Chem. Commun.* **2000**, 873-874.

5.10 P. Sanna, P. A. Sequi, S. Piras and G. Pagletti, *Heterocycles* **1995**, *41(11)*, 2459-2474.

5.11 N. D. Heinkel, P. D. Kennewell and V. B. Fish, *J. Heterocyclic Chem.* **1969**, *6*, 77-81.

5.12 J. Hénin, J. Vercauteren, C. Mangenot, B. Hénin , J. M. Nuzillard and J. Guilhem, *Tetrahedron* **1999**, *55*, 9817-9828.

5.13 S. M. McElvain and J. P. Schroeder, *J. Am. Chem. Soc.* **1949**, *71*, 40-46.

5.14 R.W. deWolfe, *Synthesis* **1974**, 153-172.

5.15 G. B. Dreyer and B. W. Metcalf, *Tetrahedron* **1988**, *29(52)*, 6885-6888.

5.16 K. Sonogashira, Y. Tohda, and N. Hagihara, *Tetrahedron Lett.* **1975**, 4467-4470.

5.17 D. Brugier, F. Outurquin and C. Paulmier, *Tetrahedron* **2000**, *56*, 2985-2993.

5.18 J. Reisch and S. Dittmann, *J. Heterocyclic Chem.* **1992**, *29*, 1857-1858.

6.1 C. Paulmier, *Bull. Soc. Chim. Fr.* **1980**, 3-4(*II*), 151-156.

6.2 C. Paulmier, *Bull. Soc. Chim. Fr.* **1979**, 11-12(*II*), 592-598.

6.3 H. R. Snyder, E. P. Merica, C. G. Force and E. G. White, *J. Am. Chem. Soc.* **1958**, *80*, 4622-4625.

6.4 V. I. Shedov, V. K. Vasil'era, Y. I. Trofimkin and A. H. Griner, *Chem. Hetero. Cmpds. (Russ.)* **1975**, 1473-1475, provided courtesy of their Latvian publisher.

6.5 E. Campaigne and P. A. Monroe, *J. Am. Chem. Soc.* **1954**, *76*, 2447-2450.

6.6 J. M. Barker, P. R. Huddleston and M. L. Wood, *Synth. Commun.* **1995**, *25(23),* 3729-3734.

6.7 M. Reinecke, H. W. Adickes and C. Pyun, *J. Org. Chem.* **1971**, *36(18),* 2690-2692.

7.1 D. D. Perrin, W. L. F. Armarego, and D. R. Perrin, *Purification of Laboratory Chemicals*, Pergamon Press, Oxford, **1997**.

7.2 W. Clarke-Still, M. Kahn and A. Mitra, *J. Org. Chem.* **1978**, *43(14)*, 2923-2925.

7.3 L. F. Audrieth, E. S. Scott and P. S. Kippur, *J. Org. Chem.* **1954**, 733-741.

7.4 C. Galvez, F. Garcia, J. Garcia and J. Soldevila, *J.Heterocyclic Chem.* **1986**, *23*, 1103-1108.

Appendix A

NMR Reference Codes

APPENDIX A: NMR REFERENCE CODES

Entry	*NMR Data File*	*Entry*	*NMR Data File*
2.5	MRK123	**4.18**	MRK259
2.6	MRK054, 131	**4.15**	MRK260
2.8	MRK112	**4.16**	MRK290
2.10	MRK059	**4.3**	MRK286
2.11	MRK304	**4.19**	MRK125/270901
2.12	MRK061	**4.21**	MRK269
2.13	MRK060	**4.23**	MRK256
2.14	MRK101	**4.24**	MRK253
2.15	MRK065	**4.26**	MRK262
2.16	MRK117	**4.27**	MRK265
2.17	MRK102	**4.28**	MRK266, 268
2.18	MRK034	**4.29**	MRK274
2.19	MRK116	**4.30**	MRK271
2.20	MRK133	**4.31**	MRK292
2.21	MRK140	**4.32**	MRK294
2.22	MRK148	**5.1**	MK3018
3.1	MRK104	**5.2**	MRK334
3.3	MRK128	**5.3**	MK3050
3.4	MRK068	**5.4**	MK1531
3.5	MRK103	**5.5**	MK4958
3.6	MRK224	**5.6**	MK2996
3.7	MRK197	**5.7**	MK2114
3.8	MRK169	**5.8**	MK4959
3.9	MRK159	**5.12**	admk2279
3.11	MRK218	**5.16**	MK0762
3.12	MRK219	**5.17**	MK1867
3.13	MRK205	**5.19**	MK5749
3.16	MRK177	**5.23**	MK1386
3.17	MRK170	**5.24**	MK0826
4.11	MRK244	**5.25**	MK1387
4.13	MRK243	**5.26**	MK2317
4.14	MRK245	**6.3**	MK0682
4.17	MRK258	**6.4**	MRK330
		6.5	MRK340
		6.9	MK0827
		6.10	MK0882
		6.11	MK2995

Appendix B

X-Ray Crystallography Tables

APPENDIX B: X-RAY CRYSTALLOGRAPHY TABLES

Table 1. Crystal data and structure refinement for $C_{17}H_{22}N_4O_3S_2$ - **3.10**.

Empirical formula	$C_{25}H_{27}N_3O_5S_2$	
Formula weight	394.51	
Temperature	293(2) K	
Wavelength	0.71073 Å	
Crystal system	triclinic	
Space group	P –1 No. 2	
Unit cell dimensions	a = 9.7291(10) Å	α= 90.00°.
	b = 10.4299(5) Å	β= 98.516(7)°.
	c = 19..7532(10) Å	γ = 90.00°.
Volume	1982.3(2) $Å^3$	
Z	4	
Density (calculated)	1.322 Mg/m^3	
Absorption coefficient	0.292 mm^{-1}	
F(000)	832	
Crystal size	0.50 x 0.30 x 0.20 mm^3	
Theta range for data collection	1.04 to 24.99°.	
Index ranges	$0 \leq h \leq 10$, $0 \leq k \leq 11$, $-20 \leq l \leq 20$	
Reflections collected	3557	
Independent reflections	3344 [R(int) = 0.0000]	
Completeness to theta = 24.99°	95.9 %	
Max. and min. transmission	0.9753 and 0.9515	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	3344 / 0 / 323	
Goodness-of-fit on F^2	1.000	
Final R indices [I>2sigma(I)]	R1 = 0.0.477, wR2 = 0.0942	
R indices (all data)	R1 = 0.1282, wR2 = 0.1258	
Largest diff. peak and hole	0.258 and -0.219 $e.Å^{-3}$	

Table 2. Atomic coordinates (x 10^4) and equivalent isotropic displacement parameters ($Å^2$x 10^3) for **3.10**. U(eq) is defined as one third of the trace of the orthogonalized U^{ij} tensor.

	x	y	z	U(eq)
C(1)	6272(4)	3165(3)	16682(17)	333(9)
C(2)	6163(4)	4574(4)	16788(18)	394(10)
C(3)	5242(4)	4414(4)	22283(19)	399(10)
C(4)	5391(4)	3034(3)	21680(17)	328(9)
C(5)	4051(4)	1854(4)	29382(18)	374(10)
C(6)	2672(5)	3001(5)	3697(2)	502(12)
C(7)	3364(4)	3213(4)	4416(2)	478(11)
C(8)	2988(7)	4224(6)	4792(3)	757(17)
C(9)	3566(9)	4398(8)	5465(4)	99(2)
C(10)	4538(9)	3559(9)	5761(4)	101(3)
C(11)	4943(8)	2552(8)	5403(3)	98(2)
C(12)	4351(7)	2401(6)	4730(3)	772(17)
C(13)	8218(4)	1951(4)	479(19)	426(10)
C(14)	9132(6)	2446(6)	-4(3)	575(13)
C(15)	7914(9)	550(5)	456(4)	709(18)
C(16)	3517(3)	11602(3)	1702(3)	28(1)
C(17)	2568(3)	12373(3)	1252(3)	33(1)
C(18)	2616(4)	12847(3)	278(3)	42(1)
C(19)	3598(4)	12539(3)	-257(3)	41(1)
C(20)	4528(3)	11771(3)	175(3)	40(1)
C(21)	4495(3)	11303(3)	1150(3)	36(1)
C(22)	9081(3)	12424(3)	6582(3)	31(1)
C(24)	7573(3)	13832(3)	5722(3)	37(1)
C(25)	7492(4)	14508(3)	6986(4)	51(1)
C(26)	8621(4)	14438(3)	5373(4)	48(1)
C(27)	6173(4)	13611(3)	4662(4)	51(1)
C(28)	11603(3)	8890(3)	7265(3)	34(1)
C(29)	11187(3)	7892(3)	7718(3)	33(1)
C(31)	11914(4)	6191(3)	8492(4)	53(1)
C(32)	12950(6)	5347(5)	8383(8)	121(3)
C(50)	1717(7)	6891(6)	2742(3)	695(15)

C(51)	-392(7)	6885(7)	1700(4)	680(16)
N(1)	6920(4)	2326(3)	13711(14)	563(8)
N(2)	7768(3)	2777(3)	866(16)	401(8)
N(3)	4920(4)	1963(3)	24488(16)	367(8)
N(4)	3624(4)	2937(4)	31896(18)	432(9)
O(1)	6554(3)	5472(3)	13711(14)	563(8)
O(2)	4645(3)	5151(3)	25712(14)	528(3)
O(3)	1338(3)	4951(2)	18954(16)	603(19)
S(1)	13826(12)	64019(9)	18750(6)	483(3)

Table 3. Bond lengths [Å] and angles [°] for **3.10**.

S(1)-O(3)	1.514(3)
S(1)-C(50)	1.771(6)
S(1)-C(51)	1.783(6)
S(2)-C(5)	1.666(4)
N(4)-C(5)	1.325(5)
N(4)-C(6)	1.464(5)
N(3)-C(4)	1.357(5)
N(3)-C(5)	1.380(5)
N(1)-C(1)	1.322(4)
N(1)-N(2)	1.396(4)
N(2)-C(13)	1.272(5)
O(1)-C(2)	1.208(4)
O(2)-C(3)	1.226(4)
C(11)-C(10)	1.357(10)
C(11)-C(12)	1.377(7)
C(12)-C(7)	1.360(7)
C(7)-C(8)	1.370(6)
C(7)-C(6)	1.495(6)
C(8)-C(9)	1.377(9)
C(9)-C(10)	1.356(10)
C(4)-C(1)	1.406(5)
C(4)-C(3)	1.452(5)
C(3)-C(2)	1.516(5)
C(2)-C(1)	1.474(5)

C(13)-C(15)	1.491(7)
C(13)-C(14)	1.490(6)
O(3)-S(1)-C(50)	105.3(3)
O(3)-S(1)-C(51)	104.9(3)
C(50)-S(1)-C(51)	98.1(4)
C(5)-N(4)-C(6)	124.1(4)
C(4)-N(3)-C(5)	129.3(4)
C(1)-N(1)-N(2)	118.8(3)
C(13)-N(2)-N(1)	116.8(3)
C(10)-C(11)-C(12)	118.7(8)
C(7)-C(12)-C(11)	122.4(6)
C(12)-C(7)-C(8)	117.3(5)
C(12)-C(7)-C(6)	122.0(4)
C(8)-C(7)-C(6)	120.8(7)
C(7)-C(8)-C(9)	121.4(7)
C(10)-C(9)-C(8)	119.4(7)
C(9)-C(10)-C(11)	120.8(7)
N(4)-C(6)-C(7)	114.5(4)
N(4)-C(5)-N(3)	116.8(3)
N(4)-C(5)-S(2)	125.4(3)
N(3)-C(5)-S(2)	117.8(3)
N(3)-C(4)-C(1)	130.1(3)
N(3)-C(4)-C(3)	137.6(3)
C(1)-C(4)-C(3)	92.3(3)
O(2)-C(3)-C(4)	136.7(4)
O(2)-C(3)-C(2)	134.8(4)
C(4)-C(3)-C(3)	88.5(3)
O(1)-C(2)-C(1)	137.5(4)
O(1)-C(2)-C(3)	135.2(4)
C(1)-C(2)-C(3)	87.2(3)
N(1)-C(1)-C(4)	132.9(3)
N(1)-C(1)-C(2)	135.1(3)
C(4)-C(1)-C(2)	92.0(3)
N(2)-C(13)-C(15)	126.7(4)
N(2)-C(13)-C(14)	116.2(4)
C(15)-C(13)-C(14)	117.0(4)

Symmetry transformations used to generate equivalent atoms:

Table 4. Anisotropic displacement parameters ($Å^2 x\ 10^3$)for **3.10**. The anisotropic displacement factor exponent takes the form: $-2\pi^2[\ h^2 a^{*2} U^{11} + ... + 2\ h\ k\ a^*\ b^*\ U^{12}\ ]$

	U^{11}	U^{22}	U^{33}	U^{23}	U^{13}	U^{12}
C(1)	39(2)	32(2)	28(2)	-1(16)	6(18)	1(18)
C(2)	45(3)	36(2)	38(2)	2(18)	9(2)	-4(2)
C(3)	44(3)	36(2)	40(2)	-1(19)	6(2)	1(2)
C(4)	36(2)	33(2)	30(2)	-1(16)	6(18)	0(17)
C(5)	39(2)	42(2)	32(2)	2(18)	7(19)	-3(19)
C(6)	50(3)	58(3)	47(3)	0(2)	21(2)	12(3)
C(7)	51(3)	52(3)	45(3)	-7(2)	20(2)	-5(2)
C(8)	76(2)	78(4)	81(4)	-21(3)	38(4)	0(3)
C(9)	114(6)	108(6)	87(5)	-52(5)	57(5)	-31(5)
C(10)	123(7)	140(7)	43(4)	-28(4)	26(4)	-43(6)
C(11)	119(6)	117(6)	54(4)	-10(4)	-5(4)	4(5)
C(12)	103(5)	77(4)	48(3)	-19(3)	1(3)	9(4)
C(13)	47(3)	45(2)	37(2)	1(2)	9(2)	0(2)
C(14)	56(4)	69(4)	52(3)	-2(3)	24(3)	3(3)
C(15)	104(6)	51(3)	65(4)	-10(3)	36(4)	2(3)
C(50)	81(5)	52(3)	75(4)	-16(3)	9(3)	3(3)
C(51)	57(4)	65(4)	82(5)	-6(4)	10(4)	-3(3)
N(1)	52(2)	30(19)	42(2)	3(16)	21(17)	-1(17)
N(2)	45(2)	42(2)	36(2)	2(15)	14(17)	-4(16)
N(3)	48(2)	28(2)	37(2)	-1(15)	15(17)	-1(17)
N(4)	53(2)	39(2)	42(2)	-2(18)	20(17)	-3(2)
O(1)	77(2)	36(16)	61(19)	11(15)	27(17)	-3(15)
O(2)	66(2)	38(17)	60(19)	-3(14)	28(17)	7(15)
O(3)	78(2)	28(15)	84(2)	-9(15)	44(18)	-5(15)
S(1)	53(7)	32(6)	65(7)	-3(5)	26(6)	-4(3)
S(2)	64(8)	43(6)	52(7)	3(5)	24(6)	-11(6)

Table 5. Hydrogen coordinates (x 10^4) and isotropic displacement parameters ($Å^2$ x 10^3) for **3.10**.

	x	y	z	U(eq)
H(3)	521	129	237	50
H(4)	391	359	3094	34
H(6A)	216	215	368	70
H(6B)	205	365	3587	47
H(9)	326	507	569	11
H(10)	490	361	617	9
H(11)	564	195	557	14
H(12)	6960	13996	7249	76
H(14A)	871	231	48	66
H(14B)	939	335	9	62
H(14C)	998	199	6	68
H(15A)	825	11	85	15
H(15B)	797	23	4	11
H(15C)	705	38	50	16
H(50A)	266	665	294	92
H(50B)	164	775	278	58
H(50C)	120	644	306	14
H(51A)	-85	642	212	16
H(51B)	63	665	138	55
H(51C)	32	781	181	13